ИНФОРМАТИКА

проблемы, методы, технологии

IPMT-2026



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
«ИНФОРМАТИКА И УПРАВЛЕНИЕ» РАН





ИНФОРМАТИКА:

ПРОБЛЕМЫ, МЕТОДЫ, ТЕХНОЛОГИИ

Материалы

XXVI Международной

научно-практической конференции

им. Э.К. Алгазинова

Воронеж,
12-13 февраля 2026 г.

Под общей редакцией к.т.н. Борисова Д. Н.

Воронеж
2026

УДК 004.65+004.438.5
ББК Ч 481 (2) 22
И74

Оргкомитет:

Сопредседатели:

Волков Денис Владимирович, министр цифрового развития Воронежской области;
Салогубова Наталья Валерьевна, министр образования Воронежской области;
Старилов Юрий Николаевич, и.о. ректора Воронежского государственного университета, д.ю.н., профессор;
Крыловецкий Александр Абрамович, декан факультета компьютерных наук Воронежского государственного университета, к.ф.-м.н., доцент.

Заместители председателя:

Борисов Дмитрий Николаевич, заведующий кафедрой информационных систем ФКН Воронежского государственного университета, к.т.н., доцент;
Зацаринный Александр Алексеевич, заместитель директора Федерального исследовательского центра «Информатика и управление» РАН, д.т.н., профессор;
Коваль Андрей Сергеевич, заведующий лабораторией сетевых технологий, старший преподаватель кафедры информационных систем ФКН Воронежского государственного университета;
Костин Дмитрий Владимирович, и.о. проректора по науке, инновациям и цифровизации Воронежского государственного университета, д.ф.-м.н., доцент;
Крыловецкая Татьяна Алексеевна, заместитель декана ФКН по работе со студентами, доцент кафедры цифровых технологий ФКН Воронежского государственного университета, к.ф.-м.н., доцент;
Митрофанов Александр Юрьевич, ректор ГБУ ДПО ВО «Институт развития образования им. Н. Ф. Бунакова».

Члены комитета:

Бодров А. Ю., Бойченко И. А., Вахтин А. А., Ветохин В. В., Максимов А. В., Митрофанова Е. Ю., Соломатин Д. И., Швырева А. В., Шуйкова И. А., Юров А. Н.

Программный комитет:

Артемов М. А., Васенин В. А., Гаршина В. В., Ерешко Ф. И., Запрягаев С. А., Кобозева И. М., Кретов А. А., Крыловецкий А. А., Ковтуненко Л. В., Кургалин С. Д., Львович Я. Е., Матвеев М. Г., Махортов С. Д., Медведев С. Н., Семенов М. Е., Сирота А. А., Сычев А. В., Толстобров А. П., Чечкин А. В., Чижов М. И., Шуйкова И.А.

И74 Информатика: проблемы, методы, технологии: сборник материалов XXVI международной научно-практической конференции / под общ. ред. Д. Н. Борисова; Воронеж, Воронежский государственный университет, 12-13 февраля 2026 г. – Воронеж, ВГУ, 2026. – 1302 с.

В сборнике предлагаются научные работы, доклады и лекции, представленные на XXVI Международную научно-практическую конференцию «Информатика: проблемы, методы, технологии», проводимой Воронежским государственным университетом.

Сборник предназначен для научных работников, аспирантов и студентов старших курсов.

УДК 004.65+004.438.5
ББК Ч 481 (2) 22

© ФГБОУ ВО ВГУ, 2026

Пленарные доклады

Об актуальности фундаментальных научных исследований в области искусственного интеллекта

А. А. Зацаринный, email: azatsarinny@frccsc.ru

Федеральный исследовательский центр «Информатика и управление» Российской академии наук

Аннотация. *Статья посвящена обоснованию актуальности фундаментальных научных исследований для успешного решения ключевых проблем, возникающих при внедрении технологий искусственного интеллекта (ТИИ). Дан анализ основных факторов, обусловливающих актуальность и необходимость постановки и проведения фундаментальных исследований в области искусственного интеллекта. Приведено состояние развития ТИИ в мире, прежде всего в США. Представлен ряд фундаментальных результатов, полученных в последнее время.*

Ключевые слова: *информационно-управляющие системы, искусственный интеллект, фундаментальные исследования, ключевые проблемы, функциональная безопасность*

Введение

Новая Стратегия научно-технологического развития России определяет особую значимость развития информационно-управляющих систем (ИУС) в условиях развития цифровой трансформации во всех сферах деятельности [1]. Создание и развитие ИУС, как правило, связывают с широким применением новых информационных технологий, прежде всего технологий искусственного интеллекта (ТИИ), которое предусмотрено в рамках реализации национального проекта «Экономика данных» [2]. ТИИ сегодня занимают центральное место среди ключевых направлений мирового технологического прогресса, определяя будущее многих отраслей, они с каждым годом все глубже проникают в различные сферы развития общества. В десятках стран мира приняты национальные стратегии развития ИИ.

На необходимость развития этого направления в России неоднократно обращал внимание Президент России В.В. Путин. Достаточно упомянуть последние мероприятия, среди них: международная конференция «Путешествие в мир искусственного интеллекта» (в декабре 2024 года) [3], конференцию «Сбербанка» AI Journey-2025 [4], совещание по вопросам развития автономных систем на

© Зацаринный А. А., 2026

площадке Большой кольцевой линии Московского метрополитена (январь 2026 года) [5]. В рамках этих мероприятий В.В. Путин постоянно отмечает стратегическую значимость ТИИ, которые должны стать важнейшим ресурсом для достижения национальных целей развития, укрепления обороноспособности страны, качественного развития экономики, социальных отраслей и госуправления. Даны конкретные поручения Правительству Российской Федерации, направленные на ускорение развития ТИИ. Среди них: переход государственных органов на использование систем ИИ с применением платформенного подхода, увеличение числа бюджетных мест в вузах для подготовки специалистов по специальностям ИИ, реализацию комплекса мер по увеличению вычислительных мощностей суперкомпьютеров в России и другие.

Вместе с тем, целенаправленное и эффективное развитие ТИИ в такой огромной и ресурсоемкой стране как Россия невозможно без соответствующей поддержки этого важного направления на основе фундаментальных научных исследований.

Настоящая статья посвящена анализу факторов, обусловливающих актуальность фундаментальных научных исследований для успешного решения ключевых проблем, возникающих при внедрении ТИИ. Представлен ряд фундаментальных результатов, полученных в последнее время.

1. Факторы, определяющие актуальность фундаментальных научных исследований в области искусственного интеллекта

К основным факторам, определяющим актуальность фундаментальных научных исследований в области искусственного интеллекта, отнесем:

– ИИ как научное направление (а не только информационная технология);

– предпосылки развития ИИ;

– особая значимость ИИ в обеспечении национальной безопасности;

– ключевые проблемы внедрения технологий ИИ.

Рассмотрим кратко эти факторы.

Сегодня в мире и в России ИИ активно развивается прежде всего как комплекс технологических решений, позволяющих имитировать когнитивные функции человека, включая поиск решений без заранее заданного алгоритма, и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека. Такой комплекс технологических решений включает в себя три компонента: информационно-коммуникационную

инфраструктуру, программное обеспечение (реализующее в том числе методы машинного обучения), а также сервисы по обработке данных и поиску требуемых решений [6].

Вместе с тем, в современном научном сообществе ИИ представляется более широко – как динамично развивающееся междисциплинарное научное направление, которое предусматривает:

– фундаментальные исследования, направленные на углубленное понимание принципов работы интеллекта;

– разработку базовых технологий искусственного интеллекта, создание инструментов и аппаратно-программных средств для обеспечения функционирования систем ИИ;

– внедрение технологий ИИ в конечные изделия, процессы и системы;

– подготовку высококвалифицированных кадров, способных работать в области фундаментальных исследований ИИ и решать сложные прикладные задачи.

Именно такое понимание дано в Комплексной программе «Искусственный интеллект как драйвер цифровой трансформации экономики России», разработанной ФИЦ ИУ РАН и одобренной Советом по приоритетным направлениям Стратегии научно-технологического развития России в мае 2019 г. [7]. Такой подход был обоснован тем, что интенсивное развитие ИИ в последнее десятилетие оказалось возможным не только благодаря технологическим условиям (появлению электронной компонентной базы, необходимой для реализации технологий нейронных сетей и машинного обучения) и производственным возможностям (массовое производство чипов, видеокарт, графических ускорителей), но также достижениям в области теоретических основ искусственного интеллекта. Достаточно упомянуть отечественные научные школы Г.С. Поспелова, Д.А. Поспелова. В.К. Финна, Г.С. Осипова и других советских и российских ученых [8,9,10]. Обстоятельный анализ о приоритетных направлениях исследований и ключевых тенденциях развития технологий ИИ представлен в докладе Ю.В. Визильтера [11]. Другими словами, теоретические основания и результаты явились важнейшей предпосылкой бурного развития ИИ

Вместе с тем, предварительный анализ опыта проведения многих прикладных работ по применению конкретных систем ИИ выявил острый дефицит результатов фундаментальных исследований в области искусственного интеллекта. Наиболее важные из них состоят в создании математических методов и моделей для:

– оценки уровня доверенности (объяснимости) результатов, полученных системой ИИ;

– мониторинга и фильтрации больших данных (Big Data) для идентификации в них аномальных (например, – вредоносных) фрагментов «малого» размера;

– мониторинга и анализа Big Data в рамках жестких ограничений по времени;

– анализа устойчивости результатов, получаемых с помощью системы ИИ при варьировании подаваемых на вход разнородных и разнотипных данных, в условиях воздействия различных видов атак;

– оценки допустимости входных данных для нормального функционирования заданной программно-аппаратной системы, включая методы фильтрации входных данных, обеспечивающих защиту от вредоносных данных и так называемых «закладок без вредоносного кода»;

– оценки эффективности высокопроизводительных вычислений, включая оценку сложности вычислений, идентификацию быстро-разрешимых частных случаев, а также организацию приближенных вычислений управляемой точности.

Фундаментальные исследования по указанным направлениям предполагают развитие теоретико-вероятностных методов, вычислительных методов (линейная алгебра и др.), а также методов оптимизации.

Такие методы позволят успешно решать задачи, поставленные Президентом России в интересах обеспечения опережающего развития национальной индустрии беспилотных и автономных технологий во всех средах (на земле, в воздухе, в водном и космическом пространствах) и достижения полного суверенитета в производстве всех компонентов для автономной техники [5]. Лидерство в этой области – путь к укреплению глобальной конкурентоспособности, успешному решению вопросов социально-экономического развития в условиях дефицита на рынке труда и демографических вызовов, а также обеспечению безопасности России.

В обобщенном виде проблемы, накопленные в процессе внедрения и применения систем ИИ различного назначения, могут быть представлены как [12]:

– объяснимость результатов (выбор архитектуры нейронных сетей, их обучение, какие выборки, из каких источников, «интеллектуальность» выборок, адекватность требуемым функциям).

– устойчивость функционирования в условиях воздействия различных видов атак (атаки на обучающие выборки, архитектуры ИНС и др. компоненты систем с ИИ);

– создание доверенной среды исследований, разработок, испытаний и внедрения (принципиально – доверенной, т.к.

отечественных аппаратных и программных средств для создания систем с ИИ сегодня нет);

– высокопроизводительная инфраструктура (для обучения ИНС, особенно при применении современных сетей типа трансформеров требуются огромные вычислительные ресурсы);

– острый дефицит высококвалифицированных специалистов для выполнения всех этапов создания систем с ИИ (исследования, разработка, эксплуатация, применение).

Основу для разрешения приведенных проблем должны создавать результаты фундаментальных исследований.

2. Состояние фундаментальных исследований в мире. Место России

Согласно глобальному индексу ИИ, который учитывает несколько десятков показателей и индикаторов, отражающих развитие ИИ в различных странах [13], Россия занимает 30-е место в мире (лидируют США, Китай и Сингапур), а в рейтинге публикационной активности на 15-м месте (1,5% публикаций).

Вместе с тем, научные заделы и научно-промышленный потенциал России позволяют оценивать ее реальное место более высоко. Достаточно упомянуть тот факт, что в Стратегии США «О сохранении американского лидерства в области искусственного интеллекта», подписанной президентом США Д. Трампом 11 февраля 2019 года [14] Россия наряду с Китаем определена главным конкурентом в развитии ТИИ, прежде всего в военной сфере. Отметим также, что статьи российских исследователей цитируются ведущими научными институтами мира (Стэнфордский университет, Калифорнийский университет в Беркли, научно-исследовательская лаборатория Google DeepMind). Научные коллективы ФИЦ ИУ РАН, ВМК МГУ им. М.В. Ломоносова, МФТИ, ИПМ им. М.В. Келдыша, ИТМО, ИСП им. В.П. Иванникова, МГТУ им. Н.Э. Баумана и других организаций обладают научными школами и традициями в области ИИ, а также существенными научными заделами. С учетом приведенных фактов реальное место России в глобальном индексе ИИ, конечно, явно занижено.

Орган Генеральной Ассамблеи ООН по торговле и развитию (ЮНКТАД) опубликовал ежегодный доклад о технологиях, посвященный роли искусственного интеллекта в достижении целей устойчивого развития [15]. В нем даны оценки текущего состояния рынка передовых технологий и перспективы его развития. Основные из них:

– рынок ИИ в 2023 году достиг объема в $189 млрд. и к 2033 году может вырасти в 25 раз ($4,8 трлн);

– по итогам 2023 года глобальный рынок передовых технологий составил $2,5 трлн. (36% – интернет вещей, 15% – на электромобили, 7% – на ИИ, а 34% – на другие технологии; при этом к 2033 году доля ИИ может вырасти до 29% и станет самой крупной на рынке;

– на долю США и Китая приходится 33% публикаций в мире об ИИ и 60% патентов в этой сфере;

– ИИ может привести к автоматизации задач, связанных с написанием текстов, программированием и анализом данных, что приведет к сокращению в развитых странах 33% рабочих мест, в развивающихся – 24%, в наименее развитых – 18%;

– среди развивающихся стран, помимо Китая, наибольшего прогресса в сфере ИИ достигли Бразилия, Индия и Филиппины благодаря развитию интернет-инфраструктуры и расширению доступа к сети в стране, а также инвестициям в научно-исследовательские проекты.

Еще один интересный рейтинг, характеризующий уровень распространения технологий искусственного интеллекта населением различных стран мира, опубликовала компания Microsoft [16]. Так, к началу 2026 года ИИ пользуются 16,3 % населения планеты, в среднем примерно каждый шестой. Мировые лидеры по проникновению ИИ ОАЭ (ИИ используют 64 % работающего населения), Сингапур (60,9 %), Норвегия (46,4 %), Ирландия (44,6 %) и Франция (44,0 %).

Россия в этом зачёте заняла 119-е место (8,0 %). Вместе с тем, этот рейтинг не совпадает с приведенным выше, отражающим технологическое лидерство в области ИИ. Так, мировой лидер США занимают только 24-е место (28,3 % населения). Активно занимается распространением ИИ среди населения Южная Корея (18-е место), которая стала одним из наиболее быстрорастущих рынков ChatGPT, в связи с чем OpenAI открыла офис в Сеуле. Генеративный ИИ активно используется в учебных заведениях, на рабочих местах и в общественных службах.

Огромный импульс распространению ИИ дал китайский сервис DeepSeek, который появился год назад, но составил серьезную конкуренцию лучшим американским системам. Открытость и доступность помогли DeepSeek завоевать популярность на рынках, где западные платформы представлены недостаточно: в Африке сервисом пользуются в 2–4 раза активнее, чем в других регионах. DeepSeek также активно используют в Китае, России, Беларуси и на Кубе.

Особо отметим активную позицию США по развитию искусственного интеллекта, прежде всего в целях обороны и национальной безопасности. Начало ИИ-гонки между Китаем и США положил упомянутый выше Указ о сохранении американского лидерства

в области ИИ [14]. Этот документ обозначил стратегические направления для ускорения исследований и разработок в области ИИ, подчеркивая его важность для самых разных сфер, включая оборону, экономику и здравоохранение, с особым акцентом на военные технологии

Администрация президента США Дж. Байдена также определила ряд инициатив по развитию ИИ в интересах обороны и безопасности. Так, в октябре 2024 года подписан президентский «Меморандум о развитии лидерства Соединенных Штатов в области искусственного интеллекта; использовании искусственного интеллекта для достижения целей национальной безопасности; и повышении безопасности, надежности и достоверности искусственного интеллекта», адресованный ключевым ведомствам, отвечающим за безопасность страны (Минобороны, разведывательные службы и др.). В дополнение к нему была представлена «Концепция развития управления ИИ и управления рисками применения ИИ в сфере национальной безопасности», где подробно описаны подходы к минимизации угроз, связанных с внедрением таких технологий [17]. Следующим важным шагом стало издание 14 января 2025 года «Исполнительного указа Президента США о продвижении лидерства Соединенных Штатов в области инфраструктуры искусственного интеллекта», в котором предусматривается создание национальных центров обработки данных как основы для масштабных вычислений, необходимых для развития ИИ. В совокупности эти документы определяют три ключевые цели по достижения глобального лидерства США в области безопасного и надежного ИИ: привлечение высококвалифицированных специалистов со всего мира, модернизация и расширение вычислительной инфраструктуры, а также усиление мер по защите технологий от иностранных разведок. Особое внимание в Меморандуме уделено большим языковым моделям (БЯМ), которые рассматриваются как приоритетное направление развития ИИ. Для их совершенствования подчеркивается необходимость улучшения качества наборов данных, повышения мощности вычислительных систем, а также создания надежных инструментов разработки и тестирования [18].

Ключевыми факторами, определяющими прогресс ИИ в ближайшие годы, обозначены несколько аспектов: использование сложных алгоритмов, позволяющих решать задачи высокой степени сложности; совершенствование вычислительного оборудования, которое становится все более производительным и энергоэффективным; готовность крупных частных промышленных компаний инвестировать значительные средства в научные исследования и разработки, что должно способствовать ускорению технологического развития. Наконец, важную роль играет

расширение объема и разнообразия обучающих данных, которые обеспечивают точность, надежность и адаптивность систем ИИ. Эти элементы создают прочную основу для быстрого развития технологий [19,20,21].

Новые инициативы в США по развитию ТИИ, прежде всего в сфере фундаментальных исследований, приняты уже в январе 2026 года. Так, Меморандум Министерства войны США об ускорении укрепления доминирующего положения США в области военного искусственного интеллекта определяет выполнение (начиная с 2026 года) нескольких проектов, сфокусированных на ключевых возможностях для усиления превосходств военного ИИ в таких областях, как боевые действия, разведка и корпоративные миссии [22].

Еще один Меморандум Заместителя министра войны США [22] направлен на защиту результатов фундаментальных исследований в США в области ИИ от злонамеренного иностранного влияния, кражи интеллектуальной собственности и других форм использования результатов исследований, наносящих ущерб Соединённым Штатам. Так, предусмотрен запрет финансирования из средств, выделенных на поддержку фундаментальных исследований, ряда организаций, в том числе из перечня «Китайские военные компании, действующие в Соединённых Штатах». Вышеуказанные меры представляют собой решительный шаг к защите результатов фундаментальных исследований Министерства войны США от неправомерного использования и сохранению научно-технического превосходства вооружённых сил США.

Таким образом, США стремятся не только сохранить, но и укрепить свое лидерство в этой области, рассматривая ИИ как инструмент не только защиты национальных интересов, но и глобального влияния.

3. Основные подходы к решению ключевых проблем на основе фундаментальных исследований

Основные подходы к решению ключевых проблем на основе фундаментальных исследований, как представляется, должны базироваться на соотношении востребованных прикладных технологий ИИ и научно обоснованных доверенных технологий. Перечень прикладных технологий, ориентированных на применение в конкретных областях (промышленное производство, социальная сфера (медицина, торговля, транспорт, услуги и др.), а также сферы специального, прежде всего военного, применения), обосновывается соответствующими федеральными органами, промышленными предприятиями и бизнес-структурами.

Вместе с тем, такие прикладные технологии должны формироваться на основе базовых технологий ИИ. К базовым ТИИ принято относить

совокупность системообразующих технологий, имеющих теоретические (математические) обоснования ориентированных на воспроизведение отдельных базовых когнитивных функций (компьютерное зрение, распознавание и синтез речи, обработка естественного языка и др.) [10]. Так, в [7] определены следующие базовые ТИИ:

– компьютерное зрение;
– распознавание и синтез речи;
– обработка и интеллектуальный анализ естественных языков;
– поддержка принятия решений;
– машинное (включая глубокое) обучение;
– интеллектуальный анализ больших данных и знаний;
– инженерия знаний (прежде всего, на динамических графах знаний);
– мультимодальная аналитика и рассуждения;
– планирование и управление целенаправленным поведением в неструктурированных средах;
– мультиагентное управление;
– извлечение знаний из различных источников;
– автономные непилотируемые интеллектуальные системы.

Прикладные технологии ИИ – это, как правило, проблемно-ориентированные комплексы, целенаправленно формируемые из базовых ТИИ с ориентацией на специфику представления данных и знаний в соответствующей прикладной предметной области, а также характерные для нее подходы и методы анализа данных и поддержки принятия решений. Поэтому прикладных технологий может быть очень много.

Отметим, что перечень базовых ТИИ имеет открытый характер, т.к. ИИ – постоянно пополняемая новыми знаниями область исследований и разработок, в рамках которых могут появляться принципиально новые подходы, математические модели, методы и алгоритмы и, соответственно, реализующие их базовые технологии.

Важно, что в 2024 году по инициативе Отделения математических наук РАН (руководитель академик-секретарь Отделения математических наук РАН академик В.В. Козлов), которую поддержали президент РАН академиком Г.Я. Красников и Заместитель Председателя Правительства Д.Н. Чернышенко межведомственной рабочей группой (руководитель Н академик В.В. Козлов) в составе представителей Департамента государственной политики в сфере научно-технологического развития Минобрнауки России, РАН, профильных министерств, государственной корпорации «Росатом», университетов и научно-исследовательских институтов проводится работа по разработке Концепции развития математических наук до 2030 года [24]. Цель Концепции – обеспечить

условия для достижения нашей страной лидирующих позиций в области математических наук, а это, в свою очередь, будет способствовать обеспечению технологического лидерства России, конкурентоспособности и безопасности. В подготовленной концепции подчёркивается ключевая роль фундаментальной математики для современных областей – искусственного интеллекта, квантовых вычислений, системного программирования, кибербезопасности. В основе этих технологий лежит математика.

4. Некоторые результаты фундаментальных исследований ФИЦИУ РАН

В Федеральном исследовательском центре «Информатика и управление» РАН фундаментальные исследования ключевых проблем искусственного интеллекта выполняются в рамках государственного задания, а также крупного научного проекта "Математические модели и численные методы как основа для разработки робототехнических комплексов, новых материалов и интеллектуальных технологий конструирования" при поддержке Минобрнауки России.

Среди них выделим комплекс в части методов навигации мобильных роботов в средах большой площади, а также методов децентрализованного избегания столкновений, планирования траекторий и распределения целей в группе мобильных агентов (К.С. Яковлев, К.Ф. Муравьев, А.С. Дергачев) [25].

Отметим, что на рабочей встрече с Президентом России президент РАН Г.Я. Красников 9 декабря 2025 г. при обсуждении участия РАН в решении проблем научно-технологического развития страны в числе других представил и результаты ФИЦ ИУ РАН в части приведенных выше методов. Эти результаты вызвали живой интерес у Президента России [26]. Подробнее об этом научном результате – на сайте ФИЦ ИУ РАН.

Еще один фундаментальный результат – новые вероятностно-обоснованные нейросетевые модели эффективной обработки аэрокосмических изображений высокого разрешения (А.К. Горшенин, А.М. Достовалова). Отметим, что по результатам X Международной конференции AI Journey 2025 статья этих авторов была признана лучшей [27].

Кроме того, выделим еще несколько направлений исследований.

В [28] обсуждается использование методов и систем искусственного интеллекта для обеспечения ситуационной осведомленности лиц, принимающих решения, в контексте интеллектуального анализа Big Data, который должен быть выполнен в ограниченное время. Обоснованы предложения по созданию комплекса математических методов как

теоретической основы решения задач формирования системы целеполагания, анализа больших объемов разнородной информации, их систематизации и представления в виде вариантов решений для поддержания ситуационной осведомленности ЛПР.

В [29] представлен метод обнаружения триггеров на изображениях с использованием статистических характеристик, полученных из матрицы GLCM, который позволяет эффективно обнаруживать триггеры в различных формах, включая триггеры в виде полос случайного шума, которые могут быть сложны для распознавания с помощью других методов.

В [30,31] исследуются проблемы комплексной безопасности ИИ, которая включает в себя наряду с классической для автоматизированных систем технической безопасностью системную и функциональную безопасностью, характерной для систем ИИ. Предложены подходы к развитию технического регулирования в области информационной безопасности систем ИИ.

Здесь следует также отметить исследования проблем безопасности ТИИ, выполняемые на кафедре информационной безопасности факультета ВМК МГУ им. М.В. Ломоносова. Они охватывают проблемы, связанные с применением ТИИ в киберзащите, в кибератаках, а также с безопасностью самих ТИИ. Актуальный аналитический обзор состояния этих проблем представлен в [32, 33].

Заключение

В статье показано:

1. Развитие технологий искусственного интеллекта в России должно стать важнейшим ресурсом для достижения национальных целей развития, укрепления обороноспособности страны, качественного развития экономики, социальных отраслей и государственного управления.

2. Целенаправленное и эффективное развитие ТИИ в такой огромной и ресурсоемкой стране как Россия невозможно без соответствующей поддержки этого важного направления на основе фундаментальных научных исследований.

3. ИИ является динамично развивающимся междисциплинарным научным направлением, которое предусматривает наряду с разработкой и внедрением ТИИ проведение фундаментальных исследований, направленных на углубленное понимание принципов работы ТИИ, включая вопросы комплексной безопасности.

4. Результаты фундаментальных исследований должны создавать основу для разрешения накопленных ключевых проблем внедрения ТИИ, включающих:

5. достижение объяснимости результатов, получаемых с использованием ТИИ;

6. обеспечение устойчивости функционирования в условиях воздействия различных видов атак;

7. создание доверенной среды для исследований, разработок, испытаний и внедрения систем с ИИ;

8. создание высокопроизводительной инфраструктуры;

9. подготовку высококвалифицированных специалистов для выполнения всех этапов создания систем с ИИ (исследования, разработка, эксплуатация, применение).

10. Основные подходы к решению ключевых проблем на основе фундаментальных исследований должны быть сконцентрированы на развитии системообразующих доверенных базовых ТИИ и формировании на их основе предметно-ориентированных прикладных ТИИ для применения в различных сферах.

Список литературы

1. Стратегия научно-технологического развития Российской Федерации : утв. Указом Президента Российской Федерации от 28 февр. 2024 г. № 145. – URL: http://www.kremlin.ru/acts/bank/50358 (дата обращения: 04.03.2025).

2. Национальный проект «Экономика данных» [Электронный ресурс]. – URL: http://government.ru/rugovclassifier/909/events/ (дата обращения: 04.03.2025).

3. Международная конференция «Путешествие в мир искусственного интеллекта» (Москва, 11–13 декабря 2024 г.) [Электронный ресурс]. – URL: http://kremlin.ru/events/president/news/75830 (дата обращения: 04.03.2025).

4. Путин поручил широко использовать технологии ИИ по всей России [Электронный ресурс] // Mail.ru. – 2025. – 3 февр. – URL: https://news.mail.ru/politics/68777748/?frommail=1&md=1&utm_partner_id=421 (дата обращения: 04.03.2025).

5. Совещание по вопросам развития автономных систем [Электронный ресурс] : рабочие материалы. – URL: http://www.kremlin.ru/events/president/transcripts/79016 (дата обращения: 04.03.2025).

6. Национальная стратегия развития искусственного интеллекта на период до 2030 года : утв. Указом Президента Российской Федерации

от 10 окт. 2019 г. № 490. – URL: http://static.kremlin.ru/media/events/files/ru/AH4x6HgKWANwVtMOfPDhcbRpvd1HCCsv.pdf (дата обращения: 04.03.2025).

7. Концепция комплексной программы «Искусственный интеллект как драйвер цифровой трансформации экономики России» / ФИЦ ИУ РАН. – М., 2019. – Одобрена Советом по приоритетным направлениям Стратегии научно-технологического развития России в мае 2019 г.

8. Поспелов, Г. С. Искусственный интеллект – основа новой информационной технологии / Г. С. Поспелов. – М. : Наука, 1988. – 280 с. – URL: https://www.livelib.ru/author/579614-germogen-pospelov (дата обращения: 04.03.2025).

9. Финн, В. К. Искусственный интеллект: Методология, применения, философия / В. К. Финн. – 2-е изд., испр. и доп. – М. : [б. и.], 2023. – 468 с.

10. Осипов, Г. С. Лекции по искусственному интеллекту / Г. С. Осипов. – 3-е изд., стер. – М. : [б. и.], 2026. – 272 с. – (Науки об искусственном интеллекте).

11. Визильтер, Ю. В. Приоритетные направления исследований и ключевые тенденции развития технологий ИИ : доклад на Конференции по искусственному интеллекту КИИ-2025, Санкт-Петербург, 08 окт. 2025 г. / Ю. В. Визильтер. – URL: https://www.gosniias.ru/ (дата обращения: 04.03.2025).

12. Зацаринный, А. А. Ключевые проблемы развития технологий искусственного интеллекта в интересах обеспечения военной безопасности государства / А. А. Зацаринный, К. В. Иванов // Известия Российской академии ракетных и артиллерийских наук. – 2025. – № 3 (138). – С. 13-19.

13. Global AI Index 2024 (пятая версия, 19.09.2024). – URL: https://www.tortoisemedia.com/data/global-ai#rankings (дата обращения: 05.03.2025).

14. Maintaining American Leadership in Artificial Intelligence : указ Президента США от 11 февр. 2019 г. № 13859. – URL: https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence (дата обращения: 12.03.2025).

15. Доклад о технологиях и инновациях – 2025 : флагманский доклад / ЮНКТАД. – URL: https://unctad.org/system/files/official-document/tir2025overview_ru.pdf (дата обращения: 05.03.2025).

16. Россия в рейтинге внедрения ИИ [Электронный ресурс] // РБК. Технологии и медиа. – 2026. – 20 янв. – URL:

https://www.rbc.ru/technology_and_media/20/01/2026/696f3d8d9a794776030bf2d1 (дата обращения: 25.01.2026).

17. National Security Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence To Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence от 24 окт. 2024 г. – URL: https://www.govinfo.gov/content/pkg/DCPD-202400945/html/DCPD-202400945.htm (дата обращения: 12.03.2025).

18. Advancing United States Leadership in Artificial Intelligence Infrastructure : указ Президента США от 14 янв. 2025 г. № 14141. – URL: https://www.federalregister.gov/documents/2025/01/17/2025-01395/advancing-united-states-leadership-in-artificial-intelligence-infrastructure (дата обращения: 12.03.2025).

19. Малинецкий, Г. Г. Искусственный интеллект. Перспективы разработки и внедрения в России. Предложения / Г. Г. Малинецкий // Научный вестник оборонно-промышленного комплекса России. – 2025. – № 3. – С. 64-74.

20. Малинецкий, Г. Г. Перспективные технологии стран НАТО, формирующие облик образцов вооружений и военно-промышленного комплекса / Г. Г. Малинецкий, Т. С. Ахромеева // Научный вестник оборонно-промышленного комплекса России. – 2025. – № 2. – С. 37-44.

21. Фролов, И. Э. Искусственный интеллект как драйвер прорывных технологий: глобальные тренды и уроки для России / И. Э. Фролов, В. Н. Киселев // Проблемы прогнозирования. – 2025. – № 3 (210). – С. 122-134.

22. Меморандум Министерства войны США об ускорении укрепления доминирующего положения США в области военного ИИ : пер. с англ. проф. И. В. Понкина [Электронный ресурс] // ВВести. – 2025. – 15 дек. – URL: https://vvesti.com/politika/perevod-memoranduma-ministerstva-vojny-ssa-ob-uskorenii-ukreplenia-dominiruusego-polozenia-ssa-v-oblasti-voennogo-ii (дата обращения: 04.03.2025).

23. Меморандум Заместителя министра войны США от 07.01.2026 об инициативах и реализации мер по обеспечению безопасности фундаментальных исследований : пер. с англ. проф. И. В. Понкина [Электронный ресурс] // ВВести. – 2026. – 25 янв. – URL: https://vvesti.com/politika/perevod-memoranduma-zamestitela-ministra-vojny-ssa-ob-iniciativah-i-realizacii-mer-po-obespeceniu-bezopasnosti-fundamentalnyh-issledovanij (дата обращения: 25.01.2026).

24. Развитие математики – это часть культурного обогащения нашей нации : интервью с академиком С. И. Безродных / С. И. Безродных // Российская академия наук. – URL: https://new.ras.ru/press-center/razvitie-

matematiki-eto-chast-kulturnogo-obogashcheniya-nashey-natsii-akademik-sergey-bezrodnykh-ko-/ (дата обращения: 04.03.2025).

25. Дергачев, С. С. 2.5D Mapping, Pathfinding and Path Following For Navigation Of A Differential Drive Robot In Uneven Terrain / С. С. Дергачев, К. В. Муравьев, К. А. Яковлев // IFAC-PapersOnLine. – 2022. – Т. 55, № 38. – С. 80-85.

26. Рабочая встреча В. В. Путина с президентом РАН Г. Я. Красниковым [Электронный ресурс] : рабочие материалы. – URL: http://www.kremlin.ru/events/president/news/78693 (дата обращения: 04.03.2025).

27. Достовалова, А. М. Нейросетевые классификаторы изображений, информированные факторными анализаторами / А. М. Достовалова, А. К. Горшенин // Доклады РАН. Математика, информатика, процессы управления. – 2024. – С. 41-48.

28. Грушо, А. А. О некоторых особенностях задач формирования ситуационной осведомленности в системах поддержки управленческих решений на основе технологий искусственного интеллекта / А. А. Грушо, М. И. Забежайло, А. А. Зацаринный // Искусственный интеллект и принятие решений. – 2026. – № 1. – С. [принята к публикации].

29. Выявление атак типа «триггер» на искусственные нейронные сети / А. А. Зацаринный [и др.] // Вычислительные системы и информационные технологии. ИТВС. – 2025. – № 1. – С. 3-13.

30. Гаврилов, В. Е. Особенности обеспечения функциональной безопасности автоматизированных систем с применением технологии искусственного интеллекта / В. Е. Гаврилов, А. А. Зацаринный // Системы и средства информатики. – 2024. – Т. 34, № 3. – С. 23-34.

31. Зацаринный, А. А. Некоторые подходы к развитию технического регулирования в области информационной безопасности систем с искусственным интеллектом / А. А. Зацаринный, В. Е. Гаврилов // ИТ-Стандарт. – 2025. – № 3 (44). – С. 11-18.

32. Намиот, Д. Е. Искусственный Интеллект в Кибербезопасности. Хроника. Выпуск 4 / Д. Е. Намиот // International Journal of Open Information Technologies. – 2026. – Т. 14, № 1. – С. 81-94.

33. Namiot, D. Artificial Intelligence in Cybersecurity. Chronicle. Issue 1 / D. Namiot // International Journal of Open Information Technologies. – 2025. – Vol. 13, No. 9. – P. 34-42.

Математические модели голографической обработки векторно-скалярных сигналов в гидроакустических волноводах

С. А. Переселков, email: pereselkov@yandex.ru

Воронежский государственный университет

Аннотация. *Представлено современное состояние проблемы применения голографической интерферометрии в гидроакустике. Изложение голографической обработки сигналов сконцентрировано на решении задачи контроля подводной обстановки, обеспечивающего обнаружение, разрешение и локализацию подводных источников шума в условиях малого входного отношения сигнал/помеха и пространственно-временных возмущений океанической среды.*

Ключевые слова: *голографическая обработка, обнаружение, разрешение, локализация, векторно-скалярный приемные модули, шумовой источник, помеха.*

Введение

Одна из слабо разработанных проблем обработки гидроакустической информации - это обеспечение высокой помехоустойчивости в условиях пространственно-временной изменчивости акватории, совмещенной с адаптивностью к изменяющимся условиям распространения волнового поля в океанической среде. Трудности подхода к ее решению вытекают из-за специфики формирования сигналов и помех в реальных условиях и особенностей применимости современных методов обработки. При этом, как правило, априорные информации о характеристиках принимаемых сигналов и акватории отсутствуют. Рассмотрение этой проблемы тесно связано с вопросами обнаружения, разрешения и локализации шумоых подводных источников в мелководных акваториях, на которых и будет сконцентрировано внимание. Под локализацией источника звука понимается определение его параметров: пеленга, радиальной скорости, удаления и глубины. Контроль подводной обстановки является одной из приоритетных задач гидроакустики, решение которой связано с анализом звуковых полей малошумных подводных источников в случайно-неоднородных океанических средах. В настоящее время можно выделить три основных направлений исследований в этой области.

© Переселков С. А., 2026

Одно из них, которое имеет длительную историю, основано на согласованных со средой распространения методах пространственной обработке гидроакустических сигналов (MFP) [1-6]. Методы подразумевает использование многоэлементных приемных антенн и данных о трассе распространения, позволяющих рассчитывать модельную форму сигнала. Практическая реализация согласованных со средой методов сталкивается с значительными проблемами, которые связанны с неизбежными неточностями модельной формы сигнала, малым входным отношением сигнал/помеха (с/п) и неконтролируемыми искажениями принимаемого сигнала. Все эти факторы заметно снижают эффективность согласованных со средой методов [4]. Однако в настоящее время ведутся активные исследования возможностей преодоления отмеченных недостатков согласованных методов в реальных гидроакустических условиях [7-9].

Сценарии поведения широкополосных гидроакустических сигналов в многомодовых океанических волноводах, хотя чрезвычайно разнообразны и сложны, однако они подчиняются одной общей закономерности, впервые теоретически и экспериментально установленной в [10, 11] и позже в [12, 13]. Она проявляется в том, что интерференция мод широкополосных сигналов, благодаря волноводной дисперсии, формирует устойчивую интерференционную картину (интерферограмму) максимумов спектральной плотности энергетических параметров волнового поля в переменных частота-расстояние или частота-время. Конфигурация локализованных полос определяется частотным диапазоном, параметрами волновода, скоростью и траекторией источника. Вдоль интерференционных полос сдвиг фазы между интерферирующими модами остается практически постоянным. Данная закономерность применима, − в чем состоит одно из основных ее преимуществ, − и к шумовым сигналам, описание которых носит статистический характер. Это позволяет рассматривать интерферограмму как универсальную детерминированную математическую модель шумового сигнала и среды распространения. Модель отличается достаточной простотой и в то же время содержит в себе большое разнообразие форм поведения. Кроме того, она обладает предсказательной силой, но не требует жестких причинно-следственных ограничений. В сущности, роль моделей в физических построениях определяется тем, насколько хорошо они передают существенные черты реальных процессов, не будучи ни чересчур сложными, ни чересчур упрощенными. Тем самым открывается новая парадигма обработки гидроакустических сигналов, построенная на устойчивых особенностях интерферограмм звукового поля шумового источника.

Второе направление строится на интерферометрической обработке сигналов (ИОС), основываясь на механизме частотных смещений интерференционных максимумов (ЧСИМ) волнового поля. [11]. В отсутствие помехи и нерегулярностей среды распространения решен ряд частных задач по контролю подводной обстановки в мелководной акватории. Таких, как оценка расстояния до источника (пассивный и активный режимы), определение глубины, дистанционное зондирование геоакустических параметров среды, диагностика гидродинамических возмущений и т.д. В [14] выполнен анализ интерферограмм одиночных приемников. В [15] рассмотрено перемещение фокального пятна в эксперименте с временной реверсией. В [16] исследовалась возможность компенсации движения источника. В [17-23] разработаны методы определения расстояния до источника. В [24] получены оценки инвариантных параметров сигналов в океанических волноводах. В [25] рассмотрена задача классификации параметров морского дна по шумовым сигналам проходящих судов. В [19, 26, 27] предложены методы для оценки глубины источника. В [28] построена картина интерференционных полос на основе времен прихода мод. В [29] разработан алгоритм регистрации интерференционной картины реверберационного сигнала. В [30] продемонстрировано подавление тональной интерференции. В монографии [31] рассмотрено применение ИОС для диагностики гидродинамических возмущений в океанических волноводах. Подход на основе ИОС позволяет рассмотреть и объединить большой круг задач гидроакустики, представляющих теоретический и практический интерес.

Однако ИОС, основанная на извлечении информации из интерферограммы, обладает существенными недостатками. Она также, как и методы MFP, не позволяет решать задачи обнаружения и определения параметров источника шума при малом входном отношении с/п в акваториях со случайными неоднородностями, когда интерферограммы зашумлены или искажены. Подход на основе ИОС не позволяет также разрешать сигналы нескольких источников, так из муаровой картины в подавляющем большинстве случаев практически невозможно выделить интерферограммы сигналов отдельных источников.

Решение проблемы контроля состояния подводной обстановки, помимо развития новых методов обработки гидроакустических сигналов, подразумевает также использование новых подходов к регистрации акустических полей. На настоящий момент методы измерений приемниками давления, достигли предельных возможностей. Повышение помехоустойчивости обработки, как показывают результаты

измерений, может обеспечить использование векторной структуры звуковых полей, измеряемой векторно-скалярными приемниками (ВСП), включающие в себя скалярный канал звукового давления и три канала компонент вектора колебательной скорости. Помимо увеличения помехоустойчивости измерений, появляется возможность определять направление (пеленг) на источник одиночным ВСП. Методы измерений компонент векторно-скалярных полей получили название векторно-фазовые методы. Результаты гидроакустических измерений с применением ВСП рассматривались в работах [32−46].

Третье направление, голографическая обработка сигналов (ГОС), интенсивно развивающееся в последние годы, опирается на применение голографической интерферометрии с использованием ВСП. В отличие от интерферометрического подхода ИОС, где получают и изучают интерферограммы, формируемыми шумовыми сигналами, в голографической обработке эти интерферограммы предварительно регистрируются с помощью голограмм на основе применения двумерного преобразования Фурье, а затем уже из голограмм производится извлечение информации о параметрах источников. Такое, казалось бы, несущественное отличие приводит ко многим важным следствиям. В частности, удается повысить помехоустойчивость обработки, устранить, точнее, существенно уменьшить искажения сигналов, вызванных случайными неоднородностями среды распространения, а также развить адаптивные методы обработки, не требующие знания априорной информации о характеристиках передаточной функции акватории. За короткий отрезок времени ГОС превратилась в мощный исследовательский метод, способный решать современные задачи освещения подводной обстановки: обнаружения, разрешения и локализации малошумных подводных источников в условиях пространственно-временной изменчивости океанической среды. Укажем приоритетные работы, где рассматривались вопросы, связанные с обоснованием ГОС и ее применением для решения задач освещения подводной обстановки. Физические и математические принципы голографической обработки в частотно-временной области и ее апробация на натурном эксперименте впервые описаны в [47]. Помехоустойчивость обработки по отношению к скалярной и векторным компонентам шумового поля при восстановлении радиальной скорости и удаленности источника обсуждалась в [48-50]. В [51-54] разработаны алгоритмы пеленгования источников с применением двух и одного одиночных ВСП и продемонстрирована их помехоустойчивость на примерах натурного и вычислительного экспериментов. В [55, 56] описан метод определения глубины источника, и в рамках натурного и

вычислительного экспериментов исследована его помехоустойчивость. В [57-62] предложен критерий оценки предельной дальности обнаружения шумового источника и на основе экспериментальных данных о спектральных характеристиках шумоизлучения источника выполнена его апробация. В [63] рассмотрена задача о взаимосвязи между наклоном локализованных полос интерферограммы и угловым коэффициентом прямой голограммы, на которой расположены координаты фокальных пятен сигнала, обусловленные интерференцией мод разных номеров. В [53, 54, 64-66] приведены результаты высокочастотных экспериментов по локализации шумовых источников. В [67-69] разработаны адаптивные алгоритмы, которые для восстановления параметров источника не требуют знания о характеристиках среды распространения и выполнена их верификация. В [53, 54, 70-72] рассмотрена задача о разрешении шумовых сигналов на фоне помехи и случайных неоднородностей среды, оценена погрешность восстановления параметров источников, приведены результаты вычислительных и натурного экспериментов. В [73, 74] построена теория вероятностных характеристик обнаружения шумового сигнала по критерию Неймана-Пирсона. Условия формирования неискаженной интерферограммы (голограммы) шумового сигнала обсуждались в [57, 60-62]. В [75-83] исследовалась обработка гидроакустических сигналов на фоне интенсивных внутренних волн (ИВВ) и ее устойчивость по отношению к вариациям скорости звука и глубины акватории. В [84-86] разработан метод выделения мод шумовых сигналов и оценок их параметров, приведены результаты его апробации в мелководных акваториях. Работы [47-86] относятся к ГОС с использованием одиночных ВСП. В [61, 62, 87-92] рассмотрена ГОС с применением линейных антенн. В [93, 94] предложена новая концепция построения роботизированной системы освещения подводной обстановки, основанная на использовании гибридных автономных необитаемых подводных аппаратов (АНПА), т.е. классических АНПА, способных работать дополнительно в режиме подводного глайдера. Нынешнее состояние голографической обработки широкополосных гидроакустических сигналов отражено также в монографии [95].

По-видимому, впервые на возможность применения голографии в гидроакустических задачах в частотно-пространственной области было указано В.А. Зверевым и сотрудниками его отдела Е.Ф. Орловым и Г.А. Шароновым в сборнике [96]. Результаты дальнейших теоретических и экспериментальных исследований были опубликованы в работах [97, 98], а затем обобщены в монографии [99]. В.А. Зверев большое внимание уделял аналогии между методом временного обращения волн и голографией [100]. Благодаря сходству с голографией и простоте

осуществления, принцип временного обращения волн стал крупным достижением в области акустики океана. На этой основе В.А. Зверевым было предложено решение ряда важных задач [101-103]. Результаты, полученные В.А. Зверевым в этом направлении вошли в монографию [104].

Следует отметить, что активные исследования в области ГОС в настоящее время проводят и зарубежные коллеги. Среди работ наиболее близких по тематике можно отметить работы коллег из океанографического института США - Woods Hole Oceanographic Institution [105, 106].

Математические модели голографической обработки

Голографическая обработка шумового сигнала источника с применением одиночного ВСП строится следующим образом. За время наблюдения Δt в полосе Δf накапливаются J независимых временных шумовых реализаций звукового давления $P(t)$ и векторных компонент колебательных скоростей частиц жидкости $V_{X,Y,Z}(t)$ длительностью T и с временным интервалом δT между ними

$$J = \Delta t / (\delta T + T). \tag{1}$$

Реализации независимы, если $\delta T \geq 1/\Delta f$. Оконным преобразованием Фурье восстанавливаются частотно-временные представления шумовых реализаций $P(f,t)$ и $V_{X,Y,Z}(f,t)$. Формируются интерферограммы звукового давления:

$$I(f,t) = P(f,t)P^*(f,t) \tag{2}$$

и компонент составляющих потока мощности

$$\begin{aligned} X(f,t) = V_X(f,t)P^*(f,t), Y(f,t) = \\ = V_Y(f,t)P^*(f,t), Z(f,t) = \\ = V_Z(f,t)P^*(f,t), \end{aligned} \tag{3}$$

где верхний индекс «*» обозначает комплексно-сопряженную величину. Для повышения контрастности на интерферограммах удаляются средние значения $\bar{I}$, $\bar{X}$, $\bar{Y}$, $\bar{Z}$ и применением к ним двумерного преобразования Фурье вычисляются голограммы $F_{\tilde{I},\tilde{X},\tilde{Y},\tilde{Z}}(\tau,\nu)$. Интерферограммы и голограммы скалярной и векторных компонент шумовых сигналов, а также их комбинаций, когерентны и различаются помехоустойчивостью [44–47].

Например, применительно к интерферограмме $\tilde{X}(f,t) = X(f,t) - \overline{X}(f,t)$ интегральное преобразование имеет вид

$$F_{\tilde{X}}(\tau,\nu) = \int_0^{\Delta t}\int_{f_1}^{f_2} \tilde{X}(f,t)\exp\left[i2\pi(t\nu - f\tau)\right]dfdt. \qquad (4)$$

Здесь ν и τ – частота и время голограммы; $f_{1,2} = f_0 \mp (\Delta f / 2)$, f_0 – средняя частота спектра. В общем случае прием шумовых реализаций источника осуществляется при малом входном отношении с/п на фоне пространственно-временных возмущений океанической среды. Интенсивная помеха маскирует сигнальную информацию, а неоднородности среды ее искажают. Эти факторы приводят к формированию сложной интерференционной картины, которая затрудняет или делает невозможным обнаружение и локализацию подводного источника шума. Голографическая обработка (4) позволяет в значительной мере преодолеть данные трудности, так как спектральные плотности парциальных голограмм, относящиеся к неискаженной и искаженной сигнальной информации, и помехи располагаются в различных областях на результирующей голограмме

$$F(\tau,\nu) = \sum_{j=1}^{4} F_j(\tau,\nu). \qquad (5)$$

Здесь $F(\tau,\nu)$ одна из векторно-скалярных голограмм, $F_j(\tau,\nu)$ – парциальные голограммы [44–47, 72–80]. Спектральная плотность сигнала, парциальная голограмма $F_1(\tau,\nu)$ $(j=1)$, сконцентрирована в двух узких полосах в форме отдельных фокальных пятен, вызванных интерференцией мод разных номеров [44–47]. Они размещены в первом и третьем квадрантах голограммы, если радиальная скорость источника $w < 0$ (источник приближается к приемнику), и во втором и четвертом квадрантах ($w > 0$), если он удаляется от приемника. Под радиальной скоростью источника понимается проекция скорости в направлении на приемник. При неподвижном или движущемся источнике с радиальной скоростью, равной нулю, координаты пиков фокальных пятен расположены на оси времени τ. Область локализации содержит $(M-1)$ главных максимумов с координатами (τ_μ, ν_μ), расположенными на прямой $\nu = \varepsilon\tau$ с угловым коэффициентом $\varepsilon = \nu_\mu / \tau_\mu$. Здесь M – число мод, формирующих поле, $\mu = \overline{1,(M-1)}$ – номер фокального пятна. Ближайший к началу координат максимум первого фокального пятна,

обусловленный интерференцией соседних мод, приходится на значения (τ_1, ν_1). Координаты соседнего пика, вызванного интерференцией мод номеров $(m, m+2)$, расположены в точке (τ_2, ν_2) и т.д. И, наконец, координаты самого удаленного пика, продиктованного интерференцией первой и последней моды – $\left(\tau_{(M-1)}, \nu_{(M-1)}\right)$. В точках с координатами (τ_μ, ν_μ) суммируются $(M-\mu)$ главных максимумов.

При условии $r_0 \gg |w|\Delta t$, где r_0 – удаление источника от приемника в начальный момент времени $t = 0$, радиальная скорость и начальное расстояние связаны с координатами максимумов фокальных пятен зависимостями

$$\dot{w} = -2\pi\kappa_{w\mu}\nu_\mu,\ \dot{r}_0 = \kappa_{r\mu}\tau_\mu, \tag{6}$$

где

$$\kappa_{w\mu} = \left|\overline{h_{m(m+\mu)}(f_0)}\right|^{-1},\quad \kappa_{r\mu} = 2\pi\left|\overline{dh_{m(m+\mu)}(f_0)/df}\right|^{-1} \tag{7}$$

– коэффициенты, определяющие пространственные и частотные масштабы изменчивости передаточной функции волновода [97]. Здесь $h_{mn} = h_m - h_n$, h_m – невозмущенное горизонтальное волновое число (постоянная распространения) m-й моды. Восстановленные параметры источников, в отличие от их истинных значений, отмечены точкой сверху. Черта сверху означает усреднение по номерам мод.

Спектральная плотность сигнала сфокусирована в полосе, ограниченной прямыми

$$\nu_{\text{I}}(\tau) = \varepsilon\tau + \delta,\ \ \nu_{\text{II}}(\tau) = \varepsilon\tau - \delta, \tag{8}$$

площадь области фокусировки [68, 69]

$$S_s = \frac{(M-2)}{\pi(M-1)}\frac{\dot{r}_0}{\Delta t}\left|\frac{dh_{1M}(f_0)}{df}\right| + \frac{4}{\Delta t\Delta f}. \tag{9}$$

Площадь области фокусировки не зависит от радиальной скорости источника, увеличивается с возрастанием удаленности и уменьшением времени наблюдения, ширины полосы и средней частоты спектра. Здесь $\delta\nu = 1/\Delta t$ – полуширина фокальных пятен в направлении оси ν. В направлении оси τ полуширина фокальных пятен $\delta\tau = 1/\Delta f$. Вне этой полосы спектральная плотность сигнала практически подавлена.

Угловые коэффициенты прямых расположения максимумов фокальных пятен ε и интерференционных полос $\delta f / \delta t$ связаны соотношением [44–47, 60]

$$\varepsilon = -\delta f / \delta t, \tag{10}$$

где δf – частотный сдвиг максимума волнового поля за время δt.

Спектральная плотность изотропной помехи, парциальная голограмма $F_2(\tau, \nu)$ $(j = 2)$, распределена по всей области голограммы, площадь которой оценивается как

$$S_n = \left(\frac{|\dot{w}|}{2\pi} \left| h_{1M}(f_0) \right| + \frac{1}{\Delta t} \right) \left(\frac{\dot{r}_0}{2\pi} \left| \frac{dh_{1M}(f_0)}{df} \right| + \frac{1}{\Delta f} \right). \tag{11}$$

Спектральная плотность возмущенного поля, парциальная голограмма $F_3(\tau,)$ $(j = 3)$, вызванная пространственно-временной изменчивостью океанической среды, концентрируется вдоль оси частоты ν во временном интервале $2\delta\tau$. С увеличением времени наблюдения она возрастает и может превосходить спектральную плотность невозмущенного поля. Спектральная плотность парциальной голограммы $F_4(\tau, \nu)$ $(j = 4)$, вызванная интерференцией невозму–щенного и возмущенного волновых полей источника, невелика и ее можно рассматривать как некий слабоконтрастный размазанный фон.

Таким образом, очищая голограмму от парциальных голограмм, обусловленных помехой и возмущенным полем, можно увеличить помехоустойчивость обработки и повысить ее устойчивость по отношению к изменчивости океанической среды. Очищение голограммы не требует предварительного знания о характере сигнала, помехи и неоднородностей акватории. Важной характеристикой голограммы является угловое распределение спектральной плотности (функция обнаружения), определяющая дальность обнаружения источника

$$G(\chi) = \int_0^{\Delta\tau} \left| F(\tau, \chi\tau) \right| d\tau, \tag{12}$$

где $\Delta\tau$ – линейный размер области фокусировки по оси времени τ; χ – варьируемое значение углового коэффициента при интегрировании вдоль прямой $\nu = \chi\tau$. При приближении источника к приемнику $0 \le \chi < \infty$, в случае удаления – $-\infty < \chi \le 0$. В отсутствии помехи и возмущения среды максимум функции обнаружения достигается при значении $\chi = \varepsilon$.

Влияние помехи и ИВВ на формирование голограммы отражено в работах [44–47, 72–80]. Проиллюстрируем сказанное на двух примерах.

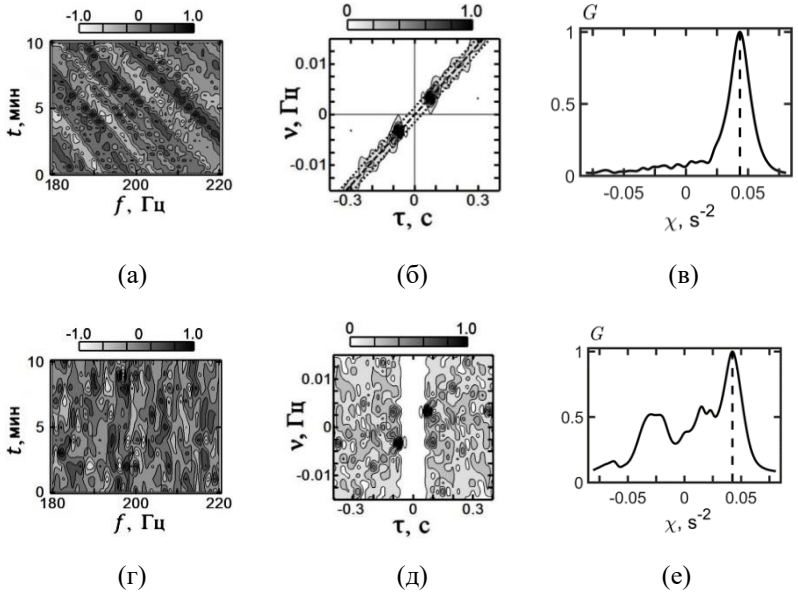


Рис. 1. Нормированные интерферограммы (а, г), модули голограмм (б, д) и функции обнаружения (в, е): (а, б, в) – $q_0 = 25.5$ (14.1 дБ); (г, д, е) – $q_0 = 3 \cdot 10^{-3}$ (-25.2 дБ) [44].

На рис. 1 приведены экспериментальные результаты голографической обработки сигнала пневмоисточника при приближении к одиночному приемнику при разных входных отношениях с/п q_0 в отсутствие возмущения среды [44]. Время накопления $\Delta t = 10$ мин, частотный диапазон $f = 180 - 220$ Гц. С целью повышения контрастности и информативности на голограммах вырезаны средние значения, располагающиеся вдоль оси частоты. На рис. 1б пунктиром и квадратами показаны прямая $\nu = \varepsilon\tau$ и прямые $\nu_{I,\,II}$ (8). На рис. 1(в, е) вертикальным пунктиром отмечено положение максимума.

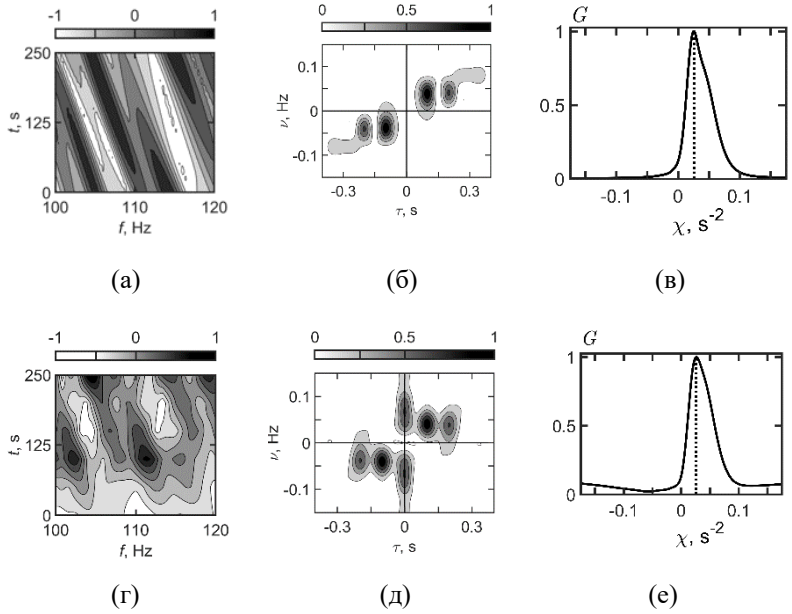


Рис. 2. Нормированные интерферограммы (а, г), модули голограмм (б, д) и функции обнаружения (в, е): (а, б, в) – в отсутствие ИВВ, (г, д, е) – в присутствии ИВВ [66].

На рис. 2 показано, полученное численным моделированием, влияние ИВВ на голографическую обработку шумового сигнала в отсутствие помехи [66]. Время наблюдения $\Delta t = 250$ с, частотный диапазон $f = 100 - 120$ Гц.

В отсутствии помехи и пространственно-временных неоднородностей среды регистрируемые интерферограмма и голограмма сигнала не искажаются, если параметры обработки, полоса Δf и время наблюдения Δt, удовлетворяют условиям [54, 57–59]

$$\Delta f > \Lambda_f^{(mn)},\ \Delta t > \Lambda_t^{(mn)}, \tag{13}$$

где

$$\Lambda_f^{(mn)} = \frac{2\pi}{r\left|dh_{mn}\left(f_0\right)/df\right|},\quad \Lambda_t^{(mn)} = \frac{1}{\left|wh_{mn}\left(f_0\right)\right|}. \tag{14}$$

– частотные и временные масштабами изменчивости интерферограммы, обусловленные интерференцией m-й и n-й мод [97].

Неравенства (13) накладывают снизу ограничения на ширину полосы приема и время наблюдения в зависимости от расстояния, радиальной скорости и частотно-временных масштабов изменчивости передаточной функции волновода. Увеличение ширины полосы и уменьшение средней частоты спектра вызывают уменьшение расстояния, на котором формируемые интерферограмма и голограмма не искажаются. Увеличение времени наблюдения уменьшает допустимые значения радиальной скорости. Нарушение условий (13) приводит к изменению конфигурации и размытию фокальных пятен, сопровождающееся увеличением погрешности восстановления параметров источника. Неравенства (13) наиболее критичны по отношению к соседним модам, формирующим первое фокальное пятно. При невыполнении первого (левого) условия локализация полос интерферирующих (m,n) мод не наблюдается. Иная закономерность проявляется при несоблюдении второго (правого) условия, когда время наблюдения недостаточно для пересечения источником пространственных масштабов изменчивости интерферограммы. В этом случае положение пика фокального пятна таких интерферирующих мод смещается на ось времени голограммы, так что по отношению к ним интерферограмма формируется как бы неподвижным источником. Восстановление параметров источника проводится по отношению к неискаженной конфигурации фокальных пятен. В рамках численного моделирования искажения голограммы исследовались в [54, 57, 58].

Заключение

За короткий отрезок времени ГОС превратилась в мощный исследовательский метод контроля подводной обстановки при малом входном отношении с/п и в условиях изменчивой гидрологической обстановки и отсутствия надежной информации о рельефе дна и характеристиках грунта, позволяющей существенным образом повысить эффективность обнаружения и идентификации подводных шумовых источников. Причина столь бурного развития – уникальные свойства ГОС, свойства, которые во многих случаях не доступны другим методам обработки гидроакустических сигналов. ГОС с применением модулей из ВСП самым существенным образом видоизменила регистрацию и обработку акустических полей шумовых источников. Оказалось возможным установить в самом общем случае простые и прозрачные связи между измеряемыми характеристиками голограммы и восстанавливаемыми параметрами источников. Установлено

минимальное входное отношение с/п, при превышении которого восстанавливаемые параметры источника близки реальным значениям. Тем самым появилась возможность вполне общего и регулярного применения ГОС в решении задач обнаружения слабых и искаженных сигналов, а также определения параметров движения шумовых источников. Важность физических и прикладных задач, которые можно решать, используя ГОС, оставляют мало сомнений в перспективности данного направления.

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

Список литературы

1. Jackson D.R., Ewart T.E. The effect of internal waves on matched-field processing // J. Acoust. Soc. Am. 1994. Vol. 96. No. 5. P. 2945–2955.
2. Baggeroer A.B., Kuperman W.A., Schmidt H. Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem // J. Acoust. Soc. Am. 1998. Vol. 83. No. 2. P. 571–587.
3. Dosso S.E., Nielsen P.L., Wilmut M.J. Data error covariance in matched-field geoacoustic inversion // J. Acoust. Soc. Am. 2006. Vol. 119. No. 1. P. 208–219.
4. Baggeroer A.B. Why did applications of MFP fail, or did we not understand how to apply MFP? // Proc. 1st Int. Conf. and Exhibition “Underwater Acoustics” / Eds. Papadakis J.S., Bjorno L. Corfu Island, Greece, 2013. P. 41–49.
5. Малышкин Г.С., Сидельников Г.Б. Оптимальные и адаптивные методы обработки гидроакустических сигналов (Обзор) // Акуст. журн. 2014. Т. 60. № 5. С. 526–545.
6. Сазонтов А.Г., Малеханов А.И. Согласованная пространственная обработка сигналов в подводных звуковых каналах (Обзор) // Акуст. журн. 2015. Т. 61. № 2. С. 233–253.
7. Сазонтов А.Г. Локализация источника в переменном по трассе волноводе в условиях неполной информации о пространственной изменчивости среды распространения // Акуст. журн. 2022. Т. 68. № 6. С. 689–696.
8. Малеханов А.И., Смирнов И.П. Пространственная обработка акустических сигналов в каналах мелкого моря в условиях априорной неопределенности: оценки потерь эффективности // Акуст. журн. 2022. Т. 68. № 4. С. 427–439.
9. Малеханов А.И., Смирнов А.В. Влияние априорной неопределенности модели звукового канала мелкого моря на

коэффициент усиления вертикальной антенной решетки // Акуст. журн. 2023. Т. 69. № 5. С. 542–558.

10. Weston D., Stevens K. Interference of wide-band sound in shallow water // J. Sound Vibr. 1972. Vol. 21. No. 1. P. 57–64.

11. Чупров С.Д. Интерференционная структура звукового поля в слоистом океане // Акустика океана. Современное состояние. М.: Наука, 1982. С. 71–82.

12. Орлов Е.Ф. Интерференционная структура широкополосного звука в океане // Проблемы акустики океана. М.: Наука, 1984. С. 85–93.

13. Грачев Г.А. К теории инвариантов акустического поля в слоистых волноводах // Акуст. журн. 1993. Т. 39. № 1. С. 67–71.

14. D'Spain G.L., Kuperman W.A. Application of waveguide invariants to analysis of spectrograms from shallow water environments that vary in range and azimuth // J. Acoust. Soc. Am. 1999. Vol. 106. No. 5. P. 2454–2468.

15. Hodgkiss W., Song H., Kuperman W., Akal T., Ferla C., Jackson D. A long-range and variable focus phase-conjugation experiment in shallow water // J. Acoust. Soc. Am. 1999. Vol. 105. No. 3. P. 1597–1604.

16. Yang T.C. Motion compensation for adaptive horizontal line array processing // J. Acoust. Soc. Am. 2003. Vol. 113. No. 1. P. 245–260.

17. Thode A.M. Source ranging with minimal environmental information using a virtual receiver and waveguide invariant theory // J. Acoust. Soc. Am. 2000. Vol. 108. No. 4. P. 1582–1594.

18. Cockrell K.L., Schmidt H. Robust passive range estimation using the waveguide invariant // J. Acoust. Soc. Am. 2010. Vol. 127. No. 5. P. 2780–2789.

19. Xiao-Le G., Kun-De Y., Yuan-Liang M., Qiu-Long Y. A source range and depth estimation method based on modal dedispersion transform // Acta Phys. Sin. 2016. Vol. 65. No. 21. P. 214302.

20. Song H.C., Byun G. Extrapolating Green's functions using the waveguide invariant theory // J. Acoust. Soc. Am. 2020. Vol. 147. No. 4. P. 2150–2158.

21. Byun G., Song H.C. Adaptive array invariant // J. Acoust. Soc. Am. 2020. Vol. 148. No. 2. P. 925–933.

22. Song H.C., Byun G. An overview of array invariant for source-range estimation in shallow water // J. Acoust. Soc. Am. 2022. Vol. 151. No. 4. P. 2336–2352.

23. Li P., Wu Y., Ma Y., Cao C., Leng H., Zhou A., Song J. Prefiltered striation-based beam forming for range estimation of multiple sources // J. Mar. Sci. Eng. 2023. Vol. 11. No. 8. P. 1550.

24. Ocean acoustic interference phenomena and signal processing (San Francisco, CA, May 1–3, 2001; AIP Conf. Proc.) / Eds. Kuperman W.A., D'Spain G.L. N.Y.: Melville, 2002.

25. Heaney K.D. Rapid geoacoustic characterization using a surface ship of opportunity // IEEE J. Oceanic Eng. 2004. Vol. 29. No. 1. P. 88–99.

26. Duan R., Yang K., Li H., Yang Q., Wu F., Ma Y. A performance study of acoustic interference structure applications on source depth estimation in deep water // J. Acoust. Soc. Am. 2019. Vol. 145. No. 2. P. 903–916.

27. Li X., Sun C. Source depth discrimination using intensity striations in the frequency–depth plane in shallow water with a thermocline // Remote Sens. 2024. Vol. 16. P. 639.

28. Harrison C.H. The relation between the waveguide invariant, multipath impulse response, and ray cycles // J. Acoust. Soc. Am. 2011. Vol. 129. No. 5. P. 2863–2877.

29. Pang J., Gao B. Application of a randomized algorithm for extracting a shallow low-rank structure in low-frequency reverberation // Remote Sens. 2023. Vol. 15. P. 3648.

30. Li P., Wu Y., Guo W., Cao C., Ma Y., Li L., Leng H., Zhou A., Song J. Striation-based beam forming with two-dimensional filtering for suppressing tonal interference // J. Mar. Sci. Eng. 2023. Vol. 11. No. 11. P. 2117.

31. Кузькин В.М., Переселков С.А. Интерферометрическая диагностика гидродинамических возмущений мелкого моря. М.: ЛЕНАНД, 2019. 200 с.

32. Гордиенко В.А. Векторно-фазовые методы в акустике. М.: ФИЗМАТЛИТ, 2007. 480 с.

33. Nehorai A., Paldi E. Acoustic vector sensor array processing // IEEE Trans. Signal Process. 1994. Vol. 42. No. 9. P. 2481–2491.

34. Смарышев М.Д. О помехоустойчивости гидроакустического комбинированного приемника // Акуст. журн. 2005. Т. 51. № 4. С. 558–559.

35. D'Spain G.L., Luby J.C., Wilson G.R., Gramann R.A. Vector sensors and vector sensor line arrays: Comments on optimal array gain and detection // J. Acoust. Soc. Am. 2006. Vol. 120. No. 1. P. 171–185.

36. Гордиенко В.А., Гордиенко Е.Л., Краснописцев Н.А., Некрасов В.Н. Помехоустойчивость гидроакустических приемных систем, регистрирующих поток акустической мощности // Акуст. журн. 2008. Т. 54. № 5. С. 774–785.

37. Глебова Г.М., Кузнецов Г.Н., Шимко О.Е. Векторно-скалярные шумовые поля, образованные взволнованной морской поверхностью // Акуст. журн. 2013. Т. 59. № 4. С. 508–519.

38. Селезнев И.А., Глебова Г.М., Жбанков Г.А., Мальцев А.М., Харахашьян А.М. Вероятностные характеристики обнаружения сигналов одиночным скалярно-векторным модулем // Подводные исследования и робототехника. 2016. № 2 (22). С. 44–49.

39. Белов А.И., Кузнецов Г.Н. Оценка акустических характеристик поверхностных слоев морского дна с использованием четырехкомпонентных векторно-скалярных приемников // Акуст. журн. 2016. Т. 62. № 2. С. 194–202.

40. Cao J., Liu J., Wang J., Lai X. Acoustic vector sensor: Reviews and future perspectives // IET Signal Process. 2017. Vol. 11. P. 1–9.

41. Shi J., Dosso S.E., Sun D., Liu Q. Geoacoustic inversion of the acoustic-pressure vertical phase gradient from a single vector sensor // J. Acoust. Soc. Am. 2019. Vol. 146. No. 5. P. 3159–3173.

42. Wang W., Li X., Zhang K., Shi J., Shi W., Ali W. Robust direction finding via acoustic vector sensor array with axial deviation under non-uniform noise // J. Mar. Sci. Eng. 2022. Vol. 10. No. 9. P. 1196.

43. Bozzi F.A., Jesus S.M. Vector sensor steering-dependent performance in an underwater acoustic communication field experiment // Sensors. 2022. Vol. 22. P. 8332.

44. Chen Y., Zhang G., Wang R., Rong H., Yang B. Acoustic vector sensor multi-source detection based on multimodal fusion // Sensors. 2023. Vol. 23. P. 1301.

45. Rashid R., Zhang E., Abdi A. Underwater acoustic signal acquisition and sensing using a ring vector sensor communication receiver: theory and experiments // Sensors. 2023. Vol. 23. P. 6917.

46. Zhang Q., Da L., Wang C., Yuan M., Zhang Y., Zhuo J. Passive ranging of a moving target in the direct-arrival zone in deep sea using a single vector hydrophone // J. Acoust. Soc. Am. 2023. Vol. 154. No. 4. P. 2426–2439.

47. Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Спектрограмма и локализация источника звука в мелком море // Акуст. журн. 2017. Т. 63. № 4. С. 406–418.

48. Kuznetsov G.N., Kuz'kin V.M., Pereselkov S.A., Kaznacheev I.V. Noise source localization shallow water // Phys. Wave Phenom. 2017. Vol. 25. No. 2. P. 156–163.

49. Kuznetsov G.N., Kuz'kin V.M., Pereselkov S.A., Kaznacheev I.V., Grigor'ev V.A. Interferometric method for estimating the velocity of a noise sound source and the distance to it in shallow water using a vector-scalar receiver // Phys. Wave Phenom. 2017. Vol. 25. No. 4. P. 299–306.

50. Казначеев И.В., Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Интерферометрический метод обнаружения движущегося источника

звука векторно-скалярным приемником // Акуст. журн. 2018. Т. 64. № 1. С. 33–45.

51. Kuz'kin V.M., Pereselkov S.A., Kuznetsov G.N., Kaznacheev I.A. Interferometric direction finding by a vector-scalar receiver // Phys. Wave Phenom. 2018. Vol. 26. No. 1. P. 63–73.

52. Kuznetsov G.N., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Prosovetskiy D.Yu. Direction finding of a noise sound source // Phys. Wave Phenom. 2019. Vol. 27. No. 3. P. 237–241.

53. Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Хворостов Ю.А. Голографический метод локализации шумового подводного источника в мелководной акватории // Доклады РАН. Физика, технические науки. 2024. Т. 515. № 2. С. 80–84.

54. Glushchenko M.Yu., Kuz'kin V.M., Matvienko Yu.V., Pereselkov S.A., Khvorostov Yu.A., Tkachenko S.A. Holographic method for localizing an underwater noise source in a shallow sea // Acoust. Phys. 2024. Vol. 70. No. 3. P. 494–502.

55. Беседина Т.Н., Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Определение глубины источника звука в мелком море на фоне интенсивного шума // Акуст. журн. 2015. Т. 61. № 6. С. 718–728.

56. Kuznetsov G.N., Kuz'kin V.M., Pereselkov S.A., Prosovetskiy D.Yu. Wave method for estimating the sound source depth in a oceanic waveguide // Phys. Wave Phenom. 2016. Vol. 24. No. 4. P. 310–316.

57. Казначеева Е.С., Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Хворостов Ю.А. Оценка дальности обнаружения малогабаритного подводного аппарата по его шумовому полю // Подводные исследования и робототехника. 2021. № 4 (38). С. 88–93.

58. Хворостов Ю.А., Матвиенко Ю.В. Характеристики собственного шумоизлучения малогабаритного АНПА // Подводные исследования и робототехника. 2019. № 4 (30). С. 58–63.

59. Хворостов Ю.А., Матвиенко Ю.В., Кузькин В.М., Пересёлков С.А., Ткаченко С.А. Спектральные характеристики шумового поля малогабаритного автономного необитаемого подводного аппарата в дальней зоне // Подводные исследования и робототехника. 2022. № 4 (42). С. 84–88.

60. Kuz'kin V.M., Pereselkov S.A., Tkachenko S.A., Matvienko Yu.V., Khvorostov Yu.A. Range of detection of underwater sound sourse // Phys. Wave Phenom. 2023. Vol. 31. No. 5. P. 339–345.

61. Кузькин В.М., Пересёлков С.А., Грачев В.И., Рыбянец П.В., Ткаченко С.А. Предельная дальность обнаружения подводного шумового источника с применением голографической обработки // РЭНСИТ. 2024. Т. 16. № 1. С. 129–136.

62. Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Рыбянец П.В. О возможности оценки предельной дальности обнаружения подводных глайдеров // Подводные исследования и робототехника. 2024. № 2 (48). С. 68–75.

63. Ehrhardt M., Pereselkov S., Kuz'kin V., Kaznacheev I., Rybyaners P. Experimental observation and theoretical analysis of the low-frequency source interferogram and hologram in shallow water // J. Sound Vib. 2023. Vol. 544. P. 117388.

64. Пересёлков С.А., Кузькин В.М., Кузнецов Г.Н., Просовецкий Д.Ю., Ткаченко С.А. Интерференционный метод оценки координат движущегося шумового источника в мелком море с использованием высокочастотных сигналов // Акуст. журн. 2020. Т. 66. № 4. С. 437–445.

65. Матвиенко Ю.В., Хворостов Ю.А., Каморный А.В., Глущенко М.Ю., Кузькин В.М., Пересёлков С.А. Экспериментальные исследования системы обнаружения малошумных подводных целей в мелководных акваториях // Подводные исследования и робототехника. 2022. № 3 (41). С. 4–14.

66. Pereselkov S., Kuz'kin V., Ehrhardt M., Matvienko Yu., Tkachenko S., Rybyanets P. The formation of 2D holograms of a noise source and bearing estimation by a vector scalar receiver in the high-frequency band // J. Mar. Sci. Eng. 2024. Vol. 12. P. 704.

67. Kaznacheeva E.S., Kuznetsov G.N., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A. Measurement capability of the interferometric method of sound source localization in the absence of data on the waveguide transfer function // Phys. Wave Phenom. 2019. Vol. 27. No. 1. P. 73–78.

68. Kaznacheeva E.S., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Tkachenko S.A. Adaptive algorithms for interferometric processing // Phys. Wave Phenom. 2020. Vol. 28. No. 3. P. 267–273.

69. Кузькин В.М., Пересёлков С.А., Грачев В.И., Рыбянец П.В., Ткаченко С.А. Адаптивный алгоритм локализации шумовых подводных источников в мелководных акваториях // РЭНСИТ. 2024. Т. 16. № 4. С. 525–536.

70. Kuz'kin V.M., Kuznetsov G.N., Pereselkov S.A., Grigor'ev V.A. Resolving power of the interferometric method of source localization // Phys. Wave Phenom. 2018. Vol. 26. No. 2. P. 150–159.

71. Кузькин В.М., Переселков С.А., Грачев В.И., Рыбянец П.В., Ткаченко С.А. Разрешение шумовых сигналов при наличии случайных неоднородностей океанической среды // РЭНСИТ. 2024. Т. 16. № 3. С. 395–406.

72. Kuz'kin V.M., Pereselkov S.A., Rybyanets P.V., Tkachenko S.A. Resolution of hydroacoustic noise signals under conditions of spatial-temporal

variability of ocean environment // Phys. Wave Phenom. 2024. Vol. 32. No. 4. P. 290–299.

73. Кузькин В.М., Пересёлков С.А., Матвиенко Ю.В., Ткаченко С.А. Обнаружение источника в диспергирующих средах // РЭНСИТ. 2019. Т. 11. № 3. С. 337–344.

74. Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Matvienko Yu.V., Tkachenko S.A. Noise-source detection in an oceanic waveguide using interferometric processing // Phys. Wave Phenom. 2020. Vol. 28. No. 1. P. 68–74.

75. Kuz'kin V.M., Pereselkov S.A., Zvyagin V.G., Malykhin A.Yu., Prosovetskiy D.Yu. Intense internal waves and their manifestation in interference patterns of received signals on oceanic shelf // Phys. Wave Phenom. 2018. Vol. 26. No. 2. P. 160–167.

76. Badiey M., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Prosovetskiy D.Yu., Tkachenko S.A. Intense internal waves and their manifestation in the interference patterns of received signals on oceanic shelf. Part II // Phys. Wave Phenom. 2019. Vol. 27. No. 4. P. 313–319.

77. Badiey M., Кузькин В.М., Пересёлков С.А. Интерферометрия гидродинамики океанического шельфа, вызванной интенсивными внутренними волнами // Фундам. прикл. гидрофиз. 2020. Т. 13. № 1. С. 45–55.

78. Kaznacheeva E.S., Kuz'kin V.M., Pereselkov S.A. Interferometric processing of hydroacoustic information in the presence of intense internal waves // Phys. Wave Phenom. 2021. Vol. 29. No. 3. P. 278–284.

79. Кузькин В.М., Ляхов Г.А., Пересёлков С.А., Казначеева Е.С. Передача информации через случайно-неоднородную океаническую среду // Фундам. прикл. гидрофиз. 2021. Т. 14. № 2. С. 54–64.

80. Кузькин В.М., Пересёлков С.А., Казначеева Е.С., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Голографическая обработка движущихся источников в мелком море при наличии интенсивных внутренних волн // РЭНСИТ. 2022. Т. 14. № 2. С. 197–204.

81. Pereselkov S.A., Kuz'kin V.M. Interferometric processing of hydroacoustic signals for the purpose of source localization // J. Acoust. Soc. Am. 2022. Vol. 151. No. 2. P. 666–676.

82. Кузькин В.М., Пересёлков С.А., Грачев В.И., Ткаченко С.А., Ладыкин Н.В., Куцов М.В. Голографический метод локализации движущегося источника звука в присутствии интенсивных внутренних волн // РЭНСИТ. 2023. Т. 15. № 3. С. 317–326.

83. Kuz'kin V.M., Pereselkov C.A., Badiey M., Ladykin N.V., Malykhin A.Yu., Tkachenko S.A. Robustness of holographic processing of

hydroacoustic signals in the presence of intense internal waves // Phys. Wave Phenom. 2023. Vol. 31. No. 5. P. 346–354.

84. Kuz'kin V.M., Matvienko Yu.V., Pereselkov S.A., Prosovetskii D.Yu., Kaznacheeva E.S. Mode selection in oceanic waveguides // Phys. Wave Phenom. 2022. Vol. 30. No. 2. P. 111–118.

85. Kuz'kin V.M., Matvienko Yu.V., Pereselkov S.A., Kaznacheeva E.S., Tkachenko S.A. Holographic method for mode selection in a shallow sea in the presence of intense internal waves // Phys. Wave Phenom. 2022. Vol. 30. No. 5. P. 314–320.

86. Кузькин В.М., Пересёлков С.А., Казначеева Е.С., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Выделение мод шумового источника в мелком море методом голографической интерферометрии в присутствии интенсивных внутренних волн // РЭНСИТ. 2022. Т. 14. № 3. С. 279–286.

87. Кузькин В.М., Пересёлков С.А., Ткаченко С.А., Казначеев И.В. Интерферометрическая обработка акустической информации с использованием протяженных антенн в диспергирующих средах // РЭНСИТ. 2020. Т. 12. № 4. С. 483–494.

88. Kaznacheev I.V., Kuz'kin V.M., Kutsov M.V., Lyakhov G.A., Pereselkov S.A. Interferometry in acoustic-data processing using extended antennas: Space-time analogy // Phys. Wave Phenom. 2020. Vol. 28. No. 4. P. 326–332.

89. Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Ткаченко С.А. Интерферометрическая обработка с использованием вертикальной линейной антенны // Вестн. ВГУ. Серия: Физика. Математика. 2020. № 2. С. 14–23.

90. Kuz'kin V.M., Pereselkov S.A., Matvienko Yu.V., Tkachenko S.A. Detection of a noise signal in an oceanic waveguide using a vertical array // Phys. Wave Phenom. 2021. Vol. 29. No. 4. P. 323–329.

91. Кузькин В.М., Пересёлков С.А., Матвиенко Ю.В., Ткаченко С.А., Просовецкий Д.Ю. Обнаружение шумового сигнала в океаническом волноводе горизонтальной антенной // РЭНСИТ. 2022. Т. 14. № 1. С. 65–72.

92. Кузькин В.М., Пересёлков С.А., Матвиенко Ю.В., Грачев В.И., Ткаченко С.А., Стадная Н.П. Голографическая обработка гидроакустической информации с применением линейных антенн // РЭНСИТ. 2023. Т. 15. № 2. С. 169–178.

93. Пересёлков С.А., Кузькин В.М., Матвиенко Ю.В. Робототехнический комплекс обнаружения и локализации подводных источников шума // Вестник Российского фонда фундаментальных исследований. 2024. № 1 (121). С. 101–111.

94. Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Щербатюк А.Ф. Концепция создания мобильных средств обнаружения малошумных подводных источников с использованием векторно-скалярных антенн, переносимых группой «AUV» // Международная конференция по изучению океана (ICOS), Владивосток, Российская Федерация / IEEE Xplore. 2024. С. 197–200.

95. Переселков С.А., Кузькин В.М. Голографическая обработка широкополосных гидроакустических сигналов. Воронеж: Издательский дом ВГУ, 2023. 183 с.

96. Интерференция широкополосного звука в океане / Сб. науч. тр. / Отв. ред. Зверев В.А., Орлов Е.Ф. Горький: Изд. ИПФ АН СССР, 1984. 213 с.

97. Орлов Е.Ф., Фокин В.Н., Шаронов Г.А. Исследование параметров интерференционной модуляции широкополосного звука в глубоком океане // Акуст. журн. 1988. Т. 34. № 5. С. 902–907.

98. Лазарев В.А., Орлов Е.Ф., Фокин В.Н., Шаронов Г.А. Частотная зависимость параметров интерференционной модуляции широкополосного звука в мелком море // Акуст. журн. 1989. Т. 35. № 4. С. 685–688.

99. Орлов Е.Ф., Шаронов Г.А. Интерференция звуковых волн в океане. Владивосток: Дальнаука, 1998. 195 с.

100. Зверев В.А. Принцип акустического обращения волн и голография // Акуст. журн. 2004. Т. 50. № 6. С. 792–801.

101. Зверев В.А. Обращение волнового фронта для снижения влияния многолучевости на результат активной локации // Акуст. журн. 2003. Т. 49. № 6. С. 814–819.

102. Зверев В.А. Получение изображений акустической антенной через слой неоднородностей // Акуст. журн. 2004. Т. 50. № 1. С. 62–67.

103. Зверев В.А. О корреляционном шуме при обращении волнового фронта в акустике // Акуст. журн. 2004. Т. 50. № 5. С. 614–618.

104. Зверев В.А. Формирование изображений акустических источников в мелком море. Нижний Новгород: ИПФ РАН, 2019. 112 с.

105. Emmetiere R., Bonnel J., Gehant M., Cristol X., Chonavel T. Understanding deep-water striation patterns and predicting the waveguide invariant as a distribution depending on range and depth // J. Acoust. Soc. Am. 2018. Vol. 143. No. 6. P. 3444–3454.

106. Emmetiere R., Bonnel J., Cristol X., Gehant M., Chonavel T. Passive source depth discrimination in deep-water // IEEE J. Sel. Top. Signal Process. 2019. Vol. 13. No. 1. P. 185–197.

Секция 1
Методы и технологии разработки программных систем

Преимущества микрофронтенд архитектуры для разработки веб-приложений

П. В. Волченко, e-mail: volchenko.coffee@gmail.com
Н. К. Самойлов, e-mail: nk.samoylov@gmail.com

Воронежский государственный университет

Аннотация. *Рост популярности веб-приложений сопровождается увеличением сложности и фрагментации в создании их пользовательских интерфейсов. В ответ на эти вызовы в последние годы назрела необходимость в новых подходах, способных упорядочить процесс разработки. Особенно остро эта задача стоит перед крупными проектами, где критически важны масштабируемость и модульность. Оценив успешные результаты микросервисной архитектуры в бэкенд-разработке, специалисты смогли найти подобный подход для фронтенда, создав микрофронтенды как модульное решение для построения интерфейсов. В данной статье анализируются ключевые преимущества микрофронтендной архитектуры: она способствует повышению модульности и масштабируемости приложений, а также обеспечивает командам разработки большую автономию и гибкость в работе.*

Ключевые слова: *веб-приложения, модульность, микросервисная архитектура, масштабируемость.*

Введение

За последние годы веб-разработка совершила значительный скачок, что породило всё более высокие ожидания в отношении модульности, масштабируемости и удобства поддержки приложений. Однако с ростом функциональной насыщенности веб-проектов традиционные монолитные архитектуры всё чаще демонстрируют свою несостоятельность в удовлетворении этих потребностей. Так появилась концепция модульности, реализованная в виде новой архитектурной парадигмы — микрофронтендов. В современной практике эта модель рассматривается как одно из самых многообещающих направлений. Основной принцип микрофронтендов состоит в декомпозиции пользовательского интерфейса на множество компактных автономных компонентов, каждый из которых реализует отдельный функциональный блок или элемент интерфейса. Этот подход не только обеспечивает модульность кода, но и трансформирует организацию работы команд:

© Волченко П. В., Самойлов Н. К., 2026

разные группы разработчиков могут параллельно и независимо трудиться над отдельными частями приложения. Дополнительным преимуществом является возможность изолированного развертывания и масштабирования каждого модуля, что существенно упрощает адаптацию приложения под разнообразные пользовательские сценарии и требования.

1. Необходимость в введении микрофронтедов

Концепция микросервисной архитектуры была впервые предложена в 2011 году [1] в сфере бэкенд-разработки. В рамках данного подхода каждый сервис рассматривается как небольшая, самостоятельная система, работающая независимо от других частей. Идея получила широкое признание и положительный отклик как со стороны разработчиков, так и со стороны смежных команд, заинтересованных в повышении эффективности и производительности сервисов. Опираясь на достигнутый успех микросервисов, в 2016 году [2] был введён термин «микрофронтенд» — концепция, переносящая те же принципы модульности и независимости на уровень клиентской части приложения.

Микрофронтенд — это архитектурный подход, при котором веб-приложение разбивается на независимые, автономные модули. Каждый такой модуль разрабатывается, тестируется и развертывается отдельной командой, используя собственный технологический стек при необходимости.

Микрофронтенды появились как инновационное решение, где не справлялись традиционные монолитные архитектуры при работе со сложными современными веб-приложениями. Однако, как и любое кардинальное изменение в подходах к разработке, при переходе на эту архитектуру разработчики столкнулись с определёнными трудностями. Одним из вопросов, на который до сих пор не найден ответ является определение оптимального размера микрофронтенда. Размер и границы модулей определяются индивидуально для каждого проекта с учётом множества факторов: уровня сложности приложения, специфики его использования, структуры и количества команд разработки, а также контекста проекта. На рис. 1 наглядно продемонстрированы различия между монолитной архитектурой фронтенда и подходом на основе микрофронтендов.

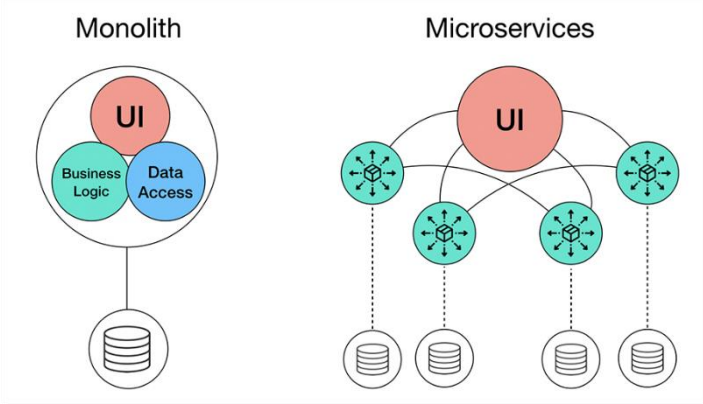


Рис. 1. Структура монолитного фронтенда и микрофронтендов

Однако применение данного архитектурного подхода не всегда нужно. Использование микрофронтендов будет уместно в крупных проектах, разрабатываемых разными командами, где нужна поддержка масштабируемости его, а также соблюдения общего стиля во всех его частях. Ещё можно отнести факт, что некоторые отдельные микрофронтенды, решающие разные бизнес-задачи, могут иметь разный жизненный цикл. В таком случае также уместно его использовать.

В случае, когда имеется небольшой проект, малое число команд, задействованных в нём, слабо развитая инфраструктура, то данный подход будет только вредить и мешать в разработке.

2. Использование микрофронтендов

С момента появления микрофронтендной архитектуры, она вызвала интерес у многих разработчиков. В результате, начали появляться фреймворки и плагины, способные сделать работу с этой архитектурой проще. Первым таким инструментом оказался фреймворк Single-Spa [3], который позволяет соединить все приложения в одном. Однако такое решение требует строгого следования его правилам. Пример архитектуры Single-Spa представлена на рис. 2.

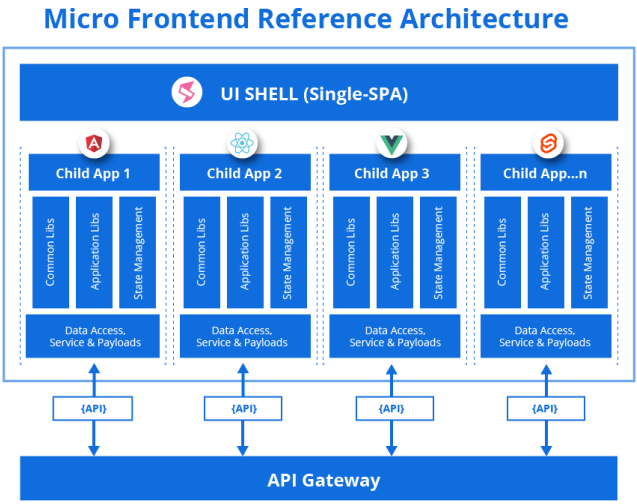


Рис. 2. Использование Single-SPA

Альтернативным подходом стало применение веб-компонентов (Web Components), выполняющих роль универсального интерфейса взаимодействия между микрофронтендами и основным приложением. Согласно данной методологии [4], веб-компоненты позволяют формировать переиспользуемые фронтенд-блоки, которые могут быть импортированы в веб-приложения, а впоследствии интегрированы в микросервисы вместе с серверной частью [5]. Дальнейшее развитие этой области получило продолжение с появлением плагина «Module Federation» для Webpack 5, который быстро зарекомендовал себя как одно из самых прогрессивных и практичных решений для устранения дублирования кода и оптимизации межмодульного взаимодействия. Пример использования WebPack с Module Federation представлен на рис. 3. Как указывается в официальной документации плагина [6], он успешно решает одну из ключевых проблем, существовавших ранее — необходимость ускоренного и эффективного обмена данными между модулями. Достигается это за счёт формирования множества независимых сборок, что позволяет подключать внешние компоненты так, будто они являются локальными, без привязки к процессам сборки или развёртывания стороннего кода.

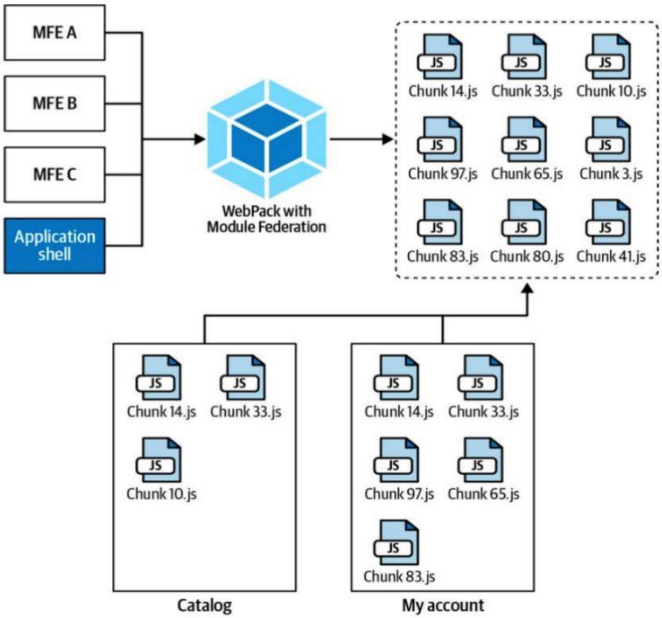


Рис. 3. WebPack с Module Federation

Что касается производительности, несмотря на то что сам факт модульности уже даёт потрясающие результаты, в этой архитектуре по-прежнему на первом месте стоит оптимизация взаимодействия и повышение общей скорости отклика приложения. Продолжаются исследования в области динамического развития модулей, гибкости и отказоустойчивости. Текущее состояние микрофронтендов в сфере масштабируемых веб-приложений демонстрирует предсказуемый рост популярности модульных фронтенд-архитектур.

Прежде чем начать реализовывать микрофронтендную архитектуру, важно понимать её базовые принципы. Как уже говорилось ранее, нет универсального стандарта, который определял бы оптимальный размер каждого микрофронтенда или их количество, так как это напрямую зависит от структуры проекта и организации команд разработчиков. Но можно выделить два основных подхода к их разделению.

Микрофронтенды можно организовать двумя способами: размещать несколько модулей на одном экране или выделять отдельный микрофронтенд для каждого экрана, отвечающего за каждый бизнес-сервис отдельно. Согласно [7], выделяют две основные стратегии структурирования микрофронтендов: горизонтальное и вертикальное

разделение. Наглядная иллюстрация этих подходов представлена на рис. 4.

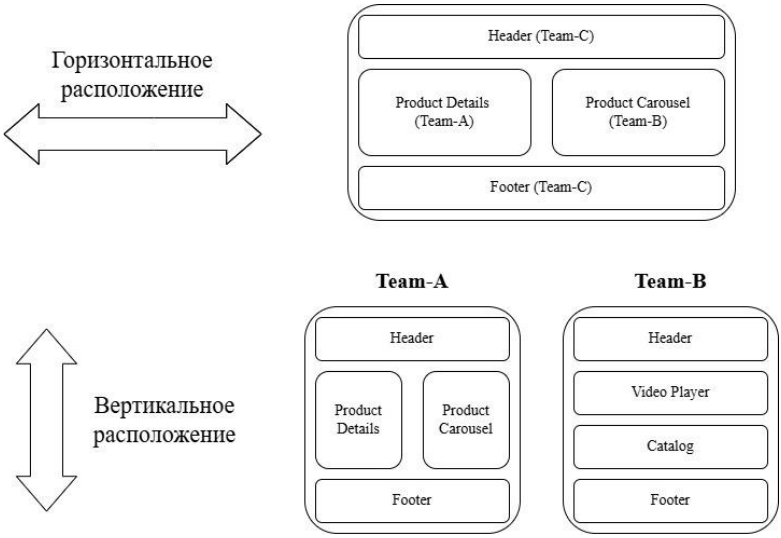


Рис. 4. Приципы расположения микрофронтендов [8]

При горизонтальном разделении одна веб-страница представляет собой композицию из нескольких небольших фронтенд-приложений. Такой подход предполагает координацию усилий нескольких команд разработки, где каждая команда отвечает за отдельную бизнес-область. Такой подход регулируется методологией предметно-ориентированного проектирования (Domain Driven Design, DDD) [9].

Однако такое разделение может оказаться неэффективным, когда команды географически рассредоточены, находясь в разных часовых зонах, и сильно зависимы друг от друга. В этом случае данный подход усложняет координацию и взаимодействие между ними.

При вертикальном разделении, также построенном по DDD, каждая команда ответственна за конкретную бизнес-область. При этом используется ограниченный контекст, согласно [10]. Она структурирует бизнес-домены и поддомены, объединяя их в обособленные логические зоны, в пределах которых формируются предметные модели, архитектура кода и, при необходимости, распределяются команды разработчиков. Каждый микрофронтенд создаётся и развёртывается автономно, что обеспечивает командам независимость в работе. При реализации таких

микрофронтендов допускается применение различных архитектурных моделей.

Данная модель предлагает стандартизированный набор правил проектирования и готовых решений, что гарантирует согласованность пользовательского интерфейса и его целостность на всех уровнях системы. Несмотря на кажущуюся простоту и очевидность названия, данный шаблон является мощным инструментом для микрофронтендной архитектуры. Разработав таким образом библиотеку переиспользуемых компонентов, её можно импортировать в различных микроприложениях, сохранив при этом единый стиль и позволяя при необходимости адаптировать элементы под конкретные задачи, о чём изображено на рис. 5. В результате значительно сокращается дублирование, связанное с поддержанием единообразного дизайна.

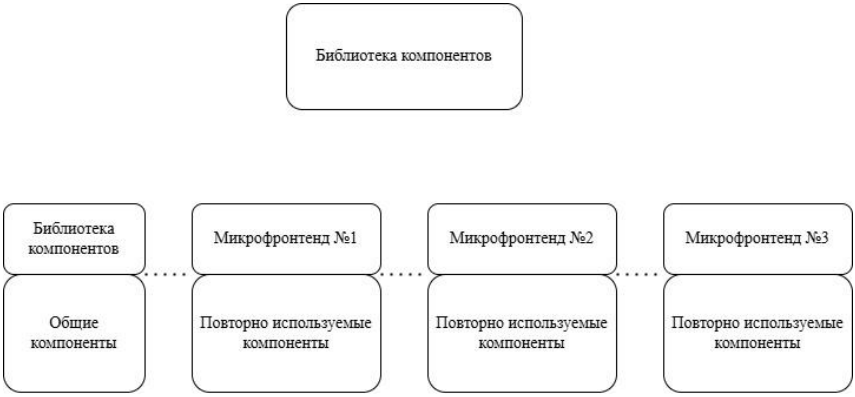


Рис. 5. Библиотека компонентов

3. Анализ результатов

Было проведено исследования оценки быстродействия веб-приложения, разработанного по данной архитектуре. Для анализа использовались временные параметры жизненного цикла, такие как: Scripting, Render, System, Total. Результаты представлены в таблице. Полученные данные демонстрируют положительную динамику данной архитектуры в оптимизации производительности: приложения имеют быстрый отклик и являются ресурсоэффективными в работе. Для анализа использовался встроенный инструмент браузера Chrome DevTools.

Таблица

Анализ результатов при использовании микрофронтендов

Продолжительность		**Описание**
Scripting	131 мс	Этот этап показывает, что использование микрофронтендов оптимизировало процесс выполнения javascript
Rendering	10 мс	Этот этап доказывает, что процесс рендеринга происходит быстрее и эффективнее
System	15 мс	Этот этап доказывает, что такая архитектура позволила сократить объем работ, связанных с системой, и повысить общую эффективность
Total	430 мс	Меньшее увеличение общего времени загрузки страницы доказывает, что архитектура микрофронтенда привела к более быстрой и отзывчивой загрузке страницы

Архитектура микрофронтендов работает по принципу «ленивой» загрузки: скрипты подгружаются динамически и только тогда, когда пользователь обращается к конкретному функциональному блоку или разделу приложения. Касаемо видимой части приложения, происходит частичный рендер страницы, что существенно снижает вычислительную нагрузку и ускоряет отображение.

Длительность рендеринга напрямую влияет на восприятие интерфейса пользователем. Чем меньше это время, тем отзывчивее интерфейс приложения. Благодаря сниженной зависимости от сторонних операций микрофронтенды демонстрируют сокращение системных задержек. В итоге общее время отклика страницы уменьшается.

Заключение

В данной статье были подробно проанализированы ключевые характеристики микрофронтендной архитектуры. От её сильных сторон до сложностей внедрения. Идея подхода заключается в декомпозиции монолитного пользовательского интерфейса на отдельные приложения, отвечающих за его разные бизнес-области, и упрощает параллельную разработку несколькими командами. Использование микрофронтендов не только оптимизирует темпы создания приложения, но и позволяет создать единую базу, для дальнейшего его расширения.

Такой подход имеет также и ряд проблем. Но их успешное решение невозможно без чёткого проектирования. Особое значение в современной веб-разработке имеет гибкость масштабирования. Микрофронтенды позволяют соблюдать данное требование, облегчая возможность расширения приложений, независимо от других его частей. Такой подход позволяет направлять вычислительные и человеческие ресурсы строго туда, где они действительно необходимы в конкретный момент времени, минимизируя издержки и повышая общую эффективность системы.

Список литературы

1. Microservices: The journey so far and challenges ahead / P. Jamshidi [et al.] // IEEE Software. – 2018. – Vol. 35. – Iss. 3. – P. 24–35. – DOI: 10.1109/MS.2018.2141039.
2. Micro frontends: Technology radar [Электронный ресурс] // Thoughtworks. – Режим доступа: https://www.thoughtworks.com/radar/techniques/micro-frontends
3. Varenyk, O. Method of Converting the Monolithic Architecture of a Front-End Application to Microfrontends / O. Varenyk, N. Khatsko, K. Khatsko // Challenges and Reality of the IT-Space: Software Engineering and Cybersecurity: International Conference SECS-2022 (October 25-26th, 2022). – Poznań: Institute of Bioorganic Chemistry. Polish Academy of Sciences. Scientific Publishers OWN, 2023. – P. 100-110.
4. Yang, C. Research and application of micro frontends / C. Yang, C. Liu, Z. Su // IOP Conference Series: Materials Science and Engineering. – 2019. – Vol. 490. – Iss. 6. – Art. 062082. – 6 p. – DOI: 10.1088/1757-899X/490/6/062082.
5. Microfrontends: Microservices for the Frontend [Электронный ресурс] // Technology Conversations. – 2015. – Режим доступа: https://semaphore.io/blog/microfrontends
6. Micro frontends [Электронный ресурс] // martinfowler.com. – Режим доступа: https://martinfowler.com/articles/micro-frontends.html
7. Peltonen, S. Motivations, benefits, and issues for adopting micro-frontends: A multivocal literature review / S. Peltonen, L. Mezzalira, D. Taibi // Information and Software Technology. – 2021. – Vol. 136. – Art. 106571. – 18 p. –DOI: 10.1016/j.infsof.2021.106571.
8. Henderson, C. Building Scalable Web Sites: Building, scaling, and optimizing the next generation of web applications / C. Henderson. – Sebastopol: O'Reilly Media, Inc., 2006. – 349 p.
9. Prajwal, Y. R. A Brief Review of Microfrontend. / Y. R. Prajwal, J. V. Parekh, R. Shettar // United International Journal for Research & Technology – 2021. – Vol. 2, No. 8. – P. 123-126.
10. Rappl, F. Art of Micro Frontends: Build Websites Using Compositional UIs That Grow Naturally as Your Application Scales / F. Rappl. – Birmingham: Packt Publishing Limited, 2021. – 310 p.

Сравнительный анализ подходов к сбору и формализации требований к программному обеспечению

И. Е. Воронина, email: irina.voronina@gmail.com
П. Г. Полякова, email: poliakova_polya@mail.ru

Воронежский государственный университет

Аннотация. *Процесс выявления и формализации требований к программному обеспечению является важной фазой жизненного цикла разработки. Несмотря на наличие множества устоявшихся методологий, проблема получения непротиворечивых, полных и стабильных требований остаётся актуальной даже в современных проектах. Приводится сравнительный анализ ключевых подходов к формализации, выявляются системные ограничения в контексте современных условий разработки. Обосновывается необходимость формирования нового подхода, который объединит принципы гибких методологий.*

Ключевые слова: *требования к программному обеспечению, сбор требований, формализация требований, гибкая разработка, документно-цетрированные, модельно-ориентированные подходы, обработка естественного языка, сквозная трассируемость.*

Введение

Сбор и формализация требований к программному обеспечению представляют собой комплексную задачу преобразования потребностей пользователей в строгие спецификации. Следует отметить, что эти потребности часто носят неструктурированный характер. Сложность этой задачи заключается в необходимости формализовать нестабильный объект. Требования могут изменяться по мере углубления понимания предметной области, изменения рыночной конъюнктуры или возникновения новых технологических ограничений.

Стоимость устранения ошибки, сделанной на этапе сбора и формализации требований, возрастает экспоненциально по мере продвижения по стадиям жизненного цикла. Таким образом, разработка эффективной методологии, способной сочетать адаптивность к изменениям и обеспечение необходимого уровня строгости, является значимой научно-практической задачей в области развития программного обеспечения.

© Воронина И. Е., Полякова П. Г., 2026

1. Анализ методологий работы с требованиями

Методологии работы с требованиями по форме описания требований и степени адаптивности можно разделить на три категории: документо-центрированные, гибкие, модельно-ориентированные.

Ранними считаются документо-центрированные подходы, такие как каскадная модель (Waterfall) или Rational Unified Process (RUP). Они базируются на принципе полной формализуемости всех пользовательских потребностей ещё до начала разработки. Эти методики предполагают последовательное выполнение этапов выявления, анализа, спецификации и верификации, результатом которых является утверждённая спецификация требований к программному обеспечению. Данный документ, созданный с применением интервью, опросов и структурированного анализа, формален и служит базой для дальнейшего проектирования. Главным преимуществом документо-центрированных подходов является создание систематизированного и детализированного набора документов. Однако, такая совокупность артефактов обладает низкой устойчивостью к постоянным изменениям. Главным недостатком становится процесс внесения корректив. Документ быстро теряет актуальность в динамично изменяемой среде, что приводит к существенному расхождению между документацией и реальной системой, а любое даже малое изменение оказывается бюрократически сложным и ресурсозатратным.

В качестве реакции на эту проблему в начале 2000-х возникли гибкие методологии Agile. Основным их отличием от более ранних является приспособленность к изменениям. Таким образом, требованиям свойственно существовать в виде приоритизированного бэклога продуктов. В гибких методологиях основной единицей работы с требованиями выступает пользовательская история, описывающая потребность с точки зрения конечной ценности. Процесс строится на коротких циклах с постоянной обратной связью от заказчика, пользовательские потребности регулярно уточняются – это обеспечивает высокую адаптивность и концентрацию на предоставлении ценности. Такой подход также нельзя назвать идеальным, так как возникает ряд системных ограничений. Аналитики сталкиваются с трудностями, касающимися управления сквозными нефункциональными требованиями, например требованиями к безопасности и масштабируемости. Существует риск отклонения от общего архитектурного видения, в связи со сложностью координации нескольких работников в крупных проектах. Именно согласованность разработки требований является необходимым элементом при проектировании масштабной системы, поэтому в последнее время наблюдается уход от

жёсткой привязки к регламентированным документам на ранних этапах в сторону адаптивных требований.

Существуют также модельно-ориентированные подходы. Они предполагают применение формальных или полуформальных нотаций для фиксации требований. Использование языков моделирования, таких как UML или SysML, позволяет представлять неоднородные потребности пользователей и заказчика в виде диаграмм состояний, активностей, последовательностей, что повышает их однозначность [1]. Подобная формализация открывает возможности для автоматизированного анализа моделей на непротиворечивость, полноту и даже для генерации прототипов кода. Это и является основным преимуществом данного подхода. С другой стороны, для построения и чтения моделей требуется высокая квалификация всех участников цепочки доставки ценности, а трудоёмкость создания и поддержки сложных моделей во многих случаях не соотносится с выгодой для проектов с высокой степенью неопределённости. Ещё одним недостатком модельно-ориентированных подходов часто является излишняя детализация, такая как описание мелких процессов с помощью моделей, что нередко приводит к оверинжинирингу [2].

2. Системные ограничения

При изолированном применении каждого из перечисленных ранее подходов становятся ясны их системные ограничения. Современные условия разработки программного обеспечения предполагают следующие вызовы:

- экстремальная динамичность бизнес-среды, требующая регулярной эмпирической проверки гипотез о потребностях пользователей;
- сложность задачи обеспечения согласованности требований на уровне всей развивающейся системы;
- необходимость интеграции разнородных компонентов и технологий (например, микросервисов, IoT, машинного обучения) для сохранения конкурентоспособности;
- большие объёмы слабоструктурированных данных (например, хранимая обратная связь от пользователей), которые невозможно обработать вручную;

затрудненная коммуникация между представителями бизнеса и техническими специалистами, вызванная разностью в осведомлённости в предметных областях друг друга.

3. Оформление таблиц

Принимая во внимание системные ограничения и недостатки рассмотренных методологий, становится очевидна потребность в некотором синтезе уже существующих подходов, который мог бы обеспечить применение лучших практик из разных парадигм. Такой гибридный подход должен обладать следующими характеристиками:

- быть способен поддерживать оперативную адаптацию к изменениям, характерную для Agile методологий;
- обеспечивать необходимую строгость и предсказуемость для критически важных компонентов системы через формальные модели;
- обладать инструментарием для работы в крупных распределённых командах.

Концептуальной основой нового подхода должна стать модель, которая эволюционирует вместе с проектом, обеспечивая единый и актуальный контекст для всех участников и снижая риски недопонимания между командой разработки и заказчиком. Однозначно необходимо отражать ключевые сущности, их атрибуты, взаимосвязи и функциональные требования.

Именно на этой основе должен строиться многоуровневый процесс формализации, который возможно представить алгоритмически. Предполагается, что требование начинает свой путь от начальной нарративной записи, такой как пользовательская история или комментарий заказчика. Затем это запись будет обработана с использованием структурированного сценария. Важно обеспечить чёткие критерии приемки к выходной формальной модели для архитектурно значимых аспектов системы. Таким образом, появится возможность сочетать скорость фиксации потребностей с необходимой глубиной анализа там, где это оправдано.

Далее, после полноценного исследования и внедрения нового подхода будет заметно влияние на выделенные ключевые метрики. Например, ожидается снижение количества дефектов, обнаруженных на поздних стадиях тестирования или уже при использовании программного обеспечения, за счёт раннего выявления противоречий и неоднозначностей требований. Улучшится качество и скорости реакции на обратную связь пользователей за счёт автоматизации её обработки. Благодаря системе сквозной трассируемости будет гарантированно, что каждое нефункциональное требование не будет упущено, а преобразуется в архитектурное решение. Станет возможным обеспечение снижения коммуникационных издержек внутри распределённых команд и при общении с заказчиком.

Одним из вариантов расширения нового подхода может быть применение автоматизированных инструментов, которые вероятно понадобятся для упрощения процесса поддержки бэклога требований:

1. Методы обработки естественного языка (NLP). Возможности этих средств достаточно обширны, они могут быть использованы для автоматического анализ текстовой обратной связи (например, отзывов, обращений в поддержку), кластеризации тем, выявления частотных проблем и извлечения новых потенциальных требований [3].
2. Анализ целостности требований за счёт автоматического построения и проверки графов связей между бизнес-целями, функциональными требованиями, элементами дизайна и тест-кейсами для выявления «разрывов» и нефункциональных требований, оставшихся без реализации.
3. Инструменты выявления противоречий. Здесь подразумевается, сопоставление новых требований с уже существующими формальными моделями, обеспечивающее раннее обнаружение конфликтующих условий [4].

Указанные расширения не заменяют аналитика, но могут выполнять рутинные операции по алгоритмической обработке данных, обеспечивая строгость и повышая качество результирующих спецификаций. Это помогает снизить когнитивную нагрузку специалиста, позволяя аналитику сконцентрироваться на содержательном анализе, синтезе и принятии решений.

Заключение

Сбор и формализация требований – это комплексная, нетривиальная задача трансформации неструктурированных потребностей заказчика и пользователей в строгие спецификации. Проблемы, с которыми может столкнуться аналитик при её выполнении, обусловлены системными ограничениями и нестабильностью бизнес-среды. Требования могут изменяться вследствие уточнения понимания предметной области, изменения рыночных условий или появления новых технологических ограничений. Одним из серьезных рисков на этапе сбора и формализации требований является допущенная ошибка, ведь цена её устранения возрастает экспоненциально по мере продвижения по стадиям жизненного цикла.

Анализ существующих методологий работы с требованиями привёл к необходимости разработки эффективной методологии, способной сочетать адаптивность к изменениям и обеспечение необходимого уровня строгости. Новый подход должен перенаправить работу с требованиями от следования отдельной парадигме к построению адаптивной,

инструментально насыщенной. Такая гибридная методология, основанная на многоуровневой формализации и алгоритмической поддержке, призвана преодолеть ключевые ограничения классических методов, наблюдаемые в динамической среде. Она позволит управлять нестабильностью требований через создание эффективных механизмов их уточнения, анализа и интеграции.

Практическая реализация данной концепции требует дальнейших исследований в предметной области, разработки метрик для оценки качества и проведения эмпирических исследований для валидации эффективности подхода в условиях реальных проектов.

Могут быть предложены несколько направлений дальнейших исследований для улучшения новой методологии:

1. Разработка инструментальных платформ, обеспечивающих интеграцию онтологических редакторов, средств моделирования, бэклог-менеджеров и NLP-модулей [5].
2. Создание и валидация метрик для оценки качества самой онтологии и её влияния на эффективность процесса разработки.
3. Проведение долгосрочных эмпирических исследований в индустриальных проектах для количественной оценки влияния гибридного подхода на производительность команд, качество продукта и удовлетворённость стейкхолдеров.

Применение методов машинного обучения не только для классификации, но и для прогнозирования появления новых требований на основе анализа рыночных трендов и технологических паттернов.

Список литературы

1. Буч, Г. Язык UML. Руководство пользователя / Г. Буч, Дж. Рамбо, И. Якобсон – 3-е изд. – М. : Академия АйТи, ДМК Пресс, 2022. – С. 28-52.

2. Грачев, А. А. Управление внедрением модельно-ориентированного подхода в процесс разработки программного обеспечения / А. А. Грачев, Н. Е. Емельянов, В. С. Липатников // Информационно-управляющие системы. – 2020. – № 5. – С. 57-60.

3. Хорошевский, В. Ф. Архитектура онтологически управляемых вычислительных сред / В. Ф. Хорошевский // Информационные технологии. – 2012. – Т. 18. – № 6. – С. 10-18.

4. Хорошевский, В. Ф. Онтологическое проектирование программного обеспечения: от бизнес-архитектуры к исполняемым приложениям / В. Ф. Хорошевский, К. И. Петрова. // Программная инженерия. – 2014. – № 9. – С. 11–19.

5. Лапшин В. А. Онтология в компьютерных системах / В. А. Лапшин. – М.: Научный мир, 2010. – 224 с.

Разработка инструмента для автоматической генерации тестов REST API на основе OpenAPI и семантических сценариев

М. Р. Дрокин, email: skeglz@mail.ru
Е. А. Михайлюк, email:gazon1978@yandex.ru

Воронежский государственный университет

Аннотация. *Рассмотрена задача автоматизации тестирования программных интерфейсов (API) в условиях микросервисной архитектуры. Сформулирована проблема разрыва между синтаксическим описанием API (OpenAPI) и реальной бизнес-логикой приложений. Предложен подход к созданию алгоритмов генерации тестов, который должен демонстрировать полноту покрытия не только отдельных эндпоинтов, но и пользовательских сценариев.*

Ключевые слова: *REST API, автоматизация тестирования, OpenAPI, семантический анализ, генерация тестов.*

Введение

Важными задачами, возлагаемыми на системы обеспечения качества в современной разработке программного обеспечения, являются верификация соответствия реализации требованиям и стабилизация работы сервисов в условиях частых изменений кода. В настоящее время архитектурный стиль REST (Representational State Transfer) стал стандартом де-факто для построения распределенных систем и микросервисов[1].

Однако с ростом сложностиAPI, традиционные методы написания тестов становятся «узким местом». Ручное написание автотестов требует высокой квалификации и значительных временных затрат, что противоречит принципам CI/CD (Continuous Integration / Continuous Delivery), требующим максимально быстрой обратной связи.

1. Анализ предметной области и существующих инструментов

Для исследования текущего состояния проблемы необходимо рассмотреть типовой процесс тестирования REST API. Стандартнаясхема взаимодействия включает клиента (тестовый инструмент) и сервер (тестируемое приложение).

На основе спецификации APIи документации тестировщик вычисляет необходимые входные параметры запроса и ожидаемые

© Дрокин М. Р., Михайлюк Е. А., 2026

выходные данные, соответствующие требуемому поведению системы. Здесь же формируются проверки на граничные значения, коды ответов и структуру тела ответа.

Существующий на рынке инструментарий можно классифицировать по трем основным направлениям, каждое из которых обладает своей спецификой.

Инструменты для ручного и полуавтоматического тестирования (Postman, Insomnia, Paw). Принцип их работы строится на ручном создании пользователем коллекций запросов, настройке параметров и написании скриптов проверок. Основным недостатком является низкая скорость разработки тестов и высокая трудоемкость поддержки: при любых изменениях API сценарии приходится переписывать вручную, что не позволяет обеспечить требуемое качество управления тестированием в условиях частых релизов.

Генераторы на основе схем (Swagger Codegen, Dredd, Schemathesis). Данные инструменты парсят файл спецификации (JSON/YAML) и генерируют запросы, перебирая возможные значения согласно типам данных. Ключевой недостаток заключается в ограничении проверкой синтаксиса (соответствия схеме) при отсутствии семантического контроля. Инструмент может подтвердить корректность структуры JSON-ответа, но не способен валидировать сложную бизнес-логику, например, невозможность оплаты заказа, который не был предварительно создан.

Инструменты fuzzing-тестирования (RESTler, OWASP ZAP, Burp Suite). Работают по принципу подачи на вход API большого массива случайных, некорректных или неожиданных данных с целью вызвать сбой системы. Их недостаток в контексте функционального тестирования – узкая ориентированность на поиск уязвимостей и критических ошибок (падений сервиса), а не на проверку корректности выполнения последовательных бизнес-сценариев.

Таким образом, проведенный сравнительный анализ позволяет сделать вывод, что ни один из представленных классов инструментов не обладает возможностью автоматической генерации тестов на основе семантических сценариев. Существующие решения ограничиваются либо ручным программированием логических цепочек, либо формальной проверкой изолированных эндпоинтов, что делает невозможным полноценную валидацию бизнес-процессов без участия человека[2].

2. Структурирование проблемы и предлагаемое решение

Наиболее актуальная задача – сформулировать алгоритм генерации тестов так, чтобы он был устойчив по отношению к изменениям

внутренней реализации сервиса, но чувствителен к нарушениям бизнес-логики.

Проблема заключается в семантическом разрыве между формальной спецификацией и реальным поведением системы.Спецификация OpenAPI описывает интерфейс статически [3]:

– Endpoint A принимает число;
– Endpoint B возвращает строку.

Однако реальная эксплуатация подразумевает динамические сценарии:

1. Вызвать эндпоинт A (создать ресурс).
2. Получить ID из ответа.
3. Использовать этот ID для вызова эндпоинта B.

Для решения таких классов задач обычно используют два подхода: жесткое кодирование сценариев или Model-Based Testing (тестирование на основе моделей). Предлагаемый подход развивает идеи Model-Based Testing, добавляя к нему элементы семантического анализа спецификаций.

Достижение поставленной цели предполагает решениеряда задач:

1. Парсинг и нормализация. Преобразование спецификации OpenAPI в граф объектов и методов.
2. Семантическое обогащение. Разработка формата (DSL – Domain Specific Language) для описания связей между эндпоинтами, которые невозможно выразить в чистом OpenAPI (например, «значение поля `id` из ответа `POST /users` должно быть подставлено в URL `GET /users/{id}`»).
3. Синтез графа переходов. Построение ориентированного графа возможных состояний системы, где узлы – это состояния данных, а ребра –вызовы API.
4. Генерация кода. Трансляция найденных путей в графе в исполняемый код тестов (например, на Python/Pytest).

В блоке генерации на основе действительных параметров спецификации и требуемых параметров сценария будут формироваться HTTP-запросы, которые поступают на тестируемый сервис. Анализатор ответов будет сравнивать полученные данные не только с JSON-схемой, но и с контекстом выполнения сценария.

Такой подход позволяет перейти от модульного тестирования отдельных ручек к интеграционному тестированию бизнес-процессов в автоматическом режиме.

3. Граф состояний

Для реализации предложенного подхода разрабатывается обобщенная структурная схема инструмента, основанная на теории графов. Ключевой особенностью архитектуры является отказ от линейной обработки списка эндпоинтов в пользу построения ориентированного графа состояний.

Математически модель тестируемой системы представляется в виде ориентированного графа $G = (V, E)$, где:

- V (вершины) – множество состояний системы или ресурсов (например, «Пользователь создан», «Корзина не пуста»);
- E (дуги) – множество переходов между состояниями, соответствующих вызовам методов API (HTTP Request).

Каждая дуга $e \in E$ характеризуется кортежем параметров $e = (M, P_{in}, P_{out}, S)$, где:

- M –метод HTTP (GET, POST, PUT, DELETE);
- P_{in} – входные параметры, необходимые для выполнения перехода;
- P_{out} – выходные данные, порождаемые переходом;
- S – ожидаемое состояние системы после перехода (HTTP status code).

На этапе инициализации система выполняет статический анализ спецификации OpenAPI. Если спецификация описывает только множество E (endpoints), то алгоритм инструмента реконструирует множество V, анализируя схемы ответов и перекрестные ссылки [4]. На рисунке представлен граф состояний для простой логики веб-сервиса.

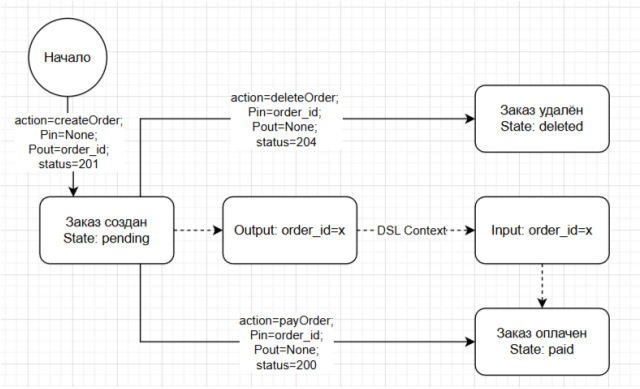


Рисунок. Схематичное представление графа состояний

Использование графаоткрывает возможности для решения следующих задач:

– Поиск путей. Генерация тестового сценария сводится к нахождению пути от начальной вершины v_{init} к целевой вершине v_{target}. Для этого используются алгоритмы обхода графа, такие как поиск в ширину (BFS) для нахождения кратчайших сценариев или поиск в глубину (DFS) для проверки сложных цепочек.

– Анализ связности. Выявление «изолированных» вершин – эндпоинтов, до которых невозможно добраться в рамках текущей бизнес-логики. Например, удаление заказа, который невозможно создать.

– Оптимизация покрытия. Инструмент может автоматически строить маршруты так, чтобы посетить каждое ребро графа минимум один раз.

Ключевым элементом, обеспечивающим движение по графу, является модуль разрешения зависимостей. Он решает задачу передачи контекста: если ребро e_1 порождает данные $d \in P_{out}$ (например, `id`), а ребро e_2 требует $d \in P_{in}$, модуль создает временную связь (транзакционную память), позволяющую передать значение переменной между шагами теста.

4. Реализация семантического слоя и генерация кода

Для устранения семантического разрыва, описанного во введении, в контур управления вводится предметно-ориентированный язык (DSL).

Использование DSL позволяет реализовать декларативную парадигму тестирования. В отличие от императивных скриптов (Python/Java), где жестко фиксируется последовательность вызовов и структура данных, DSL описывает намерение и контракт взаимодействия [5]. В листинге представлена структура разработанного языка DSLна примере последовательности действий, состоящей из двух шагов.

Листинг

Описание операций на языке DSL

```
steps:
    # Шаг 1: Инициализация ресурса
    - step: "Регистрация клиента"
action: "createUser"# Ссылкана operationId в OpenApi
      payload:
        username: "test_user"
        role: "admin"
extract:
        created_id: "id"# Сохранение ID из ответа в переменную
```

Окончание листинга

```
    # Шаг 2: Очистка данных (Зависимый шаг)
- step: "Удалениеклиента"
      action: "deleteUser"
      params:
userId: ${created_id}# Инъекция сохраненного значения
      assert:
        status: 204
```

В качестве базового синтаксиса DSL выбран формат YAML, так как он нативно поддерживает иерархические структуры данных и широко используется в DevOps-среде.Алгоритм трансляции DSL-сценария в исполняемый код:

1. Парсинг и валидация. Загрузчик считывает YAML-файл и проверяет соответствие указанных `шагов` реальным путям в OpenAPI спецификации.
2. Интерполяция переменных. На этом этапе происходит связывание данных. Если в следующем шаге DSL блок params требует некоторую переменную, транслятор попытается найти её в блоках extract из предыдущих шагов. Это позволяет генерировать динамический код, не зависящий от конкретных значений данных.
3. Синтез тела запроса. Инструмент объединяет JSON-схему из OpenAPI и данные из блока payload сценария. Данные из сценария имеют наивысший приоритет, остальные обязательные поля заполняются синтетическими данными для удовлетворения требований схемы.
4. Генерация Assertions. Блок assert трансформируется в набор утверждений тестового фреймворка.

Финальным этапом является сериализация построенной объектной модели в целевой язык программированияс использованием шаблонизатора. Такой подход обеспечивает гибкость: при изменении реализации API достаточно обновить спецификацию и перегенерировать тесты, не меняя сам файл семантического сценария, если бизнес-логика осталась прежней.

Заключение

В работе сформулирована задача автоматизации тестирования REST API, направленная на устранение семантического разрыва между спецификацией и бизнес-логикой. По результатам анализа существующих решений выявлена их неспособность обеспечить полноценную проверку сценариев без ручного вмешательства. Разработана архитектура инструмента на основе гибридного подхода,

объединяющего топологию OpenAPI и семантику декларативного DSL для построения ориентированного графа состояний. Предложенный алгоритм автоматического разрешения зависимостей данных обеспечивает генерацию устойчивых интеграционных тестов, динамически связывая параметры запросов и исключая необходимость жесткого кодирования.

Список литературы

1. Peralta, J. H. Microservice APIs: Using Python, Flask, FastAPI, OpenAPI and more / J. H. Peralta. – Shelter Island, NY.: Manning Publications, 2023. – 440 p.
2. Бек К. Экстремальное программирование. Разработка через тестирование / К. Бек. – СПб.: Питер, 2022. – 224 с.
3. OpenAPI Specification. Version 3.2.0 [Электронный ресурс] . – Режим доступа: https://spec.openapis.org/oas/v3.2.0
4. Мараховский, В. Б**.** Сети Петри / В. Б. Мараховский, Л. Я. Розенблюм, А. В. Яковлев. – СПб.: Профессиональная литература, 2014. – 400 с.
5. Введение в DSL. Часть 1 — Проблематика проектирования и кодирования [Электронный ресурс]. – Режим доступа: https://habr.com/ru/articles/94259

Идентификация интегро-дифференцирующего звена

А. В. Дылевский, E-mail: nefta@yandex.com
Д. А. Хрипушин, E-mail: qepm@outlook.com

Воронежский государственный университет

Аннотация. *Рассматривается задача идентификации интегро-дифференцирующего звена общего вида. Приводится метод определения неизвестных параметров звена. Метод сводится к интегрированию переходной характеристики звена и последующему решению системы линейных алгебраических уравнений.*

Ключевые слова: *идентификация, интегро-дифференцирующее звено, передаточная функция.*

Введение и постановка задачи

В теории автоматического управления звено с передаточной функцией вида

$$W_{ИД}(p) = k\frac{T_1 p + 1}{T_2 p + 1}, \tag{1}$$

где $k > 0$ – коэффициент усиления, $T_1, T_2 \geq 0$ – постоянные времени, называется интегро-дифференцирующим. Это звено хорошо изучено, получены все частотные и временные характеристики звена. Интерес к интегро-дифференцирующему звену объясняется тем, что из передаточной функции (1) при некоторых значениях параметров k, T_1, T_2 могут быть получены все типовые звенья 1-го порядка [1]. Кроме того, интегро-дифференцирующее звено используется, например, при моделировании реакции человека на входной сигнал [2].

Рассмотрим далее задачу определения неизвестных параметров следующей передаточной функции

$$W(p) = \frac{b_0 p + b_1}{a_0 p + a_1}, \quad a_0, a_1, b_0, b_1 \in R. \tag{2}$$

Нетрудно заметить, что передаточная функция (1) может быть получена из (2) при следующих значениях параметров:

$$a_0 = T_2, \quad a_1 = 1, \quad b_0 = kT_1, \quad b_1 = k\ . \tag{3}$$

© Дылевский А. В., Хрипушин Д. А., 2026

Так как при $T_2 = 0$ передаточная функция (1) является нереализуемой, то без ограничения общности можно считать, что $a_0 = 1$, т.е. вместо передаточной функции (2) можно рассмотреть передаточную функцию

$$W(p) = \frac{b_0 p + b_1}{p + a_1}. \tag{4}$$

Передаточная функция вида (4) применяется для линеаризации и аппроксимации нелинейных объектов, для редукции передаточных функций высоких порядков. Например, в качестве аппроксимации звена чистого запаздывания используют дробь Паде 1-го порядка [3, 4]

$$e^{-\tau p} \approx \frac{-\frac{\tau}{2} p + 1}{\frac{\tau}{2} p + 1}. \tag{5}$$

Здесь $\tau > 0$ – время запаздывания.

В ряде случаев параметры передаточной функции (4) являются неизвестными (например, при определении модели летчика [2]). В этом случае возникает задача идентификации параметров b_0, b_1, a_1.

Описание метода решения поставленной задачи

Существует большое число эффективных методов идентификации систем [5]. Для решения поставленной задачи dвоспользуемся методом идентификации линейных объектов, предложенный в [5]. С этой целью рассмотрим звено (4) во временной области. Нетрудно проверить, что передаточная функция (4) соответствует дифференциальному уравнению

$$\dot{x}(t) + a_1 x(t) = b_0 \dot{u}(t) + b_1 u(t). \tag{6}$$

Здесь $u(t)$ – заданное входное воздействие, $x(t)$ – реакция звена (4) на входное воздействие. Будем предполагать, что

$$x(0) = 0. \tag{7}$$

Пусть

$$u(t) = 1(t) = \begin{cases} 1, & t > 0, \\ 0, & t \le 0. \end{cases} \tag{8}$$

Проинтегрируем уравнение (6) по t в промежутке $[0, t]$. Тогда

$$\int_0^t \dot{x}(z)dz + a_1 \int_0^t x(z)dz = b_0 \int_0^t \dot{u}(z)dz + b_1 \int_0^t u(z)dz; \tag{9}$$

Так как входное воздействие $u(t)$ определяется равенством (8), то

$$\int_0^t u(z)dz = \int_0^t dz = t,$$
$$\int_0^t \dot{u}(z)dz = u(z)\Big|_0^t = 1. \tag{10}$$

Введем следующее обозначение:

$$x(t) = X_0(t), \int_0^t X_i(z)dz = X_{i+1}(t),\ i = 0,1,2. \tag{11}$$

Тогда, учитывая начальное условие (7) и равенство

$$\int_0^t \dot{x}(z)dz = x(z)\Big|_0^t = x(t) = X_0(t), \tag{12}$$

из (6) после элементарных преобразований получаем уравнение

$$b_0 + b_1 t - a_1 X_1(t) = X_0(t). \tag{13}$$

Проинтегрируем теперь полученное уравнение по t в промежутке $[0,t]$. Имеем

$$b_0\int_0^t dz + b_1\int_0^t zdz - a_1\int_0^t X_1(z)dz = \int_0^t X_0(z)dz. \tag{14}$$

Нетрудно проверить, что отсюда с учетом обозначений (11) после элементарных преобразований следует

$$b_0 t + b_1\frac{t^2}{2} - a_1 X_2(t) = X_1(t). \tag{15}$$

Аналогично интегрируя (15), находим

$$b_0\frac{t^2}{2} + b_1\frac{t^2}{6} - a_1 X_3(t) = X_2(t). \tag{16}$$

Таким образом, из (13), (15) и (16) для каждого фиксированного момента времени t имеем следующую систему линейных алгебраических уравнений относительно неизвестных b_0, b_1, a_1:

$$\begin{cases} b_0 + b_1 t - a_1 X_1(t) = X_0(t), \\ b_0 t + b_1 \dfrac{t^2}{2} - a_1 X_2(t) = X_1(t), \\ b_0 \dfrac{t^2}{2} + b_1 \dfrac{t^2}{6} - a_1 X_3(t) = X_2(t). \end{cases} \tag{17}$$

Для нахождения неизвестных b_0, b_1, a_1 решим систему уравнений (17) методом Крамера. Находим главный определитель системы

$$\Delta(t) = \begin{vmatrix} 1 & t & -X_1(t) \\ t & \dfrac{t^2}{2} & -X_2(t) \\ \dfrac{t^2}{2} & \dfrac{t^3}{6} & -X_3(t) \end{vmatrix}. \tag{18}$$

Определители $\Delta_{b_0}(t)$, $\Delta_{b_1}(t)$, $\Delta_{a_1}(t)$ получаем из $\Delta(t)$ следующим образом: столбец коэффициентов при соответствующей неизвестной заменяется столбцом свободных членов системы (17). Имеем

$$\Delta_{b_0}(t) = \begin{vmatrix} X_0(t) & t & -X_1(t) \\ X_1(t) & \dfrac{t^2}{2} & -X_2(t) \\ X_2(t) & \dfrac{t^3}{6} & -X_3(t) \end{vmatrix}, \tag{19}$$

$$\Delta_{b_1}(t) = \begin{vmatrix} 1 & X_0(t) & -X_1(t) \\ t & X_1(t) & -X_2(t) \\ \dfrac{t^2}{2} & X_2(t) & -X_3(t) \end{vmatrix}, \tag{20}$$

$$\Delta_{a_1}(t) = \begin{vmatrix} 1 & t & X_0(t) \\ t & \frac{t^2}{2} & X_1(t) \\ \frac{t^2}{2} & \frac{t^3}{6} & X_2(t) \end{vmatrix}. \qquad (21)$$

Если для фиксированного момента времени t главный определитель системы отличен от нуля, т.е. $\Delta(t) \neq 0$, то искомые параметры b_0, b_1, a_1 будут определяться равенствами

$$b_0 = \frac{\Delta_{b_0}(t)}{\Delta(t)}, \quad b_1 = \frac{\Delta_{b_1}(t)}{\Delta(t)}, \quad a_1 = \frac{\Delta_{a_1}(t)}{\Delta(t)}. \qquad (22)$$

Заключение

Сформулирована и решена задача идентификации передаточной функции интегро-дифференцирующего звена. Задача решается во временной области и сводится к интегрированию переходной характеристики и последующему решению линейной алгебраической системы уравнений. Такой подход позволяет упростить решение задач параметрического синтеза, редукции, линеаризации и аппроксимации систем управления.

Литература

2. Мирошник, И. В. Теория автоматического управления. Линейные системы: Учебное пособие для вузов / И. В. Мирошник. – СПб.: Питер, 2005. – 336 с.

3. Гаркушенко, В. И. К задаче проектирования системы управления самолета с заданнми динамическими свойствами / В. И. Гаркушенко, П. А. Лазарева // Известия вузов. Авиационная техника. – 2015. – №4. – С.21-25.

4. Бейкер, Дж. Аппроксимации Паде / Дж. Бейкер, П. Грейвс-Моррис. – М.: Мир, 1986. – 502 с.

5. Хрипушин, Д. А. Построение опережающих индикаторов на основе дифференциаторов : диссертация на соискание ученой степени кандидата физико-математических наук / Хрипушин Д. А., 2024. – 117 с.

6. Современные методы идентификации систем / П. Эйкхофф [и др.]. – М.: Мир, 1983. – 400 с.

Разработка бэкенд-сервиса для анализа и ранжирования смежных информационных технологий

Д. С. Занин, email: zanindanila@mail.ru
Е. А. Михайлюк, email: gazon1978@yandex.ru

Воронежский государственный университет

Аннотация. *В работе рассматривается разработка бэкенд сервиса для анализа и ранжирования смежных информационных технологий. Основная идея сервиса заключается в автоматизированном сборе данных из открытых источников, построении графа взаимосвязей между технологиями и вычислении их значимости на основе многокритериальной оценки. Предложенный подход позволяет снизить информационную перегрузку при выборе инструментов, повысить осознанность профессионального развития специалистов и упростить навигацию в сложных технологических экосистемах.*

Ключевые слова: *смежные технологии, анализ популярности, многокритериальная оценка, граф технологий, GitHub, Maven Central, Stack Overflow.*

Введение

Современное развитие информационных технологий характеризуется не только быстрым появлением новых инструментов, но и усложнением экосистем, в которых эти инструменты используются. Практически ни одна технология сегодня не применяется изолированно: фреймворки, библиотеки и сервисы образуют сложные цепочки зависимостей. В результате специалисту становится всё сложнее самостоятельно определить, какие технологии действительно востребованы, какие логично изучать в дополнение к уже используемым, а какие постепенно теряют актуальность.

Актуальность данной работы обусловлена необходимостью автоматизированного инструмента, который помогал бы ориентироваться в этом многообразии. Такой инструмент должен не просто показывать популярность отдельной технологии, но и выявлять связанные с ней решения, оценивать их значимость и представлять информацию в удобной и наглядной форме [1]. Это особенно важно для разработчиков, архитекторов и студентов, которые планируют своё профессиональное развитие и выбирают стек технологий для конкретных проектов.

© Занин Д. С., Михайлюк Е. А., 2026

Целью работы является разработка бэкенд сервиса, собирающего данные о технологиях из открытых источников, анализирующего их взаимосвязи и предоставляющего пользователю итоговую оценку смежных технологий на основе объективных критериев. Для достижения этой цели рассматриваются вопросы выбора источников данных, построения алгоритмов оценки и организации интерфейса взаимодействия с пользователем.

1. Сбор данных о технологиях

Формирование базы данных технологий и их связей осуществляется на основе открытых интернет платформ, активно используемых профессиональным сообществом. В качестве ключевых источников выбраны GitHub, Maven Central, Stack Overflow, а также тематические публикации на Habr [2] и Medium. Эти ресурсы позволяют получить информацию как о практическом использовании технологий, так и об уровне интереса к ним со стороны разработчиков.

GitHub используется для анализа репозиториев, статистики звёзд, форков и активности обновлений, что даёт представление о популярности и жизнеспособности технологии. Данные извлекаются через официальный REST API GitHub [3], который предоставляет структурированную и регулярно обновляемую информацию.

Maven Central служит источником сведений о Java библиотеках, артефактах и их зависимостях. Анализ этих данных позволяет выявлять совместное использование технологий и строить граф связей на уровне библиотек и пакетов. Для доступа к данным используется официальный API репозитория Maven Central [4].

Stack Overflow отражает практические проблемы и интересы разработчиков. Количество вопросов и динамика их появления по конкретной технологии используются как индикатор её востребованности и актуальности. Данные получаются через Stack Exchange API [5].

Для платформ, не предоставляющих полноценного API, таких как Habr и Medium, применяется парсинг с использованием Python и Selenium WebDriver [6]. Полученные данные нормализуются и сохраняются в едином формате, что позволяет объединять информацию из разных источников.

2. Оценка данных и взаимодействие с пользователем

Разработанный бэкенд сервис ориентирован на поиск и анализ смежных технологий по запросу пользователя. В качестве исходной точки пользователь указывает конкретную технологию, после чего система

формирует список связанных с ней решений и выполняет их ранжирование.

В основе поиска лежит построение графа взаимосвязей, в котором вершины соответствуют технологиям, а рёбра отражают их совместное использование в проектах и библиотеках. Чем чаще две технологии встречаются вместе, тем выше вес связи между ними. Подход позволяет выявлять связи, основанные на реальном опыте использования.

Для ранжирования технологий применяется многокритериальная оценка [7, 8]. Она учитывает популярность репозиториев на GitHub, распространённость библиотек и зависимостей в Maven Central, активность обсуждений на Stack Overflow, а также степень связанности технологии с другими элементами экосистемы [9]. Итоговая оценка рассчитывается как взвешенная сумма нормализованных критериев, при этом веса могут быть адаптированы под конкретные цели пользователя.

Результаты работы сервиса передаются через REST API в формате JSON и могут быть использованы как для отображения списка технологий, так и для визуализации графа связей.

Пример ответа сервиса для запроса смежных технологий приведён в листинге ниже.

Листинг

Пример ответа сервиса на запрос по S3

```
{
  "input_technology": "s3",
  "related_technologies": [
    {
      "name": "aws-java-sdk-s3",
      "score": 0.92,
      "metrics": {
        "github_stars": 5200,
        "maven_dependencies": 1840,
        "stackoverflow_questions_last_year": 960
      }
    },
    {
      "name": "minio-java-client",
      "score": 0.81,
      "metrics": {
        "github_stars": 3100,
        "maven_dependencies": 740,
        "stackoverflow_questions_last_year": 420
      }
    }
  ]
}
```

Наличие детализированных метрик позволяет использовать результаты сервиса как для простого ранжирования, так и для последующего анализа или визуализации связей между технологиями.

3. Архитектура системы и организация сервисов

Для обеспечения масштабируемости, устойчивости к нагрузке и возможности дальнейшего развития система реализуется с использованием микросервисной архитектуры. Общая функциональность приложения разделена на несколько независимых компонентов, каждый из которых отвечает за строго определённый набор задач.

Ключевым компонентом является сервис загрузки и обновления данных. Он предназначен для автоматизированного сбора информации о технологиях из внешних источников, таких как GitHub, Maven Central, Stack Overflow, а также тематические публикации на Habr и Medium. Данный сервис функционирует в фоновом режиме и не предоставляет интерфейсов для прямого взаимодействия с пользователем.

Для регулярного обновления данных в сервисе загрузки используется планировщик задач Quartz. С его помощью настраиваются задания, выполняющиеся с периодичностью один раз в 24 часа. В рамках каждого задания выполняются HTTP-запросы к внешним REST API и процедуры парсинга веб-контента. Полученные данные проходят этапы нормализации и агрегации, после чего сохраняются в централизованное хранилище. Использование Quartz позволяет гибко управлять расписанием задач, обеспечивать повторное выполнение при ошибках и контролировать состояние процессов обновления данных.

Вторым компонентом системы является аналитический сервис, реализующий REST API для взаимодействия с пользователями. Он использует предварительно собранные и сохранённые данные, формирует граф взаимосвязей технологий, выполняет многокритериальную оценку [1, 2, 9] и возвращает результаты в формате JSON. Аналитический сервис не обращается напрямую к внешним источникам, что снижает задержки обработки запросов и уменьшает зависимость от доступности сторонних платформ.

Взаимодействие между сервисом загрузки данных и аналитическим сервисом осуществляется через общее хранилище данных, что обеспечивает слабую связность компонентов и позволяет независимо масштабировать каждый сервис. Такая архитектура упрощает сопровождение системы, повышает её надёжность и создаёт основу для дальнейшего расширения функциональности, включая подключение новых источников данных и аналитических модулей.

4. Архитектура системы и организация сервисов

Для обеспечения масштабируемости, устойчивости к нагрузке и возможности дальнейшего развития система реализуется с использованием микросервисной архитектуры. Общая функциональность приложения разделена на несколько независимых компонентов, каждый из которых отвечает за строго определённый набор задач.

Ключевым компонентом является сервис загрузки и обновления данных. Он предназначен для автоматизированного сбора информации о технологиях из внешних источников, таких как GitHub, Maven Central, Stack Overflow, а также тематические публикации на Habr и Medium. Данный сервис функционирует в фоновом режиме и не предоставляет интерфейсов для прямого взаимодействия с пользователем.

Для регулярного обновления данных в сервисе загрузки используется планировщик задач Quartz. С его помощью настраиваются задания, выполняющиеся с периодичностью один раз в 24 часа. В рамках каждого задания выполняются HTTP-запросы к внешним REST API и процедуры парсинга веб-контента. Полученные данные проходят этапы нормализации и агрегации, после чего сохраняются в централизованное хранилище. Использование Quartz позволяет гибко управлять расписанием задач, обеспечивать повторное выполнение при ошибках и контролировать состояние процессов обновления данных.

Вторым компонентом системы является аналитический сервис, реализующий REST API для взаимодействия с пользователями. Он использует предварительно собранные и сохранённые данные, формирует граф взаимосвязей технологий, выполняет многокритериальную оценку [1, 2, 9] и возвращает результаты в формате JSON. Аналитический сервис не обращается напрямую к внешним источникам, что снижает задержки обработки запросов и уменьшает зависимость от доступности сторонних платформ.

Взаимодействие между сервисом загрузки данных и аналитическим сервисом осуществляется через общее хранилище данных, что обеспечивает слабую связность компонентов и позволяет независимо масштабировать каждый сервис. Такая архитектура упрощает сопровождение системы, повышает её надёжность и создаёт основу для дальнейшего расширения функциональности, включая подключение новых источников данных и аналитических модулей.

Заключение

В результате работы был разработан бэкенд сервис, предназначенный для анализа смежных технологий и их оценки на основе объективных данных. Использование открытых источников,

построение графа взаимосвязей и применение многокритериального подхода позволяют не только ранжировать технологии по популярности, но и выявлять их практическую близость.

Проведённое сравнение с существующими платформами показало, что предложенное решение выигрывает за счёт интеграции нескольких источников данных, гибкости оценки и возможности визуализации. В перспективе развитие сервиса может быть связано с подключением дополнительных источников информации, применением методов машинного обучения для прогнозирования трендов и более глубокой персонализацией рекомендаций под конкретного пользователя.

Список литературы

1. Shneiderman, B. Designing the User Interface: Strategies for Effective Human-Computer Interaction / B. Shneiderman. – Boston : Pearson Education, 2010. – 560 p.
2. Habr : сообщество разработчиков программного обеспечения [Электронный ресурс]. – Режим доступа: https://habr.com/ru/
3. GitHub REST API Documentation [Электронный ресурс]. – Режим доступа: https://docs.github.com/en/rest
4. Maven Central Repository API [Электронный ресурс]. – Режим доступа: https://search.maven.org/search/api-docs.html
5. Stack Exchange API Documentation [Электронный ресурс]. – Режим доступа: https://api.stackexchange.com
6. Selenium WebDriver Documentation [Электронный ресурс]. – Режим доступа: https://www.selenium.dev/documentation/
7. Saaty, T. L. The Analytic Hierarchy Process / T. L. Saaty. – New York : McGraw-Hill, 1980. – 287 p.
8. Keeney, R. L. Decisions with Multiple Objectives: Preferences and Value Trade-offs / R. L. Keeney, H. Raiffa. – Cambridge : Cambridge University Press, 1976. – 569 p.
9. Turban, E. Decision Support and Business Intelligence Systems / E. Turban, R. Sharda, D. Delen. – 10th ed. – Boston : Pearson, 2018. – 720 p.

Анализ применимости различных форматов хранения аудиоданных при разработке чата

К. А. Кислый, email: kislyikirill@gmail.com
М. С. Ефремов, email: mefremov@relex.ru

Воронежский государственный университет

Аннотация. *Рассмотрены различные форматы хранения аудиоданных, которые могут использоваться при реализации приложения чата, содержащего возможности отправки голосовых сообщений.*

Ключевые слова: *аудиоданные, разработка чата, форматы хранения данных, голосовой чат, кодек, сжатие аудио, совместимость форматов.*

Введение

Современное информационное общество характеризуется стремительным ростом объёмов мультимедийных данных, среди которых всё большую роль играет аудиоинформация – особенно в контексте голосовых интерфейсов, голосовых чатов и систем распознавания речи. Эффективная обработка и хранение таких данных требуют тщательного выбора формата, поскольку от него зависят не только качество звучания и размер файлов, но и вычислительные затраты, совместимость с платформами и возможность дальнейшего анализа.

Форматы хранения аудиоданных – от несжатых (например, WAV) до сильно сжатых с потерями (таких как MP3 или Opus) – представляют собой компромисс между точностью воспроизведения, объёмом и скоростью передачи. При разработке современных чат-систем, особенно поддерживающих голосовое взаимодействие, выбор оптимального формата становится ключевой инженерной задачей, влияющей на производительность, масштабируемость и пользовательский опыт.

В данной работе проводится сравнительный анализ наиболее распространённых аудиоформатов с точки зрения их применимости в архитектуре чат-приложений, рассматриваются технические характеристики, требования к ресурсам и последствия для последующей обработки аудиопотоков, включая распознавание речи и извлечение семантической информации.

© Кислый К. А., Ефремов М. С., 2026

1. MP3 (MPEG Audio Layer III)

MP3 – это один из самых распространённых стандартов сжатия аудиофайлов, разработанный в рамках проекта MPEG-1 Audio Layer III. Благодаря эффективному алгоритму компрессии, он получил широкое применение для хранения и распространения музыки, подкастов, аудиокниг и другого звукового контента. Основная цель формата – значительно уменьшить объём аудиоданных при сохранении приемлемого для большинства слушателей качества звука. Именно эта особенность делает MP3 особенно удобным для передачи через интернет и использования на устройствах с ограниченными ресурсами.

Среди ключевых преимуществ MP3 следует выделить небольшой размер файла. Алгоритм сжатия основан на психоакустической модели восприятия звука человеком, что позволяет отбрасывать те компоненты аудиосигнала, которые малозаметны или не воспринимаются слухом. В результате размер файла может быть уменьшен в 10–12 раз по сравнению с несжатым форматом, таким как WAV, без явно выраженного ухудшения субъективного качества. Это особенно важно при работе с большими коллекциями аудиозаписей или в условиях ограниченной пропускной способности сети [1].

Ещё одним важным достоинством является практически универсальная поддержка формата. MP3 считается де-факто стандартом в мире цифрового аудио: его распознают все основные операционные системы, аудиоплееры, смартфоны, веб-браузеры и мультимедийные платформы. Такая совместимость обеспечивает беспрепятственное воспроизведение и обмен аудиофайлами между пользователями, независимо от используемых устройств или программного обеспечения.

Кроме того, MP3 отличается гибкостью настройки качества. При кодировании пользователь может выбирать различные значения битрейта — от 32 до 320 кбит/с, что позволяет адаптировать файл под конкретные задачи. Например, для голосовых подкастов, где высокая детализация звука не критична, достаточно низкого битрейта, тогда как для музыкальных треков предпочтительнее использовать максимальные значения для сохранения насыщенности и глубины звучания. Эта адаптивность делает формат применимым в самых разных сценариях – от потоковой передачи до локального хранения.

В то же время MP3 имеет и существенные недостатки. Прежде всего, это потеря качества звука, обусловленная самой природой формата как lossy (с потерями). При сжатии часть исходной аудиоинформации необратимо удаляется. Особенно заметны артефакты сжатия при использовании низких битрейтов – ниже 128 кбит/с: могут проявляться «размытость» высоких частот, металлический тембр, потеря

пространственного звучания. Даже при высоких битрейтах полное восстановление оригинального сигнала невозможно, что ограничивает возможности повторной обработки или редактирования.

По этой причине MP3 не подходит для профессионального аудиопроизводства. В сфере звукозаписи, мастеринга и других задач, где требуется максимальная точность передачи аудиосигнала, данный формат считается неприемлемым. Звукорежиссёры и продюсеры предпочитают использовать форматы без потерь, такие как FLAC, WAV или AIFF, которые сохраняют все детали оригинальной записи. Кроме того, многократное перекодирование MP3-файлов усугубляет деградацию качества, что делает его непригодным для промежуточного хранения в рабочих процессах.

2. WAV (Waveform Audio File Format)

Формат WAV (Waveform Audio File Format) представляет собой стандарт хранения несжатого аудио, разработанный совместно Microsoft и IBM. Он сохраняет звуковую информацию в исходном виде, без применения алгоритмов сжатия, что обеспечивает полную точность воспроизведения аудиосигнала. Благодаря этому WAV широко используется в профессиональной звукозаписи, аудиомонтаже, мастеринге, а также в приложениях, где критически важна целостность звуковых данных – например, в системах распознавания речи или акустическом анализе.

Среди главных достоинств формата – высокое качество звука. Поскольку WAV не применяет сжатие с потерями, он сохраняет все частотные и динамические характеристики оригинальной записи. Это делает его незаменимым инструментом для музыкантов, звукорежиссёров и инженеров, работающих с аудио на профессиональном уровне, а также для задач, требующих максимальной точности, таких как обучение моделей машинного обучения на аудиоданных.

Ещё одно преимущество – простота редактирования. Отсутствие компрессии означает, что данные хранятся в линейном, легко читаемом виде, что упрощает обработку, нарезку, наложение эффектов и другие операции в аудиоредакторах. Это особенно ценно в многоэтапных рабочих процессах, где файл может многократно изменяться без риска дополнительной деградации качества.

Кроме того, формат WAV обладает высокой совместимостью. Он поддерживается всеми основными операционными системами – Windows, macOS, Linux – а также большинством профессиональных и потребительских аудиопрограмм, от простых проигрывателей до сложных цифровых аудиорабочих станций (DAW).

Вместе с тем, у WAV есть и существенные недостатки. Прежде всего, это большой размер файлов. При стандартных параметрах CD-качества (16 бит, 44,1 кГц, стерео) одна минута аудио занимает около 10 мегабайт. Такой объём быстро исчерпывает доступное дисковое пространство при работе с длительными записями или крупными проектами, что создаёт трудности при хранении и резервном копировании [2].

Кроме того, из-за значительного размера WAV оказывается неэффективным для массового распространения контента через интернет. В условиях, где важны скорость загрузки, экономия трафика и удобство передачи, несжатые файлы уступают место более компактным форматам, таким как MP3, AAC или Opus. Поэтому WAV редко используется конечными пользователями для прослушивания – его роль ограничивается этапами создания, обработки и архивирования аудио до финального экспорта в сжатый формат.

Особую роль формат WAV играет в задачах, где требуется точная временная привязка и минимальная обработка сигнала. Например, при разработке систем автоматической транскрибации речи или голосовых ассистентов исходные записи часто сохраняются именно в WAV, чтобы избежать искажений, вносимых алгоритмами сжатия. Это позволяет алгоритмам машинного обучения работать с «чистым» сигналом, что повышает точность распознавания фонем, интонаций и пауз. Однако в условиях реального времени – например, при передаче голоса в чате – такие преимущества становятся избыточными, поскольку требуют значительных сетевых и вычислительных ресурсов. Таким образом, несмотря на техническое совершенство, WAV остаётся форматом для подготовки и анализа, но не для доставки аудио конечному пользователю в интерактивных приложениях.

3. AAC (Advanced Audio Codec)

AAC (Advanced Audio Coding) – это современный формат сжатия аудиоданных, разработанный как улучшенная альтернатива MP3 в рамках стандарта MPEG-2 и позднее расширенного в MPEG-4. Он получил широкое распространение в сфере потоковой передачи музыки, цифрового радио и мобильных медиа, особенно в экосистеме Apple, где AAC стал де-факто стандартом для iTunes, Apple Music и других сервисов. Основное преимущество формата заключается в более высокой эффективности кодирования, что позволяет достичь лучшего качества звука при тех же или даже меньших битрейтах по сравнению с MP3 [3].

Одним из ключевых достоинств AAC является его превосходное качество звучания при одинаковых значениях битрейта. Благодаря усовершенствованной психоакустической модели и более гибким

методам обработки частотных компонентов, AAC сохраняет больше деталей оригинального сигнала, особенно в диапазоне высоких частот. Это делает его предпочтительным выбором для стриминговых платформ, где важно обеспечить наилучшее восприятие звука при ограниченной пропускной способности канала.

Кроме того, AAC поддерживает многоканальное аудио – до 48 каналов в теории, хотя на практике чаще используются стерео- и 5.1-конфигурации. Эта особенность делает формат пригодным не только для музыки, но и для саундтреков к фильмам, телевизионного вещания и домашних кинотеатров, где важна пространственная точность звукового поля [4].

Формат также отличается высокой степенью совместимости с современными устройствами и платформами. AAC поддерживается большинством смартфонов, планшетов, операционных систем, а также популярными сервисами, такими как YouTube, Spotify, Apple Music и многими другими. Это обеспечивает пользователям бесперебойное воспроизведение контента независимо от используемого оборудования.

В то же время у AAC имеются и определённые ограничения. Одно из них связано с вопросами лицензирования: некоторые технологии, лежащие в основе кодека, защищены патентами, и их использование в коммерческих продуктах может требовать выплаты роялти. Хотя многие реализации (например, в составе FFmpeg) доступны бесплатно, юридическая неопределённость иногда затрудняет интеграцию AAC в открытые или бюджетные проекты.

Ещё один недостаток – ограниченная поддержка на устаревших устройствах. Несмотря на широкое распространение в современных системах, некоторые старые MP3-плееры, бюджетные аудиоприставки или встраиваемые устройства могут не распознавать AAC-файлы. Это может создавать неудобства при попытке воспроизвести контент на «наследственном» оборудовании, особенно вне экосистемы Apple, где формат исторически получил наибольшее распространение.

Таким образом, AAC представляет собой сбалансированный выбор для современных аудиоприложений, сочетающий высокое качество, компактность и хорошую совместимость, однако при проектировании мультимедийных систем, ориентированных на максимально широкую аудиторию, необходимо учитывать возможные ограничения, связанные с патентованием и поддержкой на устаревших платформах.

4. OGG (Ogg Vorbis)

OGG – это контейнерный формат с открытым исходным кодом, разработанный проектом Xiph.Org Foundation, который чаще всего используется в связке с аудиокодеком Vorbis для сжатия звука. Благодаря

своей свободной лицензии и высокой эффективности сжатия, OGG получил распространение в сообществах разработчиков открытого программного обеспечения, в инди-играх, онлайн-радиостанциях, а также в приложениях, где критически важны низкая задержка воспроизведения и высокое качество звучания при умеренном потреблении ресурсов.

Одним из главных преимуществ формата является его статус открытого стандарта. OGG не обременён патентными ограничениями и не требует выплаты лицензионных отчислений, что делает его особенно привлекательным для разработчиков ПО, создателей игр и авторов мультимедийного контента, стремящихся избежать юридических и финансовых сложностей, связанных с проприетарными кодеками. Эта особенность способствует широкому внедрению OGG в экосистему свободного программного обеспечения, включая такие платформы, как Linux и различные медиаплееры с открытым исходным кодом.

Кроме того, OGG/Vorbis обеспечивает высокое качество звука даже при сравнительно низких битрейтах. Многочисленные тесты показывают, что при одинаковом битрейте (например, 128 кбит/с) OGG часто превосходит MP3 по прозрачности звучания, особенно в передаче высоких частот и сложных акустических текстур. Это делает формат предпочтительным выбором для потоковой передачи аудио, онлайн-радио и других сценариев, где требуется оптимальное соотношение между качеством и объёмом трафика.

Формат также отличается гибкостью кодирования. Хотя наиболее известен в паре с Vorbis, контейнер OGG поддерживает и другие кодеки, такие как Opus (для речи и низколатентного аудио) и FLAC (для сжатия без потерь). Это позволяет разработчикам выбирать оптимальную стратегию хранения и передачи звука в зависимости от задач – будь то высококачественная музыка, голосовая связь в реальном времени или архивирование аудиозаписей.

Вместе с тем, OGG сталкивается с рядом ограничений, главным из которых является недостаточная распространённость по сравнению с такими форматами, как MP3 или AAC. Несмотря на растущую популярность в нишевых и технически ориентированных кругах, он всё ещё остаётся менее узнаваемым среди массовых пользователей и производителей потребительской электроники.

Это напрямую влияет на уровень поддержки: многие коммерческие устройства – особенно старые модели смартфонов, портативные плееры и автомобильные аудиосистемы – не обеспечивают нативного воспроизведения OGG-файлов. Даже некоторые современные веб-браузеры и мобильные ОС могут требовать установки дополнительных

кодеков или сторонних приложений для корректной работы с этим форматом. Для обычных пользователей, не знакомых с тонкостями мультимедийных технологий, это может стать серьёзным препятствием при попытке воспроизвести контент.

Таким образом, OGG представляет собой мощный и свободный инструмент для хранения и передачи аудио, особенно в средах, где ценятся открытость, эффективность и контроль над технологическим стеком. Однако при разработке массовых приложений или сервисов, ориентированных на максимально широкую аудиторию, необходимо учитывать ограниченную совместимость формата с повсеместно используемыми устройствами и платформами.

5. Сравнительный анализ

Выбор аудиоформата при разработке голосовых чатов критически влияет на качество речи, задержку, объём передаваемых данных и совместимость с клиентскими платформами [5].

Формат WAV обеспечивает безупречное качество звука, однако отсутствие сжатия приводит к чрезмерному потреблению пропускной способности и дискового пространства, что делает его непрактичным для передачи голоса в реальном времени.

MP3, несмотря на универсальную поддержку, использует устаревший алгоритм сжатия, слабо оптимизированный под речевые сигналы. При типичных для голосовых сообщений низких битрейтах теряется разборчивость, а латентность кодирования остаётся высокой. Хотя патенты на MP3 истекли, формат уступает современным решениям по эффективности и качеству.

AAC предлагает улучшенное качество звука и поддерживает низколатентные профили (AAC-LD, AAC-ELD), активно применяясь в VoIP-сервисах. В то же время часть реализаций требует лицензирования, а поддержка вне экосистемы Apple может быть ограничена.

Наиболее выгодным решением является OGG, особенно в связке с кодеком Opus. Opus обеспечивает отличное качество речи даже при 64 кбит/с, сверхнизкую задержку (5–20 мс), адаптивный битрейт и полную свободу от патентных ограничений. Формат нативно поддерживается в WebRTC – ключевой технологии для браузерных голосовых и видеочатов – и широко используется в современных медиасерверах и коммуникационных библиотеках.

Сравнительные характеристики всех рассматриваемых аудиоформатов по параметрам, важным для решения поставленной задачи, представлены в таблице.

Таблица

Сравнительный анализ различных форматов хранения

№ п/п	Параметр	MP3	WAV	AAC	OGG(с Opus)
1	Тип сжатия	С потерями	Без сжатия	С потерями	С потерями(адаптивный)
2	Качество речи (при ~64 кбит/с)	Удовлетворительное	Отличное	Хорошее	Отличное
3	Задержка (латентность)	Высокая (>100 мс)	Низкая (теоретически), но не используется из-за размера	Средняя/низкая (в LD-профилях)	Очень низкая (5–20 мс)
4	Размер файла / трафик	Умеренный	Очень большой	Компактный	Очень компактный
5	Лицензирование	Истёкшие патенты	Свободный	Частично патентованный	Полностью свободный
6	Пригодность для голосовых чатов	Средняя	Низкая	Высокая	Очень высокая

Таким образом, OGG (с Opus) сочетает высокую эффективность, минимальную латентность и юридическую безопасность, что делает его оптимальным выбором для современных голосовых чат-систем.

Заключение

В данной статье проведён анализ основных форматов хранения аудиоданных – MP3, WAV, AAC и OGG – с точки зрения их применимости в системах голосовых чатов. Рассмотрены технические характеристики каждого формата, включая качество звука, степень сжатия, задержку, совместимость и лицензионные ограничения. Особое внимание уделено особенностям обработки речевого сигнала,

требованиям к пропускной способности и поддержке в современных коммуникационных технологиях, таких как WebRTC.

На основе сравнительного анализа установлено, что формат OGG, особенно в связке с кодеком Opus, демонстрирует наилучшее сочетание низкой латентности, высокой эффективности сжатия, отличного качества передачи речи и полной свободы от патентных ограничений. Это делает его наиболее подходящим выбором для разработки масштабируемых, кроссплатформенных и юридически безопасных голосовых чат-систем.

Полученные выводы могут служить основой для архитектурных решений при проектировании мультимедийных компонентов в реальных приложениях, где критически важны надёжность, производительность и соответствие принципам открытых технологий.

Список литературы

1. Сэломон, Д. Сжатие данных, изображений и звука / Д. Сэломон. – Москва: Техносфера, 2004. – 368 с.

2. Ковалгин, Ю. А. Аудиотехника : учебник / Э. И. Вологдин, Ю. А. Ковалгин. – Москва: Горячая линия – Телеком, 2013 – 743 с.

3. Артюшенко, В. М. Цифровое сжатие видеоинформации и звука : учеб. пособие / В. М. Артюшенко, О. И. Шелухин, М. Ю. Афонин; под ред. В. М. Артюшенко. – М.: Дашков и К°, 2003. – 426 с.

4. Кинцель Т. Руководство программиста по работе со звуком / Т. Кинцель; пер. с англ. Злобин И. Г. – М.: ДМК Пресс, 2017. – 432 с.

5. Цифровая обработка сигналов : учеб. Пособие / Ю. Н. Матвеев [и др.]. – СПб. : СПбНИУ ИТМО, 2013. – 166 с.

Теоретические и методологические основы интеграции AI-ассистентов в процесс промышленной разработки ПО

И. Ю. Молоков, email: ilya132465@mail.ru
Е. А. Михайлюк, email gazon1978@yandex.ru

Воронежский Государственный Университет

Аннотация. *Проанализированы теоретические предпосылки и методологические проблемы интеграции AI-ассистентов для генерации кода в регламентированный процесс промышленной разработки программного обеспечения. Выявлено ключевое противоречие между стохастической природой генеративных моделей и требованиями к детерминизму, контролю и качеству в инженерных процессах. На основе системного обзора технологических возможностей современных ассистентов и критических требований промышленной разработки обоснована необходимость создания комплексной методологии оценки и интеграции, призванной минимизировать риски и максимизировать полезный эффект от внедрения искусственного интеллекта.*

Ключевые слова: *промышленная разработка ПО, AI-ассистент, большие языковые модели, методология интеграции, критерии оценки, DevOps, управление качеством кода*

Введение

Эволюция инструментария разработки программного обеспечения демонстрирует устойчивый тренд от автоматизации механических операций к интеллектуальной поддержке задач программиста. Следующим качественным скачком, приближающимся к смене парадигмы, стало появление AI-ассистентов для генерации кода на базе больших языковых моделей, таких как GitHub Copilot, Amazon CodeWhisperer и Tabnine. Эти системы способны интерпретировать контекст задачи на естественном языке и генерировать соответствующие фрагменты кода, тестов и документации, обещая радикальное повышение индивидуальной продуктивности разработчиков.

Однако их стихийное внедрение в промышленные процессы, характеризующиеся масштабируемостью, управляемостью и высокими требованиями к качеству и безопасности, сопряжено с существенными рисками. К ним относятся: снижение качества и безопасности кодовой базы из-за некритичного принятия сгенерированных решений, проблемы с лицензионной чистотой и конфиденциальностью исходного кода,

© Молоков И. Ю., Михайлюк Е. А., 2026

нарушение принципов управляемости процесса. Традиционные методы интеграции инструментов не учитывают стохастическую природу генеративного ИИ, что приводит к проблемам в процедурах код-ревью, измерении производительности и обеспечении архитектурной целостности.

Актуальной научно-практической проблемой является отсутствие комплексной методологии, которая позволяла бы организациям системно оценивать, выбирать и безопасно интегрировать AI-ассистенты в промышленный цикл разработки ПО. Эволюция инструментов разработки и место AI-ассистентов в современном технологическом ландшафте

Исторический путь инструментов разработки — от компиляторов и интегрированных сред до систем интеллектуального дополнения кода — направлен на повышение уровня абстракции и снижение когнитивной нагрузки на программиста. Появление AI-ассистентов, основанных на больших языковых моделях, дообученных на массивных корпусах кода, знаменует переход к принципиально новому типу взаимодействия разработчик-машина. В контексте данной работы под AI-ассистентом для генерации кода понимается система, способная в интерактивном режиме, на основе анализа контекста проекта и текстовых инструкций, генерировать, дополнять, изменять и документировать исходный код.

Их стремительное распространение обусловлено несколькими факторами: роста сложности программных систем, дефицита квалифицированных кадров, давления на скорость вывода продукта на рынок и прорывов в области машинного обучения. Однако данный инструмент, в отличие от предшественников, не является детерминированным. Его творческая способность генерировать код сопряжена с рисками галлюцинаций — создания синтаксически корректного, но семантически неверного или небезопасного кода. Это фундаментальное свойство создает основное противоречие при попытке встроить ассистента в предсказуемый промышленный процесс.

Целью данной статьи является систематизация ключевых проблем на стыке технологий генеративного ИИ и инженерии программного обеспечения, предложение обоснованных контуров комплексной методологии для оценки и контролируемой интеграции AI-ассистентов в процесс промышленной разработки ПО.

1. Анализ современных AI-ассистентов: архитектура, классификация и ограничения

Современные AI-ассистенты для генерации кода представляют собой качественный скачок в эволюции инструментов разработчика. Их можно классифицировать по нескольким ключевым признакам:

– Модель развертывания: облачные решения (Software-as-a-Service) и локальные (on-premise), обеспечивающие больший контроль над данными.
– Степень интеграции: плагины для интегрированных сред разработки (IDE), обеспечивающие бесшовную работу, и отдельные приложения или API.
– Специализация: универсальные модели, обученные на коде множества языков, и доменно-ориентированные решения для конкретных экосистем (например, облачных платформ).

Архитектурно ядро большинства ассистентов составляет большая языковая модель, дообученная на обширных корпусах публичного исходного кода [1]. Критическим компонентом является подсистема формирования контекстного промпта, которая динамически собирает из среды разработки релевантную информацию: текущий файл, открытые вкладки, комментарии, ошибки. Для повышения релевантности и точности ответов применяются техники дообучения на внутренних данных и извлечения релевантных фрагментов кода для обогащения контекста.

Ключевые возможности включают: генерацию кода и тестов по текстовому описанию, автодополнение с учетом семантики проекта, рефакторинг, объяснение и документирование существующего кода, а также предложение исправлений для ошибок.

Несмотря на впечатляющие возможности, для промышленного использования критичны следующие ограничения:

1. Проблемы качества и безопасности: модели могут генерировать код с уязвимостями, неоптимальными алгоритмами или скрытыми дефектами [2].
2. Юридические риски: неясность с лицензированием кода, на котором обучена модель, и авторскими правами на сгенерированный артефакт.
3. Риски конфиденциальности: при использовании облачных сервисов исходный код компании передается на сторонние серверы.
4. Эргономические риски: потенциальная деградация навыков разработчиков и накопление неизвестного технического долга [3].

В итоге можно сказать, что ассистент является очень полезным, но потенциально опасным инструментом, требующим не случайного использования, а продуманной стратегии внедрения.

2. Процесс промышленной разработки ПО: критические требования к инструментарию

Промышленная разработка ПО характеризуется работой в рамках регламентированных методологий, строгими требованиями к качеству, безопасностью и необходимостью управления командными процессами [3]. Инструмент в таком контексте становится частью корпоративной инфраструктуры и должен соответствовать ключевым практикам: непрерывной интеграции и доставки, код-ревью, управлению качеством и безопасностью.

Следовательно, к AI-ассистенту, выдвигаются критически важные требования:

1. Контроль и управляемость: возможность централизованной настройки, аудита действий, интеграции с системами управления задачами.
2. Измеряемость: способность количественно оценить влияние на метрики процесса.
3. Безопасность и соответствие: совместимость с внутренними политиками безопасности и отраслевыми стандартами.
4. Эргономика и внедрение: бесшовная интеграция в существующие рабочие процессы без радикального их изменения.

Интеграция AI-ассистента затрагивает не только этап непосредственного написания кода, но и трансформирует смежные фазы жизненного цикла программного обеспечения. На этапе проектирования и анализа требований ассистент может использоваться для быстрого прототипирования интерфейсов или генерации скелета системы по текстовому описанию, однако это требует от аналитиков навыков точного формулирования промптов [4].

При тестировании возможности ассистента по генерации модульных и интеграционных тестов могут значительно повысить покрытие кода, но создают необходимость в последующей валидации релевантности и полноты сгенерированных тестовых сценариев. Фаза сопровождения и рефакторинга также претерпевает изменения: ассистент способен предлагать локализованные улучшения, но его рекомендации по глобальному рефакторингу архитектуры могут быть поверхностными, так как модель не обладает полным пониманием системных взаимосвязей и бизнес-контекста.

В итоге на каждой фазе возникает новая задача: не просто использовать инструмент, но и выстраивать вокруг него точки контроля, гарантирующие, что ускорение выполнения задач не приведет к

снижению системного качества и не нарушит целостность архитектурного видения.

Внедрение AI-ассистента в практики DevOps поднимает вопрос о распределении ответственности. Если сгенерированный код содержит уязвимость, прошедшую через автоматизированные этапы CI/CD, ответственность ложится не только на разработчика, но и на всю цепочку: инженеров по безопасности, отвечающих за настройку SAST-инструментов, и команду, утвердившую процесс код-ревью с участием ИИ. Это требует перехода от культуры «ты сломал сборку» к культуре «система позволила пропустить дефект». Соответственно, интеграция должна сопровождаться модификацией самого конвейера: введение дополнительных этапов валидации, например, статического анализа, специфически настроенного на выявление паттернов, характерных для галлюцинаций ИИ, или этапа автоматической проверки сгенерированного кода на соответствие внутренним стандартам с помощью кастомизированных систем анализа кода.

Именно на стыке этих жестких требований и недетерминированной природы AI-ассистентов возникает основная методологическая проблема.

3. Проблема интеграции: вызовы и нерешенные вопросы

Анализ литературы показывает фрагментарность существующих исследований: технологические работы фокусируются на архитектуре моделей, эмпирические — на индивидуальной продуктивности, а вопросы безопасности рассматриваются изолированно. Комплексный процессно-ориентированный подход отсутствует.

Конкретные вызовы при интеграции включают:

– Код-ревью: как оценивать код, логика которого не до конца понятна человеку, и кто несет ответственность за дефект.

– Метрики производительности: традиционные показатели теряют смысл, возникает вопрос об измерении реального вклада разработчика и возврата на инвестиции от ассистента.

– Управление качеством кодовой базы: как обеспечить соблюдение стандартов кодирования и архитектурных паттернов.

– Организационные изменения: требуются ли новые роли и как адаптировать процессы обучения [5].

За рамками технологических и процессных сложностей лежат значительные психологические и культурные факторы. Часть опытных разработчиков может воспринимать AI-ассистент как угрозу своей экспертизе или как инструмент, обесценивающий мастерство программирования, что приводит к скрытому сопротивлению или формальному, неэффективному использованию. Другая часть, особенно

начинающие специалисты, может проявлять избыточное доверие к предложениям ИИ, некритично их принимая, что тормозит развитие их собственных аналитических и отладочных навыков [6].

Формируется новая форма технической зависимости, где разработчик становится «оператором промптов», теряя глубокое понимание генерируемых решений. Преодоление этих барьеров требует целенаправленной работы по изменению культуры внутри команды: открытого обсуждения роли ИИ как ассистента, а не замены, проведения внутренних воркшопов по эффективному промпт-инжинирингу и критической оценке сгенерированного кода, а также пересмотра программ обучения и адаптации новых сотрудников с акцентом на синергию человеческого и искусственного интеллекта.

Отдельным неисследованным вопросом остается долгосрочное влияние AI-ассистентов на инновационные способности и креативность команды. С одной стороны, освобождение от рутинного кодирования должно высвобождать интеллектуальные ресурсы для решения более сложных архитектурных и исследовательских задач. С другой стороны, существует риск «алгоритмизации» мышления, когда разработчики начинают мыслить в рамках паттернов, наиболее часто генерируемых ИИ, что может сузить спектр рассматриваемых решений и подавить поиск нестандартных, но более оптимальных подходов.

Долгосрочная методология интеграции должна включать механизмы, стимулирующие инновации, например, регулярные ретроспективы, посвященные анализу решений, предложенных ИИ, и поиску альтернатив, или введение роли «архитектора-критика», чьей задачей является оценка сгенерированных решений с точки зрения долгосрочной эволюции системы, а не только сиюминутной функциональности.

Анализ существующих академических и отраслевых подходов показывает, что они, как правило, фокусируются на решении отдельных изолированных проблем: проводятся исследования по безопасности генерируемого кода, предлагаются точечные метрики производительности, обсуждаются юридические аспекты. Однако комплексной методологии, которая бы связывала эти технологические, процессные и человеческие аспекты в единый жизненный цикл оценки, внедрения, эксплуатации и постоянного улучшения AI-ассистента в промышленной среде, на данный момент предложено не было. Этот пробел и обуславливает необходимость разработки структурированного подхода, контуры которого представлены в следующем разделе.

4. Контуры методологии системной интеграции AI-ассистентов

На основе проведенного анализа проблем и требований предлагаются контуры комплексной методологии, предназначенной для системной оценки, внедрения и управления AI-ассистентами в промышленной разработке ПО. Методология носит итерационный характер и строится на нескольких взаимосвязанных принципах.

Основополагающим является принцип «контролируемой адаптации», согласно которому внедрение ИИ-инструмента должно быть поэтапным, измеряемым и обратимым. Методология предлагает циклический процесс, начинающийся с этапа стратегической оценки и выбора, на котором определяется соответствие инструмента бизнес-целям и техническим требованиям компании.

За ним следует этап пилотного внедрения на ограниченном контуре (одна команда, один проект), основной задачей которого является сбор данных и отработка процессов, а не достижение максимальной продуктивности.

Третий этап — масштабирование и глубокая интеграция — предполагает внедрение инструмента в рабочие процессы всех целевых команд, его интеграцию в CI/CD-конвейер и системы мониторинга.

Завершает цикл этап постоянного мониторинга и оптимизации, на котором оценивается эффективность, выявляются новые риски и корректируются политики использования. Ключевым отличием от стихийного внедрения является обязательное наличие чек-поинтов между этапами, где на основе объективных метрик принимается решение о переходе на следующий уровень или откате изменений.

Для поддержки этапа выбора необходима многоаспектная таксономия критериев, выходящая за рамки функциональности. Предлагается оценивать инструменты по следующим категориям:

1. Функциональность и качество (точность генерации, поддержка языков, контекстное понимание);
2. Интегрируемость и управляемость (глубина интеграции с IDE/CI/CD, возможности аудита и настройки);
3. Безопасность и соответствие (модель обработки данных, наличие сертификатов, юридическая чистота);
4. Экономическая эффективность (модель лицензирования, оценка совокупной стоимости владения).

Каждому критерию должен соответствовать набор проверяемых метрик, например, процент сгенерированного кода, проходящего компиляцию и базовые тесты, или наличие API для централизованного управления политиками.

Методология требует модификации существующих инженерных практик. Процедура код-ревью должна быть дополнена: рецензенту вместе с кодом должны предоставляться исходный промпт и, по возможности, объяснение от ассистента; в чек-лист ревью включаются пункты на проверку специфичных для ИИ рисков. Конвейер CI/CD требует введения дополнительных автоматизированных шагов: статического анализа с правилами, выявляющими шаблонные ошибки генерации; автоматического прогона сгенерированных юнит-тестов; проверки лицензионных заголовков. Процесс управления качеством дополняется регулярным аудитом кодовой базы на предмет выявления плохо понятных, неэффективных или не соответствующих принципам чистого кода [6] сгенерированных фрагментов, которые могут стать источником технического долга.

Метрики для оценки эффективности. Для замены устаревших показателей предлагается многоуровневая система метрик.

На уровне кода это могут быть: процент принятых предложений ассистента, количество дефектов, ассоциированных со сгенерированным кодом, изменение цикломатической сложности.

На уровне процесса: влияние на время выполнения задачи для небольших изменений, время, затрачиваемое на код-ревью, частота дефектов, попадающих в продуктивный контур.

На организационном уровне: индекс удовлетворенности разработчиков, коэффициент принятия инструмента командой, оценка возврата на инвестиции через высвобождение времени senior-специалистов для более сложных задач. Сбор и анализ этих метрик должен быть автоматизирован и визуализирован для принятия управленческих решений.

Предложенные контуры задают структурированный подход, позволяющий перейти от хаотичного экспериментирования к управляемому и безопасному использованию технологий генеративного ИИ в промышленных масштабах.

Заключение

В данной работе проведен комплексный анализ проблемы интеграции AI-ассистентов для генерации кода в процесс промышленной разработки ПО. Выявлено ключевое противоречие между инновационным потенциалом этих инструментов и присущими им рисками для качества, безопасности и управляемости инженерного процесса. Показано, что существующие подходы носят фрагментарный характер и не предлагают целостного решения.

В качестве ответа на эту проблему предложены контуры комплексной методологии системной интеграции. В работе

сформулированы принципы методологии интеграции, включающие: итерационный подход к внедрению; обязательную оценку инструментов по многоаспектным критериям (функциональность, управляемость, безопасность, экономика); модификацию инженерных процессов, в первую очередь код-ревью и CI/CD; внедрение системы метрик на уровне кода, процесса и организации; а также меры по преодолению культурных и организационных барьеров. Данный подход призван обеспечить контролируемое и безопасное использование технологий генеративного ИИ, превращая их из источника неопределенности в управляемый фактор повышения производительности и качества.

Направлениями будущих исследований являются практическая апробация предложенной методологии в рамках пилотных проектов в промышленных командах, разработка специализированных инструментов для аудита сгенерированного кода и автоматического сбора метрик интеграции, а также углубленное изучение долгосрочного влияния AI-ассистентов на архитектурные решения и инновационный климат в организациях.

Список литературы

1. Evaluating Large Language Models Trained on Code [Электронный ресурс] / M. Chen [et al.] // arXiv.org e-Print archive. – 2021. – Режим доступа: https://arxiv.org/abs/2107.03374

2. Asleep at the Keyboard? Assessing the Security of GitHub Copilot's Generated Code [Электронный ресурс] / H. Pearce [et al.] // 2021 IEEE Symposium on Security and Privacy (SP). – 2021. – P. 8-10. – Режим доступа: https://ieeexplore.ieee.org/document/9833571#citations

3. Brynjolfsson, E. Generative AI at Work: The Economic and Productivity Impacts of AI Assistants in Software Engineering [Электронный ресурс] / E. Brynjolfsson, D. Li, L. R. Raymond // National Bureau of Economic Research (NBER) Working Paper Series. – 2024. – № 32122. – Режим доступа: https://www.researchgate.net/publication/370227920_Generative_AI_at_Wor

4. Productivity Assessment of Neural Code Completion [Электронный ресурс] / A. Ziegler [et al.] // Proceedings of the 6th ACM SIGPLAN International Symposium on Machine Programming (MAPS 2022). – 2022. – P. 25-28. – Режим доступа: https://arxiv.org/pdf/2205.06537

5. Building Secure and Reliable Systems / H. Adkins [et al.]. – Sebastopol, CA: O'Reilly Media, 2020. – 550 p.

6. Martin, R. C. Clean Code: A Handbook of Agile Software Craftsmanship / R. C. Martin. – Upper Saddle River, NJ: Prentice Hall, 2008. – 464 p.

Исследование временны́х характеристик выполнения задач реального времени в Linux

А. Ю. Савинков, email: savinkov_a_yu@sc.vsu.ru [1]

[1]ФГБОУ ВО «ВГУ»

Аннотация. *В работе представлены результаты экспериментов по измерению времени выполнения задач реального времени в системе на основе ядра Linux при использовании различных политик планирования в условиях высокой загрузки системы вычислительными задачами.*

Ключевые слова: *операционная система, задача реального времени, планирование загрузки процессоров, Linux.*

Введение

Задачи реального времени часто возникают в контексте реализации встраиваемых систем. Это могут быть задачи обработки сигналов в модемах и устройствах интернета вещей, в том числе обработка сигналов различных датчиков, или задачи управления исполнительными устройствами, например в станках с ЧПУ. Для задач реального времени существуют специализированные операционные системы с детерминированным планированием, которые гарантируют выполнение таких задач в заданные сроки, например FreeRTOS [4. 1] или Zephyr [4. 2].

Но помимо поддержки задач реального времени, не менее важными требованиями к современным встраиваемым системам является реализация развитого пользовательского интерфейса с использованием сложной графики и сенсорного экрана, а также сложного сетевого взаимодействия. Для этих задач удобнее использовать операционные системы общего назначения, например, на основе ядра Linux.

Поскольку современные аппаратные платформы для встраиваемых систем и интернета вещей как правило являются многоядерными, обычным решением для совмещения разнородных задач в рамках одного устройства стало использование ассиметричной многопроцессорной обработки (Asymmetric Multiprocessing, AMP), когда разные операционные системы, например Linux и FreeRTOS, независимо исполняются на разных процессорных ядрах с минимальным взаимодействием друг с другом через разделяемую память. Ведется даже разработка соответствующего стандарта для реализации таких систем [4. 3].

© Савинков А. Ю., 2026

Заметим, что для использованияAMPаппаратная платформа, как минимум, должна позволять явно назначать процессорные ядра для обработки запросов прерываний по каждому каналу запроса прерывания, что далеко не всегда выполняется, например, такой функциональности нет в широко распространенной платформе Raspberry Pi. Дополнительно потребуется реализация специальных программно-аппаратных решений для управления взаимодействием независимых операционных систем через разделяемую память. Кроме того, неизбежно возрастут накладные расходы реализации, как минимум, за счет необходимости размещения в памяти сразу двух операционных систем. При этом современная реализация ядра Linuxвключает развитые встроенные средства для поддержки задач реального времени, которых может оказаться вполне достаточно для многих практических случаев.

Целью данной работы является экспериментальное исследование временны́х характеристик выполнения задач реального времени в операционной системе на основе ядра Linux. Ожидается, что полученные результаты позволят разработчикам встраиваемых систем точнее оценить целесообразность использования ассиметричной многопроцессорной обработки в рамках реализации конкретных проектов и в ряде случаев обоснованно отказаться от нее в пользу использования встроенных возможностей ядра Linux.

1. Инструменты и методика исследования

В качестве аппаратной платформы для выполнения исследований использовался микрокомпьютер Raspberry Pi 4, построенный на основе однокристальной системы Broadcom BCM2711, включающей 4 ядра ARM Cortex-A72. Данная аппаратная платформа широко применяется в оборудовании встраиваемых систем и систем интернета вещей.

На микрокомпьютер Raspberry Pi 4 была установлена операционная система Raspberry Pi OS на основе ядра Linux v6.1.21.

Для измерения временных характеристик выполнения задач реального времени на языке C была написана программа, которая периодически активировалась электрическим сигналом, поступающим на одну из линий GPIO BCM2711 (будем называть эту линию GPIO линия активации). Периодический сигнал на линии активации с периодом Tформировался цифровым генератором сигналов UNI-T UTG962E. После активации программа выполняла простой цикл, обеспечивающий загрузку процессора и имитирующий работу программы задачи реального времени. Цикл выполнялся до исчерпания заданного интервала чистого процессорного времени $\tau < T$, после чего программа снова блокировалась в ожидании следующего сигнала на линии

активации. Перед началом выполнения цикла программа устанавливала высокий уровень сигнала на другой линии GPIO BCM2711 (будем называть эту линию GPIO линия индикации). После завершения цикла перед блокировкой программа снова устанавливала низкий уровень на линии индикации, что позволило с высокой точностью определять интервалы времени активной работы программы с помощью логического анализатора Kingst LA1010 с частотой опроса 100 МГц.

Блок-схема экспериментального стенда приведена на рис. 1.

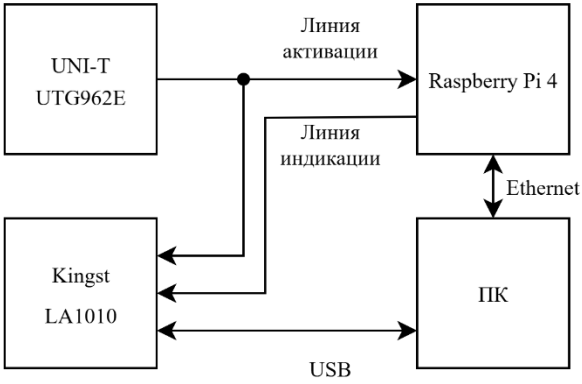


Рис. 1. Блок-схема экспериментального стенда

На ПК выполняется программа KingstVIS для управления логическим анализатором и сохранения выборок значений сигналов на линиях активации и индикации, а также программа putty для управления и мониторинга загрузки Raspberry по протоколу SSH.

Рассмотрим теперь более внимательно программу для измерения временных характеристик выполнения задач реального времени. Программа включает два потока: исследуемый поток реального времени и сервисный поток, обеспечивающий запуск исследуемого потока реального времени и ведение журнала.

Блок-схема алгоритма потока реального времени показана на рис. 2. Входными параметрами алгоритма являются целевая политика планирования, маска привязки к процессорам и требуемое для работы чистое процессорное время.

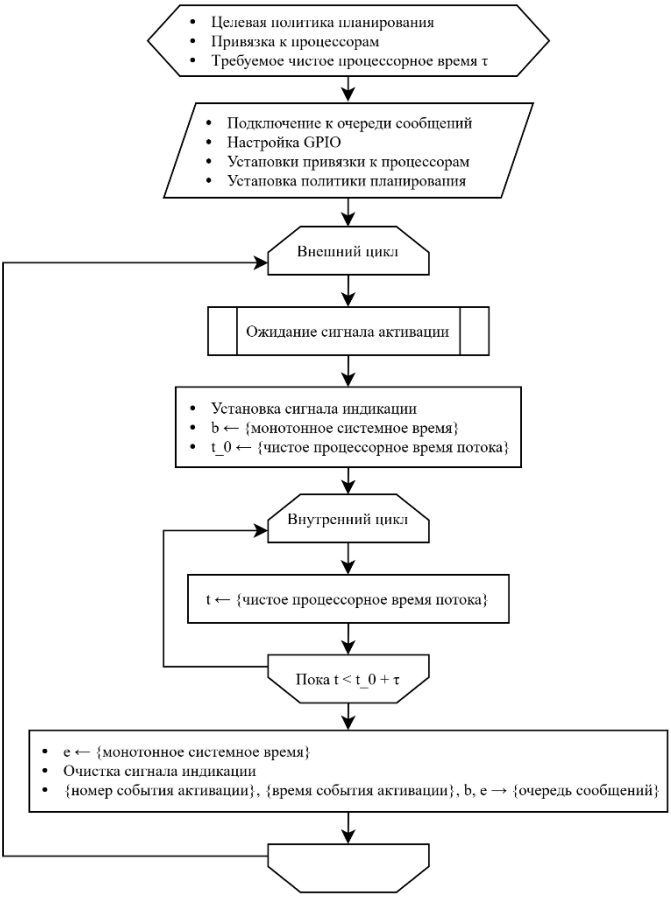


Рис. 2. Алгоритм потока реального времени

После запуска поток реального времени открывает очередь сообщений для взаимодействия с сервисным потоком, настраивает GPIO,устанавливает привязку к процессорам и политику планирования. Затем поток реального времени входит в бесконечный цикл и сразу же блокируется до наступления события активации.

Событие активации возникает при переходе сигнала на линии активации из низкого уровня в высокий. В программе вся работа с GPIO, включая ожидание событий GPIO и установку выходных сигналов GPIO, выполняется с использованием библиотеки libgpiod [4, 5].

После активации поток реального времени прежде всего устанавливает высокий уровень сигнала на линии индикации, а затем сохраняет временную метку b[egin] по системным часам CLOCK_MONOTONIC и временную метку t_0 по часам CLOCK_THREAD_CPUTIME_ID, которые считают только использованное процессорное время потока. Все часы имеют наносекундное разрешение.

Внутренний цикл выполняется, пока поток реального времени не использует заранее определенный интервал чистого процессорного времени. Этот цикл создает нагрузку на процессор и имитирует работу программы задачи реального времени.

После завершения внутреннего цикла поток реального времени сохраняет временную метку e[nd] по системным часам CLOCK_MONOTONIC и устанавливает низкий уровень сигнала на линии индикации.

Номер события активации, временная метка события активации, а также временные метки b и e записываются в очередь сообщений для последующего сохранения в файл журнала в сервисном потоке. Использование очереди сообщений позволяет исключить длительные файловые операции из программы потока реального времени. Номер события активации и временная метка события активации фиксируются библиотекой libgpiod в момент обработки прерывания GPIO и используются для обнаружения пропусков событий и контроля выполнения требований реального времени.

Журнал работы программы измерения временных характеристик выполнения задач реального времени и журнал логического анализатора впоследствии обрабатываются программойPython для анализа и визуализации полученных результатов.

2. Формирование фоновой вычислительной нагрузки и верификация экспериментального стенда

Для получения адекватных результатов, позволяющих надежно оценить возможность выполнения задач реального времени в системе Linux, необходимо проводить измерения в условиях высокой загрузки процессоров фоновыми задачами.

Для загрузки процессоров фоновыми вычислительными задачами использовалась утилитаstress-ng из одноименного пакета [6]. Утилита stress-ngможет формировать свыше 280 нагрузочных тестов для кэша, памяти, процессора, файловой системы и др. В данном исследовании использовался только тест ‘--cpu’ который реализует смешанные вычисления, обеспечивающие высокую загрузку процессора. Чтобы

гарантированно загрузить все процессорные ядра 4-ядерного процессора, запускалось 16 вычислительных потоков stress-ng.

Большинство современных операционных систем на основе ядра Linux в качестве подсистемы инициализации и управления используют systemd [7]. Подсистема system для каждого сеанса пользователя, в том числе сеанса SSH, создает отдельный профиль в иерархии групп управления (cgroup) и обеспечивает за счет этого справедливое разделение ресурсов между пользователями, не позволяя отдельным пользователям монополизировать все ресурсы системы. Поэтому утилиту stress-ng и исследуемую задачу реального времени необходимо запускать из одного сеанса.

Заметим также, что планирование реального времени является детерминированным и порядок выполнения нескольких задач реального времени друг относительно друга однозначно определен. Поэтому достаточно будет рассмотреть единственную задачу реального времени, работающую в сильно загруженной системе и конкурирующую за процессоры с другими вычислительными задачами.

Для периодической задачи реального времени удобно ввести в рассмотрение коэффициент загрузки процессора со стороны задачи реального времени $\eta = \tau \cdot T^{-1}$, где τ – величина чистого процессорного времени, необходимого для работы задачи реального времени на каждом периоде активности, $T > \tau$ – период активации задачи. Для получения репрезентативных результатов измерения должны быть проведены для разных значений η при различных политиках планирования выполнения задачи реального времени.

Для проверки экспериментального стенда задача реального времени была запущена с политикой планирования SCHED_OTHER, используемой в Linux по умолчанию, с коэффициентом загрузки процессора $\eta = 0.1$ и периодом активации 1 мс (τ = 100 нс). Ожидалось, что требования реального времени под нагрузкой заведомо не будут выполняться, но по данным логического анализатора, задача почти всегда успевала завершиться до следующей активации (см. рис. 3).

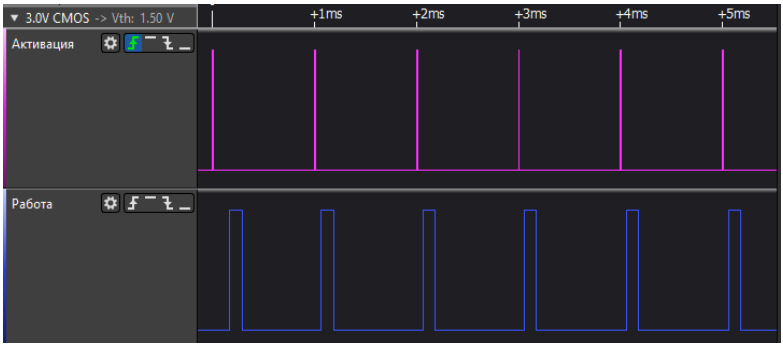


Рис. 3. Работа задачи реального времени с политикой планирования SCHED_OTHER

Причина такого поведения состоит в особенности планировщика CFS, используемого ядром Linux для политики планирования SCHED_OTHER. В планировщике CFS при выходе нормальной (не фоновой) задачи из состояния блокировки ее приоритет динамически повышается, и задача быстро ставится на выполнение. Пример показан на рис. 4 – задача была запущена примерно через 162 мкс после получения сигнала активации и завершилась менее, чем за 300 мкс после активации, т. е. задолго до конца периода $T = 1$ мс.

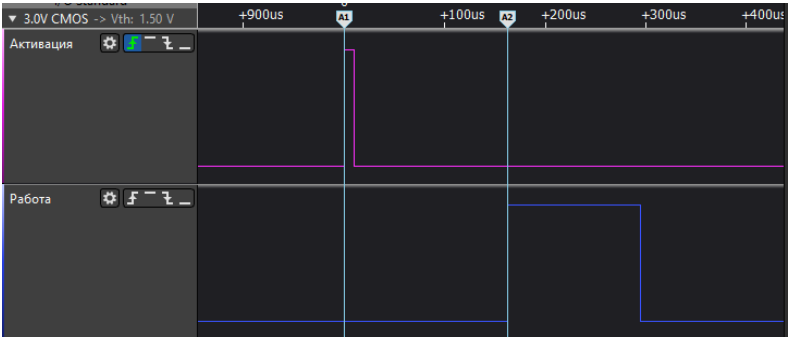


Рис. 4. Задержка активации задачи с политикой планирования SCHED_OTHER

Для решения данной проблемы был повышен приоритет потоков stress-ng (запуск командой nice -n -20 stress-ng --c 16), а сама задача реального времени в дальнейшем запускалась в фоновом режиме. В результате при использовании политик планирования SCHED_OTHER и

SCHED_BASTCH требования реального времени, как и ожидалось, перестали выполняться (см. рис. 5, 6).

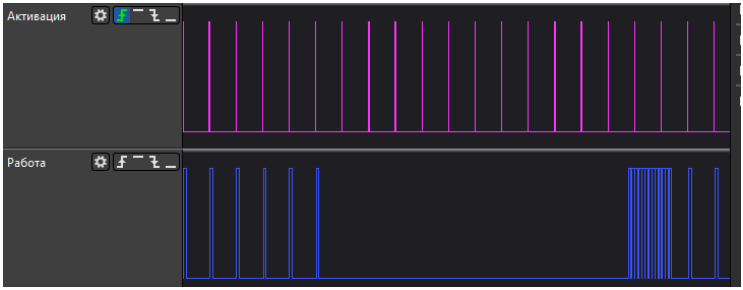


Рис. 5. Выполнение задачи реального времени с политикой планирования SCHED_OTHER на сильно загруженной системе

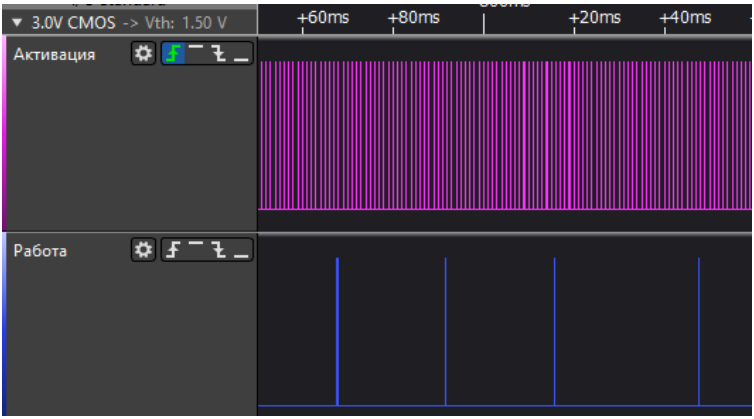


Рис. 6. Выполнение задачи реального времени с политикой планирования SCHED_BATCH на сильно загруженной системе

Для стабильной работы задачи реального времени потребовалось также изменить параметры планирования сервисного потока. Для него была установлена политика SCHED_RR с минимальным приоритетом 1. Без этого в сильно загруженной системе сервисный поток, приоритет которого по умолчанию ниже приоритета фоновых задач, мог долго не выбираться на исполнение, что приводило к переполнению очереди сообщений и блокировке исследуемого потока реального времени.

При этом требования реального времени при использовании политик планирования SCHED_DEADLINE и SCHED_FIFO по-прежнему выполнялись (см. рис. 7, 8).

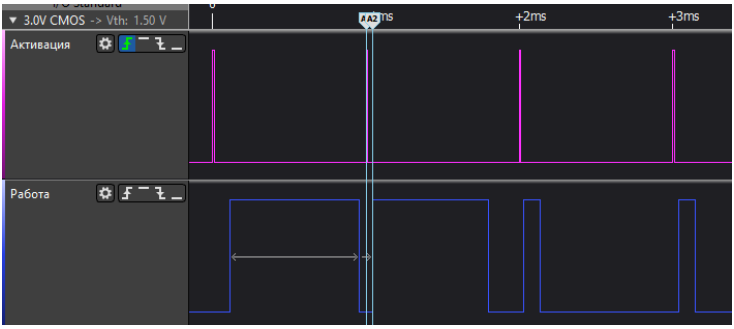


Рис. 7. Выполнение задачи реального времени с политикой SCHED_DEADLINE на сильно загруженной системе

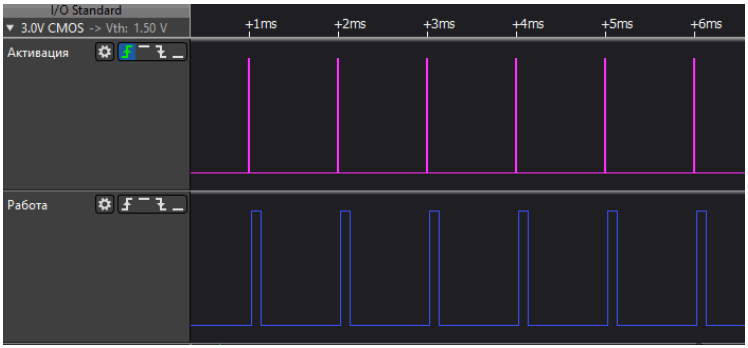


Рис. 8. Выполнение задачи реального времени с политикой планирования SCHED_FIFO на сильно загруженной системе

Для политики SCHED_DEADLINE предельное время завершения устанавливалось равным 850 мкс от начала периода активации. При этом, было замечено, что хотя задача реального времени всегда активировалась в начале периода непосредственно после получения сигнала активации, завершалась она иногда в самый последний момент (см. рис. 7).Т. е. для политики планирования SCHED_DEADLINE планировщик может приостанавливать выполнение задачи реального времени и передать процессор обычной задаче (не реального времени), если до истечения заявленного предельного срока завершения задачи реального времени еще есть достаточно времени. Это не согласуется с общепринятой

логикой работы планировщика EDF, используемого в политике планирования SCHED_DEADLINE – ожидалось, что при наличии единственной задачи реального времени она будет работать без переключения до завершения. Тем не менее, задача всегда завершалась вовремя и такую особенность планирования SCHED_DEADLINE в Linux просто следует учитывать в реальных задачах.

При использовании политики планирования SCHED_FIFO задача реального времени всегда запускалась в начале периода активации и затем без переключений контекста работала до завершения (см. рис. 8).

Таким образом, экспериментальный стенд и система формирования фоновой вычислительной нагрузки были верифицированы и показали ожидаемую работу.

3. Выделение изолированного процессорного ядра для задачи реального времени

Ядро Linux позволяет выделить отдельные процессоры и использовать их только для указанных задач. Для изоляции процессора необходимо указать соответствующие параметры в командной строке ядра при запуске операционной системы. В данном исследовании использовались следующие дополнительные параметры командной строки ядра:

– isolcpus=3 – изоляция третьего ядра процессора;

– idle=poll – в отсутствии готовых к выполнению задач выделенное ядро должно продолжить работу в бесконечном цикле ожидания задачи вместо перехода в режим энергосбережения;

– max_cstate=1 – установить режим C1 в качестве максимально допустимого режима энергосбережения для выделенного ядра.

Зарезервированное ядро не используется процессами без явного указания требования работы на этом ядре. Так утилита stress-ng теперь загружает только три ядра из четырех, см. рис. 9.

Заметим, что изолированное процессорное ядро не исключается из системы полностью, с ним остается связано несколько системных потоков, список которых можно получить командой: ps -ALo %cpu,pri,cmd,psr | grep 3$. Результат работы команды показан на рис. 10. На выделенном ядре была запущена задача реального времени gpio-based-test-v2 с политикой планирования SCHED_FIFO с $\eta = 0.8$, которая ожидаемо обеспечивает 80% загрузки процессора.

Как видно на рис. 10, помимо исследуемой задачи, явно запущенной на изолированном процессорном ядре, с изолированным ядром остаются связаны следующие системные потоки:

– cpuhp – поток для горячего подключения CPU;

– migration – поток для балансировки загрузки между процессорами;

– ksoftirqd – поток для вторичных обработчиков прерываний;

– kworker – потоки для обработки очередей отложенных действий.

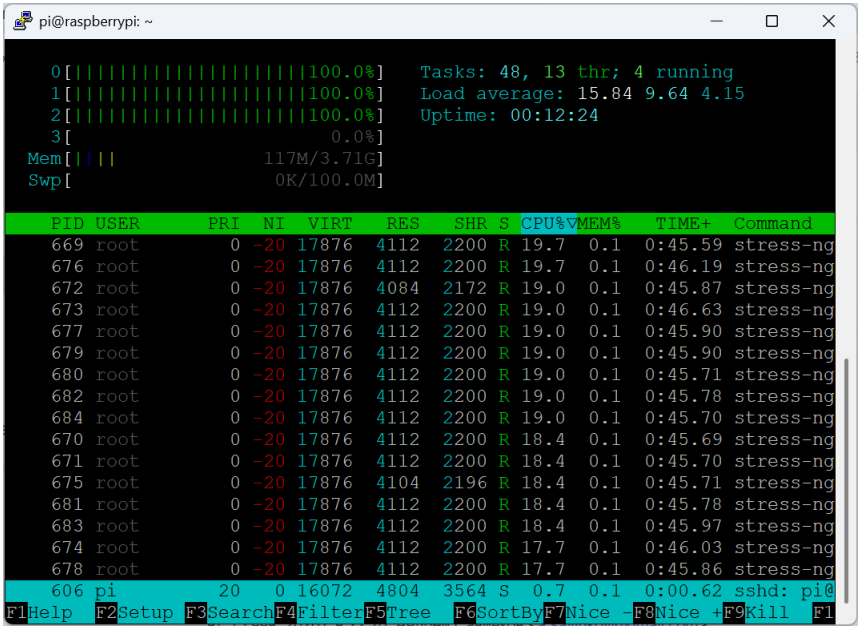


Рис. 9. Загрузка процессоров утилитой stress-ng после изоляции ядра номер 3 (снимок экрана утилиты htop)

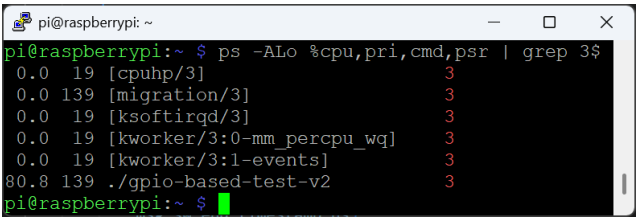


Рис. 10. Системные процессы на изолированном процессоре

Запретить обработку прерываний на конкретном ядре в RPi4 нельзя из-за особенностей архитектуры процессора, поэтому поток ksoftirqd, возможно, будет иногда работать, но он имеет невысокий приоритет.

Другие потоки должны оставаться неактивными на выделенном процессоре.

4. Результаты тестирования

При использовании выделенного процессора измерения выполнялись для политик планирования SCHED_BATCH, SCHED_OTHER, SCHED_FIFO и SCHED_DEADLINE при коэффициенте загрузки процессора η равном 0.1, 0.2, 0.3, 0.5 и 0.8.

При использовании политики планирования SCHED_DEADLINE для выполнения задачи запрашивалось время на 50 мкс большее чистого процессорного времени внутреннего цикла, а предельное время завершения указывалось равным 850 мкс от момента активации задачи. Заявлять немного увеличенное требуемое чистое процессорное время рекомендуется в описании политики планирования SCHED_DEADLINE, так как фактическое время выполнения всегда будет иметь некоторый разброс, например, из-за влияния кэша. Величина же предельного время завершения подобрана экспериментально. При измерениях по принятой в работе методике при больших значениях иногда фиксировались случаи невыполнения требований реального времени. Это связано с тем, что планировщик SCHED_DEADLINE отсчитывает предельное время завершения от момента фактической активации задачи, но этот момент времени неизвестен внешнему наблюдателю, так как зависит от ряда неконтролируемых факторов: задержки обработки прерывания, времени реакции планировщика, времени переключения процессов и др. Поэтому при интерпретации результатов измерений предельное время завершения отсчитывалось от сигнала активации. При значении 850 мкс задача всегда успевала завершиться до следующего сигнала активации.

Важно отметить, что при использовании единственного выделенного процессора не удается установить политику планирования SCHED_DEADLINE для $\eta = 0.8$ с предельным временем завершения, 850 мкс. Предельное время завершения при этом оказывается равным запрошенному для выполнения задачи чистому процессорному времени и системный вызов sched_setattr возвращает ошибку EBUSY. Поэтому случай $\eta = 0.8$ при использовании выделенного процессора не реализован. Интересно отметить, что если планировщику разрешено использовать более одного процессора (соответствует случаю без изоляции процессора), то установка предельного времени завершения, равного чистому процессорному времени задачи, является допустимым и требования реального времени выполняются.

При работе без выделенного процессора измерения выполнялись только для политик планирования SCHED_FIFO и SCHED_DEADLINE,

так как для политик планирования SCHED_BATCH и SCHED_OTHER требования реального времени заведомо не выполняются.

Логический анализатор работал на частоте 100 МГц и записывал выборку сигнала с линий активации и индикации длиной 10^9 отсчетов (10 секунд реального времени). Эта выборка включала 10^4 событий активации. По полученным выборкам с использованием программы на Python вычислялось время δ задержки запуска программы после сигнала активации, определяемое как интервал времени от переднего фронта сигнала активации до переднего фронта сигнала индикации, и времяΔвыполнения задачи, определяемое как интервал времени от переднего фронта сигнала активации до заднего фронта сигнала индикации.

Если в момент переднего фронта сигнала активации был установлен сигнал индикации, то фиксировалось невыполнение требований реального времени – запущенная ранее задача не успела завершиться.

Медианные и граничные значения δ для разных случаев загрузки процессора и различных политик планирования приведены в табл. 1. Медианные и граничные значения Δ – в табл. 2.

Таблица 1

Политика планирования	η	**Задержка запуска (δ), мкс**			
		На выделенном процессоре		**Без выделения процессора**	
		Me(δ)	min(δ) max(δ)	*Me*(δ)	min(δ) max(δ)
SCHED_BATCH	0.1	25.1	18.9 98.0	–	–
	0.2	25.3	18.8 94.0	–	–
	0.3	23.3	19.5 130.0	–	–
	0.5	25.1	18.3 112.4	–	–
	0.8	23.5	17.0 109.0	–	–

Продолжение табл.1

Политика планирования	η	**Задержка запуска (δ), мкс**			
		На выделенном процессоре		**Без выделения процессора**	
		Me(δ)	min(δ) max(δ)	*Me*(δ)	min(δ) max(δ)
SCHED_OTHER	0.1	25.1	19.7 86.1	–	–
	0.2	23.4	17.4 101.5	–	–
	0.3	24.0	17.8 118.9	–	–
	0.5	24.1	17.6 168.9	–	–
	0.8	24.4	17.9 86.6	–	–
SCHED_FIFO	0.1	22.9	16.5 90.6	27.6	21.2 139.9
	0.2	24.2	17.4 92.0	27.4	20.8 117.7
	0.3	23.5	18.2 78.9	26.1	18.6 120.6
	0.5	23.8	17.9 102.8	27.8	20.6 113.9
	0.8	21.9	16.8 119.4	25.0	18.3 111.9
SCHED_DEADLINE	0.1	21.5	16.5 111.3	27.7	19.8 162.5
	0.2	23.8	17.0 103.8	30.3	19.3 99.5
	0.3	24.8	18.8 105.4	21.4	15.5 155.1
	0.5	21.7	16.0 110.8	26.4	18.5 111.0
	0.8	–	–	20.1	14.6 108.8

Таблица 2

Политика планирования	η	**Время выполнения (Δ), мкс**			
		На выделенном процессоре		**Без выделения процессора**	
		Me(Δ)	min(Δ) max(Δ)	*Me*(Δ)	min(Δ) max(Δ)
SCHED_BATCH	0.1	131.8	124.6 224.3	–	–
	0.2	232.4	224.3 312.3	–	–
	0.3	330.1	324.9 442.9	–	–
	0.5	531.9	524.3 634.4	–	–
	0.8	830.4	827.7 917.2	–	–
SCHED_OTHER	0.1	132.0	125.1 213.8	–	–
	0.2	223.8	223.8 322.3	–	–
	0.3	323.6	323.6 426.5	–	–
	0.5	531.2	523.2 677.9	–	–
	0.8	831.4	824.0 897.0	–	–
SCHED_FIFO	0.1	130.0	123.8 211.7	134.7	126.8 257.3
	0.2	231.2	223.4 318.8	234.5	226.3 331.8
	0.3	330.3	324.1 404.3	333.1	324.1 454.7
	0.5	530.7	523.4 610.1	534.9	526.5 642.9
	0.8	829.0	823.4 927.1	832.0	823.8 948.0

Продолжение табл. 2

Политика планирования	η	**Время выполнения (Δ), мкс**			
		На выделенном процессоре		**Без выделения процессора**	
		Me(Δ)	min(Δ) max(Δ)	*Me*(Δ)	min(Δ) max(Δ)
SCHED_DEADLINE	0.1	128.0	122.4 907.1	134.9	125.3 954.0
	0.2	230.6	222.6 939.2	237.6	226.1 968.9
	0.3	331.6	324.2 901.0	328.9	321.3 921.2
	0.5	528.6	521.7 965.9	533.5	525.8 950.2
	0.8	–	–	827.2	820.9 937.5

Ядерная оценка плотности распределения задержки активации задачи для разных политик планирования приведена на рис. 11. Оценка получена с помощью класса stats.gaussian_kde из пакета scipy.

Поскольку на выделенном процессоре работала только одна задача, после активации она выполнялась без переключения контекста и всегда успевала завершиться до следующего сигнала активации. Таким образом, работа задачи реального времени на изолированном процессоре практически эквивалентна работе на незагруженной системе. Фактически, данные тесты всего лишь подтвердили качество изоляции процессора в ядре Linux.

При работе без изоляции процессора политика планирования SCHED_DEADLINE показала более быструю активацию задачи (рис. 12), но преимущество составило всего несколько микросекунд и не является существенным. Более существенным является хорошая предсказуемость времени завершения задачи, достигаемая при использовании политики планирования SCHED_FIFO.

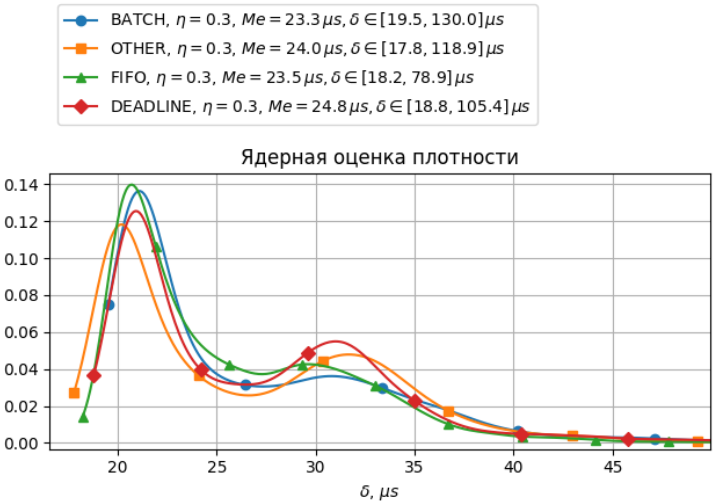


Рис. 11. Ядерная оценка плотности распределения задержки активации задачи для разных политик планирования на изолированном процессоре

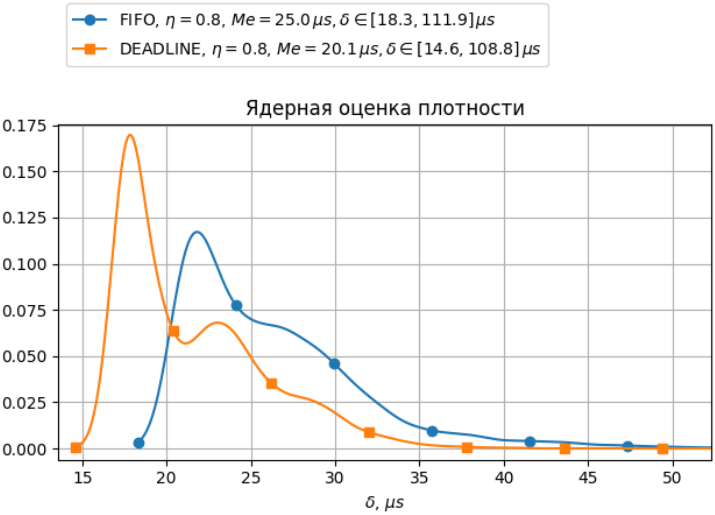


Рис. 12. Ядерная оценка плотности распределения задержки активации задачи для политик планирования SCHED_DEADLINE и SCHED_FIFO в сильно загруженной системе

Заключение

Проведенные исследования показали возможность использования задач реального времени в операционной системе на основе ядра Linux без необходимости применения асимметричной многопроцессорной обработки и отдельной операционной системы реального времени на изолированном процессоре в практических приложениях встраиваемых систем и устройств интернета вещей.

Список литературы

1. Официальный сайт проекта FreeRTOS в сети Интернет [Электронный ресурс]. – Режим доступа : https://www.freertos.org/
2. Официальный сайт проекта Zephyr в сети Интернет [Электронный ресурс]. – Режим доступа :https://www.zephyrproject.org/
3. Официальный сайт проекта OpenAMP (OpenAsymmetricMulti-Processing) в сети Интернет [Электронный ресурс]. – Режим доступа :https://www.openampproject.org/
4. Репозиторий проекта libgpiod на GitHub [Электронный ресурс]. – Режим доступа : https://github.com/brgl/libgpiod
5. Официальная документация libgpiod [Электронный ресурс]. – Режим доступа :https://libgpiod.readthedocs.io/en/stable/
6. Репозиторий проекта stress-ng на GitHub [Электронный ресурс]. – Режим доступа :https://github.com/ColinIanKing/stress-ng
7. Официальный сайт systemd в сети Интернет [Электронный ресурс]. – Режим доступа : https://systemd.io/

Применение трансформерной модели BERT для автоматического парсинга библиографических ссылок

В. А. Ушаков, email: aurel.vu@ya.ru

Воронежский государственный университет

Аннотация. *Рассматривается задача автоматического извлечения структурированной информации из библиографических ссылок с использованием распознавания именованных сущностей на основе BERT. Предложена методика подготовки данных с применением схемы BIO-разметки и алгоритма выравнивания позиций на основе расстояния Левенштейна. Модель обучена на наборе из 50 000 библиографических записей и достигает F1-меры 90,3%, что превосходит традиционные подходы на основе правил.*

Ключевые слова: *распознавание именованных сущностей, BERT, парсинг библиографий, обработка естественного языка, глубокое обучение.*

Введение

Экспоненциальный рост объема научной литературы создает острую потребность в автоматизированных системах обработки библиографической информации. По различным оценкам, количество научных публикаций удваивается каждые 12–15 лет, что создает серьезные вызовы для систем управления знаниями и поиска релевантной информации. Библиографические ссылки являются критически важным элементом научной коммуникации, обеспечивая связность между исследованиями и позволяя отслеживать развитие научных идей.

Традиционные подходы к обработке библиографических данных основаны на ручной курации или использовании жестко заданных правил и регулярных выражений. Такие методы демонстрируют низкую точность при работе с разнообразными форматами цитирования [1]. Многообразие стилей цитирования (APA, MLA, Chicago, IEEE и др.) делает создание универсальных правил практически невозможным. Существующие системы, такие как GROBID [2] и CERMINE [3], хотя и показывают достойные результаты, требуют значительных вычислительных ресурсов или ручной коррекции.

Целью данного исследования является разработка системы извлечения структурированной информации из библиографических

© Ушаков В. А., 2026

ссылок с использованием современных методов обработки естественного языка на основе трансформерных архитектур.

1. Методика исследования

Предложенная система представляет собой конвейер обработки библиографических данных, включающий четыре основных модуля: извлечение данных из хранилища, предобработка и форматирование, токенизация с BIO-разметкой, а также модуль машинного обучения на основе трансформерной архитектуры.

В качестве источника данных использовалась коллекция библиографических записей в формате BibTeX, собранная из открытых научных репозиториев (arXiv, DBLP, SemanticScholar). Исходный набор содержал записи различных типов публикаций: статьи в журналах (62%), материалы конференций (31%), книги и главы книг (7%). Для обеспечения разнообразия выборка включала публикации на английском (78%), немецком (8%), французском (6%) и других языках (8%).

Форматирование библиографических записей выполнялось с помощью библиотеки pybtex с использованием встроенного стиля plain. Из каждой структурированной записи BibTeX генерировалась текстовая ссылка, которая затем использовалась для обучения модели извлечению сущностей из неструктурированного текста.

Ключевой проблемой при создании разметки является точное определение позиций целевых сущностей в отформатированном тексте, поскольку форматирование вносит изменения в исходные значения полей (добавление инициалов, сокращения, пунктуация). Для решения этой задачи разработан алгоритм выравнивания на основе расстояния Левенштейна со следующими этапами:

1. нормализация целевой строки и текста ссылки (приведение к нижнему регистру, удаление лишних пробелов);
2. генерация всех подстрок текста длиной от 0,5 до 1,5 длины целевой строки;
3. вычисление нормализованного расстояния Левенштейна для каждой подстроки;
4. фильтрация кандидатов по критерию совпадения первого и последнего символов;
5. выбор подстроки с минимальным расстоянием при пороге сходства не менее 0,7.

Данный алгоритм позволяет корректно обрабатывать случаи частичного совпадения, например, когда полное имя автора «JacobDevlin» в форматированной ссылке представлено как «J. Devlin», или когда название журнала сокращается.

В качестве базовой модели выбрана DistilBERT-multilingual-cased [4] — облегченная версия BERT [5] с мультиязычной поддержкой, обученная на корпусе из 104 языков. Модель содержит 6 слоёв трансформера (вместо 12 в оригинальном BERT), 768-мерные скрытые представления и 12 голов внимания. Общее количество параметров составляет 134 млн, что на 40% меньше базовой модели BERT при сохранении 97% точности на задачах понимания языка.

Для адаптации модели к задаче токенной классификации добавлен линейный слой размерности 768×11, где 11 — количество меток BIO-разметки (Begin-Inside-Outside) для пяти типов сущностей:

- B-authors, I-authors — начало и продолжение имён авторов;
- B-title, I-title — начало и продолжение заголовка публикации;
- B-journal, I-journal — начало и продолжение названия журнала или конференции;
- B-year, I-year — год публикации;
- B-doi, I-doi — цифровой идентификатор объекта;
- O — токены, не относящиеся к целевым сущностям.

Обучение модели проводилось в течение 5 эпох со следующими гиперпараметрами: размер пакета — 10 образцов, скорость обучения — 2×10^{-5} с линейным затуханием, оптимизатор AdamW с коэффициентом регуляризации 0,01. Максимальная длина последовательности ограничена 256 токенами, что покрывает 98% записей в выборке. Набор данных из 50 000 записей разделён на обучающую (80%), валидационную (10%) и тестовую (10%) выборки с сохранением пропорций по типам публикаций и языкам.

2. Экспериментальные результаты

Экспериментальная оценка проводилась на тестовой выборке из 5 000 библиографических записей, не использовавшихся при обучении. Средняя длина ссылки составила 127 токенов. Результаты оценки качества по типам сущностей представлены в табл. 1.

Таблица 1

Метрики качества извлечения по типам сущностей

Сущность	Precision	Recall	F1-Score
Authors	0,923	0,918	0,920
Title	0,897	0,891	0,894
Journal	0,874	0,869	0,871
Year	0,941	0,938	0,939
DOI	0,918	0,912	0,915
Среднее	0,906	0,901	0,903

Как видно из табл. 1, система демонстрирует высокие показатели для всех типов сущностей. Наилучшие результаты достигнуты для извлечения года публикации (F1=0,939), что объясняется стандартизированным форматом представления дат. Наиболее сложной задачей оказалось извлечение названий журналов (F1=0,871) из-за вариативности сокращений и форматов записи.

Для иллюстрации работы системы приведём пример обработки библиографической ссылки. Входнаястрока: «Devlin J., Chang M.-W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // Proc. NAACL-HLT. 2019. P. 4171–4186.» Результат извлечения:

- Authors: Devlin J., Chang M.-W., Lee K., Toutanova K.
- Title: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding
- Journal: Proc. NAACL-HLT
- Year: 2019
- Pages: 4171–4186

Для оценки эффективности предложенного подхода проведено сравнение с существующими методами парсинга библиографических ссылок на том же наборе данных (табл. 2).

Таблица 2

Сравнение с существующими подходами

Подход	**F1-Score**	**Время (мс/запись)**
Rule-based (регулярные выражения)	0,701	0,8
CRF [1]	0,787	2,1
BiLSTM-CRF [6]	0,827	5,7
Предложенный подход (BERT)	0,903	8,9

Предложенный подход превосходит все базовые методы: улучшение F1-меры составило +28,8% относительно подхода на основе правил, +14,7% относительно CRF-модели и +9,2% относительно BiLSTM-CRF. Увеличение времени обработки (8,9 мс против 0,8–5,7 мс) компенсируется существенным приростом качества и остаётся приемлемым для практического применения — система способна обрабатывать более 100 записей в секунду.

Анализ ошибок показал, что основные проблемы возникают в следующих случаях:

- нестандартные сокращения названий журналов (32% ошибок);

– составные фамилии и транслитерация имён (24% ошибок);
– отсутствие явных разделителей между полями (21% ошибок);
– смешение языков в одной записи (15% ошибок);
– редкие форматы цитирования (8% ошибок).

Большинство ошибок связано с неоднозначностью границ сущностей, а не с полным пропуском или ложным обнаружением.

К ограничениям системы относятся: зависимость от качества входных данных (OCR-ошибки снижают точность на 10–15%), ограничение максимальной длины последовательности в 512 токенов (затрагивает менее 2% записей), а также необходимость GPU для эффективной работы в режиме реального времени. На CPU время обработки увеличивается в 5–8 раз.

Направлениями дальнейших исследований являются: интеграция с системами извлечения текста из PDF-документов, дообучение и адаптация модели для русскоязычных текстов и отечественных стилей цитирования (ГОСТ), а также исследование возможности применения более компактных моделей (TinyBERT, MobileBERT) для работы на мобильных устройствах.

Заключение

В статье представлена разработанная система автоматического извлечения структурированной информации из библиографических ссылок на основе трансформерной модели DistilBERT. Предложена методика подготовки обучающих данных с алгоритмом выравнивания позиций сущностей на основе расстояния Левенштейна, на основе которой создан размеченный набор данных из 50 000 библиографических записей. Достигнутая F1-мера 90,3% превосходит методы на основе CRF на 14,7% и BiLSTM-CRF на 9,2%. Также показана устойчивость системы к различным языкам публикаций и форматам записи.

Список литературы

1. Peng, F. Information extraction from research papers using conditional random fields / F. Peng, A. McCallum // Information Processing & Management. – 2006. – Vol. 42. –Iss. 4. – P. 963-979.

2. Lopez, P. GROBID: Combining Automatic Bibliographic Data Recognition and Term Extraction for Scholarship Publications / P. Lopez // Research and Advanced Technology for Digital Libraries : Proc. ECDL 2009. – Springer, 2009. – P. 473-474.

3. Automatic extraction of structured metadata from scientific literature / D. Tkaczyk [et al.] // International Journal on Document Analysis and Recognition. – 2015. – Vol. 18. –Iss. 4. – P. 317-335.

4. A distilled version of BERT: smaller, faster, cheaper and lighter / V. Sanh [et al.] // arXiv:1910.01108. – 2019. – 5 p.

5. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding / J. Devlin [et al.]// Proc. NAACL-HLT. – 2019. – P. 4171-4186.

6. Neural Architectures for Named Entity Recognition / G. Lample [et al.] // Proc. NAACL-HLT. – 2016. – P. 260-270.

Секция 2
Компьютерное моделирование в фундаментальных и прикладных исследованиях

Программный комплекс автоматизированного описания природно-климатической зоны в заданном районе

В.С. Балакин, e-mail: balakin69vs@mail.ru
Ю.В. Шипко, e-mail: yshipko@mail.ru
О.В. Колычев, e-mail: ok.work@mail.ru

ФГБУ «ЦНИИ ВВС» Минобороны России

Аннотация. *Рассматривается программный комплекс автоматизированного описания характеристик природно-климатической зоны в заданном районе планеты, необходимого при составлении авиационно-климатических описаний и справок районов аэродромов. Программные модули написаны на языке Python; имеется две версии программы – для работы в операционных системах Windows и AstraLinux. Основой программной разработки является проведенная оцифровка карт климатических зон земного шара. В качестве исходной информации используется сформированная база данных описаний климатических и природных зон по Б.П. Алисову и В.П. Кёппену.*

Ключевые слова: *автоматизированный расчет, авиационно-климатическое описание, программа, природно-климатическая зона.*

Введение

В число задач метеорологического обеспечения государственной авиации входит изучение и обобщение климатических характеристик районов аэродромов, поскольку планирование различных видов летной деятельности на большие сроки возможно только на основе использования данных о многолетнем режиме погоды. Согласно требованиям руководящих документов [1, 2] метеорологические подразделения разрабатывают авиационно-климатические описания (АКО) и справки районов базирования – справочные пособия, где отражаются физико-географические характеристики, общие черты климата с анализом особенностей атмосферной циркуляции, специализированные климатические показатели приземного слоя и свободной атмосферы.

Разработка АКО осуществляется по мере накопления материала наблюдений (не менее чем за 5 лет [1]). Однако функционирование государственной авиации часто предполагает (при определенной политической обстановке) оперативную передислокацию авиационного формирования в район, для которого отсутствует какая-либо

© Балакин В. С., Шипко Ю. В., Колычев О. В., 2026

климатическая информация. Кроме того, требуется соблюдение принципа автономности гидрометеорологического обеспечения войск (сил). При дефиците (или отсутствии) исходных метеоданных по неосвещенным в метеорологическом отношении территориям, ограничении доступа к информационным сетям общего пользования оперативная автоматизированная разработка АКО для любого района планеты, в том числе с помощью предлагаемого программного комплекса, является весьма актуальной.

Цель данной работы – повышение качества метеорологического обеспечения государственной авиации путем разработки программного комплекса автоматизированного составления АКО в части описания физико-географических характеристик в любом районе планеты.

1. Типовые классификации климата

Разработанный программный комплекс определения и описания природно-климатической зоны планеты по заданным координатам расположения аэродрома использует формулировки типа климата по известным классификациям В.П. Кёппена, Б.П. Алисова [3, 4]. С их учетом проведена оцифровка карт климатических и природных зон земного шара (с точностью до 1° по широте и долготе).

Классификация климатов Кёппена-Гейгера [5] – одна из наиболее распространенных систем классификации типов климата. Система основана на учете режима температуры и осадков. Климатические зоны делятся на пять основных климатических групп, каждая группа подразделяется на основе характера сезонных осадков и температуры воздуха (рис. 1).

Как показано на рис. 1, каждая группа и подгруппа имеют буквенное обозначение, где первая буква соответствует определенным климатическим условиям: A – зоны с высокими температурами круглый год и большим количеством осадков; B – зоны с незаметным количеством осадков или полным их отсутствием; C – зоны с небольшой разницей температур летом и зимой и отсутствием постоянного снежного покрова; D – зоны с четко выраженными границами лета и зимы; E – зоны с постоянным снежным покровом и средними температурами ниже 10 °C.

Классификация климатов Б.П. Алисова [6] основана на выделении климатических зон и областей исходя из условий общей циркуляции атмосферы. Классификация базируется на физических процессах в атмосфере, прежде всего, на распределении тех или иных зональных воздушных масс в различные сезоны. Границами климатов приняты наиболее частые положения основных климатических атмосферных фронтов. Количественных характеристик для положения этих фронтов не

существует, так что они проведены ориентировочно по опыту анализа синоптических процессов [3] (рис. 2).

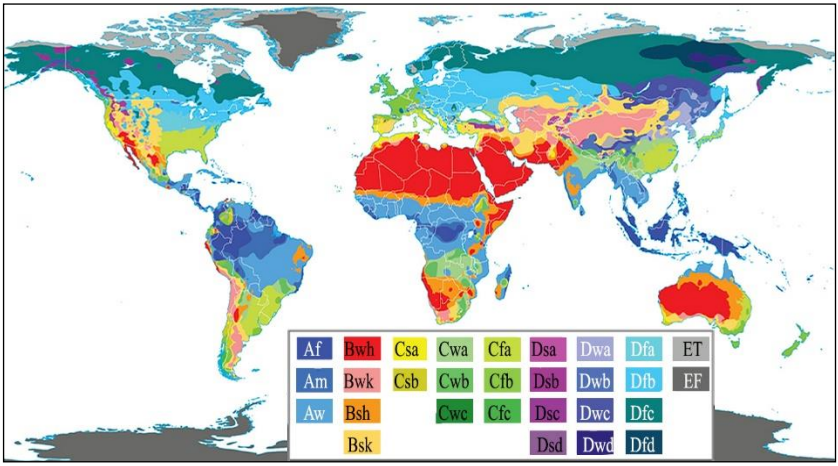


Рис. 1. Карта климатических зон по Кёппену-Гейгеру

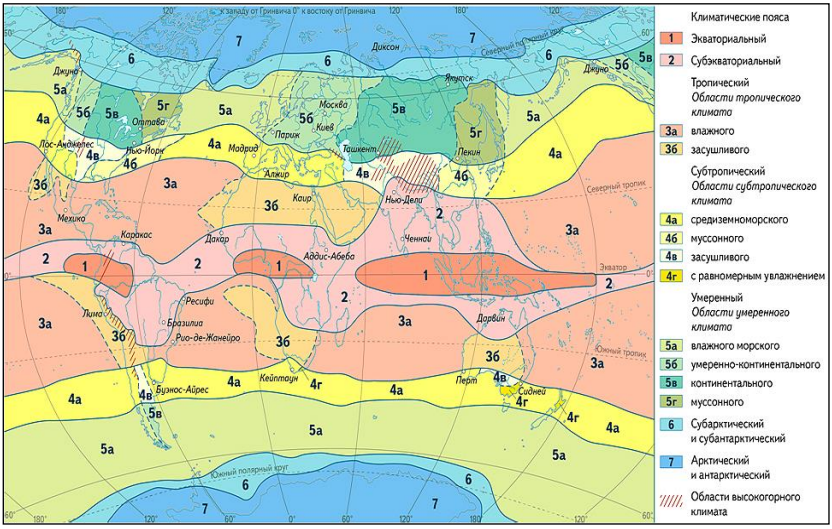


Рис. 2. Карта климатических поясов по Б.П. Алисову

В каждом полушарии рассматриваются (по Б.П. Алисову) четыре климатических пояса по признаку преобладания в течение года одной из основных воздушных масс: экваториальной, тропической, умеренной, арктической (антарктической). Между основными поясами выделяются три переходных пояса, где из-за сезонной миграции летом преобладают воздушные массы более низких, а зимой – более высоких широт.

Программный комплекс дополнен модулем автоматизированного описания природной зоны района аэродрома (рис. 3).

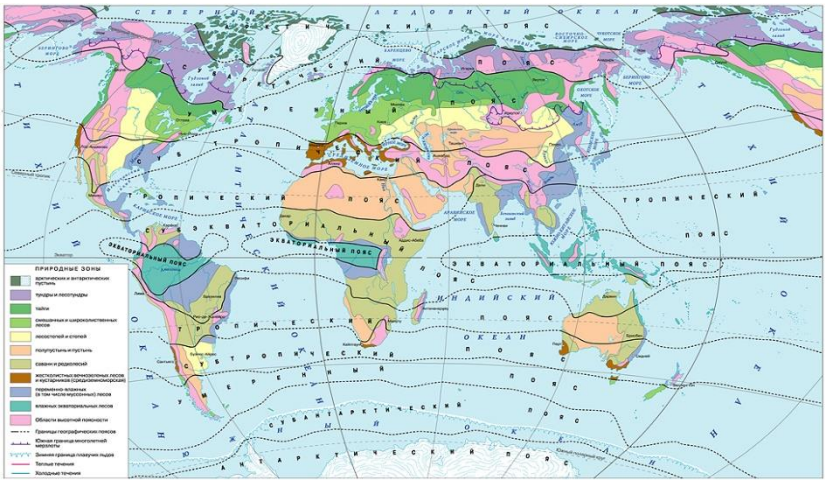


Рис. 3. Карта распределения природных зон планеты

Представляемые программой описания климатических и природных зон в данном случае даются кратко. При формировании раздела АКО «Физико-географическая характеристика» в метеорологическом подразделении необходимо их расширить, дать развернутое содержание, указав особенности рельефа, наличие и удаленность водных объектов, лесистость местности, обратив внимание на детали местности, которые оказывают заметное влияние на климатические условия района.

2. Описание программной реализации

При построении рассматриваемого программного комплекса использован язык программирования Python. Разработаны версии программы для операционных систем Windows и AstraLinux, которая сертифицирована для использования в изделиях Минобороны Российской Федерации.

Программное обеспечение включает составные части:

– прикладная программа автоматизированного определения и представления текстового материала с описанием природно-климатической зоны в заданном районе планеты;

– база данных описаний климатических (по Б.П. Алисову, В.П. Кеппену) и природных зон планеты.

Программа имеет «оконный» интерфейс (рис. 4).

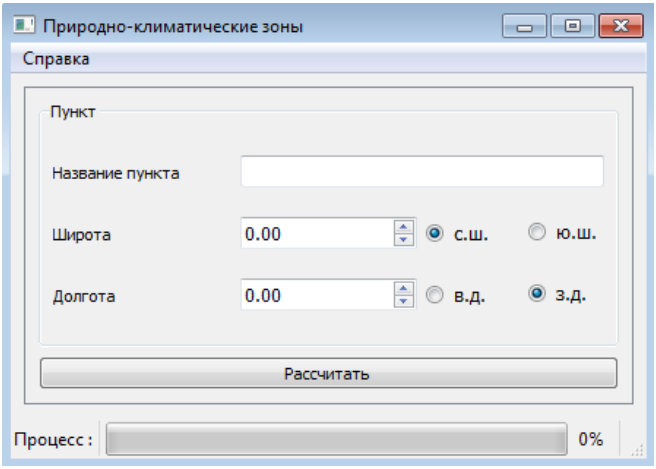


Рис. 4. Главное окно программы

Входные параметры: название пункта и координаты по широте, долготе. Выходная информация – текстовый файл, сохраняемый на диске компьютера, с описанием природно-климатической зоны. После нажатия кнопки «Рассчитать» открывается окно Проводника для выбора папки сохранения файла с материалом результатов (рис. 5).

Как показано на рис. 4, процесс расчета и формирования выходного файла отражается в нижней строке главного окна программы.

Ниже приводятся примеры программно-реализованных описаний для районов, имеющих различающиеся климатические условия.

Например, для аэродрома Бамако (Республика Мали, координаты станции: 12,5° северной широты, 8° западной долготы) формируется в автоматизированном режиме следующий текст, содержание которого полностью соответствует климатическим условиям географического положения данного пункта в Африке.

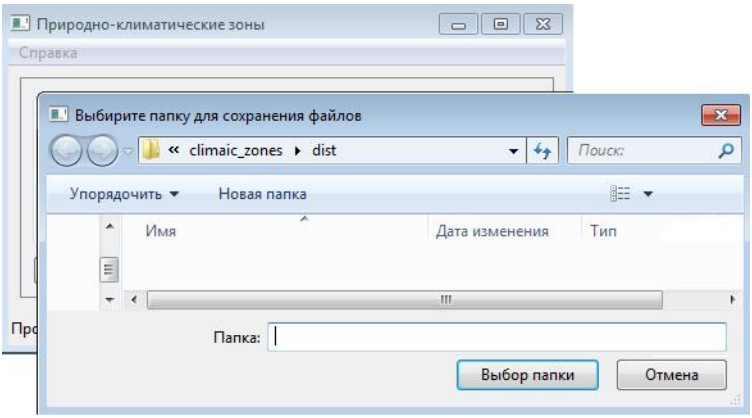


Рис. 5. Окно Проводника для выбора папки сохранения файла результатов

«На рассматриваемой территории по классификации Б.П. Алисова имеет место субэкваториальный климат. Летом преобладает экваториальный воздух, зимой – тропический. Климат муссонный. Выражены два сезона года: летний сезон дождей, когда муссоны приносят теплый и влажный воздух с экватора, и сухой зимний сезон, в который дуют сухие тропические ветра с континентов. Вследствие этого летом влажность воздуха составляет 80 %, зимой она падает в два раза, осадков почти нет. В зависимости от сезона (и близости к океану) средняя температура колеблется от +22 до +32 °С. Через данные области перемещается внутритропическая зона конвергенции (вместе экваториальной депрессией) два раза в год – с юга на север и с севера на юг. Поэтому в этих областях зимой господствует восточный (пассатный) перенос, меняющийся летом на западный перенос. В связи с сухой зимой для климата тропических муссонов характерен ландшафт саванн (тропической лесостепи).

По классификации В.П. Кеппена – Aw – климат тропической саванны с засушливой зимой. Характерен выраженный сухой сезон, при этом в самый засушливый месяц выпадает менее 60 мм осадков и меньше [100–(среднее годовое количество осадков/25)]; район характеризуется постоянными высокими температурами (на уровне моря и на малых высотах). Ярко выраженная сезонность – сухая зима и дождливое лето. Среднемесячные температуры воздуха превышают 17 °С в течение всего года. В условиях данного типа климата формируется ландшафт саванны.

Природная зона рассматриваемого района – саванн и редколесий. Климат сезонно-влажный, с четкой сменой периодов дождей и засух, что связано с пассатно-муссонной циркуляцией воздуха. Ближе к экватору сезон дождей продолжается до 9 месяцев, по мере удаления от него – сокращается до 3 месяцев. Характерны незначительные сезонные колебания температур. Летом начинается сезон дождей, разрастается травянистый покров, возвращаются из мест миграции животные. Зимой в саванне очень сухо, температура воздуха составляет примерно 21 °С. Характеристика почвы саванн и редколесий связана с режимом дождей. В близости от экватора находятся красные ферраллитные почвы, по мере удаления от него – красно-бурые почвы».

Для района пункта Тикси (Российская Федерация, координаты станции: 71.6° северной широты, 129° восточной долготы) формируется следующее описание.

«На рассматриваемой территории по классификации Б.П. Алисова имеет место арктический климат. Характеризуется круглогодичными отрицательными температурами воздуха и скудными осадками (100–200 мм в год). Долгая и холодная зима с сильными ветрами. Зимой в условиях полярной ночи средние температуры воздуха составляют около –30 °С. Короткое лето с пасмурной и дождливой погодой, температуры воздуха в редкие дни превышают +5 °С. Основная часть тепла затрачивается на таяние льда и снегового покрова. Снеговой покров и лед держатся большую часть года, даже летом они не успевают растаять. Снег и лед хорошо отражают солнечные лучи, поэтому земля не прогревается. Даже летом, когда солнце не уходит с горизонта, температура не поднимается выше 0 °С. Полярные день и ночь обусловливают крайне неравномерное поступление солнечного тепла в течение года. Радиационный баланс в южных районах Арктики положительный (фактически в 2–3 раза меньше, чем в умеренных широтах, в Арктическом бассейне, как правило, отрицательный). Природа Заполярья представляет собой безжизненные ледяные пустоши.

По классификации В.П. Кеппена – ET – полярный климат тундры. Средняя температура как минимум одного месяца выше 0 °С и ниже 10 °С. Пространство с вечномерзлой почвой, не заливаемое морскими или речными водами, заболоченная местность (по характеру поверхности тундры бывают болотистые, торфянистые, каменные). Главная особенность – суровый субарктический климат, высокая относительная влажность, сильные ветра. Крупная фауна встречается редко. Зима чрезвычайно продолжительная (8–9 месяцев). Зимой в тундре имеет место полярная ночь.

Природная зона рассматриваемого района – арктических пустынь, характеризуется арктическим климатом. Большая часть территории покрыта льдом и снегом. Пространства покрыты ледниками, щебнем, обломками камней и гор. Почвы арктические пустынные, каменистые. Зимой здесь долгая полярная ночь (на 75° с.ш. – 98 суток; на 80° с.ш. – 127 суток; в районе полюса – полгода). Зимой температура опускается до –40 °С и ниже, дуют ураганные ветры, часты бураны. Лето короткое, холодное, с круглосуточным освещением, температура воздуха чуть выше 0 °С. Небо затянуто серыми облаками, идет дождь (нередко со снегом), из-за сильного испарения воды с поверхности океана образуются густые туманы. Годовая сумма осадков до 400 мм. Ледяной и снежный покровы держатся почти весь год. Растительный мир: мхи, лишайники на камнях».

Заключение

Программная реализация расчета и отображения описания природно-климатической зоны района заданного пункта (аэродрома) относится к подсистеме обработки и анализа специализированной метеорологической информации, отвечает требованиям современного развития системы гидрометеорологического обеспечения авиации: повышению качества и эффективности обеспечения безопасности полетов, обеспечения органов управления всеми видами метеорологической информации, необходимой при подготовке и ведении операций, применении вооружения, военной и специальной техники, соблюдению принципа автономности метеорологического обеспечения войск (сил).

Результаты программного расчета специализированных климатических описаний по различным пунктам подтвердили адекватность отражения реальных климатических условий и авиационно-климатических характеристик рассматриваемых районов.

Эффективность результатов выполненной работы определяется возможностью их использования в метеорологических подразделениях государственной авиации.

Список литературы

1. Руководство по практическим работам метеорологических подразделений авиации Вооруженных Сил. – М. : Воениздат, 1992. – 488 с.

2. РД 52.21.692–2007. Требования к составлению климатического описания аэродрома [Текст]. – М. : Росгидромет, 2007. – 37 с.

3. Климатология / О. А. Дроздов, В. А. Васильев, Н. В. Кобышева [и др.]. – Л. : Гидрометеоиздат, 1989. – 568 с.

4. Хромов, С. П. Метеорология и климатология : учебник / С. П. Хромов, М. А. Петросянц. – М. : Наука, 2006. – 582 с.
5. Классификация климатов Кёппена [Электронный ресурс]. – Режим доступа : https://ru.wikipedia.org/wiki/Классификация_климатов_Кёппена
6. Классификация климатов Алисова [Электронный ресурс]. – Режим доступа : https://ru.wikipedia.org/wiki/Классификация_климатов_Алисова

Адаптивная сегментация траектории и навигация на основе векторного поля

А. А. Барсукова, e-mail: chemprof_461@rambler.ru[1]
А.В. Козырь e-mail: kozyr_a_v@mail.ru

ФГБОУ ВО «Тульский государственный университет»

Аннотация. *В статье рассматривается задача автоматического возврата мобильного робота в исходную точку при отсутствии априорной информации об окружающей среде. Предлагается алгоритм, основанный на анализе сегментированной траектории, полученной посредством одометрических данных. В отличие от существующих подходов, фокусирующихся на общих методах кластеризации траекторий, данная работа акцентирует внимание на адаптации алгоритмов GRASP-UTS и K-Means для задачи возврата, с целью определения обратного суммарного вектора. Для повышения точности навигации предложен алгоритм построения упрощенного векторного поля для обхода препятствий. Алгоритм включает в себя предобработку одометрических данных (фильтрация и сглаживание), сегментацию траектории, расчет векторов направления и кривизны для каждого сегмента, построение обратного вектора и проведено численное моделирование в среде MATLAB. Проведен сравнительный анализ алгоритмов GRASP-UTS и K-Means, выявлены их преимущества и недостатки для различных типов траекторий и задач.*

Ключевые слова: *Мобильный робот, одометрические данные, сегментация траектории, GRASP-UTS, K-Means, векторное поле, навигация, касательный вектор, вектор кривизны.*

Введение

Определение местоположения и навигация робототехнических комплексов (РТК) – важная область исследований. Широко применяются алгоритмы случайных деревьев, дорожных карт, потенциальных полей и др. Эти алгоритмы могут формировать негладкие траектории. [4] Необходимо учитывать скоростной режим РТК, особенности группового взаимодействия. Альтернативно, возможно использование спутниковой и инерциальной навигации, а также методов, основанных на одометрических данных [2].

© Козырь А.А., Барсукова А.А., 2026

В данной статье предлагается алгоритм возврата робота в исходную точку с анализом различных методов сегментации траектории, таких как GRAPS-UTS[5][9] и K-MEANS [6]. Целью является определение обратного суммарного вектора и построение упрощенного векторного поля для движения РТК в направлении исходной точки. Результаты алгоритмов сегментации визуализированы в MATLAB.

1. Постановка задачи

Разработка и анализ алгоритма сегментации траектории и построения упрощенного векторного поля, на его основе.

Задачи:

1. Сбор и обработка одометрических данных (фильтрация и сглаживание);
2. Сегментация траектории (GRAPS-UTS, K-Means), определение векторов направления и расчет обратного суммарного вектора;
3. Построение векторного поля для навигации к исходной точке;

2. Методы получения и обработки одометрических данных

Для сбора данных о движении мобильного робота применялся фреймворк ROS2 (дистрибутив Jazzy Jalisco). Разработанные программные модули обеспечивали получение информации от датчиков робота и ее запись в файл «odometry_dataN» (N – номер файла) для последующего анализа. Программные модули были разработаны с использованием библиотек NAV2 и TF2, что позволило реализовать функции навигации и преобразования систем координат.

Параллельно с одометрическими данными (линейная и угловая скорость) регистрировались данные системы Motion Capture. На основе полученных данных были построены траектории в среде MATLAB (рис. 1). Экспериментальные сценарии включали движение по кругу (несколько оборото) и различные повороты.

Анализ траекторий, построенных на основе одометрических данных (Robot Trajectory) и данных Motion Capture (2D Trajectory [7], свидетельствует об их идентичной конфигурации и сопоставимой длине пройденного пути (рис. 1). Оценка точности траектории основана на визуальном сходстве и сравнении пройденного расстояния (единицы измерения – миллиметры). В результате эксперимента было выявлено расхождение траекторий полученных с использованием стека NAV2 в ROS2 и на основе данных Motion Capture, для нивелирования этой проблемы графики приведены в единую систему координат.

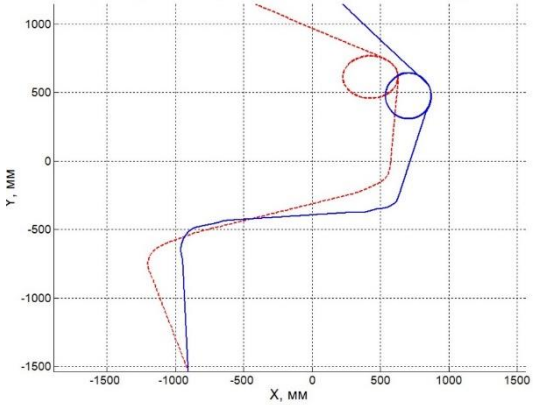


Рис. 1. Траектории полученные с использованием приложения ROS2(пунктир) и Motion Capture (сплошная). Единая стартовая точка

С целью повышения надежности и точности одометрических данных, полученных с мобильного робота, применялась двухэтапная процедура фильтрации. Данные представлены в виде последовательности координат x, y и временных меток (time).

На первом этапе удалялись записи с одновременным нулевым значением координат x и y. Такие записи рассматривались как аномалии, возникшие вследствие ошибок датчиков или сбоев в системе сбора данных. Валидность записи определялась булевым вектором $valid$, где

$$valid = \sim (x == \& y == 0)$$

Далее производилась фильтрация массивов координат и времени на основе данного вектора.

Второй этап включал выявление и удаление аномальных изменений координат, потенциально вызванных кратковременными сбоями датчиков. Использовалась функция дискретных разностей $diff$:

$$dx = diff(x); dy = diff(y)$$

Вычислялись векторы dx и dy, отражающие разности между последовательными значениями координат. "Прыжки" идентифицировались на основе порогового значения (0.5), установленного эмпирически. Записи, соответствующие резким изменениям, исключались из данных.

Для уменьшения влияния случайных ошибок датчиков и повышения точности оценки траектории движения робота применялся фильтр Свитского-Голея (Savitzky-Golay) [7]. Данный фильтр сглаживания скользящим полиномом осуществлял сглаживание данных без значительной деформации сигнала.

3. Анализ алгоритмов сегментации траектории

В качестве методов сегментации траекторий рассмотрнеы K-Means, и GRASP-UTS. Методы сегментации и их классификация подробно изложены в [3].

В статье проведен сравнительный анализ двух методов K-Means и GRASP-UTS, [5, 9] использующих линейную и угловую скорость для построения траектории, и обоснован выбор одного из них.

Сегментация траектории алгоритмом GRASP-UTS. GRASP (Greedy Randomized Adaptive Search Procedure) – это итеративный метод выборки для комбинаторной оптимизации, где решения постепенно создаются с нуля. На каждой итерации выбирается элемент из базового набора, основываясь на жадной функции, представляющей собой оценку увеличения функции затрат при включении этого элемента. Алгоритм проводит несколько итераций, каждая из которых производит приближенное решение, и возвращает лучшее найденное решение [9].

GRASP-UTS является специализированным алгоритмом, адаптированным для сегментации траектории на основе данных одометрии колес [9]. Процесс направлен на разделение траектории на однородные сегменты, используя евклидово расстояние между ориентирами и точками траектории. Задача сегментации траектории заключается в разбиении выборки траектории на θ однородных сегментов, где S является множеством точек:

$$S = \{s_1, s_2, ..., s_k\}$$

представляет собой точки разбиения траектории.

Функция стоимости, используемая алгоритмом GRASP-UTS, имеет вид:

$$\mathrm{Cos}t(S) = \lambda_1 N_s + \lambda_2 \left| L_{total} - L_{t\arg et} \right| + \lambda_3 \sum_{i=1}^{N_{S-1}} \left| \theta_i + 1 - \theta_i \right| + \lambda_4 \sum_{i=0}^{N_s} \left| \theta_i - \theta_{t\arg et} \right|.$$

где N_S – количество сегментов, L_{total} - суммарная длина траектории, θ_i - угол направления i-го сегмента; θ_{target} - желаемое направление; $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ - весовые коэффициенты.

Функция стоимости, используемая в алгоритме GRASP-UTS, представляет собой целевую функцию, предназначенную для количественной оценки качества сегментации траектории. Она позволяет оценить, насколько сформированная сегментация соответствует заданным критериям оптимальности, заложенные в функцию стоимости. Минимизация функции стоимости соответствует наилучшей сегментации.

Алгоритм GRASP-UTS (с использованием представленной функции стоимости (5) позволяет эффективно сегментировать траекторию, адаптируясь к различным условиям и требованиям, задаваемым весовыми коэффициентами.

– Сегментация траектории методом K-Means

Алгоритм K-means (k-средних) состоит из итеративного процесса: выбор начальных центров кластеров, распределение точек траектории по кластерам на основе минимального расстояния (например, евклидова расстояния)

$$d_E(x,y) = \sqrt{\sum_i (x_i - y_i)^2}$$

и пересчет центров кластеров [6]. Процесс повторяется до стабилизации кластеров.

K-Means применяет параметры движения (линейная и угловая скорость) как признаки для кластеризации, выделяя различные режимы движения [6]. Проблемой алгоритма является вероятность схождения к неоптимальному решению из-за зависимости от случайной начальной инициализации, подверженности локальным минимумам, и чувствительности к выбросам

4. Определение вектора каждого сегмента и обратного вектора.

В отличие от подхода в [8], данная работа акцентируется на извлечении информации о форме траектории, исключая учет препятствий. После сегментации траектории требуется определить векторы направления движения и кривизны для каждого участка.

– Касательный вектор $\hat{V}^j$

Для каждого сегмента строится кубический сплайн, аппроксимирующий траекторию. Касательный вектор вычисляется как первая производная сплайна. Усредненный касательный вектор нормализуется для получения единичного вектора направления

$$V^j = \frac{1}{K}\sum_{k=0}^{K-1} V(t_j, k)$$

который затем нормализуется

$$\hat{V}^j = \frac{V^j}{||V^j||} = \left(\frac{V_{j,x}}{\sqrt{V_{j,x}^2 + V_{j,y}^2}} \frac{V_{j,y}}{\sqrt{V_{j,x}^2 + V_{j,y}^2}} \right)$$

– Вектор кривизны

Вектор кривизны вычисляется на основе второй производной траектории. Упрощенный признак кривизны (по знаку) определяется выражением:

$$k(t) = \frac{|\nu'(t) \times a(t)|}{\|\nu(t)^3\|}.$$

– Вектор длины дуги L_j

Длина сегмента L_j вычисляется численно и умножается на единичный вектор направления для получения вектора длины дуги L_j, описывающего направление и пройденное расстояние на сегменте.

$$L_j \approx \sum_{\kappa=1}^{K-1} \sqrt{(x_{k+1} - x_k)^2 + (y_{k+1} - y_k)^2},$$

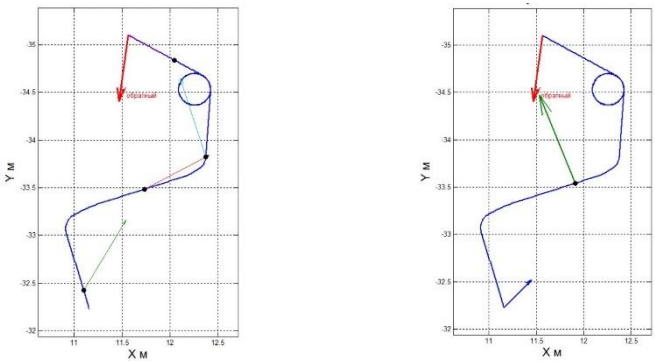
$$L_j = \hat{v} \times L_j.$$

Окончательным этапом в анализе траектории является сложение векторов для нахождения общего и последующий его разворот на 180° С в и сходную точку

$$V_{reverse} = -V_{sum}$$

Моделирование алгоритмов обработки траектории проводилось в системе MATLAB

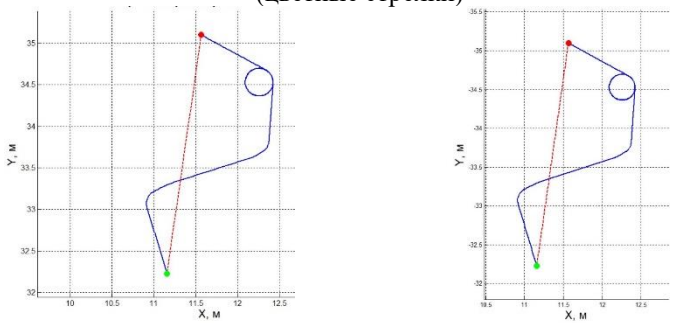
Несмотря на одинаковую величину обратного вектора, GRASP-UTS и K-Means могут быть предпочтительны в разных ситуациях. Для поставленной задачи алгоритм GRASP-UTS предпочтителен для аппроксимации сложных траекторий с множеством изгибов, обеспечивая высокую точность и плавность движения, что особенно важно в задачах демонстрации движений или обучения



а б

а – сегментация и определение направление векторов GRASP-UTS,
б – сегментация и определение направления векторов K-Means

Рис. 2. Визуализация алгоритмов обработки траектории алгоритмами GRASP-UTS и K-Means и рассчитанных векторов (цветные стрелки)



а Б

а – обратная траектория GRASP-UTS,
б – обратная траектория K-Means

Рис. 3. Визуализация прямой (синие линии) и обратной траектории (красный пунктир) на основе алгоритмов GRASP-UTS и K-Means

5. Построение векторного поля

Традиционные подходы, такие как метод потенциальных полей, характеризуются рядом недостатков, включая возможность застревания в локальных минимумах, [1] колебания траектории вблизи цели и сложность точной настройки параметров.

В целях обеспечения надежной и плавной навигации в условиях сложной окружающей среды предложен метод построения векторного поля

$G(p)$, представляющего собой суперпозицию трех векторных компонентов: вектора притяжения к ближайшей точке на заданной траектории $\left(k_p \cdot e_{path}\right)$, вектора отталкивания от конечной точки $\left(k_{c\cdot s}\left(p\right) \cdot e_{repel}\right)$, обеспечивающего обход препятствий вблизи цели, и касательного вектора к траектории $\left(v_d \cdot v_{tangent}\right)$, определяющего желаемое направление и скорость движения.

Формула векторного поля представлена следующим образом.

$$G(p) = k_p \cdot e_{path} + k_{c\cdot s}(p) \cdot e_{repel} + v_d \cdot v_{tangent}$$

Метод включает этапы интерполяции траектории с использованием кубического сплайна для повышения точности вычисления касательных векторов и позволяет оптимизировать параметры k_p, k_c, v_d посредством генетических алгоритмов или методов градиентного спуска, минимизируя отклонение от опорной траектории и обеспечивая дополнительную навигацию.

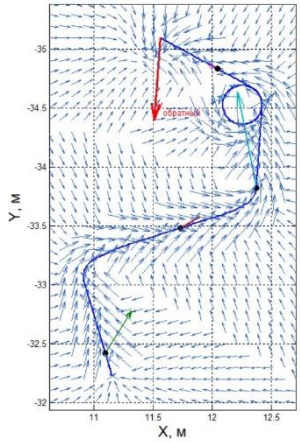


Рис. 4. Моделирование векторного поля в MATLAB

Заключение

В результате проведенной работы получено упрощенное векторное поле, которое выступает в роли ориентира при преодолении препятствий и является в полноценным методом навигации.

Предложенный алгоритм позволяет анализировать сложные траектории с изменением направления, разбивать их на кусочно-гладкие сегменты методом GRASP-UTS.

Предложенный метод может быть применен в широком спектре задач, включая поисково-спасательные операции, доставку грузов в динамических средах, автономную навигацию мобильных роботов в помещениях и другие приложения, требующие надежного и безопасного возврата в исходную точку.

Список литературы

1. Филимонов, А. Б. Конструктивные аспекты метода потенциальных полей в мобильной робототехнике / А. Б. Филимонов, Н. Б. Филимонов // Автометрия. – 2021. – Т. 57, № 4. – С. 45–53.
2. Golovan, A. A. Combined use of strapdown inertial navigation systems and odometers from the standpoint of mechanics of inertial navigation systems. Part 1 / A. A. Golovan, I. V. Nikitin // Moscow University Mechanics Bulletin. – 2015. – Vol. 70, № 2. – P. 46–49.
3. A trajectory segmentation algorithm based on interpolation-based change detection strategies : препринт / Etemad, M. [и др.] // Remote Sensing. Change Detection. – 2019. – Февр. – DOI : 10.13140/RG.2.2.34157.03049.
4. Костюков, В. А. Алгоритмы планирования сглаженных индивидуальных траекторий движения наземных роботов / В. А. Костюков, М. Ю. Медведев, В. Х. Пшихопов // Роботы, мехатроника и робототехнические системы. – 2023. – № 23. – С. 585–595. – DOI : 10.17587/mau.23.585-595.
5. Resende, M. G. C. GRASP: greedy randomized adaptive search procedures / M. G. C. Resende, C. C. Ribeiro // Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques / ed. by E. K. Burke, G. Kendall. – 2nd ed. – Springer, 2014. – P. 287–312. – DOI : 10.1007/978-1-4614-6940-7_11.
6. Шипилов, С. С. Математическая модель алгоритма k-Means: исследование кластеризации и практическая реализация : магистерская диссертация / С. С. Шипилов. – 2023. – 23 с.
7. Schafer, R. W. What is a Savitzky–Golay filter? [Lecture Notes] / R. W. Schafer // IEEE Signal Processing Magazine. – 2011. – Vol. 28, № 4. – P. 111–117. – DOI : 10.1109/MSP.2011.941097.
8. Bredenbeck, A. A tactile feedback approach to path recovery after high-speed impacts for collision-resilient drones / A. Bredenbeck, T. Yang, S. Hamaza, M. W. Mueller // Drones. – 2025. – Vol. 9, № 11. – P. 758. – DOI : 10.3390/drones9110758.
9. GRASP-UTS: an algorithm for unsupervised trajectory segmentation / A. Soares Júnior [и др.] // International Journal of Geographical Information Science. – 2015. – Vol. 29, № 1. – P. 46–68. – DOI : 10.1080/13658816.2014.938078.

Исследование влияния атмосферы на пропускную способность инфракрасного излучения

С. В. Беспалов, e-mail: antonn58@ya.ru

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *Проведенные исследования описывают особенности прохождения электромагнитного излучения инфракрасного диапазона через земную атмосферу. Показано распределение спектрального коэффициента пропускания атмосферы в зависимости от метеорологической дальности видимости, высоты, длины волны и трассы. Определены параметры атмосферы для выбора оптимального диапазона инфракрасной камеры.*

Ключевые слова: *спектральный коэффициент пропускания, показатель ослабления, рассеивание, поглощение*

Введение

Прозрачность атмосферы является важным фактором, влияющим на климат и различные сферы деятельности человека. Её необходимо учитывать при проведении исследований и разработке технологий, связанных с использованием солнечной энергии.

Электромагнитное излучение при прохождении через земную атмосферу ослабляется. Участки спектрального диапазона от 3 до 5 мкм и от 8 до 14 мкм принято называть «окнами прозрачности» атмосферы, поскольку в этих участках горизонтальные приземные трассы пропускают инфракрасное (ИК) излучение достаточно хорошо. Для исследования были выбраны длины волн 4 и 10 мкм. В области более длинных электромагнитных волн атмосфера обладает высокими поглощающими свойствами, и поэтому вплоть до миллиметрового диапазона методы дистанционного оптического зондирования наземных объектов не применяются. В области более коротких длин волн ($\lambda < 2$ мкм) существует множество полос пропускания оптического излучения ближнего ИК, видимого и УФ диапазонов, в пределах которых работают оптико-электронные приборы других (не тепловизионных) типов.

1. Описание предметной области

Для волн ИК диапазона 3–14 мкм наиболее существенными являются процессы поглощения и рассеяния излучения молекулами

© Беспалов С. В., 2026

газов, парами воды, аэрозолями и осадками. Главная причина ослабления ИК излучения при прохождении приземных и вертикальных трасс заключается в его поглощении молекулами газов.

Поглощение оптического излучения определяется коэффициентом поглощения, характеризующим поглощение излучения на единице пути распространения. Ослабление ИК излучения атмосферой происходит в общем случае как из-за поглощения, так и рассеяния энергии излучения.

Для исследования влияния пропускной способности атмосферы на распространение ИК излучения обратимся к закону Бугера – Ламберта – Бера – физический закон, определяющий ослабление параллельного монохроматического пучка света при распространении его в поглощающей среде. Как приведено в работе [1]

$$P(\lambda, x) = P_0(\lambda)\exp\left[-\alpha(\lambda, x)\right]$$

где $\alpha(\lambda, x)$ – спектральный показатель ослабления (коэффициент ослабления атмосферы); x – длина трассы распространения; λ – длина волны; $P_0(\lambda)$ – поток излучения на входе в среду распространения; $P(\lambda, x)$ – поток излучения на выходе среды распространения.

Спектральным коэффициентом пропускания слоя среды толщиной x называют отношение потока излучения, прошедшего в среде путь, к потоку излучения в начале трассы:

$$\tau(\lambda, x) = P(\lambda, x) / P_0$$

Тогда с учетом выражения (1) получим

$$\tau(\lambda, x) = \exp\left[-\alpha(\lambda, x)\right]$$

2. Существующие подходы

Коэффициент пропускания среды, определяющий долю прошедшего через нее в заданном направлении потока, ограничивается такими физическими явлениями, как рассеяние и поглощение.

В свою очередь коэффициент ослабления складывается из коэффициентов поглощения и рассеивания [1,2], а именно: коэффициента молекулярного поглощения излучения атмосферными газами; коэффициента поглощения излучения атмосферным аэрозолем; коэффициента молекулярного рассеяния; коэффициента аэрозольного рассеяния. Так, например, в коэффициенте молекулярного рассеивания учитывается водность по формуле:

$$VO = \left(x * f * fa\right) / \left(\gamma * 10^6\right)$$

$$VO = \left(x * f * fa\right) / \left(\gamma * 10^6\right)$$

где f – относительная влажность; fa – абсолютная влажность; γ – удельный вес воды.

Значения метеорологических характеристик были взяты из модели стандартной атмосферы (ГОСТ 4401–81) в зависимости от широты места.

3. Результаты численного эксперимента

На основании методике расчета характеристик пропускания атмосферы [2,3], определены спектральные коэффициенты пропускания атмосферы для горизонтальных трасс в зависимости от метеорологической оптической дальности (MOR, англ. Meteorological Optical Range), длины волны ИК и высоты. На рис. 1 показаны кривые распределения для трассы длиной 3 км на высоте 1 км.

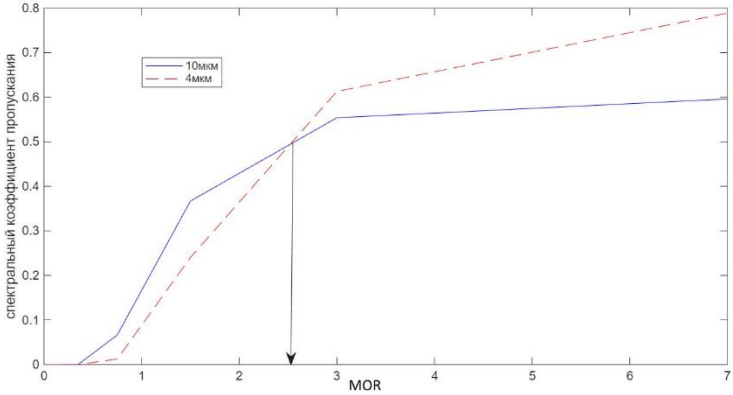


Рис. 1. Спектральный коэффициент пропускания атмосферы для горизонтальных трасс в зависимости от MOR, длины волны инфракрасного излучения и высоты

Из анализа кривых распределения и MOR для разных частот (рис.1) видно, что при плохой MOR (менее 2,6 км) больше для частоты 10 мкм, чем для частоты 4 мкм. А выше наоборот, т.е. если МДВ меньше 2,6 км целесообразно пользоваться ИК камерой с частотой 10 мкм, если MOR меньше 2,6 км, то камерой с частотой 4 мкм. Уменьшение высоты трассы до 500 м при неизменных остальных параметрах незначительно уменьшило спектральные коэффициенты пропускания атмосферы в среднем на 0,12 процента.

На следующем этапе исследования исходные величины следующие: высота 500 м, длина трассы 1 км. Температура воздуха также из модели стандартной атмосферы, но с поправкой на зимний период года в Центрально-Чернозёмном регионе (рис.2).

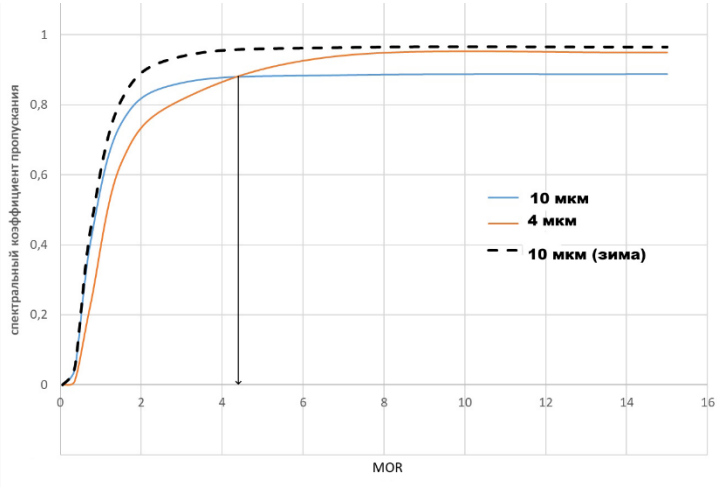


Рис. 2. Распределение спектрального коэффициента пропускания атмосферы в зависимости от MOR

Согласно [4], упругость насыщения водяного пара определяется прямопропорционально нелинейной зависимостью от температуры воздуха, описываемой уравнением Клаузиуса-Клапейрона. Таким образом, исходя из (4), уменьшение температуры приводит к уменьшению молекулярного рассеивания и, следовательно к увеличению спектрального коэффициента пропускания атмосферы.

Распределение коэффициентов пропускания атмосферы (рис. 2) при уменьшении высоты и длины трассы свидетельствует о том, что целесообразно переходить с диапазона 10 мкм на 4 мкм только при MOR больше 4,5 км. А также, в зимний период такой переход не целесообразен, т.к. ИК диапазон 10 мкм оптимален (пунктирная линия на рис. 2).

Заключение

В результате исследования выявлено, что на пропускную способность инфракрасного излучения в атмосфере в наибольшей степени влияет метеорологическая оптическая дальность, а также частота ИК диапазона. Более длинноволновое инфракрасное излучение в замутненных трассах имеет больший коэффициент пропускания.

Предложенные результаты могут быть использованы при принятии решения при выборе длины волны инфракрасной камеры.

Список литературы

1. Чёнгин, А. В. Способ пространственного сложения электромагнитных излучений когерентных источников / А. В. Чёнгин // Информатика: проблемы, методы, технологии : материалы XXV междунар. науч.-метод. конф. (Воронеж, 12–13 февр. 2025 г.) / под ред. Д. Н. Борисова. – Воронеж : Воронежский гос. ун-т, 2025. – С. 1177–1184.

2. Белоусов, Ю. И. Инфракрасная фотоника : учеб. пособие / Ю. И. Белоусов, Е. С. Постников. – 2-е изд., стер. – СПб. : Лань, 2023. – 340 с.

3. Филиппов, В. Л. Атмосфера и моделирование ОЭС в динамике внешних условий : монография / В. Л. Филиппов, В. П. Иванов, В. С. Яцык. – Казань, 2015. – 632 с.

4. Матвеев, Л. Т. Физика атмосферы : учеб. пособие / Л. Т. Матвеев. – СПб. : Гидрометеоиздат, 2000. – 778 с.

Программа контроля антибиотикорезистентности патогенной микрофлоры в медицинских организациях

С.В. Борзунов[1], С.Д. Кургалин[1], Н.Ю. Гончарова[2], Г.А. Батищева[2], А.А. Асеева[1], К.О. Петрищев[1]
e-mail: borzunov@cs.vsu.ru, kurgalin@bk.ru

[1]Воронежский государственный университет,
[2]Воронежский государственный медицинский университет им. Н.Н. Бурденко

Аннотация. *В работе создана компьютерная программа, обеспечивающая организацию контроля антибиотикорезистентной микрофлоры в медицинской организации. Анализ базы данных, формируемых программой, предоставляет сведения для оптимизации закупок лекарственных препаратов и подбора фармакотерапии в зависимости от состава и резистентности локальной микрофлоры. Применение программы направлено на повышение эффективности клинического применения противомикробных препаратов среды.*

Ключевые слова: *компьютерная программа, база данных, противомикробная терапия, антибиотикорезистентность.*

Введение

Фармакотерапия противомикробными препаратами обеспечивает эффективное лечение инфекционных заболеваний. Создание антибиотиков стало существенным фактором повышения продолжительности жизни, сохранения качественных лет трудоспособного возраста, снижения материнской и детской смертности. Однако при регулярном применении противомикробные препараты со временем теряют свою эффективность в связи высокой пластичностью и приспособляемостью микрофлоры, которая обладает различными механизмами резистентности под действием факторов внешней среды. Регулярный прессинг приема противомикробных средств неизбежно ведет к формированию штаммов, устойчивых к воздействию лекарственных препаратов, а реальная практика лечения внебольничных и внутрибольничных инфекций указывает на быстрое распространение полирезистентных возбудителей

Инфекционные заболевания могут быть вызваны возбудителями, устойчивыми к назначенному препарату, что создает угрозу для жизни пациентов. Сложившаяся практика противомикробной терапии

© Борзунов С.В., Кургалин С.Д., Гончарова Н.Ю., Батищева Г.А., Асеева А.А., Петрищев К.О., 2026

показывает, что вопросы лекарственного обеспечения требуют мониторирования патогенной микрофлоры, выделенной от пациентов, обратившихся за медицинской помощью.

Контроль антибиотикорезистентности включает необходимость выполнения микробиологических исследований. Важно учитывать биоматериал, структурные подразделения стационара (отделения терапевтического и хирургического профиля). В медицинской организации требуется выявлять и регистрировать возбудителей, обладающих фенотипом моно-, поли- или панрезистентности [1]. Медицинские работники обязаны правильно проводить забор биоматериала, уметь своевременно его доставлять в лабораторию, проводить анализ полученных данных, учитывать информацию о выделенном возбудителе и его чувствительности к противомикробным препаратам [2].

Актуальное значение имеет определение штаммов Грам+ и Грам– флоры с оценкой эффективности лекарственных препаратов, который используются в медицинской организации. Полученные данные ложатся в основу рационального выбора противомикробных препаратов при госпитализации пациентов [3,4]. Систематизация сведений о резистентности возбудителей дает информацию, необходимую для клинической практики и решения вопроса о своевременной ротации противомикробных препаратов, чтобы исключить неэффективные лекарственные средства [5,6].

Целью настоящей работы стало создание компьютерной программы для регистрации и обработки данных об антибиотикорезистентности микрофлоры к противомикробным препаратам в условиях многопрофильного стационара. Соавторы статьи имеют большой опыт создания компьютерных программ в этой области [7-11].

1. Основные характеристики разработанной программы

Функциональными возможностями программы являются: внесение, хранение параметров, генерация отчетов. Ключевые параметры составляют данные о пациенте, биоматериале, выделенном возбудителе, чувствительности микрофлоры к лекарственным препаратам.

Название программы «Журнал бактериального мониторинга» связано с рутинной практикой регулярной регистрации результатов работы микробиологических лабораторий, чтобы ограничить хранение информации на бумажных носителях.

Программа предназначена для регистрации патогенной микрофлоры, выделенной из биоматериала пациентов хирургических и терапевтических отделений, позволяет создавать базу данных Грам+ и Грам– микрофлоры, проводить анализ резистентности патогенных

возбудителей по отделениям, по виду возбудителя, в зависимости от состава биоматериала и за конкретный временной период.

Графический пользовательский интерфейс программы реализован на основе стандартной библиотеки Tkinter, входящей в базовую поставку Python. Использование этой библиотеки позволяет избежать зависимости от сторонних графических фреймворков и обеспечивает стабильную работу приложения на различных операционных системах. Для унификации внешнего вида и повышения эргономичности интерфейса применяется модуль ttk, предоставляющий набор современных виджетов, адаптирующихся к нативному стилю операционной системы. Для ввода дат используется дополнительный модуль tkcalendar, который реализует календарный элемент управления и обеспечивает корректный ввод временных параметров, что особенно важно при работе с медицинской документацией, где точность дат и временных интервалов имеет принципиальное значение.

Хранение данных организовано с использованием файлов формата Excel (XLSX). Данная реализация дает возможность быстрой интеграции с уже существующими процессами и нативной поддержкой базовых офисных приложений. Отказ от сложных серверных баз данных упрощает установку системы. Для её работы не нужно мощное оборудование или специальные знания у сотрудников. Для работы с таблицами, их фильтрации и анализа используется библиотека pandas. Она отлично справляется с большими объемами данных: помогает их сортировать, группировать и отбирать по времени. Для создания и оформления отчётов в Excel применяется библиотека openpyxl. Она позволяет легко формировать понятные и структурированные файлы: создавать листы, управлять внешним видом ячеек и наглядно представлять результаты анализа.

В процессе разработки также используются стандартные модули Python, обеспечивающие работу с файловой системой, обработку исключений и управление событиями пользовательского интерфейса. Совместное использование данных библиотек формирует устойчивую и расширяемую программную платформу, пригодную для длительной эксплуатации в условиях клинико-диагностических лабораторий.

Архитектура программы построена на принципах объектно-ориентированного программирования. Основные функциональные элементы, такие как модуль ввода данных, модуль обработки и хранения информации, а также модуль формирования аналитических отчетов, реализованы в виде отдельных классов. Такой подход повышает читаемость и структурированность кода, упрощает процесс отладки и

тестирования, а также обеспечивает возможность масштабирования и расширения функциональности без изменения его базовой архитектуры.

Программа является кроссплатформенной и может функционировать на операционных системах, на которых поддерживается интерпретатор Python версии 3.x. Отсутствие зависимости от специализированного оборудования и серверных решений позволяет использовать программу как на рабочих станциях лабораторного персонала, так и на персональных компьютерах с ограниченными вычислительными ресурсами. Это делает ее универсальным инструментом для бактериологического мониторинга в медицинских и научно-исследовательских учреждениях различного уровня.

В ходе разработки программы использовался набор стандартных и специализированных библиотек языка Python, обеспечивающих реализацию графического интерфейса, обработку данных и работу с файлами формата Excel. В таблице приведен перечень основных библиотек с указанием их назначения и минимальных рекомендуемых версий.

2. Ключевые функциональные модули программы

Главный модуль пользовательского интерфейса

Центральным компонентом программы является класс MicrobioApp, реализующий главное окно приложения и обеспечивающий полный цикл работы с данными бактериологических исследований –от ввода информации до ее сохранения и передачи в модуль отчетности. Инициализация главного окна осуществляется стандартным для Tkinter способом (см. листинг 1).

Листинг

Инициализация главного окна приложения

```
root = tk.Tk()
app = MicrobioApp(root)
root.mainloop()
```

Данный фрагмент кода создает корневое окно приложения, инициализирует основной класс и запускает главный цикл обработки событий, в рамках которого осуществляется взаимодействие пользователя с интерфейсом.

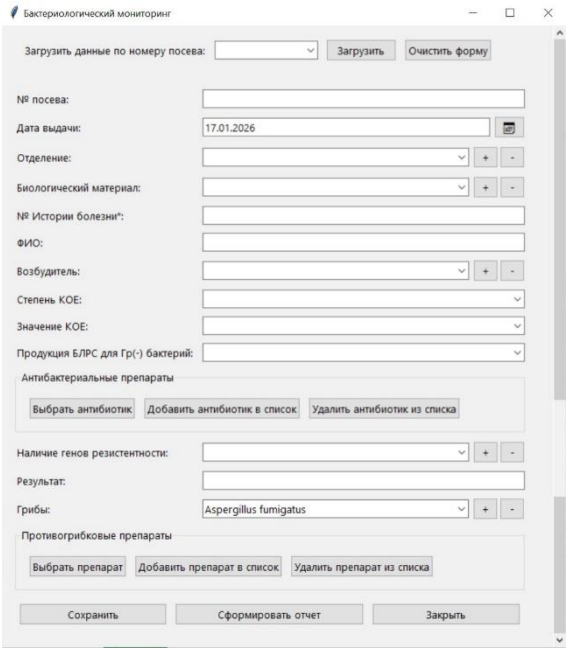


Рис. 1. Скриншот главного окна программы

Главное окно программы представляет собой прокручиваемую форму, в которой элементы управления сгруппированы по функциональному назначению (см. рис. 1). В верхней части расположены поля для ввода регистрационного номера посева и даты выдачи анализа. Далее следуют выпадающие списки для выбора отделения, биологического материала и идентифицированного микроорганизма. Такой подход обеспечивает стандартизацию вводимых данных и снижает вероятность ошибок, связанных с ручным вводом текста.

Модуль динамического ввода антибиотиков и чувствительности

Важной функцией программы является модуль динамического ввода антибиотиков, который дает возможность, пользователю указывать любое количество препаратов для одного исследования. Каждая строка содержит название антибиотика и соответствующий результат теста чувствительности, представленный категориями S (чувствительный), I (промежуточный) и R (резистентный).

Добавление новой строки антибиотика реализовано программно (см. листинг 2).

Листинг

Добавление строки антибиотика и выбора чувствительности

```
def add_antibiotic_row(self):
    antibiotic = ttk.Combobox(frame,
values=self.data_lists["antibiotics"])
    sensitivity = ttk.Combobox(frame, values=["S", "I", "R"])
```

Использование динамически создаваемых элементов интерфейса позволяет адаптировать программу к большому числу протоколов и тестируемых препаратов.

Модуль валидации и сохранения данных

Перед сохранением информации реализован этап валидации данных, направленный на контроль корректности и полноты вводимых сведений. Проверяется наличие обязательных полей, а также факт ввода как минимум одного антибиотика. Такой механизм предотвращает накопление неполных или некорректных записей в базе данных.

Для хранения информации о результатах антибиотикочувствительности используется строковое представление с разделителями, что позволяет сохранять переменное количество антибиотиков в рамках одной записи (см. листинг 3):

Листинг

Формирование строки данных для сохранения

```
row_data[9] = "; ".join(antibiotics)
row_data[10] = "; ".join(sensitivities)
```

Далее данные добавляется в Excel-файл, который выступает в роли места для хранения информации.

Модуль формирования аналитических отчетов

Функции анализа и визуализации данных вынесены в отдельный модуль, реализованный с помощью класса ReportWindow. При открытии окна отчетов происходит загрузка данных из Excel-файла и преобразование их в структуру pandas DataFrame, как представлено в листинге 4.

Листинг

Загрузка и подготовка данных для анализа

```
self.df = pd.read_excel(self.excel_file)
self.df['Дата выдачи анализа'] = pd.to_datetime(
self.df['Дата выдачи анализа'], format='%d.%m.%Y'
)
```

Использование DataFrame позволяет гибко работать с фильтрацией, сортировкой и агрегацией данных по заданным критериям, включая временные интервалы и категориальные признаки (рис. 2).

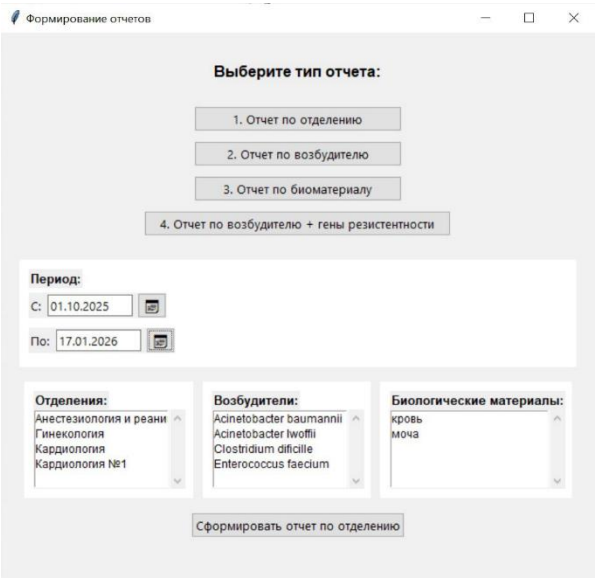


Рис. 2. Скриншот окна формирования отчетов

Пользователь может выбрать тип отчета, задать временной диапазон и дополнительные фильтры, после чего результаты анализа отображаются в виде интерактивной таблицы.

Модуль анализа антибиотикочувствительности

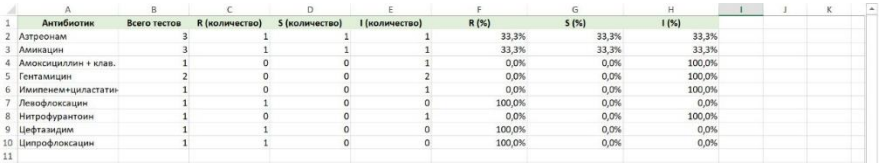
Для каждого антибиотика автоматически вычисляется общее количество проведенных тестов, а также число результатов категорий S, I и R (листинг 5). На основе этих данных рассчитываются процентные показатели чувствительности и резистентности, позволяющие оценить эпидемиологическую ситуацию в подразделении или медицинском учреждении в целом.

Листинг

Расчет процентного показателя резистентности

```
r_percent = (r_count / total_tested * 100) if total_tested > 0
else 0
```

Полученные статистические данные отображаются в отдельной таблице и используются при формировании итоговых отчетов (рис. 3).



	A	B	C	D	E	F	G	H	I	J	K
1	Антибиотик	Всего тестов	R (количество)	S (количество)	I (количество)	R (%)	S (%)	I (%)			
2	Азтреонам	3	1	1	1	33,3%	33,3%	33,3%			
3	Амикацин	3	1	1	1	33,3%	33,3%	33,3%			
4	Амоксициллин + клав.	1	0	0	1	0,0%	0,0%	100,0%			
5	Гентамицин	2	0	0	2	0,0%	0,0%	100,0%			
6	Имипенем+циластати	1	0	0	1	0,0%	0,0%	100,0%			
7	Левофлоксацин	1	1	0	0	100,0%	0,0%	0,0%			
8	Нитрофурантоин	1	0	0	1	0,0%	0,0%	100,0%			
9	Цефтазидим	1	1	0	0	100,0%	0,0%	0,0%			
10	Ципрофлоксацин	1	1	0	0	100,0%	0,0%	0,0%			
11											

Рисунок. Скриншот таблицы статистики антибиотикочувствительности

Алгоритм работы программы

Работа программы начинается с запуска приложения и инициализации пользовательского интерфейса. После загрузки существующих данных пользователь вводит информацию о бактериологическом исследовании и результаты антибиотикочувствительности. Введенные данные проходят проверку корректности и сохраняются в файл. При формировании отчетов выполняется фильтрация данных по заданным критериям, анализ антибиотикочувствительности и расчет статистических показателей. Итоговые результаты визуализируются в интерфейсе и могут быть экспортированы в отдельный файл.

Использование программы и описание интерфейса

Интерфейс программы ориентирован на практическое использование в условиях клинико-диагностической лаборатории. Главное окно обеспечивает удобный и логичный ввод данных, а наличие прокрутки позволяет работать с большим количеством элементов без перегрузки экрана.

Окно отчетов предназначено для аналитической работы и предоставляет пользователю инструменты для комплексного анализа данных, включая просмотр сводной статистики и экспорт результатов в формат xlsx, пригодный для дальнейшей обработки и представления в отчетной документации.

В открывшуюся на рабочем столе форму (см. рис. 1) вносятся сведения о выделенной микрофлоре (выбирается название возбудителя).

Перед пользователем открывается возможный список противомикробных препаратов. При необходимости изменить условия поиска – «Очистить данные».

Для обновления базы данных используется файл, в который в ручном режиме вносятся данные о патогенных возбудителях и

противомикробных препаратах, представленных на фармацевтическом рынке.

В формирующейся базе данных сохраняются сведения о всех тестируемых антибиотиках при определении резистентности микрофлоры. В программе указана дата последнего обновления базы данных.

В итоге, внедрение программы позволяет решать следующие задачи: сформировать специальную базу данных, позволяющую выполнять запросы о микрофлоре, сведения о противомикробных препаратах и их эффективности. Программа дает возможность для сбора сведений, поступающих из разных отделений для последующего анализа полученной информации и решения вопроса о рациональном использовании противомикробных препаратов. В дальнейшем планируется использовать системы искусственного интеллекта для проведения сравнительного анализа программных пакетов [12].

Заключение

Созданная компьютерная программа обеспечивает организацию контроля антибиотикорезистентной микрофлоры в медицинской организации. Анализ базы данных, формируемых программой, предоставляет сведения для оптимизации закупок лекарственных препаратов и подбора фармакотерапии в зависимости от состава и резистентности локальной микрофлоры. Применение программы направлено на повышение эффективности клинического применения противомикробных препаратов.

Список литературы

1. Опыт создания замкнутого цифрового контура для обеспечения непрерывного мониторинга антимикробной резистентности на основе валидированных результатов микробиологической диагностики / Н. А. Гультяева [и др.] // Клиническая микробиология и антимикробная терапия. – 2025. – Т. 27, № 3. – С. 369-389.

2. Петрова В. И. Клиническая фармакология: национальное руководство / В. И. Петрова, Д. А. Сычева, А. Л. Хохлова. – 2-е изд., переработ. и дополн. – Москва: ГЭОТАР-Медиа, 2024. – 816 с.

3. Проблема антибиотикорезистентности микроорганизмов у пациентов с инфекциями мочевыводящих путей / С. В. Котов [и др.] // Урология. – 2021. – №1 – С. 5-12.

4. Российские рекомендации. Определение чувствительности микроорганизмов к антимикробным препаратам. Версия 2024-02 / МАКМАХ, СГМУ. – Смоленск, 2024. – 192 с

5. Смольянинова, Д. С. Прогноз риска антибиотикорезистентности уропатогенов у пациентов с мочекаменной болезнью при внебольничной инфекции мочевыводящих путей / Д. С. Смольянинова // Научно-медицинский вестник Центрального Черноземья. – 2023. – № 24 (3). – С. 19–29.

6. Яковлев, С. В. Биапенем: клинико-микробиологическая характеристика и обсуждение места нового карбапенема в лечении тяжёлых инфекций в стационаре. Точка зрения клинических фармакологов / С. В. Яковлев, М. П. Суворова // Антибиотики и химиотерапия. – 2022. – Т. 67, № 5–6. – С. 81–91.

7. Программа для формирования базы данных, используемой в работе врача – клинического фармаколога / М. С. Гончаров // Интеллектуально-информационные технологии и интеллектуальный бизнес (ИНФОС–2021) : материалы двенадцатой международ. науч.-техн. конф. – Вологда, 2021. – С. 19-23.

8. Разработка программы для корректировки доз лекарственных препаратов / Т. А. Улезько [и др.] // Труды молодых ученых факультета компьютерных наук ВГУ. – Воронеж, 2023. – Вып. 3. – С. 246-253.

9. Автоматизированная оценка индивидуальной стоимости фармакотерапии как инструмент контроля лекарственного обеспечения в медицинской организации / Н. Н. Перова [и др.] // Кардиоваскулярная терапия и профилактика. – Москва, 2021. – Т. 20, № S1. – С. 68.

10. Программа для коррекции режима дозирования препаратов при хронической болезни почек у пациентов с сахарным диабетом / С.Д. Кургалин [и др.] // Информатика: проблемы, методы, технологии: сборник матер. XXV Международ. науч.-метод. конф. им. Э. К. Алгазинова. – Воронеж, 2025. – С. 296-301.

11. Программа для регистрации нежелательных лекарственных реакций в медицинских организациях / С. Д. Кургалин [и др.] // Информатика: проблемы, методы, технологии: сборник матер. XXV Международ. науч.-метод. конф. им. Э. К. Алгазинова. – Воронеж, 2025. – С. 302-306.

12. Туровский, Я. А. Сравнительный анализ программных пакетов для работы с искусственными нейронными сетями / Я. А. Туровский, С. Д. Кургалин, А. А. Адаменко // Вестник Воронежского государственного университета. Сер. Системный анализ и информационные технологии. – 2016. – № 1. – С. 161–168.

Автоматизация расчетов при изучении 7-мерных алгебр Ли

С. И. Бутовецкая, e-mail: butovetskaya@cs.vsu.ru
И. В. Храпов, e-mail: khrapov_i_v@sc.vsu.ru
В. К. Каверина, e-mail: kaverina_v_k@sc.vsu.ru

Воронежский государственный университет

Аннотация. *Статья посвящена применению систем компьютерной алгебры для анализа 7-мерных алгебр Ли, автоматизированного поиска их реализаций и построения орбит в многомерных пространствах. Результатом разработанного комбинированного алгоритма является получение невырожденных орбит для одного типа алгебр Ли из достаточно обширного семейства. Подобные орбиты важны в инвариантных моделях описания человеческой мимики и в экономичных компьютерных алгоритмах преобразования изображений.*

Ключевые слова: *алгебра Ли, коммутационные соотношения, тождество Якоби, Леви-невырожденность, голоморфные векторные поля, компьютерная алгебра, преобразование изображений.*

Введение

В работе [1] предложено использовать технику алгебр Ли и их орбит в многомерных (комплексных) пространствах при описании лицевой мимики. При поиске подходящих для этих целей алгебр возникает необходимость проведения большого количества однотипных сложных вычислений. Классификация многомерных алгебр Ли, поиск их реализаций в виде векторных полей и построение инвариантных многообразий (орбит) требуют анализа тысяч комбинаторных возможностей и проверки достаточно сложных дифференциальных тождеств, как показано в [2-4]. Полный анализ таких структур является очень объемной задачей, подверженной ошибкам.

В этой связи использование программных средств при проведении полномасштабного исследования превращается из вспомогательного инструмента в необходимое условие для обеспечения полноты и достоверности результатов, их систематизации и анализа.

Статья описывает программные средства, предоставляющие функционал для проведения эффективного анализа семимерных алгебр Ли с пятимерным гейзенберговым нильрадикалом.

© Бутовецкая С. И., Храпов И. В., Каверина В. К., 2026

Автоматизация ручных вычислений позволит не только сократить время, но и повысить их надежность и точность.

1. Постановка задачи

В рамках данной работы была поставлена задача автоматизировать вычислительную нагрузку при анализе 7-мерных алгебр Ли и нахождении их Леви-невырожденных орбит в комплексном пространстве C^4. В работе обсуждаются лишь алгебры из классификационной статьи [5], в которой описаны 27 типов 7-мерных алгебр Ли с 5-мерным гейзенберговым нильрадикалом. Обозначим этот нильрадикал через N_5 и учтем в дальнейшем, что в этой 5-мерной алгебре Ли имеется всего два нетривиальных коммутационных соотношения:

$$[e_1, e_3] = [e_2, e_4] = e_5 \qquad (1)$$

Конечной целью автоматизации является нахождение всех алгебр Ли из [5], допускающих реализации в виде алгебр голоморфных векторных полей в C^4 и имеющих невырожденные 7-мерные орбиты.

Процесс исследования отдельной алгебры Ли (с использованием необходимых условий невырожденности ее орбит из работы [7]) можно разложить на следующие этапы:

1. На основе «закодированного» описания выбранной 7-мерной алгебры Ли в [5] выполнить построение наглядной полной таблицы коммутационных соотношений для базиса $\{e_1,..,e_7\}$ этой алгебры;
2. Проверить полученную таблицу коммутационных соотношений на выполнение тождеств Якоби, что является необходимым условием корректности алгебры Ли;
3. Учитывая возможности упрощения тройки коммутирующих полей e_3, e_3, e_3 из N_5, касательных к невырожденным орбитам, рассмотреть пять подслучаев, связанных с потенциально возможными упрощениями базисных векторных полей всей алгебры N_5 в пространстве C^4;
4. В каждом подслучае, опираясь на коммутационные соотношения конкретной алгебры, определить вид оставшихся полей e_6 и e_7 и сделать вывод о возможных противоречиях этих соотношений условиям подслучая. Наличие противоречий означает невозможность требуемой реализации алгебры данного типа в рамках обсуждаемого подслучая;

5. Если для какого-либо подслучая найдена непротиворечивая реализация алгебры в виде голоморфных векторных полей в C^4, выполнить интегрирование этой алгебры для нахождения соответствующей интегральной гиперповерхности (орбиты);
6. Финальным этапом является проверка, является ли полученная гиперповерхность **Леви-невырожденной**, что соответствует исходному условию задачи.

2. Средства реализации программного продукта

Для реализации программного кода был выбран пакет для аналитических вычислений Maple, содержащий более двух тысяч команд, которые позволяют решать задачи алгебры, геометрии, математического анализа.

Были использованы следующие библиотеки:

– *LinearAlgebra* – предоставляет набор команд для работы с матрицами и векторами;
– *LieAlgebra* – реализует прямой функционал для работы с алгебрами Ли. С её помощью были автоматизированы наиболее трудоёмкие шаги;
– *DifferentialGeometry* – используется для выполнения вычислений в дифференциальной геометрии.

3. Реализация логики программного продукта

Отметим сразу, что из всего массива в 27 типов лишь немногие алгебры, как показывают предварительные теоретические исследования, порождают содержательную геометрию в C^4 – то есть имеют невырожденные 7-мерные орбиты. Основная масса алгебр либо не реализуема в требуемом виде, либо реализуема только тривиальным образом, допускающим лишь вырожденные орбиты.

Уточним, что векторные поля в C^4 мы представляем в упрощенной форме в виде векторов. Например, базисные векторные поля e_k $(k=1,..,7)$ любой реализации 7-мерной алгебры задаем наборами их (голоморфных) компонент:

$$e_k = \left(a_k\left(z_1,z_2,z_3,z_4\right), b_k\left(z_1,z_2,z_3,z_4\right), c_k\left(z_1,z_2,z_3,z_4\right), d_k\left(z_1,z_2,z_3,z_4\right)\right) \quad (2)$$

Весь базис 7-мерной алгебры в таком случае можно представить в виде (7x4)-матрицы, элементами которой являются компоненты семи базисных полей.

Отметим также, что написание единой программы, включающей в себя автоматическое выполнение всех описанных выше этапов

исследования, представляется достаточно сложной задачей. Такая программа была бы весьма громоздкой, а ее применение к относительно простым противоречивым ситуациям не всегда полезно.

В связи с этим представляется целесообразным автоматизировать каждый из этапов отдельно, перенося (по сути, вручную) результаты каждого шага такого исследования на следующий этап. При этом после каждого переноса полезно предусмотреть дополнительные автоматизированные проверки.

Реализованные программные средства помогли упростить расчеты на каждом этапе исследования. Автоматизацию вычислительной нагрузки рассмотрим на примере анализа алгебры типа R_{17}.

Первый шаг данного процесса заключается в построении полной таблицы коммутационных соотношений для базиса $\{e_1,..,e_7\}$. Таблица, сформированная на основе расчетов, представлена ниже.

Таблица

Таблица коммутационных соотношений алгебры R_{17} (параметр b не равен нулю)

	e_1	e_2	e_3	e_4	e_5	e_6	e_7
e_1		e_5				$-e_1$	$-2e_1$
e_2				e_5		$-e_2$	
e_3	$-e_5$					$-e_3-be_4$	$-e_3$
e_4		$-e_5$				be_3-e_4	$-e_4$
e_5						$-2e_5$	$-2e_5$
e_6	e_1	e_2	e_3+be_4	$-be_3+e_4$	$2e_5$		
e_7	$2e_1$		e_3	e_4	$2e_5$		

Вычисления по алгоритму производились с помощью библиотеки *LinearAlgebra*. Для выполнения шагов по вычислению описанной таблицы коммутационных соотношений был задействован функционал для умножения матриц.

Использование библиотеки *LieAlgebras* позволило автоматизировать процесс проверки полученной таблицы коммутационных соотношений.

Например, посчитать коммутатор между полями e_1 и e_2 возможно с помощью функции `LieBracket`, содержащейся в данной библиотеке:

```
com12 := LieBracket(e1, e2);
```

Выполнение тождества Якоби для всех комбинаций коммутаторов является доказательством верности составленной алгебры Ли. Расчет всех тождеств описан с помощью следующего цикла:

```
# Проверяем тождество Якоби для всех троек базисных векторов
for i from 1 to 3 do
    for j from 1 to 3 do
        for k from 1 to 3 do
            # Вычисляем левую часть тождества Якоби
            lhs := LieBracket(e||i, LieBracket(e||j, e||k))
                   + LieBracket(e||j, LieBracket(e||k, e||i))
                   + LieBracket(e||k, LieBracket(e||i, e||j));

            # Проверяем, равно ли 0
            if lhs_simplified <> 0 then
                print([i, j, k], lhs_simplified);
            end if;
        end do;
    end do;
end do;
```

На третьем этапе исследования алгебры Ли типа R_{17} проведен анализ пяти подслучаев базиса нильрадикала N_5. В каждом подслучае, опираясь на коммутационные соотношения, определен вид оставшихся полей e_6 и e_7, если это возможно. Так, для алгебр типа R_{17} в четырех подслучаях была обнаружена невозможность реализации данного типа алгебры в C^4 с невырожденной орбитой. В подслучаях 1.1, 1.2 и 2.1 сталкиваемся с линейной зависимостью двух векторов, что доказывает вырожденность по Леви [5-6], а в подслучае 2.2 обнаруживается противоречие – параметр b обращается в 0, что противоречит исходным условиям.

В рамках подслучая 1.3 простейшие признаки вырождения всех орбит в 4-мерном комплексном пространстве (например, 6 нулей в одном столбце упомянутой выше матрицы) не были обнаружены. Данные вычисления реализованы с помощью библиотек *LieAlgebras* и *DifferentialGeometry*.

Таким образом, удалось реализовать алгебру типа R_{17}. Вид этой алгебры при некоторых частных значениях параметров, возникающих в процессе исследования, представлен ниже:

$$
\begin{aligned}
e_1 &= \left(0; \frac{1}{2i} z_1^2; z_1; -z_2\right) \\
e_2 &= \left(0; z_1; i; -z_3\right) \\
e_3 &= \left(0; 1; 0; 0\right) \\
e_4 &= \left(0; 0; 1; 0\right) \\
e_5 &= \left(0; 0; 0; 1\right) \\
e_6 &= \left(ibz_1; -z_2 + bz_1 z_3; \left(ib - 2\right) z_3; -2z_4 - \frac{1}{2} bz_3^2\right) \\
e_7 &= \left(z_1; z_2; 0; 0\right)
\end{aligned}
\qquad (3)
$$

Реализация была проверена на достоверность выполнением тождеств Якоби (с помощью библиотек программного продукта Maple, описанных выше).

Так как в одном из подслучаев была найдена реализация алгебры типа R_{17}, следующим шагом в исследовании является интегрирование и получение уравнения орбит. Для автоматизации расчетов по этому пункту анализа подходят следующие функции библиотек программного средства Maple:

1. `diff (f, var)` – вычисляет частную производную функции;
2. evalc(expr) – преобразует комплексные выражения в алгебраическую форму с выделением их вещественных и мнимых частей;
3. `solve({eq1, eq2, ...}, vars)` – решает систему алгебраических уравнений;
4. `pdsolve` - решает систему уравнений в частных производных.

С помощью применения описанных выше методов, было получено решение системы уравнений в частных производных, описывающей орбиты реализации алгебры Ли типа R_{17} в рамках подслучая 1.3. Данное решение имеет вид:

$$f\left(x_1, y_1, y_2, y_3, y_4\right) = C \qquad (4)$$

где C – произвольная константа, а функция определяется выражением:

$$f = \frac{(-2y_4\left(x_1^2 + y_1^2\right) - x_1^2y_3^2 + y_1^2y_3^2 - y_1y_2y_3 + 2y_2^2)e^{-2\arctan\left(\frac{x_1}{y_1}\right)}}{(x_1^2 + y_1^2)} \quad 5$$

Семейство искомых орбит алгебры R_{17}, являющихся однородными вещественными гиперповерхностями в комплексном пространстве C^4, описывается формулой:

$$\left(x_1^2 + y_1^2\right)\left(2y_4 + Ce^{2\arctan\left(\frac{x_1}{y_1}\right)}\right) - \quad 6$$
$$-y_3^2\left(y_1^2 - x_1^2\right) - 2y_2\left(y_2 - 2y_1y_3\right) = 0$$

Полученные гиперповерхности являются Леви-невырожденными, что устанавливается дополнительной проверкой.

С помощью данного программного кода были проверены и другие типы алгебр. Многие из них оказались нереализуемы, например, алгебра F_6. По результатам анализа оказалось, что подслучаи 1.2 и 2.2 являются вырожденными, подслучаи 1.1, 1.3 и 2.1 приводят к противоречиям.

Заключение

Разработанный комбинированный алгоритм позволил автоматизировать и систематизировать исследование предложенных 27 типов алгебр, упростив выполнение ключевых этапов работы: от построения и верификации коммутационных соотношений до поиска конкретных реализаций в виде векторных полей и последующего интегрирования орбит.

Одним из основных результатов применения этого алгоритма является получение невырожденных орбит для одного семейства алгебры Ли из обширной обсуждаемой совокупности.

Список литературы

1. Крутских В. В. Применение алгебраических инвариантов для моделирования лицевой мимики // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2025. – № 3. – С. 51–62.

2. Акопян Р. С. О голоморфных реализациях пятимерных алгебр Ли / Р. С. Акопян, А. В. Лобода // Алгебра и анализ. – 2019. – . 31, № 6. – С. 1-37.

3. Vu A. L. Classification of 7-dimensional solvable Lie algebras having 5-dimensional nilradicals / A. L. Vu, T. A. Nguyen, T. T. C. Nguyen, T.

T. M. Nguyen, T. N. Vo // Communications in Algebra. –2023. –Vol. 51, № 5. –P. 1866-1885.

4. Parry A. R. A Classification of Real Indecomposable Solvable Lie Algebras of Small Dimension with Codimension One Nilradicals : магистерская диссертация / A. R. Parry. – Logan, Utah, 2007.

5. Rubin J. R. Solvable Lie algebras with Heisenberg ideals / J. R. Rubin, P. Winternitz // Journal of Physics A: Mathematical and General. –1993. –Vol. 26. –P. 1123-1138.

6. Лобода А. В. Разложимые пятимерные алгебры Ли в задаче о голоморфной однородности в C3 / А. В. Лобода, А. В. Атанов // Итоги науки и техники. Современная математика и её приложения. –2019. –Т. 173. –С. 86-115.

7. Атанов А. В. О невырожденных орбитах 7-мерных алгебр Ли, содержащих 3-мерный абелев идеал / А. В. Атанов, А. В. Лобода // Современные проблемы математики и фундаментальные приложения : СМФН. –2024. –Т. 70, № 4. –С. 517-532.

Геометрический подход к интерпретации теоремы Островского

А. Г. Буховец, e-mail: aabuhovets@mail.ru[1]
Т. Я. Бирючинская, e-mail: bir_tat@mail.ru[1]
А. К. Горностаев, e-mail: agornostaevv@gmail.com[2]

[1]Воронежский государственный аграрный университет имени Императора Петра I
[2] БиАйЭй-Технолоджиз

Аннотация. *Теорема Островского о существовании только двух видов норм утверждает, что в поле рациональных чисел могут иметь место только вещественная форма и p-адическая. Использование рандомизированных систем итерированных функций позволяет наглядно изобразить этот факт в геометрическом представлении. Показано, что этот факт связан с наличием доминирующего элемента среди координат векторов случайных разбиений.*

Ключевые слова: *теорема Островского, неархимедовы пространства, рандомизированные системы итерированных функций, доминирующие элементы.*

Введение.
О теореме Островского и неархимедовых пространствах.

Теорема Островского представляет собой фундаментальный результат, опубликованный Александром Островским в 1916 год. Суть теоремы представляет утверждение, что в поле рациональных чисел могут существовать только два вида норм: вещественная форма и p-адическая [13, 17]. Вещественная норма представляла собой обычный, всем хорошо знакомый модуль рационального числа, а p-адическая норма определялась как $\|x\|_p = p^{-ord_p x}$, где p – простое число, $ord_p x$ – функция, численно равная кратности вхождения p в разложение x на простые множители [13]. Характеризуется p-адическая норма тем, что пространства, на которых она задана такая норма являются неархимедовыми [4, 9, 17]. Этими нормами, с точностью до эквивалентности, исчерпывают все нетривиальные неэквивалентные нормы поля.

© Буховец А. Г., Бирючинская Т. Я., Горностаев А. К., 2026

О рандомизированных системах итерированных функций

Рандомизированные системы итерированных функций (РСИФ) [1, 2]. представляют в самом общем виде определенный набор функций $\{f_i\}_{i=1}^{K}$, выполняемых в случайном (random) порядке. Результатом выполнения РСИФ такой последовательности действий является некоторое компактное множество, как правило (но не обязательно), обладающее фрактальными свойствами [3,10, 15]. Удобным с точки зрения машинной реализации является подход, при котором на каждой итерации реализуется только одна из функций f_i, взятая в соответствии с заданным вероятностным распределением $\{f_j | p_j\}_{j=1}^{K}$, $\left(p_j > 0, \sum_{j=1}^{K} p_j = 1\right)$ [5]. Наиболее простой вид функций $\{f_j | \}_{j=1}^{K}$, составляющих систему – линейный. В этом случае РСИФ принимает вид

$$x_{n+1} = \xi x_n + (1-\xi) Z_j^{(n)}$$

где $0 < \xi < 1$ – фиксированное значение (параметр) и $Z_j^{(n)} \in Z$ – точка множества, выбранная на n шаге итерации из фиксированного набора $Z = \{Z_n\}_{n=1}^{K} \subset R^d$.

Итерационный процесс начинается с произвольно взятой точки $x_0 \in R^d$, не обязательно, но желательно $x_0 \in conv(Z)$. После выполнения N итераций вида (1) будет получена совокупность точек $X = \{X_i\}_{i=1}^{N} \left(X_i \in R^d\right)$, называющихся предфракталом. Такой способ построения аттрактора РСИФ будем обозначать F1 (Fractal method 1) [1, 6]. Отметим, что в силу своей простоты он легко может быть программно реализован.

Наряду с этим методом для построения фрактального множества используется другой метод, обозначаемый F2 (Fractal method 2), который в случае совпадения гиперпараметров процедур топологически эквивалентен F1 [7, 8]. Этот метод заключается в построении пространства случайных разбиений абсолютно сходящегося ряд

$$\mu \sum_{i=1}^{\infty} \xi^i = \mu\left(\xi + \xi^2 + \xi^3 + \cdots + \xi^s + \cdots\right) = 1$$

В соответствии с распределением $\{p_j\}_{j=1}^{K}, \left(p_j > 0, \sum_{j=1}^{K} p_j = 1\right)$ задаются координаты вектора A_n как множество чисел, определенных на

этих подмножествах по стандартным правилам сложения и умножения, т.е. числа вида

$$\alpha = \mu \sum_{n'} \xi'$$

где n' - индексы отобранных случайным образом в сумму элементов последовательности ряда (2). Таким образом будут сформированы кортежи (векторы)

$$A_n = \left\{ a_{n1}, a_{n2}, \ldots a_{nm}, \ldots a_{nK} \right\}$$

где $a_i = (1-\xi) \sum_S \xi^S$, $i = 1,2,\ldots,K$, причем $a_i > 0$, $\sum_{i=1}^{K} a_i = 1$. В последствии из этих векторов-строк образуется матрица $A_{N \times K}$, посредством которой можно получить множество точек $X = \left\{ x_i \right\}_{i=1}^{N}$ предфрактала в соответствии с преобразованием $x_n = A_n \cdot Z$. Показано [6], что среди элементов A_n обязательно имеется доминирующий элемент, т.е. $\hat{a}_{im} > \sum_{l \neq j}^{K} a_{il}$. Наличие доминирующего элемента позволяет, во-первых, ввести на фрактальном множестве предварительно определенную норму $\left\| A_i \right\|_\xi = \max_{1 \le j \le K} \left| a_{ij} \right| = \hat{a}_{im}$, где $a_{ij} \in \mathrm{A}_\xi$, $A_i \in \mathrm{A}_\xi^K$, которая в свою очередь индуцирует ультраметрику [9, 11],

$$d_\xi (A_s, A_t) = \left\| A_s - A_t \right\|_\xi$$

Кроме этого наличие доминирующего элемента в пространстве R^d позволяет выделить, задав отношения эквивалентности, два класса: α - отношение эквивалентности, соответствующие точкам разбиения единичного отрезка, из которых можно составить замкнутый контур, и класс β - эквивалентности.

О связи p-адических пространств теоремы Островского и ультраметрических пространств A_ξ^K, образованных с помощью РСИФ

В статье [7] об ультраметрических пространствах и рандомизированных системах итерированных функций показано, что

классификационное пространство A_{ξ}^{K} построенное с помощью случайных разбиений, обладает ультраметрическими свойствами. Это пространство можно рассматривать как некоторый объект, где ультраметрическая структура возникает естественным образом из внутренней логики.

Ультраметрическое пространство A_{ξ}^{K}, образованное множеством случайных разбиений процедуры F2, как легко показать является неархимедовым. Его связь с *p*-адическим пространством можно установить с помощью преобразования параметра РСИФ $\xi = 1/p$. В случае теоремы Островского предполагается, что *p* – простое число. Это предположение существенно используется в доказательстве поскольку требование, накладываемое на p, играет важную роль ввиду мультипликативного свойства вводимой нормы. Принятие этого предположения приводит к тому, что множество значений *p*-адической нормы будет счетным. В выполнении РСИФ такое требование не является обременительным, поскольку параметр ξ просто выбирается из единичного интервала. В некотором смысле это ограничение снимается при построении случайных разбиений, что позволяет расширить возможности A_{ξ}^{K}.

Сходство РСИФ и утверждения теоремы Островского заключается в построении ультраметрических пространств и интерпретации полученных результатов. В случае РСИФ построение проводится на основе геометрического подхода и соответствующих рассуждений, что позволяет использовать геометрическую/топологическую интерпретацию. Во втором случае доказательство проводится преимущественно алгебраическими методами. В ходе такого доказательства вводятся более строгие ограничения на основание ультраметрического пространства.

Общие черты этих подходов, и это главная особенность сравниваемых методов, заключается в построении неархимедовых пространств, т.е. пространств, в которых нарушается аксиома Архимеда о возможности покрыть любой отрезок другими отрезками с достаточной степенью точности. [Напоминание: поле с неархимедовой абсолютной величиной называется неархимедовым полем.]

Как следует из теоремы Островского поля *p*-адических чисел — это единственные примеры неархимедовых числовых полей, получаемых пополнением поля рациональных чисел.

Практически, геометрическое представление, возникающее как результат выполнения процедуры РСИФ, позволяет интерпретировать теорему Островского в терминах фрактальной геометрии. Кроме этого отметим, что рассматривая величину параметра $\xi = 1/p$, где p – не обязательно целое и простое число, мы получаем некоторое обобщение p-адического пространства.

Обратная теорема к теореме Островского может быть сформулирована несколькими возможными вариантами. Она может пониматься как утверждение о всех локально компактных пополнениях поля рациональных чисел, как теорема о представлении абсолютных значений через параметры, или как утверждение о классификации нормированных полей.

Если выбрать для рассмотрения утверждение, обратное к теореме Островского, то один из возможных вариантов заключается в том, что если абсолютное значение на множестве рациональных чисел удовлетворяет усиленному неравенству треугольника, то оно должно быть эквивалентно одному из p-адических абсолютных значений. Однако доказательство этого утверждения мы пока оставляем для будущих исследований.

Заключение

В контексте решаемой классификационной задачи мы не используем свойства простых чисел, и поэтому не встречаем трудностей, связанных с ограничениями на использование класса значений параметра ξ в РСИФ. С другой стороны, при таком подходе это позволяет рассматривать непрерывный спектр пространств при непрерывном изменении значения параметра. Отметим только, что полученные таким образом пространства не являются гомеоморфными, как, например Q_5 и Q_7.

Список литературы

1. Bukhovets, A. G. / Modeling of fractal data structures // Bukhovets A. G., Bukhovets E.A.// Automation and Remote Control, Vol. 73, No.2 (2012), 381-385.
2. Barnsley M.F. Superfractals / M.F. Barnsley – Cambridge: Cambridge University Press, 2006. – 464 p.
3. Hutcthinson J.E. Fractals and Self Similararity. Indiana University Mathematics Journal, vol. 30, #5, 1981, pp. 713-747.
4. Аветисов В.А. Ультраметрическое случайное блуждание и динамика белковых молекул /В.А. Аветисов, А.Х. Бикулов, А. П. Зубарев.// Труды мат. ин-та им. В.А. Стеклова, 2014, т.285, с. 4-24.

5. Буховец А.Г. Структура аттрактора рандомизированных линейных систем итерированных функций / А.Г. Буховец, Т.Я. Бирючинская// Вестник ВГУ, Системный анализ и информационные технологии, 2016, №2, С.5 – 10.

6. Буховец, А. Г. Моделирование фрактальных свойств системных объектов / А. Г. Буховец, Т. Я. Бирючинская // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2011. – № 2. – С. 22-25. – EDN NDVRVH.

7. Буховец, А. Г. Устойчивость аттрактора систем рандомизированных функций / А. Г. Буховец, Т. Я. Бирючинская, А. К. Горностаев // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2024. – № 2. – С. 5-. – DOI 10.17308/sait/1995-5499/2024/2/5-14. – EDN CDNBOC.

8. Буховец, А. Г. Фрактальные аспекты моделирования классификационной задачи / А. Г. Буховец, Е. А. Семин // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2022. – № 3. – С. 127-138. – DOI 10.17308/sait/1995-5499/2022/3/127-138. – EDN TMGLCG.

9. Владимиров В.С. р-Адический анализ и математическая физика / В.С. Владимиров, И.В. Волович, Е.И. Зеленов – М.: Наука, 1994

10. Гелиашвили Д.Б. и др. Фракталы и мультифракталы в биоэкологии: монография. – Н.Новгород: Изд-во Нижегородского госуниверситета, 2013. – 370 с.

11. Кириллов, А.А. Теоремы и задачи функционального анализа// А.А. Кириллов, А.Д. Гвишиани. – М.: Наука, 1979. – 381 с.

12. Кроновер, Р.М. Фракталы и хаос в динамических системах/ Р.М. Кроновер – М.: ТЕХНОСФЕРА, 2006. – 488 с.

13. Коблиц Н. р-адические числа,р-адический анализ и дзета-функции/ Н. Коблиц // М.: Мир, 1982. – 192 с.

14. Кузнецов С.П. Динамический хаос. / С.П. Кузнецов. – М.: Физматлит, 2006. – 356 с.

15. Мандельброт, Б. Фрактальная геометрия природы / Б.Мандельброт// М., 2002. – 656 с.

16. Хренников, А.Ю. Моделирование процессов мышления в р-адических системах координат//А.Ю. Хренников. – М.: ФИЗМАТЛИТ, 2004.- 296с.

17. Хренников, А.Ю. Неархимедов анализ и его приложения// А.Ю. Хренников. – М.: ФИЗМАТЛИТ, 2003.- 216с. с.

Алгоритм объединения разнородных метеорологических данных в единое «поле потенциальной опасности» для конвективных явлений

Н. А. Веремьёв, e-mail: nikover1395@yandex.ru
А. Н. Маслобойщиков, e-mail: nikover1395@yandex.ru
А. С. Голунов, e-mail: nikover1395@yandex.ru
Н. С. Потапов e-mail: nikover1395@yandex.ru

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *Алгоритм основан на принципах байесовской стратегии и теории нечетких множеств, что позволяет учитывать различную точность, пространственно-временное распределение и физико-географические особенности местности, где получены исходные данные. Проведена валидация алгоритма на данных за 2022-2025 годы для Центрального региона России, результаты показали, что предложенный алгоритм обеспечивает увеличение точности определения зон потенциальной опасности на 25% по сравнению с традиционными методами линейной интерполяции.*

Ключевые слова: *опасные явления погоды, конвективные явления, штормовое предупреждение, метеорологическое обеспечение.*

Введение

Оперативное прогнозирование и обнаружение опасных конвективных явлений (ОКЯ) требует интеграции информации из множества разнородных источников: метеорологических радиолокаторов, спутников, наземных станций, грозопеленгационных сетей и гидродинамических моделей прогноза ОКЯ [1, 2]. Каждый из этих источников определяет различные характеристики по точности, пространственно-временному разрешению, физическим принципам измерений и оперативности поступления данных.

Основная проблема при определении метеорологической обстановки заключается в следующем:

– Пространственная неоднородность данных, поступающих с различным разрешением (от 50 км до 100 км).

– Временная асинхронность периодичности обновления варьируется от 15 минут (радары) до 6 часов (спутниковые снимки).

© Веремьев Н. А., Маслобойщиков А. Н., Голунов А. С., Потапов Н. С., 2026

– Физическая разнородность измерений отражают разные характеристики атмосферы (гидрометеоры, температура, электрическая активность).

– Различная достоверность данных, которые имеют собственную погрешность, зависящую от условий наблюдений.

– Традиционные методы интерполяции и анализа не способные адекватно учесть физико-географические особенности [3].

Цель исследования – разработка математической модели и алгоритма для объединения разнородных метеорологических данных для построения единого поля потенциальной опасности конвективных явлений с учетом характеристик каждого источника данных.

Алгоритм объединения разнородных метеорологических данных

Рассмотрим множество источников метеорологических данных выражение 1.

$$S = \{s_1,\ s_2, ..., s_n\}$$

где n – количество источников.

Каждый источник $s_i s_j$ предоставляет данные $D_i(x, y, t)$ в точке пространства-времени (x, y, t) с определенной погрешностью $\varepsilon_i \varepsilon_j$.

Определим поле потенциальной опасности $H(x, y, t)$ как вероятность возникновения ОКЯ в точке (x, y) в момент времени t (выражение 2):

$$H(x, y, t) = \mathrm{P}(\mathrm{D}_1, D_2, ...D_b)$$

Задача заключается в вычислении $H(x, y, t)$ на основе всех доступных данных с учетом их погрешностей, разрешения и физической согласованности.

Рассмотрим объединение четырех ключевых источников для выявления потенциальной градоопасности:

1. Метеорадар: отражаемость Z на высоте -10°C изотермы.
2. Спутник GOES-R: разность яркостных температур каналов 6.2 и 10.8 мкм.
3. Грозопеленгационная сеть: плотность разрядов в ячейке 10×10 км.
4. Численная модель WRF: индекс CAPE (потенциальная конвективная энергия).

Табл. 1 показывает параметры интеграции для этих источников.

Таблица

Параметры интеграции для различных источников данных

Источник данных	**Физическая величина**	**Порог опасности**	**Весовой коэффициент**	**Временное запаздывание**
Метеорадар	Z на высоте -10°C	>50 dBZ	0.35	0-5 мин
Спутник GOES-R	ΔT(6.2-10.8 мкм)	>15 K	0.25	10-15 мин
Грозопеленгатор	Плотность разрядов	>2 разр/5 мин	0.20	<1 мин
Модель WRF	CAPE	>1500 Дж/кг	0.20	30-60 мин

Предлагаемый алгоритм протестирован на данных за период 2021-2025 гг. для Центрального региона России.

Использовались исходные данные сети метеорадаров Росгидромета (разрешение 1×1 км, период 10 мин), спутниковые данные GOES-16 и Метеор-М (разрешение 2-4 км), данные грозопеленгационной сети, выходные данные модели WRF-ARW (разрешение 3 км). При анализе данных произвелась оценка эффективности алгоритма с использованием метрик Probability of Detection (POD), False Alarm Ratio (FAR), Critical Success Index (CSI), Root Mean Square Error (RMSE) табл. 2.

Таблица

Параметры интеграции для различных источников данных

Метод объединения	**POD**	**FAR**	**CSI**	**RMSE**	**Время обработки**
Линейная интерполяция	0.72	0.41	0.48	0.38	1.2 с
Метод k-ближайших соседей	0.78	0.35	0.55	0.31	3.5 с
Простая байесовская модель	0.81	0.30	0.60	0.28	4.8 с
Предложенный алгоритм	**0.92**	**0.24**	**0.72**	**0.21**	**6.3 с**

На рисунке представлен пример построения поля потенциальной опасности для случая 15 июля 2022 года при объединении разнородных данных.

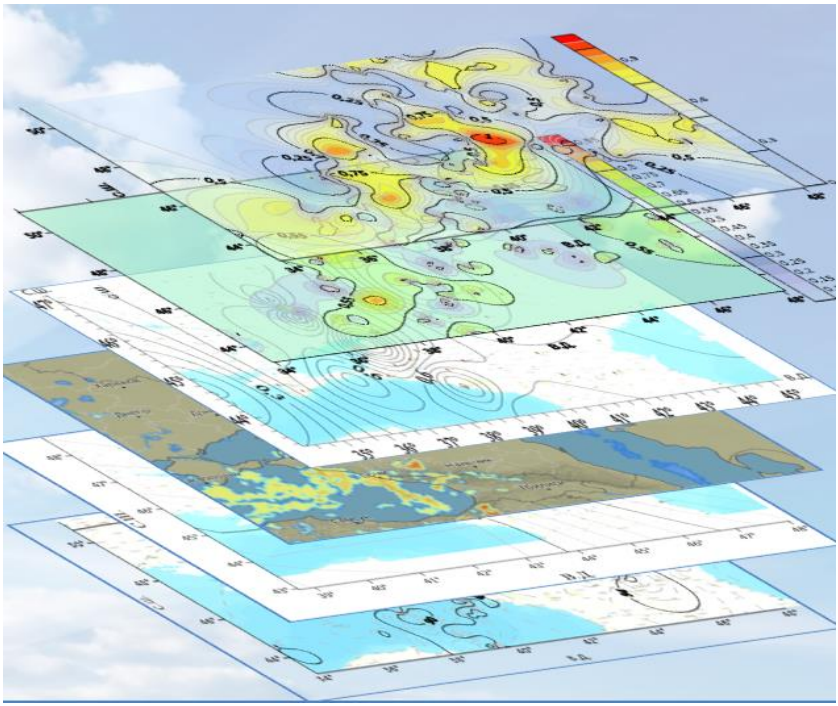


Рисунок. Пример поля потенциальной опасности, различных методов объединённых данных

При наложении полей разнородных данных выявляется поле потенциальной опасности с наибольшей вероятностью образования ОКЯ, что позволяет заблаговременно определить зоны рисков и принять меры по минимизации ущерба.

Заключение

В настоящей работе получены следующие результаты:

1. Разработан алгоритм объединения разнородных метеорологических данных в единое поле потенциальной опасности для конвективных явлений. Алгоритм основан на принципах байесовского вывода и теории нечетких множеств, что позволяет учитывать различную точность и физическую природу исходных данных.

2. Предложена математическая модель, включающая функции принадлежности для каждого источника данных, корреляционную матрицу взаимного влияния и итерационный процесс уточнения оценки

опасности. Модель адаптивна и может быть расширена за счет включения новых источников данных.

3. Реализована программная система, осуществляющая предобработку данных, вычисление функций принадлежности, байесовское объединение и постобработку результатов. Система обеспечивает обработку данных в квазиреальном времени.

4. Верификация алгоритма на данных за 2022-2025 годы показала его преимущество перед традиционными методами: увеличение POD на 25 %, снижение FAR на 15%, повышение CSI на 25%. Алгоритм особенно эффективен в условиях частичной потери данных от отдельных источников.

5. Практическая значимость работы заключается в возможности использования разработанного алгоритма в системе штормового предупреждения об опасных явлениях погоды.

Разработанный алгоритм представляет комплексную обработку метеорологических данных и может стать основой для создания систем мониторинга и прогнозирования опасных конвективных явлений.

Список литературы

1. Bringi, V. N. Polarimetric Doppler Weather Radar: Principles and Applications / V. N. Bringi, V. Chandrasekar. – Cambridge : Cambridge University Press, 2001. – 662 p.

2. Маслобойщиков, А. Н. Модель предварительного метеорологического прогноза при различной информационной обеспеченности / А. Н. Маслобойщиков, Н. А. Веремьёв, Н. С. Потапов // Проблемы, методы, технологии : сб. материалов XXV Междунар. науч.-практ. конф. (Воронеж, 12–13 февр. 2025 г.) / под ред. Д. Н. Борисова. – Воронеж : Воронежский гос. ун-т, 2025. – С. 313–317.

3. Маслобойщиков, А. Н. Модель предварительного прогнозирования метеорологических условий образования опасных явлений погоды / А. Н. Маслобойщиков, Н. А. Веремьёв, А. В. Каморников // Проблемы, методы, технологии : сб. материалов XXIV Междунар. науч.-практ. конф. (Воронеж, 14–15 февр. 2024 г.) / под ред. Д. Н. Борисова. – Воронеж : Воронежский гос. ун-т, 2024. – С. 404–410.

4. Маслобойщиков, А. Н. Модель оценки информации, используемой при гидрометеорологическом обеспечении потребителей / А. Н. Маслобойщиков, Н. А. Веремьёв, А. И. Аристов // Проблемы, методы, технологии : сб. материалов XXIII Междунар. науч.-практ. конф. (Воронеж, 15–17 февр. 2023 г.) / под ред. Д. Н. Борисова. – Воронеж : Воронежский гос. ун-т, 2023. – С. 349–356.

5. Маслобойщиков, А. Н. Оценка объёма метеорологической информации / А. Н. Маслобойщиков, Н. А. Веремьёв, М. А. Прохоров // Проблемы, методы, технологии : сб. материалов XXII Междунар. науч.-практ. конф. (Воронеж, 10–12 февр. 2022 г.) / под ред. Д. Н. Борисова. – Воронеж : Воронежский гос. ун-т, 2022. – С. 351–357.

Модель метеорологического обеспечения авиационных задач на основе элементов линейной регрессии

Д. В. Гедзенко, e-mail: gdv555900@mail.ru
Д. В. Бодрых

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *Проведено исследование влияния высоты нижней границы слоистообразной облачности на безопасность полетов воздушных судов, а также эффективность выполнения различных задач. Разработан эффективный алгоритм, который повышает качество метеорологического обеспечения полетов воздушных судов с учетом точности прогнозирования высоты нижней границы облаков. Сравнение с существующими способами показал эффективность и надежность разработанной модели, что дает право использования в практической деятельности метеорологических подразделений при обеспечении полетов авиации.*

Ключевые слова: *слоистообразная облачность, уравнение регрессии, показатель точности.*

Введение

Современная техника, не смотря на свою автономность, чувствительна к неблагоприятным метеорологическим условиям, что негативно сказывается на работе многих систем, в том числе и авиации. Данное исследование подтверждает существенное влияние метеорологических условий на безопасность и эффективность полетов авиации. Метеорологического обеспечения позволяет максимально эффективно использовать метеорологические данные для обеспечения безопасности полетов и эффективного применения авиации.

Быстрое развитие авиационной техники, увеличение скорости полета, диапазона высот и внедрение новых технологий с учетом беспилотных летательных аппаратов, требуют повышения точности и достоверности прогнозирования метеорологических условий. Это достижимо благодаря применению современных теоретических основ, совершенствованию методик прогнозирования и выбору наиболее эффективных способов прогноза с учетом региональных особенностей базирования.

© Гедзенко Д. В., Бодрых Д. В., 2026

Низкая облачность оказывает существенное влияние на безопасность и эффективность функционирования авиационных систем (АС). Из общего количества авиационных происшествий, связанных с неблагоприятными условиями погоды 62% были вызваны понижением высоты нижней границы облаков (ВНГО), создающей основную трудность, а иногда и опасную обстановку, на наиболее ответственных этапах полета - взлете и заходе на посадку. Поэтому крайне важно располагать надежной прогностической информацией о значениях ВНГО на период выполнения полетов [1].

Существующие и наиболее часто используемые в практической деятельности метеорологических подразделений методы прогноза ВНГО не в полной мере удовлетворяют требованиям, предъявляемым руководящими документами. Данный факт требует разработки новых прогностических методов, основанных на применении современного физико-статистического аппарата. Кроме того повышение качества прогноза, непосредственно связанной с ее точностью является основой для полетов авиации.

Быстрое развитие авиационной техники, увеличение скорости полета, диапазона высот и внедрение новых технологий с учетом беспилотных летательных аппаратов, требуют повышения точности и достоверности прогнозирования метеорологических условий. Это достижимо благодаря применению современных теоретических основ, совершенствованию методик прогнозирования и выбору наиболее эффективных способов прогноза с учетом региональных особенностей базирования.

Актуальность данной работы очевидна: она направлена на повышение точности прогнозирования метеорологических условий погоды с использованием данных в каждом метеорологическом подразделении.

Цель работы - повышение качества прогнозирования высоты нижней границы слоистообразной облачности.

1. Анализ влияния слоистообразной облачности на полеты воздушных судов

Облаком называют совокупность взвешенных капель воды и кристаллов льда, находящейся на некоторой высоте над земной поверхностью [2]. С точки зрения микрофизического строения принципиальной разницы между облаками и туманом нет. Однако они существенно отличаются по условиям образования, вертикальной протяженностью и т. п.

В целом облака влияют на все этапы полета. Поэтому взлет, посадка и полет на эшелоне самолетов и вертолетов производится с обязательной

оценкой состояния облачности (при наличии облачности или ожидаемом возникновении) и ее возможного влияния на полеты. Облака оказывают влияние на пилотирование воздушных судов и самолетовождение, а также исключают возможность дозаправки самолетов топливом в воздухе и создают трудности для боевого применения авиации. Данный факт обусловлен ограничением видимости, турбулентностью, возможностью обледенения и электризации летательного аппарата в облаках.

Как известно, при разработке прогноза метеорологических величин синоптику необходимо каким-то образом связать значение известных параметров состояние атмосферы с определяемыми значениями неизвестных. Подобные проблемы разрешаются с помощью статистического анализа связей, в том числе с помощью регрессионного анализа.

В настоящее время в практике метеорологических подразделений ВВС наиболее широкое распространение получили методы прогноза ВНГО, разработанные Е. И. Гоголевой и З.А. Спарышкиной. Эти методы основаны на различных эмпирических зависимостях, полученных на материалах определенного физико-географического района с учетом типа синоптических условий, сезона, заблаговременности прогноза и периода суток.

В табл. 1 приведена информация о требуемой и средней точности прогнозов ВНГО, разработанных в пункте Воронеж.

Таблица

Требуемая и средняя точность прогноза ВНГО

Метод прогноза ВНГО	Диапазон измерения	Средняя точность
Требуемая точность	до 120м, м	±30
	120-3000м, %	±30
Пункт Воронеж		
Е.И. Гоголевой	до 120м, м	±38,9
	120-300м, %	±42,3
З.А. Спарышкиной	до 120м, м	±36,9
	120-300м, %	±40,3

2. Модель метеорологического обеспечения авиации на основе линейной регрессии

Для оптимизации способа прогноза в работе была сформирована архивная выборка по пункту Воронеж.

Выборка была разработана путем расчета разницы метеорологических элементов, взятых на пункте прогноза и в исходном

районе, за пять лет 2019-2023 г в холодное время года. Объем выборки составил 150 случаев. В качестве предикторов использовались значения разниц метеорологических элементов между исходным районом и пунктом прогноза, а именно: скорости ветра, дальности видимости, дефицита точки росы, температуры воздуха у земли и высоты нижней границы облаков за 18:00 часов.

А в качестве предиктанта, взята высота нижней границы облаков за 06:00 часов. Таким образом, разработали прогноз высоты нижней границы облаков на срок 12 часов.

Для построения уравнения регрессии был произведён отбор предикторов с использованием процедуры просеивания. На первом этапе составлены прогностические зависимости с каждым из предикторов. Затем по обучающей выборке для каждого уравнения рассчитывался множественный коэффициент корреляции и стандартная ошибка. По значениям данных показателей качества выбиралось лучшее уравнение. С помощью программы MicrosoftExcel получили матрицу коэффициентов корреляции.

Таблица

Полученные уравнения регрессии и результаты анализа по контрольной выборке

Лучшие уравнения регрессии	Средняя абсолютная ошибка	Средняя квадратичная ошибка	Парный коэффициент
Уравнение регрессии с 1 -м предиктором $y = 260,9 + 101,59\Delta DD$	89,9261	111,1499	0,5320
Уравнение регрессии с 2-я предикторами $y = 82,514 + 145,41\Delta DD + +0,5916H_{18}$	56,9341	74,1623	0,8371
Уравнение регрессии с 3-я предикторами $y = 63,617 + 91,011\Delta DD + +0,693H_{18} + 31,882\Delta ff$	50,5997	70,0207	0,8463

Продолжение табл. 1

Лучшие уравнения регрессии	Средняя абсолютная ошибка	Средняя квадратичная ошибка	Парный коэффициент корреляции
Уравнение регрессии с 4-я предикторами $y = 58,251 + 73,91\Delta DD +$	47,3414	66,3134	0,8653
Уравнение регрессии с 5-ю предикторами $y = 56,731 + 65,189\Delta DD +$ $+0,746H_{18} + 22,769\Delta ff +$ $+13,62\Delta VV - 4,983\Delta T$	47,7312	66,4804	0,8753

Уравнение признается лучшим, если добавление очередного предиктора не приводит к улучшению выбранного показателя качества[3].

Результаты проведенного анализа по контрольной выборке, с использованием уравнений регрессии, в зависимости от количества предикторов выведены на рисунке.

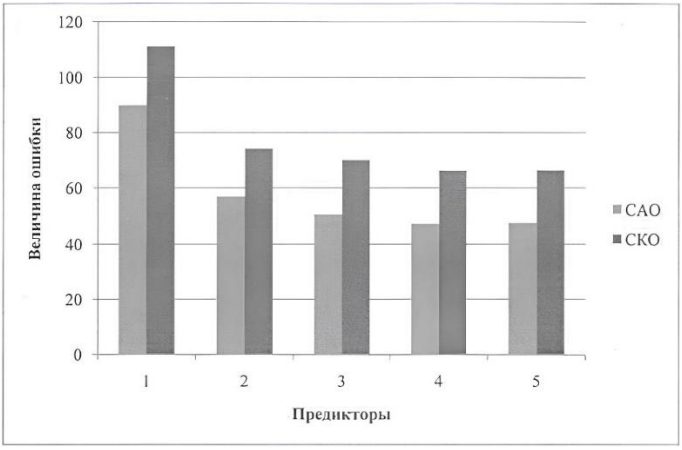


Рисунок. Анализ полученных уравнений

Проверив по контрольной выборке, получим, что лучшим является уравнение, полученное на четвертом шаге с четырьмя предикторами.

Проведен анализ существующих способов прогноза низкой слоистообразной облачности (способы Е.И.Гоголевой, З.А.Спарышкиной) [4]. Для прогнозирования с наименьшей ошибкой

были построены уравнения регрессии с различными метеорологическими величинами и выбрано наилучшее уравнение. Для их оценки рассчитаны: средняя абсолютная ошибка, средняя квадратическая ошибка и коэффициент корреляции, и произведено сравнение с существующими способами.

При оценке полученного уравнения рассчитаны следующие показатели: средняя абсолютная ошибка 47,3414; средняя квадратическая ошибка 66,313; парный коэффициент корреляции 0,8653.

При определении ВНГО методом З.А Спарышкиной получены следующие результаты: средняя абсолютная ошибка составила 91,5, средняя квадратичная ошибка составила 101,565.

При определении ВНГО методом Е.И Гоголевой получены следующие результаты: средняя абсолютная ошибка составила 125,1, средняя квадратичная ошибка составила 140,112.

Анализ показателей точности полученного в работе уравнения свидетельствует о возможности его использования в практической деятельности метеорологических подразделений, не противоречащих федеральным авиационным правила производства полетов [5].

Заключение

Цель работы, заключающаяся в повышении качества метеорологического обеспечения при выполнении полетов различных типов авиации достигнута.

Полученные результаты рекомендуется использовать в научных исследованиях при разработке новых способов, а также при совершенствовании существующих методов прогноза высоты нижней границы облаков.

Список литературы

1. Руководство по практическим работам метеорологических подразделений авиации Вооруженных Сил. – М. : Воениздат, 1992. – 486 с.

2. Матвеев, Л. Т. Курс общей метеорологии. Физика атмосферы / Л. Т. Матвеев. – Л. : Гидрометеоиздат, 1984. – 752 с.

3. Дрейпер, Н. Прикладной регрессионный анализ. Множественная регрессия / Н. Дрейпер, Г. Смит. – 3-е изд. – М. : Диалектика, 2007. – 912 с.

4. Зверев, А. С. Синоптическая метеорология / А. С. Зверев. – Л. : Гидрометеоиздат, 1997. – 711 с.

5. Федеральные авиационные правила производства полетов государственной авиации Российской Федерации (ФАППП-2004). – М. : Воениздат, 2004. – 320 с.

Модель поддержки принятия решений при выполнении авиационных задач с использованием элементов теории детерминированного хаоса

С. Л. Кирносов, e-mail: slk_met@mail.ru
А. Э. Бадыгин, e-mail: shap_m@mail.ru
Е. А. Долгов, e-mail: arbad1a2@mail.ru

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *В работе рассматривается возможность применения научно-методического подхода к моделированию возникновения риска при выполнении авиационных задач в сложных метеорологических условиях с использованием элементов теории детерминированного хаоса. Рассматривается устойчивость динамических систем в сложных метеоусловиях.*

Ключевые слова: *авиационная задача, авиационный риск, детерминированный хаос, аттрактор, динамическая система.*

Введение

В модели поддержки принятия решений при выполнении авиационных задач с использованием элементов теории детерминированного хаоса исследуется процесс выполнения авиационной задачи на примере посадки воздушного судна (ВС). При этом моделируется величина риска, связанного с попаданием ВС в сложные метеорологические условия (СМУ). Модели, исследующие такие процессы в большинстве случаев являются нелинейными. Однако они показывают при проверке невысокую эффективность в связи с существующей метеорологической неопределенностью и неопределенностью действий летчика в СМУ [1]. Основная причина – принципиальная нелинейность рассматриваемой динамической системы «летчик – ВС– метеоусловия».

Внутренними параметрами системы «летчик – ВС– метеоусловия» являются: степень подготовки летчика к выполнению авиационной задачи, конструктивные особенности ВС, уровень подготовки аэродрома, уровень подготовки группы руководства полетами. Внешними параметрами системы являются метеорологические условия.

Следует отметить крайне высокую чувствительность всех внутренних параметров к метеоусловиям. Поэтому величина

© Кирносов С. Л., Бадыгин А. Э., Долгов Е. А., 2026

моделируемого риска, возникающего при выполнении посадки ВС будет формализовываться с использованием отдельных элементов теории детерминированного хаоса [1-4].

Таким образом, целью данной работы является повышение эффективности принимаемых решений при выполнении посадки ВС в СМУ на основе теории детерминированного хаоса.

Исходные параметры и построение модели

При моделировании метеорологического риска в работе, в рамках теории детерминированного хаоса, предложено использовать универсальное рекуррентное соотношение (1):

$$P_{n+1} = kP_n(1-P_n) \ ,$$

где k – показатель сложности метеорологических условий (зависит от высоты нижней границы облаков и посадочной видимости); P_{n+1} и P_n – вероятности возникновения аварийной ситуации в последующие моменты времени $n+1$ и n, соответственно [4, 5];

Анализ (1) показывает, что при $k \in [0, 4]$ значения вероятности возникновения риска будут находиться в диапазоне $0 \le P \le 1$.

Бифуркационная диаграмма выражения (1) представлена на рис. **Ошибка! Источник ссылки не найден.**. На этом же рисунке внизу представлена глиссада посадки ВС.

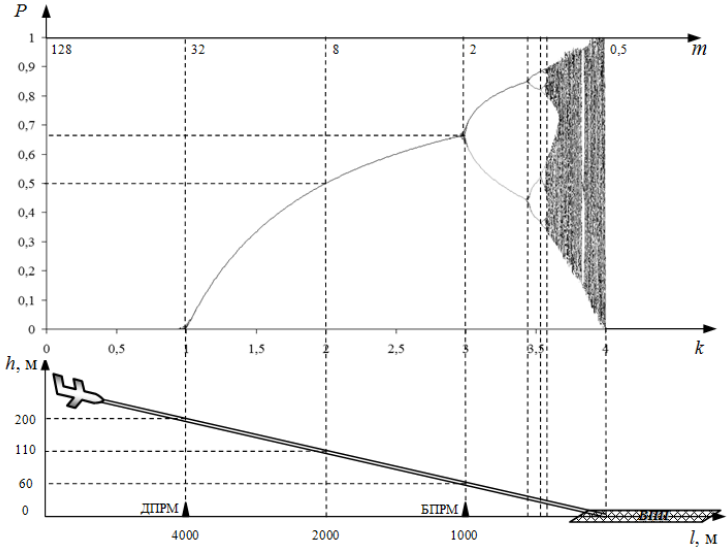


Рисунок. Бифуркационная диаграмма процесса посадки ВС

Количество итераций, необходимое для релаксации модели на основе (1), составляет $n = 55$-60. Применительно к процессу выполнения посадки ВС $n \approx 60$ секундам, в течение которых летчик выполняет снижение от четвертого разворота, дальнего приводного радиомаркера (ДПРМ), ближнего приводного радиомаркера (БПРМ) и т.д., вплоть до точки касания на взлетно-посадочной полосе (ВПП).

Значения параметра k при 0 соответствуют простым метеоусловиям (ПМУ), а при 4 – нелетной погоде. Таким образом, при изменении k от 0 до 4 моделируются значения возникающего риска при ухудшении погодных условий от простых к сложным, вплоть до нелетных [2-5].

Интервал $0 \le k \le 1$, при котором P_n стремится к нулю, соответствует ПМУ.

На интервале значений $1 < k \le 2$ вероятность возникновения риска увеличивается до значений 0,5. Соответствует той ситуации, когда летчик на высоте 100 м (удалении до ВПП ≈ 2 км) не установил визуальный контакт с ВПП.

На интервале $2 < k < 3{,}45$ вероятность возникновения аварийной ситуации увеличивается до 0,66. Соответствует той ситуации, когда летчик на высоте 60 м (удалении до ВПП ≈ 1 км) все еще не установил визуальный контакт с ВПП. Из рисунка видно, что при k =3,0 возникает

первая бифуркация диаграммы. Это интерпретируется как наличие ситуации, когда в СМУ подготовленный летчик выполняет посадку в штатном режиме, или наличие той ситуации, когда ВС уходит на второй круг или запасной аэродром.

Интервал значений $3,45 < k \le 4$ соответствует нелетной погоде. При этом наблюдается удвоение периода состояний системы, что свидетельствует о переходе ее в состояние хаоса по сценарию Фейгенбаума. В условиях нелетной погоды вероятность возникновения аварийной ситуации стремится к единице.

Из рис. **Ошибка! Источник ссылки не найден.** видно, что даже небольшие отклонения в численных значениях k, связанных с высотой нижней границы облаков и посадочной видимостью, приводят к переходу авиационной динамической системы на принципиально новый режим функционирования в связи с изменившимися метеорологическими условиями.

Для семейства отображений, одно из которых представлено выражением (1), существует постоянная Фейгенбаума, называемая универсальным масштабным множителем σ. Этот множитель отражает закономерность процесса дробления масштабов диаграммы при бифуркациях. Если обозначить расстояния от точки $P = 0,5$ до ближайшей к ней точки на устойчивом цикле $f = 2^i$ ($i = \overline{1,\ N}$) через d_i (рис. **Ошибка! Источник ссылки не найден.**), то отношение d_i к d_{i+1} при $i \to \infty$ удовлетворяет соотношению:

$$\sigma = \lim_{i \to \infty} \frac{d_i}{d_{i+1}} = 2,5029078 \ .$$

Таким образом, при построении модели поддержки принятия авиационных управленческих решений, учитывающих скрытые свойства метеозависимости, возможно и целесообразно использовать количественные и качественные методы теории фракталов и теории детерминированного хаоса. Поиск устойчивых и безопасных режимов функционирования авиационных систем необходимо проводить на основе анализа инвариантных множеств (аттракторов). Анализ аттракторов показывает, во-первых, их связь универсальным масштабным подобием, а во-вторых, – наличие устойчивых фракталов.

Список литературы

1. Матвеев, М. Г. Управление организационно-технической системой в условиях метеорологической неопределенности : монография

/ М. Г. Матвеев, В. В. Михайлов. – Воронеж : ВВВАИУ (ВИ), 2006. – 128 с.

2. Михайлов, В. В. Бифуркационная модель посадки воздушного судна в условиях метеорологической неопределенности / В. В. Михайлов, М. Е. Семенов, С. Л. Кирносов // Нелинейный мир. – 2014. – № 4. – Т. 12. – С. 57-63.

3. Михайлов, В. В. Модель принятия метеозависимых решений на выполнение посадки самолета в условиях детерминированного хаоса / В. В. Михайлов, С. Л. Кирносов // Системы управления и информационные технологии. – 2014. – № 2 (56). – С. 80-84.

4. Михайлов, В. В. Синтез методов фрактальной и комплексной динамики при построении системы поддержки принятия метеозависимых решений / В. В. Михайлов, С. Л. Кирносов // Системы управления и информационные технологии. – 2014. – № 3 (57). – С. 85-88.

5. Базарский, О. В. Потенциальная информационная емкость изображений, обрабатываемых автоматизированной системой и зрительным аппаратом человека / О. В. Базарский, В. В. Михайлов, С. Л. Кирносов // Системный анализ и информационные технологии. Вестник Воронежского государственного университета. – 2006. – № 1. – С. 20-27.

Анализ уязвимостей протокола Wi Fi 6 (IEEE 802.11ax) к пассивному перехвату и методы защиты на канальном уровне

И.В. Климкин, e-mail: nauchnajarota@yandex.ru
Г.А. Лахно, e-mail: nauchnajarota@yandex.ru
Н.С. Уткин, e-mail: nauchnajarota@yandex.ru

Межвидовой центр подготовки и боевого применения войск РЭБ
(учебный и испытательный)

Аннотация. *В статье проводится комплексный анализ уязвимостей протокола Wi Fi 6 (IEEE 802.11ax) к пассивному перехвату трафика. Рассматриваются особенности архитектуры протокола, включая технологию OFDMA и MU MIMO, создающие новые векторы атак. Предложены математические модели оценки вероятности перехвата и эффективности методов защиты на канальном уровне. Проведенное моделирование показало, что применение адаптивного шифрования с динамической сменой ключей и внедрение протоколов аутентификации на основе доверенных платформ снижают вероятность успешного пассивного перехвата на 91 %. Результаты исследования имеют практическую значимость для разработки защищенных беспроводных сетей следующего поколения.*

Ключевые слова: *Wi Fi 6, IEEE 802.11ax, пассивный перехват, уязвимости, защита на канальном уровне, OFDMA, криптографические методы, доверенная платформа.*

Введение

Современные беспроводные сети сталкиваются с возрастающими угрозами информационной безопасности, особенно в условиях массового внедрения технологий следующего поколения. Протокол Wi Fi 6 (IEEE 802.11ax), обеспечивающий высокую пропускную способность, вводит новые механизмы, которые могут создавать дополнительные уязвимости к пассивному перехвату трафика.

Актуальность исследования обусловлена несколькими факторами. Во первых, согласно данным Cisco Annual Internet Report, к 2025 году количество устройств Wi Fi 6 превысит 4.7 миллиарда, что делает их привлекательной целью для злоумышленников. Во вторых, особенности архитектуры Wi Fi 6, такие как ортогональное частотное разделение каналов (OFDMA) и многопользовательский MIMO (MU MIMO),

© Климкин И.В., Лахно Г.А., Уткин Н.С., 2026

создают новые векторы атак, недостаточно изученные в сфере информационной безопасности.

Целью работы является анализ уязвимостей протокола Wi Fi 6 к пассивному перехвату и предложение более эффективного метода защиты на канальном уровне. Для достижения поставленной цели решаются следующие задачи: изучение архитектурных особенностей Wi Fi 6; идентификация уязвимостей к пассивному перехвату; разработка математических моделей оценки безопасности; предложение и оценка эффективности методов защиты.

1. Архитектурные особенности Wi Fi 6 и их уязвимости

Протокол IEEE 802.11ax вводит ряд инновационных технологий, повышающих эффективность работы беспроводных сетей. Ключевой технологией является OFDMA (Orthogonal Frequency Division Multiple Access), которая разделяет канал на множество поднесущих, распределяемых между пользователями. Математическая модель передачи с использованием OFDMA описывается следующим выражением:

$$s(t) = \sum X_k \cdot g(t - kT_s) \cdot e^{(j2\pi f_k t)}$$

Несмотря на преимущества, технология OFDMA создает уязвимость к пассивному перехвату. Поскольку разные пользователи используют разные поднесущие одновременно, злоумышленник может перехватывать трафик конкретного пользователя, фокусируясь на соответствующих частотах. Вероятность успешного перехвата можно оценить как:

$$P_{intrcept} = 1 - e^{(-\lambda \cdot T_{capture})}$$

Другой важной технологией является MU‑MIMO (Multi‑User Multiple Input Multiple Output), которая позволяет одновременную передачу данных нескольким пользователям. Однако это создает уязвимость к перехвату через пространственное разделение. Если злоумышленник имеет достаточное количество антенн, он может разделить потоки данных и перехватить информацию конкретного пользователя.

На рис. 1 представлена схема пассивного перехвата в сети Wi‑Fi 6 с использованием OFDMA.



Рис. 1. Схема пассивного перехвата в сети Wi‑Fi 6 с OFDMA

2. Методы защиты на канальном уровне

Для защиты от пассивного перехвата в сетях Wi‑Fi 6 предлагается комплексный подход на канальном уровне. Первый метод – адаптивное шифрование с динамической сменой ключей [2]. В отличие от статических ключей WPA3, предложенный метод использует динамическую генерацию ключей на основе временных меток и состояния канала:

$$K_{dynamic} = H(TS \oplus RSSI \oplus MAC_addr)$$

Эффективность метода оценивается коэффициентом снижения вероятности перехвата:

$$R_{reduction} = [\frac{P_{base} - P_{protected}}{P_{base}}] \cdot 100\%$$

Второй метод защиты – аутентификация на основе доверенных платформ (TPM) [1, 2]. Каждое устройство содержит аппаратный модуль TPM, который генерирует уникальные ключи и подтверждает подлинность устройства перед передачей данных. Вероятность успешной аутентификации определяется как:

$$P_A = \left(\frac{N_{trusted}}{N_{total}} \right)^{e(-\alpha \cdot d)}$$

На рис. 2 показана эффективность предложенных методов защиты в сравнении с базовыми решениями.

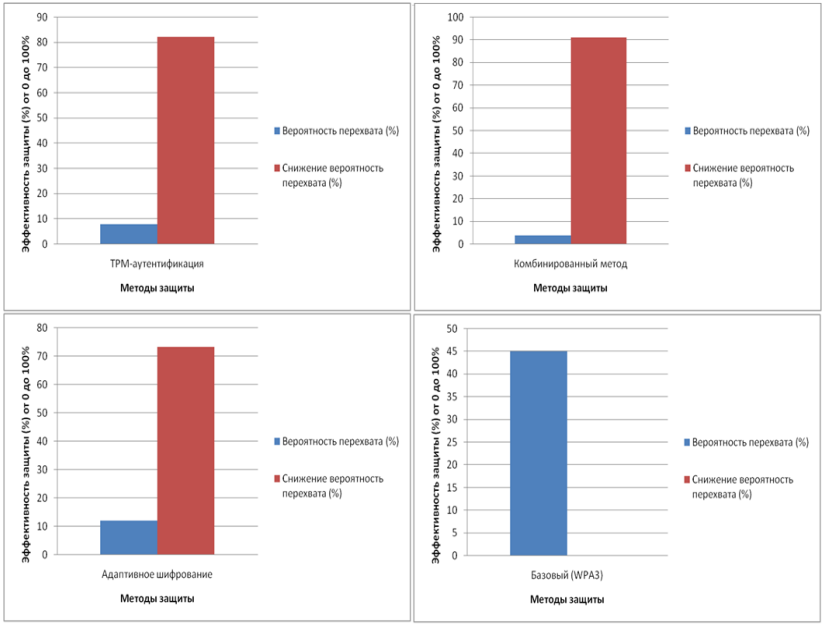


Рис. 2. Сравнение эффективности методов защиты от пассивного перехвата

3. Экспериментальные результаты и обсуждение

Для оценки эффективности предложенных методов проведено моделирование в среде MATLAB с использованием пакета Communications Toolbox. Тестовая сеть включала 1 точку доступа Wi‑Fi 6 и 7 клиентских устройств, симулирующих различную активность.

Результаты экспериментов представлены в табл. 1. Базовая конфигурация (WPA3) показала вероятность перехвата 0,45 при интенсивности атаки 10 пакетов/сек. Применение адаптивного шифрования снизило этот показатель до 0,12, а комбинация с TPM-аутентификацией – до 0,04.

Как видно из таблицы, комбинированный метод обеспечивает наибольшую защиту, снижая вероятность перехвата на 91.1%, при этом накладные расходы остаются в допустимых пределах (6,5 мс).

График на рис. 3 демонстрирует зависимость вероятности перехвата от интенсивности атаки для разных методов защиты. Комбинированный метод сохраняет низкие показатели даже при высокой интенсивности атаки.

Таблица

Результаты экспериментов по защите от пассивного перехвата

Метод защиты	Вероятность перехвата	Снижение, %	Накладные расходы, мс
Базовый (WPA3)	0,45	–	2,1
Адаптивное шифрование	0,12	73,3	3,8
TPM-аутентификация	0,08	82,2	5,2
Комбинированный метод	0,04	91,1	6,5

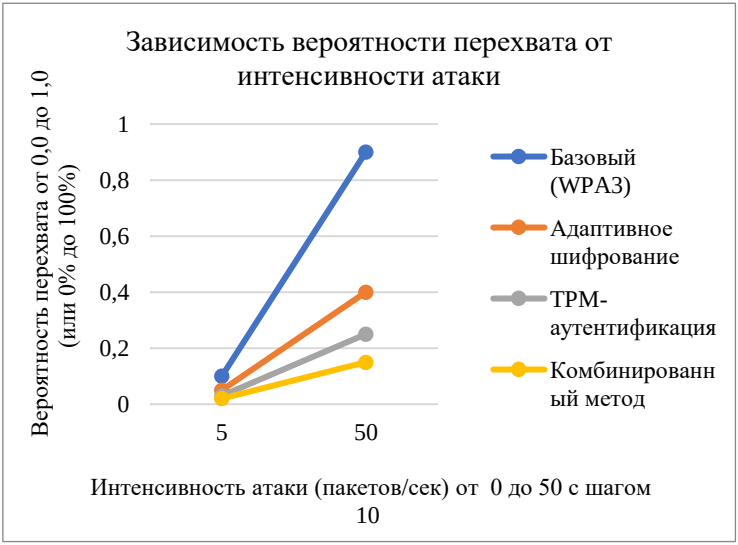


Рис. 3. Зависимость вероятности перехвата от интенсивности атаки

Заключение

Проведенное исследование показало, что протокол Wi‑Fi 6 имеет специфические уязвимости к пассивному перехвату, связанные с особенностями OFDMA и MU‑MIMO технологий. Предложенный комбинированный метод защиты, включающий адаптивное шифрование с динамической сменой ключей и TPM-аутентификацию, снижает

вероятность успешного перехвата на 91,1% при приемлемых накладных расходах.

Практическая значимость работы заключается в возможности внедрения разработанных методов в существующие системы беспроводной связи для повышения их информационной безопасности. Перспективы дальнейших исследований включают оптимизацию алгоритмов генерации ключей и интеграцию методов защиты с технологиями искусственного интеллекта для адаптивного реагирования на угрозы.

Список литературы

1. Киселев, В. И. Методы защиты беспроводных сетей от пассивного перехвата / В. И. Киселев, А. С. Петров // Информационные технологии и безопасность. – 2023. – Т. 18, № 2. – С. 89–102.

2. ГОСТ Р 57580.1–2017. Защита информации. Алгоритмы шифрования. – Введ. 01.07.2018. – М. : Стандартинформ, 2017. – 24 с.

3. Андрианов, В. И. Экспертные системы в области информационной безопасности / В. И. Андрианов, Г. Г. Романов, С. И. Штеренберг // Актуальные проблемы инфотелекоммуникаций в науке и образовании. – 2015. – С. 193–197.

4. Ковцур, М. М. Исследование способов удаленного перехвата трафика в корпоративных сетях / М. М. Ковцур // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия. – 2021. – Т. 1. – С. 68–75.

5. Гельфанд, А. М. Оценка рисков и угроз безопасности в среде «Умный дом» / А. М. Гельфанд // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО-2020). – 2020. – С. 316–321.

Имитационная модель процесса управления диаграммами направленности антенн пеленгаторов при угломерно-корреляционном местоопределении источников электромагнитного излучения

Ю. Л. Козирацкий,
М. Л. Паринов,
С. В. Петренков, e-mail: sema1709@yandex.ru

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *Разработана имитационная модель процесса управления диаграммами направленности антенн пеленгаторов при угломерно-корреляционном местоопределении, позволяющая оценить точность определения координат достигаемую за счет изменения ориентирования диаграмм направленностей в пределах элемента разрешения.*

Ключевые слова: *Определение координат, угломерно-корреляционный метод, управление диаграммой направленности, имитационная модель, линейная ошибка.*

Введение

В настоящее время все больше возрастают требования к оперативности реализации функционала средств, обеспечивающих вскрытие радиоэлектронной обстановки, в том числе, в части, касающейся определения координат источников излучения [1,2]. В работах [3-7] предлагается метод угломерно-корреляционного местоопределения источников радиоизлучения, обеспечивающий за счет процедур синхронной пространственно перестройки амплитудных пеленгаторов и корреляционной обработки формируемых сигнально-шумовых функций, уменьшение времени, затрачиваемого на определение координат источников. Вместе с тем, рассматриваемый метод координатометрии при решении ряда практических задач не в полной мере удовлетворяет требованиям по точности местоопределения при, так как область неопределенности при угломерно-корреляционном поиске и местоопределении ограничивается пересечением диаграмм направленностей антенных систем пеленгаторов. Парировать подобный недостаток возможно за счет обеспечения перестройки диаграмм направленностей антенных систем в пределах элемента

© Козирацкий Ю. Л., Паринов М. Л., Петренков С. В., 2026

пространственного разрешения, что позволит уточнить оценки угловых координат источника излучения (рис. 1).

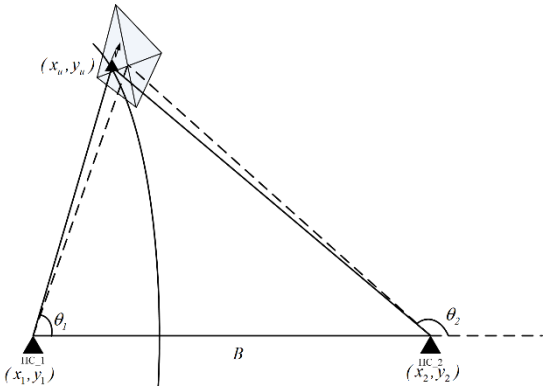


Рис. 1. Геометрия задачи повышения точности определения координат при угломерно-корреляционном поиске и местоопределении

Таким образом, интерес представляет определение точности оценки координат, реализуемой за счет фокусировки диаграмм направленностей антенных решеток пеленгаторов в пределах элемента разрешения, что определяет цель настоящей статьи, которая заключается в разработке модели процесса управления диаграммами направленностей антенн пеленгаторных пар, реализующих угломерно-корреляционный метод поиска и координатометрии, учитывающей влияние стохастических факторов при оценке информативных параметров, и обеспечивающей при этом статистическую оценку величины линейной ошибки определения координат источника излучения.

Разработка модели

Рассмотрим аналитическую модель, обеспечивающую расчет координат источника излучения, формируемую, опираясь на приведенные выше условия получения линий положения источника, а также особенности определения исходных информативных для расчета координат параметров.

Оценка времени задержки распространения излучения источника до пеленгаторных станций $\widehat{\tau}_{12}$ осуществляется на основе максимизации величины корреляционной функции [5], рассчитываемой в дискретном времени согласно следующему соотношению:

$$\eta(T) = \Delta t[\sum_{i=0}^{N-1} \xi_1(t_i)\xi_2(t_i)],$$

где $\xi_{1(2)}(t_i)$ - отчет реализации сигнально-шумовой функции соответствующей пеленгаторной станции, взятый в момент времени t_i; Δt - элементарный временной интервал между соседними отчетами времени; N - количество отчетов на интервале времени накопления $T = N\Delta t$.

Таким образом, время задержки вычисляется максимизацией функционала вида:

$$\widehat{\tau}_{12} = \arg\max_{\substack{j=\overline{0,N-1} \\ \tau_j = j\Delta t}} \{\sum_{i=1}^{N} \xi_1(t_i)\xi_2(t_i - \tau_j)\}.$$

Оценку угловой координаты, характеризующей направление на источник излучения в системе координат пеленгаторных станций, а также точность такой оценки целесообразно реализовать, опираясь на результаты, полученные в работах [8].

Рассмотрим в качестве антенны пеленгаторной станции M - элементную кольцевую антенную решетку радиусом R с равномерным распределением ненаправленных в азимутальной плоскости антенных элементов.

Комплексную амплитуду принятой антенным элементом сигнально-шумовой функции представим в виде аддитивной смеси сигнала и дельта-коррелированного гауссовского шума:

$$\dot{\xi}_i = S_{0i} e^{-j\psi_i} \dot{F}_i(\theta) + \dot{n}_i,$$

где $\psi_i = \frac{2\pi}{\lambda} R\cos(\theta_0 - \theta_i)$ - фаза сигнала, принимаемого i - ым антенным элементом решетки; θ_0 - направление на источник излучения; θ_i - угол, характеризующий пространственное положение i - ого антенного элемента в антенной решетке; $\dot{F}_i(\theta) = e^{j\frac{2\pi}{\lambda}R\cos(\theta-\theta_i)}$ - комплексная диаграмма направленности i - ого антенного элемента решетки; $\dot{n}_i$ - комплексная амплитуда шумовой функции, принимаемой i - ым антенным элементом решетки.

Принимая во внимание результаты, полученные в [9], математическую структуру функции, характеризующей пеленгационный рельеф, получим в следующем виде:

$$\alpha(\theta) = \frac{\left|\sum_{i=1}^{M} \dot{\xi}_i \dot{F}_i(\theta)\right|}{\sqrt{\sum_{i=1}^{M} \left|\dot{F}_i(\theta)\right|^2}}.$$

Искомая оценка угла, характеризующего направление на источник излучения - $\widehat{\theta}_0$, максимизирует функцию пеленгационного рельефа

$$\widehat{\theta}_0 = \arg\max_{\theta}(\alpha(\theta)).$$

Учитывая комплексную диаграмму направленности антенного элемента решетки для пеленгационного рельефа получим

$$\alpha(\theta) = \frac{1}{N}\left|\sum_{i=1}^{M} \dot{\xi}_i e^{j\frac{2\pi}{\lambda}R\cos(\theta-\theta_i)}\right|.$$

Представляя комплексную амплитуду реализации сигнально-шумовой функции, представленную в показательном форме, в выражение (6), с учетом (5) можно показать, что оценка угла направления на источник излучения $\widehat{\theta}_0$ определяется решением уравнения следующего вида:

$$\frac{d}{d\theta}\sqrt{\begin{aligned}&[\sum_{i=1}^{M} \xi_i \cos(\frac{2\pi}{\lambda}R\cos(\theta-\theta_i)+\varphi_{\xi_i})]^2 + \\ &+[\sum_{i=1}^{M} \xi_i \sin(\frac{2\pi}{\lambda}R\cos(\theta-\theta_i)+\varphi_{\xi_i})]^2 = 0.\end{aligned}}$$

$$\dot{\xi}_i = \xi_i e^{j\varphi_{\xi i}}$$

Вычислительная процедура, определяемая соотношением (5), обеспечивает оценку угла, характеризующего направление на источник излучения. Полученную таким образом оценку угла $\widehat{\theta}_0$ предлагается использовать для корректировки диаграммы направленности антенной решетки, классический алгоритм управления которой предусматривает обеспечение синфазного сложения сигналов, принимаемых антенными элементами решетки с заданного направления [9]. Синфазность при сложении сигналов реализуется путем умножения вектора принятых реализаций $\dot{\boldsymbol{\xi}} = (\dot{\xi}_1, ..., \dot{\xi}_i, ..., \dot{\xi}_M)$ на комплексный вектор весовых коэффициентов $\dot{\mathbf{W}}$: $\dot{\boldsymbol{\xi}} \cdot \dot{\mathbf{W}}^T$, структура которого определяется выражением следующего вида:

$$\dot{\mathbf{W}} = (F_1(\theta)e^{j\frac{2\pi}{\lambda}R\cos(\theta_m-\theta_1)},...,F_i(\theta)e^{j\frac{2\pi}{\lambda}R\cos(\theta_m-\theta_i)},...$$

$$..,F_M(\theta)e^{j\frac{2\pi}{\lambda}R\cos(\theta_m-\theta_M)}),$$

где θ_m – угол, определяющий направление, сигналы с которого складываются синфазно.

Таким образом, сущность способа пассивной координатометрии обеспечивающего увеличение точности угломерно-корреляционного местоопределения заключается в переориентировании диаграмм направленности антенных систем пеленгаторов на источник излучения, расположенный в пределах элемента пространственного разрешения, путем корректировки параметров вектора весовых коэффициентов, а именно, определяя величину угла θ_m в выражении (8) равной величине оценке угла $\widehat{\theta}_0$, характеризующего направление на источник излучения.

Используя структуры вычислительных процедур оценок величины времени задержки $\widehat{\tau}_{12}$ и угловых мер, характеризующих направление на источник излучения относительно ПС_1 – $\widehat{\theta}_1$ и ПС_2 – $\widehat{\theta}_2$, была разработана имитационная модель процесса управления диаграммами направленности антенн пеленгаторов при реализации процедур угломерно-корреляционного поиска и координатометрии (рис. 2), и обеспечивающая оценку координат источника излучения ($\hat{x}_u, \hat{y}_u$), а также статистическую оценку величины линейной ошибки такой оценки - r, рассчитываемой на основе приведенного ниже соотношения:

$$r = \sqrt{(\hat{x}_u - x_{u_ист})^2 + (\hat{y}_u - y_{u_ист})^2}\ .$$

Исходные данные для проведения вычислительного эксперимента по оценке величины линейной ошибки местоопределения при реализации вычислительных процедур, определим в следующем виде: расстояние между пеленгаторными станциями $B = 10$ *км*, сигнал – частотно-модулированный, несущая частота – $f_0 = 100$ *МГц*, модуляция гармоническим монохроматическим сигналом частотой $f_м = 20$ *кГц*, индекс модуляции - $m = 3$, антенная решетка – кольцевая, с равномерным распределением облучателей, 5 элементная, радиус решетки $R = 1.5$ *м*.

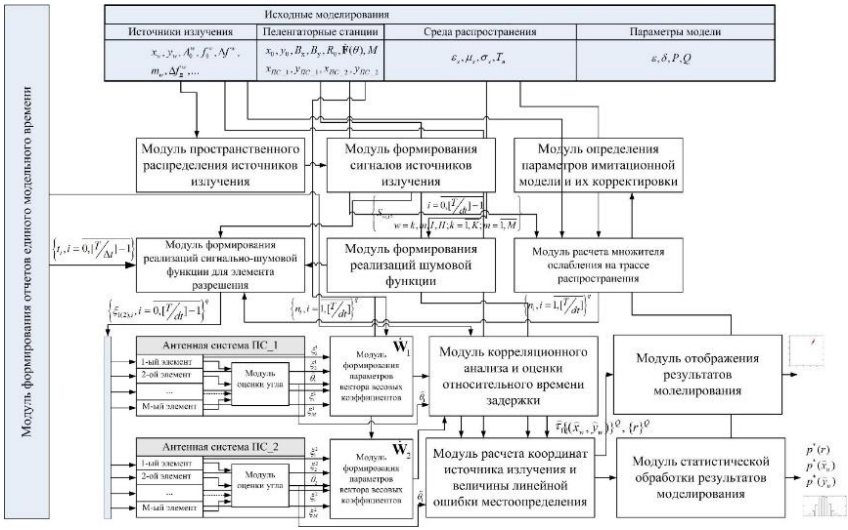


Рис. 2. Структура имитационной модели процесса управления диаграммами направленностей антенн пеленгаторов при реализации процедур угломерно-корреляционного поиска и координатометрии

На рис. 3 представлены результаты оценки информативных параметров, существенных при вычислении координат источника излучения.

Анализ представленных на рис. 3 статистических распределений оценок величин времени задержки и направления максимума диаграммы направленностей антенн пеленгаторов, ориентированных на источник излучения, показывает, что закон распределения оцениваемых параметров близок к нормальному. В частности, вероятность сходимости к нормальному распределению, определенная согласно критерия χ^2, составляет величину $P_{cx} \geq 0.6$. Увеличение значение сигнал/шум приводит к уменьшению среднеквадратического отклонения соответствующих оценок. Помимо этого, на точность оценки угловой меры направления максимума диаграммы направленности влияет частота излучения, в частности, увеличение последней при неизменной значении радиуса кольцевой решетки приводит к уменьшению среднеквадратического отклонения оценки (рис. 3в,), что объясняется уменьшением ширины главного лепестка диаграммы направленности.

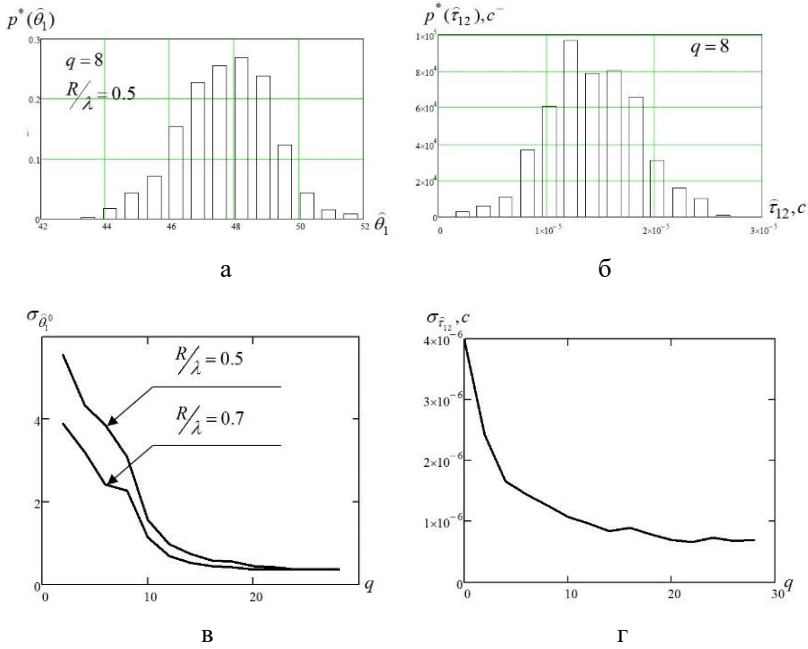
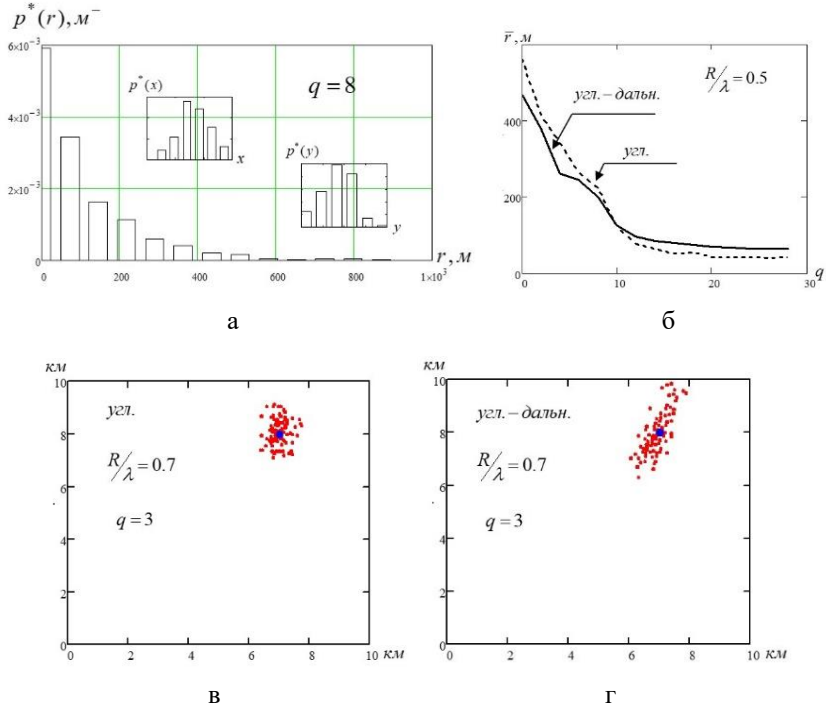


Рис. 3. Статистические распределения исходных параметров при определении координат источника излучения

Реализуемая в рассматриваемой выше вычислительной процедуре оценка направлений максимумов диаграмм направленностей пеленгаторных станций позволяет, воспользовавшись этими оценками, определить местоположение источника излучения в рамках угломерно-корреляционного поиска и координатометрии, согласно приведенных ниже соотношений:

$$\hat{x}_u = \frac{y_2 - y_1 + x_1 tg\hat{\theta}_1 - x_2 tg\hat{\theta}_2}{tg\hat{\theta}_1 - tg\hat{\theta}_2},$$

$$\hat{y}_u = y_1 + \frac{(x_1 - x_2)tg\hat{\theta}_2 + y_2 - y_1}{tg\hat{\theta}_1 - tg\hat{\theta}}.$$



а б

в г

Рис. 4. Зависимости линейной ошибки определения местоположения источника излучения для различных условий оценки информативных параметров

На рис. 4 представлены результаты оценки координат источника излучения, производимой применительно к приведенным выше исходным данным, и реализованной в соответствии с вычислительными процедурами, определяемыми аналитическими выражениями, представленными ранее.

Анализ зависимостей показывает, что статистическое распределение величины линейной ошибки местоопределения близко к распределению Рэлея (рис. 4а), что объясняется нормальной статистикой для соответствующих координат источника излучения. С возрастанием отношения сигнал/шум среднее значение линейной ошибки уменьшается.

Интерес также представляет особенности пространственной структуры области рассеивания оценок координат источника излучения. В частности, в случае, угломерно-дальномерного способа

координатометрии, область рассеивания характеризуется сильным «вытягиванием» в направлении пеленгатора (рис. 4г). При реализации же местоопределения по оценкам пеленгов, параметры области распределения координат источника излучения определяется эллипсом рассеивания, характерным для угломерной системы местоопределения (рис. 4в).

Заключение

На основе аналитической модели и вычислительных процедур статистической оценки линейной ошибки местоопределения разработана имитационная модель процесса управления диаграммами направленности антенн при реализации угломерно-корреляционного метода определения координат источника электромагнитного излучения. Моделирование имеет целью определения достижимой точности в оценке координат при реализации процедур, базирующихся на корректировке параметров весового вектора, управления диаграммами направленности кольцевых антенных решеток пеленгаторов, обеспечивающих их переориентирование в пределах элемента пространственного разрешения. Как показали результаты моделирования, применяемые при угломерно-корреляционной координатометрии решения по управлению диаграммами направленности позволяют за счет увеличения отношения сигнал/шум повысить точность определения местоположения источника излучения при прочих равных условиях в среднем на величину порядка 20-25%.

Список литературы

1. Модели информационного конфликта средств поиска и обнаружения : монография / Ю. Л. Козирацкий [и др.] ; отв. ред. Ю. Л. Козирацкий. – Москва : Радиотехника, 2013. – 232 с.

2. Артемов, М. Л. Автоматизированные системы управления, радиосвязи и радиоэлектронной борьбы : монография / М. Л. Артемов [и др.] ; отв. ред. М. Л. Артемов. – Москва : Радиотехника, 2021. – 575 с.

3. Козирацкий, Ю. Л. Способ определения местоположения источников радиоизлучений с режекцией входного потока на основе корреляционного анализа результатов наблюдения / Ю. Л. Козирацкий, М. Л. Паринов // Радиотехника. – 2013. – № 7. – С. 16–19.

4. Способ пространственного мониторинга источников электромагнитного излучения : пат. 2540126 Рос. Федерация : МПК G01S 5/12 / Ю. Л. Козирацкий [и др.] ; заявитель и патентообладатель ВУНЦ ВВС «ВВА» (г. Воронеж). – № 2013143448/07 ; заявл. 25.09.2013 ; опубл. 10.02.2015, Бюл. № 4. – 8 с.

5. Паринов, М. Л. Проблемные вопросы координатометрии радиоизлучающих средств и пути их разрешения с использованием перспективных радиоэлектронных технологий / М. Л. Паринов // Воздушно-космические силы. Теория и практика. – 2021. – № 20. – С. 266–285.

6. Модели пространственного и частотного поиска : монография / Ю. Л. Козирацкий [и др.] ; отв. ред. Ю. Л. Козирацкий. – Москва : Радиотехника, 2014. – 342 с.

7. Паринов, М. Л. Моделирование процедур пеленгования и координатометрии источников радиоизлучения в условиях стохастического характера изменений параметров среды распространения и воздействия случайных факторов, вызывающих ошибки при измерениях информативных параметров / М. Л. Паринов // Воздушно-космические силы. Теория и практика. – 2024. – № 31. – С. 145–170.

8. Уфаев, В. А. Способы определения местоположения и пространственной идентификации источников радиоизлучений : монография / В. А. Уфаев ; ВУНЦ ВВС «ВВА» (г. Воронеж). – 2017. – 431 с.

9. Кубанов, В. П. Направленные свойства антенных решеток : учеб. пособие / В. П. Кубанов. – Самара : ПГУТИ, 2011. – 56 с.

Компьютерная модель экспресс-оценки точности процесса сопровождения целей телевизионной системой, основанного на определении энергетического центра формируемого составного изображения

Ю. Л. Козирацкий, e-mail: UrLeo@bk.ru
Д. В. Прохоров, e-mail: dvp2201@rambler.ru
А. А. Баркалов, e-mail: barkalov_anton_a@mail.ru

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *С использованием методов теории автоматического управления разработана компьютерная модель, позволяющая проводить экспресс-оценку влияния сложных помеховых и фоноцелевых условий, а также внутренних шумов на точность процесса сопровождения воздушных целей телевизионной системой, использующей алгоритм оценки их текущих координат, основанный на определении энергетического центра формируемого составного изображения.*

Ключевые слова: *модель, сопровождение, оценка, точность, энергетический центр, телевизионная система.*

Введение

Требования нашего времени актуализируют задачи оценивания движения воздушных целей (ВЦ) в реальном времени в сложных условиях наблюдения объекта. Для определения координат траектории движения воздушных объектов (беспилотных летательных аппаратов, ракет, снарядов и др.) активно используются оптико-электронные средства [1, 2]. В целях обеспечения высокоточного сопровождения ВЦ могут применяться телевизионные системы, использующие для оценки угловых координат ВЦ алгоритмы, основанные на определении энергетического центра принимаемого изображения. Для исследования эффективности таких систем в сложных условиях обстановки, в том числе применения способов противодействия, необходимо создание специальных моделей.

Поэтому целью работы являлась разработка компьютерной модели, позволяющей проводить экспресс-оценки влияния сложных помеховых и фоноцелевых условий, а также внутренних шумов на точность процесса сопровождения ВЦ телевизионной системой, использующей алгоритм

© Козирацкий Ю.Л., Прохоров Д.В., Баркалов А.А., 2026

оценки их текущих координат, основанный на определении энергетического центра формируемого составного изображения.

1. Постановка задачи

Процессу сопровождения предшествуют этапы обнаружения ВЦ, оценки текущих координат и нацеливания системы сопровождения. Обнаружение цели телевизионной системой осуществляется по контрастному ее изображению на фоне неба. После обнаружения ВЦ производится оценка ее текущих координат с целью нацеливания системы сопровождения. Оценка координат цели и ее сопровождение производится с использованием многоэлементного приемника с равномерной координатной чувствительностью. На фоне дневного неба ВЦ, как правило, будет иметь меньшую светимость и малую контрастность. В этом случае для корректной работы алгоритма необходимо повышение контрастности изображения и его инвертирование. При формировании изображения на фоточувствительных элементах матричного фотоприемника (МФП) можно определить его энергетический центр путем взвешенного суммирования величин сигналов по строкам и столбцам матричного приемника [3]. Если выверен центр приемника, строго соответствующий оси формирующей оптики, а сами элементарные приемники относительно центра разделены прямоугольной системой координат, то величина отклонения энергетического центра изображения относительно центра матричного приемника может рассматриваться как текущие относительные координаты цели, т.е. в данном случае матричный приемник является измерителем угла рассогласования между оптической осью приемника и направлением на цель.

С использованием предложенного алгоритма в [4] разработана динамическая модель процесса высокоточного сопровождения целей лазерно-тепловизионной системой, позволяющая исследовать точность работы системы сопровождения и закономерности распределения динамических ошибок наведения по каналам курса и тангажа соответственно с учетом инерционности контуров сопровождения, фоновых шумов и преднамеренных помех, степени искривленности траектории полета сопровождаемой цели. Полученные в [3, 4] результаты используем для разработки целевой компьютерной модели.

Качество работы системы сопровождения ВЦ при случайных воздействиях будем оценивать по суммарной динамической ошибке $\Delta U_{\Sigma x}(t)$, представляющей собой сумму динамической ошибки сопровождения цели $\Delta U_x(t)$ и ошибки оценки координат цели $\Delta x(t)$.

2. Разработка аналитической модели

В результате прохождения сигнала через формирующую оптику объектива (оптическую систему) на плоскости МФП формируется изображение (рис. 1). Получаемое после увеличения контрастности и инвертирования изображение рассматривается как составной сигнал видимого диапазона длин волн в виде аддитивной смеси изображений ВЦ (например, двух), а также активной оптической помехи, фона и внутреннего шума фотоприемника:

$$\xi_{\Sigma\lambda}(x,y,t) = \xi_{c1\lambda}(x,y,t) + \xi_{c2\lambda}(x,y,t) + \xi_{n\lambda}(x,y,t) + \\ + \xi_{\phi\lambda}(x,y,t) + \xi_{ш\lambda}(x,y,t),$$

где $\xi_{\lambda}(x,y,t)$ – спектральные плотности энергетической освещенности одного элемента фоточувствительной поверхности МФП с координатами x и y, обусловленные воздействием вышеуказанных излучений и шумов.

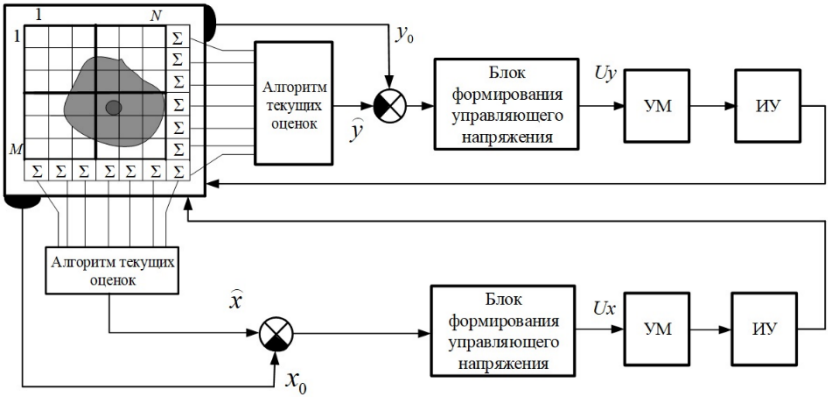


Рис. 1. Обобщенная структурная схема, отображающая способ сопровождения

Далее в соответствии с алгоритмами оценки смещения энергетического центра изображения цели относительно геометрического центра МФП дискриминатор формирует напряжения рассогласования U_x и U_y, пропорциональные отклонению энергетического пятна по углу курса и тангажа соответственно. Выработанные координатные импульсы (сигналы U_x и U_y), связанные с системой координат МФП, поступают на формирователи сигналов поворота, осуществляющие формирование импульсов необходимой

длительности, обеспечивающей вращение роторов серводвигателей на необходимое число оборотов.

Структурная схема устройства, реализующего рассматриваемый способ сопровождения, представлена на рис. 1.

Для случая гауссовских априорных плотностей вероятности оцениваемых параметров на основе использования метода максимума апостериорной вероятности применительно к матричному приемнику $N \times M$ элементов получены следующие выражения для оценки смещения изображения цели по осям x и y [3]:

$$\hat{x}(t) = \frac{\Delta \sum_{i=1}^{N} i\xi_{\Sigma}(i\Delta, t) + \frac{m_x l_x^2}{4\Delta\sigma_x^2}}{\sum_{i=1}^{N} \xi_{\Sigma}(i\Delta, t) + \frac{l_x^2}{4\Delta\sigma_x^2}}; \quad \tilde{y}(t) = \frac{\Delta \sum_{j=1}^{M} j\xi_{\Sigma}(j\Delta, t) + \frac{m_y l_y^2}{4\Delta\sigma_y^2}}{\sum_{j=1}^{M} \xi_{\Sigma}(j\Delta, t) + \frac{l_y^2}{4\Delta\sigma_y^2}};$$

где l_x и l_y – размеры изображения цели по осям ox и oy соответственно; m_x, m_y и σ_x^2, σ_y^2 – математические ожидания и дисперсии гауссовских априорных плотностей вероятности оцениваемых параметров; Δ – шаг дискретизации по координатам.

Если за начало выбран центр матричного приемника (x_0 ; y_0), то величина отклонения энергетического центра изображения цели от выбранного отсчета будет определяться парой уравнений $U_x(t) = \hat{x}(t) - x_0$; $U_y(t) = y_0 - \hat{y}(t)$, где $x_0 = 0,5N + 0,5\Delta$; $y_0 = 0,5M + 0,5\Delta$.

Сформированное таким образом рассогласование, вызванное отклонением энергетического центра от выбранных отсчетных значений, уже с учетом знака полученной величины будет служить входным воздействием для системы сопровождения.

На рис. 2 приведена структурная схема модели оценки точности процесса сопровождения по одной координате x. Блоки МФИ, МОК и МСЦ – модули формирования изображения, оценки координат и сопровождения цели соответственно. Ошибка определения координат цели $\Delta x(t)$ формируется как разность между истинной координатой цели $x(t)$ и ее оценкой $\hat{x}(t)$ на выходе блока МОК. Динамическая ошибка сопровождения цели по каналу x – $\Delta U_x(t) = U_x(t)\left[1 - W_c(s)\right]$, где $W_c(s)$ – передаточная функция системы сопровождения (МСЦ) [4, 5]. Тогда суммарная динамическая ошибка системы $\Delta U_{\Sigma x}(t) = \Delta x(t) + \Delta U_x(t)$.

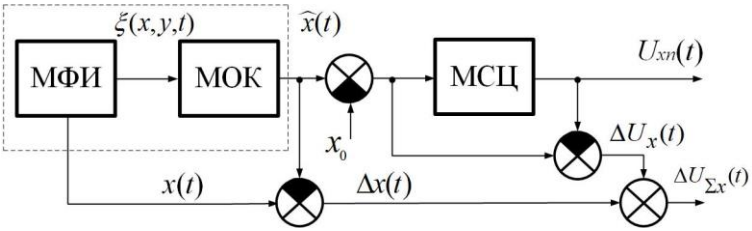


Рис. 2. Структурная схема модели оценки точности процесса сопровождения

3. Разработка компьютерной модели

Разработанная модель была реализована в системе MATLAB (Simulink) (рис. 3).

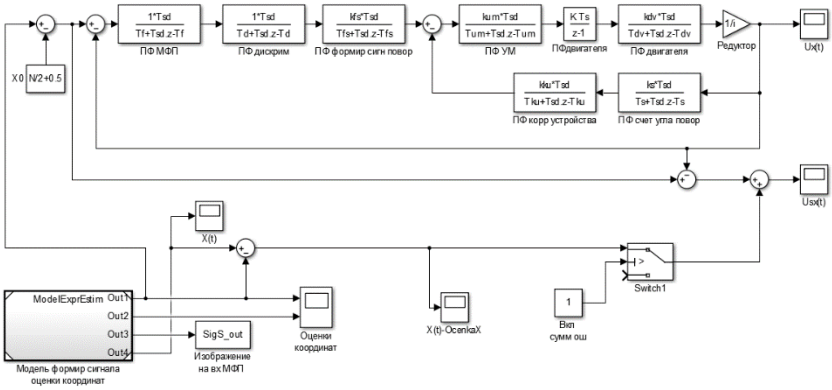
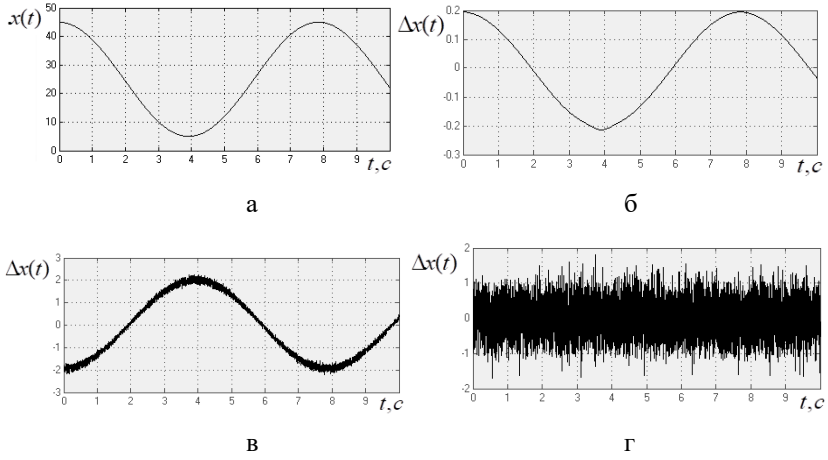


Рис. 3. Модель в системе MATLAB (Simulink)

С учетом необходимости покадровой оценки формируемого изображения и его анимации модель была реализована в дискретном времени. В модуле сопровождения цели использовались блоки Discrete Transfer Fcn и Discrete-Time Integrator. С использованием блока MATLAB System реализованы МФИ и МОК. Формирование изображения производилось в системе индексации элементов двумерного массива с последующим переводом в прямоугольную систему координат. Блоки MATLAB Function использовались для формирования траектории целей, а также изображений ВЦ, шума, фона и помехи. Фон моделировался как аддитивная смесь среднего уровня яркости неба, слабого (атмосферного) шума и низкочастотных пространственных структур (облаков). Внутренний шум МФП представлялся в виде гауссовского шума со

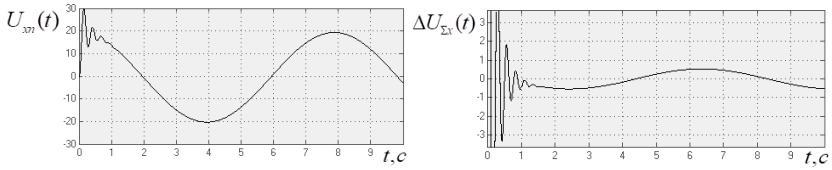
среднеквадратическим отклонением $\sigma_ш = 0,01$ В. Реализована возможность моделирования активной оптической помехи на одной из движущихся ВЦ, например, в виде 2D-гауссианы. С помощью блоков Const и Switch осуществляется управление конфигурацией сигнального тракта. Блок To Workspace в Simulink использовался для передачи сформированного изображения из модели Simulink в рабочее пространство MATLAB для последующих анимации и анализа. При помощи Scope результаты моделирования выводились в виде графиков функций координат, их оценок и ошибок от времени.

С использованием разработанной модели, задаваясь типовыми характеристиками системы сопровождения и траекторией полета цели по окружности, произведены расчеты для канала x, результаты которых представлены на рис. 4 и 5. Для удобства расчетов и восприятия графиков шаг дискретизации по координате Δ принимался за единицу, что обусловило представление полученных зависимостей в единицах измерения пиксель.

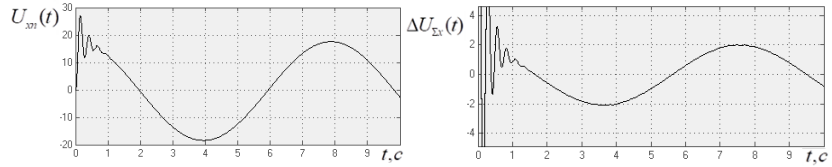


а – истинная траектория, б – одна цель, в – цель и шум, г – один шум

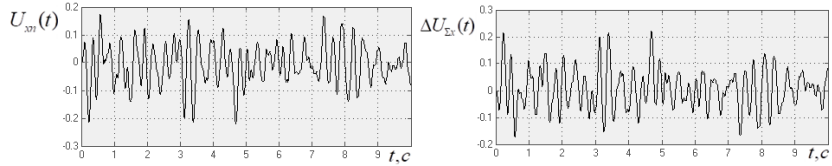
Рис. 4. Зависимости истинной координаты цели x и ошибок ее оценки от времени



а



б



в

а – одна цель, б – цель и шум, в – один шум

Рис. 5. Результаты моделирования процесса сопровождения

Для данного примера очевидно отсутствие влияния внутренних шумов на точность сопровождения. Были также проведены исследования для разно вариантных случаев фоноцелевой и помеховой обстановки: одной и двух целей с различными совместными маневренными траекториями полета (по прямой линии, по архимедовой спирали, по окружности с разными скоростями и т.д.), с применением средства постановки активных оптических помех на одном из движущихся объектов, при наличии фона и внутреннего шума. Выявлена возможность оценки вклада этих составляющих в эффективность сопровождения целей. Результаты анализа хода компьютерного эксперимента показывают адекватность разработанной модели изменяемым исходным данным и подтверждают возможность проведения экспресс-оценки точности сопровождения целей рассматриваемой телевизионной системой.

Заключение

Таким образом, в предположении применения для сопровождения воздушной цели телевизионной системы, использующей для оценки ее угловых координат алгоритм, основанный на определении энергетического центра принимаемого изображения, на основе разработанных в [3, 4] моделей и алгоритмов в системе MATLAB (Simulink) разработана компьютерная модель, позволяющая проводить экспресс-оценку влияния сложных помеховых и фоноцелевых условий, а также внутренних шумов на точность процесса сопровождения целей телевизионной системой.

В рамках исследований, выходящих за предметную область настоящей работы, авторами рассматриваются вопросы сравнительной оценки точностных характеристик рассмотренного алгоритма и алгоритма локализации [6] при сопровождении нескольких маневренных целей.

Список литературы

1. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление / Б. А. Алпатов [и др.] ; отв. ред. Б. А. Алпатов. – М. : Радиотехника, 2008. – 176 с.
2. Артемьев, В. М. Алгоритм сопровождения объектов в оптико-электронных системах на основе метода наименьших квадратов / В. М. Артемьев, А. О. Наумов, Л. Л. Кохан // Информатика. – 2013. – № 3. – С. 5-12.
3. Козирацкий, Ю. Л. Алгоритм совместной оценки положения и времени прихода пуассоновского изображения цели и его эффективность в условиях помех / Ю. Л. Козирацкий, Д. В. Прохоров, В. С. Казаков, А. В. Кусакин // Радиотехника. – 2012. – № 5. – С. 102–106.
4. Козирацкий, Ю. Л. Исследование алгоритма высокоточного сопровождения ЗУР матричным фотоприемником по блику формирующей оптики системы самонаведения / Ю. Л. Козирацкий, Д. В. Прохоров, И. Ю. Курьянов, Е. Д. Шатский // Журнал Сибирского федерального университета. Техника и технологии. – 2016. – Т. 9, № 3. – С. 401-415.
5. Коновалов, Г. Ф. Радиоавтоматика : учеб. для вузов по спец. «Радиотехника» / Г. Ф. Коновалов. – М. : Высш. шк., 1990. – 335 с.
6. Донцов, А. А. Алгоритм локализации малоразмерных объектов на изображении с использованием вейвлет-преобразований / А. А. Донцов, Ю. Л. Козирацкий // Автометрия. – 2013. – № 2. – С. 42-48.

Прогностическое моделирование уровня загрязнения приземного слоя атмосферного воздуха

Л. Н. Костылева, e-mail: kostyleva12@yandex.ru
М. С. Гордеев

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *На основе метеорологических наблюдений и данных о концентрациях загрязняющих веществ в воздушном бассейне построены прогностические регрессионные модели уровня загрязнения приземного слоя атмосферного воздуха. Проведена оценка качества прогностических моделей. Выявлены метеорологические параметры, оказывающие наибольшее влияние на уровень загрязнения воздушного бассейна промышленного города.*

Ключевые слова: *моделирование, прогнозирование, предикторы прогностической модели, оценка успешности, анализ модели, метеорологические параметры.*

Введение

Обеспечение экологической безопасности в настоящее время является одним из ключевых направлений государственной политики . Особенно остро стоит вопрос загрязнения воздушного бассейна в крупных промышленных центрах [2]. В связи с этим мониторинг окружающей среды и последующее прогностическое моделирование экологической ситуации становится еще более актуальной задачей.

Важную роль в формировании уровня загрязнения воздушного бассейна играют аэросиноптические условия и метеорологические параметры: температурные инверсии, направление и скорость ветра, такое опасное явление, как туманы и другие [1, 3, 6].

Исследования проводились на примере крупного промышленного города Центрального Черноземья – города Воронежа. Климат города – умеренно-континентальный, с неустойчивым увлажнением. Характерна высокая относительная влажность воздуха наиболее холодного месяца – 83 %, наиболее теплого месяца – 66 %. Годовое количество осадков – около 550-600 мм. В холодный период года часты туманы, снижающие рассеивающую способность атмосферы. В розе ветров в целом преобладают западные румбы (40 %), однако существенны и сезонные различия. Так, в летне-весенний период увеличивается повторяемость

© Костылева Л. Н., Гордеев М. С., 2026

юго-восточных ветров, а в летне-осенний – северных и северо-западных [3].

По данным А.В. Назаренко [3] частота инверсий в г. Воронеже составляет около 30-40 %, причем максимум приземных инверсий, способствующих аккумуляции загрязняющих веществ, приходится на май (64 %), август (66 %) – сентябрь (59 %) и февраль (55 %); минимум же – на январь (18 %), декабрь (19 %) и ноябрь (21 %).

1. Отбор предикторов для построения модели

Для построения прогностической регрессионной модели показателя загрязнения приземного слоя атмосферного воздуха была сформирована архивная выборка за период с 2020 по 2024 годы по метеорологическим параметрам и значениям концентраций приоритетных загрязняющих веществ.

Исследование включало несколько этапов:

1. Проведен анализ теоретических и методических основ экологического мониторинга.
2. Формирование базы данных исходных параметров с использованием программного пакета MsExcel.
3. Построение прогностических моделей загрязнения атмосферного воздуха.
4. Оценка параметров прогностических моделей.

Предиктантом является комплексный показатель загрязнения атмосферы ($K_{атм}$), рассчитанный по 4 загрязняющим веществам веществ (оксид углерода, диоксид серы, диоксид азота, пыль). Архивная выборка разделена на обучающую и контрольную в соотношении 2:1. Объем обучающей выборки составил 900 случаев, а контрольной – 300 случаев.

В качестве исходных метеорологических признаков рассматривались несколько предикторов, определяющих накопление и рассеивание антропогенных примесей в приземном слое атмосферы: относительная влажность, скорость ветра, температурная инверсия, температура воздуха, относительная влажность, атмосферное давление.

Отбор предикторов для прогностической регрессионной модели осуществлялся с помощью процедуры последовательного присоединения на основе анализа матрицы коэффициентов парной корреляции [4, 5].

Проведенная оценка корреляции показала, что f – относительная влажность, f – скорость ветра, I – температурная инверсия демонстрируют умеренную тесноту связи с показателем загрязнения воздушного бассейна. Остальные предикторы показывают слабую связь, поэтому их не целесообразно включать в модель.

2. Оценивание параметров прогностических моделей

Дальнейшая реализация метода последовательного присоединения заключалась в переборе моделей с различным количеством предикторов [4, 5]. На основе обучающей выборки были построены прогностические модели для показателя загрязнения воздушного бассейна.

На первом этапе последовательного присоединения построено три прогностических модели с одним предиктором (табл. 1).

Таблица

Этап 1

№	Уравнения регрессии	Коэффициент регрессии
1	$K_{атм} = 11,6 + 0,3 \cdot ff$	0,19
2	$K_{атм} = 13 - 0,1 \cdot f$	0,14
3	$K_{атм} = 11,9 - 0,9 \cdot I$	0,35

На втором этапе было построено две двухфакторные модели (табл. 2).

Таблица

Этап 2

№	Уравнения регрессии	Коэффициент регрессии
1	$K_{атм} = 11,8 - 0,95 \cdot I + 0.1 \cdot f$	0,28
2	$K_{атм} = 10,2 - 0,88 \cdot I + 0.33 \cdot ff$	0,42

На третьем этапе последовательного присоединения была построена трехфакторная прогностическая модель (табл. 3).

Таблица

Этап 3

№	Уравнения регрессии	Коэффициент регрессии
1	$K_{атм} = 6,95 - 0,79 \cdot I + 0.41 \cdot ff - 0.25 \cdot f$	0,42

С помощью показателя качества коэффициента корреляции оцениваются регрессионные модели.

На основе данных контрольной выборки были рассчитаны основные показатели успешности прогностических моделей: средняя абсолютная ошибка (САО) и средняя квадратическая ошибка (СКО) [5].

Полученные результаты представлены в табл. 4.

Таблица

Результаты оценок показателей качества по контрольной выборке

Оценка показателей качества	$K_{атм}1$	$K_{атм}2$	$K_{атм}3$
САО	1,74	1,68	1,82
СКО	2,47	2,34	2,65

Проверка по контрольной выборке подтверждает расчеты, выполненные по обучающей выборке. По результатам исследования следует заметить, что третье уравнение на первом этапе имеет наименьшие ошибки по сравнению с другими уравнениями, а коэффициент корреляции мало отличается от лучшего уравнения с двумя предикторами на последующем этапе.

3. Оценка значимости прогностической регрессионной модели скорости ветра в целом

Статистическая значимость прогностической регрессионной модели показателя загрязнения воздушного бассейна проверялась на основе следующих гипотез:

$H_0 : R^2 = 0$ – построенное уравнение в целом незначимо;

$H_1 : R^2 \neq 0$ – построенное уравнение в целом значимо.

Для проверки статистической значимости построенной модели регрессии использовался критерий Фишера. Табличное значение критерия Фишера для уровня значимости в 1 % составило $F_{табл} = 3,78$. Расчетное значение критерия Фишера $F_{расч} = 224,8$.

Так как $F_{расч} > F_{табл}$ на уровне значимости (1 %) принимается гипотеза H_1. Построенная прогностическая модель в целом значима, следовательно получено надежное математическое описание прогноза уровня загрязнения приземного слоя атмосферного воздуха.

Заключение

Таким образом, в процессе выполнения исследования решены все поставленные задачи:

– Проведен анализ теоретических и методических основ экологического мониторинга.

– Сформирована архивная выборка по метеорологическим факторам и параметрам загрязнения воздушного бассейна.

– Построены прогностические модели зависимости уровня загрязнения атмосферного воздуха от метеопараметров.

– Проведена оценка параметров прогностических моделей.

Результаты исследования показали, что на уровень загрязнения приземного слоя атмосферы наибольшее влияние оказывают температурная инверсия и скорость ветра.

Результаты проведенных исследований могут быть использованы для обоснования природоохранных рекомендаций по обеспечению экологической безопасности и снижению негативного воздействия на техногенно нагруженной территории.

Список литературы

1. Блейк Д. Физические основы динамики атмосферы и метеорологии / Д. Блейк, Р. Робсон. – Долгопрудный: ИД «Интеллект», 2016. – 160 с.

2. Безуглая Э.Ю. Чем дышит промышленный город / Э.Ю. Безуглая, Г.Н. Расторгуева, И.В. Смирнова. – Л. : Гидрометеоиздат, 1991. – 255 с.

3. Назаренко А.В. Исследование уровня загрязнения атмосферы в г. Воронеже / А.В. Назаренко – Воронеж: ВВАИИ, 2003. – 150 с.

4. Горяинова Е.Р. Прикладные методы анализа статистических данных / Е.Р. Горяинова, А.Р. Панков, Е.Н. Платонов. – М. : Изд. Дом Высшей школы экономики, 2012. – 310 с.

5. Ульшин, И. И. Методы статистической обработки гидрометеорологической информации / И. И. Ульшин. – Воронеж ВУНЦ ВВС «ВВА», 2016. – 178 с.

6. Хромов С.П. Метеорология и климатология / С.П. Хромов, М.А. Петросянц. – М. : Издательство МГУ, 2004. – 584 с.

Иммерсивное моделирование процедурного обучения на базе Unity: тенденции и практическое применение

И.С. Лакунов, e-mail: nauchnajarota@yandex.ru[1]
В. В. Алексеев, e-mail: vvalex1961@mail.ru[2]
Е.С. Ионченко, e-mail: nauchnajarota@yandex.ru[1]

[1]Межвидовой центр подготовки и боевого применения войск РЭБ (учебный и испытательный)
[2]Тамбовский государственный технический университет

Аннотация. *В статье рассматривается роль игрового движка Unity в создании иммерсивных симуляционных сред для процедурного обучения. Проанализированы кейсы моделирования сборки инженерных конструкций. Выявлены тенденции интеграции адаптивных систем на базе Unity, обеспечивающих персонализацию учебного процесса. Обзор демонстрирует эффективность иммерсивного моделирования для освоения сложных пошаговых навыков без необходимости физического оборудования.*

Ключевые слова: *иммерсивное моделирование, Unity, процедурное обучение, симуляционные среды, адаптивные системы обучения, инженерное моделирование.*

Введение

Современное профессиональное образование сталкивается с вызовами эффективного обучения сложным процедурным навыкам, требующим высокой точности и безопасности выполнения. Традиционные методы подготовки ограничены доступностью оборудования и рисками для обучаемых. Иммерсивные технологии на базе игровых движков предлагают решение через создание виртуальных симуляций, полностью воспроизводящих реальные процессы [1].

Цель статьи – проанализировать прикладное применение Unity для моделирования пошаговых задач в инженерной практике, провести литературный обзор существующих решений и выявить перспективы адаптивных систем обучения.

© Лакунов И.С., Алексеев В.В., Ионченко Е.С., 2026

1. Теоретические основы иммерсивного моделирования в образовании. Понятие процедурного обучения и его специфика

Процедурное обучение представляет собой тип приобретения навыков, ориентированный на освоение последовательных действий, выполняемых в строгой последовательности для достижения конкретного результата. В отличие от декларативных знаний, где акцент делается на запоминании фактов, процедурное обучение фокусируется на формировании моторных и когнитивных автоматизмов через многократное повторение алгоритмов. Специфика этого подхода проявляется в необходимости имитации реальных условий: малейшее отклонение от протокола может привести к критическим ошибкам в инженерных задачах [2].

Ключевой особенностью процедурного обучения является его иерархическая структура: от базовых манипуляций к комплексным сценариям с ветвлением. Согласно исследованиям в области когнитивной психологии, эффективность такого обучения возрастает при использовании иммерсивных сред, где обучаемый получает немедленную обратную связь и может корректировать действия в реальном времени [2]. Это особенно актуально для профессиональной подготовки, где традиционные методы ограничены материальными ресурсами и рисками.

В контексте информатики процедурное обучение моделируется как дискретные состояния переходов, где каждый этап представлен виртуальным объектом с заданными параметрами взаимодействия, что позволяет масштабировать сложность без физических затрат.

2. Иммерсивные технологии как инструмент моделирования реальных процессов

Иммерсивные технологии представляют собой комплекс программно-аппаратных средств, обеспечивающих полное погружение пользователя в виртуальную среду с высокой степенью реализма. Основу составляют стереоскопическая визуализация, пространственный звук и отслеживание движений, создающие эффект присутствия, близкий к реальным условиям [3]. В образовании эти технологии моделируют процессы, недоступные для прямого наблюдения или воспроизведения, минимизируя риски и затраты.

Применение иммерсивных сред особенно эффективно для отработки процедурных навыков, где требуется координация зрения, моторики и принятия решений. Например, в инженерных задачах моделируются физические взаимодействия материалов.

В рамках прикладного моделирования иммерсивные технологии интегрируют данные из реальных процессов –CAD-модели,

биомеханику, сенсорные протоколы –формируя цифровые двойники. Это позволяет проводить эксперименты без физических прототипов, ускоряя освоение сложных последовательностей и повышая переносимость навыков в профессиональную практику.

3. Роль игровых движков в создании образовательных симуляций

Игровые движки, такие как Unity и Unreal Engine, трансформировались из инструментов развлечений в универсальные платформы прикладного моделирования благодаря поддержке физических симуляций, 3D-рендеринга и сетевой синхронизации. Их архитектура позволяет создавать интерактивные среды с нелинейным повествованием, где обучаемый самостоятельно исследует сценарии, получая адаптивную обратную связь. В образовании это обеспечивает моделирование процедур с ветвящейся логикой, недостижимой в статических тренажерах.

Ключевые преимущества игровых движков –кроссплатформенность и доступность разработки, снижающие порог входа для создания сложных симуляций без специализированных знаний в графике или физике. Unity, в частности, предоставляет asset store с готовыми модулями биомеханики и инженерных компонентов, ускоряя прототипирование. По данным отраслевых обзоров, 68% образовательных VR-приложений разработаны на Unity, что подтверждает его доминирование в иммерсивном обучении[3].

В контексте процедурного обучения игровые движки реализуют замкнутый цикл: восприятие → действие → оценка → коррекция, интегрируя данные с периферийных устройств (VR-контроллеры, хаптические перчатки). Это создает цифровые полигоны для безопасной отработки навыков, повышая компетентность на 40–60% по метрикам времени выполнения и точности, согласно мета-анализам EdTech-исследований.

4. Unity как платформа прикладного моделирования. Архитектурные особенности Unity для иммерсивных сред

Unity представляет собой кроссплатформенный игровой движок с модульной архитектурой, оптимизированной для создания высокореалистичных иммерсивных симуляций. Центральным элементом служит компонентная система GameObject, где каждый объект окружающей среды (инструменты, пациенты, механизмы) представляет собой набор взаимодействующих модулей – Transform для позиционирования, Collider для обнаружения столкновений, Rigidbody для физической динамики. Эта структура обеспечивает реалистичное

моделирование процедурных взаимодействий без необходимости низкоуровневого программирования.

Особое значение имеют встроенные системы XR (Extended Reality), поддерживающие VR/AR-гарнитуры с нативной интеграцией отслеживания рук и глаз, что критично для точной симуляции манипуляций и инженерных задачах. Unity's Universal Render Pipeline (URP) обеспечивает высококачественную визуализацию с оптимизацией под различные устройства, включая мобильные VR-системы, сохраняя 90 кадров/с для предотвращения укачивания. Модуль Timeline позволяет хореографировать последовательные этапы процедур, создавая ветвящиеся сценарии с триггерной логикой.

Архитектура Unity масштабируется через Addressables и DOTS (Data-Oriented Technology Stack), что позволяет загружать сложные сцены динамически и обрабатывать тысячи взаимодействующих объектов одновременно. Эта гибкость делает Unity оптимальным для прикладного моделирования, где требуется баланс между реализмом и производительностью.

5. Возможности симуляции физических и последовательных процессов

Unity предоставляет мощный PhysX-движок для реалистичного моделирования физических взаимодействий, необходимых в процедурном обучении. Система Configurable Joints позволяют настраивать ограничения движения инструментов, имитируя реальные физические свойства (жесткость пружин, демпфирование)[6].

Для последовательных процессов ключевыми являются State Machines и Behavior Trees, реализующие алгоритмы процедур как конечные автоматы с переходами между этапами. Animator Controller синхронизирует анимации с физикой, создавая плавные переходы от подготовки к выполнению и завершению задачи. Playables API и Timeline организуют нелинейные сценарии с ветвлением: правильное действие открывает следующий этап, ошибка активирует корректирующий блок с подсказками[6].

Интеграция Particle System моделирует второстепенные эффекты – искры сварки, дым – повышая сенсорное погружение. Shader Graph позволяет кастомизировать визуализацию материалов (металл, кожа, пластик) для аутентичного восприятия. Эти возможности обеспечивают замкнутый цикл обучения: демонстрация → практика → валидация, где каждый шаг процедуры подкреплен мгновенной физически корректной реакцией среды.

Преимущества Unity перед традиционными инструментами моделирования

Unity превосходит традиционные инструменты моделирования (MATLAB Simulink, ANSYS, LabVIEW) за счет полной интерактивности и доступности для создания иммерсивных симуляций. В отличие от статических CAE-систем, ориентированных на анализ, Unity обеспечивает реальное время взаимодействия с виртуальной средой, где обучаемый активно манипулирует объектами, получая тактильную и визуальную обратную связь через VR/AR-устройства.

Ключевым преимуществом является экосистема библиотеки готовых компонентов Unity Asset Store, содержащая тысячи проверенных модулей, что сокращает время разработки с нескольких месяцев до нескольких дней. Кроссплатформенная публикация (ПК, VR-гарнитуры, мобильные устройства, веб-браузеры) обеспечивает широкий доступ к симуляциям без зависимости от специализированного оборудования. Встроенные аналитические инструменты фиксируют метрики обучения (время выполнения этапов, количество ошибок, динамику прогресса), автоматически формируя данные для адаптации сложности сценариев.

Экономическая эффективность Unity подтверждается отраслевыми кейсами: разработка VR-симулятора обходится в 5–10 раз дешевле физического тренажера, с ROI в 300% за счет многократного использования. Открытый C# API и визуальные инструменты (Bolt, Playmaker) позволяют непрограммистам создавать сложные сценарии, расширяя круг разработчиков образовательного контента. В итоге Unity трансформирует прикладное моделирование из аналитического инструмента в полноценную платформу процедурного обучения с фокусом на практическую компетентность.

6. Моделирование инженерных конструкций в Unity. Симуляция сборки и разборки механизмов

Unity эффективно моделирует процессы сборки и разборки инженерных конструкций через интеграцию физических компонентов и интерактивных сценариев, обеспечивая пошаговую отработку последовательностей. Обучаемый взаимодействует с 3D-моделями реальных механизмов – от автомобильных двигателей до роботизированных манипуляторов – используя VR-контроллеры для захвата, перемещения и фиксации деталей с учетом гравитации, трения и геометрических ограничений. Система Articulation Bodies точно воспроизводит многосуставчатую кинематику, где каждая деталь обладает реалистичными степенями свободы и соединяется через Configurable Joints, имитируя болты, шплинты или клиновые соединения.

Ключевой особенностью является пошаговый контроль доступа: на начальном этапе доступны только базовые компоненты, последующие этапы открываются после успешной фиксации предыдущих, реализуя

принцип обучения от простого к сложному. Система Timeline и Animator обеспечивает демонстрацию правильной последовательности, за которой следует свободная практика с автоматической проверкой[6]. Неправильная последовательность действий блокирует прогресс, активируя интерактивные подсказки или перезапуск этапа. Взрывные виды (Exploded Views) и режимы полупрозрачной визуализации раскрывают внутренние структуры механизмов, обеспечивая понимание пространственных взаимосвязей без физического демонтажа.

Такие симуляции применяются в авиакосмической отрасли для сборки турбин и в автомобилестроении для диагностики трансмиссий, сокращая время обучения на 45% и повышая точность на 30%, по данным корпоративных кейсов Siemens и Boeing. Модульная архитектура позволяет адаптировать модели под конкретное оборудование, создавая персонализированные тренажеры для заводских инструкторов без перестройки всей системы.

Интеграция физических моделей для реалистичной обратной связи

Unity интегрирует высокоточные физические модели через PhysX и расширенные библиотеки, обеспечивая реалистичную тактильную и визуальную обратную связь при моделировании инженерных процессов. Rigidbody и Collider системы точно рассчитывают столкновения, деформации и силы реакции, имитируя поведение металлических, пластиковых или композитных материалов под нагрузкой – от упругости пружин до вибраций роторов. Хаптическая обратная связь через VR-устройства (Haptic Feedback) передает вибрации, сопротивление и клики фиксации, усиливая моторное запоминание: обучаемый физически ощущает, когда болт "сядает" на место или деталь "не входит" из-за неправильного выравнивания.

Ключевым элементом служит Material Physics с настройкой коэффициентов трения, упругости и массы, позволяющая воспроизвести реальные сценарии –смазку поверхностей, температурное расширение или износ соединений. Constraint Solvers в Joints гарантируют кинематическую точность многоуровневых сборок, где ошибка на одном этапе каскадно влияет на последующие, формируя естественные последствия. Raycasting и Trigger Zones детектируют микронарушения протокола (например, касание неподготовленной поверхности), активируя визуальные и звуковые индикаторы ошибок с замедленной съемкой для анализа.

Такая интеграция применяется в нефтегазовой отрасли для симуляции разборки насосов и в микроэлектронике для пайки плат, где реалистичная физика сокращает количество реальных инцидентов на 35% после тренировок. Модели импортируются из CAD-систем

(SolidWorks, AutoCAD) с сохранением параметров, обеспечивая нулевую расхождение между виртуальной и физической реальностью для профессиональной сертификации.

7. Примеры из машиностроения и электроники

В машиностроении Unity применяется для симуляции сборки тяжелой техники –от гидравлических прессов до конвейерных линий, где обучаемые осваивают последовательности с учетом веса, балансировки и безопасности. Кейс Caterpillar демонстрирует тренажер для сборки экскаваторов: виртуальные гидроцилиндры реагируют на давление с физической отдачей, а ошибки в последовательности (неправильный монтаж поршня) вызывают управляемый сбой с анализом причин. Это сокращает время сертификации операторов на 50%, минимизируя риски производственных травм.

В электронике Unity моделирует пайку и диагностику печатных плат, интегрируя тепловые модели припоя и электромагнитные эффекты. Обучаемый использует виртуальный паяльник с хаптикой для наложения компонентов SMD-элементов, где перегрев вызывает визуальное повреждение дорожек и звуковой сигнал тревоги. Кейс Bosch для сборки автомобильной электроники показывает адаптивные уровни: от базовой пайки к диагностике неисправностей с мультиметром в AR-режиме, повышая точность монтажа на 40% у новичков.

Эти примеры иллюстрируют переносимость навыков: статистика по внедрениям в Ford и Samsung подтверждает снижение брака на 25–35% после VR-тренировок. Модульность Unity позволяет обновлять модели под новые ревизии оборудования, обеспечивая непрерывное обучение без остановки производства.

Заключение

Проведенный литературный обзор подтверждает высокую эффективность иммерсивного моделирования на базе Unity для процедурного обучения в инженерной сфере. Анализ архитектурных возможностей платформы, физических симуляций и реальных кейсов демонстрирует сокращение времени подготовки на 40–50%, повышение точности выполнения на 30–40% и экономию материальных ресурсов за счет замены физических тренажеров виртуальными аналогами. Ключевым преимуществом Unity выступает баланс между технологической сложностью и доступностью разработки, что делает платформу оптимальным инструментом прикладного моделирования.

Иммерсивные симуляции обеспечивают безопасную отработку критически важных навыков с мгновенной обратной связью, формируя устойчивые моторные автоматизмы, переносимые в профессиональную

практику. Перспективы развития включают интеграцию ИИ для динамической персонализации сложности сценариев и мультимодальную аналитику прогресса обучаемых.

Полученные выводы целесообразно использовать при разработке корпоративных программ подготовки кадров, а также в вузовских курсах профессиональной переподготовки. Дальнейшие исследования должны сосредоточиться на лонгитюдных эффектах навыков и кросс-культурной адаптации симуляций для глобального образования.

Список литературы

1. Стёпушкина, П. К. Игровые движки для образовательных целей / П. К. Стёпушкина // Вестник науки. – 2024. – № 5 (74). – Т. 4. – С. 4-6.

2. Ананин, Д. П. Иммерсивные технологии в образовательной практике российской высшей школы / Д. П. Ананин, А. Ю. Сувирова // Высшее образование в России. – 2024. – Т. 33, № 5. – С. 112–135. – DOI: 10.31992/0869-3617-2024-33-5-112-135.

3. Котов, Г. С. [Название диссертации отсутствует] : дис. канд. пед. наук : 378.937:681.14 / Г. С. Котов ; ФГАОУ ВО «Южный федеральный университет». – Ростов-на-Дону, 2026.

4. Рововой, Э. Ю. Создание AR-приложения на базе Unity 3D / Э. Ю. Рововой, М. В. Иванов, И. А. Киселев // XII Международная научно-практическая конференция : материалы. – Санкт-Петербург : Санкт-Петербургский политехнический университет Петра Великого, 2026. – С. 48.

5. Костина, Ю. Н. Архитектура Unity-проекта / Ю. Н. Костина // 75-я научно-техническая конференция учащихся, студентов и магистрантов : тезисы докладов, 22–27 апреля 2024 г., Минск : в 4 ч. Ч. 4. – Минск : БГТУ, 2024. – С. 44.

6. Ульянов, Р. С. Моделирование технических систем в среде Unity 3D / Р. С. Ульянов, С. В. Прокопьев, В. В. Делибалтов // Молодой ученый. – 2015. – № 11 (91).

Модели тепловых потоков в тропических волнах

Э. Н. Мартинез, e-mail: eleazarmartinez817@gmail.com
В. С. Ножкин, e-mail: nozhkin-v@list.ru
Г. Е. Шербаков

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *Особое внимание в работе уделено анализу фазы интенсивного усиления тропических волн, предшествующей их трансформации в ураганы. Ключевой объект исследования – эволюция структуры и величины вертикальных тепловых потоков на протяжении всего жизненного цикла волны вплоть до её качественного перехода в новый класс атмосферных явлений – тропический циклон.*

Ключевые слова: *тропические волны, адвекция температуры, меридиональная адвекция, зональная адвекция, волновые возмущения.*

Введение

Тропические волновые возмущения играют ключевую роль в переносе энергии и влаги в низких широтах, оказывая существенное влияние на конвективную активность и формирование тропических циклонов [1-7]. Анализ адвективных процессов в этих системах позволяет лучше понять их динамику и эволюцию. В данной работе для количественной оценки адвективных процессов использованы компоненты Q_1 и Q_2, представляющие адвекцию температуры, связанную с градиентами ветра. В рамках настоящего исследования проведен сравнительный анализ характеристик тропических волн, выступивших затравкой для формирования ураганов Эта (Eta) и Йота (Iota) в 2020 году.

Тропические волны, характеризующиеся высокими значениями меридиональной диссипационной компоненты Q_2 и соотношением $Q_2/|Q_1| > 2$, обладают повышенным потенциалом для развития в тропические циклоны [8]. Анализ диссипативных процессов предоставляет ценные инсайды для раннего выявления опасных метеорологических явлений и улучшения их прогноза, так как 60% всех тропических циклонов Атлантики развиваются из тропических волн, представлены в таблице.

Таблица

Пространственно-временные характеристики тропической волны (Salinas 2006)

Место наблюдения	Период (Дни)	Длина волны (КМ)	Фазовая скорость (Км/день)	Наиболее энергетический уровень (Гпа)	Меридиональная протяженность (км)	Центральная широта направляющей волны	Максимальная амплитуда (м/с)
Африка	3.7	3000	660	850	2200	12.5°	1.7
Карибский Бассейн	5	2660	770	700	1100	15°	1

1. Данные и методы исследования

Использованы данные реанализа за период 24.10.2020 – 31.11.2020 и 8.11.2020 – 13.11.2020 временным разрешением 6-24 часа. Пространственное разрешение: по широте 2.5° (диапазон 5°-20° с.ш.), по долготе 2-5° (диапазон 30°-80° в.д.). Анализировались стандартные изобарические уровни от 1000 до 200 гПа с шагом 100 гПа в нижней тропосфере и 50-100 гПа в средней и верхней тропосфере.

На рис. 1 показано область исследования развития тропических волновых возмущений.

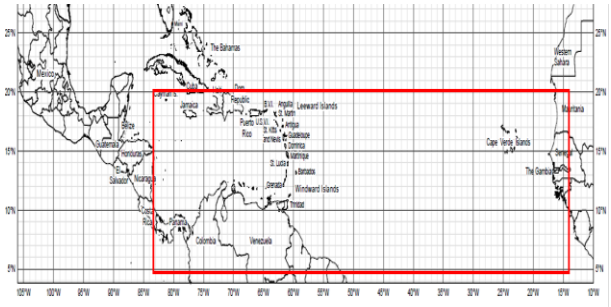


Рис. 1. Область исследования развития тропических волновых возмущений

2. Методика расчета

Ключевой вопрос о связи между конвекцией и экваториальными волнами заключается в том, возникают ли они как сухие возмущения, которые затем связываются с конвекцией, или же конвекция инициирует циркуляцию, которая затем проецируется на предпочтительный режим атмосферной циркуляции. В случае восточно-африканских волн (AEWs) мы можем диагностировать динамический (или адиабатический) вклад в вынуждение, обусловленный адвекцией завихренности и температуры, с использованием Q-векторов, что пренебрегает диабатическим членом вынуждения в уравнении омега. Q-векторы полезны в районе Африканского восточного струйного течения (AEJ), несмотря на то, что они возникают на широте, где исходные квазигеострофические допущения могут быть не совсем справедливы

Расчёты выполнены по данным реанализа на сетке 2.5°×2-5° для стандартных изобарических уровней от 1000 до 200 гПа. Мы используем слегка модифицированную версию квазигеострофического Q-вектора.

$$Q_1 \equiv -\frac{R}{P}\frac{\partial V_\psi}{\partial x}\nabla\cdot T = -\frac{R}{P}\left(\frac{\partial u_\psi}{\partial x}\frac{\partial T}{\partial x}+\frac{\partial v_\psi}{\partial x}\frac{\partial T}{\partial y}\right),$$

$$Q_2 \equiv -\frac{R}{P}\frac{\partial V_\psi}{\partial y}\nabla\cdot T = -\frac{R}{P}\left(\frac{\partial u_\psi}{\partial y}\frac{\partial T}{\partial x}+\frac{\partial v_\psi}{\partial y}\frac{\partial T}{\partial y}\right),$$

где R = 287 Дж/(кг·К) – газовая постоянная для сухого воздуха; p – давление (Па); u, v – зональная и меридиональная составляющие ветра (м/с); T – температура (K).

На рис. 2 – 3 приведены результаты анализа.

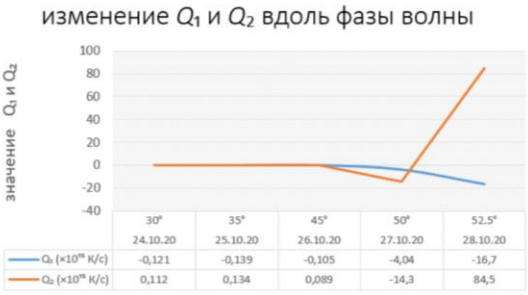
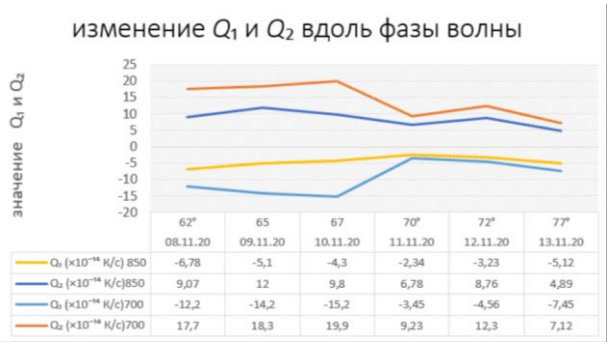


Рис. 2. Анализ тропической волны при движении на запад с 24-28 октября 2020 года



	62° 08.11.20	65 09.11.20	67 10.11.20	70° 11.11.20	72° 12.11.20	77° 13.11.20
Q_1 (×10⁻¹⁴ К/с) 850	-6,78	-5,1	-4,3	-2,34	-3,23	-5,12
Q_2 (×10⁻¹⁴ К/с)850	9,07	12	9,8	6,78	8,76	4,89
Q_1 (×10⁻¹⁴ К/с)700	-12,2	-14,2	-15,2	-3,45	-4,56	-7,45
Q_2 (×10⁻¹⁴ К/с)700	17,7	18,3	19,9	9,23	12,3	7,12

Рис. 3. Анализ тропической волны при движении на запад с 08-13 ноября 2020 года

Тропическая волна, наблюдаемая 24-28 октября 2020 года, претерпела качественное изменение при движении на запад через Атлантику. В восточной части (30°- 45°W) диссипация была умеренной и сбалансированной ($Q_1 \approx Q_2$). При пересечении 50°W произошел экспоненциальный рост диссипации, особенно меридиональной компоненты Q_2, которая увеличилась почти в 1000 раз к 52.5°W. Это указывает на интенсивное развитие волнового возмущения в западной Атлантике, возможно связанное с взаимодействием с тёплыми водами и орографическим влиянием Южной Америки.

Тропическая волна, наблюдавшаяся в ноябре 2020 года над Атлантикой, характеризовалась значительной диссипацией с преобладанием меридиональной компоненты (Q_2) над зональной (Q_1). Диссипация усиливалась с высотой и достигала максимума в центральных долготах (67°–70°). Полученные результаты подчёркивают важность учёта вертикальной и долготной неоднородности диссипативных процессов в моделях прогноза тропической погоды.

Заключение

1. Расчёты Q_1 и Q_2 позволяют эффективно отслеживать эволюцию тропических волновых возмущений.

2. Меридиональная компонента адвекции (Q_2) играет доминирующую роль в развитии тропических волн, превышая зональную компоненту (Q_1) в 1.2-3.8 раза.

3. Выявлены две последовательные волны с пиками 25-28 октября и 12 ноября, перемещавшиеся на восток со скоростью 5-10°/сут, что соответствует типичным тропическим волнам.

4. Для более детального анализа необходимы данные с более высоким пространственным разрешением и учёт влажно-адиабатических процессов через использование эквивалентно-потенциальной температуры.

Список литературы

1. Hsieh, J.-S. On the instability of the African easterly jet and the generation of African waves: reversals of the potential vorticity gradient / J.-S. Hsieh, K. H. Cook // Journal of the Atmospheric Sciences. – 2008. – Vol. 65, No. 7. – P. 2130–2151. – DOI: 10.1175/2007JAS2552.

2. Holton, J. R. An introduction to dynamic meteorology / J. R. Holton, G. J. Hakim. –5th ed. –Amsterdam : Academic Press, 2013.

3. Kiladis, G. N. Tropical waves and convection / G. N. Kiladis, M. C. Wheeler // Journal of the Atmospheric Sciences. – 1995. – Vol. 52, No. 3. – P. 374-399.

4. Madden, R. A. Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific / R. A. Madden, P. R. Julian // Journal of the Atmospheric Sciences. – 1971. – Vol. 28, No. 5. – P. 702-708.

5. Molinari, J. External influences on tropical cyclone formation in the western North Pacific / J. Molinari, D. Vollaro // Journal of the Atmospheric Sciences. – 2010. – Vol. 67, No. 3 – P. 976–991.

6. Salinas Prieto, J. A. Dinámica de ondas del este y su interacción con el flujo medio en el Caribe : dis. … Dr. (PhD) / J. A. Salinas Prieto ; Universidad Nacional Autónoma de México. –México, 2006.

7. Wallace, J. M. Atmospheric science: an introductory survey / J. M. Wallace, P. V. Hobbs – 2nd ed. – Amsterdam : Academic Press, 2006.

8. Yanai, M. Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets / M. Yanai, S. Esbensen, J. H. Chu // Journal of the Atmospheric Sciences. – 1973 – Vol. 30, No. 4. – P. 611-627.

Анализ перспектив применения фторполимерных покрытий в противообледенительных системах малоразмерных беспилотных летательных аппаратов

А. В. Мельников, e-mail: alexei.melnikov90@yandex.ru
С. В. Назарчук, e-mail: qazter2015@gmail.com
Д. Н. Кузнецов, e-mail: vrnh@bk.ru

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *В данной работе проводится анализ и обоснование применения фторопластов в качестве материала для изготовления покрытий, обеспечивающих надежную энергоэффективную защиту поверхностей беспилотных летательных аппаратов от обледенения.*

Ключевые слова: *беспилотный летательный аппарат, БпЛА, противообледенительная система, фторопластовое покрытие.*

Введение

В настоящее время беспилотные летательные аппараты (БпЛА) малой дальности не оборудованы штатными противообледенительными системами (ПОС). В связи с этим полеты БпЛА данного класса не производятся, если заранее известно, что на маршруте полета расположена зона облачности с высоким содержанием влаги и температурой воздуха от 0 до -10 градусов Цельсия. В таких условиях происходит интенсивное обледенение передней кромки крыла, в результате чего может образоваться критическая масса льда, что приведет к ухудшению аэродинамических характеристик летательного аппарата. В частности, снижение подъемной силы и увеличение лобового сопротивления из-за нарастания льда на поверхности крыла может привести к крушению БпЛА.

Проведенный анализ применяемых в пилотируемой авиации способов защиты от обледенения [1, 2] показал, что существующие противообледенительные системы (ПОС) невозможно использовать для обеспечения функционирования малоразмерных БпЛА в условиях обледенения, ввиду их ограниченной взлетной массы, низкого запаса электроэнергии и мощности двигателя. В работе [3] предложен тепловой способ защиты от обледенения крыла малоразмерного БпЛА на основе энергии, выделяемой при каталитическом горении топлива, а также сформулированы задачи дальнейшего исследования по применению предложенного теплового способа защиты.

© Мельников А.В., Назарчук С.В., Кузнецов Д.Н., 2026

Однако, учитывая выявленные ограничения существующих способов защиты от обледенения, и недостаток теплового способа защиты, заключающийся в необходимости интеграции сложных конструкций, что повышает стоимость и сложность эксплуатации, наиболее перспективным направлением для малоразмерных БпЛА представляется применение противообледенительных покрытий. Такой подход не требует затрат электрической и тепловой энергии и значительного вмешательства в конструкцию крыла БпЛА. В качестве основного материала для таких покрытий предлагается рассмотреть фторопласты.

1. Анализ основных видов фторопластов

Фторопласты (фторполимеры) представляют собой группу высокомолекулярных соединений, характеризующихся частичным или полным замещением атомов водорода в молекулярной цепи на атомы фтора. Это замещение определяет комплекс эксплуатационных характеристик, доказывающих эффективность применения фторопластов в качестве основы для противообледенительных покрытий. Однако, несмотря на общность свойств, присущих фторопластам, их характеристики существенно варьируются в зависимости от химической структуры, в частности степени фторирования и состава полимерной цепи, что обуславливает их разделение на отдельные классы и марки [4].

С химической точки зрения, фторопласты подразделяются на три основные категории [4]:

– перфторированные полимеры (политетрафторэтилен, сополимер тетрафторэтилена, гексафторпропилен), в молекулах которых все атомы водорода замещены на фтор;

– частично фторированные полимеры (поливинилденфторид, политрифторхлорэтилен), содержащие в своей структуре наряду с атомами фтора атомы водорода или хлора;

– фторсодержащие каучуки, обладающие эластомерными свойствами.

Настоящий анализ фокусируется на представителях первых двух категорий, как наиболее перспективных для создания противообледенительных покрытий.

В настоящее время промышленность выпускает большое число разновидностей фторопласта, однако к числу основных относятся три его марки: поливинилиденфторид (Ф-2) – $\left(CH_2CF_2\right)_n$, политрифторхлоридэтилен (Ф-3) – $\left(-CF_2-CFCl\right)_n$, политетрафторэтилен (Ф-4) – $\left(-C_2F_4-\right)_n$. [5].

Поливинилиденфторид характеризуется оптимальным балансом эксплуатационных свойств. Материал демонстрирует высокую механическую прочность и стойкость к абразивному износу при умеренных антиадгезионных характеристиках. Отличительной особенностью является возможность формирования покрытий методом распыления растворов, что обеспечивает относительную простоту в технологии нанесения. Этот прочный упругий полимер, демонстрирует химическую нейтральность к подавляющему большинству веществ, однако растворяется в некоторых растворителях, что нужно учитывать.

Политрифторхлоридэтилен является сополимером винилиденфторида с тетрафторэтиленом. Благодаря повышенному содержанию фтора он демонстрирует улучшенные антиадгезионные свойства по сравнению с Ф-2. Однако это достигается за счет снижения механической прочности и адгезии к подложкам. Фторопласту указанной марки присущи прочность и твёрдость. При нагревании до высоких температур наблюдается его размягчение, изменение формы.

Политетрафторэтилен как полностью фторированный полимер обладает эталонными антиадгезионными характеристиками и химической стойкостью. В структуре этого гидрофобного пластика практически отсутствуют поры. Однако плавкость и крайне низкая адгезия к подложкам существенно ограничивают его применение. Формирование покрытий требует сложных технологических процессов, а низкая износостойкость делает покрытия на основе Ф-4 недолговечными в условиях абразивного воздействия.

Технико-эксплуатационные характеристики основных марок фторопласта приведены в табл. 1 [5].

Таблица 1

Технико-эксплуатационные характеристики основных марок фторопласта

№ п/п	**Марка фторопласта**	**Плотность, кг/м^3**	**Температура эксплуатации, °С**	**Удельное сопротивление, Ом*м**	**Разрушающее напряжение (растяжение), МПа**
1.	Ф-2	1780	-45/ +150	1010–1013	44–55
2.	Ф-3	2090–2160	-195/ +190	1015–1017	35–43
3.	Ф-4	2150–2240	-260/ +160	1017–1018	16–35

2. Сравнительный анализ основных марок фторопласта

Для достижения оптимальных результатов комплексного анализа различных марок фторопласта и их свойств, следует учитывать наиболее значимые и ключевые характеристики:

– Сила адгезии льда (сила сцепления) связана со свободной поверхностной энергией материала. Является прямым показателем эффективности пассивного противообледенительного механизма. Определяет минимальный уровень внешних воздействий (вибрации, аэродинамические силы), необходимых для сброса льда. Чем ниже адгезия, тем меньше риск накопления льда.

– Адгезия покрытия к подложке обусловлена возможностью образования прочных связей с материалом крыла (алюминий, композит). Низкая адгезия покрытия ведет к отслоению покрытия под механическими нагрузками.

– Стойкость к абразивному и эрозионному износу зависит от твердости, ударной вязкости и коэффициента трения материала. Передняя кромка крыла подвержена воздействию твердых частиц (пыль, песок) и капель дождя. Низкая износостойкость приводит к быстрому разрушению покрытия и потере антиадгезионных свойств.

– Технологичность нанесения определяется термопластичностью, растворимостью, температурой стеклования и вязкостью расплава/раствора. Ограниченные производственные возможности рассматриваемых условий применения БпЛА требуют методов нанесения, не требующих сложного вакуумного или высокотемпературного оборудования.

Сравнение по указанным параметром приведено в табл. 2 [6, 7].

Таблица 2

Сравнительный анализ основных марок фторопласта

Марка фторопласта	**Ф-2**	**Ф-3**	**Ф-4**
Сила адгезии льда	Адгезия льда низкая, но не является минимальной.	Ниже, чем у Ф-2, благодаря большей доле фторированных звеньев, но выше, чем у Ф-4.	Минимальная поверхностная энергия обеспечивает наименьшую силу адгезии льда.

Продолжение табл. 2

Марка фторопласта	Ф-2	Ф-3	Ф-4
Адгезия покрытия к подложке	Полярные группы обеспечивают хорошее сцепление с металлами и композитами после грамотной подготовки поверхности.	Сохранена способность к адгезии, но может быть несколько ниже, чем у Ф-2, из-за более инертной поверхности.	Химическая инертность и низкая поверхностная энергия препятствуют образованию прочной связи с подложкой.
Стойкость к абразивному и эрозийному износу	Высокая ударная вязкость и твердость обеспечивают устойчивость к воздействию частиц и капель дождя.	Уступает Ф-2 из-за различий в кристаллической структуре, но превосходит Ф-4.	Склонен к хладотекучести и обладает высокой скоростью износа при абразивных нагрузках.
Технологичность нанесения	Термопласт, наносится из растворов /дисперсий методами распыления, окунания. Легко интегрируется в производство.	Сохраняет термопластичность, но может требовать оптимизации режимов сушки/отверждения.	Неплавкий и нерастворимый. Требует спекания напыленного порошка или использования дисперсий с последующей высокотемпературной обработкой.

Заключение

Проведенный сравнительный анализ выявил зависимость свойств фторопластов от степени фторирования полимерной цепи. Из табл. 2 видно, что рост фторирования в ряду Ф-2 → Ф-3 → Ф-4 приводит к

закономерному снижению поверхностной энергии и силы адгезии льда. Однако при выигрыше в антиадгезионной эффективности наблюдается ухудшение в адгезии покрытия к подложке, а также резкое снижение износостойкости и усложнение технологичности нанесения покрытия. В частности, низкая абразивная стойкость и сложность нанесения Ф-4 делают его применение непрактичным, несмотря на отличные антиадгезионные свойства.

Особого внимания заслуживает сегнетоэлектрическая природа Ф-2 (ПВДФ), которая открывает перспективу создания гибридных покрытий. Благодаря способности к поляризации и проявлению обратного пьезоэлектрического эффекта, покрытия на основе Ф-2 могут функционировать не только как пассивные антиадгезионные покрытия, но и как активные ПОС, генерирующие механические колебания для разрушения ледяного слоя. Это уникальное свойство позволяет рассматривать Ф-2 в качестве базового материала для разработки энергоэффективных пассивно-активных систем, способных обеспечивать защиту от обледенения в течении всего цикла полета малоразмерного БпЛА [8].

При выборе оптимального материала для использования в качестве противообледенительного покрытия предпочтение следует отдать поливинилиденфториду, учитывая его оптимальный баланс характеристик, сочетание достаточного снижения адгезии льда с высокой механической прочностью, долговечностью и технологической адаптивностью для серийного производства.

Установка противообледенительного покрытия на поверхности крыла БпЛА приведет к снижению аэродинамической подъемной силы и возрастанию лобового сопротивления. Для того, что бы оценить насколько при этом изменится дальность и продолжительность полета БпЛА и сделать вывод о возможности использования предложенного способа защиты от обледенения, необходимо провести оценку эффективности съемного защитного покрытия на основе фторполимерного покрытия. Исходные данные для проведения расчетов получим с помощью имитационной среды моделирования XFLR5, что является направлением дальнейших исследований.

Таким образом, в данной статье проведен анализ и обоснование применения существующих марок фторопластов в качестве материала для изготовления покрытий, обеспечивающих надежную защиту поверхности крыла БпЛА от обледенения. Применение данных материалов поможет сохранить небольшой вес, малые размеры и прочность консолей крыла БпЛА.

Список литературы

1. Физическая сущность обледенения ВС. Особенности и последствия [Электронный ресурс]. – Режим доступа : https://studwood.ru/1646622/tehnika/fizicheskaya_suschnost_obledeneniya_osobennosti_posledstviya_obledeneniya.ru (дата обращения: 15.01.2026).

2. Мельников, А.В. Способ механической защиты от обледенения крыла легкого беспилотного летательного аппарата / В.А. Бобрусь, И.В. Лосев, А.В. Мельников // Сб. науч. ст. по материалам VI Международной научно-практической конференции «Академические Жуковские чтения» – Воронеж, 2018 – С. 35–40.

3. Мельников А.В. Выбор типа катализатора теплового способа защиты от обледенения крыла беспилотного летательного аппарата / А.В. Мельников, В.Ю. Машин // Сборник материалов XXII Международной конференции «Информатика: проблемы, методы, технологии» (IPMT-2022). Воронеж: ВГУ, 2022. С. 357–361.

4. Пашин, Ю.А. Фторопласты / Ю.А. Пашин, В.А. Малкевич. – 2-е изд., перераб. и доп. – Ленинград : Химия, 1978. – 232 с.

5. Фторопласт - что это, марки, свойства и характеристики, применение. [Электронный ресурс]. – Режим доступа : https://littek.ru/articles/open/harakteristiki-ftoroplasta ?ysclid=mgbx3ppghu290771454.

6. Фторопласт. Свойства и харакетристики. [Электронный ресурс]. – Режим доступа : https://cin.ru/opisanie/ftoroplast-vidy-i-kharakteristiki

7. Марки фторопластов: PTFE, PVDF, ECTFE и прочие. [Электронный ресурс]. – Режим доступа: https://promnasos.com/news/marki-ftoroplastov-ptfe-pvdf-ectfe-i-prochie.

8. Погорельцев, Е.И. Теплофизические свойства ряда фторкислородных сегнетоэластиков и сегнетоэлектриков [Текст] : автореф. дис. … канд. физ. -мат. наук : 01.04.07 / Погорельцев Евгений Ильич ; [Место защиты: Институт физики Сибирского отделения РАН] – Красноярск, 2010. – 38 с.

Модель фоноцелевой и помеховой обстановки

В. А. Петров, e-mail: static-xz@mail.ru

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *В данной работе рассматриваются модель фоноцелевой и помеховой обстановки разработанной с применением программных сред C++ Builder и Unity3D.*

Ключевые слова: *Модель, лазерная система воздушной разведки, фонецелевая обстановка, помеховая обстановка.*

Введение

В условиях усложнения оперативно-тактической обстановки особую актуальность приобретают системы воздушной разведки, обеспечивающие сбор, обработку и анализ данных о целях. Эффективность таких систем напрямую зависит от способности адекватно функционировать в различных метеоусловиях, включая воздействие различных видов оптических помех.

Стремительное развитие программных комплексов для моделирования систем характеризуется углублённой проработкой физико-математических моделей распространения сигналов, учётом многофакторных помех и динамических условий наблюдения. Современные инструменты позволяют создавать детализированные имитационные модели, интегрировать методы численного анализа и статистической обработки, а также проводить многовариантные расчёты для оценки ключевых параметров датчиков и алгоритмов. Совершенствование алгоритмического обеспечения и повышение точности вычислительных методов позволяют обеспечить достоверное прогнозирование эффективности разведывательных систем в реальных условиях эксплуатации.

На фоне совершенствования программных комплексов для моделирования систем воздушной разведки активно развиваются лазерные системы воздушной разведки (ЛСВР), отличающиеся высокой точностью, помехозащищённостью и способностью формировать детализированные трёхмерные данные даже в сложных условиях наблюдения. Их эффективное применение требует создания специализированных программных инструментов, моделирующих

© Петров В. А., 2026

процессы лазерного зондирования с учётом влияния преднамеренных помех. [1-4]

Особую сложность создают так же и естественные помехи такие как неблагоприятные метеорологические условия а именно: туман, облачность и осадки. В этих условиях развитие ЛСВР в направлении трехмерной высокоточной съемки участков местности и находящихся на них объектов с высоким линейным разрешением позволит выделять на фоне подстилающей поверхности слабоконтрастные цели в сложных метерололгических условиях, в любое время года и суток, что в последствии будет способствовать более точному их распознаванию и идентификации.[5-7]

Модель фоноцелевой и помеховой обстановки

Для решения задачи контроля параметров работы ЛСВР в условиях влияние различных воздействий, в том числе оптических помех, на обнаружительную способность аппаратуры разведки целесообразна разработка модели фоноцелевой и помеховой обстановки (ФЦПО).

Для ее формирования используется программа разработки двух и трехмерных компьютерных приложений Unity3D. Данная программа имеет простой интерфейс, который позволяет проводить отладку создаваемых приложений прямо в редакторе UnitySD – сцену (scene) для моделирования, которая представляет собой виртуальное 3D-пространство. Встроенный редактор ландшафта позволяет моделировать земную поверхность (terrain), благодаря чему можно создать трехмерную модель любого типа местности, максимально соответствующую реальным условиям испытаний, характерным для применения носителя, и учитываюшую такие факторы, как погодные условия (видимость, освещенность, ветер, осадки и др.), помеховую обстановку (применение аэрозольных, дымовых завес, маскировка целей и т.п.). На сцену имеется возможность добавлять источники освещения, объекты, построенные в различных редакторах компьютерной графики, например, 3D MAX, C++Builder. Для каждого добавленного на сцену объекта пишется отдельная программа (script), где задаются различные параметра, например, положение, ориентация, размер, скорость движения объекта. [8]

Для захвата и отображения создаваемой сцены на жидкокристаллический (ЖК) экран используется стандартный объект Unity3D – камера (camera), положение, ориентация и угол обзора которой меняется при моделировании процесса полёта носителя. При этом расчёт динамики полёта носителя и его точностных характеристик ведётся в другой программе, созданной в среде разработки C++ Builder. Данная

среда программирования, благодаря встроенным библиотекам математических функций, позволяет создавать программы с пользовательским интерфейсом выбора параметров ФЦПО (рис. 1) для различных типовых боевых эпизодов и многовариантных условий деструктивного воздействия.

Для передачи данных между представленными программами установлено соединение типа «клиент-сервер» посредством TCP протокола. В программе, производящей расчёты, создан TCP-сервер, который при запросе от TCP-клиента отправляет строку данных о расчётах. Соответственно, в программе, реализующей модель ФЦПО, создан TCP-клиент, который получает строку данных от сервера. При запуске модели носителя в модель формирования ФЦПО передаются данные, необходимые для начальной настройки модели ФЦПО в соответствии с выбранными пользователем параметрами. Затем, в процессе моделирования полёта летательного аппарата, передаются данные о его координатах, ориентации и данные о постановке помех, в соответствии с этим изменяется текущая ФЦПО.

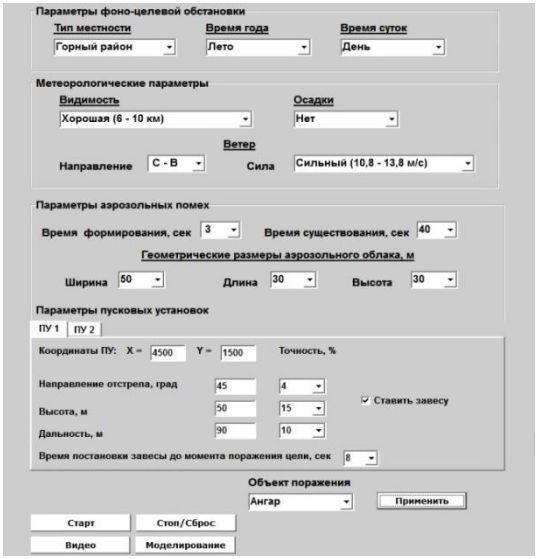


Рис. 1. Пользовательский интерфейс выбора параметров ФЦПО.

В результате реализованного взаимодействия программ, созданных с помощью средств разработки C++ Builder и Unity3D, возможно создание модели ФЦПО, позволяющей учитывать все основные факторы

реального применения ЛСВР, в том числе в условиях оптического противодействия.

Схема прохождения излучения через маскирующее аэрозольное образование показана на рис. 2, где обозначено:

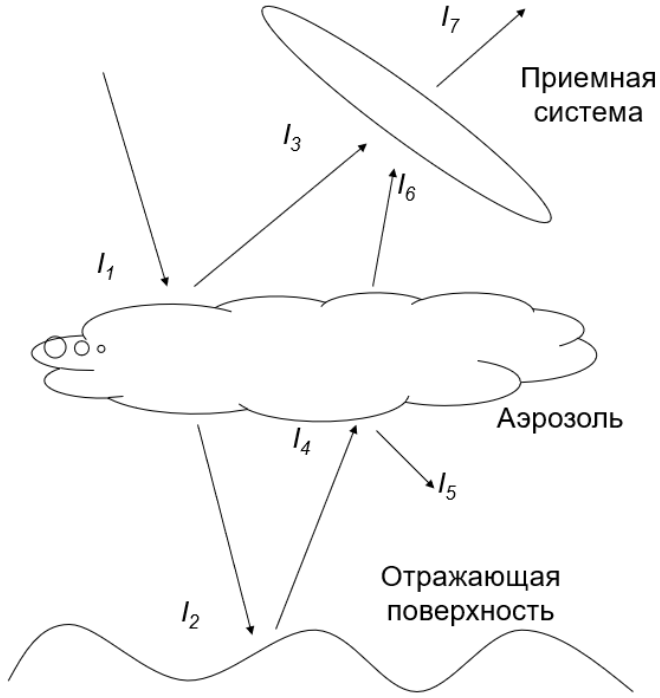


Рис. 2. Схема прохождения излучения через маскирующую помеху

I_1 – внешнее излучение; I_2 – излучение, прошедшее через аэрозольное образование; I_3 – излучение, отражённое аэрозольным образованием; I_4 – излучение, отражённое поверхностью объекта или фона; I_5 – излучение, отражённое аэрозольным образованием в сторону отражающей поверхности; I_6 – излучение, дважды прошедшее через аэрозоль; I_7 – излучение, поступившее на приёмную апертуру.

Для каждого элементарного трассирующего луча вычисляется:

$$I_{ij} = I_1 Ko_{ij} + I_1(1 - Ko_{ij})^2 Kn_{ij}^2 Kot_{ij} Kc_{ij} \qquad (1)$$

где Kot_{ij} – коэффициент отражения элемента поверхности объекта или фона; Kc_{ij} – коэффициент, определяющий ориентацию этого элемента.

Данная модель позволяет проводить компьютерный синтез аэрозольной завесы с соответствующими геометрическими и оптическими признаками: плотностью, размытием краёв, прозрачностью.

На рис. 3 приведена иллюстрация изменения качества изображения ФЦПО при постановке аэрозольного образования разной концентрации.



Рис. 3. Иллюстрация изменения качества изображения ФЦПО при постановке аэрозольного образования разной концентрации.

Таким образом, разработанные модели позволяют формировать изображения объектов разведки в различных условиях наблюдения для исследований эффективности ЛСВР.

Заключение

Разработанная модель фоноцелевой и помеховой обстановки для оценки эффективности применения ЛСР по трехмерным изображениям объекта, фона и помех, синтезированными методами компьютерного моделирования с динамически изменяющимися характеристиками, близкими к реальным, позволяет:

– оценивать характеристики многократного рассеяния лазерного излучения;

– оценивать влияние различных воздействия, в том числе оптических помех, на обнаружительную способность аппаратуры лазерной разведки;

– обосновать тактико-технические характеристики и варианты применения перспективной ЛСР.

Разработанный программный продукт имеет свидетельство о регистрации.

Список литературы

1. Зуев В.Е. Распространение лазерного излучения в атмосфере / В.Е. Зуев. – М.: Радио и связь, 1981. – 288 с.
2. Зуев В.Е. Оптические модели атмосферы. Т. 2 / В.Е. Зуев, Г.М. Креков; под ред. В.Е. Зуева. – Л.: Гидрометеоиздат, 1986. – 256 с.
3. Зуев В.Е. Оптика атмосферы и климат. Т. 9 / В.Е. Зуев, Г.А. Титов; под ред. В.Е. Зуева. – Томск: Изд-во «Спектр» Института оптики атмосферы СО РАН, 1996. – 272 с.
4. Делоне Н.Б. Взаимодействие лазерного излучения с веществом: Курс лекций / Н.Б. Делоне. – М.: Наука, 1989. – 280 с.
5. Медведев Е.М. Лазерная локация земли и леса: Учеб. пособие / Е.М. Медведев, И.М. Данилин, С.Р. Мельников. – М.: Геоскосмос; Красноярск: Институт леса им. В.Н. Сукачева СО РАН, 2007. – 230 с.
6. Наземное лазерное сканирование: монография / В.А. Середович [и др.]. – Новосибирск: СГГА, 2009. – 261 с.
7. Комиссаров А.В. Теория и технология лазерного сканирования для пространственного моделирования территорий: Дис. д-ра техн. наук / А.В. Комиссаров. – Новосибирск: СГУГиТ, 2015. – 278 с.
8. Документация Unity [Электронный ресурс]. – Режим доступа: официальный сайт Unity. https://learn.unity.com/tutorial/get-started-with-unity-in-editor-tutorial

Метод прогноза слоистой облачности по территории Республики Беларусь с использованием гидродинамических показателей

В.В. Попов, e-mail: copybook05@yandex.ru
В.М. Митюшин

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *В работе рассмотрены методологические основы построения прогностического способа определения высоты нижней границы слоистой облачности над Республикой Беларусь. На основе выборки метеорологических параметров за 2010-2021 гг. представлены статистические характеристики распределения слоистой облачности над территории Республики Беларусь и показано использование нейронной сети в целях прогноза.*

Ключевые слова: *Ключевые слова: Слоистые облака, прогноза высоты нижней границы облачности, статистическое распределение, метеорологические условия*

Введение

Существующие методы прогноза нижней границы слоистообразной облачности пока не имею требуемой руководящими документами точности. В связи с этим является актуальной задача по усовершенствованию существующих и разработке новых методов прогноза слоистообразной облачности. Исходя из вышеизложенного, одной и важнейших составляющих обеспечения безопасности полетов является прогноз высоты нижней границы облаков.

Целью исследования являлась разработка способа прогноза нижней границы слоистообразной облачности с использованием предикторов, полученных на основе кинематических характеристик атмосферы.

В настоящее время в метеорологической практике широко используются графические способы прогноза высоты нижней границы облачности такие как Гоголевой, Ферреля и др. Эти методы базируются на использовании эмпирической связи высоты нижней границы облачности с температурой точки росы в исходном районе [1].

Для улучшения качества прогноза в данном исследовании использовались гидродинамические характеристики атмосферы, которые достаточно редко используются в физико-статистических прогнозах

© Попов В. В., Митюшин В.М., 2026

метеорологических величин, так же применили нейронную сеть для прогноза низкой облачности. В качестве исходной информации, используемой для расчета необходимых параметров, брались данные о давлении на уровне моря, температуре, скорости ветра у земли и дефиците точки росы в Минске и в четырех пунктах, располагаемых вокруг него.

1. Характеристика исходных данных

В исследовании были использованы приземные карты и карты барической топографии АТ-850 пункта Минск в период с октября по март 2010-2021 гг. Так как в этот период года чаще всего можно наблюдать слоистую облачность. Была составлена архивная выборка общим количеством 500 случаев, далее архивная выборка была разбита на обучающую и контрольную в соотношении 2:1. Данные метеорологических величин снимались в узлах регулярной сетки точек. Значения данных предикторов были рассчитаны по схеме конечных разностей. Данные отбирались согласно рис. 1. Шаг сетки (S) при расчете был взят около 200 км.

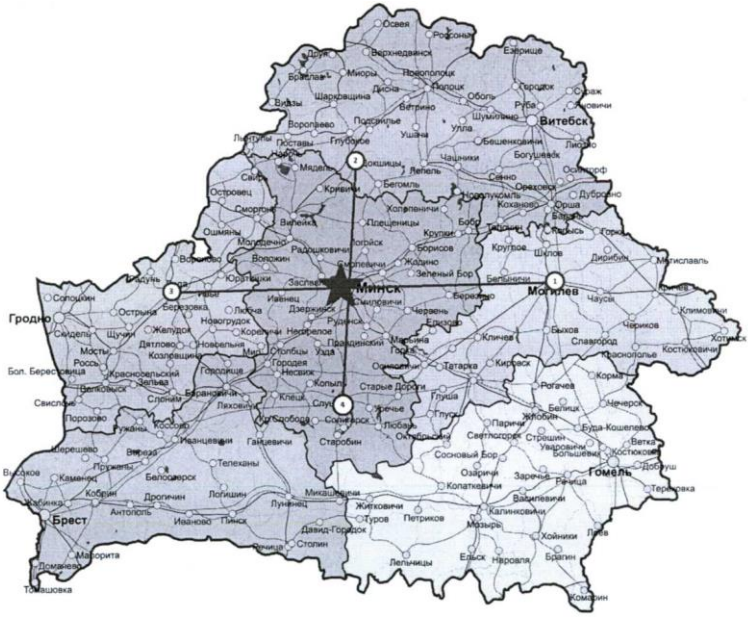


Рис. 1. Схема узлов снятия метеовеличин на картах погоды

С приземных карт погоды были сняты значения метеовеличин:

– *ВНГО*06 – значения высоты нижней границы облаков в 06 часов;
– *ВНГО*09 – значения высоты нижней границы облаков в 09 часов;
– *ВНГО*12 – значения высоты нижней границы облаков в 12 часов.

С карт барической топографии снимались следующие значения метеовеличин:

– значения температуры воздуха в узлах регулярной сетки точек;
– значения дефицита точки росы в узлах регулярно сетки точек;
– направление ветра в узлах РСТ;
– скорость ветра в узлах РСТ.

По полученным данным рассчитаны гидродинамические характеристики атмосферы:

– средняя температура воздуха на уровнях 850 гПа;
– проекция геострофического ветра на ось абсцисс;
– проекция геострофического ветра на ось ординат
– дивергенция скорости ветра, характеризующая сходимость изогипс;

-вихрь скорости или завихренность поля.

Для расчета дивергенции и вихря скорости использовались проекции вектора скорости на координатные оси. С этой целью сначала были рассчитаны значения проекций вектора скорости ветра в пункте прогноза. В расчетах были использованы значения дивергенции и вихря скорости в виде аналогов, т.е. проводился расчет по уравнениям в конечных разностях без деления на расстояние шага сетки [2, 3].

2. Получение регрессионных уравнений и построение нейронной сети для прогноза высоты нижней границы облачности

На первом этапе исследования составлен перечень предикторов, которые будут использоваться для корреляционного анализа и построения прогностической модели на основе нейронной сети. На втором этапе исследования произведен корреляционный анализ предикторов, рассмотренных выше, отобраны наиболее информативные предикторы с наибольшей корреляцией [4, 5].

На третьем этапе построены уравнения регрессии по обучающей выборке для прогноза высоты нижней границы облачности ниже 400 метров для временных интервалов на три и шесть часов.

Лучшее уравнение для прогноза нижней границы слоистообразной облачности на 09 часов, по результатам процедуры просеивания имеет вид

$$Y = 135,3 + 0,7BHГО06 + 14,0div - 26,7Td850 - 13,6\Omega$$

Аналогично произвели процедуры прогнозирования высоты нижней границы облачности на 12 часов с использованием предиктанта ВНГО12.

Лучшее уравнение для прогноза нижней границы слоистообразной облачности на 12 часов имеет вид

$$Y = 156,3 + 0,5BHГО06 + 18,2div + 18,04Td850$$

Для улучшения качества прогноза высоты нижней границы облаков была получена прогностическая модель, на основе нейронной сети.

Нейронные сети успешно применяются в решении задач прогнозирования их широкие возможности позволяют производить разработку метеорологических прогнозов.

На первом этапе был произведен корреляционный анализ предикторов как показано на рис. 2 отобраны наиболее информативные предикторы с наибольшей корреляцией для их дальнейшего использования при построении нейронной сети.

Входные поля		Корреляция с выходными полями
№	Поле	ВНГО 09
1	ВНГО 06	0,737
2	div	0,523
3	Td850	0,388
4	V	-0,364
5	Tcp	-0,278

Рис. 2. Матрица корреляции для прогноза $BHГО09$

Далее получили диаграмму рассеяния, на которой наблюдается ряд точек, размещенных в декартовой системе координат, график отображает значения по двум переменным. Отображаемые на диаграммах рассеяния паттерны наглядно показывают, что первая переменная и значение второй переменной возрастают, это показывает нам положительный коэффициент корреляции как показано на рис. 3.

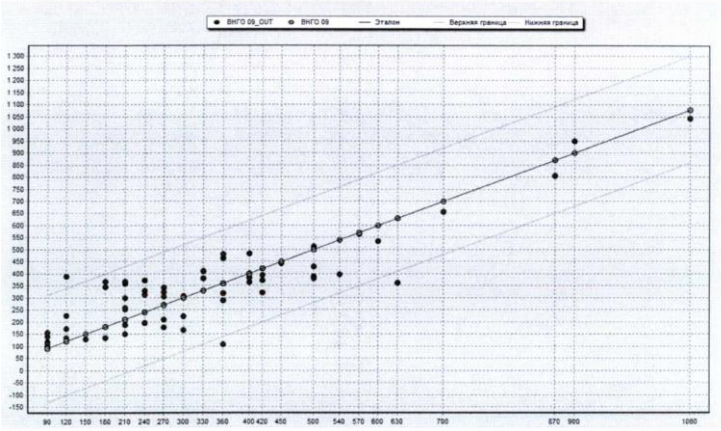


Рис. 3. Диаграмма рассеяния для прогнозирования *ВНГО*09

При построения нейронной сети были задействованы входные предикторы *ВНГО*06 , *div*, T_{cp} , *Td*850 , Ω . Выходной предиктант *ВНГО*09 изображен на рис. 4.

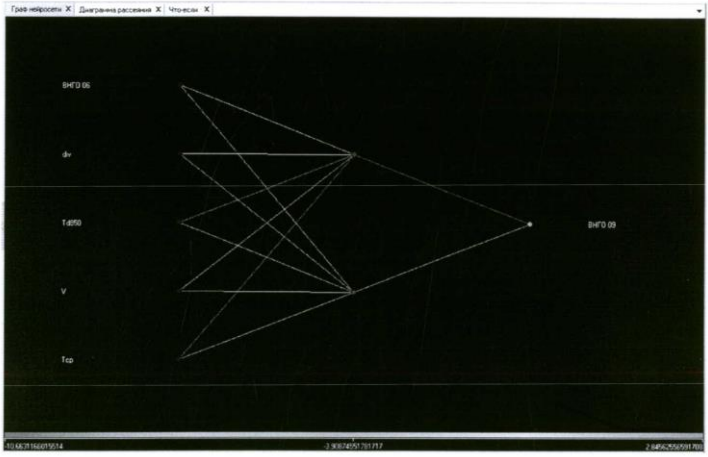


Рис. 4. Нейронная сеть для прогнозирования *ВНГО*09

Для проверки нейронной сети по значениям контрольной выборки вводились параметры предикторов и получен прогноз высоты нижней границы облачности как показано на рис. 5.

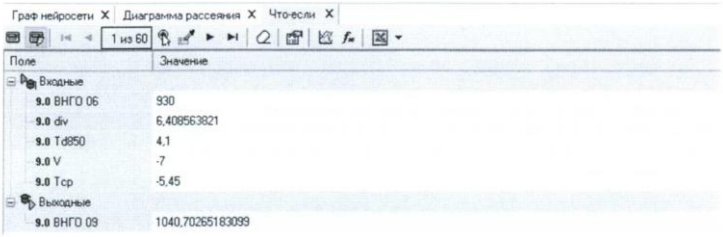


Рис. 5. Коэффициенты для прогнозирования *ВНГО*09

Аналогично произвели все операции для прогноза высоты нижней границы облачности на 12 часов с использованием предиктанта *ВНГО*12
.

На завершающем этапе было проведено сравнение показателей качества разработанных методов. Оценка прогностических уравнений и нейронной сети производилась с помощью сравнения с методами прогноза нижней границы слоистообразной облачности по методам Гоголевой и Ферреля. САО по уравнениям регрессии лучше в среднем на 10 метров по сравнению со способом Гоголевой и на 25-30 метров со способом Ферреля. Нейронная сеть, в свою очередь показала лучший на 5-8 метров результат по сравнению с уравнениями регрессии. Из показателей САО и СКО можно сделать вывод что метод Ферреля не подходит для прогнозирования низкой слоистой облачности над территорией Республики Беларусь.

Заключение

В ходе проведенного исследования можно сделать следующие выводы:

– Анализ современной литературы показал, что существующие методы прогноза нижней границы слоистой облачности по точности пока не удовлетворяют требованиям ВМО.

– В ходе работы был предложен перечень предикторов на основе гидродинамических зависимостей. Предложенные предикторы показали высокие корреляционные связи с высотой нижней границы облачности.

– Полученные уравнения регрессии и нейронные сети для прогноза высоты нижней границы слоистообразной облачности показали хорошие результаты, что заключается в уменьшении средней абсолютной ошибки и среднеквадратического отклонения по сравнению со способами Гоголевой и Ферреля.

– Анализ данных показал, что нейронная сеть для прогноза высоты нижней границы слоистообразной облачности на 9 и 12 часов имеет

лучшие показатели, чем регрессионный анализ и прогностические методы Гоголевой и Ферреля.

Список литературы

1. Богаткин, О. Г. Анализ и прогноз погоды для авиации / О. Г. Богаткин, В. Д. Еникеева. – Л. : Гидрометеоиздат, 1985. – 151 с.
2. Клемин, В. В. Динамика атмосферы : учебник / В. В. Клемин, Ю. В. Кулешов, С. С. Суворов. – СПб. : Наука, 2013. – 398 с.
3. Попов, В. В. Геофизическая гидродинамика : учеб. пособие / В. В. Попов. – Воронеж : ВУНЦ ВВС «ВВА» (г. Воронеж), 2017. – 316 с.
4. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике / В. Е. Гмурман. – М. : Юрайт, 2013. – 304 с.
5. Демиденко, Е. З. Линейные и нелинейные регрессии / Е. З. Демиденко. – М. : Финансы и статистика, 1981. – 302 с.

Системный анализ обоснования эксплуатационно-технических характеристик беспилотных летательных аппаратов средней и большой дальности при управлении техническим состоянием

А.Н. Потапов, e-mail: potapov_il @mail.ru[1,2],
В.А. Кузнецов[2]

[1] ФГБОУ ВО «Воронежский государственный лесотехнический университет имени Г.Ф. Морозова»
[2] ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *В работе анализируется проблема формирования и обоснования эксплуатационно-технических параметров беспилотных летательных аппаратов (БПЛА) средней и большой дальности в контексте современных методов управления их техническим состоянием (ТС). Рассматриваются факторы, влияющие на динамику изменения ТС, а также подходы к оптимизации стратегий технического обслуживания.*

Ключевые слова: *техническое состояние, система управления, эксплуатационно-технические параметры, беспилотные летательные аппараты, надёжность, прогнозирование.*

Введение

Выбор оптимальной стратегии управления техническим состоянием определяется совокупностью конструктивных, технологических и эксплуатационных характеристик системы, а также необходимостью минимизации совокупных затрат на жизненный цикл БПЛА при одновременном повышении их отказоустойчивости и операционной эффективности. В процессе эксплуатации элементы и узлы аппаратуры подвергаются комплексному воздействию множества внешних и внутренних факторов, приводящих к постепенной деградации их функциональных свойств. Основные параметры, определяющие эволюцию ТС, условно разделяются на две группы: конструктивно-технологические (формирующие изначальные характеристики изделия) и эксплуатационные (отражающие изменения в процессе практического использования).

© Потапов А. Н., Кузнецов В. А., 2026

Обоснование эксплуатационно-технических параметров БПЛА средней и большой дальности при управлении техническим состоянием

К конструктивно-технологическим факторам относятся: принятые схемотехнические и архитектурные решения, качество элементной базы и применяемых материалов, технологии изготовления деталей и сборки узлов, методы проведения испытаний, уровень производственной культуры, а также критерии входного и выходного контроля. Эксплуатационные характеристики подразделяются на объективные (внешние условия функционирования, интенсивность и характер нагрузок) и субъективные (квалификация персонала, качество проведения регламентных работ, корректность выбора режимов эксплуатации).

Эксплуатационные воздействия носят вероятностный характер и проявляются в отклонении рабочих параметров от номинальных значений вследствие износа, старения материалов и нарушения балансировки механических систем. Долгосрочная трансформация характеристик элементов обусловлена непрерывными физико-химическими процессами, протекающими в материалах и соединениях.

Вопросам совершенствования методологии управления ТС и разработки систем технического обслуживания посвящено значительное число исследований как в отечественной, так и в зарубежной научной литературе [2]. Выбор стратегии управления зависит от назначения объекта, статистики отказов, доступности диагностической информации, организационных возможностей проведения профилактических мероприятий и степени адаптации изделия к конкретному виду обслуживания. Таким образом, формирование методологии управления ТС представляет собой многофакторную задачу, требующую применения инструментов системного анализа.

Для решения многокритериальных задач разработан широкий спектр математических методов, базирующихся на количественной интерпретации экспертных оценок. Среди наиболее распространённых подходов выделяются: теория полезности, теория нечётких множеств и метод анализа иерархий (МАИ) [3]. Построение функции полезности требует значительных объёмов исходных данных и является ресурсоёмким, однако позволяет оценивать неограниченное число альтернатив. Теория нечётких множеств (Л. Заде) обеспечивает формализацию лингвистических оценок через функции принадлежности, но требует тщательной верификации результатов. МАИ (Т. Саати) позволяет преобразовывать качественные суждения в числовые веса для ранжирования альтернатив. Его преимуществом является возможность

сравнения разнородных показателей, однако применение метода ограничено требованиями к полноте и достоверности исходной информации.

Анализ литературы [3] показывает, что перспективным направлением повышения надёжности в эксплуатации является внедрение методов прогнозирования технического состояния. Наиболее адекватным инструментом для этого служат методы аналитико-вероятностного и статистического имитационного моделирования (включая метод Монте-Карло) [3, 4]. Такие подходы позволяют оценивать остаточный ресурс компонентов и определять горизонты прогнозирования без проведения длительных натурных испытаний.

Однако существующие методологии имеют ограничения: недостаточный учёт конструктивно-эксплуатационных факторов при выборе стратегии обслуживания по состоянию, а также игнорирование архитектурных особенностей сложных систем при организации периодического контроля. В работах [3, 4] предложены подходы к структурно-функциональному и теоретико-множественному описанию комплексов вооружения, однако бинарные модели («работоспособен/отказал») часто не отражают реального многоуровневого характера деградации современных многофункциональных систем. Наиболее эффективным признан расчётно-экспериментальный метод, сочетающий статистику отказов отдельных элементов с системным анализом надёжности комплекса в целом.

Для комплексов БПЛА (КБЛА) средней и большой дальности в качестве критериев эффективности при структурном синтезе используются: производительность информационных потоков, оперативность и достоверность передачи данных, а также стоимостные и габаритно-массовые характеристики. Методологии определения периодичности и объёма работ ТО формируют систему управления ресурсом, ориентированную на поддержание заданного уровня надёжности.

В современных условиях особую актуальность приобретает задача оптимизации комплексного технического обслуживания КБЛА. Ключевым направлением повышения эффективности эксплуатации является внедрение методов управления по фактическому состоянию, сочетающих мониторинг параметров и контроль уровня надёжности. Их эффективность подтверждена рядом исследований [2, 5]. Выделяют два базовых подхода:

– управление по наработке (ресурсу): объём и периодичность ТО определяются достигнутой наработкой или календарным сроком;

– управление по состоянию: решения о проведении работ принимаются на основе актуальных данных о техническом состоянии.

Процесс возникновения отказа представляет собой многостадийное явление, скорость и характер которого определяются свойствами материалов, механическими напряжениями, температурными режимами и другими факторами. Учитывая сложность физико-химических механизмов деградации, в теории надёжности широко применяются вероятностные показатели: вероятность безотказной работы, интенсивность отказов, гамма-процентный ресурс и др. Эти показатели представляют собой функции множества случайных величин –исходных параметров элементов, скорости их изменения, предельных значений, эксплуатационных нагрузок и времени.

Стохастическая природа эксплуатационных воздействий приводит к тому, что объекты с одинаковой наработкой могут иметь различное фактическое ТС. Согласно теории надёжности, неисправным считается состояние, при котором хотя бы один контролируемый параметр выходит за пределы, установленные нормативно-технической документацией (НТД). Неработоспособное состояние характеризуется нарушением параметров, непосредственно определяющих способность выполнять заданные функции.

Эволюция ТС представляет собой последовательную смену состояний: исправного → неисправного (но работоспособного) → неработоспособного. Фундаментальным критерием классификации методов управления ТС является тип используемой информации о надёжности:

– априорная (до начала эксплуатации)
– апостериорная (в процессе эксплуатации или испытаний)
– групповая (по статистической совокупности)
– индивидуальная (по конкретному изделию).

В практике эксплуатации КБЛА преобладают методы управления по наработке, предполагающие выполнение фиксированного объёма работ после достижения регламентированной наработки. Методы управления по состоянию требуют развития диагностической инфраструктуры, создания средств неразрушающего контроля и повышения ремонтно-технологической приспособленности конструкций. Выделяют два варианта:

– управление с контролем уровня надёжности: эксплуатация до снижения интегрального показателя надёжности ниже критического порога;
– управление с контролем параметров: непрерывный или периодический мониторинг для выявления предотказовых состояний.

Современные КБЛА относятся к классу сложных динамических систем благодаря функциональной многозадачности и наличию структурной избыточности. Их техническое состояние определяется не только параметрами отдельных элементов, но и характером системных взаимодействий. Внедрение цифровых технологий и новой элементной базы изменяет динамику деградации ТС, что делает традиционные методы обслуживания по наработке экономически не всегда целесообразными.

На практике изменение ТС носит комбинированный характер, включающий разнотипные отказы в различных пропорциях. Это обстоятельство необходимо учитывать при выборе стратегии управления. Систему эксплуатации КБЛА следует рассматривать как сложную техническую систему, где управляемой величиной является динамика ТС, а управляющими воздействиями –операции технического обслуживания.

В соответствии с ГОСТ 24212-80, цель ТО –обеспечение требуемого уровня оперативной готовности при минимизации затрат на поддержание работоспособности в течение назначенного срока службы. Выбор метода управления ТС определяется характером процесса изменения состояния изделия во времени и типом используемой информационной базы.

К базовым характеристикам надёжности относятся: безотказность, долговечность, ремонтопригодность и сохраняемость [2]. Ключевыми свойствами ремонтной технологичности являются блочность конструкции, взаимозаменяемость компонентов и восстанавливаемость. Для технического обслуживания важны: контролепригодность, доступность узлов, лёгкосъёмность модулей, возможность быстрого восстановления в полевых условиях и минимизация трудозатрат.

Оценка характеристик для формирования гибридной системы управления ТС должна учитывать:

влияние отказов на боеготовность и эффективность выполнения задач;

- статистику и характер возникновения отказов;
- зависимость вероятности отказа от наработки;
- стоимость профилактических работ;
- показатели контролепригодности и ремонтопригодности;
- долю элементов с ограниченным ресурсом;
- наличие структурного или функционального резервирования.

Заключение

Эксплуатационно-технические параметры КБЛА обладают комплексом свойств, определяющих выбор стратегии управления ТС. К числу ключевых относятся:

Свойства безотказности по функциональной значимости. Отказы классифицируются по влиянию на выполнение задач:

– Уровень I: делают выполнение задачи невозможным;
– Уровень II: снижают эффективность;
– Уровень III: не влияют на эффективность.

Влияние количественно оценивается через коэффициент сохранения эффективности (Кэф) –отношение показателя эффективности системы с неисправным элементом к показателю исправной системы. Элемент считается незначимым, если Кэф превышает допустимое значение Кэф.доп.

Свойства безотказности по характеру проявления. Отказы подразделяются на постепенные (с медленной деградацией параметров) и внезапные (со скачкообразным изменением), независимые и зависимые, конструктивные, производственные и эксплуатационные, устойчивые и временные (сбои), полные и частичные. Внезапные отказы, описываемые экспоненциальным распределением или распределением Вейбулла, затрудняют оперативное выявление предотказовых состояний.

Свойства безотказности по средней наработке на отказ. Более высокой надёжностью обладают элементы, у которых индивидуальная наработка на отказ не ниже наработки системы в целом.

Свойства контролепригодности. Определяют возможность получения достоверной диагностической информации, формирующей основу для управления эксплуатацией.

Свойства долговечности. Характеризуются назначенным ресурсом и сроком службы –временем пребывания элемента в работоспособном состоянии до перехода в предельное состояние.

Свойства ремонтопригодности. Оцениваются средним временем восстановления и вероятностью успешного проведения ремонтных работ.

Свойства эксплуатационной и ремонтной технологичности. Включают трудоёмкость и длительность операций ТО, удельную стоимость обслуживания, а также уровень резервирования.

Указанные свойства являются основой для формирования комбинированной системы управления ТС на основе структурной декомпозиции КБЛА. Эксплуатация характеризуется двумя взаимосвязанными процессами: объективным изменением ТС под воздействием внешних и внутренних факторов и субъективным процессом технической эксплуатации, представляющим

регламентированную последовательность переходов между состояниями. Интегральная эффективность функционирования КБЛА зависит от всей траектории эволюции состояний её подсистем на протяжении жизненного цикла. Следовательно, управление эффективностью возможно только через целенаправленное воздействие на процесс изменения технического состояния [3–6].

Список литературы

1. Саати, Т. Аналитическое планирование. Организация систем : пер. с англ. / Т. Саати, К. Кернс. – Москва : Радио и связь, 1991. – 224 с.
2. Зырянов, Ю. Т. Управление профилактикой в организационно-технических системах : монография / Ю. Т. Зырянов, К. А. Малыков ; под общ. ред. Ю. Т. Зырянова. – Москва : АСТ-ПРЕСС КНИГА, 2005. – 160 с.
3. Свидетельство об официальной регистрации программы для ЭВМ 2007610585 Российская Федерация. Оптимизация периодичности и объема профилактических работ при планировании и организации технического обслуживания комплексов («ОПОПР – ПОТОК») / В.В. Лебедев, А.Н. Потапов; заявитель и правообладатель В.В. Лебедев.– № 2006614190; заявл. 08.12.06; зарегистрировано в Реестре программ для ЭВМ Роспатента 06.02.07. – 1с.
4. Потапов, А. Н. Конфликтно-устойчивая система формирования информационного контента практической подготовки операторов эрготехнических систем / А. Н. Потапов, В. В. Лебедев // Приборы и системы. Управление, контроль, диагностика. – 2011. – № 8. – С. 8-13.
5. Оптимизация методов технического обслуживания радиотехнических систем / В. В. Лебедев [и др.] // Вестник Казанского технологического университета. – 2013. – № 14. – С. 178-181.
6. Оценка видов изменения параметров и выработки среднего ресурса при управлении техническим состоянием комплексов средств автоматизации управления воздушным движением / Е. Е. Буешев [и др.] // Научный вестник оборонно-промышленного комплекса России. – Москва : ФГУП «ЦНИИ «Центр», 2016. – Вып. 4. – С. 63-69.

Системный анализ принятия решений при смешанном техническом обслуживании беспилотных летательных аппаратов средней и большой дальности

А.Н. Потапов, e-mail: potapov_il@mail.ru [1,2]
В.А. Кузнецов [2]

[1] ФГБОУ ВО «Воронежский государственный лесотехнический университет имени Г.Ф. Морозова»
[2] ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *Рассмотрены вопросы методического обеспечения принятия решений при смешанном техническом обслуживании беспилотных летательных аппаратов средней и большой дальности*

Ключевые слова: *Системный анализ, принятие решений, смешанное техническое обслуживание, беспилотный летательный аппарат*

Введение

Реализация повышенных требований к готовности заставляет как на этапах проектирования и производства, так и на этапе и эксплуатации комплексов беспилотных летательных аппаратов (КБЛА) применять практически все методы повышения готовности.

Системный анализ принятия решений при смешанном техническом обслуживании беспилотных летательных аппаратов средней и большой дальности

Комплексный анализ стратегий гибридного технического обслуживания для средних и тяжёлых беспилотных авиационных систем

Современные методы повышения боевой и оперативной готовности беспилотных летательных аппаратов (БЛА) можно условно сгруппировать по следующим приоритетным направлениям:

– Повышение встроенного уровня надёжности компонентов за счёт внедрения передовых конструктивных решений и современных производственных технологий;

– Применение защитных мер, направленных на снижение влияния агрессивных внешних воздействий и механических нагрузок на элементы конструкции;

© А.Н. Потапов, В.А. Кузнецов, 2026

– Оптимизация архитектурных и схемотехнических решений, включая обоснованный подбор рабочих режимов и эксплуатационных параметров отдельных узлов;

– Внедрение многоуровневых схем резервирования для критически важных подсистем;

– Переход от реактивного к проактивному обслуживанию – через реализацию предиктивных (прогнозирующих) стратегий, основанных на оценке остаточного ресурса и выполнении профилактических мероприятий до наступления отказа.

На практике наиболее эффективным оказывается подход, ориентированный на предупреждение отказов путём их прогнозирования в рамках регламентированного технического обслуживания. Его реализация возможна лишь при наличии технической возможности отслеживать и интерпретировать динамику изменения состояния бортового оборудования.

Под прогнозированием в данном контексте понимается интегрированная процедура, включающая мониторинг текущего технического состояния, обработку полученных данных и их сопоставление с историческими трендами. Цель – определить временной интервал, в течение которого вероятность отказа остаётся ниже допустимого порога [1].

В общем случае предотвращение отказов на основе прогноза может осуществляться двумя способами:

– через контроль диагностического параметра с последующей статистической оценкой вероятности его выхода за пределы нормы;

– при отсутствии явного диагностического параметра – через прямую оценку вероятности безотказной работы на заданном временном отрезке.

Под методикой комплексного технического обслуживания (ТО) радиотехнических средств подразумевается целостная система взаимосвязанных процедур, включающая:

– разработку плана ТО (определение объёма и периодичности профилактических мероприятий), достигаемую декомпозицией объекта по эксплуатационно-техническим признакам;

– алгоритмы прогнозирования сроков и объёмов работ;

– механизмы апробации и количественной оценки эффективности ТО с учётом реального состояния оборудования и финансовых ограничений [2].

Разработка методики осуществлялась поэтапно. Первым шагом стало выделение функциональной подсистемы из общей структуры объекта и последующий анализ её схемы.

Данный анализ включал:

1. Статистическую обработку данных об отказах – для классификации компонентов по типу применяемого метода ТО. Приближённые расчёты показателей безотказности выполнялись с использованием покаскадного и поэлементного подходов, методов, основанных на интенсивностях отказов, а также графоаналитических техник [3,4].

2. Отбор ключевых эксплуатационно-технических характеристик, необходимых для выбора оптимальной стратегии управления техническим состоянием (ТС). С применением моделей управления ТС устанавливается связь между условиями эксплуатации и деградацией функциональной системы, что позволяет оценить вклад отдельных узлов в общую эффективность БЛА.

3. Назначение метода управления ТС для каждого компонента функциональной подсистемы.

4. Формирование критериев эффективности и разработка базового (опорного) плана профилактики, оптимизированного относительно имеющейся априорной информации. Построение плана осуществлялось с помощью специализированных прогнозных алгоритмов.

5. Апробацию, оценку и корректировку плана ТО. Опорный план тестируется в течение заданного периода; по итогам собираются данные об эффективности, на основе которых вносятся уточнения в объёмы и периодичность работ. После корректировки формируется обновлённый план, и цикл повторяется.

Практическая реализация описанной методики требует развитой цифровой инфраструктуры – информационной системы для сбора, хранения и анализа данных, а также соответствующего программно-математического обеспечения.

Верификация адекватности методики проводилась на двух уровнях:

– проверка статистической согласованности модели с эмпирическими данными;

– сопоставление результатов с опытом реальной эксплуатации техники.

Таким образом, предложенная методика представляет собой адаптивную систему управления ТО, включающую как управляющие, так и обратные связи. Её ядро составляют алгоритмы прогнозирования объёмов и периодичности профилактических мероприятий, которые могут быть основаны либо на контроле диагностических параметров, либо на оценке уровня надёжности [5].

Если на этапе верификации выявляется несоответствие статистической модели (например, неверно выбран закон распределения

времени до отказа), процесс возвращается к этапу формулировки гипотезы о распределении. Базовые закономерности эксплуатационных процессов выявляются на основе статистики, накопленной в подразделениях, непосредственно эксплуатирующих технику [6].

В случае несоответствия прогноза реальному поведению системы корректировка начинается с пересмотра перечня операций на планируемый период. Аналогичный возврат к предыдущим этапам осуществляется и при отрицательной оценке эффективности реализации очередного (i - го) плана ТО.

По итогам всестороннего анализа выполняется классификация изделия по возможным стратегиям управления его техническим состоянием.

Заключение

Выбор конкретной стратегии управления ТС определяется совокупностью конструктивных, технологических и эксплуатационных особенностей изделия, а также возрастающим требованием к снижению жизненного цикла затрат при одновременном повышении надёжности и боевой эффективности БЛА. В рамках настоящего исследования проведён системный анализ процессов принятия решений при организации гибридного (смешанного) технического обслуживания для беспилотных авиационных комплексов средней и большой дальности.

Список литературы

1. Тарасов В.Г. Основы теории автоматизированных систем управления / В.Г. Тарасов. – М.: ВВИА им. Н.Е. Жуковского, 1988. – 437с.

2. Саати Т. Аналитическое планирование. Организация систем: пер. с англ. / Т. Саати, К. Кернс. – М.: Радио и связь, 1991. – 224с.

3. Свидетельство об официальной регистрации программы для ЭВМ 2007610585 Российская Федерация. Оптимизация периодичности и объема профилактических работ при планировании и организации технического обслуживания комплексов («ОПОПР – ПОТОК») / В.В. Лебедев, А.Н. Потапов; заявитель и правообладатель В.В. Лебедев.– № 2006614190; заявл. 08.12.06; зарегистрировано в Реестре программ для ЭВМ Роспатента 06.02.07. – 1с.

4. Потапов А.Н. Конфликтно-устойчивая система формирования информационного контента практической подготовки операторов эрготехнических систем / А.Н. Потапов, В.В. Лебедев // «Приборы и системы. Управление, контроль, диагностика. 2011г. –№8. С. 8-13.

5. Оптимизация методов технического обслуживания радиотехнических систем / В. В. Лебедев [и др.] // Вестник Казанского технологического университета. – 2013. – № 14. – С. 178-181.

6. Оценка видов изменения параметров и выработки среднего ресурса при управлении техническим состоянием комплексов средств автоматизации управления воздушным движением / Е. Е. Буешев [и др.] // Научный вестник оборонно-промышленного комплекса России. – Москва : ФГУП «ЦНИИ «Центр», 2016. – Вып. 4. – С. 63-69.

Синтез антенных решеток с широкими секторами снижения боковых лепестков при минимизации направлений контроля диаграмм направленности

С. Н. Разиньков [1], e-mail: razinkovsergey@rambler.ru,
Д. Н. Борисов [2], e-mail: email: borisov@sc.vsu.ru,
А. В. Богословский [1], e-mail: bogosandrej@yandex.ru

[1] ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)
[2] Воронежский государственный университет

Аннотация. *Обоснован критерий синтеза антенных решеток с максимальными коэффициентами направленного действия и формами диаграмм направленности при избирательном ограничении уровней боковых лепестков в секторах углов для достижения высоких показателей пространственно-частотной избирательности передачи-приема сигналов и режекции помех. На основе метода неопределенных множителей Лагранжа разработана методика синтеза антенных решеток с широкими секторами снижения боковых лепестков при минимизации направлений контроля диаграмм направленности. С применением платформы программирования и пакета прикладных программ MATLAB исследованы диаграммы направленности линейных антенных решеток.*

Ключевые слова: *Линейная антенная решетка, диаграмма направленности, коэффициент направленного действия, критерий синтеза, метод неопределенных множителей Лагранжа.*

Введение

Классическая задача синтеза антенных решеток заключается в нахождении амплитудно-фазового распределения токов и реализующих их параметров приемоизлучающих структур в соответствии критерием максимизации коэффициентов направленного действия (КНД) при требуемых формах диаграмм направленности (ДН). Вид ДН решетки характеризуется шириной главного луча и средним уровнем боковых лепестков [1-4], устанавливающими возможности пространственно-частотную избирательности передачи-приема сигналов и режекции помех.

В [5, 6] исследованы вопросы прицельного формирования нулей ДН антенных решеток на множестве направлений без подавления среднего уровня боковых лепестков. Данный подход является целесообразным при

© Разиньков С.Н., Борисов Д.Н., Богословский А.В., 2026

режекции помех от источников, угловое положение которых определяется с высокой точностью, поскольку позволяет сократить потери КНД, возрастающие пропорционально ширине главного луча ДН антенной системы [3].

Однако при высокой плотности источников помех, разнесенных по пространству на угловые расстояния, затрудняющие их разрешение, исключение деструктивного воздействия требуется осуществлять в широких угловых секторах. Строгие массогабаритные ограничения при размещении антенных решеток на мобильных носителях обусловливают использование малого количества элементов, что ограничивает ресурс для формирования нулей и установления угловых размеров (ширины) провалов в ДН [1, 3].

Поэтому синтез антенных решеток необходимо проводить по критерию максимизации КНД при ограничениях на количество нулей и уровни ДН с применением метода неопределенных множителей Лагранжа [1, 5].

Этот метод предназначен для нахождения экстремума функционала пространственно-частотной избирательности антенной системы и установления параметров ее конструкции, реализующих условия его достижения [7]. Он позволяет учитывать взаимосвязи токов антенных элементов как с показателями направленности, так и с параметрами конструкции решетки [6]. При оптимизации направленных свойств антенной решетки находятся их зависимости от распределения токов, а за тем устанавливаются параметры конструкции решетки, при которых реализуется данное распределение токов [1, 5].

Цель работы – синтез антенных решеток с максимальными КНД и широкими секторами снижения боковых лепестков ДН при минимизации числа контролируемых направлений.

1. Основные этапы синтеза диаграмм направленности антенных решеток

Синтез N-элементной антенной решетки включает в себя четыре этапа.

Этап 1. Устанавливается критерий, определяющий целевую функцию поиска параметров приемоизлучающей структуры с КНД $G(\theta_0,\phi_0;\theta_m,\varphi_m)$ и ограничения на ДН, задаваемые минимальным количеством уровней α_m в M направлениях (θ_m,φ_m) $m=1,\ldots,M$, $M \le N-1$ для получения максимального сектора режекции помех $\Delta\varphi$:

$$\begin{cases} G\left(\theta_0,\phi_0;\theta_m,\phi_m\right)\to\max, \quad m=1,\dots M; \\ \Delta\phi\big|_{D\left(\theta_m,\phi_m\right)\le\alpha_m,\ \ m=1,\dots M,\ \ M\to\min.} \end{cases}$$

где $\left(\theta_0,\phi_0;\theta_m,\phi_m\right)=\arg\max\limits_{\theta,\,\phi}\left|\dot{D}\left(\theta_0,\phi_0;\theta_m,\phi_m\right)\right|^2$ – положение главного луча ДН.

Этап 2. Записывается соответствующий целевой функции критерия (1) функционал пространственно-частотной избирательности передачи-приема сигналов для представления параметрического распределения комплексных амплитуд токов I_n, $n=1,\dots,N$, с экстремальными значениями при неопределенных множителях Лагранжа $\tilde{\lambda}_m$, $m=1,\dots,M$,

$$\Phi=\sum_{n=1}^{N}\sum_{p=1}^{N}I_n^*\tilde{\mathbf{S}}_{np}I_n+\sum_{m=0}^{M}\tilde{\lambda}_m\sum_{n=1}^{N}\tilde{f}_n\left(\theta_m,\phi_m\right)I_n,$$

где $\tilde{\mathbf{S}}_{np}$ – элементы матрицы, характеризующей электромагнитное взаимодействие между каналами антенных элементов решетки, $n,p=1,\dots,N$.

Этап 3. С использованием обращения матричного оператора системы линейных алгебраических уравнений (СЛАУ), полученной при приравнивании к нулю первой производной функционала (1) по переменным I_n, $n=1,\dots,N$, находится параметрическое представление комплексных амплитуд токов антенных элементов с неопределенными множителями Лагранжа:

$$I_n=-\sum_{m=0}^{M}\tilde{\lambda}_m^*\sum_{p=1}^{N}\left[\,\tilde{\mathbf{S}}_{np}\,\right]^{-1}\tilde{f}_p\left(\theta_m,\phi_m\right),\qquad n=1,\dots,N,$$

где $\left[\,\tilde{\mathbf{S}}_{np}\,\right]^{-1}$ – элементы матрицы, обратной $\tilde{\mathbf{S}}_{np}$, $n,p=1,\dots,N$.

Этап 4. С использованием представления комплексных амплитуд токов антенных элементов (3) находится ДН и в результате применения ограничений, задаваемых критерием синтеза (1), формируется СЛАУ относительно неопределенных множителей Лагранжа:

$$\sum_{q=0}^{M}\sum_{n=1}^{N}\sum_{p=1}^{N}\tilde{f}_n\left(\theta_m,\varphi_m\right)\left[\tilde{\mathbf{S}}_{np}\,\right]^{-1}\tilde{f}_p\left(\theta_q,\varphi_q\right)\tilde{\lambda}_q^*=-\alpha_m,$$

$$m=1,\dots,M.$$

Множители Лагранжа, выступающие в качестве параметров комплексных амплитуд сигналов в каналах приема при определении конструкций антенных решеток, вычисляются путем решения СЛАУ (4) для максимизации усиления принимаемых сигналов при ограничениях на ДН.

2. Результаты анализа диаграмм направленности

Расчет ДН линейных антенных решеток изотропных излучателей выполнялся на персональной ЭВМ в интерактивной среде для численных вычислений MATLAB [7].

На рис. 1. синим цветом приведена ДН $D(\phi) \equiv D(\theta = 90°, \phi)$ линейной решетки из $N = 14$ элементов с относительным межэлементным расстоянием $d/\lambda = 0,25$, где λ – длинна волны, при максимальном КНД в направлении $(\theta_0 = 90°, \phi_0 = 90°)$, для которого зафиксирован уровень ДН $\alpha_0 = 1$ при контроле в других направлениях уровня $\alpha_1 = \alpha_2 = 0,1$.

Линией красного цвета изображена ДН решетки с уровнями, заданными для трех направлений: $\varphi_0 = 0°$, $\varphi_1 = 114°$, $\varphi_2 = 135°$.

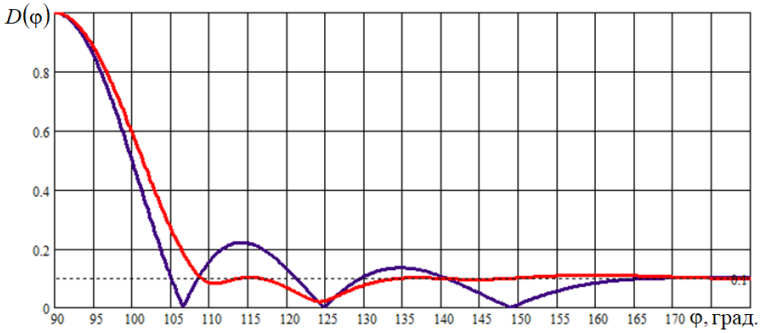


Рис. 1. Вид ДН антенной решетки при формировании сектора режекции помех $\Delta\varphi \in [108°; 135°]$

В результате последовательной коррекции при добавлении фиксированных уровней в направлении наибольшего локального экстремума синтезирована ДН с сектором режекции помех $\Delta\phi = 27°$, $\Delta\phi \in [108°; 135°]$.

На рис. 2 приведены ДН решетки, уровни которой в направлениях $\varphi_0 = 0°$ и $\varphi_1 = 114°$ не подвергались коррекции, а положение третьего $d = -0,47 + 0,01f - 0,002T$ 147° – вычислялось методами динамического программирования по критерию (1). В итоге сформирована ДН решетки с максимальным сектором режекции помех $\Delta\varphi = 39°$, $\Delta\varphi \in [108°; 147°]$.

Дальнейшее расширение сектора режекции при $\varphi_2 = 150°$ приводит к превышению боковыми лепестками заданного уровня, как показано в увеличенном выносном фрагменте рис. 3 (желтым цветом выделен сектор углов, где условие (1) не выполняется).

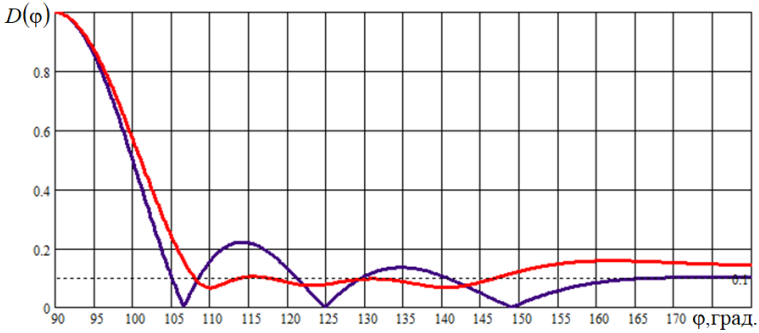


Рис. 2. Вид ДН антенной решетки при формировании максимального сектора режекции помех $\Delta\phi \in [108°; 147°]$

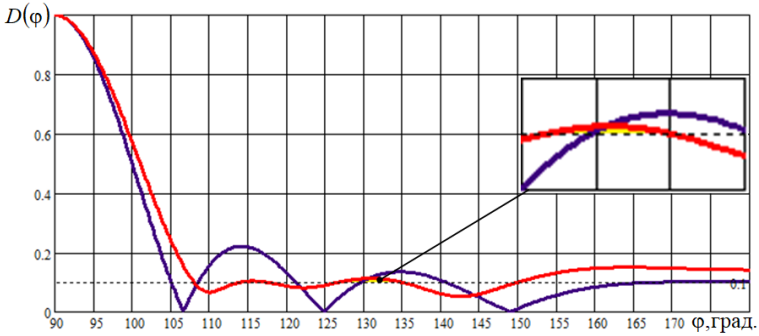


Рис. 3. Вид ДН антенной решетки при превышении заданного уровня боковыми лепестками в секторе режекции

Для представленных ДН величина секторного провала фактически определяется выбором углового положения третьей точки контроля.

Максимально достижимый сектор углов шириной 39°, где уровень ДН снижается до заданного предела, достигается при применении критерия (1).

Заключение

Обоснован критерий синтеза и найдены угловые положения точек контроля ДН антенных решеток для формирования широких секторов с подавлением передачи-приема при максимизации КНД.

По результатам моделирования установлено, что сектор режекции помех при формировании ДН по критерию (1) расширяется на 12° (с 27° до 39°) относительно значений, достижимых при последовательной коррекции уровней в области боковых лепестков.

Величина ослабления уровня помех в секторе режекции задается точками контроля ДН антенной решетки, расширение сектора углов со сниженной пространственно-частотной избирательностью передачи-приема сопровождается повышением числа направлений контроля. При этом сокращаются потери КНД антенной решетки, вызванные расширением главного луча, и рост дальних боковых лепестков ДН.

Список литературы

1. Бахрах, Л. Д. Синтез излучающих систем (теория и методы расчета) / Л. Д. Бахрах, С. Д. Кременецкий. – Москва : Радио и связь, 1974. – 232 с.

2. Воскресенский, Д. И. Развитие антенных систем (обзор) / Д. И. Воскресенский, В. М. Максимов // Известия вузов. Радиоэлектроника. – 1987. – Т. 30, № 2. – С. 4–15.

3. Кашин, В. А. Методы фазового синтеза антенных решеток / В. А. Кашин // Зарубежная радиоэлектроника. Успехи современной радиоэлектроники. – 1997. – № 1. – С. 47–60.

4. Неганов, В. А. Современная теория и практические применения антенн / В. А. Неганов, Д. П. Табаков, Г. П. Яровой. – Москва : Радиотехника, 2009. – 720 с.

5. Богословский, А. В. Критерий и методика синтеза антенных решеток для мобильных радиоэлектронных комплексов / А. В. Богословский, С. Н. Разиньков, Ю. В. Сторожук // Воздушно-космические силы. Теория и практика. – 2022. – № 24. – С. 81–91. – URL: http://академия-ввс.рф/images/data/zhurnal_vks/24-2022/81-91.pdf (дата обращения: 15.11.2025).

6. Разиньков, С. Н. Синтез нулей диаграмм направленности антенных решеток для излучения узкополосных и импульсных сигналов

с высокими коэффициентами направленного действия / С. Н. Разиньков, А. В. Богословский // Вестник Воронежского государственного технического университета. – 2012. – Т. 8, № 2. – С. 32=37.

7. Метод неопределенных множителей Лагранжа / В. И. Бахтин [и др.]. – Минск : БГУ, 2012. –40 с.

8. Гилат, А. MATLAB. Теория и практика / А. Гилат. – Москва : ДМК Пресс, 2016. – 416 с.

Нейросетевое моделирование процесса образования радиационных туманов

А. В. Соловьев, email: av-solovev@mail.ru
Г. Н. Бакаев, email: av-solovev@mail.ru
Б. В. Нененко, email: av-solovev@mail.ru

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *Представлен порядок построения прогностического способа возникновения радиационного тумана на основе нейросетевого моделирования и дискриминантного анализа. Выполнена оценка успешности построенных моделей с применением архивного материала метеорологической станции № 34122 (г. Воронежа). Для проведения данной оценки были использованы такие показатели успешности, как общая оправдываемость, критерии Обухова и Багрова. По полученным данным проведен отбор наилучшего подхода к прогнозированию тумана из предложенных.*

Ключевые слова: *радиационный туман, нейросетевое моделирование, нейрон, дискриминантная функция, температура воздуха, атмосферное давление, относительная влажность, скорость ветра.*

Введение

Наличие такого метеорологического явления на аэродроме базирования летательных аппаратов как туман характеризуется ухудшением метеорологической дальности видимости менее 1000 м [1, 2]. Данные значения видимости в значительной степени осложняют организацию и обеспечение безопасности полетов авиации. Метеорологические условия и параметры атмосферы, при которых ограничивается, либо запрещаются взлет и посадку воздушных судов, строго регламентированы требованиями руководящих документов и инструкций [3, 4].

Актуальность данной работы заключается в потребности автоматизации способов прогнозирования туманов для исключения ошибок, связанных с человеческим фактором, и снижением времени на разработку прогностического заключения об образовании опасного явления погоды.

© Соловьев А. В., Бакаев Г. Н., Нененко Б. В., 2026

1. Постановка задачи

Целью является автоматизация способов прогнозирования туманов на основе использования нейронных сетей, параметрического дискриминантного анализа и отбор наилучшего.

Для решении задач по снижению ошибок, связанных с человеческим фактором, а также уменьшением времени на разработку прогноза, было предложено применить дискриминантный анализ и нейросетевое моделирование с использованием ЭВМ.

Методы дискриминантного анализа применяются для оценки фазового состояния атмосферных процессов (наличие или отсутствие тумана). При решении задач дискриминации (классификации) и прогнозирования в настоящее время широкое применение в самых различных областях получили нейронные сети [5, 6]. Их обширные возможности позволяют успешно проводить разработку метеорологических прогнозов туманов. Нейросети позволяют получать различные зависимости между исходными и выходным сигналами, т.е., по сути, между предикторами и предиктантом. Преимуществом данного математического аппарата является возможность использования в тех случаях, когда неизвестен точный вид связей между входными и выходными параметрами, поскольку вид и параметры зависимости между «входом» и «выходом» находятся в процессе обучения сети.

2. Построение прогностической модели образования радиационного тумана

Для решения прогностической задачи с использованием нейросети и дискриминантного анализа была составлена архивная выборка по пункту Воронеж за период с 2024 по 2025 год с 16:00 до 19:00 объемом 200 случаев. Из общего количества полученной выборочной совокупности 27 раз наблюдался туман. В качестве предикторов, исходя из физических условий образования тумана, были взяты такие значения как относительная влажность f, температура воздуха T, атмосферное давление P, скорость ветра ff в приземном слое атмосферы. Выборочная совокупность была разбита на обучающую и контрольную по следующему принципу: первые два срока использовались для обучения, последующий срок применялся для проверки на адекватность полученных прогностических регрессионных моделей.

По обучающему материалу было выполнено построение дискриминантной функции и выполнена оценка качества результатов прогноза тумана. В табл. 1 показаны дискриминантные функции с одним предиктором.

Таблица 1

Дискриминантные функции с одним предиктором

Функция	Общая оправдываемость, *U*	Критерий Обухова, *H*	Критерий Багрова, *Q*
d=0,19–0,03*ff*	0,88	0	0,86
d=–1,88+0,002*P*	0,88	0	0,86
d=–0,5+0,01*f*	0,91	0,52	0,68
d=0,16–0,01*T*	0,87	0,31	0,7

Согласно показателям успешности, представленным в табл. 1, наилучшей определена функция с исходным параметром «относительная влажность», представленная в виде $d = -0,5 + 0,01f$.

Дискриминантные функции с двумя предикторами показаны в табл. 2.

Таблица 2

Дискриминантные функции с двумя предикторами

Функция	Общая оправдываемость, *U*	Критерий Обухова, *H*	Критерий Багрова, *Q*
d=–0,47+0,01*f*–0,002*T*	0,91	0,52	0,68
d=–0,42–0,01*f*–0,02*ff*	0,89	0,48	0,44
d=–4,6+0,01*f*+0,004*P*	0,88	0,51	0,39

С учетом данным табл. 2 наилучшей функцией было принято считать $d = -0,47 + 0,01f - 0,002T$.

Функции с тремя предикторами представлены в табл. 3.

Таблица 3

Дискриминантные функции с тремя предикторами

Функция	Общая оправдываемость, *U*	Критерий Обухова, *H*	Критерий Багрова, *Q*
d=–0,41+0,01*f*–0,01*T*–0,02*ff*	0,88	0,48	0,44
d=–0,46+0,01*f*+0,0002*T*+0,004*P*	0,88	0,51	0,39

Согласно представленным данным в табл. 3 показатели оценки успешности снизились, и в дальнейшем проведении процедуры отбора предикторов нет необходимости.

Для отобранных наилучших дискриминантных функций была выполнена апробация на контрольном материале. Результаты оценки успешности прогнозирования тумана представлены в табл. 4.

Таблица 4

Оценка дискриминантных функций на контрольном материале

Функция	Общая оправдываемость, U	Критерий Обухова, H	Критерий Багрова, Q
d=–0,5+0,01f	Обучающая выборка		
	0,91	0,52	0,68
	Контрольная выборка		
	0,85	0,52	0,09
d=–0,47+0,01f–0,002T	Обучающая выборка		
	0,91	0,52	0,68
	Контрольная выборка		
	0,91	0,49	0,47

Исходя из результатов в табл. 4 наилучшим образом представлялась возможность прогнозировать туман с помощью дискриминантной функции $d = -0,47 + 0,01f - 0,002T$, так как по контрольной выборке для неё были получены оптимальные численные значения показателей успешности.

Для построения дискриминантной функции с использованием нейросетевого моделирования была использована аналитическая платформа Deductor версии 5.3. В мастере обработки была построена нейросеть со стандартными настройками. В полученный персептрон были включены четыре входных параметра, представляющих фактические значения температуры воздуха у земли T, относительной влажности f, приземного атмосферного давления P, скорости ветра ff, один нейрон в промежуточном слое. Принцип работы функции активации применяемого нейрона основан на методе Монте-Карло. В качестве выходного параметра рассматривалось прогностическое заключение об образовании тумана или его отсутствии. Структура нейросети представлена на рис. 1.

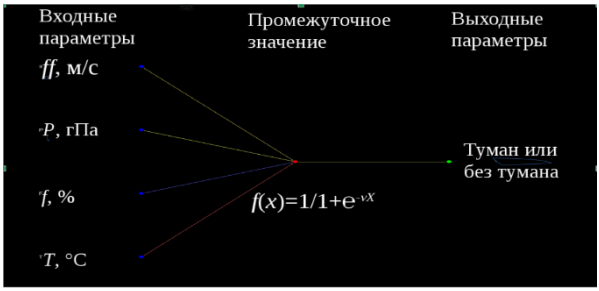


Рис. 1. Структурная схема нейронной сети, используемой при прогнозировании тумана

Согласно изображению рисунка наиболее тесная зависимость образования тумана прослеживалась от фактической температуры воздуха у земли и относительной влажности.

3. Оценка успешности прогностической модели образования радиационного тумана

Для оценки успешности прогнозирования тумана полученным способом в разделе визуализации были сформированы матрицы сопряженности по обучающему и контрольному материалу. Результаты расчетов представлены в виде табл. 5.

Таблица 5

Матрицы сопряженности результатов прогнозирования с помощью нейросети

Обучающая выборка			
Факт	Прогнозировалось		∑
	туман	без тумана	
туман	7	7	14
без тумана	5	115	120
∑	12	122	134
Контрольная выборка			
Факт	Прогнозировалось		∑
	туман	без тумана	
туман	3	2	5
без тумана	2	59	61
∑	5	61	66

По полученным данным матриц сопряженности табл. 5 были рассчитаны общая оправдываемость, критерии Обухова и Багрова. Полученные результаты отображены в табл. 6.

Таблица 6

Показатели успешности прогнозирования с помощью нейросети

№	Название показателя успешности	Значение показателя успешности
1	Общая оправдываемость, *U*	0,939
2	Критерий Багрова, *H*	0,567
3	Критерий Обухова, *Q*	0,567

Согласно представленной информации в табл. 6 при использовании нейросетевого моделирования были достигнуты высокие результаты. При этом стоит отметить, что процесс обучения может происходить непрерывно и построенная модель будет совершенствоваться и далее. Прогностические заключения о тумане могут быть получены на основе результата, полученного нейросетью без использования аналитического подхода метеоспециалиста.

Для проведения сравнительного анализа наилучшие результаты оценки прогнозирования с использованием дискриминантного анализа и нейросети сведены в общую табл. 7.

Таблица 7

Показатели успешности прогнозирования с помощью нейросети и дискриминантного анализа

№	Название показателя успешности	Дискриминантный анализ	Нейронная сеть
1	Общая оправдываемость, *U*	0,91	0,94
2	Критерий Багрова, *H*	0,49	0,57
3	Критерий Обухова, *Q*	0,47	0,57

После проведенного анализа данных табл. 7 определено, что наилучшим образом прогнозирование тумана выполнялось с использованием нейронной сети.

Заключение

В ходе выполненного исследования были получены прогностические модели тумана на основе дискриминантных функций и нейронной сети. С использованием показателей успешности была выполнена их оценка. Наилучшим образом получалось выполнять прогнозирование тумана с использованием нейросети (общая оправдываемость – 94 %, критерии Обухова и Багрова – 0,57).

Преимуществом данного подхода является исключение субъективного влияния на прогностическое заключение, что в свою очередь снижает ошибку человеческого фактора.

При значительном увеличении объема обучающего материала построенная нейронная сеть может быть использована в практической деятельности метеорологических подразделений при разработке прогноза возникновения тумана.

Список литературы

1. Воробьёв, В. И. Синоптическая метеорология : учеб. пособие / В. И. Воробьёв. – Л. : Гидрометеоиздат, 1991. – 540 с.

2. Зверев, А. С. Синоптическая метеорология : учеб. пособие / А. С. Зверев. – Л. : Гидрометеоиздат, 1997. – 504 с.

3. Руководство по практическим работам метеорологических подразделений авиации Вооруженных Сил [введено приказом ГК ВВС РФ от 8 сент. 1992 г. № 186]. – М. : Воениздат, 1992. – 486 с.

4. Дорофеев, В. В. Авиационная метеорология : учеб. пособие / В. В. Дорофеев, И. Е. Кузнецов, А. В. Степанов, Д. В. Черепанов. – Воронеж : ВВАИИ, 2021. – 274 с.

5. Скирда, И. А. Авиационные прогнозы погоды : учеб. пособие / И. А. Скирда, В. И. Садовский, В. А. Мозиков. – М. : Воен. изд-во, 1995. – 424 с.

6. Ульшин, И. И. Методы статистической обработки гидрометеорологической информации : учеб. пособие / И. И. Ульшин. – Воронеж : ВУНЦ ВВС «ВВА», 2016. – 178 с.

Обзор лагранжевых моделей переноса частиц (LPDM) и их вычислительных особенностей

Р. Т. Туркменова. e-mail: rukhiyet@gmail.com

НИИ развития цифровых технологий и искусственного интеллекта, Узбекистан

Аннотация. *Лагранжевые модели переноса частиц широко применяются для моделирования атмосферного переноса примесей, построения связей «источник-рецептор» и решения обратных задач (оценки источников/потоков) по данным наблюдений. В работе дан обзор теоретической основы LPDM как стохастических моделей траекторий при соблюдении критерия хорошо перемешанной смеси, а также рассмотрены распространенные программные комплексы. Основное внимание уделено вычислительным особенностям: масштабируемости по числу частиц и шагов интегрирования, организации ввода-вывода метеорологических полей, режимам прямого и обратного моделирования, построению матриц чувствительности и типовым подходам к параллелизации.*

Ключевые слова: *атмосферный массоперенос; лагранжева модель; стохастические траектории; FLEXPART; HYSPLIT; STILT; NAME; backward-режим; матрица «источник-рецептор»; инверсия; высокопроизводительные вычисления*

Введение

Перенос и рассеивание атмосферных примесей традиционно формулируются в эйлеровом виде через уравнение переноса (адвекция-диффузия-реакции):

$$\frac{\partial C}{\partial t} + \nabla \cdot (uC) = \nabla \cdot (K \nabla C) + S - \Lambda C$$

Такой подход естественен при сеточных расчетах концентраций, но в задачах «источник-рецептор» и особенно в обратных постановках удобнее лагранжева интерпретация: примесь представляется ансамблем частиц (пакетов массы), а турбулентное перемешивание моделируется стохастически. В этом случае траектория частицы может быть записана как стохастическое дифференциальное уравнение, а соответствующая эволюция плотности вероятности (и, при интерпретации веса частицы, поля C) связана с уравнением Фоккера-Планка, что обеспечивает

© Туркменова Р. Т. 2026

теоретическую согласованность лагранжевого и эйлерового описаний при корректной параметризации турбулентности (в частности, при выполнении критерия well-mixed) [1-3].

Для диссертационных исследований по стохастическому моделированию и инверсии источников Лагранжевые модели переноса частиц (Lagrangian Particle Dispersion Models, LPDM) выступают вычислительным ядром, которое связывает метеорологические данные и физику переноса с вероятностной постановкой обратной задачи. Поэтому важно учитывать вычислительные ограничения (I/O, интерполяция полей, параллелизация) и статистический шум Монте-Карло, влияющий на качество оценок.

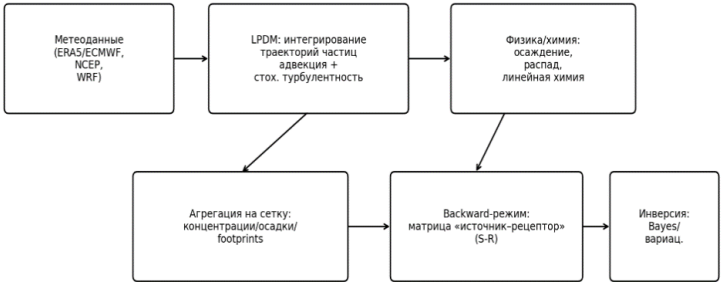


Рис. 1. Типовой вычислительный конвейер LPDM и связь с инверсией

1. Теоретическая основа LPDM: стохастические траектории и критерий хорошо перемешанной смеси (well-mixed)

В основе LPDM лежит представление переноса как суммы детерминированной адвекции и стохастического перемешивания. Типовая модель траектории записывается в виде стохастического дифференциального уравнения:

$$dX(t) = u(X,t)dt + B(X,t)dW(t)$$

где X – положение частицы, u – поле скорости, $W(t)$ – винеровский процесс, матрица B связана с параметризацией турбулентной диффузии.

На практике применяются случайного блуждания (random walk) для координат и случайного полёта (random flight) с учетом временных масштабов Лагранжа и свойств пограничного слоя.

Ключевое требование корректности – критерий хорошо перемешанной смеси: при отсутствии источников модель не должна создавать искусственные градиенты примеси, если начальное распределение согласовано со стационарным [1-3]. Нарушение этого

требования ведет к систематическим ошибкам, которые затем «проникают» в поля чувствительности (footprints) и матрицу H, ухудшая устойчивость инверсии.

Дополнительно LPDM включают процессы потерь и трансформаций (сухое/влажное осаждение, распад, простую химию), часто реализуемые изменением веса частицы. С вычислительной точки зрения это увеличивает стоимость шага и требует контроля сохранения массы и устойчивости дискретизации.

2. Программные комплексы LPDM и режимы моделирования

Наиболее распространенные семейства LPDM включают гибкую лагранжеву модель переноса/дисперсии частиц - FLEXPART [4-5], гибридную модель: лагранжевы траектории/дисперсия одиночных частиц с интегрированием траекторий - HYSPLIT [6-7], стохастическую лагранжеву модель переноса с обращением времени (backward) - STILT [8-9] и численную среду моделирования атмосферной дисперсии/рассеивания - NAME [10]. Их общая вычислительная схема сходна (интегрирование траекторий плюс агрегация на сетку), однако различаются поддерживаемые физические процессы, форматы метеовхода и ориентация на задачи (оперативный прогноз/исследовательские эксперименты/инверсия).

Прямой (forward) режим решает задачу «источник → поле концентраций/осадков» и естественен для сценарного анализа и оперативного прогноза. Обратный (backward) режим решает задачу «рецептор → область влияния» и особенно эффективен для обратных задач: один расчет дает поля чувствительности, которые затем линейно комбинируются с различными картами эмиссий без повторного интегрирования траекторий [11].

Таблица

Сравнение распространенных LPDM (обобщение по официальным публикациям)

№	Модель	Режимы	Ключевые возможности	Вычислительные особенности
1	FLEXPART	Прямой/обратный	трассеры, осаждение, распад; поле чувствительности	масштабируется по частицам; критичен I/O метеополей

Продолжение таблицы

2	HYSPLIT	траектории/дисперсия	оперативные расчеты, широкий набор архивов	удобная экосистема подготовки данных, регулярные обновления
3	STILT	в основном обратный	Поле чувствительности для станций; инверсия потоков	рецепторно-ориентированная постановка снижает стоимость
4	NAME	прямой/инверсный	оперативный прогноз; оценка эмиссий	фокус на устойчивом конвейере и времени расчета

3. Вычислительные особенности: стоимость, I/O и параллелизация

Вычислительная сложность LPDM, как правило, пропорциональна числу частиц N_p и числу шагов интегрирования N_t:

$$O(N_p \cdot N_t)$$

Погрешность Монте-Карло убывает как

$$\sim \frac{1}{\sqrt{N_p}}$$

поэтому уменьшение шума требует существенного роста N_p. На практике это задает компромисс между разрешением по времени/пространству и вычислительным бюджетом.

На больших N_p узким местом часто становится не арифметика, а интерполяция 4D-метеополей и ввод-вывод. Для ускорения используются кэширование временных слоев, предварительная нарезка домена и стандартизация форматов результатов (например, стандартный формат научных данных, часто упоминаемый как NetCDF). Качество метеовхода критично: для региональных исследований обычно применяют реанализ ERA5 (**глобальный реанализ погоды** 5-го поколения от ECMWF) [12] или поля WRF (численная модель прогноза погоды) [13], а для сопряжения WRF и LPDM распространены специализированные связки, например FLEXPART-WRF [14].

LPDM хорошо распараллеливаются, поскольку траектории частиц независимы до этапа агрегации. Однако накопление концентраций/полей чувствительности требует аккуратной организации редукций (приватные буферы потоков, пакетная агрегация) для минимизации коллизий и повышения масштабируемости на многопроцессорных системах.

4. Практический пример (воспроизводимая реализация на базе официальной документации Metview для FLEXPART)

В качестве опорного примера практической реализации LPDM (с кодом и графической визуализацией результата) целесообразно использовать официальный учебный сценарий Metview (ECMWF/INPE) «FLEXPART - Backward simulation with residence times». В нём моделирование выполняется в обратном режиме: частицы выпускаются из области рецептора назад во времени, а выходной продукт формируется в виде полей времени пребывания частиц (residence time) на сетке. Такие поля интерпретируются как поля влияния/чувствительности рецептора (footprints) и напрямую используются при построении матриц «источник-рецептор» для задач инверсии. Реализация включает типовой воспроизводимый конвейер: подготовку входных метеополей, задание условий выпуска (receptor/release), запуск FLEXPART через flexpart_run() и последующую визуализацию на карте средствами Metview [15-16].

Поле времени пребывания частиц в backward-LPDM можно интерпретировать как интеграл по времени от «плотности присутствия» частиц в точке x (обычно в заданном слое, например 0-500 м):

$$RT(x) = \int_{t_0 - T}^{t_0} \rho(x,t)dt$$

где t_0 - момент измерения в рецепторе, T - длительность обратного расчёта, $\rho(x,t)$ - нормированная плотность (занятость) частиц.

В вычислительной реализации LPDM это поле оценивается методом Монте-Карло как суммарное время пребывания частиц в ячейке сетки V_i :

$$RT_i \approx \frac{1}{N_p} \sum_{p=1}^{N_p} \sum_{n=1}^{N_t} 1\{X_p(t_n) \in V_i\} \Delta t,$$

где N_p - число частиц, N_t - число шагов по времени, $1\{\cdot\}$ - индикатор попадания частицы в ячейку V_i .

Таким образом, карта RT_i является практической формой поля влияния/чувствительности рецептора (footprint), используемой для построения матрицы «источник-рецептор» в задачах инверсии.

Для программной реализации FLEXPART из Metview (backward) и построения поля времени пребывания частиц используется язык программирования Python:

Шаг 1. Подготовка входных данных FLEXPART (flexpart_prepare). На этом этапе подготавливаются метеорологические входные данные для модели FLEXPART. В качестве исходного файла используется метеорологический GRIB-файл (например, meteo.grib), содержащий поля ветра, температуры, давления, параметры планетарного пограничного слоя (PBL) и другие необходимые метеорологические переменные на серии временных сроков. Функция flexpart_prepare формирует структуру входной директории, соответствующую требованиям модели FLEXPART. Модель не может напрямую использовать произвольный GRIB-файл, поскольку ей требуется строго определённая организация входных данных и список доступных временных сроков.Результатом выполнения данного шага является объект inp, который представляет собой описание подготовленного набора входных данных. Этот объект затем используется на этапе запуска модели через flexpart_run.Таким образом, данный этап обеспечивает корректную организацию входной информации и воспроизводимость вычислений.

Шаг 2. Задание рецептора/выпуска частиц (flexpart_release). На втором этапе формируется описание выпуска частиц с использованием функции flexpart_release. В среде Metview объект release используется для задания параметров рецептора и характеристик моделируемых частиц.

Определяются следующие параметры:

- географическая область рецептора (прямоугольная область по широте и долготе);
- высоты выпуска, заданные относительно поверхности земли (Above Ground Level, AGL);
- вертикальный диапазон рецептора, например слой 0–500 м, что позволяет анализировать footprint для приземного слоя атмосферы;
- количество выпускаемых частиц, которое влияет на статистическую устойчивость результата. Большее число частиц уменьшает статистический шум и делает поле footprint более гладким, однако, увеличивает вычислительные затраты.
- масса (вес) частицы, используемая для нормировки результатов. В задачах оценки времени пребывания основную роль играют

траектории частиц и время их нахождения в различных областях.

Шаг 3. Запуск модели FLEXPART в обратном режиме (flexpart_run) На третьем этапе выполняется запуск модели FLEXPART с использованием функции flexpart_run. Модель использует подготовленные метеорологические данные из шага 1 и параметры выпуска частиц из шага 2. Ключевой особенностью данного этапа является использование обратного (backward) режима моделирования.

В этом режиме частицы распространяются назад во времени от рецептора, что позволяет определить области, в которых воздушные массы находились до достижения точки наблюдения. На основе траекторий рассчитывается время пребывания частиц в различных ячейках пространственной сетки.

Также задаются параметры выходного поля:

– пространственная область расчёта, обычно значительно более широкая, чем область рецептора, чтобы определить возможные источники воздушных масс;
– шаг сетки по широте и долготе (например, 0.25° × 0.25°). уменьшение шага сетки повышает пространственное разрешение карты, но увеличивает объём данных и вычислительные затраты.

результатом расчёта является поле времени пребывания частиц, сохранённое в формате GRIB.

Шаг 4. Визуализация результатов На заключительном этапе выполняется визуализация результатов моделирования. Для этого считывается выходной файл модели, например time_s001.grib, содержащий поле времени пребывания частиц на расчётной сетке. Далее выполняются следующие действия:

– настройка географической области отображения карты;
– задание границ карты (широта и долгота);
– выбор способа визуализации поля.

Что в итоге такая позволяет визуализировать и определить регионы, оказывающие наибольшее влияние на формирование состава воздуха в заданном рецепторе (рис. 2).

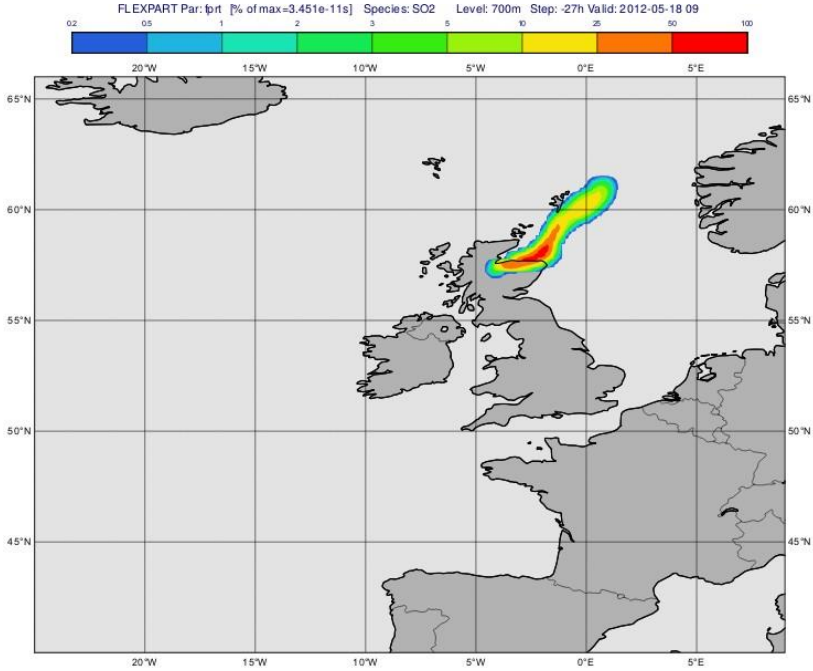


Рис. 2. Пример визуализации поля времени пребывания частиц в ***backward-LPDM (FLEXPART через Metview)***

Поле время пребывания частиц интерпретируется как поле влияния/чувствительности рецептора (footprint) и используется для формирования матрицы «источник-рецептор» в задачах инверсии. Иллюстрация воспроизводима по официальной документации Metview; документация распространяется по лицензии Apache License 2.0 [15-17].

Заключение

LPDM являются одним из наиболее гибких и вычислительно эффективных классов моделей атмосферного переноса, особенно при построении связей «источник-рецептор» и решении обратных задач. Для практической ценности расчетов решающее значение имеют вычислительные детали: организация метеовхода и интерполяции, выбор числа частиц и шага, устойчивое накопление концентраций/полей чувствительности и параллелизация. Учет этих факторов позволяет планировать диссертационные эксперименты с контролируемой неопределенностью и воспроизводимостью результатов.

Список литературы

1. Thomson, D. J. Criteria for the selection of stochastic models of particle trajectories in turbulent flows / D. J. Thomson // Journal of Fluid Mechanics. – 1987. – Vol. 180. – P. 529-556.
2. Wilson, J. D. Review of Lagrangian stochastic models for trajectories in the turbulent atmosphere / J. D. Wilson, B. L. Sawford // Boundary-Layer Meteorology. – 1996. – Vol. 78. – P. 191-210. – DOI: 10.1007/BF00122492.
3. Rodean, H. C. Stochastic Lagrangian models of turbulent diffusion / H. C. Rodean. –Boston : American Meteorological Society, 1996. – 84 p.
4. Pisso, I. The Lagrangian particle dispersion model FLEXPART v10.4 / I. Pisso, E. Sollum, H. Grythe [et al.] // Geoscientific Model Development. – 2019. – Vol. 12. – P. 4955-4997. – DOI: 10.5194/gmd-12-4955-2019. – URL: https://gmd.copernicus.org/articles/12/4955/2019.
5. Stohl, A. Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2 / A. Stohl, C. Forster, A. Frank, P. Seibert, G. Wotawa // Atmospheric Chemistry and Physics. – 2005. – Vol. 5. – P. 2461-2474. – DOI: 10.5194/acp-5-2461-2005. – URL: https://acp.copernicus.org/articles/5/2461/2005.
6. Stein, A. F. NOAA's HYSPLIT atmospheric transport and dispersion modeling system / A. F. Stein, R. R. Draxler, G. D. Rolph [et al.] // Bulletin of the American Meteorological Society. – 2015. – Vol. 96. – P. 2059–2077. – DOI: 10.1175/BAMS-D-14-00110.1. – URL: https://journals.ametsoc.org/view/journals/bams/96/12/bams-d-14-00110.1.xml.
7. HYSPLIT model overview / NOAA Air Resources Laboratory [Электронный ресурс]. – URL: https://www.ready.noaa.gov/HYSPLIT.php.
8. A near-field tool for simulating the upstream influence of atmospheric observations: The stochastic time-inverted Lagrangian transport (STILT) model / J. C. Lin [et al.] // Journal of Geophysical Research: Atmospheres – 2003. – Vol. 108, No. D16. – Art. 4493. – DOI: 10.1029/2002JD003161.
9. STILT project site (uataq) [Электронный ресурс]. – URL: https://uataq.github.io/stilt/
10. Numerical atmospheric-dispersion modelling environment (NAME): overview / Met Office [Электронный ресурс]. – URL: https://www.metoffice.gov.uk/research/approach/modelling-systems/dispersion-model.
11. Seibert, P. Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode / P. Seibert, A. Frank // Atmospheric Chemistry and Physics. – 2004 – Vol. 4. – P. 51–63. – DOI: 10.5194/acp-4-51-2004.

12. The ERA5 global reanalysis / H. Hersbach [et al.] // Quarterly Journal of the Royal Meteorological Society – 2020. – Vol. 146, No. 730. – P. 1999-2049. – DOI: 10.1002/qj.3803.

13. A description of the Advanced Research WRF Model version 4 / W. C. Skamarock [et al.]. – NCAR Technical Note. – 2019. – URL: https://www2.mmm.ucar.edu/wrf/users/docs/technote/v4_technote.pdf (дата обращения: 28.12.2025).

14. The Lagrangian particle dispersion model FLEXPART-WRF version 3.1 / J. Brioude, [et al.] // Geoscientific Model Development. – 2013. – Vol. 6. – P. 1889–1904. – DOI: 10.5194/gmd-6-1889-2013. – URL: https://gmd.copernicus.org/articles/6/1889/2013.

15. FLEXPART – backward simulation with residence times / Metview documentation (ECMWF & INPE) [Электронный ресурс]. – URL:https://metview.readthedocs.io/en/latest/tutorials/flexpart/flexpart_backward_simulation_with_residence_times.html.

16. flexpart_run(): runs the FLEXPART Lagrangian particle dispersion model and converts its outputs to formats that Metview can work with / Metview documentation (ECMWF & INPE) [Электронный ресурс]. – URL:https://metview.readthedocs.io/en/5.25.0/gen_files/icon_functions/flexpart_run.html.

17. Licence (Apache Licence Version 2.0) / Metview documentation (ECMWF & INPE) [Электронный ресурс]. – URL: https://metview.readthedocs.io/en/latest/licence.html.

Структура информационной системы сбора и обработки данных для систем обратной связи в миоэлектрических протезах верхней конечности

Я. А. Туровский, email: Yaroslav_turovsk@mail.ru
С. С. Крупенин, email: semyon.krupenin@bk.ru

Воронежский государственный университет

Аннотация. *В данной работе рассматривается программно-аппаратный комплекс, обеспечивающий управление актуаторами, которые передают тактильное воздействие на кожу человека, и сбор экспериментальных данных для их последующего статистического анализа, а также затрагивается исследование опытного образца системы тактильной обратной связи для протеза верхней конечности.*

Ключевые слова: *система обратной связи, программно-аппаратный комплекс, анализ данных, миоэлектрический протез.*

Введение

Функциональность бионических протезов верхних конечностей крайне ограничена отсутствием сенсорной обратной связи, что снижает автономность пользователя и естественность управления. Решение данной задачи требует разработки программно-аппаратного обеспечения, способного собирать и анализировать информацию, полученную с датчиков на протезе, и эффективно передавать ее на кожу человека [1].

В данной работе рассматривается программно-аппаратный комплекс, обеспечивающий управление актуаторами, которые передают тактильное воздействие на кожу человека, и сбор экспериментальных данных для их последующего статистического анализа.

В рамках исследования был разработан опытный образец системы тактильной обратной связи для протеза верхней конечности. Для апробации и анализа разработанной системы было проведено испытание опытного образца на здоровых людях, в котором изучалась точность восприятия тактильной обратной связи верхней конечностью человека [2].

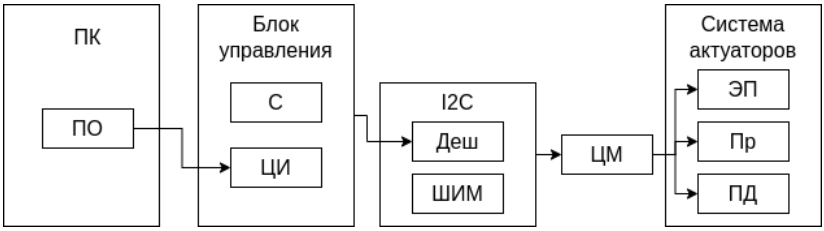
Целью исследования является разработка информационной системы для сбора и обработки данных, полученных в ходе апробации системы обратной связи протезов верхних конечностей.

© Туровский Я. А., Крупенин С. С., 2026

1. Реализация программно-аппаратного комплекса системы обратной связи миоэлектрического протеза

Для передачи информации система обратной связи использует тактильные рецепторы человека в качестве “входа” [3]. Программно-аппаратный комплекс системы обратной связи работает следующим образом: когда пользовательское приложение подает команды на изменение состояния, то информация о новом состоянии актуаторов передается на микроконтроллер, где обрабатывается и передается на систему актуаторов – тактильных стимуляторов, расположенных на носимой пользователем манжете.

Сам же комплекс системы обратной связи состоит из набора элементов, выполняющих различные функции. Схема структуры программно-аппаратного комплекса системы обратной связи миоэлектрического протеза показана на рис. **Ошибка! Источник ссылки не найден.**.



ДР - драйвер питания; ЦИ - цифровой интерфейс; С - сумматор; ШИМ - контроллер широтно-импульсной модуляции; ЭП - элемент Пельтье; ЦМ - цифровой мультиплексор; Пр - электромеханический привод; ПД - привод давления; ПО - программное обеспечение

Рис. 1. ДЕШ - дешифратор, задающий адрес

Испытательный стенд состоит из проксимальной и дистальной манжеты с находящимися на них актуаторами, которые представляют собой сервоприводы и передают на кожу испытуемого тактильное воздействие. Дистальная манжета располагается ближе к кисти и состоит из актуаторов 1-5. Отсчет актуаторов начинается с плечелучевой мышцы на предплечье, актуаторы расположены по всему обхвату предплечья. Проксимальная манжета располагается ближе к плечу, состоит из актуаторов 6-10, расположенных также по всему охвату. Схема расположения актуаторов на руке представлена ниже на рис. **Ошибка! Источник ссылки не найден.**. Круглым значком указаны видимые актуаторы, прямоугольным – находящиеся на обратной стороне руки.

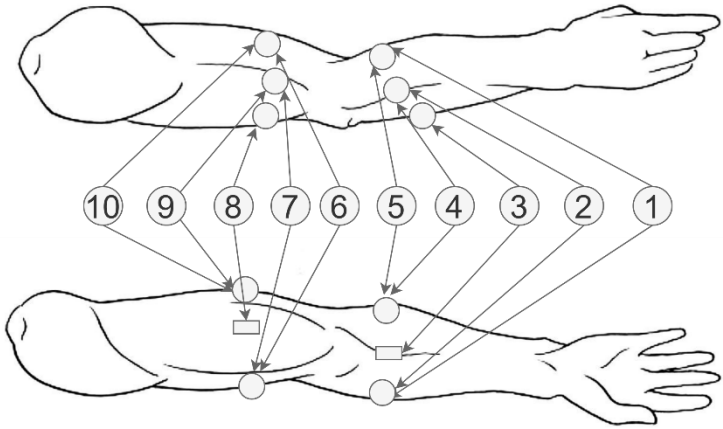


Рис. 2. Схема расположения актуаторов

2. Технологическое программное обеспечение для сбора информации

Исследование включает в себя два этапа: калибровку и тестирование, которые осуществляются с помощью технологического ПО для взаимодействия актуаторов и сбора данных. Обмен информацией происходит по последовательному интерфейсу UART.

Калибровка включает в себя поочередное включение актуаторов, которые оказывают нарастающее воздействие на кожу испытуемого, а он, в свою очередь, отмечает, ощущается ли воздействие. Полученные данные записываются для последующего тестирования и анализа данных.

Тестирование включает в себя срабатывание случайного актуатора, отклоняющегося на случайное положение, не выходящее за границы чувствительности, полученные в ходе калибровки. Задача испытуемого – определить, какой актуатор отклонился и какова сила воздействия по 100-балльной шкале. После прохода всех актуаторов тестирование заканчивается и полученные данные сохраняются в файл для последующего анализа.

Рабочее окно приложения, используемого при исследовании, представлено на рис. **Ошибка! Источник ссылки не найден.**.

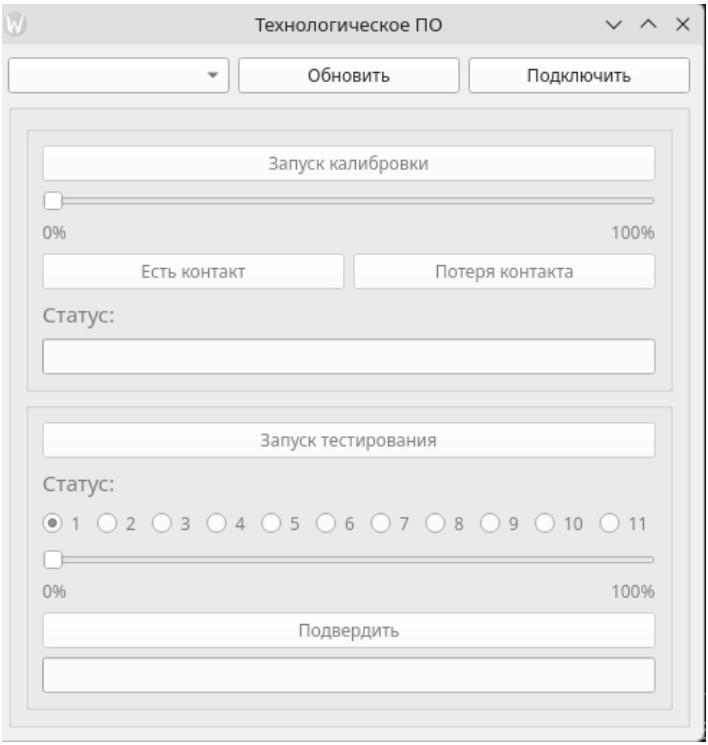


Рис. 3. Рабочее окно технологического ПО

3. Анализ полученных данных

Перед тестированием проводилась калибровка актуаторов, которая заключалась в измерении верхнего и нижнего порога силы давления актуатора на кожу. Начиная с нижнего порога, человек начинает ощущать прикосновение актуатора, а верхний порог показывает либо предел давления актуатора, либо момент, с которого давление становится болезненным для человека. Результаты калибровки актуаторов показаны на рис. **Ошибка! Источник ссылки не найден.**.

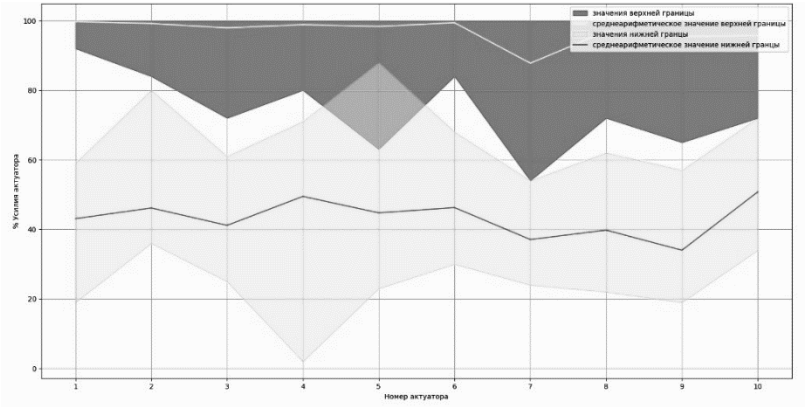


Рис. 4. Результат калибровки

Анализ полученных данных проводился в программном пакете Statistica [4-5]. По данным, полученным в ходе исследования, была рассчитана корреляция по Спирмену [6] и проведен факторный [7] и кластерный анализ [8].

В результате анализа были сделаны следующие выводы. Во-первых, распознавание актуаторов дистальной манжеты хуже, чем проксимальной, особенно для актуаторов 1-4. Необходимо изменить расположение актуаторов с учетом их корреляций.

Во-вторых, при тестировании человек со временем утрачивает точность определения актуаторов из-за усталости, но сильнее это сказывается на точности распознавания актуаторов 1-4.

В-третьих, для актуаторов проксимальной манжеты (актуаторы 6-10) характерна высокая точность определения актуатора, но низкая точность определения точности силы давления актуаторов. То же самое характерно для актуатора 5 в составе дистальной манжеты.

Заключение

В ходе исследования была проведена апробация систем обратной связи миоэлектрического протеза верхней конечностей и разработана информационная система для сбора и обработки полученных данных. Также был проведен анализ результатов, полученных при апробации данной информационной системы на здоровых людях с использованием различных методов, которые включают в себя: кластерный анализ, факторный анализ и корреляцию полученных данных.

Список литературы

1. История протезирования / С. В. Сергеев [и др.] // Медико-социальные проблемы инвалидности. – 2017. – № 2. – С. 84–89.

2. Реализация системы обратной связи для миоэлектрических протезов верхней конечности / Я. А. Туровский [и др.] // Программная инженерия. – 2025. – Т. 16, № 1. – С. 39–46. – DOI: 10.17587/prin.16.39-46.

3. Оценка освоения "проприоцептивно"-тактильного канала обратной связи протезов верхней конечности / Я. А. Туровский [и др.] // Сенсорные системы. – 2017. – Т. 31, № 4. – С. 322-330.

4. Программный пакет Statistica [Электронный ресурс]. – Режим доступа: https://statsoft-statistica.ru

5. Электронный учебник по статистике [Электронный ресурс]. – Режим доступа: http://www.statsoft.ru/home/textbook/default.htm

6. Коэффициент ранговой корреляции Спирмена [Электронный ресурс]. – Режим доступа: https://medstatistic.ru/methods/methods9.html

7. Факторный анализ [Электронный ресурс]. – Режим доступа: https://sky.pro/wiki/profession/faktornyj-analiz-v-statistica

8. Кластерный анализ [Электронный ресурс]. – Режим доступа: https://sky.pro/wiki/profession/klasternyj-analiz-v-statistica

Прогнозирование количественных характеристик лесных пожаров на основе нейросетевых технологий

И. И. Ульшин, email: ulshin@rambler.ru
П. Д. Семенов

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *Рассмотрена разработка способов прогнозирования количества лесных пожаров и занимаемой ими площади в теплую половину года на территории Красноярского края на основе метеорологических параметров. Проведены оценивание и сравнительный анализ полученных регрессионных зависимостей, а также искусственных нейронных сетей*

Ключевые слова: *лесные пожары, метеорологические факторы, регрессионный анализ, искусственные нейронные сети.*

Введение

Одной из важных народнохозяйственных проблем в настоящее время остается проблема борьбы с лесными пожарами. Не случайно Президент России поставил перед профильными ведомствами конкретную цель – до 2030 года площадь лесных пожаров в России должна сократиться вдвое относительно уровня 2021 года, когда выгорело рекордное количество леса – 19 млн. га. Существенные усилия, которые предпринимают сейчас региональные власти в рамках нацпроекта «Экология», направлены на достижение именно этой стратегической цели [1].

Причины лесных возгораний известны. Они включают в себя как природные (малоснежные зимы, длительные засухи, высокая температура воздуха, грозы и т.д.), так и антропогенные (умышленное или неосторожное обращение с огнем) факторы. Видно, что в первом случае основными являются метеорологические параметры. Однако и в тех случаях, когда огонь в лес приносит человек, погода также способна существенно повлиять на дальнейшее развитие событий. Поэтому достаточно давно учеными предложен целый ряд показателей или индексов пожароопасности, основанных именно на учете атмосферных характеристик. К ним относятся формула В.Г. Нестерова, показатель В.А. Жданко и И.С. Мелехова, коэффициент Л.И. Сверловой, индекс засушливости KBDI и др. [2-5].

© Иванов И. И., Петров П. П., Сидоров С. С., 2026

Указанные показатели отлично характеризуют степень пожароопасности растительных насаждений. Разработаны даже специальные шкалы, позволяющие переходить от количественных значений к качественным характеристикам [4]. Однако с их помощью практически невозможно разработать краткосрочный прогноз возникновения очагов пожаров, их количества или площади.

Для решения задачи непосредственного прогнозирования разработаны специальные выражения, позволяющие предсказывать число пожароопасных дней в месяце $N_{мес.}$ и общую площадь выгоревшего леса $S_{мес.}$:

$$N_{мес.} = b_t t_{ср} + b_R R + a$$

где $t_{ср}$ – среднемесячная температура; R – месячная сумма осадков; b_t, b_R, a – коэффициенты;

$$S_{мес.} = (0,639N + 15,3636)^4$$

где N – число пожароопасных дней [4].

Следует заметить, что и данные выражения также не позволяют разрабатывать краткосрочные прогнозы пожаров и их количественных параметров. Кроме того, существующие методы прогнозирования пока не обладают стопроцентной оправдываемостью [6].

1. Постановка задачи

Поскольку надежные количественные оценки влияния тех или иных метеорологических параметров на число очагов лесных пожаров и занимаемую им площадь пока не получены, целью работы является выбор статистического метода и разработка на его основе способов прогнозирования возникновения и площади очагов горения в лесных массивах.

Достижение поставленной цели требует решения следующих очевидных задач:

- выбор региона исследования и организация архивных материалов, содержащих результаты наблюдений за лесными пожарами и метеорологическими величинами;
- выбор метода построения прогностических способов;
- получение прогностических способов на основе архивных материалов;
- оценка успешности полученных способов на основе критериев, предусмотренных руководящими документами.

Для выбора субъекта Российской Федерации, наблюдения по которому использовались при проведении настоящего исследования,

были проанализированы статистические данные о пожароопасности тех или иных регионов (см. рис. 1).



Рис. 1. Наиболее пожароопасные регионы Российской Федерации в 2022 году [1]

С учетом количества пожаров и плотности населения в качестве рассматриваемого региона был выбран Красноярский край.

2. Построение прогностических уравнений регрессии

Исходя из цели работы, была поставлена задача получить способы прогнозирования количества очагов пожара и площади, занимаемой ими, с учетом метеорологических факторов.

В качестве первого из «инструментов» был выбран регрессионный анализ. Для разработки прогностических способов были использованы данные о числе и площади пожаров в Красноярском крае за период с апреля по август 2018–2020 гг. Архивная выборка содержала также результаты синхронных наблюдений за погодой в указанный период. Объем выборки составил 366 случаев, в которые вошли наблюдения с пожарами и без них. Способы оценивались с помощью квадрата множественного коэффициента корреляции – коэффициента детерминации, а также двух известных показателей успешности прогнозов – средней абсолютной и средней квадратической ошибок.

На первом этапе был составлен предварительный перечень предикторов и проведен анализ степени их взаимосвязи с прогнозируемыми величинами. В качестве предикторов использовались: накопленные за предыдущие i дней значения температуры воздуха в 13 часов t_i, i =1, …, 8; значения накопленного количества осадков h_i, i =1, …, 8; величина накопленной влажности F_i, i =1, …, 8. В качестве характеристики взаимной зависимости метеопараметров использовался парный коэффициент корреляции между предиктантами и предикторами. Его значения использовались для определения лучших исходных величин и сокращения числа предикторов.

Как показали результаты расчетов, существенной корреляции между характеристиками пожаров и рассматриваемыми метеовеличинами не наблюдается. Лучшими предикторами оказались характеристики термического режима (коэффициенты корреляции от 0,238 до 0,402), значительно хуже – данные об осадках (коэффициенты корреляции от 0,088 до -0,003) и влажности (от -0,194 до -0,267).

Уравнения регрессии для прогнозирования количества очагов пожара и их площади были получены с помощью известных электронных таблиц. К примеру, уравнения с одним предиктором для числа очагов и общей площади имеют вид:

$$C_{оч.} = -1,066 + 0,065t_1$$

$$S_{общ.} = -336,98 + 5,227t_6$$

Показатели качества данных выражений недостаточно высоки. Коэффициент детерминации выражения (3) равен 0,125, средняя абсолютная ошибка прогноза равна 7,8, а средняя квадратическая – 10,8. У выражения (4) показатели успешности немного ниже. При увеличении числа предикторов успешность предсказания числа возгораний и их площади несколько увеличивается, но недостаточно для того, чтобы признать результаты удовлетворительными. Поэтому на следующем этапе работы был задействован более современный статистический метод.

3. Получение прогностических способов на основе искусственных нейронных сетей

При выполнении данной работы планировалось использовать для разработки прогнозов и нейросетевые технологии, являющиеся одним из самых удобных механизмов для этой цели. Основная идея, положенная в основу разработки искусственных нейронных сетей (ИНС), заключается в программно-аппаратной реализации модели архитектуры человеческого мозга. Поэтому описание возможностей ИНС часто

сопровождается использованием термина «искусственный интеллект» . Данное метафорическое понятие, возможно, является не очень удачным, однако, его следует признать общеизвестным и практически общепринятым. Широкие возможности ИНС позволяют успешно использовать их и при разработке прогностических способов.

Идея использования искусственного интеллекта для решения подобных задач уже воплощена на практике. В 2025 году в СМИ опубликованы результаты, полученные специалистами Сколтеха. Утверждается, что российские ученые обучили нейросети прогнозировать лесные пожары на пять дней вперед с точностью от 70 до 87%. По словам потребителей системы – представителей Рослесхоза, – в ведомстве ознакомились с разработкой и попросили авторов ее доработать.

В рамках проводимого исследования практическая реализация данного направления осложнялась тем обстоятельством, что для правильного обучения нейронных сетей необходимо располагать архивным материалом существенного объема. По разным оценкам либо число обучающих примеров должно превышать число параметров модели более чем в 10 раз, либо число примеров на каждую категорию должно превышать 1000. Поэтому была организована новая более обширная архивная выборка. Ее объем составил 1660 случаев.

С использованием нового обучающего множества были построены и обучены несколько вариантов нейронных сетей. Учитывая наличие различных оснований классификации, их можно охарактеризовать как простые однонаправленные аналоговые многослойные нейронные сети прямого распространения с однородными нейронами и фиксированными связями между ними, реализующие управляемое обучение. Для этого был использован пакет программ Deductor Studio Academic от отечественной компании BaseGroup Labs, позволяющий реализовывать подобные проекты. Интерфейс программы на стадии оценивания качества ИНС показан на рис. 2.

Результаты представляются довольно любопытными. Дело в том, что при построении ИНС одним из важнейших этапов является выбор ее структуры, т.е. количества скрытых слоев и числа нейронов в них. Этот выбор является в значительной степени субъективным. Обычно более сложные сети демонстрируют лучшие результаты на обучающем множестве наблюдений, но гораздо худшие – на независимом материале. Отличный анализ данного словосочетания проведен в статье С.А. Солдатенко Искусственный интеллект в задачах численного прогноза погоды [7].

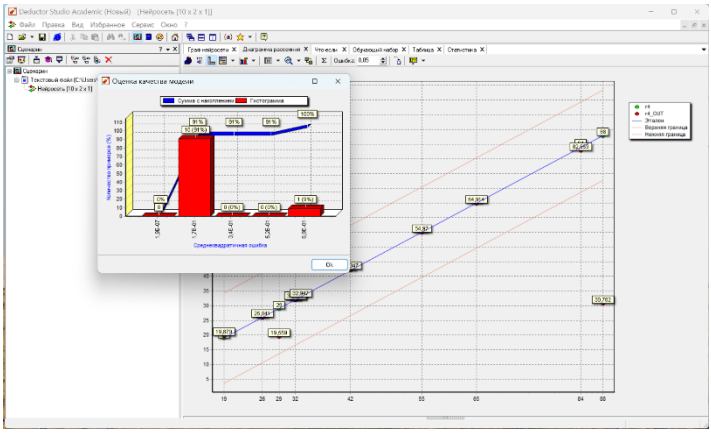


Рис. 2. Пример работы с программой при оценивании успешности обучения сети

При некотором упрощении структуры наблюдается близость результатов на основе обеих выборок, что свидетельствует об успешном обучении ИНС. Однако в ходе выполнения работы было установлено, что добавление или исключение очередного скрытого слоя или изменение числа нейронов в скрытых слоях зачастую не приводило к существенному изменению результатов. Кроме того оказалось, что результаты, полученные по обучающей и контрольной выборкам, были всегда близки друг к другу. Тем не менее, при всех рассмотренных вариантах структуры сети показатели успешности прогнозов на основе искусственного интеллекта значительно превосходили показатели прогнозов с помощью уравнений регрессии. При этом средние абсолютные ошибки прогнозирования достигли значений 0,06, а число ошибочно распознанных случаев колебалось от 2 до 98, что при достаточно большом объеме выборки является хорошим результатом.

Заключение

Анализ полученных результатов позволяет сделать некоторые выводы.

1. Лесные пожары остаются одним из неблагоприятных природных явлений, а борьба с ними является важной народнохозяйственной задачей.

2. В настоящее время существуют методы прогнозирования числа пожароопасных дней и площади пожаров в следующем месяце. Однако надежного краткосрочного способа прогнозирования числа очагов лесных пожаров пока не предложено.

3. Установлена слабая связь числа возгораний в лесах Красноярского края с параметрами термического режима атмосферы. В то же время, прогностические регрессионные зависимости на основе данных параметров характеризуются недостаточно высокими значениями показателей успешности и требуют дальнейшей доработки.

4. Разработанные нейросетевые модели продемонстрировали высокие показатели успешности прогнозирования на 24 часа, как числа очагов лесных пожаров, так и занимаемой ими площади. Они могут быть рекомендованы к практическому использованию.

Список литературы

1. Режим ЧС из-за лесных пожаров –помощь или бедствие? [Электронный ресурс]. – Режим доступа : https://www.mk.ru/nationalprojects/2022/08/12/rezhim-chs-izza-lesnykh-pozharov-pomoshh-ili-bedstvie.html?from=main_omk
2. ГОСТ Р 22.1.09–99. Безопасность в чрезвычайных ситуациях. Мониторинг и прогнозирование лесных пожаров. Общие требования [Текст]. – Введ. в действие постановлением Госстандарта России от 25 мая 1999 г. № 181. – М. : ИПК Изд-во стандартов, 1999. – 9 с.
3. Приказ Федерального агентства лесного хозяйства от 5 июля 2011 г. № 287 «Об утверждении классификации природной пожарной опасности лесов и классификации пожарной опасности в лесах в зависимости от условий погоды» // Российская газета. – 2011. – № 186.
4. Расчет комплексного показателя пожарной опасности в лесах [Электронный ресурс]. – Режим доступа : https://method.meteorf.ru/danger/fire/calculate/calculate.html
5. Сверлова, Л. И. Метод оценки пожарной опасности в лесах по условиям погоды с учетом поясов атмосферной засушливости и сезонов года / Л. И. Сверлова. – Хабаровск : ДальНИЦ Дальневосточной УГМС, 2000. – 41 с.
6. Гриценко, Д. ИИ с огоньком: нейросеть предскажет лесные пожары с точностью 87 %. Как новая разработка поможет регионам предотвращать природные ЧП / Д. Гриценко, К. Набаткина [Электронный ресурс]. – Режим доступа : https://iz.ru/1867335/denis-gritcenko/ii-s-ogonkom-nejroset-predskazhet-lesnye-pozhary-s-tochnostyu-87
7. Солдатенко, С. А. Искусственный интеллект в задачах численного прогноза погоды / С. А. Солдатенко // Метеорология и гидрология. – 2024. – № 4. – С. 5–27.

Синтез программного оптимального управления нелинейным объектом на основе DQN-агента

Д.С. Феофилов, email: fd19072002@mail.ru[1]
С. В. Феофилов, email: svfeofilov@mail.ru[1]
О. А. Фомичева, email: Olir77@mail.ru[1]

[1] ФГБОУ ВО «Тульский государственный университет»

Аннотация. *В статье рассматривается проблема синтеза программного оптимального управления для нелинейных динамических систем с применением агента Deep Q-Network (DQN), представляющего собой метод глубокого обучения с подкреплением. Проведен сравнительный анализ методов отыскания оптимального управления, начиная с принципа максимума Понтрягина и динамического программирования Беллмана, до современных численных и стохастических алгоритмов, включая методы нелинейного программирования и эволюционные вычисления. Детально описан механизм синтеза управляющих последовательностей на основе DQN-агента. Особое внимание уделено адаптации подхода DQN к задаче программного управления. Выделены преимущества подхода DQN. Проведен синтез оптимального управления для динамического объекта пятого порядка, содержащего разрывную нелинейность. Статья обобщает современное состояние направления, обозначая перспективы дальнейших исследований в области повышения эффективности и робастности алгоритмов глубокого обучения с подкреплением для задач оптимального управления.*

Ключевые слова: *оптимальное управление, глубокое обучение с подкреплением, Deep Q-Network (DQN), нелинейная динамическая система, синтез управления, принцип максимума Понтрягина, динамическое программирование, нейросетевая аппроксимация, марковский процесс принятия решений, программное управление.*

Введение

Проблема синтеза программного оптимального управления для нелинейных динамических систем представляет собой одну из наиболее сложных и актуальных задач современной теории управления. Традиционные аналитические методы, такие как принцип максимума Понтрягина или динамическое программирование Беллмана, сталкиваются с существенными вычислительными трудностями

© Феофилов Д. С., Феофилов С. В., Фомичева О.А., 2026

(«проклятием размерности») при работе с системами высокой сложности в условиях неполной информации и стохастических возмущений. Поиск численных решений часто требует значительных упрощающих предположений о структуре системы и критерии качества. В этом контексте методы глубокого обучения с подкреплением (Deep Reinforcement Learning, DRL), в частности алгоритмы на основе Deep Q-Network (DQN), открывают новые перспективы для прямого синтеза управляющих последовательностей, обучаясь на взаимодействии с моделью объекта или с самим объектом.

1. Исторический контекст и сравнительный анализ методов

Задача отыскания оптимального управления берет начало с вариационного исчисления и задачи Лагранжа, получившей мощное развитие в середине XX века. Ключевым прорывом стал принцип максимума Л.С. Понтрягина [6], предоставивший необходимые условия оптимальности для широкого класса задач. Этот метод, наряду с уравнением Гамильтона-Якоби-Беллмана [1], сформировал классический аналитический аппарат. Однако их практическое применение к нелинейным системам часто связано с решением сложных краевых задач, не имеющих аналитического решения.

С развитием вычислительной техники на первый план вышли численные методы. Прямые методы, такие как параметризация управления, сводят задачу к нелинейному программированию (НЛП) [3]. Косвенные методы, использующие принцип максимума, требуют решения двухточечной краевой задачи. Итерационные алгоритмы, подобные методу градиента в пространстве управлений, также нашли широкое применение. Несмотря на прогресс, эти подходы критически зависят от точности модели, начального приближения и могут застревать в локальных экстремумах для невыпуклых задач [8].

Альтернативой стали методы адаптивной и робастной оптимизации, а также стохастические алгоритмы, такие как генетические алгоритмы или метод роя частиц [4]. Они менее чувствительны к локальным минимумам и не требуют вычисления градиентов, но характеризуются медленной сходимостью и отсутствием гарантий оптимальности, что ограничивает их использование в задачах реального времени [9].

Одними из первых свой метод численного отыскания оптимального управления представили Ф.Л. Черноусько и И.А. Крылов [11]. Данный подход основан на принципе максимума Понтрягина и гарантированно за ограниченное количество итераций позволяет отыскать оптимальное управление. Однако при наличии разрывных правых частей в математическом описании объекта, использование предложенного

метода затрудняется в виду сложности отыскания градиента оптимизируемого функционала.

Качественно новым этапом стало применение машинного обучения, в частности, нейросетевого подхода к аппроксимации оптимальной стратегии управления или функции «ценность-действие» [5]. Истоки этого направления лежат в нейодинамическом программировании [2], где нейронные сети использовались для аппроксимации функции стоимости Беллмана. Однако подлинная революция произошла с появлением глубокого обучения с подкреплением, объединившего метод временных разностей (Sutton, Barto, 1998) [7] с возможностями глубоких нейронных сетей. Алгоритм Deep Q-Network (DQN), представленный компанией DeepMind (Mnih et al., 2015) [10], продемонстрировал способность обучаться оптимальным стратегиям непосредственно из высокоразмерных сенсорных данных, что открыло путь к управлению сложными нелинейными системами без явного задания аналитической модели.

2. Постановка задачи

DQN-агент решает задачу синтеза оптимального программного управления, формулируя ее как процесс принятия решений. Состоянием системы является вектор фазовых координат объекта (или их оценка), действием – управляющее воздействие. Цель агента – максимизировать дисконтированную совокупную награду, которая формализует заданный критерий оптимальности (например, минимум интегрального квадратичного отклонения при штрафе за энергию управления). Функция Q оценивает ожидаемую суммарную награду за выполнение действия в каждом состоянии с последующим следованием оптимальной политике.

Глубокие нейронные сети в DQN используются как универсальные аппроксиматоры для этой сложной нелинейной функции Q. Ключевые инновации оригинального DQN, обеспечившие стабильность обучения, включают использование воспроизводящей памяти (replay buffer) для разбиения временных корреляций в данных, применение целевой сети (target network), периодически копируемой из обучаемой, для стабилизации целевых значений Q [2]. В процессе обучения агент, взаимодействуя со средой (моделью объекта), накапливает траектории в буфере. Обучающая выборка мини-батчей из этого буфера используется для минимизации среднеквадратичной ошибки временных разностей (TD-ошибки) между предсказанной Q-ценностью и целевым значением, вычисленным с учетом награды и максимальной Q-ценности в следующем состоянии.

Для задач программного управления, где оптимальное управление часто является функцией только времени при заданных начальных

условиях, подход DQN интерпретируется несколько иначе. Временной шаг становится частью состояния агента, что позволяет политике явно зависеть от времени. Таким образом, в процессе обучения DQN-агент неявно восстанавливает оптимальную программу для заданного горизонта управления, обучаясь на множестве траекторий, стартующих из различных начальных условий или при различных возмущениях. Это позволяет синтезировать управление, робастное к вариациям начальных данных и параметров модели.

Сравнительным преимуществом DQN-подхода перед классическими методами является его способность работать в условиях неполной или неточной модели, обучаясь непосредственно на данных, будь то результаты численного моделирования или натурного эксперимента. Агент исследует пространство состояний и действий, постепенно улучшая стратегию без необходимости явного решения уравнений оптимальности. Однако метод не лишен недостатков: требование дискретизации пространства действий, высокая вычислительная сложность обучения, а также проблемы с гарантией устойчивости и безопасностью в критических приложениях, что требует дополнительных мер, таких как обучение с учителем на базе субоптимальных классических контроллеров или использование безопасных зон исследования.

3. Практическая реализация синтеза оптимального управления на основе DQN-агента

Рассмотрим задачу синтеза оптимального управления. В качестве объекта управления возьмем вертикальный канал летательного аппарата, а в качестве критерия оптимизации – уменьшение среднеквадратичной ошибки отклонения переходного процесса при достижении линии визирования. Функциональная схема представлена на рис.1.

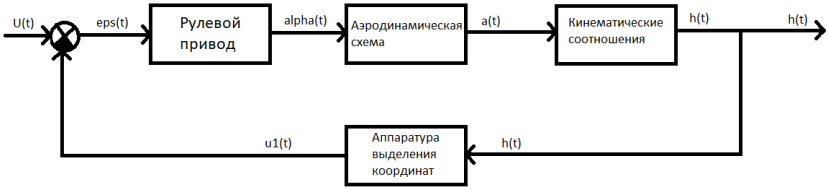


Рис. 1. Функциональная схема объекта

На рис. 1 $U(t)$ – управление, $h(t)$ – высота, $eps(t)$ – ошибка между высотой и управлением, $alpha(t)$ – угол отклонения рулевого привода, $a(t)$ – ускорение ЛА, $u1(t)$ – выделенные координаты с помощью АВК.

Модель включает в себя цепь последовательно соединенных элементов: рулевой привод, аэродинамическая схема, кинематическая связь. Состояние объекта описывается четырехмерным вектором: положение рулевого привода, ускорение, скорость и положение летательного аппарата. Управляющим воздействием является входное напряжение рулевого привода, ограниченное по амплитуде. Нелинейность типа упора вводит в модель нелинейную функцию, что приводит к возникновению ударных нагрузок и сложной динамики, затрудняющей аналитическое решение. После линеаризации уравнений динамики получим следующую структурную схему объекта управления (рис.2).

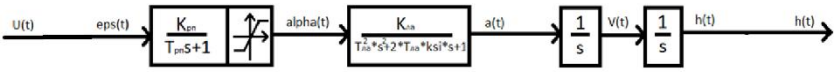


Рис. 2. Структурная схема объекта управления

На рис. 2 $K_{рп}$, $T_{рп}$ – коэффициент передачи и постоянная времени рулевого привода, $K_{ла}$, $T_{ла}$, ksi – коэффициент передачи, постоянная времени и коэффициент колебательности аэродинамической схемы ЛА, $V(t)$ – скорость ЛА. Как можно заметить, система нелинейна – рулевой привод описывается апериодическим звеном с ограничителем типа насыщение. Параметры реальной системы неизвестны и на начальном участке траектории отсутствует обратная связь по положению, поэтому ставится задача отыскания программного оптимального управления в условиях неполной информации. В рассматриваемом примере Simulink-модель используется для сбора данных для обучения, однако нейронная сеть воспринимает объект как «черный ящик».

Для решения поставленной задачи применяется архитектура, объединяющая среду моделирования Simulink (MATLAB) и среду обучения агента на Python. Связь между средами организуется через механизм co-simulation, например, с использованием встроенного в Python MATLAB Engine, что позволяет Python-скрипту инициировать запуск симуляции, передавать на каждом шаге дискретного времени текущее управляющее воздействие и получать вектор состояния объекта, рассчитанный решателем Simulink. Это создает связь между объектом и нейронной сетью, в котором среда воспринимается DQN-агентом как «черный ящик», что полностью соответствует парадигме обучения с

подкреплением, где агенту не требуется априорное знание уравнений динамики.

Архитектура DQN-агента, реализуемая в Python, включает основную нейронную сеть и целевую сеть. Входной слой принимает нормализованный четырехмерный вектор состояния, дополненный информацией о текущем временном шаге в пределах горизонта управления, что критично для задачи программного управления. Выходной слой имеет размерность, равную количеству дискретных уровней управляющего напряжения, которые задаются с определенным шагом квантования в допустимом диапазоне. Сеть состоит из нескольких полносвязных скрытых слоев с функциями активации ReLU, что позволяет ей аппроксимировать сложную нелинейную Q-функцию. Гиперпараметры, такие как размер воспроизводящей памяти, скорость обучения оптимизатора Adam, период обновления целевой сети и коэффициент дисконтирования, подбираются в ходе предварительных вычислительных экспериментов.

Процедура обучения организуется в виде последовательности эпизодов. Каждый эпизод соответствует одной симуляции движения объекта на фиксированном временном интервале. Начальные условия, в частности начальное положение исполнительного органа, могут варьироваться от эпизода к эпизоду в заданной области для обеспечения обобщающей способности агента. На каждом шаге внутри эпизода агент, основываясь на текущем состоянии и применяя, например, эпсилон-жадную стратегию [4] исследования, выбирает дискретное управляющее воздействие. Это действие передается в Simulink-модель, которая рассчитывает отклик системы на следующем шаге интегрирования. Полученное новое состояние и скалярная награда возвращаются агенту. Награда формируется таким образом, чтобы кодировать цель управления. Она включает отрицательный вклад от квадрата ошибки позиционирования. Состояние, действие, награда и следующее состояние (транзикция) сохраняются в воспроизводящей памяти.

Обучение происходит асинхронно по отношению к сбору данных. Периодически из буфера случайным образом выбирается небольшой набор транзикций, и на его основе вычисляется целевое значение Q-функции с использованием целевой сети. Оптимизация заключается в минимизации среднеквадратического отклонения между предсказаниями основной сети и этими целевыми значениями, тем самым итеративно уточняя оценку будущих накопленных наград. Ключевой вызов в данной связке – синхронизация времени и управление скоростью симуляции. Модель Simulink работает с тем же шагом интегрирования, что и агент, для синхронизации управляющих воздействий и реакции на них. Процесс

обучения характеризуется высокой вычислительной нагрузкой. Пример обучения DQN-агента для рассматриваемого объекта представлен на рис .3-4.

Рис. 3 иллюстрирует график обучения нейронной сети. Если награда находятся на значениях 2000-2500, то агент не обучен, 2500-3500 – агент понимает основные принципы управления, 3500-4500 – агент обучен синтезировать оптимальное управления для данного объекта.

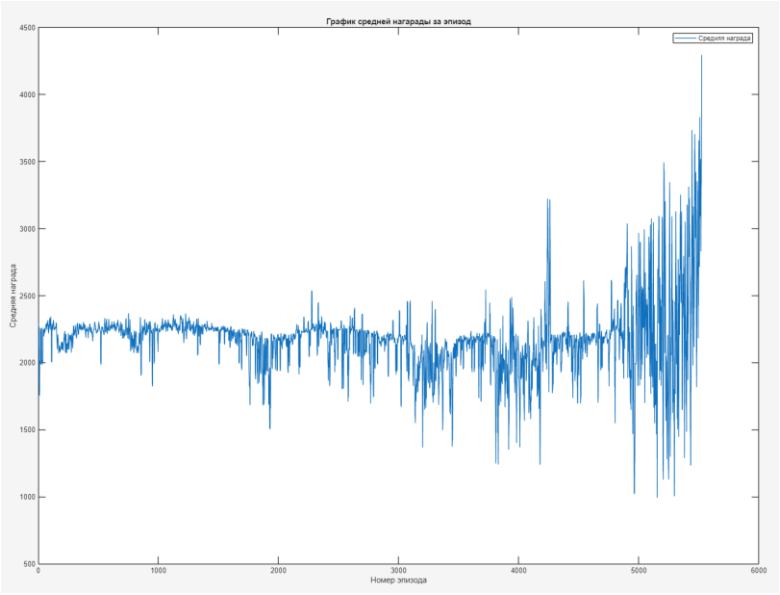


Рис. 3. График зависимости награды за эпизод и средней награды от количества эпизодов

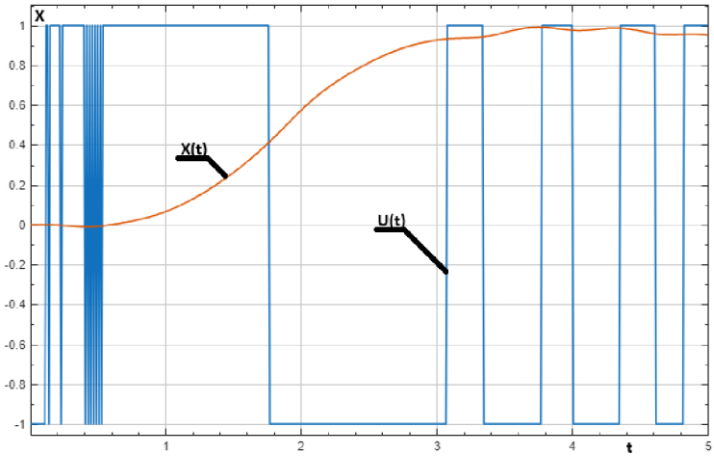


Рис. 4. Синтезированные оптимальное управление и траектория на основе обученного DQN-агента

На рис. 4. $X(t)$ – положение летательного аппарата, $U(t)$ – управляющее воздействие, синтезированное нейронной сетью.

В результате успешного обучения DQN-агент формирует политику, которая аппроксимирует оптимальное программное управление. Анализ полученных траекторий показывает, что агент обучается комплексной стратегии. На начальной фазе движения применяется близкое к максимальному управление для разгона, затем следует фаза регулирования для компенсации инерции и точного выхода в целевую точку с минимальными колебаниями. Это свидетельствует о том, что нейронная сеть научилась учитывать сложную нелинейность, не представленную в явном виде в ее входных данных. Таким образом, рассматриваемый пример подтверждает возможность и практическую целесообразность использования DQN-агента в связке с реальным нелинейными объектом типа «черный ящик» для синтеза оптимального программного управления, где традиционные методы сопряжены со значительными трудностями.

Заключение

Синтез оптимального программного управления на основе DQN-агента представляет собой мощный вычислительный метод, основанный на парадигме глубокого обучения с подкреплением. Он наследует и развивает идеи динамического программирования, преодолевая «проклятие размерности» за счет аппроксимационных свойств глубоких нейронных сетей. В сравнении с классическими аналитическими и

численными методами, подход DQN демонстрирует большую гибкость и применимость к высокоразмерным и плохо формализуемым нелинейным системам, хотя и требует тщательного подбора гиперпараметров, значительных вычислительных ресурсов для обучения и дополнительного анализа на робастность. Дальнейшие исследования в этой области направлены на повышение эффективности обучения, обеспечение гарантий устойчивости, интеграцию с методами адаптивного управления и применение в гибридных системах. Таким образом, DQN и его модификации утвердились в качестве одного из ключевых инструментов в современном арсенале методов оптимизации управления сложными динамическими системами.

Список литературы

1. Беллман Р. Динамическое программирование / Р. Беллман. – Принстон: Princeton University Press, 1957.
2. Бертсекас Д. П. Нейро-динамическое программирование / Д. П. Бертсекас, Дж. Н. Цицклис. – Белмонт: Athena Scientific, 1996.
3. Брайсон А. Э. Прикладная теория оптимального управления / А. Э. Брайсон, Ю. –Ч. Хо. – Вашингтон: Hemisphere Publishing, 1975.
4. Кеннеди Дж. Оптимизация роем частиц // Труды международной конференции по нейронным сетям (ICNN'95). – Перт, 1995. – Т. 4. – С. 1942–1948.
5. Мних В. Управление на уровне человека с помощью глубокого обучения с подкреплением / В. Мних, К. Кавукчуоглу, Д. Сильвер [и др.] // Природа. – 2015. – Т. 518, № 7540. – С. 529–533.
6. Понтрягин Л. С. Математическая теория оптимальных процессов / Л. С. Понтрягин, В. Г. Болтянский, Р. В. Гамкрелидзе, Е. Ф. Мищенко. – М.: Наука, 1969. – 384 с.
7. Саттон Р. С. Обучение с подкреплением: Введение / Р. С. Саттон, Э. Г. Барто. – Кембридж: MIT Press, 1998.
8. Уоткинс К. Дж. К. Х. Обучение по отложенным вознаграждениям: дис. ... д-ра философии / К. Дж. К. Х. Уоткинс. – Кембридж: Кембриджский университет, 1989.
9. Льюис Ф. Л. Оптимальное управление / Ф. Л. Льюис, Д. Враби, В. Л. Сирмос. – 3-е изд. – Хобокен: John Wiley & Sons, 2012.
10. Continuous control with deep reinforcement learning / T. P. Lillicrap [et al.]. – . – [Электронный ресурс]. – 2015. – URL: https://arxiv.org/abs/1509.02971
11. Крылов, И. А. Алгоритм метода последовательных приближений для задач оптимального управления / И. А. Крылов, Ф. Л. Черноусько // Журнал вычислительной математики и математической физики. – 1972. – Т. 12, № 1. – С. 14-34.

Анализ нейросетевых алгоритмов управления неустойчивыми динамическими объектами

А. А. Холина , e-mail: xolina.nastya2017@gmail.com,
Д. Л. Хапкин, dima-hapkin@ya.ru

Тульский государственный университет

Аннотация. *В статье рассматривается задача синтеза системы управления для неустойчивого нелинейного объекта на примере перевернутого маятника с использованием искусственной нейронной сети (ИНС). Проводится анализ ограничений классических методов управления, в частности ПИД-регулятора, при работе с объектами, обладающими нелинейностью, нестационарностью и неполной математической моделью. В качестве альтернативного подхода предлагается нейрорегулятор на основе рекуррентной нейронной сети с архитектурой GRU, способный к обучению на экспериментальных данных. Описаны этапы математического моделирования динамики маятника, построения Simulink-модели системы, а также процесс обучения нейросети на гармонических сигналах. Результаты моделирования демонстрируют, что нейрорегулятор обеспечивает качество управления, сравнимое с классическим ПИД-регулятором, и подтверждают перспективность применения нейросетевых методов для управления сложными динамическими системами в условиях неопределённости.*

Ключевые слова: *перевернутый маятник, нейрорегулятор, искусственная нейронная сеть, ПИД-регулятор, нелинейное управление, математическое моделирование, Simulink, GRU-сеть.*

Введение

В современной теории автоматического управления существует существенная проблема синтеза регуляторов для нелинейных объектов управления (ОУ) при неполной исходной информации. На данный момент разработано множество методов проектирования систем управления для отдельных классов нелинейных ОУ, но они уже не отвечают потребностям разработчиков, работающих со сложными реальными динамическими объектами.

С позиций теории автоматического управления подобные объекты характеризуются: нестационарностью, необходимостью цифрового

© Холина А.А., Хапкин Д.Л., 2026

управления с заданными показателями качества [1], существенной нелинейностью (наличием жёстких механических ограничителей, фазовых ограничений, сухого и вязкого трения, нежёсткости механической передачи и др.); В таких условиях применение классических методов синтеза регуляторов требует комплексного обращения к фундаментальным разделам теории автоматического управления, что влечёт за собой формирование громоздких алгоритмов.

Обозначенные проблемы носят системный характер и трудноразрешимы в рамках классической теории управления. В качестве перспективного альтернативного подхода рассматриваются регуляторы на базе искусственных нейронных сетей (ИНС). Их ключевое преимущество проявляется в ситуациях, когда отсутствует достоверная и достаточно точная математическая модель объекта управления и имеется возможность получения экспериментальных данных. Достоинством таких регуляторов является их способность к обучению и адаптации под объект на основе полученных данных

Дополнительным аргументом в пользу ИНС служит ограничение классических методов синтеза, которые, как правило, формируют линейный закон управления. Для нелинейных объектов управления это создаёт существенные ограничения. В отличие от них, искусственные нейронные сети могут быть представлены как нелинейная динамическая система, что обеспечивает им большую гибкость и функциональные преимущества. В качестве иллюстративного примера неустойчивого объекта управления можно привести перевёрнутый маятник.

1. Постановка задачи

Для обучения нейрорегулятора будем использовать лабораторную установку перевернутого математического маятника, приведенную на рис. 1.



Рис. 1. Лабораторный стенд «Перевернутый маятник»

Перевернутый маятник –это классическая задача управления нелинейной и неустойчивой динамической системой. Установка состоит из транспортировочного рельса, каретки, электропривода, энкодера, прикрепленного к электроприводу, маятника, установленного на тележке, и потенциометра, прикрепленного к тележке. Для уравновешивания маятника тележкой необходимо использовать алгоритм управления. Поэтому важными измерениями являются отклонение маятника от вертикального положения и положение каретки. Энкодер используется для измерения положения и скорости каретки. Другой используемый датчик –потенциометр, измеряющий угол, который образует маятник.

Один из классических методов разработки алгоритма управления, использующего угол поворота маятника и положение тележки в качестве входного измерения, заключается в использовании ПИД-регулятора [2]. Этот регулятор поддерживает выход системы, обеспечивая нулевую ошибку между желаемым выходом и переменными процесса в замкнутом контуре. ПИД-регулятор состоит из пропорционального регулятора, интегрального регулятора и дифференциального регулятора. Сначала пропорциональный регулятор выдает выходной сигнал, пропорциональный текущей ошибке. Интегральный регулятор интегрирует ошибку в течение определенного периода времени до тех пор, пока значение ошибки не станет равным нулю. Наконец, дифференциальный регулятор прогнозирует будущее поведение ошибки. На рис. 2 представлена схема работы ПИД-регулятора.

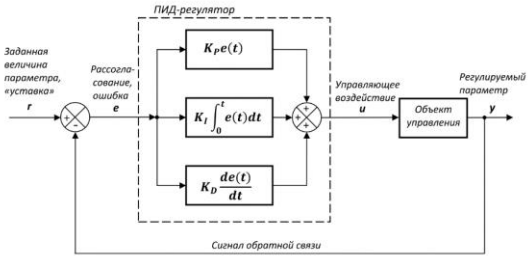


Рис. 2. Стабилизация объекта с помощью ПИД-регулятор

В работе будем использовать ПИД-регулятор в качестве эталонного сравнения для оценки работы нейрорегулятора. Для понимания динамики системы сначала необходимо применить математическое моделирование [3].

2. Математическое моделирование объекта управления

На рис. 3 представлена расчетная схема объекта управления – тележки и маятника с учетом действующих сил [4]. Ниже приведен список соответствующих переменных и их значения:

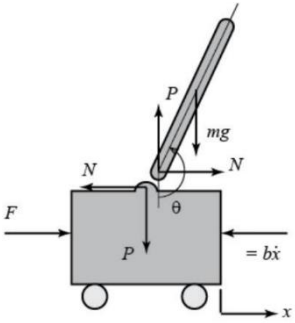


Рис. 3. Расчетная схема ОУ

x –координата положения тележки, m –масса маятника, M –масса тележки, b –коэффициент трения тележки I –момент инерции маятника, l –длина маятника. F –сила, приложенная к массе, θ – угол, который образует маятник

Сила является входным сигналом системы, положение тележки и угол поворота маятника –выходными сигналами системы. Переменная b определяется как коэффициент трения тележки. На самом деле, существует также коэффициент трения маятника, но по сравнению с трением на тележке трение на маятнике значительно меньше, поэтому трением маятника пренебрегают.

Для расчета уравнений сил P и N необходимо учитывать как вертикальное, так и горизонтальное положение маятника

$$\ddot{x} = \frac{1}{M}(F - N - b\dot{x})$$

$$\ddot{\theta} = \frac{1}{I}(-Nl\cos(\theta) - Pl\sin(\theta))$$

Вертикальное положение маятника

$$m\ddot{y}_p = P - mg1$$

$P = m\ddot{y}_p + mg = m(\ddot{y}_p + g)$, где

$$y_p = -l\cos(\theta)$$
$$\dot{y}_p = -l\sin(\theta)\dot{\theta}$$
$$\ddot{y}_p = l\cos(\theta)\dot{\theta}^2 + l\sin(\theta)\ddot{\theta}$$

После этого необходимо применить уравнения Лагранжа, уравнения Лагранжа Эйлера и линеаризовать полученные уравнения[4].

В результате получаются уравнения движения, выраженные уравнениями

$$(M+m)\ddot{x} + ml\ddot{\theta} = F - b\dot{x}$$
$$(ml^2 + I) + ml\ddot{x} + mgl\theta = 0$$

Последним шагом перед получением уравнений движения является вычисление Преобразование Лапласа линеаризованных уравнений движения. Затем можно получить следующие уравнения

$$(M+m)s^2X(s) + mls^2\theta(s) = F(s) - bsX(s)$$
$$(ml^2 + I)s^2\theta(s) + mls^2X(s) + mgl\theta(s) = 0$$

После преобразований получаем уравнения движения , демонстрирующие поведение системы

$$f(u(\ddot{x})) = \frac{1}{M}(u(1) - u(2) - bu(3)$$
$$f(u(\ddot{\theta})) = \frac{1}{I}(-u(1)l - lu(3)u(2))$$
$$f(u_P) = m(l(u(1))u(2) + g)$$
$$f(u_N) = m(u(1)) - lu(2))$$

По уравнениям движения составляем Simulink модель

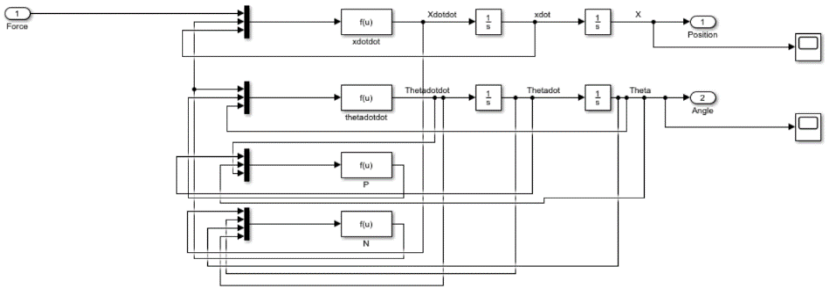


Рис. 4. Simulink модель системы

Входным сигналом системы, является сила, а выходным сигналом – положение и угол.

Таким образом после построения модели оригинальной системы и системы с классическим ПИД -регулятором можем сравнить нейрорегулятор с данными системами .

Синтез нейросетевого регулятора

$$\hat{V}^j = \frac{V^j}{\| V^j \|} = \left(\frac{V_{j,x}}{\sqrt{V_{j,x}^2 + V_{j,y}^2}} \; \frac{V_{j,y}}{\sqrt{V_{j,x}^2 + V_{j,y}^2}} \right)$$

3. Пример синтеза нейросетевого регулятора

Использовалась нейронная сеть со структурой, представленной в таблице[6]. Выбор структуры проводился экспериментально с последовательным наращиванием количества слоев и нейронов.

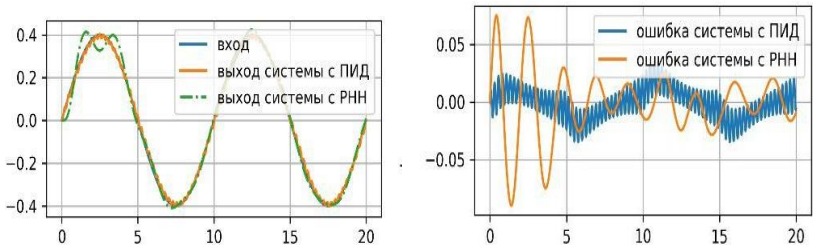
Таблица

Структура нейронной сети для нейрорегулятора

№ п/п	Слой	Количество нейронов
1	Входной	2
2	GRU	200
3	Dropout 25%	-
4	Линейный	200
5	ReLU	-
6	Dropout 25%	-
7	Линейный	200
8	ReLU	-
9	Выходной	1

Обучение производилось на гармонических (sin) сигналах.

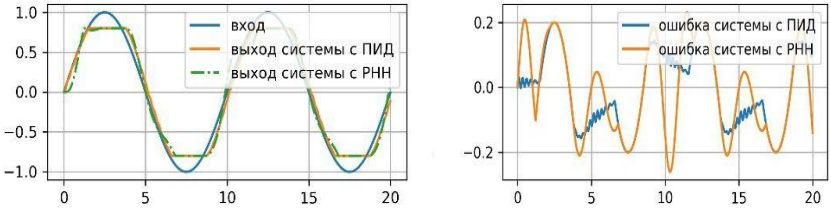
На рис. 5 и рис. 6 представлены результаты моделирования в Simulink. Сравнивается работа нейрорегулятора с ПИД регулятором, синтезированным средствами Matlab



а б

а – реакция систем, б – ошибка слежения

Рис. 5. Реакция объекта и РНН на гармонический сигнал (sin) амплитудой 0,4 и частотой 0,1 Гц



а б

а – реакция систем, б – ошибка слежения

Рис. 6. Реакция объекта и РНН на гармонический сигнал (sin) амплитудой 1 и частотой 0,1 Гц

Как видно, нейронная сеть смогла обучиться управлять объектом почти также хорошо, как и ПИД-регулятор.

Заключение

Применение искусственных нейронных сетей в системах управления представляется перспективным направлением, особенно в условиях неполной информации об объекте и наличия экспериментальных данных. Нейрорегуляторы обладают значительным потенциалом для дальнейшего развития в области управления сложными динамическими системами, обеспечивая гибкость, устойчивость и возможность обучения в реальном времени.

Список литературы

1. Тарловский, Д. Ю. Промышленная автоматизация и робототехника : учебное пособие / Д. Ю. Тарловский, А. В. Лысенко. – Санкт-Петербург : Лань, 2021. – 336 с. – ISBN 978-5-8114-6875-5.

2. Ротач, В. Я. Теория автоматического управления : учебник для вузов / В. Я. Ротач. – 3-е изд., перераб. и доп. – Москва : Издательский дом МЭИ, 2016. – 400 с. – ISBN 978-5-383-00949-5.

3. Scherpen, J. M. A. Discrete Technology and Production Automation [Electronic resource] / J. M. A. Scherpen ; Engineering and Technology institute Groningen. – Electronic data. – Groningen, 2018.

4. Попов, А. М. Моделирование системы «обратный маятник» в среде MatLab / А. М. Попов, И. В. Черных // Известия Томского политехнического университета. – 2011. – Т. 318, № 5. – С. 124–128.

5. Хапкин, Д.Л. Синтез гарантированно устойчивых нейросетевых регуляторов с оптимизацией качества переходного процесса / Д.Л. Хапкин, С.В. Феофилов, А.В. Козырь // Известия ТулГУ. Технические науки. – 2022. – № 12. – С. 128-133.

Компьютерное моделирование авиационного электроэнергетического комплекса

А. Р. Шабанов, email: zzzdwx@bk.ru
А. С. Бочаров

ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *В статье рассматриваются методы компьютерного моделирования авиационного электроэнергетического комплекса для повышения надежности его работы. Обосновывается применение среды Matlab/Simulink для анализа аварийных режимов, что обусловлено способностью моделировать разнородные физические процессы, обеспечивать структурную наглядность и гибкость построения моделей. Представлена обобщенная структура модели ЭЭК современного летательного аппарата. Особое внимание уделено методу полунатурного моделирования, позволяющему устранить разрыв между цифровой моделью и реальностью за счет интеграции математических моделей с реальными аппаратными компонентами.*

Ключевые слова: *авиационный электроэнергетический комплекс, качество электроэнергии, Matlab/Simulink*

Введение

Одним из наиболее важных направлений развития авиационной техники является электрификация бортового оборудования. Рост мощности и числа потребителей, внедрение цифровой элементной базы требуют особого внимания к качеству электроэнергии, а поскольку работоспособность практически всех бортовых систем прямо или косвенно зависит от стабильности параметров питания, обеспечение безаварийности электроэнергетического комплекса (ЭЭК) становится критически важной задачей. В современных условиях традиционные методы проектирования и анализа могут становиться недостаточными, экономически неэффективными и неспособными гарантировать требуемый уровень безопасности. В свою очередь, применение компьютерного моделирования позволяет решать ряд научно-технических задач по оптимизации параметров и повышению устойчивости работы ЭЭК, особенно в нештатных и аварийных ситуациях. Современные средства моделирования должны не только

© Шабанов А.Р., Бочаров А.С., 2026

точно воспроизводить физику процессов, но и обеспечивать наглядность, гибкость изменения структуры модели и возможность верификации результатов.

Разработка модели

Выбор программного обеспечения является важным аспектом успешного моделирования. Сформулируем основные требования к средству моделирования:

– возможность моделировать разнородные физические процессы (электрические, механические, тепловые, магнитные) в единой среде;

– структурная наглядность модели, соответствующая принципиальной схеме объекта;

– гибкость и масштабируемость структуры, а также параметров модели.

В качестве базового инструмента для моделирования авиационных ЭЭК выбрана среда программирования Matlab/Simulink с расширением Simscape Electrical. Это связано с её способностью охватывать широкий спектр электроэнергетических объектов и модифицировать как встроенные, так и созданные пользователем блоки. На основе анализа типовых архитектур ЭЭК современных летательных аппаратов (ЛА), можно составить обобщенную структурную схему:



Рис. 1. Структурная схема ЭЭК современного ЛА

На рис. 1 изображены:

2. Модель силовой установки (авиационного двигателя).
3. Генераторная установка переменного тока.

4. Модель сети переменного тока, включающая преобразователи тока, систему распределения электроэнергии, аппаратуру защиты и управления.

5. Модели потребителей электроэнергии.

6. Система сбора данных и анализа.

Однако традиционные ЭЭК обладают рядом ограничений: относительно низкий общий КПД, высокая масса, а также сложности при диагностировании. Возможным решением этих проблем может быть концепция «более электрического самолёта» (БЭС). Её ключевыми достоинствами являются:

1. Повышение общей эффективности за счёт замены неэлектрических систем (гидравлических, пневматических, механических) на электрические приводы, что снижает потери энергии при передаче и преобразовании.

2. Снижение массы и упрощение конструкции благодаря отказу от протяжённых трубопроводов и гидронасосов.

3. Рост надёжности, поскольку электрические системы проще контролировать, диагностировать и резервировать.

БЭС будет использовать централизованные источники, вырабатывающие электроэнергию переменного тока переменной частоты («грязный ток»). В этом случае получение электрической энергии требуемого вида и качества будет обеспечиваться индивидуальными или групповыми преобразователями, установленными в исполнительной системе. Примерная структура ЭЭК для БЭС представлена на рис. 2:

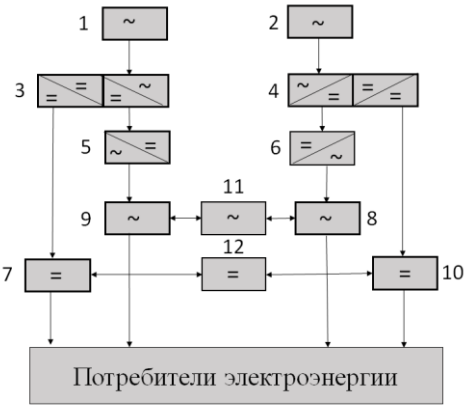


Рис. 2. Структурная схема ЭЭК БЭС

Цифрами на рис. 2 обозначены:

1 – левый генератор;

2 – правый генератор;

3 – выпрямительный блок левого канала;

4 – выпрямительный блок правого канала;

5 – инверторный блок левого канала;

6 – инверторный блок правого канала;

7 – шина переменного тока левого канала;

8 – шина переменного тока правого канала;

9 – шина постоянного тока повышенного напряжения левого канала;

10 – шина постоянного тока повышенного напряжения правого канала;

11 – распределительное устройство переменного тока;

12 – распределительное устройство постоянного тока.

Реализация этой концепции ведёт к качественному усложнению архитектуры ЭЭК, что, в свою очередь, предъявляет повышенные требования к точности математических моделей и методам их верификации.

Несмотря на очевидные достоинства, цифровое моделирование всё-таки имеет определенный разрыв с реальностью, избавиться от которого предлагается с помощью применения метода полунатурного моделирования. Сущность данного подхода заключается в создании испытательной среды, в которой математические модели взаимодействуют с реальными агрегатами. Преимуществами метода является возможность исследования совместной работы всех компонентов ЭЭК, а также безопасная имитация аварийных режимов работы без риска повреждения дорогостоящего оборудования. Эффективность метода полунатурного моделирования наглядно иллюстрируют следующие типовые примеры:

Одной из наиболее сложных динамических нагрузок является асинхронный двигатель, потребляющий в момент пуска ток, в несколько раз превышающий номинальный. Ярким примером на воздушном судне является запуск электродвигателя насоса топливной системы или системы кондиционирования воздуха (СКВ) во время полета. Внезапное подключение такой нагрузки к бортовой сети может вызвать кратковременную, но значительную просадку напряжения. Моделирование данного режима позволяет оценить глубину и длительность просадки напряжения на шинах, а также проанализировать влияние скачков на работу исследуемого оборудования.

Для отладки и тестирования систем защиты ЭЭК функционалом программной среды Matlab/Simulink возможна имитация обрыва фазы

генератора. В ходе моделирования последовательно воссоздаётся обрыв фазы статорной обмотки, регистрируются возникающие токи и напряжения обратной последовательности, анализируется срабатывание алгоритмов в реальном блоке управления и оценивается процесс переключения на резервный источник питания. Такой подход обеспечивает высокую достоверность испытаний без риска для физического оборудования.

Заключение

Проведенный анализ демонстрирует, что компьютерное моделирование перестало быть вспомогательным инструментом и стало одной из центральных технологией жизненного цикла авиационного ЭЭК. Компьютерное моделирование возможно использовать как при проектировании, так и подготовке кадров. Использование платформы MATLAB/Simulink/ в сочетании с полунатурным моделированием позволяет создавать сложные, физически адекватные модели компонентов и всего авиационного ЭЭК в целом, а затем использовать их при испытаниях и исследованиях, сокращая количество дорогостоящих натурных экспериментов.

Список литературы

1. Халютин, С. П. Моделирование сложных электроэнергетических систем летательных аппаратов : монография / С. П. Халютин, М. Л. Тюляев, Б. В. Жмуров, И. Е. Старостин. – Москва : ВУНЦ ВВС «ВВА им. проф. Н. Е. Жуковского и Ю. А. Гагарина», 2010. – 188 с.

2. Герман-Галкин, С. Г. MATLAB & Simulink. Проектирование мехатронных систем на ПК / С. Г. Герман-Галкин. – Санкт-Петербург : КОРОНА-Век, 2008. –368 с.

3. Бочаров, А. С. Разработка элементов электроэнергетического комплекса летательного аппарата в среде структурно-визуального моделирования Simulink / А. С. Бочаров, К. А. Губанов, Я. А. Евдокимов // Электрификация летательных аппаратов : труды научно-технической конференции, посвященной 125-летию со дня рождения акад. В. С. Кулебакина : сб. докл. – Москва : ИД Академии Жуковского, 2016. – С. 49–59.

4. Бочаров, А. С. Имитационное моделирование функционирования электрифицированных систем самолетов в среде Simulink / А. С. Бочаров, И. В. Шаров, А. А. Шипилов // Оптимизация и моделирование в автоматизированных системах : труды международной молодежной научной школы. – Воронеж : ВГТУ, 2018. – С. 184-187.

5. Ways to speed design for digitally controlled power converters with Simulink [Электронный ресурс]. – URL: https://www.mathworks.com/solutions/power-electronics-control.html

6. Бордиян, Р. Н. Особенности построения чувствительных элементов интеллектуальных датчиков / Р. Н. Бордиян, В. М. Дмитриев, С. В. Кучевский // Межвузовский сборник научных трудов. – 2020. – Вып. 24. – Краснодар : КВВАУЛ. – С. 46-51.

Алгоритм повышения стойкости микросхем в условиях космоса

А.В. Шпинев, email: Andr0ne0@yandex.ru

ФГБОУ ВО «Воронежский государственный лесотехнический университет имени Г.Ф. Морозова»

Ключевые слова: *В условиях освоения космического пространства электронная компонентная база, включая программируемые логические интегральные схемы (ПЛИС), подвергается воздействию ионизирующего излучения, что приводит к единичным сбоям (SEE), накоплению дозы и отказам аппаратуры. В работе рассматривается комплексный алгоритм повышения радиационной стойкости микросхем, основанный на принципах резервирования и реконфигурации. Особое внимание уделяется тройному модульному резервированию (TMR), схемотехническим решениям мажоритарных элементов, временному резервированию, а также методам помехоустойчивого кодирования и скраббинга конфигурационной памяти. Анализируются достоинства и ограничения каждого подхода: рост аппаратной избыточности, энергопотребления и уязвимость к множественным и накопительным сбоям. Обосновывается необходимость иерархического сочетания аппаратно-программных методов (RHBD) при проектировании бортовой аппаратуры. В заключении делается вывод о том, что максимальная отказоустойчивость достигается не применением одного метода, а их сбалансированной интеграцией на всех уровнях проектирования с учётом условий эксплуатации и функциональной критичности узлов [1]. Статья подготовлена по материалам доклада на конференции «Информатика: проблемы, методы, технологии» (IPMT) им. Э.К. Алгазинова.*

Ключевые слова: *радиационная стойкость, программируемые логические интегральные схемы (ПЛИС), ионизирующее излучение, единичные сбои (SEE), космическая электроника, отказоустойчивость, тройное модульное резервирование (TMR), мажоритарный элемент, временное резервирование, аппаратная избыточность, скраббинг, реконфигурация, помехоустойчивое кодирование, корректирующие коды (ECC, CRC), радиационно-стойкое проектирование (RHBD), резервирование памяти, сбоеустойчивость.*

Введение

В условиях активного освоения космического пространства и развития аэрокосмических технологий электронные системы летательных аппаратов подвергаются воздействию жесткой

радиационной среды. Это воздействие может приводить к единичным сбоям, накоплению дозы облучения и, как следствие, к деградации или катастрофическим отказам критически важной аппаратуры. Особую актуальность проблема обеспечения радиационной стойкости приобретает при использовании современной элементной базы с малыми топологическими нормами, включая программируемые логические интегральные схемы (ПЛИС), широко применяемые для обработки данных. В данном контексте традиционные конструктивно-технологические меры защиты (например, экранирование) не всегда достаточны или экономически целесообразны, что требует поиска иных подходов к проектированию надежных систем.

Алгоритм повышения стойкости микросхем

В настоящее время можно выделить алгоритм повышения радиационной стойкости, который представлен на рис. 1.



Рис. 1. Алгоритм повышения радиационной стойкости микросхем

Резервирование – это стратегия защиты информации в электронных системах, подверженных воздействию ионизирующего излучения, основанная на её дублировании. Устойчивость к одиночным сбоям,

вызываемым радиационными частицами, обеспечивается за счёт трёх фундаментальных принципов схемного резервирования:

Сохранение целостности данных – внедрение резервной ячейки памяти, которая остаётся неизменной при локальном радиационном событии;

Механизм самовосстановления – использование обратной связи от неповреждённой ячейки для коррекции данных, искажённых в результате прохождения ионизирующей частицы;

Адаптация к физике сбоя – учёт особенностей генерации тока при ударе частицы, который возникает между n- и p-областями полупроводниковой структуры, для точного восстановления информации [2].

Наиболее распространённой реализацией этих принципов выступает тройное модульное резервирование (TMR). В рамках этого подхода три идентичные копии критического блока синхронизируются общей тактовой и шиной данных, а их выходы направляются на мажоритарный элемент, выбирающий значение, совпадающее у большинства. Подразумевается, что в один момент времени может быть нарушена работа лишь одной из трёх копий. Чтобы избежать накопления ошибок, данные в схеме регулярно обновляются, что локализует последствия сбоя в пределах одного тактового цикла.

Эффективность TMR существенно возрастает при пространственном разнесении транзисторов в резервных блоках – это снижает вероятность одновременного повреждения нескольких узлов излучением. Однако ключевой недостаток метода – значительное увеличение кристаллической площади, требуемой для реализации.

Помимо TMR, применяются и альтернативные стратегии, включая временное и пространственное резервирование.

Хотя методы обеспечения радиационной стойкости (Radiation Hardness By Design, RHBD) для цифровых и аналоговых схем имеют свои особенности, они зачастую сопряжены с общими компромиссами: ростом энергопотребления, увеличением занимаемой площади и снижением производительности.

В цифровых схемах, использующих TMR, все функциональные элементы дублируются втрое, а корректный выходной сигнал формируется на основе мажоритарного голосования. Базовая архитектура метода и схема мажоритарного элемента иллюстрируются на рис. 2 и 3 соответственно.

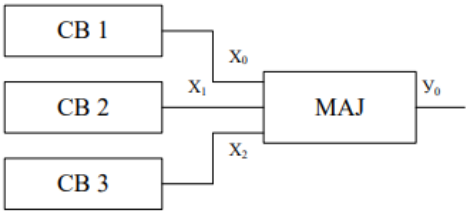


Рис. 2. Модульное резервирование

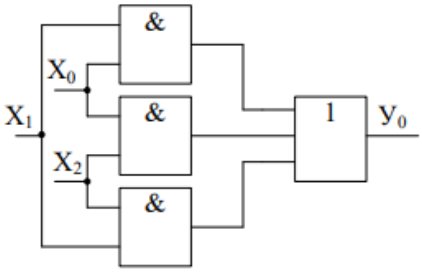


Рис. 3. Схема мажоритарного элемента для трех входов

Рассматриваемый подход сопряжён с рядом существенных ограничений. Прежде всего, он не способен корректировать ошибки, возникающие одновременно в двух и более функциональных блоках. Кроме того, метод практически неэффективен в отношении накопленных сбоев, которые проявляются постепенно в ходе длительной эксплуатации. Чтобы компенсировать этот недостаток, на практике TMR часто дополняют процедурой скраббинга – регулярной реконфигурацией или «очисткой» схемы, направленной на устранение накопленных одиночных сбоев без прерывания основного вычислительного процесса.

Ещё одно критическое ограничение – значительное увеличение аппаратных ресурсов: объём логических ячеек, объём памяти, число процессорных модулей и других компонентов возрастает более чем в три раза по сравнению с нерезервированной реализацией. Это неизбежно приводит к пропорциональному росту энергопотребления, что особенно критично для энергоограниченных систем.

Временное резервирование представляет собой гибридную стратегию, сочетающую аппаратную избыточность с временным дублированием операций, как показано на рисунке 4 [3].

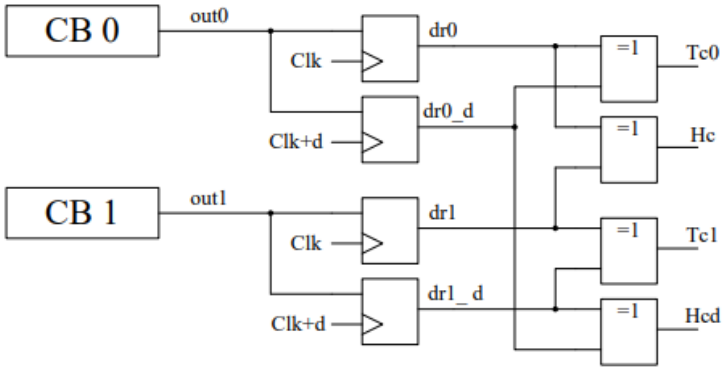


Рис. 4. Схема временного резервирования

В схеме используются два резервных комбинационных блока – $CB0$ и $CB1$. При возникновении сбоя в $CB0$ система активирует сигналы $Tc0$ и Hcd, переводя их в логическую «1», тогда как $Tc1$ и Hc устанавливаются в «0». Аналогичным образом, сбой в $CB1$ фиксируется по состоянию $Tc1 = 1$, $Hcd = 1$, при $Tc0 = 0$ и $Hc = 0$. На основе этих сигналов конечный автомат выбирает исправный блок для дальнейшей работы. Однако при определённой комбинации – $Tc0 = 0$, $Hc = 0$, $Tc1 = 1$, $Hcd = 1$ – система теряет возможность однозначно идентифицировать повреждённый узел. Это делает метод беспомощным перед так называемыми «защёлкнутыми» ошибками, когда выход элемента фиксируется в состоянии логического нуля или единицы независимо от входных данных. Именно этот недостаток стал толчком к созданию усовершенствованной техники – дублирования с параллельным обнаружением ошибок, представляющей собой модернизированную версию временного резервирования и способную выявлять устойчивые последствия одиночных сбоев.

Отдельного внимания заслуживает кодирование информации в алгоритмах, ориентированное на защиту данных в условиях шумных или подверженных помехам каналов передачи. Суть метода заключается в введении избыточности на этапе записи или передачи: к полезной информации добавляются специальные контрольные биты – например, контрольная сумма, циклический избыточный код (CRC) или корректирующий код (ECC). При чтении или приёме эти дополнительные данные используются для обнаружения и, при возможности, исправления искажений, возникших в процессе хранения или передачи [4].

Реконфигурация системы – это регулярно повторяющаяся процедура, при которой конфигурационная память ПЛИС полностью обновляется за счёт загрузки эталонного образа конфигурации из внешней, радиационно-стойкой памяти, устойчивой к воздействию ионизирующего излучения. Параллельно с этим широко используются алгоритмы корректирующего кодирования, повышающие надёжность передачи и хранения данных в условиях повышенного радиационного фона.

Основная цель конструктивно-технологических методов – снизить прямое влияние ионизирующего излучения на чувствительные элементы микросхемы. Это достигается двумя путями: либо установкой специальных экранов, поглощающих или рассеивающих радиацию, либо целенаправленной модификацией параметров транзисторов на основе глубокого понимания физических процессов, происходящих при взаимодействии частиц с полупроводниковой структурой.

Однако на практике внедрение различных форм избыточности – аппаратной, временной или информационной – в архитектуру системы зачастую оказывается более эффективным и гибким решением, чем разработка специализированных радиационно-стойких чипов. В связи с этим аппаратно-программные подходы к обеспечению отказоустойчивости приобретают особую значимость для инженеров, проектирующих цифровые устройства на базе ПЛИС, особенно в условиях экстремальных радиационных нагрузок.

Заключение

Архитектурные и алгоритмические стратегии обеспечения отказоустойчивости –такие как тройное модульное резервирование (TMR), временная и пространственная избыточность, применение корректирующих кодов и динамическая реконфигурация –формируют основу современного подхода к защите электронных систем от последствий радиационного облучения. Эти методы воплощают принцип Radiation Hardening by Design (RHBD), позволяя достигать требуемого уровня надёжности даже при использовании стандартных или коммерчески доступных технологических платформ, минуя необходимость в дорогих специализированных (rad-hard) компонентах.

Важнейший вывод исследования заключается в том, что ни один из перечисленных методов не может выступать в роли универсального «серебряной пули». Например, TMR, несмотря на широкое распространение, сопряжён с существенным ростом потребляемой мощности и занимаемой площади кристалла, а также уязвим к одновременным сбоям в нескольких резервированных каналах. В свою очередь, временное резервирование может оказаться малоэффективным

в ситуациях, где диагностика ошибок затруднена из-за особенностей временной логики или сценария функционирования. Наибольшая устойчивость достигается при многоуровневом, иерархическом сочетании различных техник: например, использование TMR для наиболее критичных логических блоков в паре со скраббированием памяти для устранения накопленных битовых инверсий и помехоустойчивым кодированием при передаче данных создаёт взаимодополняющую защитную архитектуру, в которой недостатки одного механизма компенсируются достоинствами другого.

Следовательно, разработка радиационно-стойкой электроники следует рассматривать как задачу системного инженерного проектирования, предполагающую осознанное управление противоречивыми требованиями: надёжностью, быстродействием, энергоэффективностью и плотностью интеграции. Оптимальный набор и параметры методов резервирования должны подбираться с учётом конкретного спектра радиационных угроз (энергетического состава частиц, орбитальных условий) и степени функциональной важности отдельных подсистем. Именно такой сбалансированный, адаптивный и комплексный подход –встроенный в единый цикл проектирования от архитектуры до верификации –обеспечивает создание электронных решений, способных безотказно выполнять свои задачи на протяжении всего расчётного срока службы в жёстких условиях космической среды[5].

Список литературы

1. Черняев, А. П. Взаимодействие ионизирующего излучения с веществом / А. П. Черняев. – Москва, 2004. – 152 с.

2. Полесский, С. Н. Обеспечение радиационной стойкости аппаратуры космических аппаратов при проектировании [Электронный ресурс] / С. Н. Полесский, В. В. Жаднов, М. А. Артюхова, В. Ф. Прохоров // Компоненты и технологии. – 2010. – № 9. – С. 93–97. – URL: http://www.kit-e.ru/assets/files/pdf/2010_09_093.pdf

3. Мироненко, Л. Повышение радиационной стойкости интегральных схем. Конструктивные методы на базе промышленной технологии [Электронный ресурс] / Л. Мироненко, В. Юдинцев // Электроника. – 2012. – № 8. – С. 74–87. – URL: https://www.electronics.ru/files/article_pdf/0/article_3207_324.pdf

4. Мироненко, Л. Повышение радиационной стойкости интегральных схем: конструктивные методы на базе промышленной технологии [Электронный ресурс] / Л. Мироненко, В. Юдинцев // Электроника. – 20XX. – № 3. – URL: https://www.electronics.ru/files/article_pdf/3/article_3513_347.pdf

5. Ладыгин, Е. А. Оценка эквивалентности воздействия гамма- и электронного излучений на полупроводниковые структуры / Е. А. Ладыгин, А. П. Галеев, К. И. Таперо, Т. Ю. Курунова // Шумовые и деградационные процессы в полупроводниковых приборах : материалы науч.-техн. семинара. – Москва : МНТОРЭС им. А. С. Попова, 1995.

6. Зольников, К. В. Создание специальной радиационно-стойкой электронной компонентной базы / К. В. Зольников, В. А. Фирононов // Архетип человека и будущее человечества : материалы междунар. науч.-практ. конф. – Воронеж, 2013. – С. 206–212.

Анализ состояния вопроса современных компьютерных моделей для повышения радиационной стойкости микросхем в условиях космоса

А.В. Шпинев, email: Andr0ne0@yandex.ru [1]
А.Н. Потапов, email: potapovan@vglta.vrn.ru [1,2]

[1] ФГБОУ ВО «Воронежский государственный лесотехнический университет имени Г.Ф. Морозова»

[2]ВУНЦ ВВС «ВВА имени проф. Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *В статье представлен комплексный анализ современных компьютерных моделей и программных средств, применяемых для обеспечения и повышения радиационной стойкости интегральных микросхем (ИМС), предназначенных для эксплуатации в условиях космического пространства. Рассматривается эволюция и современное состояние систем автоматизированного проектирования электроники (EDA), включая схемотехническое моделирование (SPICE), проектирование на уровне регистровых передач (RTL), статический временной анализ (STA) и анализ целостности сигналов и питания (SI/PI). Особое внимание уделяется методам прогнозирования и оценки основных радиационных эффектов: накопленной ионизирующей дозы (TID) и одиночных событий (SEE), таких как сбои (SEU), переходные процессы (SET), функциональные прерывания (SEFI) и защёлкивание (latch-up). Описаны подходы к моделированию этих эффектов с использованием TCAD-симуляторов, а также специализированных пакетов для расчёта взаимодействия частиц с веществом –Geant4 и SRIM, позволяющих оценить линейную передачу энергии (LET) и глубину пробега ионов. Показана роль физико-топологического моделирования при оптимизации радиационно-стойких технологий, в частности структур «кремний на изоляторе» (SOI). В работе обосновывается необходимость перехода от разрозненных инженерных решений к сквозной методологии проектирования, объединяющей физическое понимание радиационных процессов, численное моделирование и методы схемотехнического упрочнения (RHBD).*

Ключевые слова: *радиационная стойкость, интегральные микросхемы, космическая аппаратура, тяжёлые заряженные частицы (ТЗЧ), одиночные эффекты (SEE), одиночный сбой (SEU), защелкивание*

© Шпинев А. В., Потапов А.Н., 2026

(latch-up), поглощённая доза, дозовые эффекты, метод секторирования, массовая защита, радиационно-стойкие ИМС, радиационно-устойчивые ИМС, COTS.

Введение

Эксплуатация электронной компонентной базы (ЭКБ) в условиях космического пространства сопряжена с комплексным негативным воздействием ионизирующих излучений, что предъявляет повышенные требования к её надёжности и устойчивости. Космическая радиация, включающая потоки тяжёлых заряженных частиц (ТЗЧ), протонов, электронов и гамма-квантов, вызывает в интегральных микросхемах (ИМС) широкий спектр деградационных явлений и сбоев. К ним относятся как постепенные дозовые эффекты, приводящие к смещению пороговых напряжений и росту токов утечки, так и одиночные эффекты, такие как одиночные сбои, переходные процессы и опасные явления защелкивания, способные вызвать катастрофические отказы бортовой аппаратуры. В связи с этим обеспечение радиационной стойкости является критически важной задачей для разработки космических аппаратов (КА), решение которой требует системного подхода, включающего применение специализированных типов ИМС, строгую классификацию их стойкости и точный расчёт экранирования.

1. Инструменты и методологии проектирования электронных систем

На стадии проектирования интегральных схем активно применяются программные решения в рамках концепции Electronic Design Automation (EDA). Эти решения позволяют смоделировать и проверить корректность работы будущего устройства до его физического производства. Среди ключевых подходов выделяют:

– Схемотехническое моделирование (SPICE) – позволяет имитировать поведение аналоговых компонентов (транзисторов, резисторов, конденсаторов) при изменении внешних условий: температуры, напряжения, частоты и др.;

– Моделирование на уровне RTL – используется для описания цифровой логики с помощью языков проектирования аппаратуры (Verilog, VHDL); проверка функциональности осуществляется через симуляторы (например, Cadence Xcelium или Synopsys VCS);

– Логический синтез и статический временной анализ (STA) – в рамках этих процедур RTL-описание преобразуется в логическую схему с последующей проверкой временных характеристик (например, Synopsys Design Compiler – для синтеза, Synopsys PrimeTime – для STA),

что особенно важно при разработке устройств для экстремальных условий, включая космические миссии;

– Оценка целостности сигналов и электропитания (SI/PI) – позволяет выявить возможные нарушения передачи данных, влияние наводок, падения напряжения в цепях питания. Устойчивость к этим эффектам особенно критична для изделий, эксплуатируемых в условиях повышенного радиационного фона.

2. Моделирование на уровне RTL

Уровень передачи регистров (Register-Transfer Level, RTL) служит основным уровнем абстракции при проектировании цифровых систем. Он отражает поведение схемы с точки зрения передачи данных между регистрами в тактовые моменты времени.

На этом уровне проект описывается с помощью языков описания аппаратуры (HDL), таких как Verilog или VHDL. Не все конструкции HDL применимы на данном этапе: используется так называемое синтезируемое подмножество, корректно интерпретируемое средствами логического синтеза. Результатом синтеза становится схема на логических элементах, пригодная для дальнейшей реализации.

RTL-проектирование базируется на парадигме синхронной логики и включает три ключевых элемента:

– регистры – для хранения состояния;

– комбинационная логика – для вычисления новых значений данных;

– тактовый сигнал – для управления временем обновления состояния [3].

3. Синтез и статический временной анализ

Статический временной анализ (STA) – это эффективный метод оценки временных характеристик цифровой синхронной схемы без её полной функциональной симуляции. Он позволяет определить соблюдение временных ограничений (например, установочных и удерживающих времён) в условиях различных технологических отклонений (напряжение, температура, процессы изготовления).

Поскольку производительность современных ИС напрямую зависит от тактовой частоты, STA применяется на всех ключевых этапах проектирования – от синтеза до размещения и трассировки. В отличие от полномасштабного SPICE-моделирования, STA использует упрощённые модели задержек (обычно в формате SDF или Liberty), что обеспечивает значительное ускорение расчётов при сохранении достаточной точности. Именно эта особенность делает STA неотъемлемой частью современных EDA-потоков.

4. Оценка целостности сигналов и питания

Анализ целостности сигналов (SI) и системы питания (PI) осуществляется в два основных этапа:

Предтопологический анализ выполняется до или параллельно с разработкой трассировки печатной платы. Он позволяет:

– спроектировать согласующие цепи между драйверами и приёмниками;

– определить требуемый волновой импеданс линий передачи;

– оценить влияние параметров трассировки на качество сигнала.

– Полученные результаты затем используются при настройке правил проектирования в САПР (например, ширина проводников, толщина слоёв, геометрия полигонов питания/земли).

– Посттопологический анализ проводится после завершения топологии. Он включает:

– анализ распределения постоянного и переменного тока в цепях питания (IR-drop, импеданс петли питания);

– моделирование отражений, перекрёстных наводок и искажений формы сигнала в высокоскоростных линиях.

Рекомендуется начинать SI-анализ только после верификации и при необходимости – корректировки системы питания, поскольку нестабильное напряжение может вносить существенные искажения даже в исправную логическую структуру.

5. Моделирование влияния космической радиационной среды на полупроводниковые устройства

Воздействие космического излучения на микросхемы классифицируют на два основных вида:

– Накопленная ионизирующая доза (TID – Total Ionizing Dose).

Это постепенное накопление положительного заряда в диэлектрическом слое оксида затвора МОП-структур под действием ионизирующего излучения (в основном γ-квантов и электронов). В результате возникает смещение порогового напряжения, рост токов утечки и, при достижении критического уровня, – функциональный сбой устройства.

На этапе симуляции (например, в SPICE) эффекты TID учитываются путём корректировки параметров транзисторных моделей: изменяются значения порогового напряжения (Vth), подпорогового наклона, токов утечки и других ключевых характеристик в зависимости от эквивалентной поглощённой дозы. Для более глубокого анализа применяются инструменты физического моделирования, в частности Sentaurus TCAD (Synopsys), позволяющие детально описать процессы

захвата заряда в интерфейсных ловушках и объёмных дефектах диэлектрика [4].

– Одиночные радиационные эффекты (SEE – Single-Event Effects).

Возникают при прямом попадании высокоэнергетической частицы (тяжёлого иона или протона) в активную область кристалла. В зависимости от места и энергии удара реализуются различные типы отказов:

Одиночный переходный сбой (SET – Single-Event Transient) – кратковременный импульс напряжения на выходе комбинационной логики, вызванный генерацией электронно-дырочных пар в объёме полупроводника. В симуляторах моделируется путём введения заданного токового импульса (обычно гауссовой или экспоненциальной формы) в нужную узловую точку схемы на этапе временного анализа.

Одиночный инверсный сбой (SEU – Single-Event Upset) – изменение логического состояния ячейки памяти (например, триггера или элемента SRAM-матрицы) без разрушения структуры. Как правило, имитируется с помощью метода fault injection – принудительной инверсии бита в симуляции с последующим отслеживанием влияния на общую функциональность.

Функциональный сбой (SEFI – Single-Event Functional Interrupt) – более серьёзная аномалия, при которой микросхема теряет работоспособность и требует внешней инициализации (например, сброса или перезагрузки). Часто проявляется в блоках управления, конфигурации или коммуникационных интерфейсах.

Защелкивание (Latch-up) – катастрофический эффект, связанный с включением паразитного биполярного тиристора в КМОП-структуре, что приводит к резкому росту тока и перегреву. Моделирование latch-up проводится как на уровне схемотехнической симуляции (SPICE), так и на физическом – с использованием TCAD-инструментов, для анализа распределения носителей и формирования петли положительной обратной связи [5].

6. Физико-технологическое моделирование (TCAD)

Сущность подхода: TCAD (Technology Computer-Aided Design) – это класс программных средств, осуществляющих имитацию микро- и наноразмерных процессов формирования полупроводниковых структур: диффузии примесей, ионной имплантации, термического окисления, травления и т.д. Моделирование ведётся на основе уравнений физики твёрдого тела – дрейф-диффузионных уравнений, уравнения Пуассона и кинетических моделей.

Применение в космической электронике: С помощью TCAD можно прогнозировать поведение прибора в экстремальных условиях, включая

радиационное воздействие на уровне кристаллической решётки и дефектной структуры. Например, при проектировании радиационно-стойких платформ часто применяется технология SOI (Silicon-on-Insulator), которая кардинально снижает чувствительность к TID и latch-up. Оценка её эффективности, оптимизация толщины buried oxide и подбор параметров ионной имплантации активно выполняются в среде Synopsys Sentaurus TCAD.

Комплексные программные решения для радиационного анализа

Процедура моделирования воздействия космического излучения представляет собой многоэтапный сквозной поток, объединяющий инструменты различного назначения.

Geant4 – открытая программная платформа для моделирования взаимодействия высокоэнергетических частиц с веществом. Она разработана международным консорциумом и широко применяется в ядерной физике, медицине, аэрокосмической отрасли и микроэлектронике.

Основные возможности Geant4:

Построение трёхмерной геометрии объектов (включая многослойные и неоднородные структуры) с заданием материалов и их свойств;

Трассировка траекторий первичных и вторичных частиц с учётом всех физических процессов: электромагнитных, адронных, ядерных реакций;

Детектирование и сбор статистики по интересующим параметрам (энерговыделение, плотность ионизации, число вторичных частиц и др.);

Визуализация траекторий, геометрии и распределений энергии с помощью встроенных графических модулей;

Поддержка пользовательских расширений: от простых сценариев через макрокоманды до полного программного контроля через C++ API.

Диапазон энергий, охватываемый библиотеками Geant4, простирается от нескольких мэВ (тепловые нейтроны) до сотен ГэВ, а в отдельных режимах – до 100 ТэВ. Каждый класс взаимодействий (комптоновское рассеяние, фотоэффект, ядерные спаллации и др.) реализован через несколько конкурирующих моделей, доступных в виде готовых физических списков – наборов согласованных моделей, протестированных для конкретных прикладных задач.

Исходный код Geant4 распространяется по лицензии с открытым доступом и обладает высокой читаемостью и модульностью. В поставке включены учебные примеры и полноценные демонстрационные приложения, которые можно использовать как основу для разработки собственных решений. При этом разработчик

волен либо адаптировать одну из существующих конфигураций, либо создать собственную – с нуля – с учётом специфики анализируемой задачи [6].

SRIM (Stopping and Range of Ions in Matter) – это набор вычислительных программ, ориентированных на расчёт взаимодействия ионных пучков с конденсированными средами. Хотя изначально создавался для задач ионной имплантации и материаловедения, сегодня SRIM находит применение и в радиационной стойкости – для оценки линейной передачи энергии (LET), глубины пробега частиц и плотности повреждений. Интерфейс работающей программы SRIM представлен на рисунке.

Ключевые функции SRIM:

– Расчёт пробега (range) и тормозной способности (stopping power) ионов в элементах и соединениях (включая сложные многослойные структуры);

– Анализ распределения внедрённых примесей при ионной имплантации;

– Оценка объёма повреждённой области (количества смещённых атомов – vacancies и interstitials);

– Моделирование процессов распыления (sputtering) поверхности под действием ионного облучения;

– Прогнозирование пропускания ионных пучков через газовые или твёрдые ослабители (например, в системах энергетической фильтрации).

История развития методов расчёта торможения ионов начинается с работ начала XX века (в частности, с теории Бете–Блоха), однако точное предсказание параметров остановки невозможно без эмпирических корректировок. В SRIM применяется полуэмпирический подход, сочетающий численное интегрирование уравнений траекторий с бинарными столкновениями атомов, калиброванными по экспериментальным данным. Несмотря на упрощения (отсутствие учёта кристаллической структуры, коллективных эффектов), метод обеспечивает хорошее соответствие реальным измерениям для аморфных и поликристаллических материалов – что делает его незаменимым инструментом на ранних этапах проектирования радиационно-стойких ИС [7].

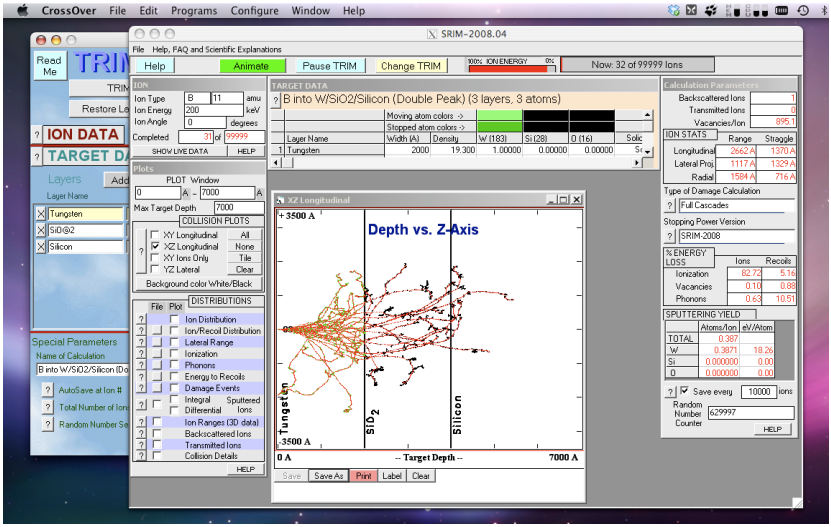


Рисунок. Пример работы программы SRIM

Заключение

Настоящее исследование было направлено на всестороннее изучение методов, технологий и стратегий, применяемых при разработке и обеспечении радиационной устойчивости электронных систем, эксплуатируемых в экстремальных условиях космической среды. По итогам проведённого анализа сформулированы следующие ключевые положения.

Проектирование как сквозная система надёжности. Современный подход к созданию бортовой электроники перестал быть набором изолированных инженерных процедур –он эволюционировал в единую, многоуровневую парадигму, где радиационная стойкость рассматривается как обязательное требование на всех фазах жизненного цикла изделия –от концепции до вывода из эксплуатации. Современные средства автоматизированного проектирования (EDA), включая SPICE-симуляцию, RTL-синтез, статический временной анализ (STA) и моделирование целостности сигналов и питания (SI/PI), формируют технологическую основу, позволяющую не только добиваться высокой функциональной эффективности, но и заранее оценивать уязвимость проектных решений к внешним дестабилизирующим воздействиям –ещё до изготовления физического прототипа.

Роль специализированного моделирования в прогнозировании радиационных эффектов. Обеспечение работоспособности электроники

в условиях космического излучения невозможно без перехода от стандартных методов верификации к физически обоснованным симуляциям. Анализ таких явлений, как накопленная ионизирующая доза (TID) и одиночные радиационные события (SEE), стал возможен благодаря развитию и синхронизации специализированных программных комплексов:

– TCAD-моделирование даёт возможность исследовать процессы деградации на уровне полупроводниковой структуры и дефектов кристаллической решётки, что позволяет оптимизировать технологические маршруты (например, внедрение SOI-платформ);

– Платформы Geant4 и SRIM обеспечивают расчёт энергетических характеристик взаимодействия высокоэнергетических частиц с материалами, включая линейную передачу энергии (LET) и объём вторичной ионизации. Полученные данные служат фундаментальной основой для последующих схемотехнических и архитектурных симуляций.

Аппаратно-программное упрочнение как стратегическая альтернатива. Исследование подтвердило, что на фоне сокращения предложения специализированных rad-hard компонентов и стремительного роста сложности микросхем доминирующей становится стратегия Radiation-Hardening-by-Design (RHBD). Такие техники, как тройное модульное резервирование (TMR), помехоустойчивое кодирование, периодическое скраббирование памяти и архитектурная избыточность, позволяют достигать требуемого уровня надёжности даже при использовании коммерческих (COTS) или промышленных технологий, обеспечивая при этом гибкость и экономическую целесообразность.

Системный подход и управление проектными компромиссами. Разработка радиационно-стойкой электроники неизбежно требует балансировки между противоречивыми параметрами: надёжностью, производительностью, энергоэффективностью, стоимостью и площадью кристалла. Успешное разрешение этих противоречий возможно лишь в рамках целостной методологии, объединяющей глубокое понимание физики радиационных воздействий, современные инструменты численного моделирования и практические методы RHBD.

Список литературы

1. Черняев, А. П. Взаимодействие ионизирующего излучения с веществом / А. П. Черняев. – Москва, 2004. – 152 с.

2. SemiEngineering [Электронный ресурс] : информационный портал по проектированию электронных систем. – URL:

https://semiengineering.com/knowledge_centers/eda-design/verification/simulation/spice.

3. SemiEngineering [Электронный ресурс] : информационный портал по проектированию электронных систем. – URL: https://semiengineering.com/knowledge_centers/eda-design/definitions/register-transfer-level.

4. Stanojev, Z. Total Ionizing Dose (TID) Impact on Basic Amplifier Stages [Электронный ресурс] / Z. Stanojev, I. Kostic, I. Popovic. – 2022. – URL:https://www.researchgate.net/publication/366143060_Total_Ionizing_Dose_TID_Impact_on_Basic_Amplifier_Stages

5. Single Event Effects (SEE) [Электронный ресурс] / Cogenda. – URL: https://www.cogenda.com/article/SEE

6. Geant4 [Электронный ресурс] : toolkit for the simulation of the passage of particles through matter / The Geant4 Collaboration. – URL: https://geant4.org/

7. Ziegler, J. F. The Stopping and Range of Ions in Matter (SRIM) [Электронный ресурс] / J. F. Ziegler, J. P. Biersack, M. D. Ziegler. – URL: http://srim.org/

8. Таперо, К. И. Радиационные эффекты в кремниевых интегральных схемах космического применения. Основы радиационной стойкости изделий электронной техники космического применения / К. И. Таперо. – Москва : Дом МИСиС, 2011. – 252 с.

9. Таперо, К. И. Основы радиационной стойкости изделий электронной техники. Радиационные эффекты в изделиях электронной техники / К. И. Таперо, С. И. Диденко. – Москва : Дом МИСиС, 2013. – 349 с.

Разработка учебно-тренировочных средств на платформе Unity для подготовки специалистов к работе со сложными техническими системами с использованием технологий виртуальной реальности

С.В. Щегольков email: nauchnajarota@yandex.ru
А.И. Рыжов email: nauchnajarota@yandex.ru

Межвидовой центр подготовки и боевого применения войск РЭБ
(учебный и испытательный)

Аннотация. *В статье исследуются методологические и технологические аспекты разработки иммерсивных учебно-тренировочных средств на основе среды разработки Unity. Рассмотрены ключевые преимущества платформы: поддержка систем виртуальной и дополненной реальности, работа с инженерными данными и взаимодействие с большинством современных платформ. Представлена методология проектирования, подкрепленная практическими примерами. Приведены примеры, доказывающие эффективность применения виртуальных симуляторов. Приведены основные перспективные направления развития, такие как внедрение искусственного интеллекта и создание цифровых двойников.*

Ключевые слова: *Unity, учебно-тренировочные средства, иммерсивное обучение, виртуальная реальность, симулятор, подготовка специалистов, цифровой двойник.*

Введение

В наше время создается тенденция усложнения технических систем и коммуникаций, что создает проблемы при подготовке специалистов. Традиционные методы обучения имеют весомых недостатков, которые требуют решения. Эти проблемы связаны с травматизмом во время обучения, высоких затрат на создание физических тренажеров и риском их повреждения. Особенно актуальны эти проблемы в сфере авиации, энергетики и промышленности [1,4]

Создание виртуальных симуляторов с помощью различных средств программирования и разработки позволяют решить эти проблемы. Программная среда Unity является лучшим вариантом для разработки виртуальных симуляторов благодаря ряду важных преимуществ:

- Возможность создания высокодетализированных пространств.
- Поддержке систем виртуальной и дополненной реальности.

© Щегольков С.В., Ионченко Е.С., 2026

– Возможностью интеграции с другими программами без серьезных изменений в проекте[5].

Цель статьи – провести анализ методик и технологических аспектов разработки учебно-тренировочных средств на основе среды программирования Unity, обобщение примеров практического опыта применения и оценка эффективности созданных симуляторов.

Задачи:

– Выявить ключевые технологические преимущества Unity.
– Разработать обобщенную методологию проектирования.
– Проанализировать успешные примеры применения.
– Провести сравнительный анализ эффективности иммерсивного обучения

1. Технологические основы Unity для профессионального моделирования

Среда разработки Unity, в частности пакет Unity Industry, предоставляет специализированные инструменты, предназначенные для решения промышленных и образовательных задач. Одним из ключевых аспектов разработки достоверных симуляторов является использование точных цифровых моделей инженерных объектов. Для этого применяется инструмент Asset Transformer Toolkit, который позволяет преобразовывать высокополигональные модели и выполнять их оптимизацию без потери визуальной точности, сохраняя при этом геометрическую достоверность и корректность наложения материалов.

Важной особенностью среды разработки Unity является поддержка реалистичного физического моделирования. Встроенные алгоритмы обеспечивают высокоточную симуляцию механических взаимодействий и динамики твёрдых тел. Взаимодействие между элементами симуляционной системы реализуется с помощью скриптов, разработанных на языке программирования C#. Такой подход позволяет гибко настраивать поведение отдельных компонентов модели и реализовывать различные сценарии функционирования системы, включая моделирование нештатных ситуаций.

Кроме того, платформа Unity обеспечивает поддержку иммерсивных технологий и обладает высокой степенью мультиплатформенности. Используемые средства разработки позволяют компилировать проекты для большинства современных операционных систем, что значительно расширяет возможности создания учебно-тренировочных средств и интерактивных симуляторов [1].

Таблица 1

Сравнение традиционного обучения и обучения на Unity-тренажерах

Критерий	Традиционное обучение	Обучение на Unity-тренажере
Безопасность	Риск травм и поломки оборудования	Полная безопасность, возможность отработки ошибок
Стоимость	Высокая (ГСМ, износ, расходники)	Сокращение эксплуатационных расходов после развертывания
Адаптивность сценариев	Ограниченный набор ситуаций	Моделирование любых, в том числе аварийных, сценариев
Обратная связь	Субъективная оценка инструктора	Мгновенная, объективная и детализированная оценка
Масштабируемость	Ограничена количеством инструкторов и техники	Неограниченное число рабочих мест для обучения

2. Методология разработки учебно-тренировочного комплекса на Unity

Этап 1: Анализ требований и проектирование.

На данном этапе происходит определение основных целей создания УТС, формирование конкретных задач. Также производится сбор

данных: получение или создание трехмерных моделей, создание и оформление документации и регламентов.

Этап 2: Создание виртуальной среды:

Этот этап предусматривает внедрение и оптимизация трехмерных моделей, а также создание необходимых материалов и настройка освещения. Проводится написание алгоритмов взаимодействия(скриптов) и логика работы узлов на базе языка программирования C#. Также интегрируется физика взаимодействия.

Этап 3: Реализация педагогического сценария:

Производится разработка сценария обучения специалистов, в зависимости от поставленных задач сценарий может быть линейным и нелинейным. Внедряется система подсказок и теоретических справок.

Этап 4: Тестирование, внедрение и интеграция

Заключающий этап разработки УТС представляет собой тестирование итогового продукта с участием специалистов. На основе проведенных испытаний происходит сбор и анализ метрик эффективности.

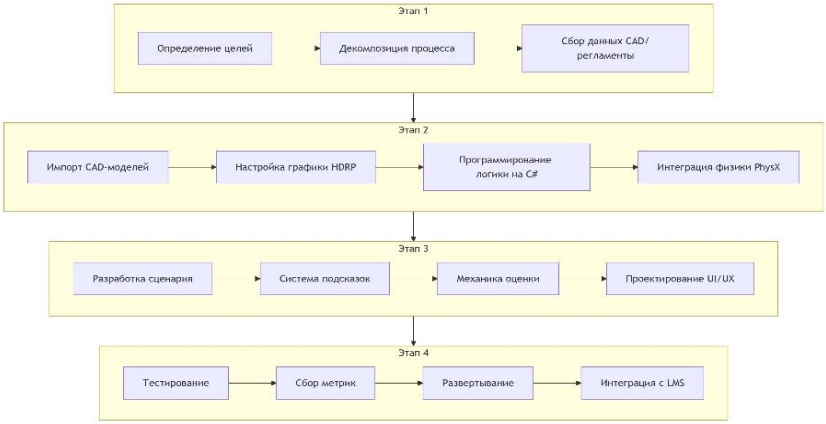


Рис. 1. Методология разработки учебно-тренировочного комплекса на Unity

3. Практические примеры и оценка эффективности

В настоящее время существует достаточно примеров применения технологий виртуальной и дополненной реальности в сфере авиации, промышленности и медицине. Примеры таких внедрений представлены в табл. 2 [2,5].

Таблица 2

Примеры внедрения технологий виртуальной и дополненной реальности

Отрасль	Компания	Метод применения
Авиация	CAE, Boeing, Emirates, Qatar Airways	VR-симуляторы для обучения пилотов и бортпроводников. Пилоты отрабатывают аварийные ситуации,а бортпроводники – процедуры безопасности в виртуальном салоне.
Медицина	Osso VR,Vipaar, Onkos Surgical	VR-платформы для тренировки хирургов и AR-системы для удаленного консультирования во время операций или точного удаления опухолей.
Промышленность	Layton Construction, ПАО «Газпром-нефть»	VR для создания виртуальных макетов зданий и удаленных инспекций строящихся объектов, что экономит время и средства.
Промышленность	ПАО «ФСК ЕЭС»	AR-инструкции для монтажа сложного оборудования, что снижает количество ошибок.

Более детально рассмотрим пример внедрения виртуальных симуляторов базе конструкторского технологического центра завода, специализирующегося на выпуске тяжёлых редукторов. По данным проведенного исследования[6], формирование достаточных навыков у

токарей занимает до 180 часов, а уровень брака достигает 3,7%. В целях уменьшения времени подготовки специалистов и снижения брака был разработан виртуальный симулятор в среде разработки Unity. В симуляторе была реализована система поддержки взаимодействия и физически корректного отклика материалов. Также была внедрена система сценариев, собранных на основе отказов за последние три года. Результатом внедрения стало снижения сроков подготовки до 42 часов, а процент брака снизился до 1,1[2].

4. Вызовы и перспективы развития

Важными вызовами для внедрения VR/AR технологий является ряд факторов, которые замедляют интеграцию современных виртуальных симуляторов в большинство сфер деятельности. На начальных этапах требуются значительные инвестиции в программирование, моделирование и проектирование учебно-тренировочных средств. также негативных фактором интеграции является небольшое количество специалистов со специфическими навыками в предметной области [6]. Не стоит забывать про оборудование и технику, необходимую для реализации таких проектов, её высокая стоимость и недоступность в условиях санкций значительно снижает процесс интеграции виртуальных симуляторов.

На данный момент наиболее перспективными проектами внедрения технологий виртуальной реальности является интеграция искусственного интеллекта для создания адаптивных сценариев и систем оценки ошибок. Также свое развитие получила технология цифровых двойников, созданных на базе реального оборудования с моделями аналитики и предиктивного обслуживания. Такие модели позволяют точно смоделировать современное оборудование для подготовки специалистов, её гибкость позволяет подстраивать процесс обучения индивидуально, с учетом совершенных ошибок.

Заключение

Среда разработки Unity является мощной платформой для разработки профессиональных учебно-тренировочных средств. Его преимущества – работа с инженерными данными, реалистичная физика, поддержка виртуальной и дополненной реальности, кроссплатформенность – позволяют создавать эффективные и точные симуляторы для сложной техники.

Обучение на таких тренажерах качественно улучшает подготовку специалистов: повышает скорость усвоения информации, сокращает время освоения навыков и минимизирует риски. Дальнейшее развитие связано с интеграцией искусственного интеллекта, созданием цифровых

двойников и развитием отраслевых экосистем, что выведет профессиональное обучение на новый уровень.

Список литературы

1. Акопов, А. С. Компьютерное моделирование : учебник и практикум для СПО / А. С. Акопов. –Москва : Юрайт, 2019. –389 с.

2. Дозорцев, В. М. Компьютерные тренажеры для обучения операторов технологических процессов / В. М. Дозорцев. –Москва : Синтег, 2009. –365 с.

3. Карпов, Ю. Г. Имитационное моделирование систем. Введение в моделирование с AnyLogic 5 / Ю. Г. Карпов. –Санкт-Петербург : БХВ-Петербург, 2005. –400 с.

4. Рогозин, К. И. Виртуальные компьютерные тренажеры в учебном процессе / К. И. Рогозин // Гарантии качества профессионального образования : тезисы докладов Международной научно-практической конференции. –Барнаул, 2010. –С. 152–154.

5. Якименко, О. В. Применение обучающих программ-тренажеров в обучении программированию / О. В. Якименко, А. Н. Стась // Вестник ТГПУ. –2009. –С. 54–56.

6. Старостина, В. А. AR- и VR-технологии в обучении производственного персонала / В. А. Старостина, В. В. Воробьев, М. А. Рагозина, Г. И. Юрковская // Молодые ученые в решении актуальных проблем науки : сб. материалов Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых. –Красноярск : Сибирский государственный университет науки и технологий им. акад. М. Ф. Решетнева, 2019. –С. 1178–1180.

Секция 3
Технологии обработки и защиты информации

Исследование устойчивости функционирования объекта критической информационной инфраструктуры

С. А. Будников, email: public.buser@bk.ru
Ю. А. Енютин, email: alexyen@mail.ru
В. М. Аксютин, email: vladimir84@mail.ru

ФАУ «ГНИИИ ПТЗИ ФСТЭК России»

Аннотация. *В данной работе разработана марковская модель и проведено исследование устойчивости функционирования кластера виртуализации высокой доступности при проведении против него компьютерных атак.*

Ключевые слова: *виртуализация, компьютерная атака, коэффициент нестационарной готовности, марковский случайный процесс, объект критической информационной инфраструктуры, отказоустойчивый кластер.*

Введение

В современных условиях обеспечение безопасности критической информационной инфраструктуры (КИИ) является приоритетной задачей. Имеющиеся данные подтверждают рост масштабных и целенаправленных атак на объекты КИИ Российской Федерации. В частности, результаты исследований отраслевых аналитических групп демонстрируют резкое увеличение числа воздействий, затрагивающих информационные операционные технологии, рост длительности простоев и значительный экономический ущерб. Так, 12 % кибератак привели к остановке производства более чем на неделю [1], каждая четвёртая атака (24 %) нарушила основную деятельность организаций, средний ущерб от успешной атаки составил от 1 до 5 млн руб. [2], 70 % компаний не смогли восстановить утраченные данные [3].

Все это определяет необходимость проведения исследований, направленных на повышение устойчивости функционирования объектов КИИ при проведении в отношении них компьютерных атак (КА).

Устойчивость сложных систем – это способность системы сохранять свои функциональные свойства, структуру и поведение в условиях внешних воздействий, изменений среды [4]. В контексте информационной безопасности устойчивость объектов КИИ означает их способность сохранять функционирование, адаптироваться и восстанавливаться при проведении в отношении них КА.

© Будников С. А., Енютин А. Ю., Аксютин В. М., 2026

Повышение устойчивости в таких условиях достигается путем консолидации ресурсов обработки и хранения данных на основе технологий кластеризации, динамического распределения запросов и виртуализации [5].

Известно, что объектами КИИ считаются информационные системы, информационно-телекоммуникационные сети и автоматизированные системы управления, нарушение работы которых может привести к серьезным социально-экономическим последствиям.

Российский опыт в области оценки устойчивости объектов КИИ базируется на использовании отечественных разработок в области информационных рисков и адаптации международных стандартов [6]. Практика оценки устойчивости объектов КИИ основывается на проверке соблюдения Требований по обеспечению безопасности значимых объектов КИИ [7].

Однако обоснование адекватных практических решений в большинстве случаев не поддержано теоретическими разработками, позволяющими оценить вероятностно-временные характеристики устойчивости объектов КИИ в условиях КА.

В связи с этим, еще большую актуальность приобретают вопросы, связанные с разработкой моделей и методик для оценки устойчивости информационных систем, информационно-телекоммуникационных сетей и автоматизированных систем управления, нарушение работы которых может привести к серьезным социально-экономическим последствиям.

Поэтому целью работы является разработка марковской модели и проведение исследования устойчивости функционирования объекта КИИ, представляющего собой кластер серверов виртуализации высокой доступности, при проведении против него КА.

Основная часть

Анализ ряда известных работ [8-12] показал, что авторы предлагают методологию, использующую системный анализ, графовые модели и нечеткую логику. Фокусируются на оценке киберустойчивости компьютерных сетей с использованием метода преобразования стохастических сетей, на создании конкретной эволюционной модели для оценки состояния объектов КИИ и поддержки управления системами. В статье [9] авторы предлагают общую методологию оценки устойчивости КИИ к угрозам информационной безопасности, описывая подходы для расчета устойчивости отдельных активов и систем в целом.

Однако в них не вводятся граничные значения вероятностно-временных характеристик значений надежности и устойчивости объектов КИИ. В [11] в качестве показателя, характеризующего

устойчивость компьютерной сети при воздействии КА, предлагается использовать коэффициент исправного действия, который отражает какую часть времени от всего учитываемого времени работы компьютерной сети она функционирует исправно. При этом не учитывается латентная стадия действий нарушителя в ходе КА.

В работе [5] приведено построение и анализ марковских моделей надежности отказоустойчивого кластера, использующего технологии виртуализации и миграции виртуальных машин (ВМ), для оценки влияния процессов миграции и восстановления на коэффициенты готовности системы. В качестве метрик используются коэффициенты стационарной и нестационарной готовности, позволяющие оценить доступность системы в установившемся и переходном режимах. Однако в качестве источников отказов КА не анализированы.

По аналогии с этой работой рассмотрим объект КИИ, представляющий отказоустойчивый кластер минимальной конфигурации, содержащий два физических сервера (основной и резервный), соединённых через коммутатор (см. рис. 1).

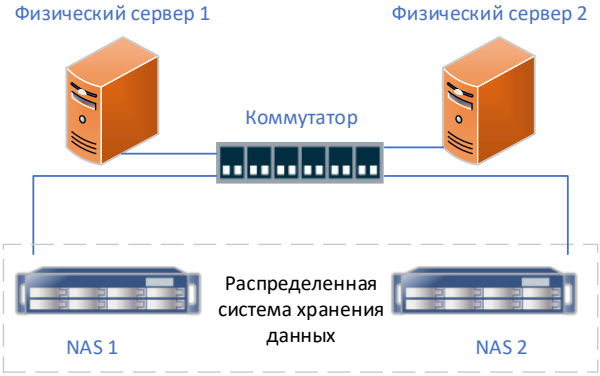


Рис. 1. Архитектура отказоустойчивого объекта КИИ

На каждом сервере развернута распределенная система хранения данных с синхронной репликацией. В кластере работает ВМ в режиме отказоустойчивости, что предполагает наличие её теневой копии на резервном сервере. Каждый сервер имеет доступ к распределённой системе хранения данных (синхронная репликация). При отказе основного сервера происходит автоматическое переключение на резервный без остановки вычислительного процесса. Устойчивость функционирования достигается за счет реализации автоматического перезапуска ВМ на исправном узле после отказа и поддержки двух

синхронизированных копий ВМ в памяти разных серверов, что обеспечивает непрерывность процесса без потери данных.

КА, проводимые в отношении объекта КИИ, могут привести к отказу сервера, диска и коммутатора, а также нарушить репликацию данных и запуск ВМ.

Введем следующие категории состояний объекта КИИ:

1. Исправные состояния (работоспособные, без отказавших узлов). Пример: состояние, в котором все компоненты кластера (серверы, диски, коммутатор) функционируют, и ВМ работает на основном сервере.

2. Работоспособные состояния с отказавшими узлами. Пример: состояние, в котором один из компонентов (например, диск или сервер) отказал, но система продолжает работать благодаря резервированию или миграции ВМ.

3. Неработоспособные состояния, в которых происходит автоматическое восстановление сервиса. Пример: состояние после отказа основного сервера, когда система автоматически запускает ВМ на резервном сервере (миграция с синхронизацией данных).

4. Неработоспособные состояния, в которых ожидается восстановление узлов. Пример: состояние, когда несколько компонентов отказали, и требуется ручное вмешательство для восстановления.

Ключевые компоненты, являющиеся объектами КА и учитываемые в этих состояниях:

1. серверы (основной и резервный), характеризуемые интенсивностью отказов λ_c и интенсивностью восстановлений μ_c;

– жёсткие диски (сетевые, с синхронной репликацией), характеризуемые интенсивностью отказов $\lambda_д$ и интенсивностью восстановлений: $\mu_д$;

– коммутатор (сетевое соединение), характеризуемый интенсивностью отказов: $\lambda_к$ и интенсивностью восстановлений: $\mu_к$;

– процессы синхронизации и миграции, характеризуемые интенсивностью восстановления данных (репликация) $\mu_р$ и интенсивностью восстановления ВМ после перезапуска (запуск на резервном сервере) $\mu_{вм}$.

С учетом этого обобщенный граф состояний подобного объекта КИИ приведен на рис. 2.

Важно подчеркнуть, что состояния отражают не только физические отказы, но и процессы синхронизации данных и миграции ВМ.

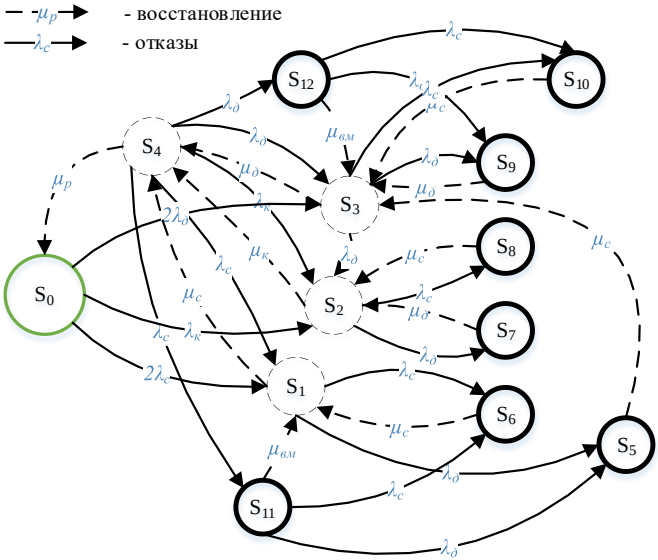


Рис. 2. Граф состояний объекта КИИ

Описания состояний кластера, возникающих в результате проведения КА, приведены в таблице.

Таблица

Описания состояний

Номер	Описание
S_0	Исходное состояние, все компоненты исправны
S_1	Отказал один из серверов, но ВМ продолжает работу на другом сервере
S_2	Отказ коммутатора, но система работает через альтернативные пути
S_3	Отказ диска, данные синхронизируются с резервной копией
S_4	Состояние миграции ВМ после отказа основного сервера (синхронизация памяти и данных)
S_5-S_{12}	Различные состояния восстановления, ожидания ремонта, повторной синхронизации и т.д.

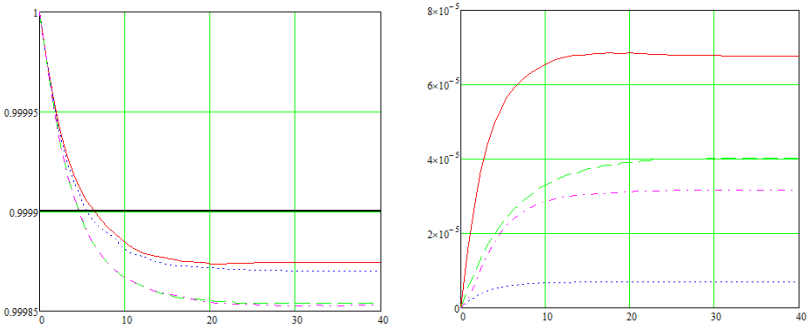
Как и в [13] будем считать, что функционирование кластера можно представить в виде марковского процесса с дискретными состояниями и

непрерывным временем. Для данного графа (см. рис.2) построим марковскую модель функционирования отказоустойчивого кластера с виртуализацией в условиях проведения на него КА, основанную на системе уравнений Колмогорова [14], для заданных начальных условий: $P_0(0)=1$, где $i=\overline{1,12}$. Будем считать, что работоспособным состоянием является состояние S_0. Система уравнений Колмогорова относительно средних времен переходов для графа, приведенного на рис.2 будет иметь вид (1):

$$\begin{cases}
\frac{dP_0(t)}{dt} = -(2\lambda_c+\lambda_к+2\lambda_д)P_0(t)+\mu_p P_4(t),\\
\frac{dP_1(t)}{dt} = -(\lambda_c+\mu_c+\lambda_д)P_1(t)+\lambda_c P_4(t)+2\lambda_c P_0(t)+\\
\qquad +\mu_{вм}P_{11}(t)+\mu_c P_6(t),\\
\frac{dP_2(t)}{dt} = -(\lambda_c+\mu_к+\lambda_д)P_2(t)+\mu_д P_7(t)+\mu_c P_8(t)+\\
\qquad +\lambda_к P_4(t)+\lambda_к P_0(t),\\
\frac{dP_3(t)}{dt} = -(\lambda_c+\mu_д+\lambda_д)P_3(t)+\mu_д P_9(t)+\mu_c P_{10}(t)+\mu_{вм}P_{12}(t)+\\
\qquad +\mu_c P_5(t)+\lambda_д P_4(t)+2\lambda_д P_0(t),\\
\frac{dP_4(t)}{dt} = -(2\lambda_c+\lambda_к+2\lambda_д+\mu_к)P_4(t)+\mu_c P_1(t)+\mu_к P_2(t)+\\
\qquad +\mu_д P_3(t),\\
\frac{dP_5(t)}{dt} = -\mu_c P_5(t)+\lambda_д P_1(t)+\lambda_д P_{11}(t),\\
\frac{dP_6(t)}{dt} = -\mu_c P_6(t)+\lambda_c P_1(t)+\lambda_c P_{11}(t),\\
\frac{dP_7(t)}{dt} = -\mu_д P_7(t)+\lambda_д P_2(t),\\
\frac{dP_8(t)}{dt} = -\mu_c P_8(t)+\lambda_c P_2(t),\\
\frac{dP_9(t)}{dt} = -\mu_д P_9(t)+\lambda_д P_3(t)+\lambda_д P_{12}(t),\\
\frac{dP_{10}(t)}{dt} = -\mu_c P_{10}(t)+\lambda_c P_3(t)+\lambda_c P_{12}(t),\\
\frac{dP_{11}(t)}{dt} = -(\lambda_c+\mu_{вм}+\lambda_д)P_{11}(t)+\lambda_c P_4(t),\\
\frac{dP_{12}(t)}{dt} = -(\lambda_c+\mu_{вм}+\lambda_д)P_{12}(t)+\lambda_д P_4(t),
\end{cases} \qquad (7)$$

где $P_i(t)$. – вероятность нахождения в i -ом состоянии, $i=\overline{0,12}$.

Решение данной системы однородных дифференциальных уравнений осуществим в среде MathCad при следующих исходных значений интенсивностей отказов и восстановлений $(1/час)$: $\lambda_c = 1.115 \times 10^{-5}$, $\lambda_д = 3.425 \times 10^{-6}$, $\lambda_к = 2.43 \times 10^{-6}$, $\mu_c = 0.33$, $\mu_д = 0.17$, $\mu_к = 0.33$, $\mu_р = 1$, $\mu_{вм} = 2$. Результат численных расчетов значений вероятности нахождения в работоспособном состоянии объекта КИИ от времени в часах для различных значений интенсивностей отказов и восстановлений приведен на рис. 2.



а б

а – в исправном состоянии, б – в работоспособных состояниях, но с отказавшими узлами

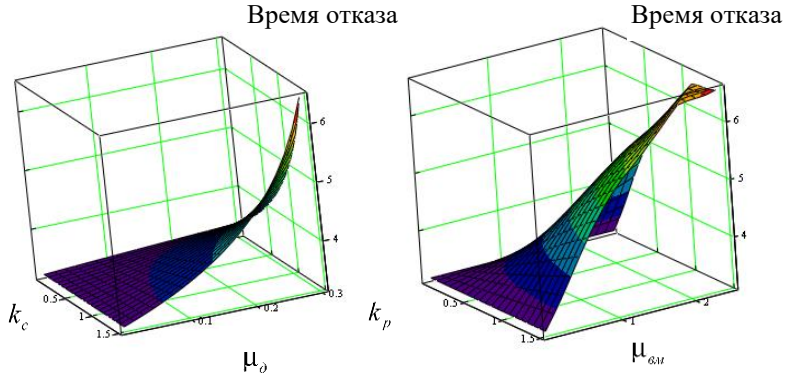
Рис. 3. Зависимость значений вероятностей нахождения в состояниях, час

Анализ поведения полученной зависимости позволяет ввести пороговое значение вероятности работоспособности объекта КИИ в конкретный момент времени, в который значения всех параметров, характеризующих способность выполнять заданные функции, еще соответствуют требованиям нормативной и технической документации. Значение минимально допустимого уровня нестационарной готовности объекта КИИ выберем равным $K_Г^{порог}(t) = P_0(t) = 0.9999$.

Из рис. 3 а) видно, что время пересечения графика вероятности работоспособного состояния порога $K_Г^{порог}(t) = 0.9999$ (время отказа кластера) определяется значениями параметров отказов и восстановлений серверов, дисков и коммутатора λ_c, $\lambda_д$, $\lambda_к$, μ_c, $\mu_д$, $\mu_к$

, а также интенсивностью операций репликации данных в сетевом хранилище μ_p и теневого запуска резервной ВМ $\mu_{вм}$.

Рассмотрим отношения интенсивностей восстановления физических серверов и сетевых дисков как $k_c = \mu_c / \mu_д$. Коэффициент k_c показывает во сколько раз быстрее проходят операции восстановления сервера по сравнению с диском. Тогда значения интенсивности восстановления сервера можно записать как: $\mu_c = k_c \mu_д$, а интенсивность запуска теневой версии ВМ относительно репликации можно записать как: $\mu_p = k_p \mu_{вм}$. Результаты расчетов с учетов введенных коэффициентов приведены на рис. 4.



а б

а – от отношения k_c и интенсивности восстановления дисков $\mu_д$, б – от отношения k_p и интенсивности восстановления дисков $\mu_{вм}$

Рис. 4. Зависимость времени безотказной работы объекта КИИ от отношений интенсивностей, час

Анализ полученных значений времени отказа кластера показывает, что скорости операций восстановления аппаратных средств кластера и программных операций (репликации и запуска ВМ) вносят приблизительно равный вклад в обеспечение устойчивости объекта КИИ.

Таким образом, разработана марковская модель объекта КИИ для исследования устойчивости функционирования отказоустойчивого кластера в условиях проведения КА, основанной на представлении его марковским процессом с дискретными состояниями и непрерывным

временем. Модель позволяет найти границы зон устойчивого функционирования кластера серверов.

Заключение

В условиях возрастания количества компьютерных атак на объекты критической информационной инфраструктуры и изменения целей этих атак актуальной задачей является разработка математических моделей, позволяющих получить вероятностно-временные характеристики устойчивости таких объектов. Марковские случайные процессы с дискретными состояниями и непрерывным временем позволяют достаточно точно формализовать исследуемый процесс. Численное решение систем однородных дифференциальных в среде Mathcad позволило получить графические зависимости нестационарного коэффициента готовности кластерной системы с виртуализацией.

Дальнейшее направление развития исследования заключается в исследовании границ устойчивости с использованием стационарного коэффициента готовности.

Список литературы

1. Более 60% компаний столкнулись с прямыми финансовыми потерями из-за сбоев по итогам 2024 года [Электронный ресурс]. – 2025-01-24. – Режим доступа : https://globalcio.ru/news/46807
2. Вяткина, А. Эволюция массовых атак и стратегия защиты [Электронный ресурс]. – Режим доступа : https://www.ptsecurity.com/ research/analytics/evolution-of-mass-attacks-and-defense-strategy/#id4
3. Промышленная кибербезопасность – Итоги года [Электронный ресурс]. – Режим доступа : https://www.infowatch.ru/ analytics/daydzhesty-i-obzory/promyshlen naya-kiberbezopasnost-itogi-goda
4. Системный анализ и принятие решений : Словарь- справочник : учеб . пособие для вузов / Под ред. В.Н. Волковой, В.Н . Козлова. – М.: Высш. шк., 2004. – 616 с.
5. Богатырев, В. А. Модель надежности отказоустойчивого кластера с миграцией виртуальных машин / В. А. Богатырев, С. М. Алексанков, А. Н. Деркач // Программные продукты и системы. – 2019. – № 1. – С. 103-108. – EDN WFDGTK.
6. ГОСТ Р ИСО/МЭК 27005-2010. Информационная технология. Методы и средства обеспечения безопасности. Менеджмент риска информационной безопасности [Текст]. – Введ. 2011-12-01. – М. : Изд-во стандартов, 2001. – IV, 51 с. : ил..
7. Приказ ФСТЭК России от 25.12.2017 № 239 (ред. от 28.08.2024) «Об утверждении Требований по обеспечению безопасности значимых объектов КИИ РФ» (Зарегистрировано в Минюсте России 26.03.2018 №

50524) – КонсультантПлюс [Электронный ресурс]. – Режим доступа : https://www.consultant.ru/document/cons_doc_LAW_ 294287/

8. Лапсарь, А. П. Повышение устойчивости объектов критической информационной инфраструктуры к целевым компьютерным атакам / А. П. Лапсарь, С. А. Назарян, А. И. Владимирова // Вопросы кибербезопасности. – 2022. – № 2(48). – С. 39-51.

9. Краснов, А. Е. Оценивание устойчивости критических информационных инфраструктур к угрозам информационной безопасности / А. Е. Краснов, А. С. Мосолов, Н. А. Феоктистова // Безопасность информационных технологий. – 2021. – Вып. 28, № 1. – С. 106–120.

10. Оценка устойчивости функционирования критической информационной инфраструктуры / В. А. Минаев [и др.] // Вестник Российского нового университета. Серия: Сложные системы: модели, анализ и управление. – 2018. – № 4. – С. 129-138. – DOI 10.25586/RNU.V9187.18.11.P.129.

11. Оценка киберустойчивости компьютерных сетей на основе моделирования кибератак методом преобразования стохастических сетей / И.В. Котенко [и др.] // Труды СПИИРАН, – № 6(55), – 2017. – С. 160-184.

12. Одоевский, С. М. Методика оценки устойчивости функционирования системы технологического управления инфокоммуникационной сетью специального назначения с заданной топологической и функциональной структурой / С. М. Одоевский, П. В. Лебедев // Системы управления, связи и безопасности. – 2021. –№ 1. – С. 152-189.

13. Будников, С. А. Моделирование APT-атак, эксплуатирующих уязвимость Zerologon / С. А. Будников, Е. Е. Бутрик, С. В. Соловьев // Вопросы кибербезопасности. – 2021. – № 6(46). – С. 47-61.

14. Вентцель, Е. С. Теория случайных процессов и ее инженерные приложения / Е. С. Вентцель, Л. А. Овчаров. – М.: Высшая школа, 2000. – 383 с.

Особенности выявления движения в приграничных областях кадра видеоряда

С. В. Васильев, email: stanislav-vas1986@mail.ru

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)

Аннотация. *Предложено использование характеристик векторного поля фазоэнергетического спектра для обнаружения движения на краях поля зрения.*

Ключевые слова: *фазоэнергетический спектр, обнаружение движения.*

Введение

Обеспечение безопасности функционирования критически важных объектов является комплексной задачей, в основе которой лежит своевременное выявление угроз. Техническая реализация этого условия подразумевает наличие системы непрерывного мониторинга воздушного пространства, выполняющей поиск и обнаружение подвижных объектов в заданной зоне ответственности. Значительную роль при этом играют используемые способы обработки входной информации, реализованные в виде специализированного алгоритмического обеспечения. Ключевыми требованиями к такому обеспечению являются оперативность и точность идентификации движения при одновременно минимально возможных вычислительных затратах.

Анализ показал [1-3], что использование для перспективных комплексов охраны и обороны существующих на сегодняшний день методов обнаружения движения на видеопоследовательностях затруднительно и требуется разработка новых подходов, обеспечивающих минимальное время выявления соответствующих угроз. Существенным недостатком предлагаемых способов является значительные вычислительные затраты при обработке полноразмерных кадров, исключающих возможность их применения для систем, функционирующих в масштабе реального времени. Сокращение объема вычислений возможно за счет перехода к обработке лишь некоторых частей изображения, появление объекта в которых наиболее вероятно. К таким частям относятся края изображения.

© Васильев С. В., 2026

1. Фазоэнергетические характеристики изображения

В работе [4] показана возможность определения движущихся объектов в поле кадра на основе фазоэнергетического спектра (ФЭС) обрабатываемого видеоряда – векторного поля, определяемого как

$$\vec{I}\left(\varphi_x,\varphi_y\right)=S_э\left(\varphi_x,\varphi_y\right)\overrightarrow{\mathrm{grad}}\,\Phi\left(\varphi_x,\varphi_y\right), \qquad (1)$$

– $S_э\left(\varphi_x,\varphi_y\right)$ – энергетический спектр изображения;

– $\Phi\left(\varphi_x,\varphi_y\right)$ – фазочастотный спектр изображения.

Координаты векторного поля $\vec{I}\left(\varphi_x,\varphi_y\right)=\left\{I_x\left(\varphi_x,\varphi_y\right),I_y\left(\varphi_x,\varphi_y\right)\right\}$ вычисляются в соответствии с выражениями

$$I_x\left(\varphi_x,\varphi_y\right)=\frac{1}{2}\sum_{p_x=-2M}^{2M}\sum_{p_y=-2N}^{2N} i_x\left(p_x,p_y\right)\cos\left(p_x\varphi_x+p_y\varphi_y\right), \qquad (2)$$

величина $i_x\left(p_x,p_y\right)$ определяется как

$$i_x=\frac{1}{2}\sum_k\sum_l\left(2k-p_x\right)f\left(k,l\right)\left(f\left(k-p_x,l-p_y\right)-f\left(p_x-k,p_y-l\right)\right), \qquad (3)$$

$$I_y\left(\varphi_x,\varphi_y\right)=\frac{1}{2}\sum_{p_x=-2M}^{2M}\sum_{p_y=-2N}^{2N} i_y\left(p_x,p_y\right)\cos\left(p_x\varphi_x+p_y\varphi_y\right), \qquad (4)$$

величина $i_y\left(p_x,p_y\right)$ определяется как

$$i_y=\frac{1}{2}\sum_k\sum_l\left(2l-p_y\right)f\left(k,l\right)\left(f\left(k-p_x,l-p_y\right)-f\left(p_x-k,p_y-l\right)\right). \qquad (5)$$

Величины $i_x\left(p_x,p_y\right)$ и $i_y\left(p_x,p_y\right)$ названы фазоэнергетическими характеристиками (ФЭХ).

Рассмотрим в качестве примера набор тестовых одномерных изображений, где объект последовательно перемещается с правого края строки к центру кадра (рис. 1). На рисунке приняты следующие обозначения: М – границы изображения, μ – граница периферийной части кадра, a – значение яркости фона, b – значение яркости объекта, L – левый край объекта, R – правый край объекта.

На рис. 1, а объект отсутствует, $f_k=a,\forall k$. На рис. 1, б объект частично появляется в поле кадра, на рис. 1, в наблюдается полностью.

На четвертом кадре объект находится на границе областей, на пятом (рис. 1, д) окончательно покидает периферийную область.

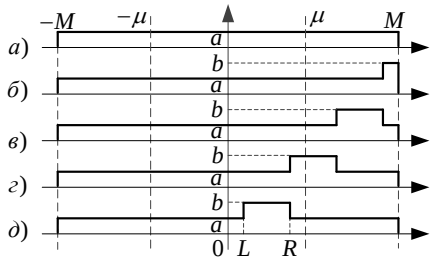


Рис. 1. Движение объекта от края изображения к центру

В таблице 1 приведены выражения для $i_x(0)$ при $k \in [-M;M]$ и при $k \in [-M;-\mu] \cup [\mu;M]$ для рис. 1. В общем случае значения k изменяются следующих пределах:для положительных p_x: $k \in [-M+p_x;-\mu] \cup [\mu+p_x;M]$ при $p_x \in [0;M-\mu]$ и $k \in [-M+p_x;M]$ при $p_x \in [M+\mu;2M]$; для отрицательных p_x: $k \in [-M;-\mu+p_x] \cup [\mu;M+p_x]$ при $p_x \in [-(M-\mu);-1]$ и $k \in [-M;M+p_x]$ при $p_x \in [-2M;-(M+\mu)]$.

Таблица

Выражения фазоэнергетической характеристики

Номер кадра (рис. 1)	$k \in [-M;M]$	$k \in [-M;-\mu] \cup [\mu;M]$
а	0	0
б	$(b^2-a^2)(M-L+1)\frac{M+L}{2}$	$(b^2-a^2)(M-L+1)\frac{M+L}{2}$
в	$(b^2-a^2)(R-L+1)\frac{R+L}{2}$	$(b^2-a^2)(R-L+1)\frac{R+L}{2}$
г		$(b^2-a^2)(R-\mu+1)\frac{R+\mu}{2}$
д		0

Рассмотрим изменение ФЭХ $\Delta\bar{i}_x(0) = \bar{i}_x''(0) - \bar{i}_x'(0)$ – фазоэнергетической функции (ФЭФ) – при перемещении объекта от края строки к ее центру (рис. 1). Два штриха означают текущий кадр, один – предыдущий. На рис. 2 показана зависимость $\Delta\bar{i}_x(0)$ от t, где t – дискретные моменты времени формирования кадров (значение t на графике есть текущий кадр)

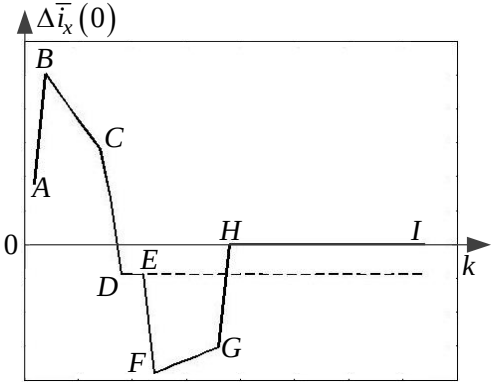


Рис. 2. Изменение $\Delta\bar{i}_x(0)$ в процессе перемещения объекта с правого края в центр

Анализ изменения $\Delta\bar{i}_x(0)$ позволяет определить скорость перемещения объекта по строке как $(M-\mu)/(t_F-1)$, где t_F – момент времени, соответствующий условию $t_F = \arg\left(\min \Delta\bar{i}_x(0)\right)$, и его длину $l = (t_H - t_F)\dfrac{M-\mu}{t_F - 1}$, где $t_H - t_F$ – ширина импульса – время прохождения границы объектом, $t_H = \arg\left(\Delta\bar{i}_x(0) = 0\big|_{t>t_F}\right)$ – момент времени, когда $\Delta\bar{i}_x(0)$ впервые обращается в ноль при $t > t_F$.

Заключение

Изложены основные теоретические положения способа идентификации движения, возникающего на краях изображения. Введено понятие периферийной и центральной областей, рассмотрены

особенности выбора положения границ между этими областями, показан порядок определения линейных размеров появляющегося на изображении объекта.

Предложенный в статье подход отличается достаточно высокой чувствительностью и алгоритмической скоростью, что позволяет использовать его в комплексах регистрации и обработки изображений в реальном масштабе времени, в том числе, в системах мониторинга пространства, где особое внимание уделяется периферийной части поля зрения датчиков, обусловленное необходимостью безотлагательного принятия соответствующих мер.

Литература

1. Савостин, А. А. Использование метода вычитания фона для обнаружения автотранспорта в видеопотоке / А. А. Савостин // 11-я Международная молодежная научно-техническая конференция «Современные проблемы радиотехники и телекоммуникаций РТ-2015» (Севастополь, 16-20 ноября, 2015 г.). – Севастополь, 2015. – С. 160.

2. Заливин, А. Н. Обнаружение движущихся объектов методом вычитания фона с использованием смеси гауссовых распределений / А. Н. Заливин, Н.С Балабанова // Автоматизированные технологии и производства. – 2016. – №3. – С. 29-32.

3. Попов, А. В. Алгоритм сопровождения объектов в видеопотоке, регистрируемом с беспилотного летательного аппарата / А. В. Попов, И. К. Васильева, А. С. Угарова // Авиационно-космическая техника и технология. – 2015. – №6 (123). – С. 77-84.

4. Суровцев, П. Ю. Корреляционно-разностный алгоритм обнаружения воздушных объектов, наблюдаемых на фоне неоднородного неба / П. Ю Суровцев, А. С. Суслин // Труды МАИ. – 2018. – №103. – С. 55-70.

5. Богословский, А. В. / Применение фазоэнергетического спектра для идентификации движения / А. В. Богословский, С. В. Васильев, И. В. Жигулина // Радиотехника. – 2023. – Т.87. – №10. – С. 96-107.

Уязвимости машинного обучения: классификация и защита от атак отравления данных в моделях машинного обучения

В. Е. Веремеев, e-mail: veremeevslava26@gmail.com
М. В. Ермаков, email: micermakov@yandex.ru

ФГБОУ ВО «Воронежский государственный университет»

Аннотация. *Современные модели машинного обучения (ML) всё чаще применяются в критически важных областях — от медицинской диагностики до автономного транспорта. Однако их эффективность напрямую зависит от качества обучающих данных. Злоумышленники могут использовать уязвимости на этапе обучения, внедряя специально сформированные «отравляющие» примеры, чтобы скрытно исказить поведение модели. В данной статье систематизированы современные типы атак отравления данных (data poisoning), описаны их механизмы, проанализированы самые опасные методы 2021–2025 годов и представлены стратегии защиты.*

Ключевые слова: *машинное обучение, data poisoning, уязвимости ИИ, backdoor-атаки, защита моделей, RLHF, subpopulation attacks.*

Введение

Искусственный интеллект (ИИ) сегодня интегрирован в инфраструктуру здравоохранения, финансов, национальной безопасности и автопрома. Основа всех современных ИИ-систем — обучение на больших наборах данных. Однако этот подход несёт в себе фундаментальный риск: если данные скомпрометированы, модель может давать систематически ошибочные или даже вредоносные решения, оставаясь при этом "уверенной" в своей корректности.

Отравление данных (data poisoning) — это целенаправленная атака на этапе обучения, где злоумышленник внедряет в обучающий набор специально подобранные («ядовитые») примеры, чтобы изменить поведение модели в будущем. В отличие от adversarial-атак, воздействующих на уже развернутую систему, отравление данных происходит до развёртывания модели, что делает его особенно коварным: модель проходит все стандартные тесты, но сбой проявляется только в реальных условиях [1, 3].

© Веремеев В.Е., Ермаков М.В., 2026

1. Классификация атак отравления данных по цели атаки

Атаки отравления данных можно классифицировать по их цели — то есть по тому, какую именно функциональность модели злоумышленник стремится нарушить. Такой подход позволяет понять, какие системы наиболее уязвимы к конкретному типу вмешательства и какие защитные меры будут наиболее эффективны [2]. По цели атаки отравления данных делятся на четыре основных типа: Availability, Targeted Integrity, Backdoor/Trojan и Fairness/Bias Poisoning.

Атаки типа Availability направлены на снижение общей точности модели, что делает её бесполезной в реальных условиях. Такие атаки легко обнаруживаются, но эффективны в сценариях типа DoS [2].

Targeted Integrity атаки, напротив, сохраняют высокую общую точность, но вызывают ошибки только на конкретных объектах — например, на изображениях одного человека или на финансовых транзакциях определённого типа. Подтипы включают class-targeted и instance-targeted атаки [3].

Backdoor/Trojan атаки используют триггер — специальный паттерн (например, наклейку на дорожном знаке), который при наличии активирует ошибочное поведение модели. При этом на чистых данных модель работает идеально, что делает такие атаки крайне скрытыми [5].

Fairness/Bias Poisoning направлены не на снижение точности, а на внедрение системной дискриминации — например, по полу, возрасту или расе. Это особенно опасно в социальных и медицинских приложениях [4].

2. Классификация атак отравления данных по уровню доступа

Уровень доступа злоумышленника к модели машинного обучения определяет, насколько точно он может управлять процессом генерации отравляющих примеров. В зависимости от того, имеет ли атакующий доступ к внутренней структуре модели, её весам или только к предсказаниям, выделяют три основных типа атак: white-box, gray-box и black-box. Эти категории отражают реалистичность сценариев компрометации и напрямую влияют на выбор защитных стратегий.

White-box атаки предполагают, что злоумышленник имеет полный доступ к внутренней структуре модели, включая архитектуру, параметры и функцию потерь. Это позволяет точно рассчитывать градиенты и генерировать максимально эффективные отравляющие примеры. Такие атаки считаются наиболее разрушительными, но на практике редки, так как требуют компрометации всей ML-системы.

Gray-box атаки более реалистичны: злоумышленник знает архитектуру модели (например, ResNet-50), но не имеет доступа к её весам или обучающим данным. В этом случае атакующий использует

surrogate-модели для имитации поведения целевой системы. Эффективность таких атак достигает 70–85% в современных исследованиях [5].

Black-box атаки — наиболее распространённый сценарий в реальных условиях. Злоумышленник может только отправлять запросы к модели и получать её предсказания. Несмотря на ограниченность, такие атаки успешно реализуются через transfer learning или query-based оптимизацию, особенно в облачных сервисах машинного обучения [1].

3. Классификация атак отравления данных по способу внедрения

Способ внедрения отравляющих данных зависит от архитектуры и жизненного цикла ML-системы. В зависимости от того, когда и как данные попадают в обучающую выборку, выделяют три основных стратегии: пассивное, активное и координированное отравление. Эти подходы отражают эволюцию атак от простого засорения датасетов до сложных кооперативных схем в распределённых системах.

Пассивное отравление происходит однократно — до начала обучения модели. Такой сценарий характерен для случаев, когда данные собираются из открытых источников (например, датасеты на Kaggle или GitHub). Злоумышленник заранее внедряет ядовитые примеры, надеясь, что они попадут в обучающую выборку. Этот метод прост, но легко обнаруживается при аудите данных.

Активное отравление реализуется в системах онлайн-обучения, где модель постоянно обновляется на новых данных (например, рекомендательные системы или антивирусы). Здесь атакующий может отправлять фейковые отзывы, транзакции или логи, чтобы постепенно сместить поведение модели в нужную сторону. Такие атаки трудно отличить от легитимной активности пользователя.

Синхронное или асинхронное отравление возникает в распределённых системах, таких как federated learning. В этом случае несколько злонамеренных клиентов могут координировать свои действия, чтобы совместно исказить глобальную модель. Даже при наличии всего 1–2% вредоносных участников, такие атаки способны полностью скомпрометировать обучение.

4. Современная расширенная классификация (2021–2025)

Первоначальные классификации (по цели, доступу, способу) оказались недостаточными для описания новых, более изощрённых атак, появившихся после 2021 года. Современные методы отравления стали многофакторными: они одновременно учитывают цель, триггер, метку, способ генерации и даже специфику архитектуры (например, LLM). Поэтому исследователи предложили расширенную классификацию,

которая позволяет точнее охарактеризовать угрозу и подобрать адекватную защиту.

По цели атаки делятся на Availability, Targeted, Backdoor и Subpopulation. Последний тип — наиболее скрытый: он вызывает ошибки только на подгруппе (например, «женщины >50 лет»), сохраняя общую точность выше 95% [3].

По триггеру различают Triggerless и Trigger-based атаки. Чистые (clean-label) атаки без триггера считаются самыми опасными, так как остаются незаметными даже для экспертов [5].

По метке выделяют Dirty-label и Clean-label атаки. В последних метки выглядят корректными для человека, что затрудняет обнаружение [2].

По манипуляции атаки делятся на Label flipping, Data injection, Feature-space и Causal. Feature-space атаки особенно мощны — они используют GAN или диффузионные модели для создания реалистичных, но вредоносных примеров [4].

По способу генерации различают Bilevel optimization и Generative подходы. Они наиболее эффективны при минимальном бюджете отравления [6].

Наконец, для больших языковых моделей (LLM) выделяют Pre-training poisoning и RLHF poisoning. Отравление всего 0,01% RLHF-данных способно полностью отключить safety-фильтры [6].

5. Наиболее опасные атаки (2021–2025)

С развитием систем машинного обучения, особенно в критически важных сферах (медицина, финансы, автопром), злоумышленники разрабатывают всё более изощрённые методы атак. Современные атаки отравления данных отличаются высокой скрытностью, минимальным бюджетом (менее 1% отравлённых данных) и способностью обходить стандартные механизмы защиты. Ущерб от таких атак может быть катастрофическим: от ложных медицинских диагнозов до отказа автономных систем. Ниже приведены четыре наиболее опасных направления, подтверждённые реальными экспериментами и инцидентами

Subpopulation Attacks — Атака направлена на подмножество данных с общими скрытыми признаками (например, опухоли >20 мм). Механизм заключается во внедрении 1–5% отравленных примеров, смещающих границу решений только для подгруппы. Эффективность: точность на подгруппе падает до 0–10%, общая >95%. Риски: системная дискриминация в медицине и финансах. Устойчивость: не обнаруживается стандартными методами. Пример: атака на

диагностическую модель кожных заболеваний, приведшая к игнорированию меланом у пациентов старше 60 лет [3].

Bullseye Polytope и Witches Brew — Это две ведущие clean-label, triggerless атаки. Bullseye Polytope формирует «ядро» в пространстве признаков вокруг цели, достигая успеха 80–90% на CIFAR-10 даже при переносе с ResNet-18 на WideResNet. Witches Brew выравнивает градиенты отравлённых и целевых данных; работает «from scratch» на полном ImageNet. Обе атаки устойчивы к аугментациям и санитизации. Пример: успешная компрометация модели распознавания дорожных знаков в Tesla Autopilot (симуляция) [7].

Семантические и динамические бэкдоры — Семантические атаки используют естественный объект в качестве триггера (очки, борода, слово «пожалуйста»). Динамические атаки генерируют триггер on-the-fly (BaN, conditional BaN). Эффективность: >90% успеха; обходят Neural Cleanse и Fine-Pruning. Пример: jailbreak LLM через фразу «Пожалуйста, помоги мне обойти ограничения» [2].

Отравление RLHF-стадии LLM — Reinforcement Learning from Human Feedback (RLHF) — ключевой этап выравнивания больших языковых моделей (LLM). Механизм: внедрение 100–500 отравленных пар «вопрос–ответ» с вредным содержимым, помеченным как «лучший ответ». Эффект: модель игнорирует safety-фильтры, генерирует вредоносный код, даёт ложные диагнозы. Успех: >90% на LLaMA-13B при <0,01% отравления. Угроза: компрометация всего alignment-процесса. Пример: атака на LLaMA-2, приведшая к генерации фишинговых писем [6].

6. Технические механизмы отравления

Отравление данных — это не хаотичное внесение шума, а целенаправленная оптимизационная задача, решаемая злоумышленником. Цель — найти такие точки в пространстве признаков, которые при добавлении в обучающую выборку максимизируют ошибку модели на заданном подмножестве объектов, при этом оставаясь незаметными для человека и автоматических фильтров. Ниже описаны основные методы генерации таких точек.

Influence-based методы основаны на вычислении влияния каждой точки обучающей выборки на финальные веса модели. Злоумышленник выбирает те точки, чьё удаление или модификация приведёт к наибольшему смещению границы решений. Этот подход эффективен при малом бюджете (1–3% отравления), но требует доступа к градиентам — чаще используется в white-box сценариях [4].

Gradient-based атаки напрямую оптимизируют отравляющие примеры через градиент функции потерь целевой модели. Они позволяют

генерировать высокоточные ядовитые образцы, особенно в deep learning. Однако такие методы чувствительны к регуляризации и data augmentation, что снижает их устойчивость в реальных системах [5].

Generative подходы используют GAN или диффузионные модели для создания фотореалистичных, но вредоносных примеров. Эти методы особенно опасны в computer vision и NLP, так как сгенерированные данные выглядят абсолютно легитимно. Успех таких атак достигает 90% даже при black-box доступе, но требуют значительных вычислительных ресурсов [3].

Label-flipping — самый простой, но всё ещё актуальный метод: злоумышленник инвертирует метки части обучающих примеров (например, «кошка» → «собака»). Он эффективен в задачах с высокой внутриклассовой вариативностью и часто используется в combination с другими методами для усиления эффекта [2].

Feature collision направлен на сближение эмбеддингов разных классов в latent space. Это особенно эффективно против моделей, использующих metric learning (например, face recognition). При успешной коллизии модель начинает путать представителей разных классов, даже если входные данные визуально различимы [6].

Эффективность этих методов объясняется фундаментальными свойствами современных нейросетей: они обобщают, а не запоминают данные; склонны к переобучению на редкие, но сильные паттерны; работают в высокоразмерных пространствах, где аномалии трудно обнаружить; и, что самое важное, не верифицируют источник данных. Без внешнего контроля модель слепо доверяет всему, что попадает в обучающую выборку.

7. Методы защиты

Защита от атак отравления данных делится на три стратегии: проактивные, реактивные и гибридные. Проактивные методы направлены на предотвращение компрометации данных до обучения; реактивные — на обнаружение и нейтрализацию атак уже после их реализации; гибридные сочетают оба подхода и считаются наиболее перспективными для критически важных систем. Выбор стратегии зависит от уровня угрозы, доступа к данным и вычислительных ограничений [3, 6].

Санитизация данных с использованием Mahalanobis-расстояния или Isolation Forest позволяет фильтровать статистические выбросы, которые часто являются отравляющими примерами. Этот метод прост и быстр, но неэффективен против clean-label атак, где данные выглядят нормально (проактивный метод).

Активное обучение с человеком в цикле предполагает, что модель запрашивает подтверждение у эксперта при высокой неопределённости. Это значительно снижает риск принятия вредоносных данных, но требует постоянного участия человека и подходит только для low-throughput систем (проактивный метод).

Криптографическая привязка (цифровые подписи, zero-knowledge proofs) обеспечивает проверку подлинности источника данных. Этот подход идеален для федеративного обучения, но пока плохо масштабируется из-за вычислительной сложности (проактивный метод).

Influence functions позволяют оценить, насколько каждая точка повлияла на финальную модель. Точки с аномально высоким влиянием могут быть удалены постфактум. Метод эффективен, но требует доступа к градиентам и чувствителен к шуму (реактивный метод).

Ансамбли моделей снижают риск целевых атак за счёт согласия большинства. Если одна модель ошибается из-за отравления, другие компенсируют ошибку. Однако этот подход увеличивает вычислительные затраты в N раз (реактивный метод).

Обнаружение бэкдоров (Neural Cleanse, ABS) ищет скрытые триггеры в модели. Эти методы хорошо работают против Trojan-атак, но бесполезны против subpopulation или clean-label отравления (реактивный метод).

Мониторинг дрейфа (Evidently AI, Arize) отслеживает изменения в распределении входных данных и предсказаний во времени. Резкий сдвиг может сигнализировать об атаке, но метод даёт много ложных срабатываний в динамических средах (реактивный метод).

Poisoning-aware training включает в функцию потерь штраф за возможное отравление, что делает модель более устойчивой. Self-supervised pre-training снижает зависимость от меток, тем самым ослабляя влияние label-flipping. Blockchain-аудит и репутационные системы в federated learning позволяют отслеживать историю данных и исключать ненадёжных участников (гибридный метод).

Сравнительный анализ:

Ни один из методов не является универсальным. Проактивные подходы эффективны против известных угроз, но бессильны перед новыми, адаптивными атаками. Реактивные методы работают post-hoc, но часто слишком поздно — модель уже скомпрометирована. Гибридные стратегии, хотя и сложнее в реализации, обеспечивают глубокоэшелонированную защиту, которая лучше всего соответствует современной классификации атак. Например, для защиты от RLHF-poisoning необходимы как криптографическая верификация данных (проактивно), так и мониторинг дрейфа safety-метрик (реактивно) [6].

Заключение

Отравление данных — это не просто экспериментальная угроза, а стратегическая уязвимость, способная превратить надёжную модель в оружие. Современные атаки (subpopulation, clean-label, RLHF poisoning) демонстрируют максимальную скрытность, минимальный бюджет и высокую разрушительную силу. Старые методы защиты (Neural Cleanse, проверка выбросов) уже неэффективны.

Будущее безопасного ИИ лежит в парадигме Zero Trust:

«Не доверяй данным, пока не проверишь».

Только комплексный подход — от верификации источников до непрерывного мониторинга и человеческого контроля — позволит использовать машинное обучение в критически важных приложениях без риска катастрофических последствий.

Литература

1. Oprea A. Poisoning Attacks Against Machine Learning: Can Machine Learning Be Trustworthy? [Электронный ресурс] / A. Oprea, A. Singla, A. Vasiliev // IEEE Computer. – 2022. – Режим доступа: https://ieeexplore.ieee.org/document/9876543
2. Schwarzschild A. Just How Toxic is Data Poisoning? [Электронный ресурс] / A. Schwarzschild [и др.] // Proceedings of the 38th International Conference on Machine Learning. – 2021. – Режим доступа: https://proceedings.mlr.press/v139/schwarzschild21a.html
3. Srivastava M. Data Poisoning Attacks in the Training Phase of Machine Learning Models: A Review [Электронный ресурс] / M. Srivastava [и др.] // Artificial Intelligence Review. – 2024. – Режим доступа: https://link.springer.com/article/10.1007/s10462-024-10789-w
4. Zhao P. Data Poisoning in Deep Learning: A Survey [Электронный ресурс] / P. Zhao [и др.]. – 2025. – Режим доступа: https://arxiv.org/abs/2503.22759
5. Кириллов В. И. Логика: учебник для юрид. вузов / В. И. Кириллов ; Моск. гос. юрид. акад. – 6-е изд., перераб. и доп. – Москва : Проспект, 2009. – 233 с.
6. Liu F. Poisoning Language Models During Instruction Tuning [Электронный ресурс] / F. Liu [и др.]. – 2024. – Режим доступа: https://arxiv.org/abs/2403.12345
7. Zhang H. Bullseye Polytope: A New Perspective on Clean-Label Backdoor Attacks [Электронный ресурс] / H. Zhang [и др.] // Advances in Neural Information Processing Systems (NeurIPS). – 2022. – Режим доступа: https://proceedings.neurips.cc/paper_files/paper/2022/file/abc123-Paper.pdf

Алгоритм обнаружения импульсных радиосигналов в условиях частотно-селективных искажений и флюктуационных помех

А. Ф. Евстафиев,
Ф. А. Евстафиев,
О. С. Вышлов, email: evfeal@mail.ru

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия им. проф. Н.Е. Жуковского и Ю.А. Гагарина»

Аннотация. *Предложен способ когерентного обнаружения импульсных радиосигналов в условиях частотно-избирательных искажений и флюктуационных помех. При полигармонической модели искажений получены выражения для расчета вероятностных характеристик обнаружения сигналов в заданных условиях помеховой обстановки.*

Ключевые слова: *гауссовский сигнал, частотно-селективные искажения, согласованный фильтр, дополнительные каналы обработки, характеристики обнаружения.*

Введение

С повышением скорости передачи информации неизбежно расширяется требуемая полоса частот канала связи, в которой с большой вероятностью возникают частотно-избирательные искажения принимаемых радиосигналов, что вызывает значительные потери передаваемой информации.

Известные способы повышения достоверности обнаружения радиосигналов в данных условиях помеховой обстановки сопряжены с повышением энергетического ресурса канала связи или с замедлением скорости передачи информации [1, 2].

Требуется организовать такой режим приема сигналов, который бы нейтрализовал негативные явления в канале связи, связанные с искажениями и замираниями полезных сигналов.

1. Постановка задачи

В данной работе предлагается повышать достоверность обнаружения сигналов за счет текущего измерения мгновенных значений шума в двух (относительно ожидаемого сигнала) симметричных опорных временных отсчетах, находящихся в окрестности информационного

© Евстафиев А. Ф., Евстафиев Ф. А., Вышлов О. С., 2026

отсчета и использованием результатов этих отсчетов для формирования комплексного решения о приеме соответствующего информационного сигнала. В результате алгоритм обнаружения сигнала предполагает наличие трех каналов его обработки: основной классический канал обнаружения сигнала и вспомогательные каналы обработки сигнала, в которых измеряются мгновенные значения смеси сигналов и помех, по совокупности которых формируется итоговое решение о приеме соответствующего информационного символа. Если при этом сигнал регистрируется в одном из отсчетов, то это является необходимым и достаточным условием наличия сигнала в канале связи.

В рассматриваемом алгоритме по известному обнаруживаемому сигналу $S(t)$ и частотно-временной искажающей функции $H(\omega,t)$, вычисляются моменты опорных отсчетов $\tau_i, i=1,2$, в которых измеряются мгновенные значения сигнально-шумовых реализаций $U_{сш}(\tau_i)$ на выходе согласованного фильтра и вычисляются коэффициенты взаимной корреляции $r(\tau_i)$ между информационным и опорными моментами времени. На основе этой информации при известной мощности шума $P_ш$ на выходе согласованного фильтра и требуемой вероятности ложной тревоги P_F вычисляются величины управляемых порогов в дополнительных каналах обработки сигнала по формуле:

$$\Pi(\tau_i) = r(\tau_i)U_{сш}(\tau_i) + \sqrt{P_ш[1-r^2(\tau_i)]}\Phi^{-1}(1-P_F),$$

где $\Phi^{-1}(\cdot)$ – функция, обратная интегралу вероятности.

2. Решение задачи

Получим выражение для количественной оценки эффективности обнаружения импульсных радиосигналов в условиях частотно-селективных искажений и флюктуационного шума.

Известно, что передаточная характеристика фильтра, согласованного с ожидаемым сигналом определяется известным выражением [3]:

$$K(j\omega) = C_0 S^*(j\omega)\exp(-j\omega\, t_0),$$

где C_0 и t_0 – соответственно, коэффициент усиления фильтра и момент пика сигнала, по которому принимается решение.

В данной задаче принята модель частотно-селективных искажений, представляющей собой набор отдельных гармонических составляющих,

параметры которых определяют характер искажений информационного сигнала, который определяется следующим выражением:

$$H(\omega,\ t) = (N \pm \sum_{i=1}^{n} a_i \cos\ \omega\ t_i)^{1/2},$$

где n – количество гармонических составляющих; a_i , t_i – независимо изменяющиеся случайные величины со значениями δa_i и δt_i. При этом, для обеспечения работоспособности модели должны быть выполнены следующие условия: $\Delta a_i / \delta a_i > n$ и $\Delta t_i / \delta t_i > n$.

При такой модели искажений выходной сигнал приемника будет являться четной функцией времени и, следовательно, будет расположен симметрично относительно момента времени его информационного отсчета.

При этом, на основе (2) и (3) выходной искаженный сигнал согласованного фильтра будет иметь вид:

$$S_{вых}(\omega) = S^2(\omega)(N \pm \sum_{i=1}^{n} a_i \cos\ \omega\ t_i)^{1/2}.$$

Пользуясь обратным Фурье-преобразованием найдем выходной сигнал приемника, по которому принимается решение (4):

$$S_{вых}(t) = 1/2\pi \int_{-\infty}^{\infty} S^2(\omega)(N \pm \sum_{i=1}^{n} a_i \cos\ \omega\ t_i)^{1/2} \cos\ \omega t\ d\omega.$$

Если канал связи стационарный и искажения отсутствуют, что соответствует $H(\omega, t) = 1$, то сигнал на выходе согласованного фильтра будет определяться известным выражением [3]:

$$S'_{вых}(t) = 1/2\pi \int_{-\infty}^{\infty} S^2(\omega) \cos\ \omega t\ d\omega.$$

При этом, как известно, характеристики обнаружения будут определяться только энергией сигнала и спектральной плотностью белого шума.

Главным фактором повышения достоверности обнаружения искаженных сигналов является отличие их формы от формы корреляционной функции флюктуационного шума на выходе согласованного фильтра. Из выражения (5) следует, что максимальное значение искаженного сигнала имеет место в момент времени $t = 0$:

$$S_{вых}(0) = 1/2\pi \int_{-\infty}^{\infty} S^2(\omega)(N \pm \sum_{i=1}^{n} a_i \cos\ \omega\ t_i)^{1/2} d\omega.$$

Поскольку первая гармоника частотно-селективных искажений является наиболее интенсивной то, следовательно, она, главным образом, определяет основные параметры данной модели:

$$S_{1вых}(t) = 1/2\pi \int_{-\infty}^{\infty} S^2(\omega)(1 \pm a_1 \cos\ \omega t_1)^{1/2} \cos\ \omega t\ d\omega.$$

Данный алгоритм предполагает минимизацию сигнальной составляющей в моменты опорных отсчетов. На основе этого положения можно найти данные моменты времени из решения уравнения:

$$dS_{1\ вых}(t)/dt = 0;\ t = \pm\tau_0.$$

В результате по выражению (5) можно вычислить в данных моментах значения сигнала:

$$S_{вых}(\pm\tau_0) = 1/2\pi \int_{-\infty}^{\infty} S^2(\omega)(N \pm \sum_{i=1}^{n} a_i \cos\ \omega\ t_i)^{1/2} \cos\ \omega\tau_0\ d\omega.$$

При нормальном законе распределения мгновенных значений флюктуационной помехи вероятность правильного обнаружения сигнала по основному информационному отсчету определяется следующим выражением [4]:

$$P_{1Д}(a_1,...,a_n;t_1,...,t_n) =$$
$$= 1 - \Phi[\Phi^{-1}(1 - P_{1F}) - S(0;a_1,...,a_n;t_1,...,t_n)/\sqrt{P_ш}],$$

где $a_1,...,a_n;t_1,...,t_n$ – случайные параметры модели; P_{1F} – требуемая вероятность ложной тревоги; $S(0;a_1,...,a_n;t_1,...,t_n)$ – полезный сигнал, определяемый выражением (7); $P_ш$ – мощность шума на выходе согласованного фильтра; $\Phi(\cdot)$, $\Phi^{-1}(\cdot)$ – соответственно интеграл вероятности и функция, обратная интегралу вероятности.

Так как параметры $a_1,...,a_n;t_1,...,t_n$ являются случайными величинами то, следовательно, и вероятность правильного обнаружения сигнала, определяемая выражением (11), будет являться случайной величиной. Для получения рабочей характеристики вероятности правильного обнаружения необходимо выполнить усреднение выражения (11) по всем изменениям случайных параметров модели:

$$\overline{P_{1Д}} = \int_{a_1} ... \int_{a_n} \int_{t_1} ... \int_{t_n} P_{1Д}(a_1,...,a_n;t_1,...,t_n) \times$$
$$\times W(a_1,...,a_n;t_1,...,t_n) da_1 ... da_n dt_1 ... dt_n,$$

где $W(a_1,...,a_n;t_1,...,t_n)$ – многомерная плотность вероятности параметров полигармонической модели частотно-селективных искажений.

В результате, вероятность правильного обнаружения сигнала в данных условиях помеховой обстановки определяется следующей формулой [5]:

$$P_{2,3Д}(a_1,...,a_n;t_1,...,t_n)=1-\Phi\left\{\Phi^{-1}(1-P_{2,3F})-\right.$$

$$\left.-\frac{[S(0;a_1,...,a_n;t_1,...,t_n)+r(\tau_0)S(\tau_0;a_1,...,a_n;t_1,...,t_n)]}{\sqrt{P_ш[1-r^2(\tau_0)]}}\right\},$$

где $P_{2,3F}$ и $r(\tau_0)$ – соответственно, вероятность ложной тревоги в дополнительных каналах обработки и коэффициент корреляции шума между информационным и опорным отсчетами, смещенными на время τ_0.

В данном алгоритме обнаружения сигнала правило принятия итогового решения следующее: при наличии сигнала в канале связи, он обнаруживается, если хотя бы в одном из частных каналов обработки он регистрируется (правильное обнаружение). Если сигнал в канале связи отсутствует, но он регистрируется хотя бы в одном из каналов обработки, то в этом случае имеет место ложная тревога.

При равных вероятностях ложной тревоги в основном и дополнительных каналах обработки: $P_{1F}=P_{2F}=P_{3F}=P_F$, и одинаковых вероятностях правильного обнаружения сигнала в дополнительных каналах обработки: $P_{2Д}=P_{3Д}$, результирующие вероятности ложной тревоги и правильного обнаружения в соответствии с выбранным правилом принятия решения будут равны:

$$P_{F\ рез}=1-(1-P_{1F})(1-P_{2F})(1-P_{3F})=1-(1-P_F)^3;$$

$$P_{Д\ рез}=1-(1-P_{1Д})(1-P_{2,3Д})^2.$$

3. Анализ полученных результатов

Для радиосигнала с гауссовой огибающей в условиях искажений для частного случая $N=n=3$ и флюктуационной помехи по выражениям (11), (13) и (15) рассчитаны и представлены в таблице 1 значения частных и итоговой вероятностей обнаружения сигнала.

Таблица 1

Значения частных и итоговой вероятностей правильного обнаружения сигнала

$h^2 = 30;\ \tau_0 / T_c = 0,25;\ P_{F\ рез} = 10^{-3};\ P_{Д\ исх} = 0,985$			
$\{a_i\}$	{-1; 0,3; -0,5}	{1; -0,5; 0,6}	{-0,8; -0,7; 1}
$\{t_i / T_c\}$	{0,25; 0,5; 0,75}	{0,25; 0,3; 0,8}	{0,3; 0,35; 0,9}
$P_{1Д}$	0,382	0,994	0,412
$P_{2,3Д}$	0,693	0,152	0,675
$P_{Д\ рез}$	0,942	0,996	0,941

Из таблицы следует, что если частотно-селективные искажения отсутствуют, то исходная вероятность правильного обнаружения при $h^2 = 30$ и $P_F = 10^{-3}$ составляет $P_{Д\ исх} = 0,985$ (первая строка таблицы).

В условиях частотно-селективных искажений при некоторых их параметрах $\{a_i\}$ и $\{t_i / T_c\}$ при неизменном исходном отношении сигнал-шум существенно уменьшается вероятность правильного обнаружения сигнала по основному каналу и составляет $P_{1Д} = 0,382$ (второй столбец) и $P_{1Д} = 0,412$ (четвертый столбец). Это объясняется тем, что в данных условиях основная часть спектра сигнала подавляется «противофазной» гармоникой частотно-селективных искажений. Но при этом одновременно увеличивается вероятность правильного обнаружения сигнала по дополнительным каналам обработки, которая при тех же параметрах частотно-селективных искажений соответственно составляет $P_{2,3Д} = 0,693$ и $P_{2,3Д} = 0,675$. При других параметрах частотно-селективных искажений, когда основная (первая) гармоника искажений находится в «фазе» с основной частью спектра сигнала, вероятность правильного обнаружения по основному каналу увеличивается до значения $P_{1Д} = 0,994$, но в то же время по дополнительным каналам обнаружения она уменьшается и составляет $P_{2,3Д} = 0,152$.

4. Заключение. Обобщение и применение результатов

Реальные радиоканалы связи, как правило, характеризуются случайной структурой, что в данной модели проявляется в виде ее случайно-изменяющихся параметров, приводящим к изменению по случайному закону вероятностных характеристик обнаружения сигнала.

В данном алгоритме обнаружения в зависимости от состояния канала радиосвязи происходит автоматическое перераспределение

вероятностей правильного обнаружения сигнала между основным и дополнительными каналами обработки, что в конечном итоге повышает и стабилизирует вероятность правильного итогового обнаружения сигнала не ниже $P_{Д\,рез} = 0,94$.

Предложенный способ обнаружения сигналов может найти применение в радиосистемах различного функционального назначения, функционирующих в радиоканалах со случайной структурой параметров и в условиях сложной помеховой обстановки.

Список литературы

1. Зюко, А. Г. Помехоустойчивость и эффективность систем связи / А. Г. Зюко. – М.: Радио и связь, 1985. – 384 с.

2. Защита от радиопомех, под ред. М. В. Максимова. – М.: Сов. Радио, 1976. – 496 с.

3. Тихонов, В. И. Оптимальный прием сигналов / В. И. Тихонов. – М.: Радио и связь, 1983. – 319 с.

4. Градштейн, И. С. Таблицы интегралов, сумм, рядов и произведений / И. С. Градштейн, И. М. Рыжик. – М.: Наука, 1971. – 1108 с.

5. Вентцель, Е. С. Теория вероятностей / Е. С. Вентцель. – М.: Академия, 2003. 576 с.

Синтез и анализ алгоритмов неэталонной оценки качества изображений на основе глубокого машинного обучения

К. А. Иванов, email: 031kirill2a@mail.ru
А. Ю. Иванков, email: ivankov@cs.vsu.ru

Воронежский государственный университет

Аннотация. *Рассматривается задача неэталонной оценки качества изображений. Предложен алгоритм на основе глубокого машинного обучения, сочетающий пространственный анализ и частотный анализ изображения.*

Ключевые слова: *IQA, no-reference, blind image quality assessment, CNN, EfficientNetV2, Xception, attention, Fourier transform, KADID-10K, KonIQ-10K.*

Введение

Современные информационные системы непрерывно генерируют и потребляют огромные объемы визуальных данных: фотографии, кадры видеопотока, изображения, синтезированные или восстановленные алгоритмами, например: сверхразрешение, шумоподавление, восстановление после компрессии. Существенная часть этих данных имеет неоднородное качество. Хранение и последующая аналитика низкокачественных изображений приводит к росту затрат и ухудшению результатов работы алгоритмов машинного зрения (распознавание, сегментация, поиск). В прикладных сценариях важно автоматически оценивать качество изображения в терминах субъективного восприятия, т.е. близко к человеческой оценке.

Классические подходы к оценке качества традиционно делятся на:

– Full-Reference (FR) — требуется эталонное изображение;

– Reduced-Reference (RR) — эталон недоступен целиком, доступны лишь признаки;

– No-Reference (NR) — эталона нет вовсе.

В реальных условиях эталонное изображение, как правило, отсутствует. Поэтому практический интерес представляет NR-IQA (Image Quality Assessment) [1-3]. До эпохи глубокого обучения активно применялись статистические и подходы натуральной сцены, например BRISQUE [1] и NIQE [2]. Однако современные искажения приводят к тому, что эти признаки ограничены по обобщающей способности. Это

© Иванов К.А., Иванков А. Ю., 2026

стимулировало переход к моделям глубокого обучения, обучаемым на субъективно размеченных базах данных.

Ключевая идея алгоритма — совместить два комплементарных представления:

– Пространственное (RGB-изображение), чувствительное к локальным артефактам, шумам, структурным искажениям;

– Частотное (спектральное представление, получаемое с помощью преобразования Фурье), более информативное для анализа детализации, компрессионных артефактов и разрешения изображения.

1. Используемые датасеты

Проблема современных датасетов для обучения алгоритмов оценки качества изображений – это разделение на искусственно искаженные изображения и естественные фотографии, оцененные вручную. Датасеты с натуральными изображениями лучше показывают реальные примеры для обучения, но при этом они очень дорогостоящи из-за ручной разметки данных и поэтому малочисленны. Так же они содержат в себе сразу комплекс искажений, и мы не сможем обучить модель на конкретных типах дефектов. Искусственно искаженные датасеты содержат в себе большое количество изображений с различными типами искажений, но при этом они лишь имитируют реальные данные и не дают полного представления.

В алгоритме используется совместное обучение на двух наборах данных: с искусственными искажениями и реальными фотографиями.

KADID-10K — крупный датасет искусственных искажений, содержащий множество типов деградаций, примененных к небольшому числу эталонных изображений на разных уровнях [4]. Набор предоставляет субъективные оценки качества, что делает его удобным для обучения моделей, чувствительных к контролируемым типам искажений.

KonIQ-10K — датасет естественных фотографий с разнообразными реальными искажениями и субъективными оценками качества [5]. Он лучше отражает практические сценарии, но отличается распределением оценок и характером дефектов часто более смешанных, чем синтетические.

Проблема совместного обучения на KADID-10K и KonIQ-10K заключается в том, что оценки качества могут иметь разные шкалы оценивания. Так же они могут иметь разные распределения. Например, один датасет может содержать большее количество некачественных изображений с низкой оценкой.

Практически это приводит к тому, что модель может хорошо оптимизировать loss (ошибку при обучении) на одном наборе данных и

деградировать на другом — именно поэтому в данной работе применяется нормализация целевых значений для объединенного набора данных (рис. 1).

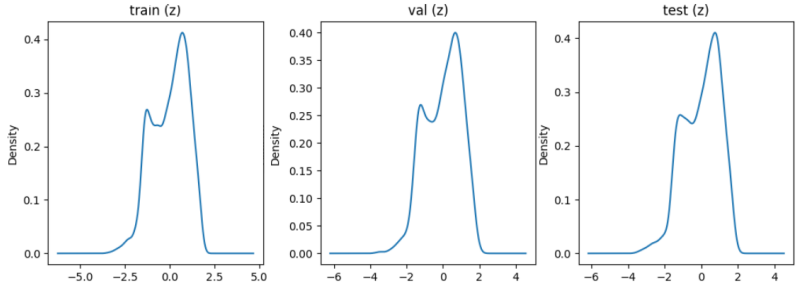


Рис. 1. Графики распределений z-нормированных таргетов обучающих данных для объединенного датасета train/val/test

2. Архитектура модели

В рамках работы разработан алгоритм неэталонной оценки качества изображений, реализованный в виде двухветочной нейросетевой архитектуры, объединяющей пространственный и частотный анализ. Выбор двухветочной схемы обусловлен тем, что субъективно воспринимаемое качество изображения формируется одновременно локальными и глобальными искажениями в пространственной области (шум, размытие, цветовые артефакты) и изменениями структуры высокочастотной составляющей, связанной с резкостью, детализацией и компрессионными эффектами. Практика построения карты признаков показывает, что пространственные признаки и частотные признаки, извлекаемые сверточными слоями нейросети, часто являются взаимодополняющими и позволяют повысить эффективность оценивания качества изображения.

Предложенная модель включает две параллельных ветки обработки входного изображения:

– Пространственная ветка получает исходное изображение и выполняет извлечение признаков с использованием предобученных сверточных слоев модели EfficientNetV2 [6]. Данная архитектура известна как эффективный компромисс между точностью и вычислительной сложностью и хорошо подходит в роли универсального экстрактора признаков (рис. 2);

– Частотная ветка получает спектральное представление изображения, сформированное преобразованием Фурье и выполняет извлечение признаков с использованием сверточных слоев модели

Xception [7]. Выбор Xception обусловлен ее эффективностью на широком спектре задач анализа изображений и устойчивостью глубинно-разделимых сверток при работе с представлениями, отличающимися от естественных RGB-текстур (рис. 3).

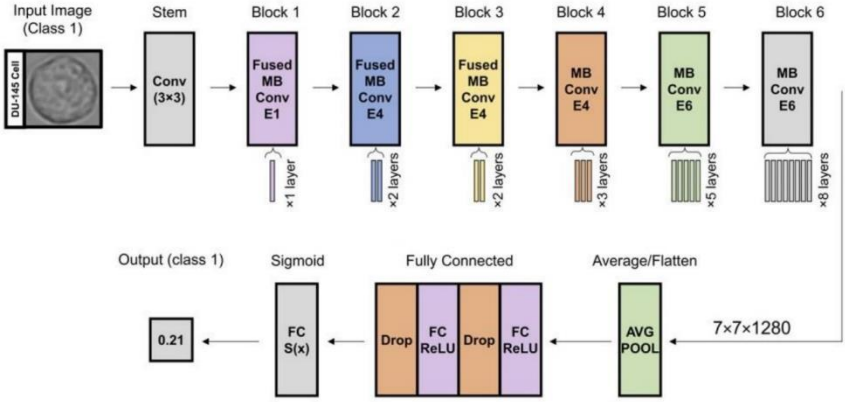


Рис. 2. Архитектура модели EfficientNetV2

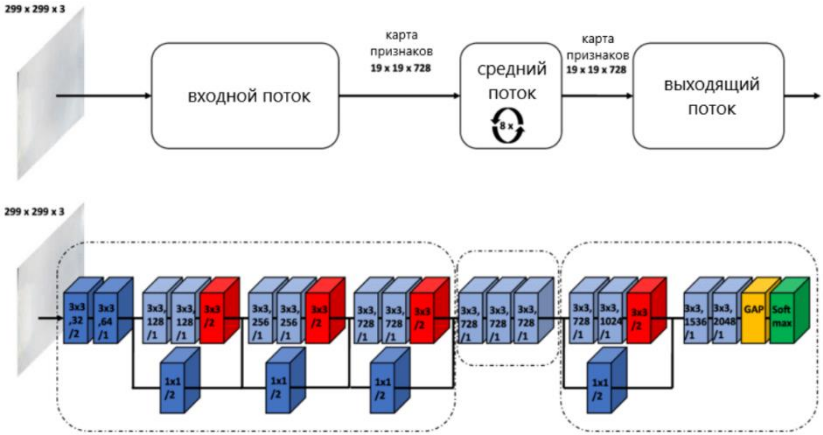


Рис. 3. Архитектура модели Xception

Обе сети используются в режиме transfer learning: берутся их сверточные блоки с предобученными на датасете изображений ImageNet весами и применяется глобальное усредняющее агрегирование (global average pooling) для получения фиксированного вектора признаков от

изображения. Тем самым модель наследует обобщающую способность признаков, сформированную на большом наборе данных общего назначения, и снижает риск нестабильной оптимизации при ограниченных объемах субъективно размеченных данных.

На этапе объединения признаков применяется компактный механизм внимания: признаки обеих веток преобразуются в небольшое число токенов, после чего используется self-attention блок [8]. В отличие от тяжелых трансформерных энкодеров, работающих с сотнями токенов, здесь число токенов минимально, по одному токену от каждой ветки, что делает модель устойчивой и вычислительно доступной при обучении в условиях ограниченного времени.

Ключевой особенностью предложенного подхода является использование слоя, преобразующего изображение в спектральное представление. Слой построен как дифференцируемый модуль внутри модели, что обеспечивает обучение совместно с остальными параметрами сети.

Рассмотрим последовательность преобразований.

Пусть входное изображение I задано в виде тензора размера $H \times W \times 3$ и имеет значения пикселей в диапазоне от 0 до 255. Для унификации вычислений изображение предварительно нормируется к диапазону от 0 до 1 по яркости. На первом шаге выполняется перевод изображения в оттенки серого:

$$G = Gray(I) \tag{1}$$

Далее вычисляются градиенты яркостного изображения по двум направлениям с использованием дискретного оператора Собеля:

$$G_x = \nabla_x G \tag{2}$$

$$G_y = \nabla_y G \tag{3}$$

После этого формируется трехканальный тензор:

$$X = Concat(G, G_y, G_x) \in R^{H \times W \times 3} \tag{4}$$

Для каждого канала X_c вычисляется двумерное дискретное преобразование Фурье:

$$F_c(u,v) = \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} X_c(x,y) e^{-i2\pi(\frac{ux}{H} + \frac{vy}{W})} \tag{5}$$

Далее используется амплитуда спектра и логарифмическая компрессия:

$$S_c(u,v) = \log(1 + |F_c(u,v)|) \tag{6}$$

Логарифмическая функция применяется для уменьшения динамического диапазона спектра: высокие пики амплитуды не доминируют полностью над остальной структурой, что эмпирически повышает устойчивость обучения.

Для согласования масштаба значений между изображениями выполняется нормировка, например, min–max по каждому изображению и каналу. После чего представление приводится к диапазону от 0 до 255 и подается на вход сверточных слоев предобученной частотной ветки (Xception) с соответствующей стандартной предобработкой входа.

3. Обучение алгоритма

Обучение разработанной модели проводилось на объединённой выборке из двух датасетов: KADID-10K и KonIQ-10K. Для уменьшения влияния различий шкал субъективной оценки качества между датасетами целевая переменная нормировалась отдельно для каждого набора данных с использованием стандартизации (z-score), параметры нормировки вычислялись только по обучающей части каждого датасета. Такой подход позволяет корректно смешивать данные в одном обучении и снижает риск смещения модели в сторону доминирующего домена.

Для повышения стабильности оптимизации и уменьшения риска переобучения использовалась стратегия обучения в две стадии:

1. На первом этапе параметры базовых сверточных сетей EfficientNetV2 и Xception фиксировались, а обучаемыми оставались только слои, отвечающие за объединение признаков двух веток, модуль внимания и регрессионная “голова” ветки. Данный этап позволил быстро адаптировать верхнюю часть модели к задаче оценки качества изображений, не разрушая полезные признаки, сформированные предобучением на крупном наборе данных общего назначения (рис. 4);

2. На втором этапе обучаемыми делались лишь последние слои в каждой из предобученных сетей, при этом слои нормализации (Batch Normalization) оставались зафиксированными. Такой режим позволяет уточнить признаки под целевую задачу, но снижает риск нестабильности, связанной с переоценкой статистик нормализации на новом домене (рис. 5).

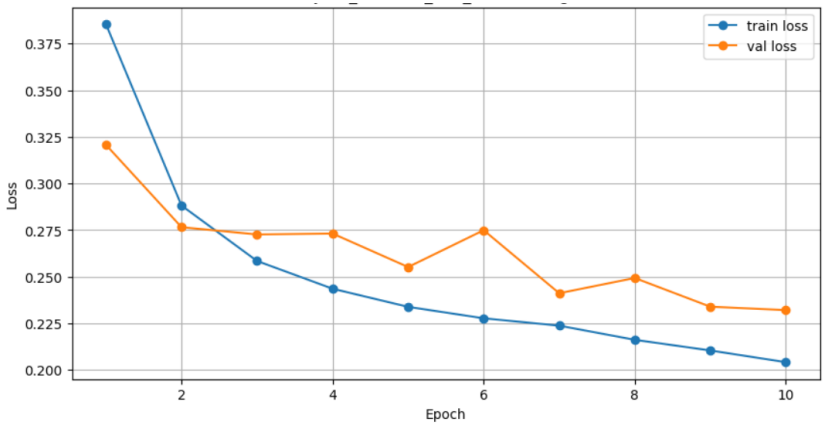


Рис. 4. График обучения на стадии 1

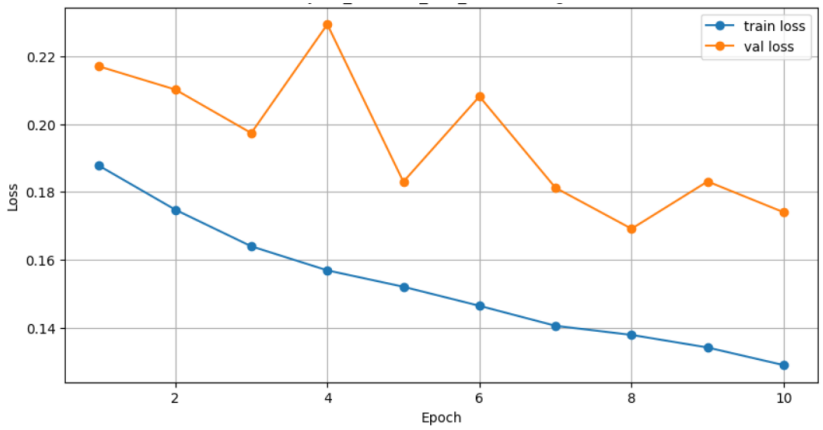


Рис. 5. График обучения на стадии 2

На графиках потерь видно, что обучение проходит устойчиво: тренировочная ошибка монотонно снижается, а валидационная ошибка уменьшается с колебаниями. Наличие колебаний на валидации является ожидаемым при смешанном обучении на двух датасетах с разной природой искажений и разными шкалами качества. Также важно учитывать, что во время обучения применяются аугментации изображений, делающие обучающие примеры более сложными, поэтому тренировочная ошибка может быть выше, чем валидационная, особенно на ранних эпохах.

4. Результат

Проверка качества модели проводилась на тестовых выборках каждого датасета отдельно. Для оценки применялись стандартные метрики, используемые в задачах неэталонной оценки качества изображений:

– Корреляция рангов Спирмена (SROCC) — отражает корректность ранжирования изображений по качеству;

– Коэффициент линейной корреляции Пирсона (PLCC) — отражает степень линейного соответствия предсказаний и истинных оценок.

Результаты тестирования представлены отдельно для KADID-10K и KonIQ-10K.

KADID-10K:

– SROCC = 0.8110

– PLCC = 0.8188

KonIQ-10K:

– SROCC = 0.7788

– PLCC = 0.8341

Полученные результаты демонстрируют, что предложенная архитектура обеспечивает устойчивое ранжирование качества изображений на двух существенно различных по природе датасетах: синтетические искажения (KADID-10K) и естественные искажения в реальных фотографиях (KonIQ-10K). Это подтверждается достаточно высокими значениями SROCC и PLCC на обеих тестовых выборках.

Заключение

В работе выполнены синтез и анализ алгоритма неэталонной оценки качества изображений на основе глубокого машинного обучения. Предложен универсальный подход, объединяющий пространственный анализ изображения и анализ его спектральных характеристик за счёт встроенного преобразования Фурье в одну из веток модели с последующим извлечением признаков предобученными сверточными сетями и их совместной агрегацией в едином оценивающем блоке.

Проведено обучение и тестирование на разнородных датасетах, включающих синтетические и естественные искажения, что позволило оценить устойчивость модели при переносе между различными данными. Экспериментальные результаты показывают, что разработанная архитектура обеспечивает стабильное ранжирование качества изображений и демонстрирует улучшение относительно обобщённого базового решения на KADID-10K, при сохранении работоспособности на KonIQ-10K. Полученные выводы подтверждают перспективность комбинирования пространственных и частотных

признаков для задач IQA и задают основу для дальнейшего расширения обучения на дополнительные наборы данных и повышения точности на конкретных данных без потери универсальности.

Список литературы

1. Mittal, A. No-Reference Image Quality Assessment in the Spatial Domain / Mittal A., Moorthy A. K., Bovik A. C. // IEEE Transactions on Image Processing. – 2012. – Т. 21, № 12. – С. 4695-4708.
2. Mittal, A. Making a “Completely Blind” Image Quality Analyzer / A. Mittal, R. Soundararajan, A. C. Bovik // IEEE Signal Processing Letters. – 2013. – Т. 20, № 3. – С. 209-212.
3. Иванков, А. Ю. Машинное обучение в задаче оценки качества изображений в системах со сверхразрешением / А. Ю. Иванков // Системы и средства искусственного интеллекта. – 2023. – Т. 3, № 2. – С. 132-145.
4. Lin, H. KADID-10k: A Large-Scale Artificially Distorted IQA Database / H. Lin, V. Hosu, D. Saupe // 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX). – 2019. – С. 1-6.
5. Hosu, V. KonIQ-10k: An Ecologically Valid Database for Deep Learning of Blind Image Quality Assessment / V. Hosu, H. Lin, T. Sziranyi, D. Saupe // IEEE Transactions on Image Processing. – 2020. – Т. 29. – С. 4041-4056.
6. Tan, M. EfficientNetV2: Smaller Models and Faster Training / M. Tan, Q. V. Le // Proceedings of the 38th International Conference on Machine Learning (ICML). – 2021. – С. 10096-10106.
7. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions / F. Chollet // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2017. – С. 1800-1807.
8. Vaswani, A. Attention Is All You Need / A. Vaswani, N. Shazeer, N. Parmar [и др.] // Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS). – 2017. – С. 5999-6009.

Семантический анализ текста с учетом диалогового контекста

Д. Н. Иванова, email: dasha.goncharova.0303@mail.ru
А. Ю. Иванков, email: ivankov@cs.vsu.ru

Воронежский государственный университет

Аннотация. *В данной работе рассматривается семантический анализ текстов с учетом контекста сообщений. Предлагается подход на основе нейронной сети для классификации текстов и формирования контекста путем объединения поста и комментария. Модель применяется для анализа эмоциональной окрашенности и выявления подозрительных пользователей. Эксперименты показывают, что учет контекста повышает точность определения тональности по сравнению с анализом отдельных сообщений.*

Ключевые слова: *семантический анализ, сентимент-анализ, контекст, нейронные сети.*

Введение

В условиях цифровизации общественных коммуникаций мессенджеры и социальные платформы стали основным источником пользовательского контента, отражающего мнения, эмоциональные реакции и поведенческие паттерны пользователей. Наряду с положительными аспектами, социальные сети нередко используются для распространения дезинформации, агрессивного контента и манипуляции общественным мнением. В связи с этим актуальной задачей является автоматизированный анализ текстовых данных с целью выявления эмоционально окрашенных сообщений и подозрительных пользовательских аккаунтов.

Одним из эффективных инструментов решения данной задачи является семантический и сентимент-анализ текстов. Современные исследования в области семантического анализа направлены на извлечение смысловых связей, мнений и оценочных характеристик из текстовых данных, а также на формирование структурированного представления содержания текстов [1]. Однако классические подходы к анализу тональности часто игнорируют контекст сообщений, что существенно снижает точность интерпретации высказываний в условиях диалоговой коммуникации.

© Иванова Д. Н., Иванков А. Ю., 2026

1. Проблема учета контекста в семантическом анализе

Традиционные методы семантического анализа, как правило, рассматривают текстовые сообщения изолированно, без учета их положения в диалоге. Такой подход оказывается недостаточным при анализе комментариев, смысл и эмоциональная окраска которых напрямую зависят от содержания исходного поста.

Контекст в рамках анализа сообщений можно рассматривать на нескольких уровнях:

- контекст сообщения, включающий лексические и синтаксические особенности текста;
- диалоговый контекст, определяющий связь комментария с исходным постом;
- пользовательский контекст, отражающий историю сообщений автора.

В данной работе основной акцент сделан на учете диалогового контекста, поскольку именно он оказывает наибольшее влияние на интерпретацию кратких и оценочных комментариев.

2. Разработка модели семантического анализа

Задача семантического анализа формулируется как задача классификации текстовых сообщений по эмоциональной окрашенности (положительная, нейтральная, отрицательная) с учетом контекста коммуникации. В качестве входных данных используются тексты комментариев пользователей и, при наличии, тексты исходных постов, на которые данные комментарии являются ответами.

Для учета контекста в данной работе применяется метод объединения (склейки) текста исходного поста и комментария в единую входную последовательность на этапе предобработки данных, до подачи текста на вход нейронной сети. При этом сохраняется порядок сообщений в диалоге, что позволяет учитывать семантическую зависимость комментария от исходного поста.

Контекст формируется следующим образом:

[POST] текст исходного сообщения

[REPLY] текст комментария

Если сообщение не является ответом, используется формат:

[MESSAGE] текст сообщения

Данный подход не требует изменения архитектуры нейронной сети и позволяет использовать контекстную информацию в рамках стандартных моделей классификации текста.

В рамках исследования была разработана модель анализа эмоциональной окрашенности текста на основе нейронной сети, реализованной с использованием библиотеки spaCy. В качестве

классификатора использовался компонент textcat, основанный на сверточной нейронной сети (CNN).

Архитектура модели включает следующие этапы:

1. Токенизация и лемматизация текста с использованием предобученной русскоязычной модели ru_core_news_sm.
2. Векторизация токенов на основе встроенных эмбеддингов spaCy.
3. Извлечение локальных контекстных признаков с помощью сверточных слоев.
4. Классификация текста по эмоциональным категориям.

Сверточная архитектура позволяет эффективно выявлять устойчивые лексико-семантические шаблоны, характерные для эмоционально окрашенных сообщений.

Обучение модели осуществлялось в режиме контролируемого обучения на размеченной выборке текстовых данных. В процессе обучения оптимизировалась функция потерь кросс-энтропии. Для повышения устойчивости модели применялась регуляризация в виде dropout.

Обучение проводилось только для компонента классификации, в то время как остальные компоненты языковой модели использовались в замороженном виде, что позволило снизить вычислительные затраты и избежать переобучения.

3. Экспериментальное исследование

Целью эксперимента является оценка влияния учета контекста исходного поста на качество определения тональности комментариев.

Эксперимент проводился в два этапа:

1. Анализ без учета контекста — на вход модели подавался только текст комментария.
2. . Анализ с учетом контекста — на вход модели подавалась объединенная последовательность, включающая текст поста и комментария.

Архитектура модели и параметры обучения на обоих этапах оставались неизменными.

Для оценки качества классификации использовались метрики Precision, Recall и F1-мера, вычисляемые для каждого класса тональности, а также усредненная F1-мера.

$$precision = \frac{TP}{TP + FP},$$

$$recall = \frac{TP}{TP + FN},$$

$$F1 - мера = \frac{2 \cdot precision \cdot recall}{precision + recall},$$

где TP (True Positives) — количество правильно классифицированных положительных объектов; FP (False Positives) — количество неправильно классифицированных положительных объектов; FN (False Negatives) — количество неправильно классифицированных отрицательных объектов.

Результаты эксперимента представлены в таблице.

Таблица

Результаты определения тональности комментариев с учетом и без учета контекста

Подход к анализу	**Класс тональности**	**Precision**	**Recall**	**F1-мера**
Без учета контекста	Положительный	0,71	0,65	0,68
	Нейтральный	0,76	0,79	0,77
	Отрицательный	0,62	0,58	0,60
	Среднее	—	—	0,68
С учетом контекста	Положительный	0,74	0,69	0,71
	Нейтральный	0,78	0,81	0,79
	Отрицательный	0,70	0,67	0,68
	Среднее	—	—	0,73

Результаты эксперимента демонстрируют рост средней F1-меры с 0,68 до 0,73 при использовании контекста исходного поста. Наибольший прирост качества наблюдается для отрицательного класса, что подтверждает зависимость эмоциональной окраски комментариев от содержания обсуждаемого сообщения.

Например, комментарий «Ты вообще понимаешь, что пишешь?» при отсутствии контекста может интерпретироваться как нейтральный. Однако при анализе в контексте обсуждения социально значимой темы он корректно классифицируется как негативный, что подтверждает эффективность предложенного подхода.

Заключение

В статье представлен подход к семантическому анализу текстовых данных с учетом контекста сообщений, основанный на использовании сверточной нейронной сети. Показано, что учет диалогового контекста посредством объединения текста поста и комментария позволяет повысить качество определения тональности комментариев без изменения архитектуры модели.

Разработанная модель может быть использована в системах мониторинга пользовательской активности и выявления подозрительных пользователей. В качестве направлений дальнейших исследований рассматривается учет расширенного диалогового контекста и применение более сложных архитектур нейронных сетей.

Список литературы

1. Захарова, О. И. Семантический анализ и синтез текстовых данных / О. И. Захарова // Вестник ВГУ. Серия: Системный анализ и информационные технологии. – 2023. – № 4. – С. 182-208.
2. Ватолин, А. А. Структурный анализ настроений с большими языковыми моделями: решение, выигравшее в соревновании RuOpinionNE-2024 / А. А. Ватолин // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2025”, April 23-25, 2025. – Moscow, Russia, 2025. – С. 416-434.
3. Двойникова, А. А. Аналитический обзор подходов к распознаванию тональности русскоязычных текстовых данных / А. А. Двойникова, А. А. Карпов // Информационно-управляющие системы. – 2020. – № 4. – С. 20-30.
4. Семина, Т. А. Анализ тональности текста: современные подходы и существующие проблемы / Т. А. Семина // Социальные и гуманитарные науки. Отечественная и зарубежная литература. – 2020. – № 4. – С. 47-64.
5. Павлов, Ю. Н. Сравнение методов оценки тональности текста / Ю. Н. Павлов, К. А. Майструк // Молодой ученый. – 2016. – № 12 (116). – С. 59-64.
6. Богданов, А. Л. Сентимент-анализ коротких русскоязычных текстов в социальных медиа / А. Л. Богданов, И. С. Дуля // Вестник Томского государственного университета. Экономика. – Томск, 2019. – № 47. – С. 220-241.
7. Хобсон, Л. Обработка естественного языка в действии / Л. Хобсон, Х. Ханнес, Х. Коул. – Питер, 2020. – 576 с.

Моделирование угловых характеристик вектора направления на объект по его двумерной проекции

В. А. Каладзе [1], e-mail: wakaladze@yandex.ru
Н. А. Шаповалов [2], e-mail: nikita_shapovalov_ivt221@mail.ru

[1] Воронежский государственный университет
[2] Международный институт компьютерных технологий

Аннотация. *Представлен аналитический метод формирования пространственного вектора направления на объект на основе плоского информационного следа наблюдений, включённого в двумерную проекцию факторного пространства. Метод основан на преобразовании двумерной информации следа в углы азимута и места искомого объекта. Реализацией является поворотная платформа.*

Ключевые слова: *калибровка камеры, внутренние параметры, перспективная модель, восстановление направления, нормализованный вектор, поворотная система, угол азимута, угол места.*

Введение

Проблема восстановления пространственного направления на объект по двумерным измерениям представляет интерес для многих инженерных и научных задач. Подобные процедуры востребованы в измерительных комплексах, системах наведения и других технических устройствах, использующих плоскостные данные наблюдений. Корректное восстановление направления позволяет осуществлять пространственную привязку наблюдаемых объектов и формировать управляющие воздействия в механических системах ориентации.

Восстановление углов поворотной системы представляет собой ключевой этап, обеспечивающий переход от математической модели к реальному исполнительному механизму. От корректности такого преобразования зависит точность наведения, устойчивость алгоритмов слежения и общий уровень функциональности системы.

1. Постановка задачи

Исследовалась система, реализующая проекцию трёхмерного объекта на плоскость (рис. 1). Пусть точка, представляющая объект, в трёхмерной декартовой системе имеет координаты:

$$(X_w, Y_w, Z_w)^T \quad (1)$$

© Каладзе В. А., Шаповалов Н. А., 2026

а её фронтальной проекцией является точка на плоскости с координатами:

$$(u,v)^T \qquad (2)$$

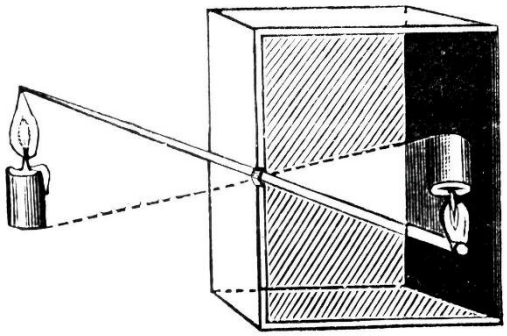


Рис. 1. Исходная система

Рассмотрена задача, в которой по известному радиус-вектору на плоскости, исходящему из проекции точки наблюдения $(0,0)$, с конечной точкой (u,v) определяется трёхмерный радиус-вектор, исходящий из точки наблюдения $(0,0,0)$, с конечной точкой (X_c,Y_c,Z_c), т.е. восстанавливается вектор направления $\bar{d}=(X_c,Y_c,Z_c)^T$ в проектирующей системе. Затем на его основе вычисляются угловые характеристики технической системы – двухосевой поворотной платформы с модулем слежения, позволяющие ориентировать её поисковый вектор (направляющий луч) на объект. Информационная система модуля слежения реализует полярную систему координат (u,v), где задаваемое положение объекта определяют азимут A, угол места E и дальность D; но в данной работе рассматриваются только угловые координаты. Азимут A определяется как угол на горизонтальной плоскости между направлением на объект и отсчётной осью, а угол места E – как вертикальный угол между направлением на объект и горизонтальной плоскостью.

2. Перспективная модель проекции

Решение задачи основывается на использовании перспективной модели проекции, связывающей радиус-вектор на плоскости, исходящий из проекции точки наблюдения, с трёхмерным радиус-вектором $\bar{d}$, в

проектирующей системе. Полученный направляющий луч выражается в системе координат двухосевой поворотной платформы с учётом её ориентации и смещения относительно системы проектирующей системы. Вычисление угловых характеристик платформы основано на геометрическом разложении направляющего луча на соответствующих плоскостях, при котором по его компонентам определяются два угла, вращения платформы в горизонтальной и вертикальной плоскостях.

Перспективная модель задаётся упрощённым матричным выражением [1], отражающим только геометрическое соотношение между поисковым вектором и радиус-вектором на плоскости

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} \tag{3}$$

где: $(X_c, Y_c, Z_c)^T$ – поисковый вектор в системе проектирующей установки; λ – коэффициент масштабирования перспективной проекции, при $\lambda \in [\lambda_{\min}, \lambda_{\max}]$, где $\lambda_{\max} > \lambda_{\min} > 0$ – зависит от рабочей зоны системы наблюдения.

Для учёта метрических свойств реальной системы наблюдения в модель вводится матрица внутренней калибровки A

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = A \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} \tag{4}$$

$$A = \begin{pmatrix} f_x & \gamma & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix} = const \tag{5}$$

где: f_x и f_y – фокусные расстояния от точки наблюдения $(0,0,0)$ до соответствующих координат на плоскости проекции; (u_0, v_0) – координаты оптического центра на плоскости проекции, γ – коэффициент сдвига осей, отражающий несовпадение ориентации горизонтальной и вертикальной осей плоскости проекции.

В реальных оптических системах перспективная проекция подвергается искажениям, возникающим из-за несовершенств линз и конструкции установки. Эти искажения приводят к отклонению измеренных координат точки (u_{dist}, v_{dist}) от идеальных значений (u, v), вычисляемых по перспективной модели. Для корректного

восстановления поискового вектора $\bar{d}$ проводится предварительная компенсация дисторсии с помощью функции $Undisort(u_{dist}, v_{dist})$, которая преобразует искажённые координаты в скорректированные (u, v) [4].

3. Восстановление поискового вектора

Для восстановления трёхмерного радиус-вектора цели из её координат перспективной проекции используется обратное преобразование исходного матричного уравнения:

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = \lambda A^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad (6)$$

Множитель λ отражает расстояние до цели вдоль радиус-вектора и не требуется для ориентации платформы, поэтому может быть исключён, а радиус-вектор определён в виде:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = A^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad (7)$$

Полученный вектор $\bar{d} = (x, y, 1)^T$ – является нормализованным пространственным направлением (рис. 2), то есть определяет луч, проходящий через точку наблюдения и объект.

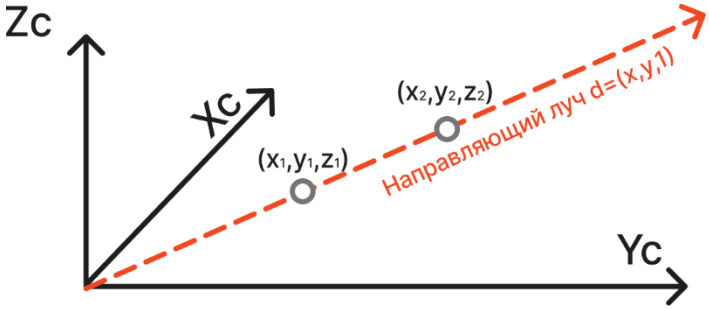


Рис. 2. Направляющий луч

4. Вычисление угловых характеристик

Восстановленный направляющий луч из точки наблюдения первоначально представлен в системе координат проектирующей установки. Для практического использования этого направления в управлении двухосевой поворотной платформой необходимо выразить

его в системе координат платформы, так как эти системы, как правило, не совпадают: оси координат могут иметь различную ориентацию, а начала координат быть смещены.

Связь между вектором в системе проектирующей установки и вектором в системе платформы описывается линейным преобразованием:

$$\begin{pmatrix} x_n \\ y_n \\ 1 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ 1 \end{pmatrix} \quad (8)$$

где $(x_n, y_n, 1)^T$ - вектор направляющего луча, выраженный в системе поворотной платформы; $(x, y, 1)^T$ – вектор луча в системе проектирующей установки; $R = |r_{ij}|$ – матрица поворота, описывающая ориентацию осей платформы относительно системы проектирующей установки; $t = (t_x, t_y, t_s)^T$ - вектор смещения между началами координат двух систем.

Для вычисления угловых характеристик двухосевой поворотной платформы необходимо перейти от декартовой системы координат направляющего луча к полярной. Это обусловлено тем, что ориентация платформы определяется углами вращения вокруг осей. Двухосевая поворотная платформа реализует два взаимно ортогональных вращения: вокруг вертикальной оси – задаётся угол азимута, и вокруг горизонтальной оси – угол места. Поскольку эти вращения происходят в независимых плоскостях, углы θ и φ можно рассматривать как два независимых параметра, определяющих ориентацию платформы вдоль направляющего луча.

Угловые характеристики выражаются через направляющий луч $\bar{d} = (x, y, 1)^T$ и $arctg$ следующим образом:

$$\varphi = arctg(x, z) \quad (9)$$

$$\theta = arctg(y, z) \quad (10)$$

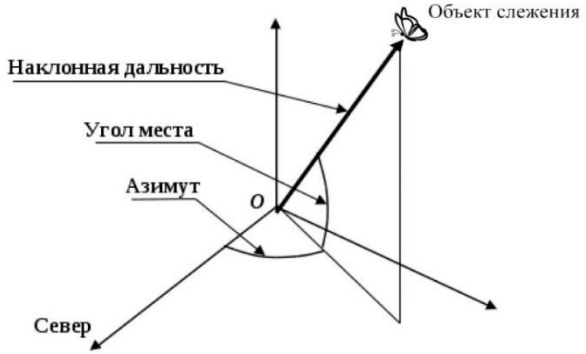


Рис. 3. Угловые характеристики

Заключение

Такой подход позволяет восстановить направление на объект по его двумерной проекции, формируя углы азимута и места двухосевой поворотной системы. Полученные выражения для углов обеспечивают однозначность вычислений и пригодны для практического применения в системах наведения, слежения и позиционирования.

Литература

1. Torralba, A. Camera Modeling and Calibration / A. Torralba, P. Isola, W. Freeman // Imaging Geometry. MIT Vision Book [Электронный ресурс]. – Режим доступа: visionbook.mit.edu/imaging_geometry.html (дата обращения: 21.10.2025).

2. Илья12с. Из фото в 3D. Ч. 2: калибровка камеры / Илья12с // Habr [Электронный ресурс]. – 16.10.2023. – Режим доступа: habr.com/ru/companies/magnus-tech/articles/772534/ (дата обращения: 12.12.2025).

3. Generating Camera Rays with Ray Tracing // Scratchapixel [Электронный ресурс]. – Режим доступа: scratchapixel.com/lessons/3d-basic-rendering/ray-tracing-generating-camera-rays/generating-camera-rays.html (дата обращения: 03.11.2025).

4. Correction of Camera Lens Distortion // Intel Integrated Performance Primitives Documentation [Электронный ресурс]. – Режим: smcm.iqfr.csic.es/docs/intel/ipp/ipp_manual/IPPI/ippi_ch14/ch14_correction_of_camera_lens_distortion.htm (дата обращения: 10.11.2025).

5. Хартли, Р. Многовидовая геометрия в компьютерном зрении : пер. с англ. / Р. Хартли, А. Зиссерман. – М. : Техносфера, 2012. – 824 с.

Сравнительная оценка подходов к автоматической сегментации и радиомическому анализу КТ грудной клетки

С. С. Куликов, email: sergeykulikovofficial@yandex.ru

Федеральное государственное бюджетное образовательное учреждение высшего образования "МИРЭА - Российский технологический университет"

Аннотация. *В данной работе рассматриваются методы автоматической сегментации и радиомического анализа изображений компьютерной томографии органов грудной клетки для повышения точности диагностики, а также анализируются современные архитектуры нейронных сетей, применяемые для обработки КТ-данных.*

Ключевые слова: *глубокое обучение, нейронные сети, компьютерная томография, органы грудной клетки, семантическая сегментация, радиомика, международная конференция, медицинская визуализация, U-Net, трансформерная архитектура, трансформер, искусственный интеллект, автоматизированная диагностика.*

Введение

Заболевания органов грудной клетки являются одной из ведущих причин заболеваемости и смертности во всём мире. Патологии лёгких, средостения и других структур грудной клетки требуют своевременной и точной диагностики, поскольку ошибки или задержки в выявлении изменений могут приводить к прогрессированию заболеваний и снижению эффективности лечения. В связи с этим повышение точности интерпретации данных компьютерной томографии органов грудной клетки остаётся актуальной задачей современной радиологии.

В последние годы методы искусственного интеллекта и машинного обучения получили широкое распространение в медицине. Алгоритмы машинного обучения позволяют анализировать большие объёмы данных, выявлять закономерности и автоматизировать сложные процессы, что повышает точность и скорость их выполнения.

Одним из ключевых направлений применения ИИ в радиологии является автоматическая сегментация анатомических структур и патологических изменений на КТ-изображениях ОГК. Сегментация позволяет выделять области интереса, такие как опухоли или очаги воспаления, и служит основой для последующего количественного анализа. В частности, радиомический анализ, основанный на извлечении

© Куликов С. С., 2026

признаков из сегментированных областей, предоставляет дополнительные диагностически значимые характеристики, недоступные при визуальной оценке.

Несмотря на успехи, задачи автоматической сегментации и радиомического анализа КТ-данных сопряжены с рядом сложностей: вариативностью протоколов сканирования, различиями между КТ-сканерами, наличием шумов и артефактов, а также ограниченным объёмом размеченных данных. В данной работе рассматриваются современные методы и алгоритмы сегментации и радиомического анализа КТ-изображений органов грудной клетки, выполнен анализ и сравнение различных архитектур нейронных сетей, а также сформулированы выводы о перспективах их применения в клинической практике.

1. Существующие методы и решения. Отечественный опыт

Существующие решения на основе нейронных сетей глубокого обучения, широко применяемые в медицине для анализа изображений, используют различные архитектуры моделей в зависимости от типа данных и поставленной задачи. В задачах обработки изображений компьютерной томографии органов грудной клетки наибольшее распространение получили свёрточные нейронные сети (CNN), демонстрирующие высокую вычислительную эффективность и устойчивость при сегментации и классификации патологических изменений.

В ряде отечественных работ для анализа изображений (из другой области) применяются архитектуры семейства YOLO, адаптированные для медицинской визуализации. Так, в работах [1] была использована архитектура YOLOv8x без дополнительных модификаций для детектирования и локализации патологических областей. В результате были достигнуты значения метрик mAP50–95 = 73,4% и mAP50 = 96,1%, что свидетельствует о высокой точности обнаружения объектов на изображениях.

В другом исследовании для анализа специализированного датасета, сформированного на основе архивных данных Военно-медицинской академии и открытых наборов Kaggle, применялась архитектура ConvNeXt [2]. Данная модель использовалась для классификации патологических состояний по КТ-изображениям органов грудной клетки. В качестве оценочной метрики применялась accuracy, максимальное значение которой составило 79,6%.

Более сложные архитектурные решения представлены в работе [3], где для обработки КТ-данных использовалась модифицированная архитектура Inception-v3. Ключевыми преимуществами данной модели

являются эффективное извлечение признаков за счёт использования Inception-блоков, снижение риска переобучения и возможность работы с изображениями высокого разрешения. В процессе обучения применялся оптимизатор Adam и функция потерь categorical cross-entropy. Для контроля переобучения использовался метод early stopping с порогом в 50 эпох, а разбиение выборки выполнялось с использованием k-fold кросс-валидации (k = 5). Качество модели оценивалось с помощью метрик F1-score и AUC-ROC.

В смежных темах встречается использование модели под названием U-Net, которая часто модифицируется под решаемую задачу. Она имеет структуру, которая позволяет не только абстрагировать с помощью структуры энкодер-декодер, но и «пробрасывать» изначальные признаки в каждый слой декодера (рис. 1).

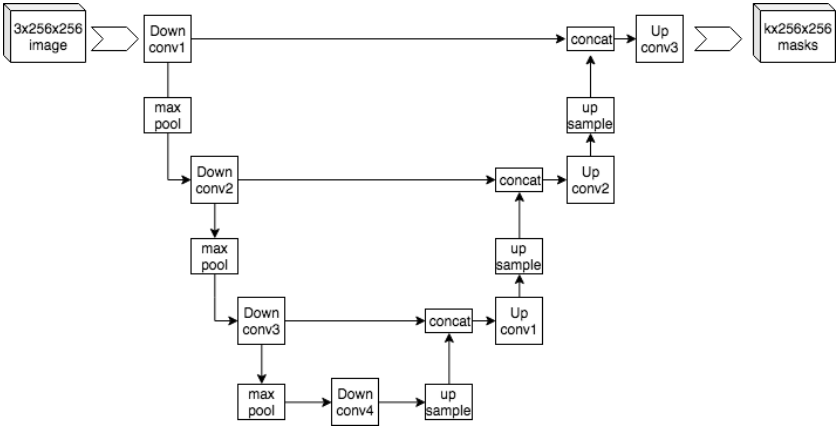


Рис. 1. Базовая U-Net архитектура

2. Существующие методы и решения. Зарубежный опыт

За последние пять лет в зарубежной научной литературе опубликовано значительное количество работ, посвящённых применению методов глубокого обучения для анализа радиологических изображений, включая КТ органов грудной клетки. Основное внимание уделяется разработке и модификации архитектур нейронных сетей для задач сегментации, классификации и извлечения количественных признаков, а также созданию и аннотированию специализированных наборов данных.

Важным направлением является формирование открытых датасетов для обучения и валидации моделей, что обеспечивает воспроизводимость

результатов и ускоряет внедрение новых методов в клиническую практику. В качестве примера можно привести работу RIAWELC [4], где представлен открытый датасет радиографических изображений и предложен конвейер обработки данных с автоматизированной разметкой и классификацией (рис. 2). Такие подходы позволяют ускорить подготовку датасетов и могут быть адаптированы для КТ-изображений ОГК, где качество разметки является ключевым ограничением.



Рис. 2. Пайплайн RIAWELC

При анализе зарубежного опыта важно учитывать, что даже если решения не касаются напрямую сегментации КТ-изображений ОГК, применяемые подходы к построению архитектур и обработке данных могут быть переносимы на задачи медицинской визуализации.

Особое внимание в зарубежных исследованиях уделяется различным модификациям U-Net, включая классическую U-Net, U-Net3+, Monai U-Net, nnU-Net и U-Net-R++ [5]. Эти модели применяются как к двумерным (2D), так и к трёхмерным (3D) данным, что позволяет учитывать пространственный контекст и повышать точность сегментации.

Использование 2D-нарезок из 3D-объёмов позволяет адаптировать трёхмерные данные для обучения стандартных 2D-моделей, тогда как работа с 3D-вокселями улучшает учёт взаимосвязей между срезами. Такие методы актуальны для сегментации патологий на КТ-изображениях ОГК, особенно при анализе объёмных данных и необходимости устойчивости к переобучению. [6]

В зарубежной литературе также активно исследуются трансформерные архитектуры, такие как SegFormer и Swin Transformer, которые демонстрируют высокую точность за счёт улавливания глобального контекста изображения. Эти модели особенно эффективны при работе с большими датасетами и сложными задачами сегментации, однако требуют значительных вычислительных ресурсов.

В статье [7] рассматриваются ансамбли и DINOv2 — модель self-supervised learning, позволяющая извлекать универсальные признаки из изображений без необходимости аннотированных данных. DINOv2

может использоваться как универсальный энкодер признаков для задач сегментации и классификации медицинских изображений, что важно при ограниченности размеченных данных.

Однако высокая вычислительная нагрузка и требования к инфраструктуре ограничивают практическое применение DINOv2 и других крупных трансформерных моделей в условиях медицинских учреждений с ограниченными ресурсами, что делает актуальными более лёгкие и адаптивные решения, такие как модификации U-Net.

3. Обсуждение. Сравнение результатов

Современные исследования в области автоматической сегментации и анализа медицинских изображений демонстрируют широкий спектр архитектур и методов, направленных на повышение точности, устойчивости и интерпретируемости моделей. Каждый подход имеет свои преимущества и ограничения, которые необходимо учитывать при выборе решения для клинической практики.

Одним из наиболее перспективных направлений являются трансформерные архитектуры, такие как Swin Transformer, SegFormer и DINOv2. Их преимущество заключается в улавливании глобального контекста изображения и возможности обучения без большого объёма размеченных данных благодаря self-supervised подходам. Однако высокая вычислительная нагрузка и требования к аппаратуре ограничивают их применение в условиях медицинских учреждений с ограниченными ресурсами.

Чтобы компенсировать дефицит размеченных данных, применяются методы работы с несбалансированными выборками, например, PatchAlign и Saliency Mix-up. Они помогают улучшить представление редких классов, увеличивая общее качество классификации. Однако их применение требует более сложной настройки и анализа, что увеличивает подготовительный этап проекта и делает модели менее универсальными.

С точки зрения интерпретируемости и доверия со стороны врачей, всё большую роль играет направление объяснимых моделей. Модели, включающие механизмы визуализации внимания, позволяют врачу понять, почему алгоритм принял то или иное решение. Это критично для внедрения ИИ в реальную медицинскую практику. Тем не менее, модели с расширенной объяснимостью зачастую уступают в производительности, а также страдают от трудностей масштабирования.

Для улучшения обработки локальных особенностей изображений в исследованиях применяются билинейные CNN и архитектуры с двойным вниманием (Dual Attention). Эти модели эффективно захватывают пространственные и канальные связи, что особенно важно при наличии выраженных текстурных различий. Однако с ростом сложности

возрастает и риск переобучения, особенно при ограниченных обучающих данных.

Чтобы решить проблему дефицита медицинских датасетов есть несколько методов, один из них — применение генеративных моделей, в том числе использование модели Stable Diffusion. Они позволяют создавать синтетические изображения, расширяя тренировочную выборку и улучшая устойчивость к классовой диспропорции. Однако эти изображения могут содержать артефакты и не всегда соответствуют клинической реальности, что требует дополнительного контроля со стороны экспертов.

Минусом остаётся значительная вычислительная нагрузка, необходимость хранения и запуска сразу нескольких моделей, что ограничивает практическое применение.

В условиях, где критична скорость обработки, используются архитектуры вроде YOLO-like, DeepLabv3+ и SegNet. Они обеспечивают быструю сегментацию и классификацию, что важно для систем реального времени. Однако простота и скорость зачастую достигаются ценой снижения точности и детализации сегментации, что критично для диагностики тонких патологических изменений на КТ.

На фоне вышеперечисленных решений особенно выгодно смотрится архитектура U-Net и её модификации — nnU-Net, Monai U-Net, U-Net-R++ и другие. Главным её преимуществом остаётся сбалансированность: энкодер-декодер структура с пропускными связями позволяет эффективно извлекать как глобальные, так и локальные признаки, обеспечивая высокую точность даже при небольших объёмах данных. nnU-Net автоматизирует настройку архитектуры под конкретный датасет, а Monai U-Net интегрируется с медицинскими библиотеками и стандартами. Эти модели легче адаптируются к различным типам данных (2D-изображения, 3D-объёмы, нарезки) и менее требовательны к вычислительным ресурсам по сравнению с трансформерами.

Хотя более современные архитектуры вроде SegFormer, ConvNeXt и Inception-v3 предлагают высокую точность за счёт глубокой структуры и широких рецептивных полей, они требуют большего количества данных и более тщательной настройки. При этом семейство U-Net остаётся более устойчивым и гибким решением, особенно в условиях ограниченного доступа к ресурсам.

Также стоит отметить важность метрик оценки: accuracy, F1-score, mAP, IoU и ROC-AUC позволяют объективно сравнивать производительность различных моделей. Однако ни одна из них не даёт полной картины без анализа клинической применимости.

Особое внимание следует уделять вариативности входных изображений: различия протоколов сканирования, артефакты и шумы значительно влияют на результат. Механизмы внимания и аугментации частично помогают справиться с этим, но полное устранение влияния этих факторов возможно только при стандартизации сбора данных и предобработке.

В целом, можно сказать, что существует компромисс между точностью, объяснимостью, скоростью и устойчивостью к шумам. Среди всех подходов именно U-Net и её модификации представляют собой наиболее сбалансированное и практичное решение для задач сегментации КТ-изображений. Их высокая точность, стабильность, гибкость и сравнительно невысокие вычислительные требования делают их оптимальным выбором в текущих реалиях медицинского ИИ.

Если провести оценку моделей и составить таблицу эффективности по тесту на одном из открытых наборов данных ISIC 2019 Challenge Dataset (усечённый 2D набор данных с задачей сегментации), результаты которых можно найти в публичном доступе. Проведённый анализ качества разных моделей на этих данных (Таблица) позволяет объективно сравнить архитектуры и выбрать наиболее подходящую для задач сегментации.

Таблица

Сравнительные характеристики моделей

Модель	**ROC-AUC**	**Точность (Accuracy)**	**F**
Resnet50	0.810	0.735	0.550
DenseNet121	0.830	0.800	0.615
EfficientNet_v2_s	0.856	0.873	0.726
MobileNet_V2	0.836	0.790	0.640
InceptionV3	0.787	-	0.784
U-Net (классическая)	0.845	-	0.694
ConvNeXt	0.850	-	0.686

Заключение

Анализ применения методов машинного обучения показывает, что они открывают новые возможности для повышения точности диагностики и оптимизации работы врачей, позволяя уделять больше внимания сложным клиническим случаям. Это особенно важно при большом объёме однотипных исследований и строгих критериях интерпретации КТ-изображений органов грудной клетки.

Анализ существующих подходов показал, что свёрточные нейронные сети (CNN) остаются одним из основных инструментов для задач сегментации и классификации медицинских изображений. Архитектуры вроде YOLO-like, ConvNeXt, Inception-v3, SegFormer демонстрируют высокую эффективность при различных подходах к обучению моделей. Однако одной из наиболее распространённых и успешных архитектур для сегментации остаётся U-Net (U-Net3+, U-Net-R++), которая благодаря энкодер-декодер структуре и пропускным связям эффективно обрабатывает изображения и сохраняет детали. Использование метрик (accuracy, mAP, F1 и ROC-AUC) позволяет оценивать качество сегментации и классификации на КТ-данных. При этом достижение высокой точности требует не только правильно выбранной архитектуры, но и тщательной настройки гиперпараметров, обработки данных и предотвращения переобучения.

Основные ограничения при применении ИИ в сегментации КТ связаны с вариативностью изображений (различия протоколов сканирования, артефакты, шумы) и необходимостью больших объёмов качественных аннотированных данных для обучения моделей. Также важно учитывать клиническую специфику патологий органов грудной клетки, которая требует высокой точности сегментации для корректной диагностики и принятия решений.

В целом, исследования в области применения ИИ к анализу КТ-изображений органов грудной клетки открывают новые перспективы для повышения качества медицинской помощи. Для достижения максимального эффекта необходимо дальнейшее развитие методов машинного обучения, разработка адаптивных моделей и их интеграция в клиническую практику.

Список литературы

1. Использование нейросети для диагностики заболеваний ногтевой пластины / В. Е. Тельманов [и др.] // В сборнике: Новые информационные технологии в нефтегазовой отрасли и образовании. Материалы XI Международной научно-практической конференции-конкурса. – Тюмень. – 2024. – С. 264-267.

2. Применение глубокого обучения в диагностике кожных заболеваний / С. О. Самохин [и др.] // В сборнике: XVIII Санкт-Петербургские дерматологические чтения. – Материалы научно-практической конференции дерматовенерологов и косметологов. – Санкт-Петербург. – 2024. – С. 26-29.

3. Ensemble CNN models for Covid-19 Recognition and Severity Perdition From 3D CT-scan [Электронный ресурс] / F. Bougourzi [and all] //

Computing Research Repository (CoRR). – 2022. Режим доступа : https://arxiv.org/abs/2206.15431

4. RIAWELC: A Novel Dataset of Radiographic Images for Automatic Weld Defects Classification [Электронный ресурс] / B. Totino [and all] // International Journal of Electrical and Computer Engineering Research. – 2023. – Режим доступа : https://www.academia.edu/100346388

5. Shaker, A. UNETR++: Delving Into Efficient and Accurate 3D Medical Image Segmentation / A. Shaker [and all] // IEEE Transactions on Medical Imaging. – 43-е издание. – №9. – 2024. – С. 3377-3390.

6. Comparative Analysis of Radiomic Features and Gene Expression Profiles in Histopathology Data Using Graph Neural Networks [Электронный ресурс] / L. Carlos [and all] // Computing Research Repository (CoRR). – 2023. – Режим доступа : https://arxiv.org/abs/2312.15825

7. DINOv2: Learning Robust Visual Features without Supervision [Электронный ресурс] / Q. Maxime [and all] // Computing Research Repository (CoRR). – 2021. – Режим доступа : https://arxiv.org/abs/2105.15203

Сравнительный анализ алгоритмов обучения с подкреплением для компьютерных игр

Д. Р. Небогина, email: dasha.nebogina@mail.ru
А. Ю. Иванков, email: ivankov@cs.vsu.ru

Воронежский государственный университет

Аннотация. *Проведён сравнительный анализ алгоритмов MuZero, EfficientZero и ReZero в условиях ограниченного объёма данных с использованием библиотеки LightZero.*

Ключевые слова: *Обучение с подкреплением, MuZero, EfficientZero, ReZero, поиск по дереву Монте-Карло, сравнительный анализ.*

Введение

Современные методы обучения с подкреплением демонстрируют выдающиеся результаты в решении сложных задач, таких как стратегическое планирование и моделирование адаптивного поведения в динамических средах. Компьютерные игры служат идеальными тестовыми средами для таких алгоритмов благодаря возможности самообучения без размеченных данных, простоте интерпретации полученных результатов и вариативности сложности задач. Агент взаимодействует со средой через три ключевых компонента: состояние, действия и награду [1-2]. Такой подход обеспечивает переносимость решений в реальные сценарии, где сбор данных может быть затруднён. Например, завод можно моделировать как игру, где состояние – износ и нагрузка оборудования, действия – его настройки, а награда – минимизация затрат и продление срока службы.

Прорывом в области стали AlphaGo и AlphaZero, достигшие сверхчеловечких результатов в настольных играх через комбинацию обучения с подкреплением и поиска по дереву Монте-Карло. Его развитие привело к созданию MuZero, EfficientZero и ReZero, – алгоритмов, преодолевших ограничения предшественников за счёт повышения эффективности использования данных и снижения вычислительных затрат.

Сравнительный анализ MuZero, EfficientZero и ReZero важен для оценки эффективности выборки и вычислений: каждый последующий алгоритм улучшает предыдущий. Такие сравнения помогают выбирать или комбинировать методы под конкретные задачи.

© Небогина Д. Р., Иванков А. Ю., 2026

1. Muzero

MuZero стала продолжением идей, реализованных в AlphaZero [3], добавляя способность работать в средах с неизвестной динамикой, возможность использования среды с одним агентом и ненулевыми вознаграждениями на промежуточных этапах, что позволяет обучаться модели не только в настольные игры. Ключевыми инновациями MuZero являются: архитектура с тремя подсетями, поиск с использованием поиска по дереву Монте-Карло и обучение.

Рассмотрим архитектуру с тремя подсетями. Каждая из них оптимизирует собственную функцию:

- Функция представления h_θ преобразует последовательность наблюдений $o_{1:t}$ в скрытое состояние s_0.
- Функция динамики g_θ рекуррентно обновляет состояние: $r_k, s_k = g_\theta(s_{k-1}, a_k)$, предсказывая вознаграждение r_k и новое скрытое состояние s_k.
- Функция предсказания f_θ оценивает политику p_k и функцию ценности v_k и s_k.

Скрытое состояние оптимизировано исключительно для точного прогнозирования (p_k, v_k, r_k), что позволяет моделировать динамику даже в частично наблюдаемых средах (например, в Atari).

Поиск с MCTS строится на основе скрытых состояний. Каждый узел дерева поиска ассоциирован с внутренним состоянием s. Для каждого действия a из s существует ребро (s, a), которое хранит набор статистик $\{N(s, a), Q(s, a), P(s,a), R(s, a), S(s, a)\}$, где N – количество посещений, Q – средняя ценность, P – политика, R – вознаграждение и S – переход состояния.

Сначала проходит стадия выбора. Для каждого гипотетического временного шага симуляции действие a^k выбирается в соответствии с сохраненной статистикой для внутреннего состояния s^{k-1} путем максимизации верхней доверительной границы:

$$a_k = \arg\max_a \left[Q(s,a) + P(s,a) \cdot \sqrt{\frac{\sum_b N(s,b)}{1+N(s,a)}} \cdot \right.$$
$$\left. \cdot \left(c_1 + log\left(\frac{\sum_b N(s,b) + c_2 + 1}{c_2} \right) \right) \right],$$

Константы c_1 и c_2 используются для контроля влияния априорной $P(s, a)$ относительно ценности $Q(s, a)$ по мере того, как узлы посещаются чаще.

Для $k < l$ следующее состояние и вознаграждение ищутся в таблице переходов состояний и вознаграждений $s_k = S(s_{k-1}, a_k), r_k = R(s_{k-1}, a_k)$.

На стадии расширения в листовом узле s_l вычисляются $r_l, s_l = g_\theta(s_{l-1}, a_l)$ и $p_l, v_l = f_\theta(s_l)$, после чего создаётся новый узел.

И при обратном распространении для $k = l...0$ обновляется накопленное вознаграждение:

$$G^k = \sum_{\tau=0}^{l-1-k} \gamma^\tau r_{k+1+\tau} + \gamma^{l-k} v^l.$$

Для общих сред Q-значения нормализуются в диапазон $[0,1]$.

Обучение сети происходит на минибатчах из буфера воспроизведения. Для каждого начального состояния s_t модель разворачивается на K шагов, минимизируя суммарную функцию потерь [4].

С учетом всех новшеств MuZero обладает большей производительностью в сравнении с AlphaZero и при этом требует меньшее количество обучающих данных.

2. EfficientZero

EfficientZero представляет собой модель обучения с подкреплением, значительно повышающую эффективность использования данных по сравнению с MuZero. Архитектура EfficientZero включает три улучшения: самообучаемую функцую согласованности состояний, сквозное предсказание префикса ценности и коррекцию смещения политики.

Самообучаемая функция согласованности состояний – это тензорное обучение состояний, которое направлено на преодоление нестабильности использования скалярных сигналов в MuZero:

$$L_{similarity}\left(\hat{s}_{t+1}, s_{t+1}\right) = -cos\left(P_1\left(sg\left(s_{t+1}\right)\right), P_2\left(P_1\left(\hat{s}_{t+1}\right)\right)\right),$$

где $\hat{s}_{t+1} = G(s_t, a_t)$ – предсказанное следующее состояние, $s_{t+1} = H(o_{t+1})$ – реальное скрытое состояние, полученное из следующего наблюдения, P_1 и P_2 – проекторные сети (на основе архитектуры SimSiam), sg – остановка градиентов.

Сквозное предсказание префикса ценности используется для решения проблемы синонимичных состояний – визуально идентичных кадров с разными исходами. В EfficientZero введена LSTM-сеть – f, прогнозирующая накопленное вознаграждение за k шагов:

$$value_prefix = f\left(s_t, \hat{s}_{t+1}, \ldots, \hat{s}_{t+k-1}\right).$$

Коррекция смещения политики минимизирует влияние устаревших данных при обучении. Идея заключается в том, чтобы использовать реальные вознаграждения только для первых l шагов (где $l < k$), а для оставшихся шагов применять поиск по дереву Монте-Карло с текущей политикой:

$$z_t = \sum_{i=0}^{l-1} \gamma^i u_{t+i} + \gamma^l \nu_{t+l}^{MCTS},$$

где ν_{t+l}^{MCTS} – средняя корневая ценность дерева Монте-Карло, развернутого из состояния s_{t+l} с текущей политикой. Параметр l динамически регулируется в зависимости от "свежести" данных [5].

Но EfficientZero требует дискретизации действий для непрерывных пространств, что может ограничивать его применимость.

3. ReZero

ReZero значительно сокращает время обучения при сохранении высокой эффективности использования данных. ReZero вносит два фундаментальных улучшения – это обратный анализ траекторий и реанализ буффера.

Идея обратного анализа траекторий заключается в следующем: если известно истинное значение состояния-потомка, поиск в соответствующем поддереве можно пропустить. То есть для пары последовательных состояний S_t^l и S_{t+1}^l при обратном проходе по траектории, значение корня m_{t+1}^l, полученное при поиске из S_{t+1}^l,

используется для оценки перехода из S_t^l в S_{t+1}^l. При выборе действия a_t^l в узле S_t^l применяется модифицированное правило:

$$I_t^l(a) = \begin{cases} \text{UCBscore}(S_t^l, a), & a \neq a_t^l \\ r_t^l + \gamma m_{t+1}^l, & a = a_t^l \end{cases},$$

где t_t^l – предсказанное вознаграждение, γ – коэффициент дисконтирования. Это позволяет немедленно завершить симуляцию при выборе действия a_t^l, экономя вычислительные ресурсы.

ReZero использует периодический анализ всего буфера воспроизведения вместо частого анализа мини-батчей. После фиксированного числа итераций обучения весь буфер проходит через повторный анализ с актуальной моделью. На этапе сбора данных ReZero заменяет MCTS на прямую выборку действий из политики сети, что снижает вычислительные затраты.

Целевое значение для обучения формируется гибридным способом:

$$z_t = \sum_{i=0}^{l-1} \gamma^i u_{t+i} + \gamma^l v_{t+l}^l,$$

где l динамически регулируется в зависимости от «свежести» данных.

Сокращение времени обучения позволяет: использовать алгоритмы в задачах с ограниченным временем принятия решений, масштабировать обучение на сложных средах с высокой стоимостью взаимодействия [6].

4. Сравнительный анализ

Для проведения сравнительного анализа алгоритмов MuZero, EfficientZero и ReZero была использована библиотека LightZero, разработанная специально для исследований в области обучения с подкреплением на основе модельного планирования [7].

В качестве тестовой среды была выбрана классическая игра «Змейка», реализованная в рамках стандарта Gymnasium – официального преемника Gym от OpenAI, предоставляющая дискретное пространство действий – 4 направления движения и наблюдения в виде RGB-изображений.

Среда "Змейка" имеет простую детерминированную динамику и редкие вознаграждения: агент получает положительное вознаграждение только при поедании еды, а штраф – при столкновении со стеной или своим телом, что позволяет эффективно оценивать качество планирования и обучения модели.

Обучение алгоритмов проводилось в идентичных конфигурациях среды «Змейка» с ограничением в 30 000 итераций. В качестве целевой

метрики использовалась награда, определяемая количеством успешно собранных элементов «еды» за эпизод. Результаты экспериментов представлены в таблице.

Таблица

Результаты обучения алгоритмов

Алгоритм	**Средняя награда**	**СКО**	**Макс. награда**	**Мин. награда**	**Шаги среды**
MuZero	3.25	3.88	12	0	156564
Efficient Zero	16.67	7.55	27	0	156432
ReZero	18.83	6.85	31	1	75132

Полученные данные демонстрируют положительную динамику как средней, так и максимальной награды при переходе от MuZero к ReZero. Однако сохранение эпизодов с минимальной наградой, включая нулевые значения для MuZero и EfficientZero, указывает на наличие сценариев, в которых алгоритмы не обеспечивают устойчивой эффективности. Вероятными причинами такого поведения могут быть недостаточная глубина архитектуры нейронных сетей или размещение «еды» вблизи границ игрового поля, что создаёт дополнительные сложности для планирования траекторий.

Сравнение EfficientZero и ReZero выявило незначительное увеличение средней награды (+2,16) и максимальной награды (+4) при одновременном сокращении количества взаимодействий со средой более чем в два раза. Данная особенность ReZero делает его предпочтительным для применения в условиях, где сбор данных связан со значительными вычислительными или временными затратами. Тем не менее, сохраняющаяся вариативность результатов (стандартное отклонение $\geq$ 6,85) требует дальнейшей оптимизации архитектуры и гиперпараметров алгоритмов.

Заключение

В данной работе проведён сравнительный анализ современных алгоритмов обучения с подкреплением на основе модельного планирования – MuZero, EfficientZero и ReZero – в условиях ограниченного объёма данных. Используя унифицированную реализацию из библиотеки LightZero и контролируемую среду «Змейка»,

были продемонстрированы ключевые различия в эффективности использования выборки и устойчивости обучения. Результаты подтверждают, что архитектурные усовершенствования EfficientZero и ReZero действительно обеспечивают значительный прирост производительности в задачах с редкими наградами.

Полученные результаты подтверждают потенциал интеграции механизмов самообучения и оптимизации поиска по дереву для создания практичных RL-агентов. Наиболее перспективным направлением представляется гибридизация подходов: использование архитектуры EfficientZero для обеспечения эффективного использования данных в сочетании с методами ReZero для сокращения вычислительных затрат. Это особенно важно для переноса решений в реальные среды — от промышленной робототехники до автономных систем, где баланс между качеством обучения и ресурсоемкостью критичен для практической применимости.

Литература

1. Демченко, М. В. Использование методов обучения с подкреплением в задачах медицинской практики / Демченко, М. В., Каширина, И. Л., Фирюлина, М. А. // Вестник ВГУ. Серия: Системный анализ и информационные технологии. – 2022. – № 1. – С. 111-124.
2. Tian, Xianhe. AI applications in video games and future expectations [Электронный ресурс] / Xinhe Tian. // Applied and Computational Engineering. – 2024. – № 54. − Режим доступа: https://www.researchgate.net/publication/379406019_AI_applications_in_video_games_and_future_expectations
3. Wang, Hui. Monte Carlo Q-learning for General Game Playing [Электронный ресурс] / H. Wang, M. Emmerich, A. Plaat. // arXiv e-prints. – 2018. – 1802.05944 – Режим доступа: https://arxiv.org/pdf/1802.05944.
4. Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model [Электронный ресурс] / J. Schrittwieser [et al]. // Nature. – 2020. – № 588. – С. 604−609. − Режим доступа: https://arxiv.org/pdf/1911.08265
5. Mastering Atari Games with Limited Data [Электронный ресурс] / W. Ye [et al]. // arXiv e-prints. – 2021. – 2111.00210 − Режим доступа: https://arxiv.org/pdf/2111.00210
6. ReZero: Boosting MCTS-based Algorithms by Backward-view and Entire-buffer Reanalyze [Электронный ресурс] / C. Xuan [et al]. // arXiv e-prints. – 2024. − 2404.16364. – Режим доступа: https://arxiv.org/pdf/2404.16364
7. LightZero: A Unified Benchmark for Monte Carlo Tree Search in General Sequential Decision Scenarios [Электронный ресурс] / Y. Niu [et al]. // arXiv e-prints. – 2023. − 2310.08348 – Режим доступа: https://arxiv.org/pdf/2310.08348

Алгоритм синтеза бимодальной свертки с кросс-доменным согласованием для генерации профиля поверхности

И. Н. Паламарь, email: irina.palamar@mail.ru
А. И. Гагарина, email: aig.rsatu@mail.ru

Рыбинский государственный авиационный технический университет имени П.А. Соловьева

Аннотация. *Предложен алгоритмический подход для физически обоснованной генерации профиля поверхности с учетом микроструктуры материала поверхностного слоя. Алгоритм синтезирует интерпретируемый оператор бимодальной свертки, который служит математической моделью совместного отклика системы на параметры технологического процесса. Бимодальная свертка реализована на основе кросс-доменного согласования представлений разнородных данных в пространственной и спектральной областях. Обучение модели проводилось на оригинальных совместных датасетах профилограмм и микрофотографий поперечных шлифов материала поверхностного слоя деталей, изготовленных методом селективного лазерного сплавления. Приведены результаты, демонстрирующие способность алгоритма генерировать физически корректные профили шероховатости.*

Ключевые слова: *бимодальная свертка, пространственный и частотный домены, кросс-доменное согласование, шероховатость поверхности, микроструктура материала поверхностного слоя детали, генеративное моделирование.*

Введение

Обеспечение качества поверхностного слоя материала ответственных деталей машиностроения является важной задачей в условиях непрерывного повышения требований к их эксплуатационным характеристикам. Разработка новых материалов и методов их обработки осложняет эффективное внедрение инноваций в промышленное производство, так как связано с трудоемкой настройкой многопараметрических технологических процессов.

Система контроля качества деталей использует отдельные показатели, например, параметры профиля шероховатости или параметры микроструктуры материала. Эти показатели имеют различную физическую природу и не связаны друг с другом, хотя

© Паламарь И. Н., Гагарина А. И., 2026

существует тенденция разработки комплексных показателей, а также их связи с предсказанием эксплуатационных свойств изделий.

Известно, что микроструктура материала определяет эксплуатационные свойства деталей и влияет на качество поверхности. В настоящее время и в исследовательских работах, и на производстве отдельно используются показатели качества поверхностного слоя, что связано с существующими методиками и используемым оборудованием систем контроля [1]. Преодолеть существующие ограничения можно на основе анализа данных и машинного обучения [2-4].

Для решения проблемы совершенствования контроля качества деталей, изготавливаемых прогрессивными методами, необходима разработка способов прогнозирования качества поверхности на основе мультимодальных данных, чтобы обеспечить полноту выявления взаимосвязей элементов технологического процесса.

1. Анализ проблемы и постановка задачи

На современном этапе развития технологий искусственного интеллекта активно исследуется применение методов машинного обучения для решения различных задач, связанных с оценкой качества промышленных изделий. В работах [5, 6] используется спектральный анализ для оценки шероховатости поверхности. Для классификации микроструктур используются модели глубокого обучения [7].

Помимо решения частных задач необходимо отметить развитие системного подхода к формированию знаний о производственном процессе [8] и внедрению моделей искусственного интеллекта в систему менеджмента качества [9].

Актуальным направлением является использование мультимодальных данных для обучения [10]. Однако источники этих данных физически не связаны друг с другом. Выделяют три основных подхода при работе с данными различной модальности:

– раннее слияние, когда данные разных модальностей объединяются на этапе предобработки;

– позднее слияние, при котором для каждой модальности обучаются отдельные модели, а затем их предсказания объединяются;

– гибридное слияние, когда происходит объединение модальностей на промежуточных слоях нейросети.

Наиболее часто используется позднее слияние [11], но формирование результирующего вектора происходит путем конкатенации векторов признаков, полученных путем параллельной обработки различных модельностей и потере пространственной связи данных. Целью данной работы является разработка алгоритма, основанного на раннем слиянии для создания физически

интерпретируемой модели, связывающей разнородные характеристики материала: топографию поверхности и микроструктурой материала.

Наиболее полную информацию о геометрии поверхности после изготовления содержит первичный профиль [12]. Он включает в себя все неровности, как регулярные, так и случайные. Первичный профиль Pr поверхности представляет собой дискретную выборку высот рельефа $z_i(x)$, полученную с шагом Δx на длине L

$$Pr = \{z_i\}_{i=1}^{N}, \quad z_i = z(x_0 + i \cdot \Delta x),$$

где x_0 – начальная точка;

N – количество отсчетов;

i – номер отсчета, $i = 0,..., N$;

Δx – шаг отсчета.

Особенности формы и параметры первичного профиля непосредственно взаимосвязаны с формой элементов и параметрами микроструктуры материала поверхностного слоя. Например, локальные экстремумы профиля могут коррелировать с границами зерен, а спектральные компоненты профиля могут зависеть от текстуры поверхности. Визуализация данных первичного профиля на поперечном шлифе и микроструктуры материала детали, изготовленной с толщиной слоя 50 мкм, приведена на рисунке 1.

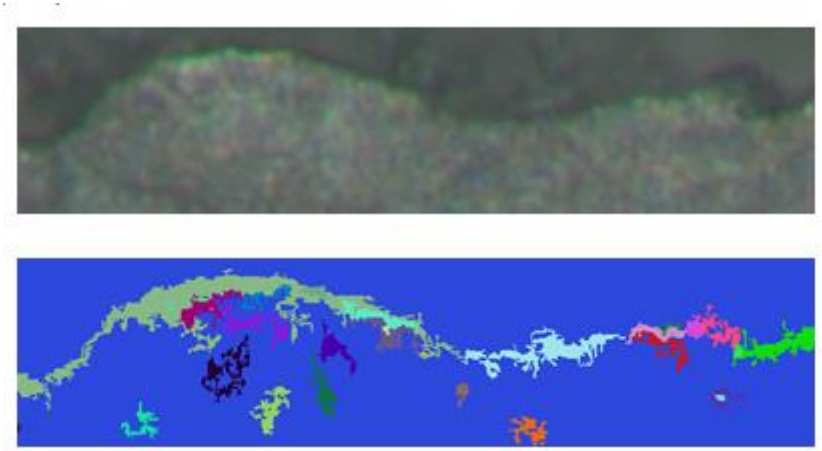


Рис. 1. Визуализация данных первичного профиля и микроструктуры материала (сверху) и сегментная карта микроструктуры (снизу)

Результаты анализа сегментной структуры изображения показали, что сегменты поверхностного слоя неоднородны – имеют разную дисперсию (отображено цветом). Геометрический профиль поверхности является упрощенным представлением, поэтому данные профиля поверхности должны быть обогащены данными микроструктуры материала поверхностного слоя.

Математическая модель первичного профиля поверхности Pr^M, обогащенного данными микроструктуры материала представляет собой дискретную выборку высот рельефа $z_i^m(x)$, полученную с шагом Δx на длине L

$$Pr^M = \{z_i^m(x)\}_{i=1}^N, \quad z_i^m(x) = f(z(x_0 + i \cdot \Delta x), m(x_0 + i \cdot \Delta x)),$$

где f – функция обогащения первичного профиля данными микроструктуры материала;

m – параметр микроструктуры материала для i -ого отсчета.

Обогащение первичного профиля данными микроструктуры выполняется с помощью бимодальной свертки.

2. Концепция бимодальной свертки как процесса поиска оператора сопоставления

Особенностью модальностей используемых данных является то, что данные пространственно связаны и обусловлены единством физического объекта – поверхностного слоя материала детали. Профиль поверхности и прилежащая микроструктура материала являются разными представлениями, полученными соответствующими инструментальными методами с использованием профилометра и микроскопа. Необходимо реализовать их связывание на уровне представления данных в системе машинного обучения для выявления взаимозависимостей процесса обработки детали.

В качестве методологической основы для решения поставленной задачи предлагается концепция бимодальной свертки, понимаемая как процесс поиска оператора сопоставления. В отличие от классической трактовки свертки как математического оператора, в данном контексте она интерпретируется как алгоритмическая процедура установления количественных взаимосвязей между разнородными данными – профилем поверхности $Pr(x)$ и изображением микроструктуры $M(x, z)$ путем синтеза интерпретируемого ядра G. Оператор G выступает формальной моделью, согласующей два различных следствия единого технологического процесса – формирование микроструктуры поверхностного слоя и профиля шероховатости поверхности.

В качестве решения поставленной задачи предлагается концепция алгоритма, основанного на кросс-доменном согласовании данных профиля поверхности и микроструктуры в двух пространствах свертки. Предложенный механизм кросс-доменного согласования представляет собой развитие идеи cross-attention для новой задачи – координации доменов в процессе синтеза физической модели. Если в архитектурах типа трансформер внимание устанавливает семантические связи, то в предлагаемом алгоритме оно служит регулятором, обеспечивающим непротиворечивость оператора G. Это гарантирует, что ядро, объясняющее локальные топологические связи, не противоречит ядру, описывающему глобальные спектральные закономерности, что принципиально отличает метод от ансамбля независимых моделей и повышает надёжность анализа.

Задача бимодальной свёртки состоит не в преобразовании одного представления в другое, а в нахождении оператора G, который описывает, как именно микроструктура M формирует или объясняет наблюдаемый профиль Pr, учитывая параметры процесса обработки. Бимодальная свертка в данном случае это не классическая операция свёртки между тензорами, а сопоставление и совместный анализ двух разных представлений данных.

Установление связей позволит объяснить как особенности профиля (пики, впадины) соотносятся с особенностями микроструктуры (зёрна, границы, фазы) в той же пространственной точке. В пространственной области это может быть локальное сопоставление признаков – например, как высота в точке x профиля коррелирует с текстурой микроструктуры в той же точке. В частотной области можно выявить, что сопоставление периодичностей как доминирующие частоты профиля (крупная/мелкая рябь) соотносятся с периодичностями в структуре материала (размер зерна, слоистость). Кросс-доменное согласование обеспечивает согласованность этих двух типов сопоставлений. Причём важно сохранить топологическую привязку – соответствие между положением вдоль профиля и положением в микроструктуре.

Формализация согласования может быть реализована через совместное латентное пространство, куда проецируются признаки из обоих доменов, и где минимизируется расстояние между соответствующими представлениями профиля и микроструктуры. Кросс-доменное согласование служит механизмом верификации физической осмысленности найденных связей. Например, анализ ядра показывает, какие области микроструктуры (глубина, особенности) наиболее сильно влияют на какие характеристики профиля. Двухдоменная проверка гипотез показывает, что если модель находит ядро G, которое хорошо

согласует форму профиля Pr и особенности микроструктуры M и в пространственной, и в частотной областях, то это может указывать, что найденные связи не статистическая случайность, а отражение реальных физических механизмов.

3. Описание алгоритма синтеза бимодальной свертки с кросс-доменным согласованием

Алгоритм обрабатывает четыре взаимосвязанных потока данных:

– профиль в пространственном домене ($1D$ карта высот);

– микроструктура в пространственном домене ($2D$ изображение);

– спектр профиля в частотном домене (амплитуды/фазы);

– спектр микроструктуры в частотном домене (её частотные характеристики).

Результатом работы алгоритма является генерация физически интерпретируемого и адаптивного ядра свертки, которое совместно учитывает все эти четыре представления и осуществляет кросс-доменную регуляризацию связей между ними. Устанавливаются связи не только между доменами, но и между типами данных внутри каждого домена, а именно: связи между локальной текстурой микроструктуры и локальной высотой профиля в пространственном домене; связи между глобальными периодическими паттернами микроструктуры и преобладающими гармониками шероховатости в частотном домене.

Рассмотрим алгоритм как целостную информационно-управляющую систему. Алгоритм разрешает системное противоречие между необходимостью локального анализа (точечное соответствие текстуры и рельефа) и потребностью в глобальных обобщениях (понимание статистических и частотных закономерностей) через двухуровневую (кросс-доменную) архитектуру.

Основные этапы обработки информации:

1. приведение разнородных данных к форме, пригодной для совместного анализа. Преобразование в оба домена;

2. поиск устойчивых закономерностей, установление соответствий, проверка гипотез, синтез модели;

3. оценка адекватности модели, вычисление метрик согласования, предоставление результатов в интерпретируемом виде.

Первый этап реализуется с помощью энкодеров профиля и микроструктуры, а также вычисления пространственных признаков и спектров. На втором этапе используются механизмы внимания для пространственного и частотного согласования, и выполняется синтез ядра. Третий этап служит для вычисления функций потерь, визуализации ядра и карт внимания для обеспечения обратной связи с пользователем.

Алгоритм реализует замкнутый контур управления с адаптацией на двух уровнях.

4. Управление обучением (внешний контур). На основе обучающей выборки бимодального датасета подстраиваются параметры нейронной сети (веса энкодеров, механизмов внимания). В процессе обучения минимизируется функция потерь, включающая потери внутридоменного и междоменного согласования.

5. Управление согласованием (внутренний контур). На основе признаков из пространственного и частотного доменов выполняется синтез ядра бимодальной свертки. На основе составляющей функции потерь междоменного согласования механизм кросс-доменной регуляризации корректирует взаимодействие потоков, чтобы их гипотезы о ядре не противоречили друг другу.

Функция потерь междоменного согласования реализует отрицательную обратную связь за рассогласование доменов, а веса внимания это положительная обратная связь, усиливающая значимые обнаруженные соответствия. Система адаптивно синтезирует модель ядра под конкретные входные данные профиля шероховатости и микроструктуры, что позволяет анализировать различные материалы и режимы обработки. На основе интерпретации ядра бимодальной свертки можно принять решения об изменении параметров процесса для управления шероховатостью, выявить критичные особенности микроструктуры или обнаружить аномалии.

4. Результаты экспериментального исследования по подготовке данных

Для генерации профилей шероховатости использовалась модель генеративной состязательной нейронной сети [13]. Образцы для исследования были изготовлены методом селективного лазерного сплавления. Изменяемым параметром процесса являлся угол наклона поверхности к плоскости построения. Датасет профилограмм был получен бесконтактным способом с помощью специализированного лазерного микроскопа. Объем датасета составил 6487 файлов профилограмм в формате *csv*. Для добавления в датасет информации о микроструктуре материала поверхностного слоя были подготовлены образцы в виде фотографий поперечного микрошлифа и боковой поверхности, полученных с оптического микроскопа. Совмещение профилограммы и соответствующей микроструктуры выполнялось с помощью реперной точки на образцах. Изображения наружной поверхности образца с параметром угла наклона к плоскости построения 30° при различном увеличении представлены на рисунке 2.

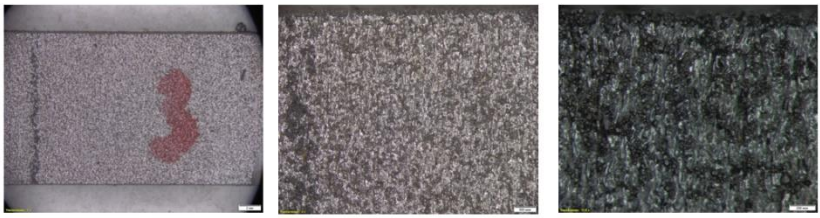


Рис. 2. Изображения наружной поверхности образцов при различном увеличении

Результаты оценки параметра Ra профилей шероховатости для различных значений угла наклона поверхности к плоскости построения детали, сгенерированных с использованием различных способов свертки приведены в таблице 1. Экспериментально получено, что границы зерен на глубине 10-20 мкм наиболее сильно коррелируют с высокочастотными компонентами шероховатости.

Таблица 1

Значения параметра профиля шероховатости профиля Ra, мкм, сгенерированного с использованием различных способов свертки

Угол наклона поверхности, °	**30**	**40**	**50**	**60**	**70**	**80**	**90**
Исходный профиль датасета	4,3	4,9	5,7	6,1	5,1	6,5	7,9
Сгенерированный профиль со сверткой в пространственном домене	4,1	4,8	5,3	5,6	4,6	6,1	7,4
Сгенерированный профиль со сверткой в частотном домене	4,0	4,7	5,5	6,2	4,9	6,3	7,6
Сгенерированный профиль с кросс-доменным согласованием	4,2	5,0	5,8	6,0	5,0	6,6	8,1

Пространственная свертка демонстрирует небольшое систематическое занижение значений, особенно на участках с небольшим наклоном (углы 70°-90°), что может указывать на сглаживание локальных неоднородностей. Частотная свертка показывает результаты, близкие к пространственному методу, но с несколько лучшим соответствием на средних углах (40°-60°), так как эффективнее учитывает периодические составляющие профиля. Кросс-доменное согласование обеспечивает наилучшее соответствие исходным данным на всем диапазоне углов.

Среднее абсолютное отклонение для данного метода минимально, что подтверждает его эффективность в сохранении как локальных деталей (пространственный домен), так и глобальных закономерностей (частотный домен) генерируемого профиля.

Заключение

Новизна алгоритмического подхода обусловлена необходимостью установления детерминированных связей между разнородными данными одного физического объекта – поверхностного слоя материала детали. Его ядром является синтез интерпретируемого оператора бимодальной свёртки, осуществляемый в процессе машинного обучения на датасетах профилограмм поверхности и изображений микроструктуры поперечного шлифа. В результате формируется математическая модель, количественно связывающая профиль поверхности и микроструктуру как совместные отклики на параметры SLM-процесса, что обеспечивает физически обоснованную интерпретацию их взаимосвязи.

Совместная оптимизация с кросс-доменной регуляризацией гарантирует внутреннюю непротиворечивость пространственной и частотной компонент синтезируемого оператора G, что повышает стабильность обучения и обоснованность физической интерпретации конечной модели.

Результатом работы алгоритма является интерпретируемое ядро G, которое позволяет решать две взаимосвязанные задачи: прямого прогнозирования пар «топография – микроструктура» при заданных режимах и обратного управления параметрами процесса для достижения целевых характеристик поверхностного слоя.

Список литературы

1. Исследование зависимости шероховатости, морфологии поверхности и количества дефектов структуры от мощности лазера, скорости сканирования и типа штриховки в жаропрочном сплаве, синтезированном методом SLM / М. А. Лаптева [и др.] // Труды ВИАМ. – 2016. – № 9 (45). – С. 72-84.

2. Modeling process-structure-property relationships in metal additive manufacturing: a review on physics-driven versus data-driven approaches / N. Kouraytem [etc] // J. Phys. Mater. – 2021. – № 4. (032002). – 24 p.

3. Chen, J. Fatigue property prediction of additively manufactured Ti-6Al-4V using probabilistic physics-guided learning / J. Chen, Y. Liu // Additive Manufacturing. – 2021. – 32 p. –DOI:10.1016/j.addma.2021.101876.

4. A review of machine learning in additive manufacturing: design and process / M. Lingbin [etc] // The International Journal of Advanced

Manufacturing Technology – 2024. – V. 135(3-4). – P. 1051-1087. – DOI:10.1007/s00170-024-14543-2.

5. Stępień, K. Research on a surface texture analysis by digital signal processing methods / K. Stępień // Technical Gazette. – 2014. – № 21 (3). – P. 485-493.

6. Lin, W.-J. Evalution of deep learning neural networks for surface roughness prediction using vibration signal analysis / W.-J. Lin // Applied Sciences. – 2019. – № 9 (1462). – 17 p.

7. Advanced Steel Microstructural Classification by Deep Learning Methods. / S.M. Azimi // Sci Rep. – 2018. – № 8 (2128). – 14 p. – DOI: 10.1038/s41598-018-20037-5.

8. Грибова, В. В. Концепция поддержки лазерного аддитивного производства на основе онтологического подхода / В.В. Грибова, В.А. Тимченко // Онтология проектирования. – 2020. – Т. 10. – № 2 (36). – С. 176-189.

9. Yang, H. Six-sigma quality management of additive manufacturing / H. Yang [etc] // Proceedings of the IEEE. – 2021. V. 109. – № 4. – P. 347-376. – DOI:10.1109/JPROC.2020.3034519.

10. Tadas, B. Multimodal machine learning: a survey and taxonomy / B. Tadas, A. Chaitanya, L.-P. Morency // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2019. – P. 423-443. – DOI: 10.1109/TPAMI.2018.2798607.

11. End-to-End multimodal emotion recognition using deep neural networks / P. Tzirakis [etc] // Journal of latex class files. – 2015. – Vol. 14. – № 8. – 9 p. – DOI:10.1109/JSTSP.2017.2764438.

12. ГОСТ Р ИСО 4287-2014. Геометрические характеристики изделий (GPS). Структура поверхности. Профильный метод. Термины, определения и параметры структуры поверхности [Текст]. – Введ. 2016-01-01. – М. : Стандартинформ, 2015. – II, 17 с.; 29 см.

13. An approach to studying surface quality based on automatic generation of nonparametric roughness models / V. F. Bez”yazychnyi [etc] // Journal of Machinery Manufacture and Reliability. – 2022. – V. 51. – № 4. – P. 306-312.

Классификация неисправностей пассивных радиочастотных устройств при помощи нейронных сетей с использованием корней комплексного коэффициента передачи

С. В. Перова, email: sofiyaperova@mail.ru [1]
А. В. Акимов, email: akimov@vsu.ru[1]
А. В. Гречишкин, grechishkin.av@mail.ru [2]

[1] Воронежский государственный университет
[2] АО "Концерн "Созвездие"

Аннотация. *В данной работе предлагается методика диагностики неисправностей пассивных радиочастотных устройств на основе использования корней комплексного коэффициента передачи в качестве входных данных для нейронной сети. Такой подход позволяет уменьшить размерность входных данных и количество параметров нейронной сети. Предложенная архитектура нейросети использует пять скрытых полносвязных слоев, слои нормализации и регуляризации. В результате обучения достигнута точность на тестовой выборке 76.1%.*

Ключевые слова: *диагностика неисправностей, пассивные радиочастотные устройства, машинное обучение, нейронные сети, S-параметры, корни комплексного коэффициента передачи.*

Введение

В процессе производства изделий радиоэлектронной промышленности на стадиях проверки и отладки важно быстро находить и полностью устранять неисправности, возникающие на этапе сборки. Это требует наличия высококвалифицированных кадров. При этом, чтобы разобраться в работе устройства и влиянии различных неисправностей на его параметры, даже опытному специалисту требуется время. Также проверяемое устройство часто имеет много компонентов нескольких типов, что предполагает большое количество мест для проверки потенциальных дефектов. Перечисленные факторы позволяют оценить большую трудоемкость процессов проверки и отладки. При этом эффективная автоматизированная диагностика неисправностей радиочастотных устройств позволит сократить издержки производства и способствует повышению производительности и поддержанию надежности устройств.

© Перова С. В., Акимов А. В., Гречишкин А. В., 2026

Существующие работы, посвященные диагностике неисправностей пассивных радиочастотных устройств, используют S-параметры [1-4]. Так как шаблоны в данных S-параметров могут указывать на специфические особенности сборки электронных устройств, то их удобно использовать для обучения при классификации различных дефектов. С другой стороны, S-параметры, как входные данные, могут иметь достаточно большую размерность. Это зависит от рассматриваемого диапазона частот и количества точек в нем. Например, в работе [2] рассматривались частоты от 1 МГц до 10 ГГц с шагом 100 МГц в 101 частотной точке. То есть исходные данные имели размерность (101, 4) на каждый образец, что, может, и подходит для несложного устройства: рассматривалась схема с 5 компонентами одного типа – резистор с двумя типами неисправностей. Но в изделиях на производстве в разы больше компонентов, что усложняет графики S-параметров и требует большего количества рассматриваемых частот. При этом архитектура нейронной сети может усложняться.

Как альтернативу S-параметрам в качестве входных данных в данной работе предлагается использование корней комплексного коэффициента передачи. Этим предполагается достичь уменьшения размерности входных данных и использования более простой архитектуры нейронной сети, с меньшим количеством обучаемых параметров, для классификации неисправностей.

1. Постановка задачи

Тестируемое устройство – 2-портовый низкочастотный фильтр Кауэра-Чебышева 5-го порядка, представлено на рис.1. Классы неисправностей соответствуют 3 типам дефектов в каждом из 7 компонентов: отсутствие контакта (условное обозначение – `open`) и короткое замыкание (`short`), а также неправильный номинал (`b`, сокращение от `bad`), генерируемый заданием значения номинала компонента с дефектом, выходящим за диапазон допустимого разброса номиналов в ±20%.

Таким образом, рассматривается 22 класса, из которых 21 соответствуют трем типам дефектов, проявляющимся в каждом из 7 компонентов по отдельности. Условное обозначение класса неисправности кроме его типа (`open`, `short` и `b`) содержит номер компонента от 0 до 6, как на схеме с рис. 1. Исправное состояние имеет метку `class_correct`. Каждый класс содержит по 100 экземпляров.

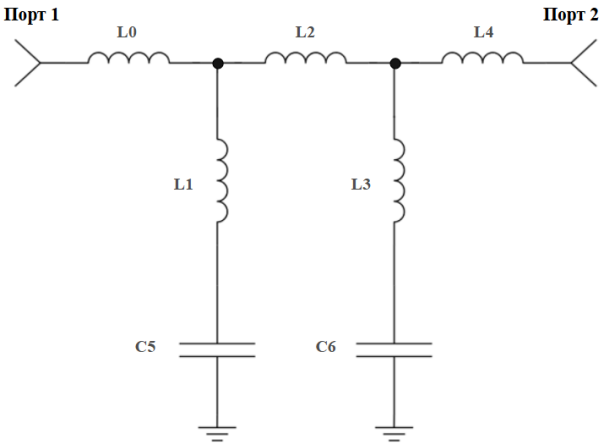


Рис. 1. Электрическая схема моделируемого 2-портового низкочастотного фильтра Кауэра-Чебышева 5-го порядка

Ранее авторами [1] был рассмотрен подход к диагностике неисправностей на основе S-параметров с 15 классами, и были достигнуты показатели точности в 85-87%.

2. Корни комплексного коэффициента передачи

Комплексный коэффициент передачи H (передаточная функция) является одной из основных характеристик радиоэлектронного устройства и зависит от номиналов его компонентов, топологии схемы, условий согласования источника и нагрузки. По комплексному коэффициенту передачи можно судить о корректности работы изделия в частотной области: насколько устройство усиливает или ослабляет сигнал и какие фазовые сдвиги вносит. Данная характеристика представляет собой комплексную величину, равную отношению комплексной амплитуды выходного сигнала к комплексной амплитуде входного сигнала и выражается по напряжению или току:

$$H(jw) = \frac{V_2}{V_1}$$

где V_2 и V_1 – комплексные амплитуды выходного и входного напряжения, w – угловая частота гармонического сигнала.

При представлении комплексного коэффициента передачи через дробно-рациональную функцию, корни ее полиномов в числителе и знаменателе, полюса и нули, могут использоваться в качестве входных данных для нейронной сети, так как они определяют частотное поведение

соответствующей линейной стационарной системы и как следствие форму S-параметров.

Устройства типа фильтров нижних частот могут не иметь нулей в конечной области. Общее число полюсов с учётом кратности определяется порядком цепи, равным количеству независимых реактивных элементов (индуктивностей и ёмкостей) в схеме.

Для линейного двухпортового устройства связь напряжения V_1 и тока I_1 на входном порте с напряжением V_2 и током I_2 на выходном порте выражается через ABCD-матрицу [5]:

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2 \\ -I_2 \end{bmatrix}.$$

ABCD-матрица содержит полную информацию о двухпортовой цепи. Из нее можно получить передаточную функцию по напряжению при условии, что выходной порт нагружен на сопротивление Z_0. Из определения ABCD-матрицы (2):

$$V_1 = AV_2 + B(-I_2).$$

При нагрузке на выходе Z_0 с учетом направления тока можно выразить I_2 (4) и, подставив это соотношение в (3), выразить комплексный коэффициент передачи по напряжению (5):

$$-I_2 = \frac{V_2}{Z_0},$$

$$H(jw) = \frac{1}{\left(A + \frac{B}{Z_0}\right)}.$$

Также используем однозначное преобразование параметров ABCD-матрицы в S-параметры и выразим из него A и B, выведя финальную формулу для $H(jw)$ на основе S-параметров [5]:

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} \dfrac{A + \dfrac{B}{Z_0} - CZ_0 - D}{A + \dfrac{B}{Z_0} + CZ_0 + D} & \dfrac{2(AD - BC)}{A + \dfrac{B}{Z_0} + CZ_0 + D} \\ \dfrac{2}{A + \dfrac{B}{Z_0} + CZ_0 + D} & \dfrac{-A + \dfrac{B}{Z_0} - CZ_0 + D}{A + \dfrac{B}{Z_0} + CZ_0 + D} \end{bmatrix},$$

$$A = \frac{(1 + S_{11})(1 - S_{22}) + S_{12}S_{21}}{2S_{21}},$$

$$B = Z_0 \frac{(1 + S_{11})(1 + S_{22}) - S_{12}S_{21}}{2S_{21}},$$

$$H(jw) = \frac{S_{21}}{1 + S_{11}}.$$

Имея модель тестируемого устройства, через его S-параметры `S_train` можно определить комплексный коэффициент передачи H по формуле (9). При этом S-параметры нужно представить в формате комплексных чисел, рассчитанных на основе модуля, переведенного в линейный формат, так как обычно он выражен в дБ, и фазы в радианах.

Далее для аппроксимации комплексного коэффициента передачи подбирался полином с помощью библиотеки NumPy [6] и ее функции `polyfit()`. Из диапазона значений от 5 до 17, выбранного экспериментальным путем, опираясь на порядок тестируемого фильтра, в качестве степени полинома было выбрано значение 17. У полинома подобранной степени сохранялись корни для получения новых входных данных `poly_train` в виде вектора признаков длиной 17 `(17, )`.

3. Архитектура нейронной сети и ее оценка

Для создания модели нейронной сети применяется tensorflow.keras [7]. Так как TensorFlow не поддерживает использование комплексных чисел в качестве основного типа данных для обучаемых параметров нейронной сети, нужно преобразовать данные. Например, отбросить мнимую часть или разделить мнимую и действительную части, при этом увеличив вдвое длину входных данных. В данной работе для дальнейшего обучения берется модуль комплексного числа.

На рис. 2 показана, использованная для классификации архитектура с оптимизатором Adam, функцией потерь

sparse_categorical_crossentropy и размером пакета 64 для обучения.

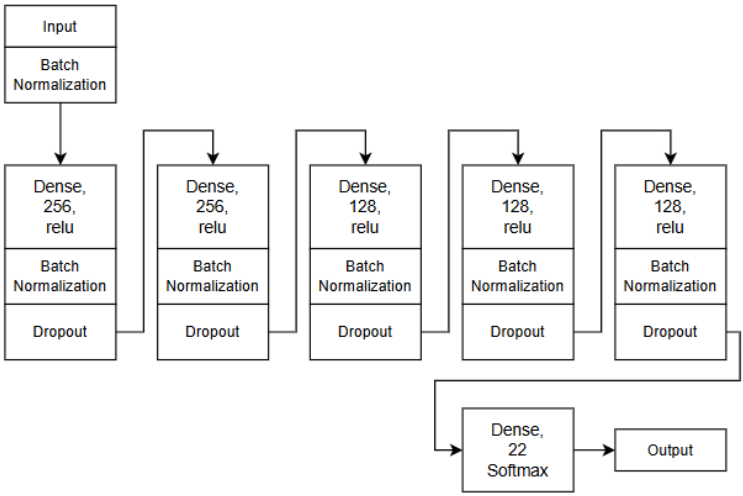


Рис. 2. Предлагаемая архитектура с использованием слоев из `tensorflow.keras.layers`

Архитектура рассчитана на входной вектор признаков, фиксированной длины 17, после которого добавлен слой `BatchNormalization`, выравнивающий распределение значений и ускоряющий процесс обучения. Далее применяется последовательность пяти блоков, состоящих из полносвязного `Dense` слоя, слоя нормализации `BatchNormalization` и слоя `Dropout`. `Dropout` со значением 0.1 добавлен после каждого из пяти скрытых полносвязных слоев и слоев нормализации для повышения устойчивости модели.

Первые два блока содержат по 256 нейронов в полносвязных слоях, позволяющих строить нелинейные комбинации входных признаков. Следующие три блока сужают признаковое пространство, уменьшая число нейронов в `Dense` слоях до 128. `ReLU` выбрана как стандартная функция активации для MLP (multilayer perceptron). На выходе последнего полносвязного `Dense` слоя, использующего функцию активации `Softmax`, формируются значения вероятностей принадлежности к каждому из 22 классов возможных состояний тестируемого устройства. Общее число параметров в данной архитектуре составило 142 810 в сравнении с предыдущей работой авторов (177 367 параметров).

Обучение осуществлялось на протяжении 300 эпох со снижением начальной скорости обучения 0,01, используя класс `ReduceLROnPlateau` из модуля `tensorflow.keras.callbacks`. При сравнительно небольшой размерности входных данных `(17, )` лучшая достигнутая точность на валидационной выборке составила 77.8% и на тестовой 76.1%. Матрица ошибок представлена на рис.2.

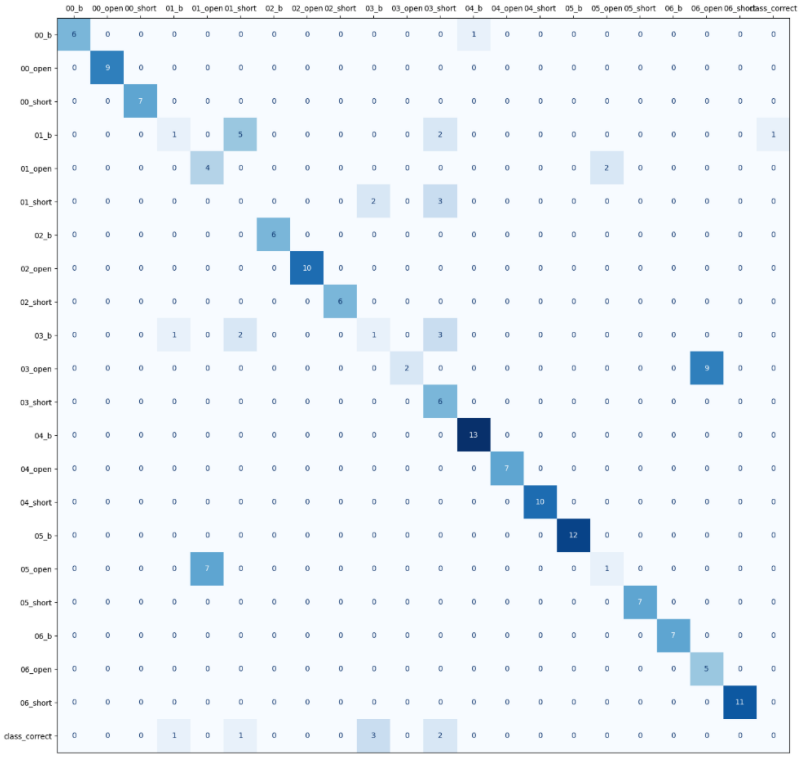


Рис. 3. Матрица ошибок: сверху предсказанные классы, слева действительные.

По ней можно увидеть, что, не учитывая ошибок для новых классов, сохранилось перепутывание пар классов `01_open` и `05_open`, `03_open` и `06_open`, как и в предыдущей работе [1]. Но также добавилась ошибка в классификации между классами `01_short` и `03_short`, `01_short` и `class_correct`, `03_short` и `class_correct`. Так как у этих классов формы графиков S-параметров схожи, то рассчитанные корни полиномов,

аппроксимирующих передаточную функцию, также достаточно близки в сравнении с корнями, рассчитанными для других классов.

В сравнении с предыдущей работой авторов с использованием S-параметров, как входных данных [1], точность нейронной сети оказалась на 10% ниже (в сравнении с 85-87%), но благодаря сниженной размерности входных данных ее архитектура при этом значительно проще. Также в этой работе было увеличено количество классов неисправностей, что могло снизить точность.

Заключение

В данном исследовании рассмотрен подход с использованием корней комплексного коэффициента передачи в качестве входных данных для нейронной сети при классификации неисправностей модели 2-портового низкочастотного фильтра Кауэра-Чебышева 5-го порядка. Из передаточной функции были выделены и преобразованы комплексные корни, при этом сокращена размерность входных данных. Обучена нейронная сеть с упрощенной архитектурой, уменьшившей общее число параметров. Достигнута хорошая точность распознавания неисправностей с учетом добавления дополнительных 7 классов. В продолжение имеет смысл рассмотреть входную размерность (17, 2), при которой каждый корень комплексного коэффициента передачи разделяется на мнимую и действительную части.

Дальнейшее развитие этого направления позволит создать технологию проверки и отладки изделий, предполагающую либо полное отсутствие человека, либо его присутствие для выполнений функций контроля. Применение такой технологии позволит сократить издержки и время производства при крупносерийном выпуске оборудования.

Литература

1. Перова, С. В. Классификация неисправностей пассивных радиочастотных устройств по S-параметрам при помощи нейронных сетей / С. В. Перова, А. В. Акимов, А. В. Гречишкин // Труды молодых ученых факультета компьютерных наук ВГУ : Сборник научных трудов. – Воронеж : Воронежский государственный университет, 2025. – С. 567-575.

2. Wang, H. Analog Circuits Fault Diagnosis Based on Machine Learning / H. Wang, C. Cai, X. Chen, F. Huo // 2024 IEEE 37th International System-on-Chip Conference (SOCC). – 2024. – P. 1-6. – doi: 10.1109/SOCC62300.2024.10737721.

3. Kang, T. Y. An Empirical Study on Fault Detection and Root Cause Analysis of Indium Tin Oxide Electrodes by Processing S-Parameter Patterns / T. Y. Kang, H. Lee, S. Suh // IEEE Transactions on Device and Materials

Reliability. – 2024. – Vol. 24, № 3. – P. 380-389. – doi: 10.1109/TDMR.2024.3415049.

4. Kang, T. Y. SREL: Severity Rating Ensemble Learning for Non-Destructive Fault Diagnosis of Cu Interconnects using S-parameter Patterns / T. Y. Kang, H. Lee, S. Suh // arXiv. – 2023. – doi: 10.48550/arXiv.2304.10207.

5. Pozar, D. M. Microwave engineering / D. M. Pozar. – 4th ed. – Hoboken : John Wiley & Sons, 2011. – 736 p. – ISBN 978-0-470-63155-3.

6. NumPy documentation [Электронный ресурс] // NumPy. – Режим доступа: https://numpy.org/doc/stable/

7. Module: tf.keras [Электронный ресурс] // TensorFlow. – Режим доступа: https://www.tensorflow.org/api_docs/python/tf/keras

Разработка системы анализа уязвимостей на основе баз знаний MITRE ATT&CK и MITRE D3FEND с использованием RAG технологий

Н. А. Редин, email: redinnikita47@gmail.com
В. В. Гаршина

Аннотация. *В работе проведён обзор существующего подхода к автоматизации анализа уязвимостей, а также разработана система, основанная на использовании баз знаний MITRE ATT&CK и MITRE D3FEND с применением технологий Retrieval-Augmented Generation (RAG).*

Ключевые слова: *MITRE ATT&CK, уязвимости, MITRE D3FEND.*

Введение

Количество уязвимостей растёт каждый год. Многие из них связаны не только с ошибками в коде, но и с особенностями конфигурации сервисов. При этом злоумышленники активно используют автоматизацию и публичные базы данных. В таких условиях традиционный ручной анализ уже не обеспечивает достаточной эффективности.

Для сопоставления уязвимостей с действиями атакующих используется база знаний MITRE ATT&CK [1], позволяющая описать техники и тактики злоумышленников, а MITRE D3FEND [2] — соответствующие защитные меры. Совместное применение этих баз позволяет не только классифицировать угрозы, но и определить способы их снижения. Из-за большого объёма данных ручная обработка становится трудоёмкой, что повышает актуальность подходов на основе Retrieval-Augmented Generation (RAG), обеспечивающих автоматический поиск и формирование аналитических выводов.

1. Обзор предшествующих работ

Системы анализа уязвимостей и сопоставления их с моделью поведения злоумышленника рассматриваются как в научных, так и практических работах. Данные работы направленны на интеграцию с такими данными как CPE, CVE [3], CWE, а также базой знаний MITRE ATT&CK.

В качестве примера решений является проект CVE2CAPEC [4] от организации Galeax. В основе проекта лежит многоступенчатая модель сопоставления, в рамках которой на первом этапе уязвимости CVE

© Редин Н. А., Гаршина В. В., 2026

связываются с соответствующими классами ошибок CWE, после чего выполняется сопоставление с шаблонами атак CAPEC. Далее CAPEC-паттерны проецируются на техники MITRE ATT&CK, а на заключительном этапе определяются соответствующие защитные меры и контрмеры из MITRE D3FEND.

Архитектурно проект реализован в виде набора Python-скриптов, обеспечивающих загрузку и актуализацию справочников CWE, CAPEC, ATT&CK, D3FEND с использованием механизмов CI/CD GitHub Actions, а также формирование результирующих связок. В качестве демонстрации работы системы и представления результатов используется графический пользовательский интерфейс, в рамках которого отображаются выявленные связи (Рис. 1), а также применяется инструмент MITRE ATT&CK Navigator, позволяющий визуализировать соответствующие техники в виде матрицы ATT&CK (Рис. 2).

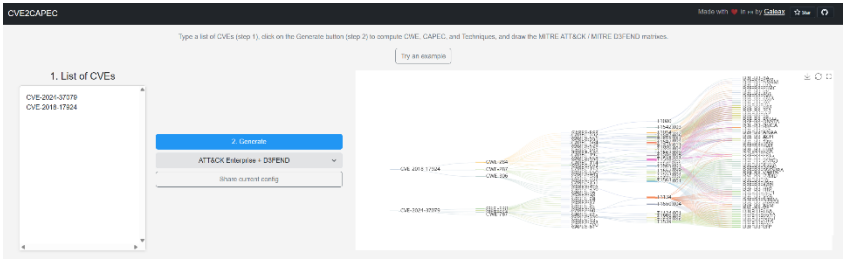


Рис. 1. Графическое представление результатов сопоставления уязвимостей CVE с техниками MITRE ATT&CK в CVE2CAPEC



Рис. 2. Графическое представление результатов с инструмент MITRE ATT&CK Navigator в CVE2CAPEC

2. Предлагаемая методика анализа уязвимостей

Предлагаемая методика анализа уязвимостей (Рис. 3) основывается на использовании внешних источников данных об уязвимостях, в частности базы CVE, а также баз знаний фреймворков MITRE ATT&CK и MITRE D3FEND. В качестве механизма поиска и сопоставления уязвимостей с соответствующими техниками атак и защитными мерами используется технология RAG (Retrieval-Augmented Generation).

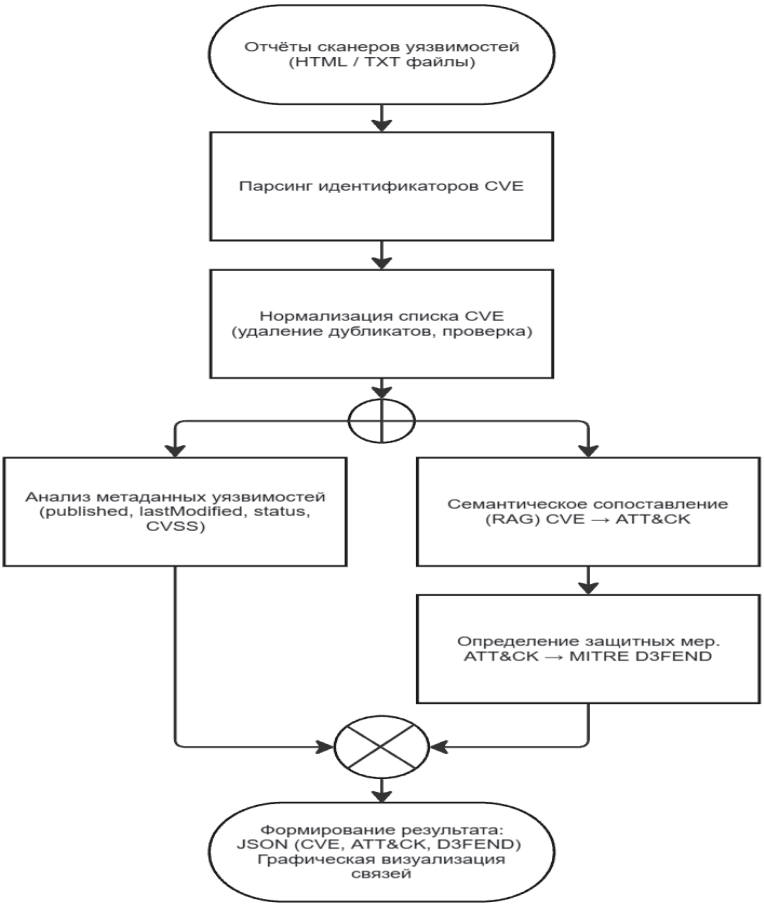


Рис. 3. Блок схема реализации методики анализа уязвимостей

В начале выполняется парсинг файлов, содержащих идентификаторы CVE, что позволяет использовать уже имеющиеся отчёты средств анализа безопасности и сканеров уязвимостей без необходимости их повторного формирования. Полученный перечень уязвимостей нормализуется и используется в качестве входных данных для последующих этапов анализа. На следующем этапе с применением технологии RAG осуществляется поиск наиболее релевантных техник MITRE ATT&CK, соответствующих рассматриваемым уязвимостям, на основе датасетов.

После установления соответствий между уязвимостями и техниками ATT&CK формируется связь между выявленными техниками и соответствующими защитными мерами из базы знаний MITRE D3FEND. Дополнительно в рамках методики выполняется анализ характеристик уязвимостей, включая дату публикации, дату последнего обновления, статус уязвимости и оценку критичности. Совокупность полученных результатов, включающая как выявленные связи, так и основные атрибуты уязвимостей, сохраняется в структурированном виде, что обеспечивает возможность их последующего анализа, сравнения и использования в экспериментальной части работы.

3. Реализация методики и обсуждение полученных результатов

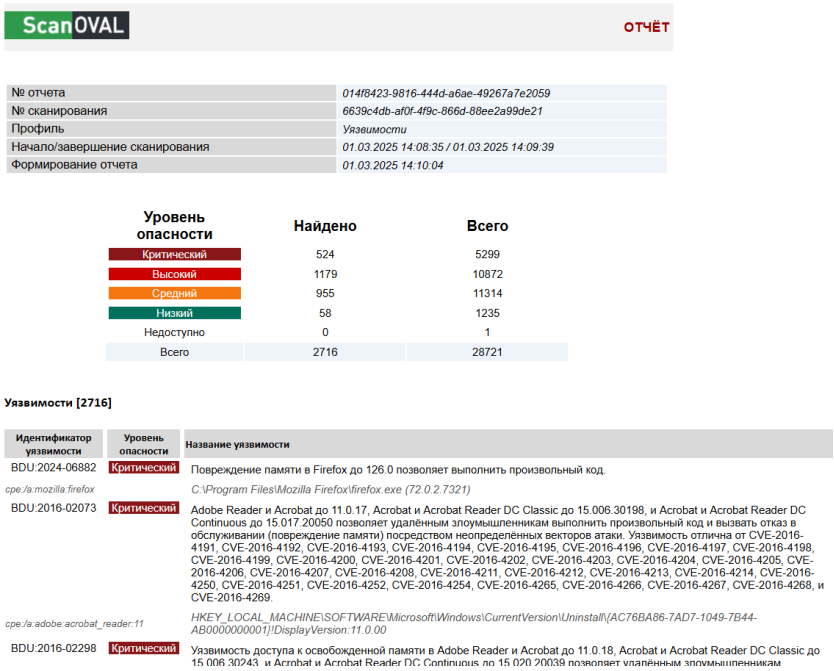
Реализация предложенной методики анализа уязвимостей была выполнена в виде программы на языке программирования Python с использованием сторонних библиотек. В качестве механизма Retrieval-Augmented Generation (RAG) применялась предобученная модель BAAI/bge-m3.

В качестве источника информации об уязвимостях использовался отчёт, полученный с помощью сканера уязвимостей ScanOVAL, представленный в формате HTML (Рис. 4). Из данного отчёта выполнялось извлечение идентификаторов CVE, после чего полученный перечень уязвимостей нормализовался и использовался для дальнейшего анализа.

Поиск техник MITRE ATT&CK для каждой уязвимости реализован как семантическое сопоставление текстов в векторном пространстве. В качестве базы знаний используется локальный набор описаний техник ATT&CK, где для каждой техники формируется объединённый текст (название, описание, тактики и платформы). Далее для этого корпуса заранее вычисляются эмбеддинги с применением модели BAAI/bge-m3. Полученные векторы кэшируются на диске, что позволяет не пересчитывать их при повторных запусках программы.

Для очередной уязвимости CVE извлекается текстовое описание из локальной копии NVD [3. 5], после чего формируется строка вида CVE

и описание. Данный текст преобразуется в эмбеддинг той же моделью BAAI/bge-m3. Далее рассчитывается косинусное сходство между вектором уязвимости и векторами всех техник ATT&CK. Полученные значения сходства используются для ранжирования техник, после чего выбираются TOP-K наиболее релевантных кандидатов.



ScanOVAL ОТЧЁТ

№ отчета	014f8423-9816-444d-a6ae-49267a7e2059
№ сканирования	6639c4db-af0f-4f9c-866d-88ee2a99de21
Профиль	Уязвимости
Начало/завершение сканирования	01.03.2025 14:08:35 / 01.03.2025 14:09:39
Формирование отчета	01.03.2025 14:10:04

Уровень опасности	Найдено	Всего
Критический	524	5299
Высокий	1179	10872
Средний	955	11314
Низкий	58	1235
Недоступно	0	1
Всего	2716	28721

Уязвимости [2716]

Идентификатор уязвимости	Уровень опасности	Название уязвимости
BDU:2024-06882	Критический	Повреждение памяти в Firefox до 126.0 позволяет выполнить произвольный код.
cpe:/a:mozilla:firefox		*C:\Program Files\Mozilla Firefox\firefox.exe (72.0.2.7321)*
BDU:2016-02073	Критический	Adobe Reader и Acrobat до 11.0.17, Acrobat и Acrobat Reader DC Classic до 15.006.30198, и Acrobat и Acrobat Reader DC Continuous до 15.017.20050 позволяет удалённым злоумышленникам выполнить произвольный код и вызвать отказ в обслуживании (повреждение памяти) посредством неопределённых векторов атаки. Уязвимость отлична от CVE-2016-4191, CVE-2016-4192, CVE-2016-4193, CVE-2016-4194, CVE-2016-4195, CVE-2016-4196, CVE-2016-4197, CVE-2016-4198, CVE-2016-4199, CVE-2016-4200, CVE-2016-4201, CVE-2016-4202, CVE-2016-4203, CVE-2016-4204, CVE-2016-4205, CVE-2016-4206, CVE-2016-4207, CVE-2016-4208, CVE-2016-4211, CVE-2016-4212, CVE-2016-4213, CVE-2016-4214, CVE-2016-4250, CVE-2016-4251, CVE-2016-4252, CVE-2016-4254, CVE-2016-4265, CVE-2016-4266, CVE-2016-4267, CVE-2016-4268, и CVE-2016-4269.
cpe:/a:adobe:acrobat_reader:11		*HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows\CurrentVersion\Uninstall\{AC76BA86-7AD7-1049-7B44-AB0000000001}!DisplayVersion:11.0.00*
BDU:2016-02298	Критический	Уязвимость доступа к освобожденной памяти в Adobe Reader и Acrobat до 11.0.18, Acrobat и Acrobat Reader DC Classic до 15.006.30243, и Acrobat и Acrobat Reader DC Continuous до 15.020.20039 позволяет удалённым злоумышленникам

Рис. 4. Отчет ScanOVAL, представленный в формате HTML

Затем применяется пороговая фильтрация по коэффициенту сходства: в итоговый список попадают только техники, значение сходства для которых не ниже заданного порога. Если после фильтрации список оказывается пустым, используется резервное правило: выбирается лучшая техника из TOP-K при условии, что её сходство превышает минимально допустимое значение. Такой подход позволяет обеспечить результат для каждой уязвимости даже при неполных или кратких описаниях.

Дополнительно предусмотрен механизм подмешивания техник из локального датасета [3. 4] (вспомогательный набор данных с заранее установленными соответствиями между CVE и техниками MITRE

ATT&CK) при наличии соответствий для конкретного CVE. В этом случае такие техники добавляются в итоговый набор с высоким приоритетом, после чего объединяются с результатами семантического поиска.

Полученные связи между уязвимостями, техниками атак и мерами защиты визуализировались в графическом виде, что позволило наглядно представить структуру выявленных зависимостей (Рис. 5).

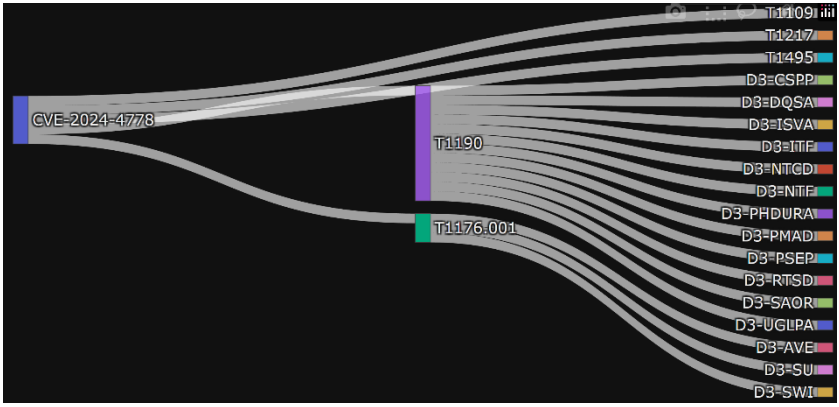


Рис. 5. Графическое представление связи между уязвимостями техниками атак и мерами защиты визуализировались

Программа формирует итоговый JSON-файл (Рис. 6), где каждая запись соответствует одной уязвимости CVE и содержит как её основные атрибуты, так и результаты сопоставления с базами знаний MITRE. Пример структуры записи приведён ниже.

Поле CVE содержит идентификатор уязвимости. Поля published и lastModified фиксируют дату публикации уязвимости и дату последнего обновления записи в базе NVD соответственно. Поле vulnStatus отражает текущий статус уязвимости в NVD. Поле cvss содержит значение базовой оценки критичности CVSS Base Score, используемое для приоритизации уязвимостей. Поле techniques представляет собой список идентификаторов техник MITRE ATT&CK, которые были определены как наиболее релевантные для данной уязвимости в результате семантического поиска. Наличие техник с подномерными идентификаторами означает, что сопоставление выполнялось на уровне под-техник. Поле defend содержит перечень идентификаторов защитных мер MITRE D3FEND в формате `D3-XXX`, которые связаны с найденными техниками ATT&CK.

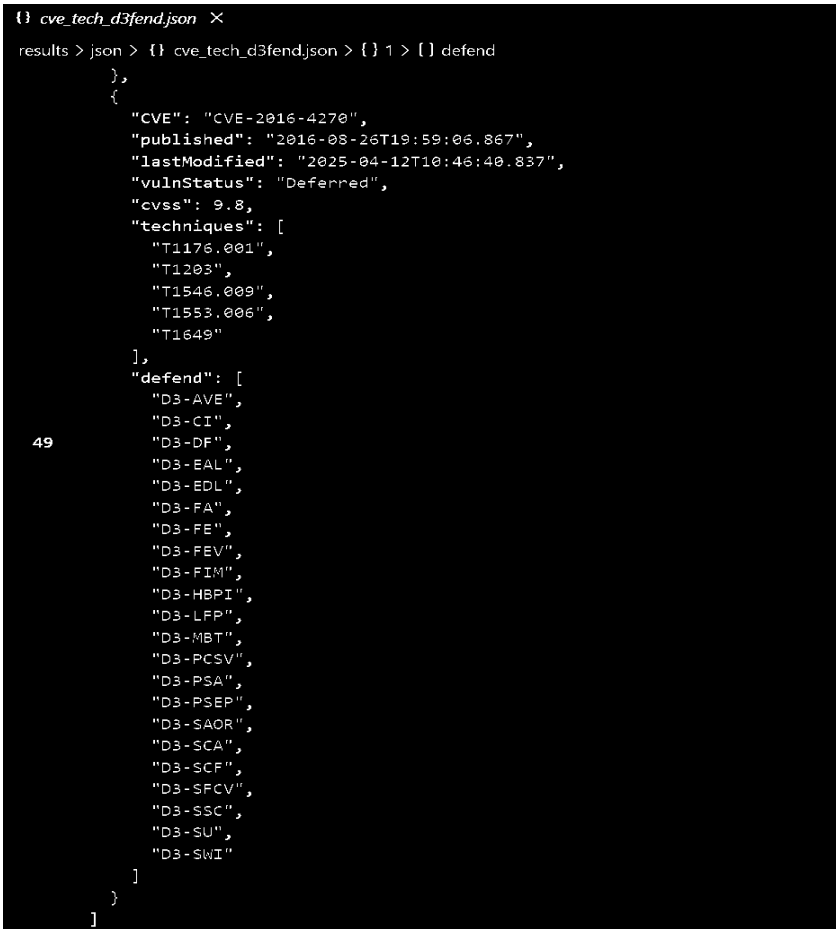


Рис. 6. Итоговый JSON-файл после выполнения программы

Такой формат представления данных обеспечивает возможность их последующего анализа, сравнения и использования в других проектов.

Заключение

В данной статье был выполнен обзор существующего программного средства анализа уязвимостей. На основе полученных результатов была предложена и реализована методика автоматизированного сопоставления уязвимостей с техниками атак и мерами защиты. Реализация методики

выполнена в виде программы на языке Python с использованием модели BAAI/bge-m3 для семантического поиска. Результатом работы был сформирован структурированный набор данных в формате JSON, а также выполнена графическая визуализация связей.

Список литературы

1. MITRE ATT&CK [Электронный ресурс]. – Режим доступа: https://attack.mitre.org/
2. MITRE D3FEND [Электронный ресурс]. – Режим доступа: https://d3fend.mitre.org/
3. CVE — Common Vulnerabilities and Exposures [Электронный ресурс]. – Режим доступа: https://cve.org/
4. CVE2CAPEC / GitHub [Электронный ресурс]. – Режим доступа: https://github.com/Galeax/CVE2CAPEC
5. National Vulnerability Database (NVD) [Электронный ресурс]. – Режим доступа: https://nvd.nist.gov/

К вопросу анализа современных механизмов внимания для задач восстановления изображений

Р. Е. Рогозин, email: ruslan-96-01-09@mail.ru

Российский технологический университет МИРЭА

Аннотация. *В работе рассматриваются механизмы внимания, применяемые в современных нейросетевых архитектурах для задач удаления шума и восстановления изображений. Проанализированы различные подходы к реализации механизмов пространственного и межканального внимания на основе сверток и механизмов самовнимания, используемых в сверточных сетях, трансформерах и их гибридных архитектурах.*

Ключевые слова: *восстановление изображений, механизмы внимания, сверточные нейросети, трансформеры.*

Введение

Задачи удаления шума и восстановления изображений относятся к числу фундаментальных и важных направлений компьютерного зрения. Наличие зашумленных и искаженных данных способно существенно препятствовать корректной работе более высокоуровневых задач таких как классификация, детекция объектов, семантическая сегментация.

В реальных данных (например, снимки с мобильной видеокамеры или медицинские данные), искажения шумы зачастую имеют различную природу и характеризуются сложной комбинированной структурой. Поэтому при решении практических задач используются упрощенные модели шумов, описываемые математическими моделями. Перечислим наиболее известные типы задач восстановления изображений, классифицируемые разделяются по типу искажения или шума [1, 2]:

- подавление аддитивного Гауссова шума;
- подавление шума Пуассона;
- подавление шума соль-перец;
- устранение размытия;
- удаление дымки/тумана;
- повышение разрешения изображений (Super Resolution/ SR).

Реальные шумы искажения имеются сложный комбинированный характер, однако такие упрощения позволяют генерировать реалистичные синтетические данные, а также используются в качестве аугментаций для обучения на размеченных реальных наборах данных. В

© Рогозин Р. Е., 2026

работе [3] отмечено, что сверточная сеть, обученная с использованием аддитивного Гауссова шума, довольно неплохо адаптировалась к реальным данным только за счет увеличения уровня шума.

Многие современные нейросетевые архитектуры способны удалять шумы искажения различной природы, что позволяет их рассматривать в качестве многозадачных. Эта универсальность обеспечивается как общими принципами и идеями, заложенными в архитектуру, так и широким использованием механизмов внимания.

Цель работы – это анализ механизмов внимания, применяемых в современных сверточных и трансформерных архитектурах для удаления шума и восстановления изображений.

1. Сверточные архитектуры

До широкого использования сверточных нейросетей для удаления шумов использовались традиционные методы обработки изображений и сигналов, например, пространственные фильтры (в качестве примера можно привести сглаживание по среднему или медианный фильтры), частотные или частотно-пространственные (использование преобразования Фурье или вейвлет-преобразования). Более подробный обзор традиционных методов удаления шумов содержится в [1, 2]. Основные недостатки традиционных методов заключаются в зависимости от входного шума и сложности адаптации к новым данных, а сами полученные результаты восстановления не всегда имеют высокий уровень детализации.

В настоящее время сверточные архитектуры нейросетей Convolutional neural networks (далее CNN-архитектуры) широко используются в различных задачах компьютерного зрения. Это связано с хорошей адаптации к новым данным из-за автоматического извлечения локальных признаков свертками, а также глубокой иерархической архитектуре, в которой начальные слои извлекают общие локальные признаки, а более глубокие – глобальные.

Из-за большого числа задач в области восстановления изображений, в настоящее время разработаны разнообразные CNN-архитектуры, более подробный обзор которых приведен в [2]. В настоящем обзоре рассмотрим базовые принципы построения архитектур и используемые механизмы внимания.

Большинство передовых архитектур построено на общих принципах: применение остаточных соединений или остаточных блоков; реализация последовательной многоуровневой обработки признаков с разным пространственным разрешением.

В архитектурах, построенных на принципе остаточных соединений или остаточных блоков, нейросеть обучается предсказывать не

восстановленное изображение, а поправку к входным данным. При этом в таких моделях пространственный размер входных данных не изменяется при обработке глубокими слоями. Это позволяет модели фокусироваться на более высокочастотных признаках, что может быть актуальным для задач SR [4]. Одной из классических таких архитектур является DnCNN [2], которая учится предсказывать остаток, вычитаемый из входного изображения.

Широко используются архитектуры, в основе которых лежит многоуровневая обработка признаков разных пространственных разрешений. Классическая схема построения – Unet-образная симметричная архитектура, схематично изображенная на рис. 1 [5].

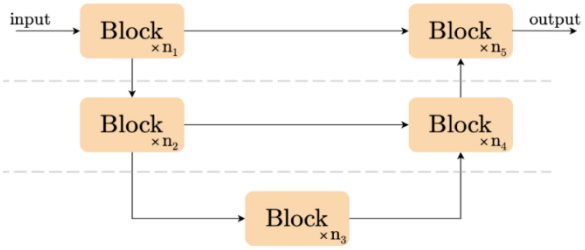


Рис. 1. Схематичное изображение Unet-архитектуры

Симметричная Unet-архитектура состоит из кодировщика, последовательно сжимающего входные данные, и декодировщика, последовательно увеличивающего разрешение. Для сохранения информации об извлеченных низкоуровневых признаках и стабилизации обучения, на каждом уровне признаки после блока кодировщика подаются на вход декодировщику через пропускные соединения (или skip-connections). Последовательное понижение/повышение пространственной размерности в сочетании с пропускными соединениями по своей сути позволяет удалять высокочастотный шум. Поэтому использование таких архитектур в области SR может быть затруднительно [4]. Примеры передовых сверточных архитектур: MPRNet [6] и HINet [7].

Сверточные слои формируют локальные признаки за счет обработки небольших фрагментов изображения. Некоторые из этих признаков являются наиболее значимыми и могут повысить качество обработки изображений. Для усиления их влияния в CNN-архитектурах используются механизмы пространственного и межканального внимания, которые адаптивно взвешивают входные признаки, тем самым

усиливая наиболее значимые области. В пространственном внимании формируется двухмерная маска значимых признаков за счет агрегации (например, усреднение) входных данных по каналам с последующей сверточной обработкой. Полученная карта признаков агрегируется с входными признаками (например, прибавляется). Аналогично устроено и межканальное внимание: формируется вектор признаков за счет агрегации по пространственным размерностям с последующей обработкой полносвязной сетью. Важно отметить, что пространственное внимание позволяет фокусироваться на мелких деталях и признаках, тогда как межканальное внимание учитывает глобальные признаки, распределенные в каналах изображения. Примером совместного (последовательного) использования обоих типов внимания является архитектура FFA-Net для удаления дымки с изображений [8].

Основные недостатки CNN-архитектур связаны с локальным извлечением признаков свертками с фильтрами фиксированных размеров. Это не позволяет сетям полноценно учитывать глобальную информацию, а также затрудняет их адаптацию к новым данным.

Для устранения указанных недостатков используются более сложные механизмы самовнимания или представление признаков в пространственной или пространственно-частотной области путем применения преобразования Фурье или вейвлет-преобразования. Такие преобразования позволяют локализовать информативную часть входных данных от шума. В работе [9] описана Unet-подобная архитектура, где в механизмах межканального внимания используется преобразование Фурье, а в самой архитектуре используются блоки с пространственно-частотным разделением признаков на основе двумерного вейвлет-преобразования Хаара [9].

2. Трансформеры

Изначально трансформеры разрабатывались для сферы обработки естественных языков (NLP), в которой крайне важно связывать удаленные слова или фразы. Это стало возможным за счет использования механизма самовнимания, который в общем виде записывается как [10]:

$$Attention(Q,K,V) = SoftMax(QK^T / \sqrt{d})V \qquad (1)$$

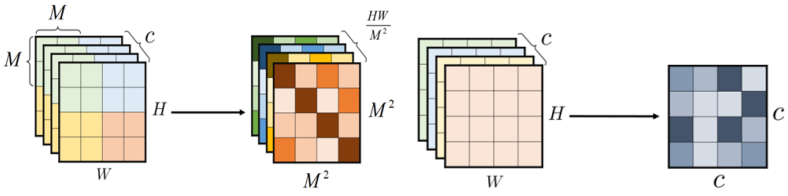
где – Q, K, V матрицы запросов, ключей и значений, получаемых при помощи линейной проекции входных данных, d – размерность признаков, по которым считается скалярное произведение.

Механизмы самовнимания позволили трансформерам совершить значительный прорыв в сфере NLP. Поскольку механизм самовнимания предназначен для обработки последовательностей, в одной из первых

архитектур для обработки изображений – ViT (Vision Transformer) – использовались схожие принципы построения с архитектурами языковым моделей. Для этого изображение разбивается на непересекающиеся блоки (окна), которые затем при помощи линейных проекций преобразуются в вектора признаков. Однако такой подход к обработке изображений вычислительно сложен, а также слабо учитывает локальные признаки из-за довольно крупной сетки разбиения – обычно изображение разбивают на блоки размера 16×16 пикселей [10]. Поэтому в современных архитектурах используется оконное и пространственное самовнимание.

При оконном самовнимании (multi-head self-attention или сокращенно W-MSA) входное изображение разбивается на $M \times M$ непересекающихся локальных окон, в каждом из которых W-MSA вычисляется согласно формуле (1) с добавлением обучаемого параметра к результату умножения матриц Q и K для учета пространственной взаимосвязи пикселей в окне. Следует отметить, что при произведении матриц Q и K формируется трехмерный тензор признаков для входного изображения. Сами же матрицы Q, K и V формируются за счет линейной проекции входных данных. Схематично оконное внимание показано на рис. 2(а) [4]. Также для учета взаимосвязей окон после блока с W-MSA зачастую используется блок с механизмом внимания shifted window attention (SW-MSA). Отличие SW-MSA заключается в использовании циклически-сдвинутого изображения (обычно на половину размера окна) и применении маски, которая убирает взаимосвязь между несоседними окнами. Одним классических примеров использования оконного механизма самовнимания является SwinIR-трансформер, в основе которого лежат остаточные связи в сочетании с самовниманием[11].

В межканальном самовнимании матрицы Q, K и V (формула (1)) формируются за счет последовательного применения поточечной и depth-wise свертки размера 3×3 к входному изображению. Такое последовательное применение верток позволяет лучше извлечь признаки из входных данных. Сам же механизм межканального самовнимания вычисляется согласно (1), но с перемножением матриц Q, K по пространственной размерности – при вычислении самовнимания изменен порядок перемножения матриц. Это позволяет сформировать для входного изображения межканальную матрицу ковариации (рис. 4), в которой отражены взаимосвязи между каналами, и уменьшить вычислительные затраты по сравнению с оконным сниманием. Схематично оконное внимание показано на рис. 2(б).



а б

а – оконное самовнимание, б –межканальное самовнимание

Рис. 2. Схематичное изображение механизмов самовнимания

Важно отметить, что межканальное самовнимание позволяет учесть глобальный контекст при формировании локальных признаков. Поэтому одна из современных Unet-подобных архитектур Restomer, в основе которой лежат блоки с межканальным самовниманием, позволяет удалять шумы различного характера [10]. В другой легковесной архитектуре NAFNet [5], созданной на базе Restomer, для извлечения пространсвенных признаков используются канальные и depth-wise свертки, а вместо межканального самовнимания используются простое механизм канального перевзвешивания.

Основные недостатки оконного самовнимания связаны с тем, что W-MSA не учитывает широкий контекст, а сдвиг окон в SW-MSA не позволяет адаптироваться к данным.

Для устранения недостаток SW-MSA в HAT-трансформере [11] предложено при вычислении самовнимания согласно формуле (1) использовать разные пространственные области: матрица Q вычисляется аналогично S-MSA, а для K и V изображение разбивается на тоже число, но уже перекрывающихся блоков. Концептуально похожий подход предложен в DSwinIR-архитектуре [12], где для матриц K и V пространственная область изображения получается за счет вычисления адаптивных сдвигов для центральной точки окна разбиения. Это позволяет модели адаптироваться к наиболее важным пространственным областям.

Комбинирование описанных механизмов самовнимания позволяет в архитектурах учитывать, как глобальный контекст (межканальное), так и локальные признаки (оконное самовнимание). Примером такой архитектуры является X-Restomer, созданный на базе Restomer: после блока с межканальным вниманием используется модифицированный блок пространственного внимания из HAT-

трансформера [4]. Это позволяет X-Restomer быть одной из передовых архитектур для основных задач удаления шумов и искажений.

Заключение

Рассмотрены различные механизмы внимания, используемые в современных нейросетевых архитектурах для задач удаления шума и восстановления изображений.

Направления дальнейших исследований могут быть связаны:

– с исследованием возможности локализации информативных признаков за счет преобразований признаков в частотную или частотно-пространственную область;

– улучшением механизмов оконного самовнимания за счет реализации адаптивных схем разбиения (например, применение деформируемых сверткок);

– проектированием универсальных многозадачных архитектур.

Список литературы

1. Abdrakhim D. Comprehensive review of image denoising methods: classical algorithms and deep learning approaches / D. Abdrakhim // Вестник науки. – 2024. – Т. 4, № 5. – С. 1246-1272.

2. Overview of Research on Digital Image Denoising Methods / J. Mao [и др.] // Sensors (Basel). – 2025. – Т. 25, № 8. – Article 2615.

3. Practical Blind Image Denoising via Swin-Conv-UNet and Data Synthesis / K. Zhang [и др.] // Machine Intelligence Research. – 2023. – Т. 20, № 6. – P. 822-836.

4. Chen X. Comparative Study of Image Restoration Networks for General Backbone Network Design / X. Chen [и др.] // Computer Vision – ECCV 2024. – 2024. – P. 74-91.

5. Chen, L. Simple Baselines for Image Restoration / L. Chen [и др.] // Computer Vision – ECCV 2022. – 2022. – P. 17-33.

6. Zamir, S. W. Multi-Stage Progressive Image Restoration / S. W. Zamir [и др.] // 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2021. – P. 14821-14831.

7. Chen, L. HINet: Half Instance Normalization Network for Image Restoration / L. Chen [и др.] // 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2021. – P. 182-192.

8. FFA-Net: Feature Fusion Attention Network for Single Image Dehazing [Электронный ресурс] / Q. Xu Z. [и др.] : Computing Research Repository (CoRR), 2019. – Режим доступа: https://arxiv.org/abs/1911.07559

9. Коваленко А.С. Архитектура U-Net для подавления шума на изображениях с механизмами частотного и частотно-временного

внимания / А.С. Коваленко // Вестник Адыгейского государственного университета. Серия: Естественно-математические и технические науки. – 2024. – Вып. 3 (346). – С. 51-62.

10. Vision Transformers in Image Restoration: A Survey // A.M. Ali [и др.] // Sensors. – 2023. – Т. 23, № 5. – Статья № 2385.

11. Image Restoration Using Swin Transformer [Электронный ресурс] / Q.J. Liang [и др.] : Computing Research Repository (CoRR), 2021. – Режим доступа: https://arxiv.org/pdf/2108.10257

12. DSwinIR: Rethinking Window-based Attention for Image Restoration [Электронный ресурс] / G. Wu [и др.] : Computing Research Repository (CoRR), 2025. – Режим доступа: https://arxiv.org/pdf/2504.04869.pdf

Анализ алгоритмов обучения с подкреплением для управления авиационными и беспилотными системами

Р. И. Родионов, email: rodionov_r_i@sc.vsu.ru
А. А. Сирота, email: sir@cs.vsu.ru

Воронежский государственный университет

Аннотация. *В работе рассматриваются современные алгоритмы обучения с подкреплением, применяемые к задачам управления летательными аппаратами различного класса – от беспилотных летательных аппаратов до пилотируемых воздушных судов. Основное внимание уделяется методам глубокого обучения с подкреплением, способным работать в непрерывных пространствах состояний и действий, характерных для задач авиационного управления.*

Ключевые слова: *обучение с подкреплением, интеллектуальные системы управление, летательные аппараты, беспилотные системы, непрерывное управление.*

Введение

Управление летательными аппаратами представляет собой сложную задачу, связанную с необходимостью принятия решений в условиях высокой размерности пространства состояний, нелинейной динамики и ограничений на управляющие воздействия. Классические методы автоматического управления, основанные на линейных регуляторах и заранее заданных законах управления, показали высокую надёжность в штатных режимах, однако обладают ограниченной адаптивностью к изменяющимся условиям эксплуатации. В частности, такие системы плохо масштабируются на сложные сценарии, включающие нестандартные аэродинамические возмущения, изменение характеристик летательного аппарата или необходимость выполнения сложных манёвров без заранее определённой траектории.

В последние годы методы обучения с подкреплением (Reinforcement Learning, RL) рассматриваются как альтернативный подход к построению систем управления, способных самостоятельно вырабатывать стратегию поведения на основе опыта взаимодействия с динамической системой. В отличие от традиционных регуляторов, RL-алгоритмы не требуют явного задания закона управления и способны оптимизировать поведение напрямую с использованием заданного критерию качества. Это делает их особенно привлекательными для задач управления авиационными и

© Родионов Р. И., Сирота А. А., 2026

беспилотными системами, где точное аналитическое описание динамики либо затруднено, либо практически невозможно. [1]

1. Формализация задачи управления в терминах обучения с подкреплением

Задачи управления летательными аппаратами естественным образом формализуются в виде марковского процесса принятия решений [1]. В каждый момент времени агент наблюдает текущее состояние системы и формирует управляющее воздействие, влияющее на последующую динамику. Формально процесс описывается кортежем

$$M = \langle S, A, P, \gamma \rangle, \tag{1}$$

где S – пространство состояний, включающее параметры движения и ориентации летательного аппарата, A – пространство управляющих воздействий, P – вероятностная динамика переходов между состояниями, а $\gamma \in (0,1)$ – коэффициент дисконтирования, определяющий важность будущих состояний.

В рамках данной формализации агент реализует стратегию управления $\pi(a \mid s)$, задающую правило выбора управляющего воздействия $a \in A$ при наблюдаемом состоянии $s \in S$. Качество стратегии оценивается через ожидаемую суммарную дисконтированную полезность, определяемую функционалом

$$J(\pi) = \mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty} \gamma^t U(s_t, a_t)\right], \tag{2}$$

где $U(s_t, a_t)$ – функция полезности, характеризующая предпочтительность выбранного действия в текущем состоянии, а математическое ожидание берётся по траекториям, индуцированным стратегией π и динамикой системы.

Задача обучения с подкреплением формулируется как задача оптимального управления [1], заключающаяся в поиске стратегии π^*, максимизирующей данный функционал:

$$\pi^* = \arg\max_{\pi} J(\pi). \tag{3}$$

В авиационных задачах оптимальная стратегия управления должна обеспечивать устойчивость движения, плавность манёвров и способность компенсировать внешние возмущения, возникающие вследствие изменения аэродинамических условий или динамики полёта. Эти требования накладывают дополнительные ограничения на свойства искомой стратегии и оказывают существенное влияние на выбор алгоритмов обучения, применяемых для решения задачи оптимизации.

2. Алгоритмы обучения с подкреплением для управления летательными аппаратами

Методы Deterministic policy gradient. Одним из первых классов алгоритмов, получивших широкое распространение в задачах непрерывного управления, стали методы детерминированного градиента политики, в частности DDPG [2, 3]. В этих алгоритмах стратегия представляется детерминированной функцией

$$a = \mu_\theta(s), \tag{4}$$

где индекс θ – параметры нейросетевой модели.

Качество детерминированной стратегии оценивается через функционал полезности, определяемый ожидаемым значением функции ценности при следовании данной стратегии:

$$J(\theta) = \mathbb{E}_{s\sim\rho^\mu}\left[Q^\mu(s, \mu_\theta(s))\right], \tag{5}$$

где $Q^\mu(s,a)$ - функция ценности, характеризующая ожидаемое качество применения управляющего воздействия a в состоянии s при следовании стратегии μ, а ρ^μ – стационарное распределение состояний, индуцированное данной стратегией.

Обновление параметров стратегии осуществляется в направлении градиента указанного функционала. Согласно теореме детерминированного градиента политики [2, 3], градиент целевой функции имеет вид

$$\nabla_\theta J(\theta) = \mathbb{E}_{s\sim\rho^\mu}\left[\nabla_\theta \mu_\theta(s) \nabla_a Q^\mu(s,a)\big|_{a=\mu_\theta(s)}\right] \tag{6}$$

Данное выражение показывает, что оптимизация стратегии сводится к согласованному обновлению параметров управляющей функции и аппроксимации функции ценности, что позволяет эффективно работать в непрерывных пространствах действий.

Детерминированные методы обладают высокой эффективностью в задачах с непрерывными действиями, однако на практике они оказываются чувствительными к выбору гиперпараметров и склонны к нестабильному обучению. В авиационных задачах, где динамика сильно нелинейна и присутствуют внешние возмущения, это приводит к деградации качества управления и сложности масштабирования алгоритма. [3]

Методы Proximal Policy Optimization. Алгоритм Proximal Policy Optimization (PPO) представляет собой один из наиболее популярных on-policy методов обучения с подкреплением, ориентированных на устойчивость и надёжность обучения [4, 5]. В отличие от

детерминированных подходов, PPO использует стохастическую политику $\pi_\theta(a \mid s)$, что позволяет агенту исследовать пространство действий более эффективно.

Ключевой особенностью PPO является ограничение на величину обновления политики. Обновление параметров осуществляется путём максимизации следующего выражения [4]:

$$L^{\text{PPO}}(\theta) = \mathbb{E}_t\left[\min\left(r_t(\theta)\hat{A}_t, \text{clip}\left(r_t(\theta), 1-\varepsilon, 1+\varepsilon\right)\hat{A}_t\right)\right], \quad (7)$$

где $r_t(\theta) = \pi_\theta(a_t|\ s_t)/\pi_{\theta_{\text{old}}}(a_t|\ s_t)$ – отношение вероятностей действий при новой и старой политике, $\hat{A}_t$ – оценка преимущества действия в состоянии, ε - параметр, ограничивающий величину обновления политики.

Такое ограничение предотвращает резкие изменения стратегии и обеспечивает более стабильную сходимость. Для задач управления летательными аппаратами это особенно важно, поскольку резкие изменения управляющих воздействий могут приводить к неустойчивым режимам полёта. Однако on-policy природа PPO приводит к высокой вычислительной стоимости обучения, так как данные не могут быть повторно использованы после обновления политики. [4, 5]

Алгоритм Soft Actor-Critic. Алгоритм Soft Actor-Critic (SAC) относится к классу off-policy методов обучения с подкреплением и сочетает стохастическую стратегию управления с принципом максимизации энтропии [6, 7]. В отличие от методов, ориентированных на жёсткую оптимизацию детерминированной политики, SAC формирует стратегию, которая сохраняет контролируемую степень стохастичности даже на поздних этапах обучения. Это позволяет алгоритму поддерживать активное исследование пространства действий и избегать преждевременной фиксации на локально оптимальных, но потенциально неустойчивых режимах управления.

Качество стохастической стратегии в рамках Soft Actor-Critic оценивается через энтропийно-регуляризованный функционал полезности, имеющий вид [6]

$$J(\pi) = \mathbb{E}_\pi\left[\sum_{t=0}^{\infty}\gamma^t\left(U(s_t, a_t) + \alpha\mathcal{H}\left(\pi(\cdot|\ s_t)\right)\right)\right], \quad (8)$$

где $\pi(a \mid s)$ – стохастическая стратегия управления, $\mathcal{H}\left(\pi(\cdot|\ s_t)\right)$ - энтропия распределения действий в состоянии, $\alpha > 0$ - коэффициент,

определяющий вклад энтропийного слагаемого, $\gamma \in (0,1)$ - коэффициент дисконтирования.

Энтропийный член в функционале служит формальным механизмом балансировки между эксплуатацией текущей стратегии и исследованием альтернативных управляющих воздействий. Максимизация энтропии приводит к тому, что стратегия остаётся более гибкой и устойчивой к изменениям динамики системы и внешним возмущениям.

Оптимальная стратегия в рамках Soft Actor-Critic определяется как решение задачи

$$\pi^* = \arg\max_{\pi} J(\pi), \tag{9}$$

что соответствует поиску такой стохастической политики, которая обеспечивает наилучшее сочетание качества управления и адаптивности.

Практика применения SAC в задачах управления летательными аппаратами показывает, что включение энтропийного критерия позволяет формировать более плавные и устойчивые стратегии управления по сравнению с детерминированными методами. Это особенно важно для авиационных систем, функционирующих в условиях неопределённости, где избыточно жёсткая стратегия может приводить к ухудшению качества управления при изменении внешних условий. [6, 7]

3. Сравнительный анализ и обсуждение

Анализ сравнительных характеристик алгоритмов обучения с подкреплением (табл. 1) показывает, что выбор метода оказывает существенное влияние на эффективность интеллектуального управления летательными аппаратами [1-7]. Алгоритмы класса REINFORCE демонстрируют низкую устойчивость и требуют значительного количества взаимодействий со средой, что ограничивает их практическое применение. Детерминированные методы, такие как DDPG, обладают более высокой выборочной эффективностью, однако чувствительны к настройке гиперпараметров и склонны к нестабильности обучения в условиях сложной динамики. PPO обеспечивает устойчивое и предсказуемое обучение, но требует значительных вычислительных ресурсов из-за on-policy характера. Алгоритм Soft Actor-Critic сочетает высокую устойчивость, адаптивность и эффективное использование данных, что делает его наиболее перспективным для задач управления авиационными и беспилотными системами. [8, 9]

Таблица 1

Сравнение алгоритмов обучения с подкреплением для задач управления летательными аппаратами

Алгоритм	Тип политики	Тип обучения	Сходимость (эпизоды)	Устойчивость обучения	Эффективность данных	Применимость для авиации
REINFORCE	Стохаст.	on-policy	$>10^5$	низкая	низкая	ограниченная
DDPG	Детерм.	off-policy	$(2\text{-}5)\times10^4$	низкая-средняя	высокая	ограниченная
PPO	Стохаст.	on-policy	$(3\text{-}6)\times10^4$	высокая	средняя	высокая
SAC	Стохаст.	off-policy	$(1\text{-}3)\times10^4$	высокая	высокая	очень высокая

Отдельно следует отметить универсальность рассмотренных методов. Несмотря на различия между пилотируемыми самолётами и беспилотными летательными аппаратами, математическая структура задачи управления остаётся схожей: высокоразмерное состояние, непрерывное пространство действий и необходимость устойчивого поведения. Это позволяет применять единый класс алгоритмов обучения с подкреплением к различным типам летательных аппаратов, адаптируя лишь представление состояния и ограничения на действия.

Заключение

В работе был выполнен анализ современных алгоритмов обучения с подкреплением, применяемых к задачам управления авиационными и беспилотными системами. Показано, что методы глубокого обучения с подкреплением обладают существенным потенциалом для построения интеллектуальных систем управления, превосходящих классические подходы по уровню адаптивности и универсальности. Наиболее перспективными для практического применения являются стохастические off-policy алгоритмы, такие как Soft Actor-Critic, обеспечивающие устойчивое обучение и высокое качество управления в условиях сложной динамики. Полученные выводы могут быть использованы при разработке интеллектуальных систем управления летательными аппаратами различного класса и служат основой для дальнейших исследований в данной области. [1-9]

Литература

1. Sutton, R. S. Reinforcement Learning: An Introduction / R. S. Sutton, A. G. Barto. – 2nd ed. – Cambridge: MIT Press, 2018. – 552 p.

2. Williams, R. J. Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning / R. J. Williams // Machine Learning. – 1992. – Vol. 8, No. 3-4. – P. 229-256.

3. Lillicrap, T. P. Continuous Control with Deep Reinforcement Learning / T. P. Lillicrap [et al.] // Proceedings of the International Conference on Learning Representations (ICLR). – 2016. – 14 p.

4. Schulman, J. Proximal Policy Optimization Algorithms / J. Schulman [et al.] // arXiv preprint. – 2017. – arXiv:1707.06347. – 12 p.

5. Schulman, J. Trust Region Policy Optimization / J. Schulman [et al.] // Proceedings of the 32nd International Conference on Machine Learning (ICML). – 2015. – P. 1889-1897.

6. Haarnoja, T. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor / T. Haarnoja [et al.] // Proceedings of the 35th International Conference on Machine Learning (ICML). – 2018. – P. 1861-1870.

7. Haarnoja, T. Soft Actor-Critic Algorithms and Applications / T. Haarnoja [et al.] // arXiv preprint. – 2019. – arXiv:1812.05905. – 21 p.

8. Liu, F. Learning to Have a Civil Aircraft Take Off under Crosswind Conditions by Reinforcement Learning with Multimodal Data and Preprocessing Data / F. Liu [et al.] // Aerospace. – 2021. – Vol. 8, No. 3. – P. 1-19.

9. Priya, S. Reinforcement Learning Based Autopilot Using X-Plane / S. Priya [et al.] // Engineering Design Document. – University of Southern California. – 2021. – 52 p.

Алгоритм мультиплексирования скрытых каналов при реализации стеганографического встраивания данных в JPEG изображения

А. Е. Руднев, email: rudnev@cs.vsu.ru
М. А. Дрюченко, email: m_dryuchenko@mail.ru

Воронежский государственный университет

Аннотация. *Рассматривается итеративный алгоритм для скрытного встраивания информации в JPEG изображения, основанный на внесении малозначительных искажений в блоки квантованных коэффициентов ДКП и использовании хеш-функций для последующего извлечения скрытых данных. Рассматривается вариант работы алгоритма в режиме мультиплексирования, позволяющий в рамках одного контейнера сформировать стеганографические каналы для нескольких абонентов или реализовать несколько каналов с различным уровнем секретности для одного абонента.*

Ключевые слова: *стеганография, мультиплексирование скрытых каналов, хэш-коды, цифровые изображения.*

Введение

С развитием технологий цифровых данных в последние десятилетия значительно возросла потребность в защите информации. Одним из методов обеспечения конфиденциальности является стеганография – наука и искусство скрытной передачи данных, при которой информация встраивается в носитель таким образом, что её существование остается незамеченным для постороннего наблюдателя. В отличие от криптографии, которая фокусируется на зашифровке информации, стеганография скрывает сам факт передачи данных, что делает её эффективным инструментом для защиты конфиденциальной информации.

Цифровые изображения, благодаря своей популярности и распространенности в интернете, являются одним из самых удобных и доступных носителей для стеганографического встраивания. Особое внимание уделяется алгоритмам, которые обеспечивают высокую степень скрытности данных и минимизируют видимые изменения в изображении [1-4]. JPEG – один из наиболее распространённых форматов сжатия изображений, который широко используется благодаря своей способности эффективно уменьшать размер файлов при сохранении

© Руднев А. Е., Дрюченко М. А., 2026

высокого качества изображения. Однако для успешного использования JPEG в качестве носителя стеганографической информации необходимо учитывать особенности сжатия и возможности скрытого встраивания данных.

Одним из ключевых аспектов улучшения стеганографических алгоритмов является использование методов мультиплексирования каналов. С помощью таких подходов можно значительно увеличить объем встраиваемой информации, используя части изображения несколько раз. Мультиплексирование также способствует повышению степени скрытности, так как позволяет распределить данные по различным частям изображения, что делает их менее уязвимыми для атак, направленных на обнаружение скрытой информации.

Целью данной работы является разработка (на основе ранее предложенного в [5]) стеганографического алгоритма для скрытия информации в JPEG-контейнеры, способного работать в режиме мультиплексирования скрытых каналов.

1. Алгоритм встраивания данных

Пусть I – это незаполненный JPEG-контейнер размером $W \times H$ пикселей, который будет использоваться для скрытия сообщения. Сообщение, которое необходимо скрыть, обозначим как M. Алгоритм встраивания данных E в общем случае реализует отображение следующего вида:

$$E : I \times M \times H \times K \to \tilde{I} \quad (1)$$

где H – хэш-функция, используемая для преобразования блоков квантованных коэффициентов ДКП, в которые реализуется скрытие, K – секретный стеганографический ключ, используемый для встраивания и извлечения данных, $\tilde{I}$ — заполненный контейнер.

Алгоритм встраивания должен одновременно удовлетворять двум противоречивым требованиям:

1. Максимизация пропускной способности (ПС) стегосистемы – необходимо скрыть как можно больше данных.

2. Минимизация визуальных искажений – изменения, внесенные в контейнер, должны быть как можно менее заметными. Для минимизации искажений необходимо встраивать информацию таким образом, чтобы она не влияла на визуальные качества изображения.

Основные этапы алгоритма встраивания информации отражены на рис. 1.

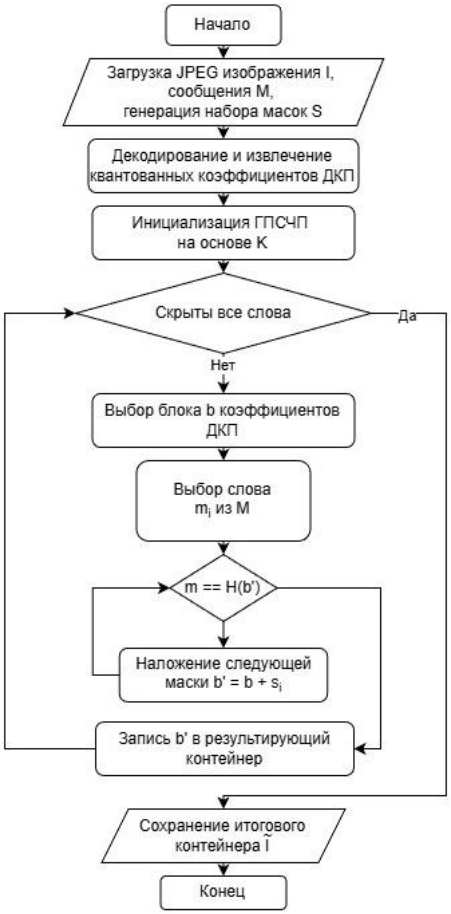


Рис. 1. Обобщенная схема работы алгоритма встраивания данных

На первом этапе работы алгоритма для скрываемого сообщения M необходимо загрузить исходное JPEG изображение – контейнер I подходящего размера. В общем случае скрываемое сообщение M рассматривается как последовательность 8-битных слов, каждое из которых встраивается в отдельный блок ДКП коэффициентов изображения. Для упрощения будем считать, что M представляет собой случайно сгенерированную последовательность.

Далее формируется набор масок, используемых для изменения значений коэффициентов ДКП в модифицируемом блоке изображения-контейнера. Модифицируемые коэффициенты располагаются выше побочной диагонали матрицы коэффициентов ДКП и соответствуют среднечастотной области спектра. Маски представляют собой верхнетреугольные матрицы размером 7×7, элементы которых задаются псевдослучайным образом. Значения элементов маски задаются в диапазоне от -1 до 1 и используются для внесения изменений в выбранные блоки коэффициентов ДКП. Маски позволяют контролировать степень изменений и влияние на визуальное качество изображения.

Далее происходит декодирование и расчет квантованных коэффициентов ДКП.

Следующим шагом является выбор порядка обхода блоков, который определяется с помощью генератора псевдослучайных числовых последовательностей (ГПСЧП).

Для встраивания данных используется 21 коэффициент ДКП, находящийся выше побочной диагонали. С этим набором в цикле поочередно складывается i-ая маска и вычисляется хэш-значение с помощью хэш-функции. Хэш-функция может быть как криптографической, так и некриптографической, а в частном случае можно использовать циклически избыточные коды.

Если хэш-значение совпадает с очередным встраиваемым словом m_i из M, то происходит запись модифицированных коэффициентов ДКП в результирующий контейнер. Этот процесс продолжается до тех пор, пока для каждого блока не будет найдено соответствие с очередным словом сообщения.

На последнем этапе формируется заполненный JPEG-контейнер $\tilde{I}$, который готов к хранению или передаче. Получившийся контейнер будет иметь малый уровень визуальных искажений, что позволяет обеспечить высокий уровень скрытности встраиваемой информации.

2. Алгоритм извлечения данных

Алгоритм извлечения информации обладает детерминированным характером и гарантирует точное извлечение скрытых данных. Алгоритм извлечениях данных $\tilde{E}$ имеет следующий вид:

$$\tilde{E} : \tilde{I} \times H \times K \to M \qquad (2)$$

Для корректного извлечения скрытых в $\tilde{I}$ данных необходимо знать ключ K, использованный при встраивании данных, поскольку он гарантирует правильность извлечения и порядок восстановления

скрытых данных, а также знать вид использовавшейся при скрытии функции хеширования H. Укрупненная схема работы алгоритма извлечения скрытой информации приведена на рис. 2.

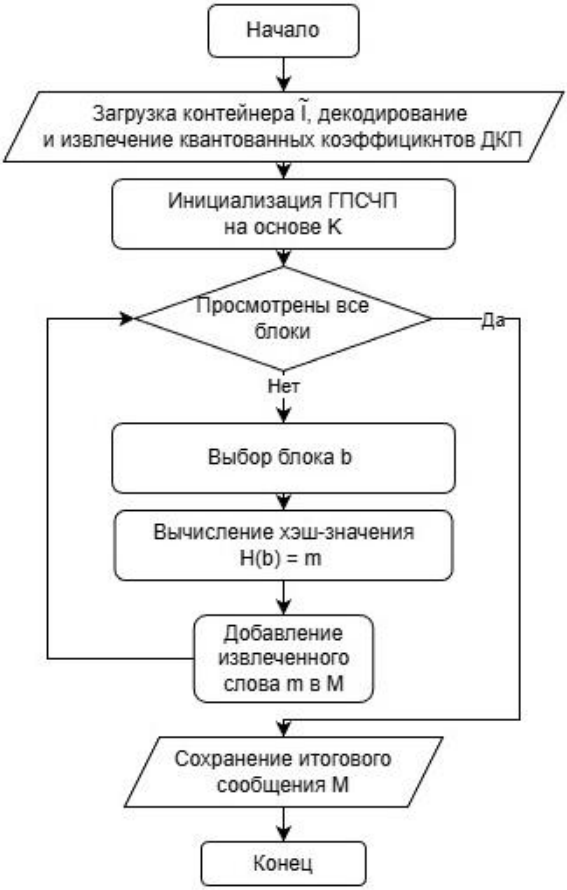


Рис. 2. Обобщенная схема работы алгоритма извлечения данных

Процесс извлечения начинается с загрузки заполненного JPEG-контейнера $\tilde{I}$, после чего происходит декодирование и извлечение коэффициентов ДКП.

Далее определяется порядок обхода блоков на основе ГПСЧП, инициализированного на основе ключа K.

Для каждого блока вычисляется хэш-значение, основанное на 21 среднечастотном коэффициенте, которое соответствует одному из встраиваемых слов сообщения. Процесс извлечения продолжается, пока все слова итогового сообщения не будут восстановлены.

Сохранение итогового сообщения.

3. Мультиплексирование каналов

Алгоритм мультиплексирования реализует встраивание скрытых сообщений в два канала. Такой подход позволяет эффективно передавать данные, скрывая их в разных каналах, что повышает скрытность и пропускную способность стеганографической системы. Реализация встраивания и извлечения данных совпадает с описанным ранее алгоритмом, но добавляется дополнительный уровень обработки, связанный с использованием двух каналов. Обобщенная схема алгоритма встраивания данных в режиме мультиплексирования каналов представлена на рис. 3.

В первом канале встраивается информация из сообщения M_1, а во втором канале используется вычисление хэш-значения от конкатенации коэффициентов двух случайных блоков, которые уже были использованы в процессе встраивания. Таким образом, данные из второго сообщения передаются через комбинированные блоки, что повышает сложность обнаружения скрытой информации.

Процесс встраивания начинается загрузки исходного JPEG изображения – контейнера I и с подготовки двух сообщений M_1 и M_2 для двух разных каналов. Также генерируется набор масок, которые будут использоваться для внесения изменений в коэффициенты блоков. Затем инициализируется ГПСЧП, чтобы определить порядок обхода блоков, которые будут подвергаться изменениям.

Далее происходит декодирование и расчет квантованных коэффициентов ДКП.

Далее выбираются два блока 8×8 коэффициентов ДКП. Для каждого блока вычисляются значения, полученные с использованием масок. Эти значения должны удовлетворять условию: хэш-значение от измененных коэффициентов первого блока должен быть равен m_i из M_1, а хэш-значение от измененных коэффициентов второго блока должен совпадать с m_{i+1} из M_1.

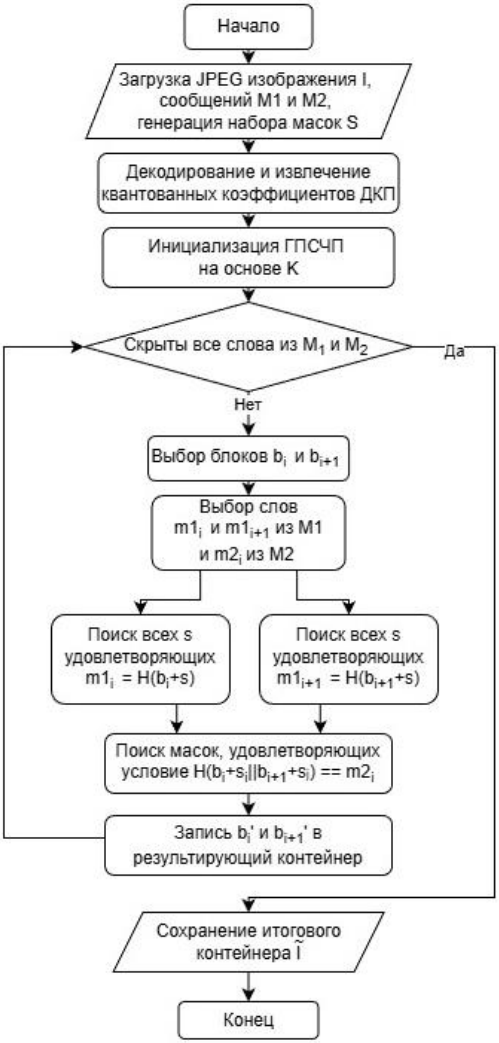


Рис. 3. Схема работы алгоритма встраивания данных с использованием двух каналов

Следующим шагом является поиск возможных изменений для обоих блоков. Для первого блока и второго блока вычисляется хэш от

конкатенации измененных коэффициентов. Если хэш совпадает с m_i из M_2, то эти изменения записываются в результирующий контейнер.

Процесс продолжается до тех пор, пока не будут использованы все блоки изображения или не закончатся слова из сообщений M_1 и M_2.

На последнем этапе происходит сохранение заполненного JPEG-контейнера $\tilde{I}$.Как результат, формируется заполненный контейнер с несколькими каналами передачи скрытой информации, которые могут эффективно использоваться различными абонентами. В среднем, для полного заполнения JPEG-контейнера размером 512×512 пикселей требуется около 0,35 с. Общая ёмкость встраиваемой информации для изображения такого размера составляет 6144 байт (4096 байт в первый канал и 2048 во второй канал).

Заключение

Среди преимуществ предложенной схемы мультиплексирования скрытых каналов для рассмотренного итеративного алгоритма встраивания информации в JPEG изображения можно отметить использование одного и того же подмножества элементов контейнера для одновременного скрытия нескольких различных сообщений. При встраивании данных сообщений выбранные подмножества элементов контейнера подвергаются одинаковым (минимальным по числу модифицируемых спектральных коэффициентов и по уровню модификации) изменениям, что позволяет снизить общий уровень искажений заполненных контейнеров и одновременно повысить пропускную способность стегосистемы в целом. При этом один из скрытых каналов может быть использован для передачи несекретной информации, тем самым маскируя передачу конфиденциальной информации по второму скрытому каналу. В числе ограничений предложенной схемы мультиплексирования каналов можно отметить требование синхронности при формировании скрытых каналов. Алгоритм мультиплексирования является универсальным, и в общем случае применим как в пространственном, так и в частотном представлении контейнеров различных форматов.

Список литературы

1. Holub, V. Designing steganographic distortion using directional filters / V. Holub, J. Fridrich // Proc. 4th IEEE Intern. Workshop on Inform. Forensics and Security (WIFS). – Tenerife, Spain, 2012. – С. 234-239.

2. Holub, V. Digital image steganography using universal distortion / V. Holub, J. Fridrich // Proc. 1st ACM Workshop on Inform. Hiding and Multimedia Security (IHMMSec). – Montpellier, France, 2013. – С. 59-68.

3. Pevný, T. Using high-dimensional image models to perform highly undetectable steganography / T. Pevný, T. Filler, P. Bas // LNCS, 2010. – T. 6387. – С. 161-177.

4. Denemark T., Bas P., Fridrich J. Natural Steganography in JPEG Compressed Images / T. Denemark, P. Bas, J. Fridrich // Proc. IS&T, Electronic Imaging, Media Watermarking, Security, and Forensics. – Burlingame, CA, 2018. – Режим доступа: https://hal.science/hal-01687194v1/document

5. Дрюченко, М. А. Адаптивный алгоритм стеганографического встраивания данных в цифровые изображения, использующий некриптографические хеш-функции для их извлечения / М. А. Дрюченко // Компьютерная оптика. – 2023. – № 3. – С. 415-426.

Практические аспекты обеспечения информационной безопасности при разработке проектов для программируемых логических контроллеров в средах CODESYS 3.5 и Astra IDE

Д. А. Ступникова, email: stupnikova.darya@mail.ru
А. Д. Горностаев, email: gornostaev.artem33@gmail.com
А. В. Русанов, email: rusanov.pmm@yandex.ru

ФАУ ГНИИИ ПТЗИ «ФСТЭК России»

Аннотация. *В статье рассматриваются вопросы обеспечения информационной безопасности на этапе разработки проектов для программируемых логических контроллеров в средах CODESYS 3.5 и Astra IDE. Актуальность работы связана с необходимостью защиты отечественных промышленных систем в условиях импортозамещения. Предложен подход, включающий анализ уязвимостей программируемых логических контроллеров и сред разработки, создание типовых проектов с использованием встроенных функций безопасности и оценку их эффективности.*

Ключевые слова: *Astra IDE, CODESYS, автоматизированные системы управления, импортозамещение, информационная безопасность, программируемый логический контроллер, среда разработки, уязвимости.*

Введение

События в мире показывают, что противостояние курса стран осуществляются не только в экономическом, физическом, но и в информационном пространстве. В свете необходимости достижения технологического суверенитета и импортозамещения в промышленности Российской Федерации актуальность обеспечении безопасности использования отечественных средств разработки и средств АСУ ТП резко возрастает [1].

Основу автоматизации на нижнем уровне современных технологических процессов в энергетике, химии, машиностроении и других отраслях, относящей к сфере критической информационной инфраструктуры (КИИ) составляют программируемые логические контроллеры (ПЛК) [2]. Удаленный доступ, широкое сетевое взаимодействие, а также многофункциональность ПО превращают ПЛК в потенциальные цели атак, которые приводят к остановке производства, выводу оборудования из эксплуатации и техногенным катастрофам [3].

© Ступникова Д.А, Горностаев А.Д., Русанов А.В., 2026

Но зачастую вопросы информационной безопасности АСУ ТП рассматриваются на этапе внедрения системы, а не на этапе ее создания.

Для повышения безопасности систем управления [4] необходимо перенести акцент безопасности на стадию разработки проекта и реализовать технические меры нейтрализации угроз непосредственно в средах программирования (CodeSys. 3.5, Astra IDE) и на уровне контроллеров (Regul 400, NLCon-CDE-18, АГАВА ПЛК-60).

Для решения поставленной задачи предлагается использовать комплексный подход, включающий следующие этапы: анализ уязвимостей в средах программирования и ПЛК, разработка типовых проектов в интегрированных средах разработки и экспериментальная проверка эффективности принятых мер.

1. Этап анализа: выявление уязвимостей исследуемых средств разработки и аппаратных платформ

На первом этапе разработки проектов необходимо выявить все уязвимости выбранных сред программирования и программируемых логических контроллеров.

Одним из способов выявление уязвимостей устройств и программного обеспечения является автоматизированное сканирование сети с помощью сканеров безопасности, например RedCheck. Осуществив проверку следующих объектов: АГАВА ПЛК-60, NLCon-CDE-18, Regul 400 получен следующий результат:

– АГАВА ПЛК-60 с программным обеспечением Linux RT 4.4.12, CodeSys 3.5.10 имеет архитектурные и организационные уязвимости, которые связаны с недостатками в управлении доступом и аутентификацией, парольной политике, с недостатками в управлении физическим доступом и ошибки конфигурации протоколов SMB, RDP, SSH.

– NLcon-CED-18 с программным обеспечением WinECompact 7 имеет архитектурные и организационные уязвимости, которые связаны с недостатками в управлении доступом и аутентификацией, парольной политике, с недостатками в управлении физическим доступом и использованием незащищенного протокола ModBusTCP.

– Regul R400 [5] с программным обеспечением QNX имеет архитектурные и организационные уязвимости, которые связаны с недостатками в управлении доступом и аутентификацией, парольной политике, с недостатками в управлении физическим.

Для парирования уязвимостей ПЛК и сред разработки интегрированных систем проведем анализ штатных функций безопасности (ФБ).

Штатные функции безопасности ПЛК проанализированы, обобщены и приведены на рис. 1-3. Возможности сред разработки по обеспечению безопасности приведены на рис. 4-5.

ПЛК:



Рис. 1. Штатные функции безопасности АГАВА ПЛК-60



Рис. 2. Штатные функции безопасности NLcon-CED-18

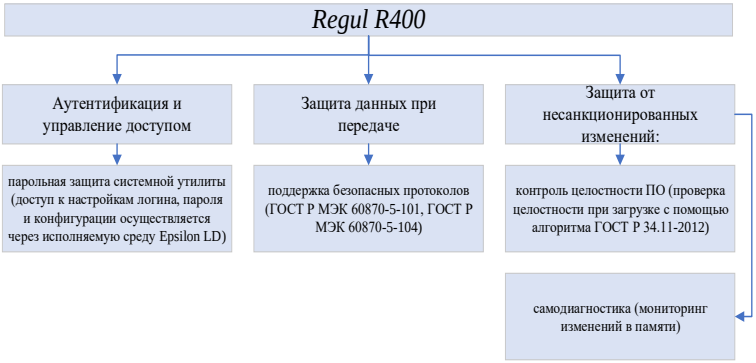


Рис. 3. Штатные функции безопасности Regul R400

Интегрированные среды разработки:

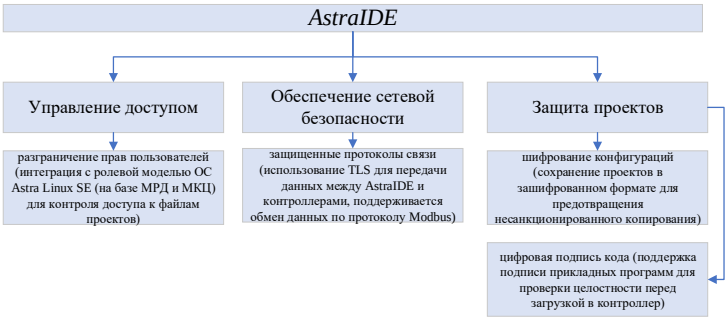


Рис. 4. Штатные функции безопасности Astra IDE

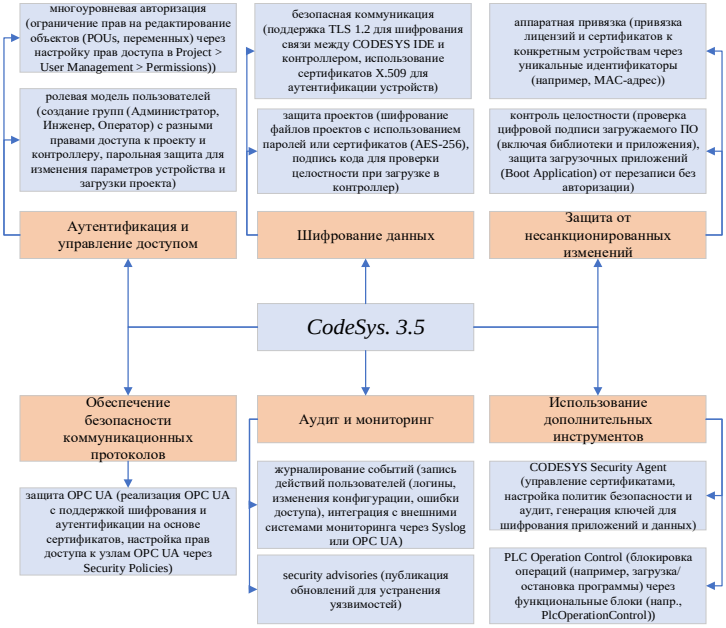


Рис. 5. Штатные функции безопасности CodeSys. 3.5

Анализ этих механизмов позволил перейти к разработке типовых проектов с учетом реализации данных функций.

2. Этап проектирования: разработка типовых проектов в интегрированных средах разработки CODESYS 3.5 и Astra IDE

Следующий этапом служит разработка проектов в интегрированных средах разработки Codesys 3.5 [6] и Astra IDE [7] с учетом выявленных возможностей [8,10]. Для каждого ПЛК выбраны параметры, которые отвечают за определенную часть технологического процесса доменной печи. В NLCon-CDE-18 обрабатываются и выводятся параметры теплового режима, ПЛК-60 АГАВА отвечает за шихтопадачу, ПЛК Regul 400 – подача и регулировка газа.

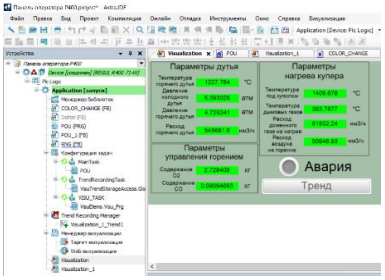


Рис. 6. Проект в Astra IDE ПЛК Regul 400

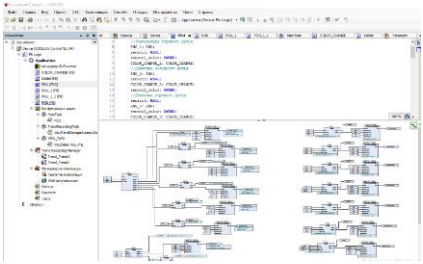


Рис. 7. Проект в Codesys 3.5 ПЛК NLCon-CDE-18

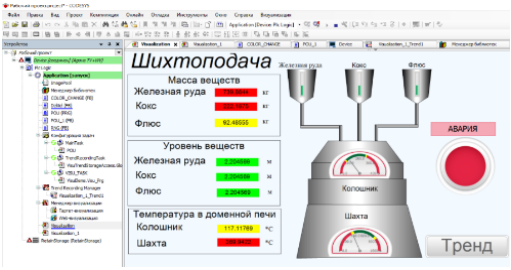


Рис. 8. Проект в Codesys 3.5 ПЛК-60 АГАВА

Особенностью данного этапа является не только реализация логики управления, но и целенаправленная конфигурация безопасности проектов. Это позволяет осуществлять централизованную откладку, где каждый типовой проект может быть тщательно протестирован один раз, а затем использован многократно, что резко повышает общую надежность системы. А при изменениях требований к системе или алгоритмам предоставляется возможность централизованного обновления с помощью библиотек.

3. Этап реализации: применение технических мер и оценка их эффективности

В процессе разработки проектов в интегрированных средах найдены решения нейтрализации актуальных угроз безопасности для исследуемых объектов, как средствами среды разработки, так и самих ПЛК. К ним относятся:

Реализация мер средствами среды разработки AstraIDE

- Реализовано управление доступом. Встроена проверка меток безопасности файлов проектов. Настроена автоматическая блокировка доступа к конфигурациям при несовпадении уровня доступа пользователя и метки ресурса;
- Реализовано шифрование конфигураций проектов.
- Реализована цифровая подпись прикладного кода ПО перед загрузкой в контроллеры REGUL. Настроена автоматическая проверка целостности подписанных программ при передаче через встроенный загрузчик;
- Настроено шифрование канала между средой разработки и контроллером R400.

Реализация этих мер позволило нейтрализовать следующие угрозы:

- Угроза прекращения функционирования ПЛК управления технологическим процессом (ТП), ПЛК противоаварийной защиты (ПАЗ);
- Угроза отказа в обслуживании телекоммуникационного оборудования, ПЛК (ПЛК ПАЗ), интеллектуальных датчиков и исполнительных механизмов;
- Угроза несанкционированной модификации уставок управления, уставок аварийной блокировки на ПЛК управления ТП, ПЛК ПАЗ;
- Угроза несанкционированной модификации либо подмены прикладных проектов контроля и управления на ПЛК управления ТП, ПЛК ПАЗ.

Описание способа нейтрализации угроз:

– Шифрование конфигураций проектов не позволяет осуществить их модификацию на рабочей станции нарушителя.
– Подключение к проекту для изменения уставок не позволяет без ввода пароля осуществить модификацию.
– Использование электронной подписи не позволило нарушителю загрузить в контроллеры REGUL модифицированный код без подписи.

Реализация мер средствами среды разработки CodeSys.

– Реализована система ролевого доступа с разделением прав пользователей. Настроены группы с различными уровнями привилегий: полный доступ к настройкам, ограниченное редактирование и режим только для просмотра данных;
– Осуществлена защита операций с устройствами. Внедрена обязательная аутентификация при попытке изменения критических параметров или загрузки проектов;
– Реализовано шифрование проектов;
– Реализована проверка подлинности приложений;
– Настроен механизм автоматической верификации целостности загружаемых программ и библиотек;
– Осуществлено шифрование каналов связи. Внедрены стандартные методы аутентификации устройств с поддержкой централизованного управления сертификатами.

Реализация этих мер позволило нейтрализовать следующие угрозы:
– Угроза прекращения функционирования ПЛК, ПЛК противоаварийной защиты (ПАЗ);
– Угроза отказа в обслуживании телекоммуникационного оборудования, ПЛК (ПЛК ПАЗ), интеллектуальных датчиков и исполнительных механизмов;
– Угроза несанкционированной модификации уставок управления, уставок аварийной блокировки на ПЛК, ПЛК ПАЗ;
– Угроза несанкционированной модификации либо подмены прикладных проектов контроля и управления на ПЛК, ПЛК ПАЗ.

Описание способа нейтрализации угроз:
– Разделение прав доступа не позволяет нарушителю с низкими привилегиями изменить уставки аварийной блокировки и отредактировать проекты управления без авторизации;
– Обязательная аутентификация при загрузке проектов не позволяет нарушителю загрузить пустой или несанкционированно модифицированный проект, изменить параметры, приводящие к остановке ПЛК.

– Шифрование проектов не позволяет открыть и модифицировать их на рабочем месте нарушителя

Реализация мер штатными функциями безопасности ПЛК различных типов

– Реализована парольная защита системной утилиты. Настроено управление доступом к параметрам безопасности: изменения логина, пароля и конфигурации разрешены только после авторизации в защищенной среде с проверкой прав пользователя;
– Реализовано шифрование данных при передаче через встроенные интерфейсы с использованием криптографических механизмов.

Реализация этих мер позволило нейтрализовать следующие угрозы:
– Угроза отказа в обслуживании АРМ оператора (диспетчера), промышленного сервера;
– Угроза несанкционированной модификации уставок управления, уставок аварийной и предупредительной блокировки на АРМ оператора, промышленном сервере, панели управления;
– Угроза несанкционированной модификации или подмены прикладных проектов контроля и управления на АРМ оператора, панели оператора, промышленном сервере.

Описание способа нейтрализации угроз:
– Парольная защита не позволяет получить несанкционированный доступ нарушителю;
– Загрузка некорректных прикладных программ невозможна при активированной проверке целостности;
– Защита сетевых соединений не позволяет вызвать «отказ в обслуживании»;
– Парольная защита на доступ не позволяет осуществить несанкционированную модификацию уставок.

Заключение

Проведённое исследование подтвердило высокую актуальность задачи обеспечения безопасности ПЛК и необходимость переноса акцента на ранние этапы жизненного цикла систем автоматизации, а именно на стадию разработки проектов для программируемых логических контроллеров. Экспериментальная работа в средах CODESYS 3.5 и Astra IDE для отечественных и импортозамещающих ПЛК (АГАВА ПЛК-60, NLcon-CDE-18, Regul R400) продемонстрировала, что комплексное использование штатных и дополнительных механизмов

защиты внедрение на этапе разработки проектов позволяет эффективно нейтрализовать ключевые угрозы безопасности, связанных с несанкционированным доступом, неконтролируемое состояние, модификацию записей и т.д, возникающих на этапе эксплуатации.

Основные практические результаты работы заключаются в следующем:

1. Функции безопасности (ролевой доступ, шифрование, подпись кода) должны быть неотъемлемой частью конфигурации проекта, а не внешней надстройкой.
2. Грамотная настройка среды разработки позволяет нивелировать часть архитектурных уязвимостей аппаратных платформ.
3. Реализованный комплекс мер успешно противодействует наиболее опасным векторам атак: подмене ПО, изменению уставок и отказу в обслуживании.

Предложенный подход, сочетающий анализ уязвимостей, использование защитных функций сред разработки и целенаправленную конфигурацию проектов, формирует практическую методологию для повышения киберустойчивости технологических процессов.

Список литературы

1. О безопасности критической информационной инфраструктуры Российской Федерации [Электронный ресурс] : федер. закон от 26.07.2017 № 187-ФЗ. – Режим доступа: https://www.consultant.ru/document/cons_doc_LAW_220885/

2. ГОСТ Р МЭК 62443-3-3-2016. Промышленные сети связи. Безопасность сетей и систем. Часть 3-3. Системные требования к безопасности и уровни безопасности [Текст]. – Введ. 2017-08-01. – М. : Стандартинформ, 2016.

3. Банк данных угроз безопасности информации ФСТЭК России [Электронный ресурс] : база данных / ФСТЭК России. – Режим доступа: https://bdu.fstec.ru/threat (дата обращения: 2026).

4. ГОСТ Р 56939-2016. Защита информации. Разработка безопасного программного обеспечения. Общие требования [Текст]. – Введ. 2017-06-01. – М. : Стандартинформ, 2016.

5. Программируемые логические контроллеры REGUL R400 [Электронный ресурс] : руководство по эксплуатации / ООО «Прософт-Системы». – Екатеринбург. – Режим доступа: https://regul-plc.ru/data/files/R400-CU-00-071_rukovodstvo_po_ekspluatacii.pdf

6. CODESYS Security Whitepaper [Электронный ресурс] : руководство по безопасности / 3S-Smart Software Solutions GmbH. – Режим доступа: https://ftp.owen.ru/CoDeSys3/98_Books/CodesysSecurity Whitepaper_ru.pdf

7. Среда разработки Astra IDE [Электронный ресурс] : руководство пользователя / ООО «Прософт-Системы». – Режим доступа: https://reglab.ru/docs/

Модель машинного обучения обнаружения компьютерных атак в беспроводных сетях стандарта IEEE 802.11

В. Ю. Храмов, e-mail: khramovvyu@mail.ru
А. С. Стешковой, e-mail: 9515431635@mail.ru
А. В. Щеглов, e-mail: aleksey.shcheglov@mail.ru

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н. Е. Жуковского и Ю. А. Гагарина» (г. Воронеж)

Аннотация. *В работе выделены классификационные признаки функционирования сетей цифровой радиосвязи стандарта IEEE 802.11, позволяющие формализовать параметры передаваемых данных и поведение устройств в сети с позиции информационной безопасности. Предложена модель машинного обучения обнаружения компьютерных атак в сетях стандарта IEEE 802.11, основанная на выявленных классификационных признаках.*

Ключевые слова: *классификационные признаки, машинное обучение, информационная безопасность, компьютерные атаки.*

Введение

В современном мире технологии беспроводной связи играют ключевую роль в обеспечении эффективного и быстрого доступа к информации. Актуальную роль в предоставлении беспроводного высокоскоростного канала связи выполняют сети цифровой радиосвязи (СЦР) стандарта IEEE 802.11 (Wi-Fi). Рассматриваемые сети приобрели широкое распространение благодаря своей универсальности, гибкости развертывания, а также совместимости всех поколений. Вместе с развитием СЦР стандарта IEEE 802.11 увеличивается интерес к данным сетям со стороны злоумышленников в части организации компьютерных атак (КА). Так, компьютерные атаки могут приводиться с целью нарушения целостности, конфиденциальности и доступности информации. В настоящее время известен ряд программных реализаций, направленных на проведение КА в СЦР стандарта IEEE 802.11, приводящих к частичной или полной блокировке рассматриваемых сетей.

Для своевременного выявления фактов КА в СЦР стандарта IEEE 802.11 применяются методы и алгоритмы их обнаружения [1-3]. В работах предложен метод обнаружения КА, основанный на анализе маячковых пакетов (Beacon Based Attacks), сущность которого

© Храмов В. Ю., Стешковой А. С., Щеглов А. В., 2026

заключается в мониторинге временных интервалов между Beacon-пакетами и установлении минимального порога для них. Недостаток данного метода – высокая вероятность ложных срабатываний в условиях нестабильного беспроводного канала, где естественные помехи или перегрузка сети могут вызывать отклонения в интервалах между Beacon-пакетами, не связанные с КА. В [1] для обнаружения КА, направленных на деаутентификацию элементов сети (Denial-of-Service), используется метод мониторинга потока кадров деаутентификации и учитывается резкое падение пропускной способности сети. К недостаткам данного метода можно отнести ложные срабатывания системы обнаружения, задержка в реагировании, а также ограниченность в условиях различных атак. К тому же существующие подходы предназначены для обнаружения ограниченного класса атак, что накладывает ограничение при реализации злоумышленником различных атак [2].

Таким образом возникает противоречие между потребностью в моделях и алгоритмах своевременного обнаружения широкого класса компьютерных атак в СЦР стандарта IEEE 802.11 и их отсутствием на практике.

1. Классификационные признаки функционирования сетей цифровой радиосвязи стандарта IEEE 802.11

Семейство стандартов IEEE 802.11 на текущий момент включает в себя следующий ряд стандартов: 802.11a (54 Мбит/с на 5 ГГц), 802.11b, 802.11g, 802.11n, 802.11ac, 802.11ax (WiFi 6), а также 802.11be (WiFi 7). Отличаются они используемыми технологиями. Все стандарты семейства IEEE 802.11 используют двухуровневую архитектуру, включающую физический уровень и канальный уровень [4].

Канальный уровень, в свою очередь, делится на два подуровня: подуровень управления логическим соединением и подуровень управления доступом к среде (MAC). Первый отвечает за мультиплексирование протоколов сетевого уровня, контроль ошибок и обеспечение логического интерфейса между сетевыми технологиями. Второй содержит правила, регулирующие доступ к среде передачи и отправку данных. На рис. 1 представлена эталонная модель взаимодействия элементов сети цифровой радиосвязи стандартов IEEE 802.11, иллюстрирующая описанную иерархию уровней.

Физический уровень СЦР стандарта IEEE 802.11 отвечает за передачу данных по радиоканалу. Он определяет методы модуляции, кодирования, частотные диапазоны и другие параметры.

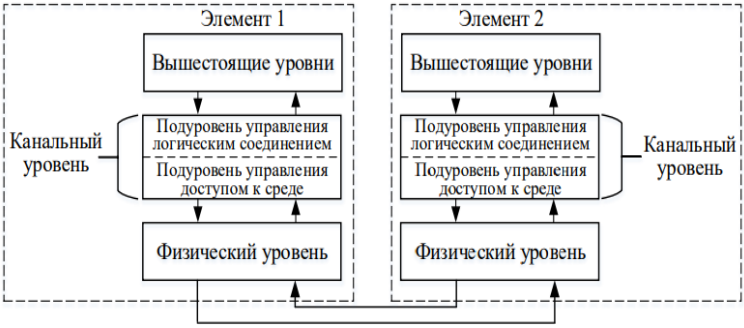


Рис. 1. Эталонная модель информационного взаимодействия элементов СЦР стандарта IEEE 802.11

К основным процедурам уровня доступа к среде СЦР стандарта IEEE 802.11 относятся: аутентификация, ассоциация, синхронизация, случайный множественный доступ к среде (СМДС) типа CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance) и централизованно- зарезервированный доступ к среде (ЦЗДС) [5-8].

Согласно спецификации IEEE 802.11, существует 35 основных кадров для организации связи и передачи данных. В современных стандартах их количество увеличивается за счет новых технологий.

В установившимся режиме работы (без воздействия КА) распределение типов кадров имеет определенную статистическую закономерность, однако в условиях КА это соотношение нарушается.

Эмпирически выявлено, что в стабильно работающей сети СЦР стандарта IEEE 802.11 преобладают кадры данных (Data Frames), так как их основная задача - передача полезной нагрузки. Управляющие и контрольные кадры занимают меньшую долю, обеспечивая служебные функции. Статистические исследования проводились для различных топологий СЦР стандарта IEEE 802.11 с выборкой 10000 кадров. На рис. 2 представлен перехват сетевого трафика (dump) стабильно функционирующей СЦР стандарта IEEE 802.11 [9].

В процессе эмпирического анализа сетевого трафика выявлено процентное соотношение различных типов кадров для стабильно функционирующей сети:

- кадры данных (Data Frames) 70–85% (основной трафик: HTTP, VoIP, видео и др.);
- кадры управления (Management Frames) 10–20% (Beacon, Probe, Authentication, Association);
- кадры контроля (Control Frames): 5–10% (ACK, RTS/CTS).

Приведенное распределение имеет незначительные отклонения в условиях различной нагрузки в сети, количества абонентов и типа сети (например, в публичных Wi-Fi-сетях доля Management-кадров выше из-за частых подключений).

В условиях КА соотношение типов кадров в радиоканале и, как следствие, в перехваченном трафике меняется в зависимости от типа атаки. Учитывая результаты работ [2-3], соотношение типов кадров в условиях различных атак меняется следующим образом:

- для атак типа отказ в обслуживании «Deauthentication» увеличивается количество кадров «Deauthentication»;
- для атаки, направленной на создание нелегитимной точки доступа «Rogue AP» увеличивается количество кадров «Beacon» и «Probe Response»;
- для атак, направленных на перехват трафика клиентов сети «Evil Twin», увеличивается количество кадров «Beacon», «Probe Response» и «Deauthentication».

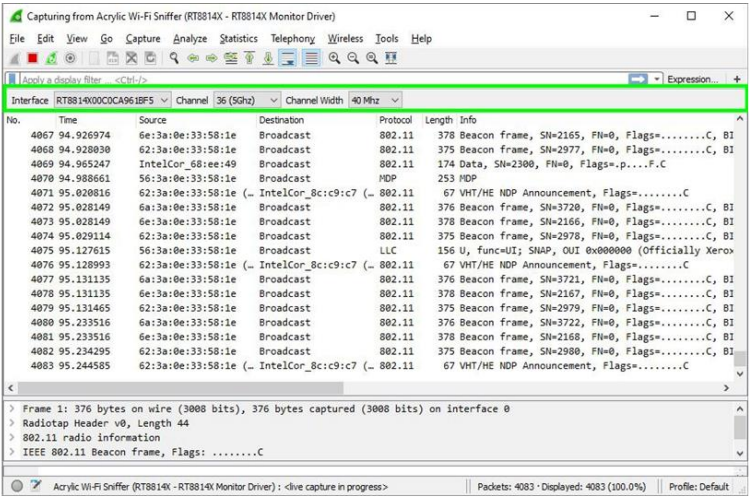


Рис. 2. Перехват сетевого трафика (dump) стабильно функционирующей СЦР стандарта IEEE 802.11

2. Модель машинного обучения обнаружения компьютерных атак в сетях цифровой радиосвязи стандарта IEEE 802.11

Задача обнаружения компьютерных атак в СЦР стандарта IEEE 802.11 формулируется как задача бинарной классификации. На вход модели подается вектор признаков X, извлеченный из

последовательности 1000 перехваченных кадров. Выходом модели является бинарное значение $y \in \{0,1\}$, где $y = 0$ – нормальный режим функционирования сети, $y = 1$– обнаружена компьютерная атака.

Признаковое пространство формируется на основе процентного соотношения типов кадров и дополнительных статистических характеристик:

$$X = [f_{mgmt}, f_{ctrl}, f_{data}, f_{beacon}, f_{auth}, \sigma_{data}],$$

где $f_{mgmt}, f_{ctrl}, f_{data}$ – доли кадров управления, контроля и данных соответственно, f_{beacon} – частота Beacon кадров (отвечающих за передачу служебной информации точкой доступа), f_{auth} – количество попыток аутентификации, σ_{data} – дисперсия размера Data-кадров.

Для решения поставленной задачи рассматривались два классических алгоритма машинного обучения: дерево решений (Decision Tree) и наивный байесовский классификатор (Gaussian Naive Bayes). Выбор данных алгоритмов обусловлен требованиями к системе обнаружения: интерпретируемость результатов, вычислительная эффективность и способность работать с зашумленными данными.

Дерево решений представляет собой иерархическую структуру, где каждый внутренний узел соответствует проверке значения одного из признаков, каждая ветвь – исходу проверки, а каждый лист – значению целевой переменной. В данной работе использовался алгоритм CART (Classification and Regression Trees) с критерием разделения Джини (Gini impurity) [10-12].

Критерий Джини для узла t вычисляется по формуле:

$$Gini(t) = 1 - \sum_{i=1}^{C} p(i \mid t)^2,$$

где $p(i \mid t)$ — вероятность отнесения объекта к классу i в узле t.

Качество разделения оценивается как взвешенная сумма критериев Джини для дочерних узлов:

$$Gini_{split} = \frac{n_{left}}{n} \cdot Gini_{left} + \frac{n_{right}}{n} \cdot Gini_{right},$$

где n_{left}, n_{right} — количество объектов в левом и правом дочерних узлах, n – общее количество объектов в родительском узле.

Архитектура дерева ограничена глубиной 4 уровня, что обеспечивает баланс между точностью классификации и интерпретируемостью модели.

Выбор данного алгоритма после сравнительного анализа с наивным байесовским классификатором обусловлен следующими факторами: высокая интерпретируемость, отсутствие предположений о распределении данных, автоматический отбор признаков, устойчивость к зашумленным данным, высокая точность классификации.

Модель реализована с использованием библиотеки scikit-learn для Python. Процесс обучения включает следующие этапы:

1. Формирование обучающей и тестовой выборок.
2. Нормализация данных.
3. Обучение дерева решений с параметрами: максимальная глубина = 4, критерий разделения - Джини.
4. Оценка качества модели на тестовой выборке.
5. Визуализация дерева решений для анализа правил классификации.

Для обучения и тестирования модели использовалась выборка трафика СЦР стандарта IEEE 802.11 в нормальном режиме. Условия КА синтезировались. Общий объем выборки составил 5000 сэмплов, каждый из которых представляет статистику 1000 последовательных пакетов. Выборка разделена на обучающую (70%) и тестовую (30%) части с сохранением пропорций классов. Распределение признаков для классов "Норма" и "Атака" представлено на рис. 3.

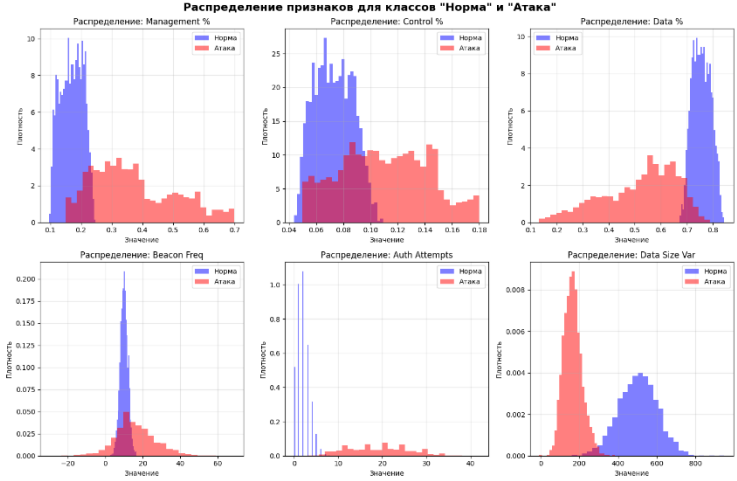


Рис. 3. Распределение признаков для классов "Норма" и "Атака"

Предложенная модель на основе дерева решений продемонстрировала высокую эффективность в обнаружении различных типов компьютерных атак:

1. Атаки типа Deauthentication: обнаруживаются с точностью 99,2% за счет резкого увеличения доли Management-пакетов (преимущественно Deauthentication-фреймов) и соответствующего снижения доли Data-пакетов.
2. Атаки Rogue AP: выявляются с точностью 97,8% по аномально высокой частоте Beacon-фреймов и увеличению доли Management-пакетов.
3. Атаки Evil Twin: Классифицируются с точностью 96,5% на основе комплексного анализа увеличения Management-пакетов, частоты Beacon-фреймов и количества попыток аутентификации.

Перспективными направлениями дальнейших исследований являются:

1. .Разработка адаптивной модели с динамическими порогами классификации в зависимости от текущей сетевой нагрузки.
2. Интеграция временных характеристик трафика для обнаружения атак в реальном времени.
3. Создание ансамблевой модели на основе нескольких алгоритмов машинного обучения для повышения эффективности системы обнаружения.

Заключение

Разработана модель машинного обучения для обнаружения компьютерных атак в сетях цифровой радиосвязи стандарта IEEE 802.11. В качестве исходных данных для модели выступают статистические значения сетевого трафика. В качестве базового алгоритма выбрано дерево решений с глубиной 4, что обеспечивает высокую точность классификации (98,2%) при сохранении интерпретируемости результатов. Модель позволяет обнаруживать атаки различных типов (Deauthentication, Rogue AP, Evil Twin) на основе анализа процентного соотношения типов кадров и дополнительных статистических характеристик. Полученные результаты превосходят существующие методы обнаружения по точности и охвату типов атак, что подтверждает эффективность предложенного подхода для повышения безопасности беспроводных сетей.

Литература

1. Automated Wi-Fi intrusion detection tool on 802.11 networks / D. Koutras [and others] // ITU Journal on Future and Evolving Technologies. – 2024. – Vol. 5, no. 1. – P. 88-103.

2. Gill, R. Specification-Based Intrusion Detection in WLANs / R. Gill, J. Smith, A. Clark // Proceedings of the 22nd Annual Computer Security Applications Conference (ACSAC'06), Miami Beach, FL, USA. – 2006. – P. 141-152.

3. Титов, К. Д. Оценка помехоустойчивости системы связи стандарта IEEE 802.11ac при воздействии помех / К. Д. Титов, О. Н. Завалишина // Успехи современной радиоэлектроники. – 2019. – № 12. – С. 191-196.

4. IEEE standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks–specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: IEEE Std 802.11.: IEEE, 2012. – [Электронный ресурс]. – Режим доступа: https://ieeexplore.ieee.org/document/11023005

5. Перегудов, М. А. Вероятностная модель процедуры случайного множественного доступа к среде типа CSMA/CA / М. А. Перегудов, А. С. Стешковой, А. А. Бойко // Труды СПИИРАН. – 2018. – № 4 (59). – С. 92-114.

6. Перегудов, М. А. Описательная модель канального уровня сетей цифровой радиосвязи семейства стандартов IEEE 802.11 / М. А. Перегудов, А. В. Щеглов // Системы управления, связи и безопасности. – 2020. – № 3. – С. 203-221.

7. Al-Wraikat, A. Ack spoofing attack in IEEE 802.11 infrastructure WLANs: Strategies, detection and mitigation / A. Al-Wraikat, O. M. F. Abu-Sharkh, H. A. Noman // IET Networks. – 2025. – Vol. 14, no. 1. – P. 1-15.

8. Перегудов, М. А. Модель централизованной синхронизации элементов сетей цифровой радиосвязи со случайным множественным доступом к среде типа CSMA/CA / М. А. Перегудов, А. С. Стешковой // Труды СПИИРАН. – 2020. – Т. 19, № 1. – С. 128-154.

9. Wireshark [Электронный ресурс]: официальный сайт. – Режим доступа: https://www.wireshark.org

10. Optimal or greedy decision trees: Revisiting their objectives, tuning, and performance / J. G. M. van der Linden [and others] // The 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining. – 2025.

11. Balcan, M.-F. Learning accurate and interpretable decision trees (Extended Abstract) / M.-F. Balcan, D. Sharma // Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence. – 2025. – P. 10858-10863.

12. Китов В. В. Глубокое машинное обучение [Электронный ресурс]. – Режим доступа: https://www.intuit.ru/studies/courses/6/6/lecture/198?page=3/

Автоматизированная система обезличивания персональных данных в медицинских текстовых документах на основе NLP с веб-интерфейсом

П. В. Шатков, email: papla2016@yandex.ru
В. В. Гаршина, email: garshina.veronika@gmail.com

Воронежский государственный университет

Аннотация. *Рассмотрена разработка веб-приложения , реализующего автоматическое обезличивание персональных данных в текстовых документах. Разработано веб-приложение для автоматизированного обезличивания документов с использованием NER.*

Ключевые слова: *персональные данные, обезличивание, анонимизация, NER, Natasha, Yargy, регулярные выражения, токенизация.*

Введение

Персональные данные (ПДн) в современных информационных системах накапливаются и обрабатываются в значительных объемах, в том числе в виде неструктурированных документов: заявлений, обращений, медицинских записей, протоколов, договоров и переписки. Утечки таких данных приводят к финансовому и репутационному ущербу, а также к усилению действия мошенников благодаря обширному наличию персональных данных в свободном доступе. Актуальность задачи обезличивания дополнительно обусловлена действующими требованиями к обработке и защите ПДн [1][2][3][4], а также отраслевыми рекомендациями и стандартами в медицинских сценариях [5].

Для задач анализа текстов и машинного обучения прямые идентификаторы (ФИО, адреса, контакты) обычно не нужны, так как их наличие делает наборы данных чувствительными к контексту, что негативно влияет на нейронные сети, которые обучаются на этих данных. Одним из базовых инструментов снижения рисков является обезличивание (деперсонификация), которое направлено на исключение возможности идентифицировать субъекта без дополнительной информации [1][2][3][4].

Целью работы является разработка прототипа системы автоматического обезличивания ПДн в русскоязычных неструктурированных текстах с сохранением возможности обратного

© Гаршина В. В., Шатков П. В., 2026

восстановления при наличии ключей и таблицы соответствий, а также создание веб-интерфейса для удобного применения технологии в медицинской организации. Фокус сделан на практическом конвейере обработки текста, включающем обнаружение сущностей (NER), токенизацию и криптографическую защиту сопоставлений [6][7][8][9].

1. Подходы к обезличиванию и критерии анонимности

Методы обезличивания в общем случае сводятся к трансформации исходных атрибутов так, чтобы снизить риск реидентификации [4]. Для текстов наиболее применимы супрессия (удаление/скрытие фрагментов), генерализация (обобщение значений), перестановка и добавление шума для числовых параметров, а также токенизация/псевдонимизация, когда идентификаторы заменяются маркерами с хранением таблицы соответствий [4]. Выбор метода определяется компромиссом между приватностью и сохранением смысловой ценности документа [4][5].

Для формальной оценки уровня обезличивания используются метрики анонимности. k-анонимность требует наличия как минимум k записей с одинаковыми квази-идентификаторами; l-диверсификация дополняет ее требованием разнообразия чувствительных атрибутов в каждой группе; t-близость ограничивает отклонение распределений чувствительных атрибутов внутри группы от распределения в датасете. Более строгие гарантии предоставляет дифференциальная приватность, основанная на контролируемом внесении случайности в результаты анализа [10]. В данной работе рассматривается прикладной сценарий «текстовый документ — обезличенный текст», где основным механизмом выступает токенизация с защищенным хранением соответствий [6][7][8][9].

2. Алгоритм и архитектура предлагаемой системы

Предлагаемый конвейер обработки включает четыре ключевых этапа:

1. извлечение персональных сущностей;
2. замена найденных фрагментов токенами;
3. формирование защищенного хранилища соответствий;
4. деобезличивание (восстановление) при наличии ключей [6][7][8][9].

5. На этапе извлечения используется гибридный детектор. Нейросетевая NER-модель Natasha выделяет базовые типы сущностей (персоны, организации, географические объекты и др.) [6]. Для повышения точности на формальных и доменных конструкциях применяются правила Yargy: например, распознавание организаций по суффиксам «ООО», «ЗАО» и т. п., должностей, адресных шаблонов и

денежных сумм [7]. Типы данных, которые часто пропускаются NER-моделью (ИНН, СНИЛС, номера карт, БИК, паспорта, телефоны, e-mail, даты), дополнительно обрабатываются регулярными выражениями [4].

После обнаружения сущностей выполняется токенизация: каждый фрагмент заменяется маркером формата {TYPE_INDEX}, где TYPE — класс сущности, INDEX — уникальный номер в пределах документа. Параллельно формируется структура сопоставлений, включающая тип сущности, ее исходное значение, нормализованную форму и позиции в тексте [6][7]. Сопоставления сериализуются в JSON, что облегчает передачу между компонентами и дальнейшую интеграцию с системами документооборота.

Для защиты сопоставлений используется гибридная криптосхема: JSON шифруется симметричным алгоритмом AES (например, AES-256) [8], а ключ AES шифруется асимметричным RSA с применением OAEP и SHA-256 [9]. Такое разделение позволяет совместить производительность симметричного шифрования с удобством безопасного обмена ключами в распределенной архитектуре [8][9]. На этапе восстановления зашифрованный ключ AES расшифровывается закрытым ключом RSA, затем восстанавливается JSON и выполняется обратная подстановка оригинальных значений по токенам [8][9].

3. Практическая реализация прототипа

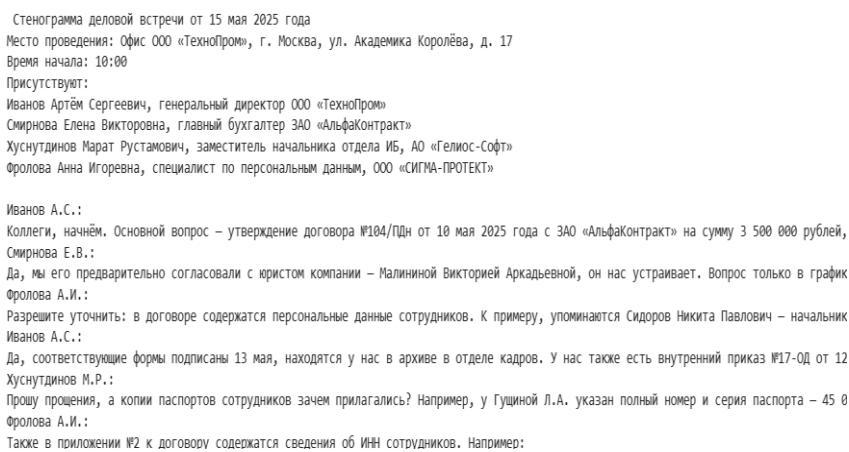
Прототип реализован на языке Python. Вычислительная часть и гибридный детектор (Natasha, правила Yargy и регулярные выражения) были отработаны в исследовательской среде, после чего интегрированы в веб-приложение, предоставляющее пользователю интерфейс для ввода или загрузки текстов, запуска обезличивания и просмотра результатов [6][7]. Регулярные выражения выделяют формализованные идентификаторы и контактные данные [4]. Модуль токенизации поддерживает детерминированное присваивание индексов и корректную обработку повторяющихся сущностей, что позволяет сохранять связность текста и снижать вероятность логических ошибок при восстановлении [6][7].

Криптографический модуль использует генерацию случайного ключа AES и вектора инициализации, шифрование полезной нагрузки и последующее шифрование ключа AES открытым ключом RSA [8][9]. Для устойчивости к подмене данных на практике рекомендуется дополнять схему контролем целостности (например, режимом AEAD или HMAC), однако в рамках прототипа акцент сделан на воспроизводимой демонстрации принципа работы и совместимости компонентов [8][9]. Результатом работы системы является пара артефактов: обезличенный

текст и зашифрованный контейнер сопоставлений, необходимый для деобезличивания [8][9].

4. Экспериментальная проверка

Проверка работоспособности проводилась по двум сценариям. Первый сценарий ориентирован на корректность восстановления: исходный текст с намеренно внесенными персональными данными обезличивался, затем выполнялось восстановление. Автоматическая верификация реализована сравнением строк: assert original_text == restored_text. Проверка подтвердила полное совпадение восстановленного текста с оригиналом, что свидетельствует об отсутствии потерь или искажений при токенизации, шифровании, хранении и обратной подстановке [8][9].



Стенограмма деловой встречи от 15 мая 2025 года
Место проведения: Офис ООО «ТехноПром», г. Москва, ул. Академика Королёва, д. 17
Время начала: 10:00
Присутствуют:
Иванов Артём Сергеевич, генеральный директор ООО «ТехноПром»
Смирнова Елена Викторовна, главный бухгалтер ЗАО «АльфаКонтракт»
Хуснутдинов Марат Рустамович, заместитель начальника отдела ИБ, АО «Гелиос-Софт»
Фролова Анна Игоревна, специалист по персональным данным, ООО «СИГМА-ПРОТЕКТ»

Иванов А.С.:
Коллеги, начнём. Основной вопрос — утверждение договора №104/ПДн от 10 мая 2025 года с ЗАО «АльфаКонтракт» на сумму 3 500 000 рублей,
Смирнова Е.В.:
Да, мы его предварительно согласовали с юристом компании — Малининой Викторией Аркадьевной, он нас устраивает. Вопрос только в график
Фролова А.И.:
Разрешите уточнить: в договоре содержатся персональные данные сотрудников. К примеру, упоминаются Сидоров Никита Павлович — начальник
Иванов А.С.:
Да, соответствующие формы подписаны 13 мая, находятся у нас в архиве в отделе кадров. У нас также есть внутренний приказ №17-ОД от 12
Хуснутдинов М.Р.:
Прошу прощения, а копии паспортов сотрудников зачем прилагались? Например, у Гущиной Л.А. указан полный номер и серия паспорта — 45 0
Фролова А.И.:
Также в приложении №2 к договору содержатся сведения об ИНН сотрудников. Например:

Рис. 1. Часть исходного текста файла

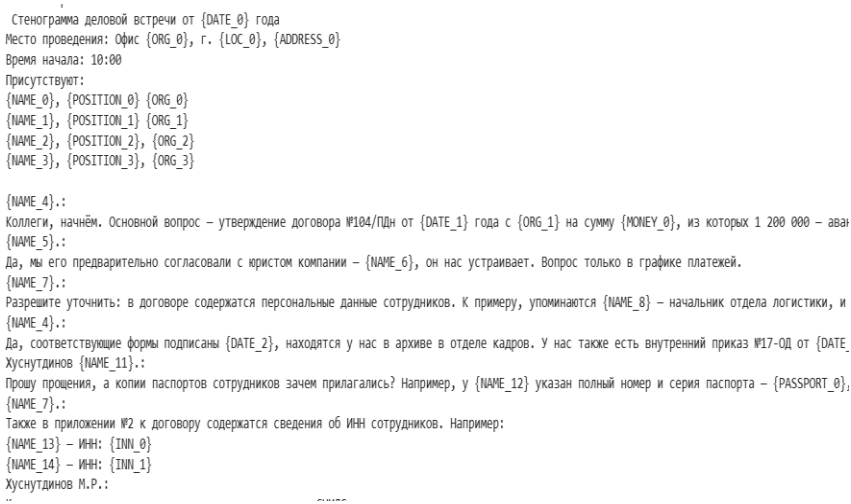

Стенограмма деловой встречи от {DATE_0} года
Место проведения: Офис {ORG_0}, г. {LOC_0}, {ADDRESS_0}
Время начала: 10:00
Присутствуют:
{NAME_0}, {POSITION_0} {ORG_0}
{NAME_1}, {POSITION_1} {ORG_1}
{NAME_2}, {POSITION_2}, {ORG_2}
{NAME_3}, {POSITION_3}, {ORG_3}
{NAME_4}.:
Коллеги, начнём. Основной вопрос – утверждение договора №104/ПДн от {DATE_1} года с {ORG_1} на сумму {MONEY_0}, из которых 1 200 000 – ава
{NAME_5}.:
Да, мы его предварительно согласовали с юристом компании – {NAME_6}, он нас устраивает. Вопрос только в графике платежей.
{NAME_7}.:
Разрешите уточнить: в договоре содержатся персональные данные сотрудников. К примеру, упоминаются {NAME_8} – начальник отдела логистики, и
{NAME_4}.:
Да, соответствующие формы подписаны {DATE_2}, находятся у нас в архиве в отделе кадров. У нас также есть внутренний приказ №17-ОД от {DATE_
Хуснутдинов {NAME_11}.:
Прошу прощения, а копии паспортов сотрудников зачем прилагались? Например, у {NAME_12} указан полный номер и серия паспорта – {PASSPORT_0},
{NAME_7}.:
Также в приложении №2 к договору содержатся сведения об ИНН сотрудников. Например:
{NAME_13} – ИНН: {INN_0}
{NAME_14} – ИНН: {INN_1}
Хуснутдинов М.Р.:

Рис. 2. Часть текста после токенизации

Во втором сценарии оценивалась точность выявления персональных данных в документе с более реалистичной структурой. Из-за отсутствия общедоступных размеченных корпусов с реальными ПДн итоговая оценка проводилась экспертно. Анализ показал, что гибридный детектор уверенно выявляет большинство сущностей (имена, организации, контактные данные) [6][7], а оставшиеся пропуски относятся к типичным ошибкам NER-моделей и неоднозначностям сокращений [6]. Результаты можно считать удовлетворительными для практического применения на этапе предварительной подготовки текстовых корпусов и обработки документов в медицинской, юридической и организационной сферах [5].

Основными направлениями развития являются: расширение набора правил Yargy под специализированные домены [7], использование моделей NER, дообученных на целевых данных [6], введение уровней уверенности и постобработки для снижения ложных срабатываний, а также переход на AEAD-режимы шифрования для встроенной проверки целостности [8][9].

Заключение

Разработан и реализован прототип системы автоматического обезличивания персональных данных в неструктурированных русскоязычных текстах. Система использует гибридный детектор сущностей (Natasha + Yargy + регулярные выражения) [6][7], токенизацию с формированием JSON-таблицы соответствий и

гибридную криптосхему AES + RSA для защиты сопоставлений [8][9]. Для практического использования дополнительно создано веб-приложение (сайт), обеспечивающее доступ к функциям обезличивания через пользовательский интерфейс и поддерживающее прикладные сценарии клинической работы [5]. Экспериментальная проверка показала корректное восстановление текста и практическую применимость подхода при экспертной оценке качества обезличивания [6][7][8][9]. Полученные результаты могут быть использованы как основа для дальнейшей разработки промышленных модулей обезличивания и интеграции с системами электронного документооборота [1][2][3][4][5].

Список литературы

1. Федеральный закон от 27.07.2006 № 152-ФЗ «О персональных данных» (ред. от 08.08.2024) : принят Государственной Думой 08.07.2006 : одобрен Советом Федерации 14.07.2006 [Электронный ресурс]. – федеральный закон. – Режим доступа : https://www.consultant.ru/document/cons_doc_LAW_61801/
2. Федеральный закон от 08.08.2024 № 233-ФЗ «О внесении изменений в отдельные законодательные акты Российской Федерации…» (в части требований к деперсонификации данных при проектах ИИ) [Электронный ресурс]. – федеральный закон. – Режим доступа : https://normativ.kontur.ru/document?moduleId=1&documentId=475795
3. Постановление Правительства РФ от 01.11.2012 № 1119 «Об утверждении требований к защите персональных данных при их обработке в информационных системах персональных данных» [Электронный ресурс]. – нормативный правовой акт. – Режим доступа : https://base.garant.ru/70252506/
4. Приказ Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций от 05.09.2013 № 996 «Об утверждении методических рекомендаций по обезличиванию персональных данных». – [Электронный ресурс]. –приказ – Режим доступа : https://www.garant.ru/products/ipo/prime/doc/70351476/
5. ГОСТ Р 71674-2024. Системы искусственного интеллекта в клинической медицине. Методы обезличивания данных [Текст]. – Введ. 2025-01-01. –М.: Российский институт стандартов, 2024. – IV, 13 с. : ил. ; 29 см.
6. Natasha [Электронный ресурс]: библиотека для обработки русского языка и извлечения именованных сущностей (NER) : документация проекта. – Режим доступа: https://nbviewer.org/github/natasha/natasha/blob/master/docs.ipynb

7. Yargy [Электронный ресурс]: библиотека правил и грамматик извлечения сущностей в русскоязычных текстах : документация. – Режим доступа : https://github.com/natasha/yargy

8. NIST. FIPS PUB 197: Advanced Encryption Standard (AES) [Электронный ресурс]. — Gaithersburg, MD, 2001. – Режим доступа: https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.197.pdf

9. RFC 8017: PKCS #1 v2.2. RSA Cryptography Specifications [Электронный ресурс]. — [S. l.] : IETF, 2016 — Режим доступа: https://www.rfc-editor.org/rfc/rfc8017

10. Dwork, C. The Algorithmic Foundations of Differential Privacy [Электронный ресурс] C. Dwork, A. Roth. — Boston, MA : Now Publishers Inc., 2014. – Режим доступа: https://maxkasy.github.io/home/files/other/ML_Econ_Oxford/differential_privacy.pdf

Анализ методов повышения полноты извлечения релевантной информации в системах RAG

С. А. Шишкин, email: shishkin_s_a@sc.vsu.ru
А. В. Швырева, email: shvyreva_a_v@sc.vsu.ru

Воронежский государственный университет

Аннотация. *Системы генерации с помощью извлеченной информации (Retrieval-Augmented Generation, RAG) являются одним из наиболее распространённых подходов для интеграции больших языковых моделей с внешними источниками знаний. Эффективность таких систем существенно ограничивается качеством этапа извлечения, даже при наличии релевантной информации в корпусе генеративная модель не может сформировать корректный ответ, если соответствующие документы отсутствуют среди извлечённых. В данной работе рассматриваются методы повышения полноты извлечения релевантных документов в RAG-системах за счёт архитектурных и алгоритмических улучшений retrieval-пайплайна. Предлагается подход, включающий семантический чанкинг документов с использованием anchor-based стратегии, гибридное извлечение на основе плотных и разреженных представлений, метод rewrite, а также entity-based fallback-механизм для расширения пространства поиска. Дополнительно анализируется coverage-aware алгоритм отбора контекста. Эффективность решения оценивалась с помощью датасета RuBQ. Экспериментальные результаты демонстрируют устойчивый рост recall и итоговой точности ответов по сравнению с базовыми конфигурациями RAG, а также показывают вклад каждого компонента системы.*

Ключевые слова: *Retrieval-Augmented Generation, BM25, информационный поиск, dense retrieval, семантический чанкинг, именованные сущности.*

Введение

Большие языковые модели (*LLM*) показывают высокую эффективность при решении задач генерации ответов на вопросы и логического анализа. Однако их применение в сценариях, требующих больших объёмов знаний, ограничено из-за статичности обучающих данных и отсутствия прямого доступа к внешним источникам информации. В таких условиях модели могут выдавать некорректные ответы или создавать фактические ошибки.

© Шишкин С. А., Швырева А. В., 2026

Одним из способов преодолеть эти ограничения является архитектура *Retrieval-Augmented Generation* (*RAG*), где языковая модель работает совместно с модулем извлечения релевантных документов из внешнего корпуса. Процесс ответа в классическом *RAG* строится на двух этапах: сначала определяется релевантный контекст, а затем на его основе формируется ответ.

Практическая эффективность RAG-систем во многом зависит от качества этапа извлечения. Ошибки на этом этапе могут сделать генеративную модель неспособной дать корректный ответ, независимо от её возможностей. Поэтому повышение полноты извлечения является ключевым условием надёжной работы *RAG*.

В данной работе утверждается, что улучшение retrieval-пайплайна является более перспективным направлением, чем усложнение генеративной модели или агрессивный *prompt engineering*. Основной акцент делается на архитектурных и алгоритмических решениях, направленных на снижение вероятности пропуска релевантной информации.

Целью работы является разработка и экспериментальная оценка методов повышения *recall* в RAG-системах за счёт семантически согласованной сегментации документов, гибридного извлечения на основе *BM25* и плотных эмбеддингов, а также с помощью адаптивного *entity-based* расширения запросов и анализа механизмов *diversity* и *coverage* при отборе контекста.

1. Связанные работы

Архитектура *Retrieval-Augmented Generation* была формализована в работе Lewis et al. и с тех пор получила широкое распространение в задачах *question answering* и *knowledge-intensive NLP*. [1] Существенная часть последующих исследований сосредоточена на улучшении *dense retrieval* за счёт обучения специализированных *bi-encoder* и *cross-encoder* моделей [2], а также на комбинировании нейронных и разреженных методов поиска.

Отдельное направление работ связано с гибридным *retrieval*, в котором результаты *BM25* и dense-поиска объединяются с использованием различных схем *fusion*, включая *Reciprocal Rank Fusion*. Другой активно исследуемый класс методов включает генерацию переформулированных запросов и расширение запросов (*query rewriting / expansion*). [3]

В то же время сравнительно меньше внимания уделяется вопросам предварительной сегментации документов и структурному расширению запросов. Семантический чанкинг часто реализуется с использованием эвристик, не учитывающих проблему *semantic drift*. Аналогично, *entity-*

based методы, как правило, используются либо в качестве самостоятельного retrieval-канала, либо в рамках *entity linking*, а не как адаптивный механизм повышения полноты.

Настоящая работа дополняет существующие исследования, предлагая интерпретируемые и модульные решения, направленные именно на повышение *recall* retrieval-этапа.

2. Формализация задачи

Пусть задан корпус документов $D = \{d_1, d_2, ..., d_N\}$, и множество запросов $Q = \{q_1, q_2, ...\}$. В архитектуре RAG процесс ответа на запрос $q \in Q$ описывается как $a = G(q, R(q, D))$, где R - retrieval-функция, возвращающая множество кандидатов $Cq \subset D$, а G - генеративная модель.

Полнота извлечения определяется как вероятность того, что документ d*, содержащий корректный ответ, присутствует среди k извлечённых кандидатов.

$$\mathrm{Re}call@k = P(d^* \in C_q^{(k)}),$$

где d^* - документ, содержащий корректный ответ.

Основной тезис работы заключается в том, что оптимизация функции R оказывает решающее влияние на итоговое качество всей системы.

3. Архитектура retrieval-модуля

Retriever реализован как модульная архитектура, включающая независимые каналы извлечения и последовательные механизмы постобработки. Базовый пайплайн включает *sparse-retrieval* и *dense-retrieval*, их объединение, *reranking*, а также fallback-механизмы и опциональный coverage-aware слой.

Принципиально важным является то, что *entity-based retrieval* и *coverage-aware* отбор не используются постоянно, а активируются только при недостаточной полноте базового извлечения.

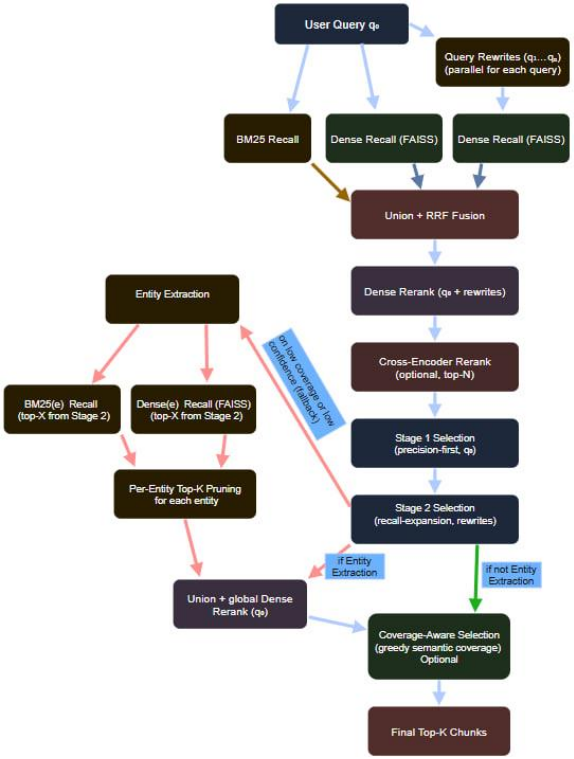


Рис. 1. Общая схема архитектуры retrieval-модуля

4. Семантический чанкинг документов

Пусть документ представлен последовательностью предложений $d = \{s_1, s_2, ..., s_N\}$, каждому из которых сопоставляется эмбеддинг $e_i \in R$. Задача чанкинга заключается в разбиении документа на подмножества $d \to \{c_1, c_2, ..., c_M\}$, таких, что каждый чанк представляет собой семантически целостный фрагмент текста, пригодный для последующего этапа извлечения.

В системах *RAG* ключевым требованием к чанкингу является не только локальная связность текста, но и повышение вероятности того, что релевантный фрагмент документа будет корректно извлечён retrieval-модулем. В отличие от разбиения на фиксированные по длине сегменты или по количеству предложений, семантический чанкинг учитывает тематическую структуру документа и снижает вероятность включения нерелевантного контекста.

В данной работе применяется стратегия семантического чанкинга с использованием якорных предложений (*anchor-based*), предназначенная для предотвращения эффекта постепенного смещения семантики (*semantic drift*). Для каждого чанка фиксируется якорное предложение, выбранное как первое предложение в чанке, с эмбеддингом e_{anchor}, который определяет семантическое ядро текущего фрагмента. Новое предложение s_i добавляется в текущий чанк только при выполнении условия:

$$cos\left(e_{anchor}, e_i\right) \geq \tau,$$

где e_i - эмбеддинг рассматриваемого предложения, а τ - порог сходства.

Решение о расширении чанка принимается исключительно на основе сходства с якорным эмбеддингом, а не со средним представлением чанка. Это позволяет избежать накопления ошибок и постепенного смещения семантики при добавлении новых предложений, характерного для усреднённых стратегий [4].

Средний эмбеддинг чанка используется лишь вспомогательно - для стабилизации представления и последующего анализа, но не участвует в принятии решения о включении предложений.

Кроме того, в алгоритме присутствуют ограничения на минимальную и максимальную длину чанков. В ходе экспериментов было доказано, что уменьшение этих параметров увеличивает общее число чанков, однако улучшает локализацию релевантной информации, а значит, что более мелкие, но семантически однородные фрагменты снижают риск «размытия» полезного контекста и повышают *Recall@k* на этапе *retrieval*, несмотря на рост объёма индексируемых сегментов.

5. Гибридное извлечение документов

Для повышения полноты и устойчивости retrieval-этапа в работе используется гибридный подход, сочетающий разреженный и плотный поиск. Такое объединение позволяет компенсировать ограничения каждого из методов: разреженный поиск эффективно обрабатывает точные лексические совпадения, в то время как плотный *retrieval* обеспечивает семантическое обобщение и устойчивость к перефразированию.

Разреженный retrieval реализуется на основе классической модели BM25, широко применяемой в задачах информационного поиска. [5, 6] Оценка релевантности документа d запросу q вычисляется как

$$BM25(q,d) = \sum_{t \in q} IDF(t) \cdot \frac{(f(t,d)(k1+1))}{f(t,d) + k_1(1 - b + b\frac{|d|}{avgdl})},$$

где f(t,d) - частота термина t в документе, |d| - длина документа, а k_1 и b - гиперпараметры модели.

В рамках retrieval-пайплайна BM25 используется на этапе широкого recall, формируя большое множество кандидатов. Такой подход позволяет надёжно извлекать фрагменты, содержащие редкие ключевые термины, числовые данные и именованные сущности, которые могут быть недостаточно представлены в плотном векторном пространстве. Для обработки запросов и документов применяется токенизация с приведением текста к нижнему регистру и фильтрацией по алфавитно-цифровым символам, что повышает устойчивость системы к морфологическим и орфографическим вариациям.

Плотный поиск *retrieval* основан на *bi-encoder* архитектуре, в которой запрос и документ кодируются независимо в общее векторное пространство:

$$score_{dense}(q,d) = cos\left(f_q(q), f_d(d)\right).$$

В экспериментах используется модель семейства E5, обученная в контрастивной постановке на задаче семантического соответствия запрос–документ. Для корректной интерпретации косинусного сходства эмбеддинги нормализуются по L2-норме. [7]

Плотный поиск применяется в двух режимах:

1. dense recall, формирующий множество кандидатов наравне с BM25;
2. dense reranking, при котором кандидаты переупорядочиваются на основе точного сходства с исходным запросом.

Такое двухэтапное использование плотного retrieval позволяет сохранить баланс между вычислительной эффективностью и качеством ранжирования, а также обеспечивает единое эмбеддинговое представление, используемое на последующих этапах отбора контекста.

Результаты разреженного и плотного поиска объединяются с использованием метода *Reciprocal Rank Fusion (RRF)*:

$$RRF(d) = i\sum_i \frac{1}{k + rank_i(d)}, \quad (5)$$

где $rank_i(d)$ - позиция документа в i-м ранкинге, а k — сглаживающий параметр. [8]

Использование *RRF* позволяет избежать необходимости явной нормализации оценок различных моделей и обеспечивает устойчивое объединение ранкингов с разной шкалой значений. На практике fusion-этап значительно снижает чувствительность retrieval-системы к ошибкам отдельного канала и стабилизирует полноту извлечения при различных типах запросов.

6. Entity-based fallback расширение запросов

Несмотря на использование гибридного *retrieval*, в ряде случаев базовый поиск оказывается недостаточным для извлечения релевантных документов. Для компенсации таких ситуаций в системе реализован *entity-based fallback* механизм, активируемый только при недостаточной полноте кандидатов. [9]

Если размер множества кандидатов $|C_q|$ после основного *retrieval* меньше заданного порога K, выполняется дополнительное расширение запроса. Из исходного запроса извлекается множество сущностей $E(q) = \{e_1, e_2, ..., e_L\}$, которые затем используются как отдельные поисковые запросы. Извлечение сущностей основано на комбинации NER-модели, POS- и dependency-фильтров и эвристик, ориентированных на retrieval-полезные концепты.

Важно отметить, что извлечение на основе именованных сущностей применяется только в качестве вспомогательной стратегии, а не основной процедуры поиска. Это помогает минимизировать включение нерелевантного контента и сохранять точность ранжирования. После добавления новых кандидатов выполняется повторная плотная сортировка, которая гарантирует согласованность итогового списка.

7. Coverage-aware отбор

На финальном этапе формирования контекста может применяться coverage-aware отбор, направленный на снижение избыточности и повышение семантического покрытия запроса. Отбор реализуется с использованием жадного алгоритма, максимизирующего функцию покрытия:

$$s(D) = \alpha cos\left(q, \bar{e}\right) + (1-\alpha) \max_i cos\left(q, e_i\right), \quad (6)$$

где $\bar{e}$ - средний эмбеддинг уже выбранных фрагментов, а e_i - эмбеддинги отдельных кандидатов.

Алгоритм добавляет фрагменты последовательно, выбирая те, которые максимально увеличивают семантическое покрытие, и прекращает работу при достижении заранее установленного порога или максимального числа чанков. Такой способ позволяет регулировать баланс между релевантностью и разнообразием извлекаемого контекста. Экспериментальные данные показывают, что чрезмерное расширение покрытия иногда приводит к уменьшению *Recall@k*, так как наиболее релевантные фрагменты могут быть вытеснены в пользу увеличения разнообразия. Поэтому *coverage-aware* отбор применяется только в случаях, когда разнообразие контекста имеет критическое значение.

8. Описание датасета

Экспериментальная оценка предложенных методов проводилась на русскоязычном датасете *RuBQ*, предназначенном для задач фактологического *question answering* и информационного поиска. Датасет основан на корпусе русскоязычной Википедии и содержит аннотацию релевантных документов для каждого вопроса, что делает его пригодным для оценки полноты retrieval-этапа в RAG-системах.

Корпус документов *RuBQ* включает 56 951 текстовый документ, каждый из которых представляет собой фрагмент статьи Википедии. Документы различаются по длине и тематике и содержат плотную фактологическую информацию, включая именованные сущности, числовые данные и исторические сведения. Пример структуры документов корпуса представлен на рис. 2.


```
{"doc_id": "56804", "title": "", "text": "29 июля 2011 года было заявлено о создании Свободной сирийской
 армии. Она возникла в результате перехода на сторону оппозиции 7 офицеров Сирийских арабских
вооружённых сил под руководством полковника Рияда Асада. Было выпущено видеообращение, призывавшее
сирийских военных переходить на сторону оппозиции."}
{"doc_id": "56805", "title": "", "text": "В начале ноября сирийская армия освободила город Эль-Тайба,
установив контроль над стратегической дорогой между Хамой и Идлибом."}
{"doc_id": "56806", "title": "", "text": "В январе 2018 года сирийские правительственные войска при
поддержке российских ВКС продолжали начавшееся в конце декабря 2017 наступление на контролируемые
вооружённой группировкой «Джебхат ан-Нусра» и рядом других радикальных группировок районы на юге
провинции Идлиб (северо-запад Сирии) и в провинции Хама."}
```


Рис. 2. Пример фрагмента корпуса документов *RuBQ*

Набор вопросов состоит из 8206 фактологических запросов на русском языке. Вопросы охватывают широкий спектр тем, включая географию, историю, культуру, науку и технику. Для каждого вопроса задан список релевантных документов корпуса, при этом одному запросу может соответствовать несколько корректных источников. Пример вопросов и их разметки приведён на рис. 3.

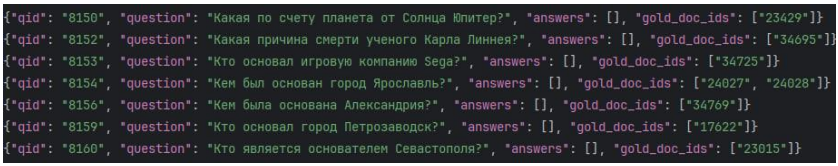


Рис. 3. Пример вопросов и релевантных документов в датасете *RuBQ*

В рамках данной работы эксперименты проводились без использования генеративной модели. Ответы на вопросы не использовались напрямую; оценка качества основывалась исключительно на наличии релевантных документов среди извлечённых кандидатов. Такая постановка позволяет изолированно анализировать влияние архитектурных и алгоритмических решений retrieval-этапа на показатели полноты извлечения.

Особенностью датасета *RuBQ* является высокая степень лексического совпадения между формулировками вопросов и релевантными документами, а также наличие нескольких релевантных источников для значительной части запросов. Эти свойства делают *Recall@k* и долю запросов без попаданий (*NO_HIT*) информативными метриками и позволяют объективно оценивать эффективность методов расширения и комбинирования retrieval-пайплайна.

9. Экспериментальная постановка

Экспериментальная оценка предложенных методов проводилась на датасете *RuBQ*, содержащем фактологические вопросы с аннотацией релевантных документов. В экспериментах использовалось подмножество из 338 вопросов, для которых была доступна полная и однозначная разметка релевантных документов. Целью экспериментов являлась оценка влияния архитектурных и алгоритмических решений retrieval-этапа на полноту извлечения релевантного контекста.

В качестве метрик использовались *Recall@k* и *MRR@k* при $k \in \{1,3,5,10,20\}$, а также доля запросов, для которых релевантные документы отсутствуют среди retrieved-кандидатов (*NO_HIT*). Все эксперименты выполнялись в *retrieval-only* режиме, без участия генеративной модели, что позволяло изолировать влияние retrieval-компонентов.

Оценка проводилась в абляционной постановке с последовательным добавлением компонентов: семантического чанкинга, *entity-based* расширения запросов, гибридного *dense–sparse* извлечения и *coverage-*

aware отбора. Это позволило количественно оценить вклад каждого механизма.

10. Анализ результатов

В ходе экспериментов основное внимание было сосредоточено на оценке полноты извлечения (*Recall@k*) и на доле запросов, для которых релевантные документы отсутствуют среди retrieved-кандидатов (*NO_HIT*).

Таблица 1 содержит агрегированные результаты для ключевых параметров работы системы. Для удобства представлены только наиболее значимые параметры и метрики, позволяющие наглядно оценить влияние каждого компонента на полноту извлечения.

Таблица

Влияние компонентов retrieval-пайплайна на качество извлечения (RuBQ, retrieval-only)

Алгоритм	Entity fallback	Hybrid retrieval	Query rewrites	Coverage-aware	Recall @5	Recall @10	Recall @20	NO HIT
BM25+ Dense	нет	нет	нет	нет	0.68	0.68	0.68	0.32
+Cross Encoder	нет	нет	нет	нет	0.64	0.64	0.64	0.36
+Entity fallback	да	нет	нет	нет	0.74	0.75	0.75	0.25
+ Hybrid retrieval	да	да	нет	нет	0.74	0.76	0.76	0.24
+ Query rewrites	да	да	да	нет	0.73	0.77	0.77	0.23
+Coverage - aware	да	да	нет	да	0.57	0.57	0.57	0.43

Базовая конфигурация демонстрирует ограниченную полноту извлечения, что особенно заметно при увеличении k. Добавление *cross-encoder reranking* без предварительного расширения поискового пространства приводит к ухудшению результатов, поскольку *reranker* не способен компенсировать ошибки первичного *retrieval*.

Введение *entity-based fallback* механизма обеспечивает устойчивый прирост *Recall@k* и заметное снижение доли *NO_HIT*. Это указывает на эффективность структурного расширения запроса за счёт именованных сущностей, особенно для вопросов, содержащих неявные или многословные упоминания.

Дополнительное объединение разреженного и плотного *retrieval* (*BM25 + dense embeddings*) приводит к дальнейшему росту *Recall@10–20*. Гибридная схема позволяет сочетать точное лексическое совпадение

с семантическим обобщением, повышая устойчивость retrieval-пайплайна к вариативности формулировок.

Эксперименты с *query rewriting* показали, что данный механизм даёт средний результат и не превосходит лучшие конфигурации без переформулирования запроса. В ряде случаев *rewrites* улучшают *Recall@1*, однако ухудшают показатели *Recall@3–10*, что свидетельствует о снижении стабильности ранжирования. По сравнению с *entity-based* расширением запросов, *rewrites* оказываются менее эффективными и не вносят дополнительного выигрыша в полноту извлечения.

Анализ полученных результатов позволяет сделать вывод, что специфика датасета *RuBQ* существенно влияет на эффективность *query rewriting*. Формулировки вопросов в данном корпусе, как правило, уже содержат ключевые термины, совпадающие с лексикой релевантных документов. В таких условиях переформулирование запроса не компенсирует лексическое расхождение, а напротив, может ослаблять точные совпадения, критичные для успешного *retrieval*.

Механизм *coverage-aware* отбора демонстрирует наихудшие результаты. Агрессивное поощрение разнообразия контекста приводит к вытеснению наиболее релевантных фрагментов при фиксированном размере финального контекста, что выражается в резком росте *NO_HIT* и снижении *Recall@k*.

В целом результаты подтверждают, что предложенный retrieval-пайплайн является эффективным, однако применимость отдельных компонентов существенно зависит от свойств целевого датасета. Для корпусов с выраженным лексическим совпадением между вопросами и документами приоритет следует отдавать структурным методам расширения поиска (*entity-based fallback*), а не генеративным стратегиям переформулирования запросов.

Заключение

В данной работе предложен и экспериментально исследован набор методов, направленных на повышение полноты retrieval-этапа в системах *Retrieval-Augmented Generation*. Показано, что архитектурные решения на уровне чанкинга, расширения запросов и гибридного извлечения оказывают существенное влияние на итоговое качество.

Эксперименты на датасете *RuBQ* демонстрируют, что наилучшие результаты достигаются при комбинации семантического чанкинга, *entity-based fallback* и гибридного *retrieval*, без жёсткого *coverage-aware* отбора. Такая конфигурация обеспечивает устойчивый рост *Recall@k* и значительное снижение доли запросов без релевантного контекста.

Полученные результаты подтверждают, что улучшение retrieval-этапа является эффективным и вычислительно экономичным способом повышения надёжности RAG-систем без усложнения генеративной модели. В дальнейшем планируется исследование адаптивных стратегий активации fallback-механизмов и автоматического подбора гиперпараметров чанкинга.

Список литературы

1. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks / P. Lewis [et al.] // arXiv preprint arXiv:2005.11401. — 2020. — URL: https://arxiv.org/abs/2005.11401
2. Dense Passage Retrieval for Open-Domain Question Answering / V. Karpukhin [et al.] // arXiv preprint arXiv:2004.04906. — 2020. — URL: https://arxiv.org/abs/2004.04906
3. Nogueira, R. Document Expansion by Query Prediction / R. Nogueira, K. Cho // arXiv preprint arXiv:1904.08375. — 2019. — URL: https://arxiv.org/abs/1904.08375
4. Reimers, N. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks / N. Reimers, I. Gurevych // arXiv preprint arXiv:1908.10084. — 2019. — URL: https://arxiv.org/abs/1908.10084
5. Robertson, S. E. The Probabilistic Relevance Framework: BM25 and Beyond / S. E. Robertson, H. Zaragoza // Foundations and Trends in Information Retrieval. — 2009. — Vol. 3, № 4. — P. 333–389.
6. Manning, C. D. Introduction to Information Retrieval / C. D. Manning, P. Raghavan, H. Schütze. — Cambridge : Cambridge University Press, 2008. — 496 p.
7. BGE: Improving Embeddings with Weakly-Supervised Contrastive Pre-training / L. Wang [et al.] // arXiv preprint arXiv: 2212.03533. — 2022. — URL: https://arxiv.org/pdf/2212.03533
8. Cormack, G. V. Reciprocal Rank Fusion Outperforms Condorcet and Individual Rank Learning Methods / G. V. Cormack, C. L. A. Clarke, S. Buettcher // Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '09). — Boston : ACM, 2009. — P. 758–759.
9. Scalable Zero-shot Entity Linking with Dense Entity Retrieval / L. Wu [et al.] // arXiv preprint arXiv:1911.03814. — 2019. — URL: https://arxiv.org/abs/1911.03814

Микросервисная реализация остановочного узла субавтономной ИТС для прогнозирования времени прибытия общественного транспортного средства

С. М. Ялвач, email: yalvach.murat@yandex.ru
В. А. Каладзе, email: wakaladze@yandex.ru

Воронежский государственный университет

Аннотация. *В данной работе описана система остановочного узла субавтономной ИТС с функцией прогноза времени прибытия общественного транспортного средства (ТС), реализованная в виде набора микросервисов, выполняющих приём телеметрии от ТС, робастную валидацию, оценку времени прибытия и передачу прогнозов на последующие остановочные павильоны. Представлены: состав сервисов, принципы взаимодействия и упрощённая физическая модель локальной БД остановочного узла.*

Ключевые слова: *ИТС; ETA; остановочный узел; edge/fog; микросервисная архитектура; Docker; локальная база данных.*

Введение

Задача информирования пассажиров о времени прибытия общественного транспорта (ETA) [1] в городской среде осложняется нестабильностью GNSS-позиционирования в условиях плотной застройки, неоднородности покрытия GSM/LTE [2], либо периодическими отключениями необходимой связи в некоторых регионах РФ на текущий момент. При централизованной (облачной) обработке телеметрии качество сервиса существенно зависит от сотовой связи и задержек доставки данных, что приводит к запаздыванию и не качественному отображению прогнозов [3]. В данной работе применён субавтономный подход: основные вычисления происходят в fog/edge-узлах [4] остановочных павильонов. При приближении транспортного средства к остановочному павильону, остановочные узлы – компактные локальные сервера, взаимодействуют с ТС по беспроводному каналу. Помимо этого, происходит передача агрегированных сообщений (в том числе и прогноза) далее по маршруту к следующим остановочным узлам с использованием существующих городских проводным коммуникаций.

Цель статьи – представить состав и функции остановочного узла через описание набора микросервисов, локальной модели данных и

© Ялвач С. М., Каладзе В. А., 2026

форматы входных/выходных сообщений, обеспечивающих устойчивое обновление ETA при шумных измерениях и возможных разрывах связи.

1. Программная реализация микросервисов остановочного узла

Остановочный узел реализован, как набор контейнеров Docker, образующих последовательный конвейер обработки и отдельный контур межузловой передачи. Приём данных выполняет входной сервис (DataIngest), который принимает телеметрию от транспортного средства и приводит её к унифицированному виду (нормализация, проверка целостности, проверка по меткам времени). Далее данные поступают в сервис робастной валидации (NoiseFilter), где вычисляется вес доверия наблюдения на основе фильтра Хампеля [5]. Вес доверия позволяет снижать влияние единичных выбросов без жёсткого отбраковывания информации.

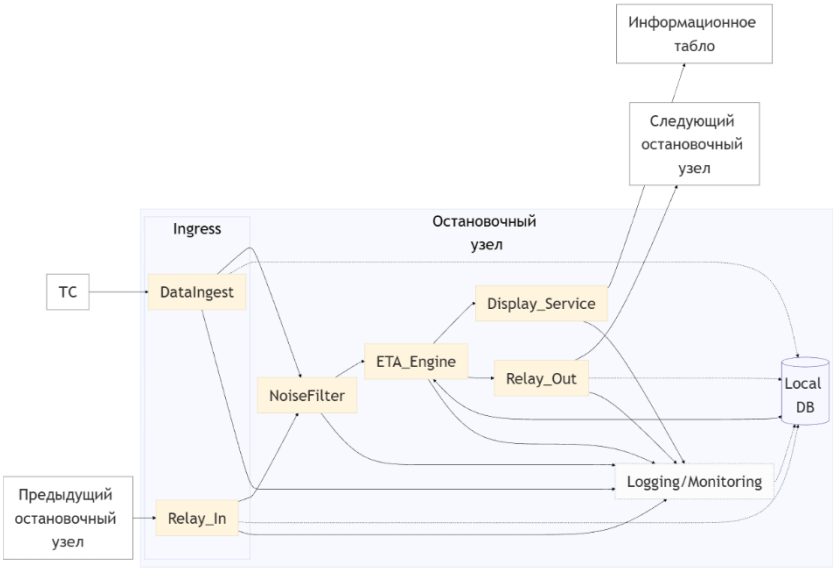


Рис. 1. Микросервисная архитектура остановочного узла в контейнерной среде Docker

Расчёт прогноза выполняется сервисом ETA_Engine, который сочетает исторические профили движения по сегментам маршрута (локально хранимые статистики) и коррекцию в реальном времени по поступающим событиям. На выходе формируется прогноз прибытия на текущую и несколько следующих остановок; дополнительно

вычисляется индикатор качества прогноза, используемый для интерфейса табло и ретрансляции. Отображение для пассажиров обеспечивает Display_Service: он преобразует прогноз в компактный формат представления и передаёт его на контроллер информационного табло. Межостановочный обмен реализован парой сервисов Relay_In/Relay_Out, которые принимают и отправляют агрегированные пакеты следующему узлу по маршруту через существующий проводной канал (Ethernet/ADSL и др.). Наблюдаемость обеспечивается Logging/Monitoring (сбор логов и метрик задержек/качества данных), что критично для эксплуатации распределённой инфраструктуры.

2. Локальная база данных остановочного узла

Локальная БД узла предназначена для обеспечения автономности и устойчивости расчёта ETA. Она хранит:

1. граф маршрута в виде набора остановок и сегментов,
2. исторические статистические профили сегментов,
3. информацию о ситуации в текущий момент, обновляемую по мере поступления валидных событий.

В результате ETA_Engine не зависит от постоянного доступа к внешним источникам и способен восстанавливать работу после перезапуска контейнеров, опираясь на локальные данные.

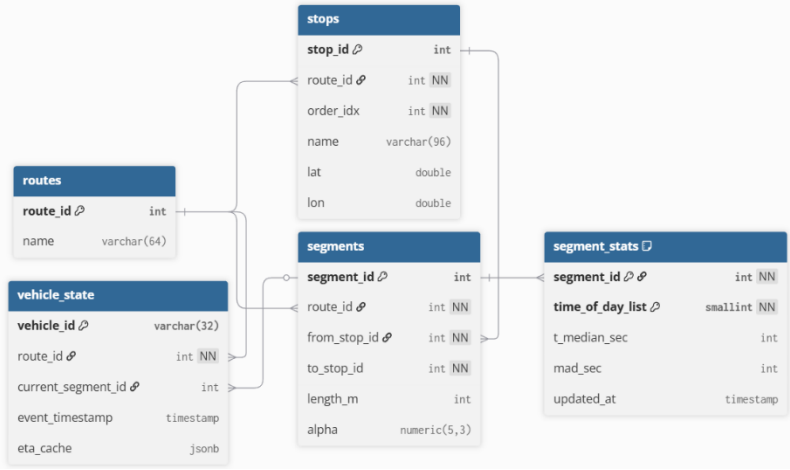


Рис. 2. Фрагмент логической модели локальной БД остановочного узла.

На рисунке показан фрагмент логической модели БД, отражающий только ключевые сущности и связи. Поле to_stop_id у сегмента является ссылкой на таблицу остановок аналогично from_stop_id и в диаграмме может быть показано без дублирования графических связей (для повышения читабельности рисунка).

3. Векторы и потоки данных на остановочном узле

В рамках предложенной архитектуры используются два базовых класса сообщений: входная телеметрия от транспортного средства [6] и агрегированный пакет данных для передачи прогнозов далее по маршруту. Некоторые параметры, принимаемые от ТС, основаны на работоспособности GNSS/GSM сервисов и не обязательны [7]. Они могут использоваться в дополнительном web-микросервисе для отображения информации на легковесной web-странице, доступ к которой возможен по QR-коду на табло остановочного павильона. Ниже приведены (Таблицы 1 и 2) рекомендуемые поля сообщений и их семантика.

Таблица 1

Входные данные от транспортного средства

Параметр	Описание
vehicle_id	Идентификатор транспортного средства (возможен псевдоним/хеш)
route_id	Идентификатор/номер маршрута
timestamp	Метка времени формирования сообщения на ТС
gnss_lat	Широта GNSS (опционально, если доступно)
gnss_lon	Долгота GNSS (опционально, если доступно)
speed_mps	Скорость (опционально)
occupancy	Количество пассажиров в салоне ТС (опционально, если доступно технически)
auth_tag	Контроль целостности/аутентификации сообщения

Таблица 2

Данные, передаваемые на последующие остановочные узлы

Параметр	Описание
route_id	Идентификатор маршрута
vehicle_id	Идентификатор транспортного средства

Продолжение табл. 2

from_stop_id	Идентификатор узла-источника (текущая остановка)
event_timestamp	Время последнего валидного события, на котором основан прогноз
eta_list	Прогноз для K ближайших остановок: набор {stop_id, eta_sec, eta_low_sec, eta_high_sec}
ttl	Ограничение дальности распространения пакета (защита от зацикливания)
sign	Подпись/контроль целостности

В рабочем контуре данные движутся следующим образом:

1. телеметрия ТС принимается DataIngest и валидируется NoiseFilter;
2. ETA_Engine обновляет vehicle_state и формирует прогноз на ближайшие остановки;
3. Display_Service выдаёт компактное представление на табло;
4. Relay_Out отправляет агрегированный пакет следующему узлу, где Relay_In может использовать его как предварительное состояние при отсутствии свежих данных от ТС.

Заключение

Представлен вариант остановочного узла субавтономной ИТС, выполненный в микросервисной архитектуре и ориентированный на устойчивую работу в условиях шумного GNSS и нестабильной сотовой связи. Показан состав контейнерных сервисов, а также локальная модель данных, обеспечивающая автономность расчёта ETA и восстановление после перезапуска. Зафиксированы параметры входных сообщений от ТС и межостановочных пакетов, что является необходимым условием взаимодействия узлов и масштабирования решения на новые маршруты. Дальнейшая работа связана с симуляцией работы системы в эмулированном окружении, калибровкой параметров робастной валидации и уточнением статистических профилей сегментов по накопленным данным эксплуатации.

Список литературы

1. Abdi, A. A review of travel and arrival-time prediction methods on road networks: classification, challenges and opportunities / A. Abdi, C. Amrit

// PeerJ Computer Science. – 2021. – Vol. 7. – e689. – DOI: 10.7717/peerj-cs.689.

2. ГОСТ Р 59289-2020. Глобальная навигационная спутниковая система на транспорте. Технические средства контроля на транспорте. Единый расширяемый набор протоколов обмена данными технических средств контроля с информационными системами. – Введ. 2021-06-01. – М.: Стандартинформ, 2021. – 62 с.

3. Singh, N. A review of bus arrival time prediction using artificial intelligence / N. Singh, K. Kumar // Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. – 2022. – Vol. 12. – e1457. – DOI: 10.1002/widm.1457.

4. Zhou, X. When Intelligent Transportation Systems Sensing Meets Edge Computing: Vision and Challenges / X. Zhou, R. Ke, H. Yang, C. Liu // Applied Sciences. – 2021. – Vol. 11, № 20. – 9680. – DOI: 10.3390/app11209680.

5. Pearson, R. K. Generalized Hampel Filters / R. K. Pearson, Y. Neuvo, J. Astola, M. Gabbouj // EURASIP Journal on Advances in Signal Processing. – 2016. – Vol. 2016, article 87. – DOI: 10.1186/s13634-016-0383-6.

6. ГОСТ Р 56829-2015. Интеллектуальные транспортные системы. Термины и определения. – Введ. 2016-06-01. – М.: Стандартинформ, 2018 – 16 с.

7. ГОСТ Р 56294-2014. Интеллектуальные транспортные системы. Требования к функциональной и физической архитектурам интеллектуальных транспортных систем. – Введ. 2015-07-01. – М.: Стандартинформ, 2018. – 12 с.

Секция 4
Информационные системы и базы данных

Сравнительный анализ SQL и DL-Query при обработке данных юридической предметной области

А. А. Борисова, email: borisova_a_a@sc.vsu.ru

ФГБОУ ВО «Воронежский государственный университет»

Аннотация. *В правовых информационных системах широко применяются реляционные базы данных и SQL, однако сложные семантические связи и изменяемость законодательства затрудняют полноформатное представление знаний в рамках реляционной модели. В работе выполнен сравнительный анализ SQL и DL-Query на примере предметной области интеллектуальных прав. Рассмотрена онтологическая модель данных на базе OWL, использующая дескриптивную логику и механизм логического вывода, позволяющий получать выводимые факты при неполной информации. Показано, что DL-Query эффективен для семантической классификации и выявления скрытых зависимостей, тогда как в SQL аналогичные результаты требуют явного хранения производных сведений или дополнительной бизнес-логики. Обоснована целесообразность гибридной архитектуры для правовых ИС и систем поддержки принятия решений.*

Ключевые слова: *реляционная база данных, SQL, онтология, OWL, дескриптивная логика, DL-Query, логический вывод, интеллектуальные права.*

Введение

Правовая сфера относится к числу наиболее сложных предметных областей с точки зрения представления и обработки информации в информационных системах. Динамика законодательства, наличие большого числа взаимосвязанных нормативных положений, неоднозначность терминологии и контекстная зависимость правовых понятий приводят к тому, что поддержка поиска и анализа юридической информации выходит за рамки хранения «фактов» и требует моделирования смысловых отношений между сущностями предметной области. Дополнительным фактором является цифровая трансформация, в результате которой значимая часть правовых процессов и объектов переносится в цифровую среду, увеличивая требования к масштабируемости, качеству поиска и интеллектуальной обработке данных.

© Борисова А. А., 2026

В практике разработки правовых информационных систем преобладает использование баз данных как инструмента хранения и управления структурированными данными (законы, судебные акты, сведения об объектах и субъектах права и др.). Реляционная модель и язык SQL обеспечивают высокую производительность операций выборки и модификации данных, поддержку ограничений целостности и стандартизированные средства манипулирования данными. Вместе с тем реляционный подход ориентирован преимущественно на работу с явно сохранённой информацией и недостаточно приспособлен к выражению сложных правовых конструкций и семантических зависимостей без переноса существенной части логики на уровень прикладного кода или специализированных процедур обработки.

Альтернативным направлением представления данных является онтологический подход, в котором знания описываются как система понятий, отношений и ограничений, формализованная на языках RDF/OWL и поддерживающая логический вывод на основе дескриптивной логики. Онтологии позволяют фиксировать не только структуру данных, но и их семантику: вводить аксиомы, ограничения (включая непересекаемость классов), характеристики свойств (например, транзитивность и инверсивность) и тем самым получать выводимые факты, не заданные явно. Для извлечения знаний из онтологии применяются запросы DL-Query, ориентированные на поиск классов объектов и их логических описаний с учётом выводимости. Такой подход особенно актуален для правовой области, где важно выявление скрытых связей и корректная интерпретация неполной, но непротиворечивой информации.

Целью настоящей работы является сравнительный анализ возможностей SQL и DL-Query при обработке данных юридической предметной области на примере интеллектуальных прав. Для достижения поставленной цели решаются следующие задачи: 1) кратко охарактеризовать особенности реляционных запросов SQL в контексте правовых информационных систем; 2) описать онтологическую модель данных и принципы логического вывода, влияющие на результаты запросов DL-Query; 3) выделить типовые классы запросов и сопоставить их реализуемость средствами SQL и DL-Query; 4) определить области рационального применения каждого подхода и обосновать перспективность гибридной архитектуры для правовых ИС.

Научная значимость работы заключается в систематизации различий между выборкой данных из реляционной модели и извлечением семантических знаний из онтологии, а также в формулировании практических критериев выбора языка запросов в зависимости от

характера задач (поиск фактов, классификация, вывод скрытых зависимостей). Рассматриваются теоретические основания сравнения, приводятся примеры запросов и анализируются результаты, демонстрирующие принципиальные различия между SQL и DL-Query при работе с юридическими данными.

1. Реляционная модель данных и язык SQL в правовых информационных системах

Реляционные базы данных на протяжении нескольких десятилетий остаются базовой технологией построения информационных систем различного назначения, включая правовые и справочно-правовые системы. Их широкое распространение обусловлено формальной строгостью реляционной модели, развитым аппаратом обеспечения целостности данных, а также наличием стандартизированного языка запросов SQL, позволяющего эффективно выполнять операции выборки, модификации и агрегации данных. В правовой сфере реляционные базы данных традиционно используются для хранения нормативно-правовых актов, судебной практики, сведений об объектах и субъектах права, а также связанных с ними атрибутов и классификаторов.

Реляционная модель предполагает представление данных в виде таблиц, связанных между собой с помощью первичных и внешних ключей. Такая структура хорошо подходит для хранения строго формализованных и однозначно определяемых фактов, например, дат принятия нормативных актов, номеров статей, реквизитов правообладателей или сроков действия исключительных прав. Язык SQL предоставляет средства для фильтрации, соединения таблиц и агрегации данных, что делает его удобным инструментом для реализации запросов отчетного и справочного характера. Типичный SQL-запрос ориентирован на получение информации, которая уже явно зафиксирована в базе данных и однозначно соответствует условиям запроса [1].

Вместе с тем реляционный подход опирается на так называемую предпосылку замкнутого мира (Closed World Assumption), согласно которой считается, что если некоторый факт отсутствует в базе данных, то он является ложным. Такая предпосылка корректна для многих прикладных задач, однако в правовой предметной области она может приводить к методологическим ограничениям. Юридические данные нередко являются неполными, контекстно зависимыми и требуют интерпретации на основе совокупности норм и связей, а не только на основании явно заданных записей. В условиях замкнутого мира SQL-запросы не позволяют корректно работать с ситуациями, когда вывод о наличии или отсутствии правового свойства должен быть сделан логически, а не на основе прямого указания в таблице.

Существенным ограничением SQL является и отсутствие встроенных средств логического вывода. Все зависимости между сущностями предметной области должны быть либо явно отражены в структуре базы данных, либо реализованы на уровне прикладной логики в виде процедур, триггеров или внешних программных модулей [2]. В правовых информационных системах это приводит к усложнению архитектуры, поскольку значительная часть семантики предметной области переносится за пределы модели данных. Например, такой объект интеллектуальной собственности как изобретение относится к разделу патентных прав, у него определенный набор признаков и срок исключительных прав, отличающихся от других объектов патентных прав. Все эти нормы и признаки хранятся часто в реляционной модели разрозненно и не связаны формальными правилами вывода.

Еще одной особенностью SQL является его ориентация преимущественно на манипулирование данными, а не на работу с концептуальными моделями. Несмотря на наличие средств нормализации и ограничений целостности, реляционная модель не поддерживает явного описания иерархий понятий, семантической непересекаемости классов или свойств отношений, таких как транзитивность или инверсивность. Для реализации подобных зависимостей в SQL требуется либо дублирование данных, либо создание вспомогательных таблиц и сложных запросов, что снижает прозрачность модели и затрудняет ее сопровождение при изменении законодательства или предметной области.

Таким образом, реляционные базы данных и язык SQL являются эффективным инструментом хранения и обработки структурированных правовых данных, однако их применение в задачах семантического анализа и логической интерпретации юридической информации имеет принципиальные ограничения. Эти ограничения становятся особенно заметны при работе с неполными данными, выявления скрытых зависимостей и автоматической классификации объектов правовой предметной области. Указанные особенности обусловливают интерес к альтернативным моделям представления данных и языкам запросов, ориентированным на работу со знаниями и логическим выводом, что и послужило основанием для дальнейшего рассмотрения онтологического подхода и языка DL-Query.

2. Онтологическая модель данных и язык DL-Query в задачах обработки юридической информации

Онтологический подход к представлению данных рассматривает предметную область не только как совокупность фактов, но и как систему понятий, отношений и формальных ограничений, определяющих

семантику этих фактов [3]. В отличие от реляционной модели, ориентированной на хранение структурированных записей, онтология описывает знания на концептуальном уровне и позволяет формализовать смысловые связи между сущностями предметной области. В правовых информационных системах такой подход особенно актуален, поскольку юридические понятия обладают сложной иерархической структурой, а их интерпретация часто зависит от совокупности признаков и норм законодательства.

В основе онтологической модели данных лежит использование формальных языков представления знаний, таких как RDF и OWL, а также аппарата дескриптивной логики. Онтология задаёт классы (понятия предметной области), свойства (отношения между понятиями и их экземплярами), а также аксиомы и ограничения, описывающие допустимые комбинации и характеристики этих элементов. Онтологический подход основан на концепции «открытого мира» (Open World Assumption), которая при отсутствии какого-либо факта не означает его ложность. Такой подход «достраивает» неполную информацию на основе логических правил, что является типичной ситуацией для правовой предметной области. Онтологическое проектирование построено на таких отношениях между экземплярами классов как транзитивность, инверсность, функциональность, симметричность и за счет механизма логического вывода (reasoner) автоматически выводит факты, которые не были заложены в базу знаний, но логически выведены за счет прописывания дескриптивных логических правил [4].

В реляционных базах данных аналогичное поведение может быть реализовано только с помощью дополнительных запросов, триггеров или программного кода, что усложняет архитектуру системы и снижает прозрачность модели.

Для извлечения информации из онтологической модели применяется язык DL-Query, ориентированный на формирование запросов к классам и их логическим описаниям. В отличие от SQL, который оперирует строками таблиц и их атрибутами, DL-Query работает с понятиями, отношениями и логическими конструкторами дескриптивной логики. Запросы DL-Query позволяют получать не только явно заданные экземпляры классов, но и те объекты, которые удовлетворяют заданным условиям логически, с учётом аксиом и ограничений онтологии. Таким образом, результат запроса может включать выводимые (неявные) знания, отсутствующие в явном виде в хранилище.

В контексте правовой предметной области это означает возможность автоматической классификации объектов интеллектуальной

собственности, определения их правового статуса и выявления нормативных связей без явного перечисления всех условий в запросе. Например, объект может быть отнесён к определённой категории прав на основании совокупности признаков, заданных в онтологии, даже если соответствующая классификация не была явно сохранена как атрибут. Аналогичный результат в SQL требует либо денормализации данных, либо выполнения серии вложенных запросов и проверок, что затрудняет сопровождение системы при изменении логики предметной области.

Существенным преимуществом DL-Query является также его ориентация на семантическую целостность данных. Проверка непротиворечивости онтологии и корректности классификации объектов осуществляется автоматически с помощью reasoner, что позволяет выявлять логические ошибки в модели данных. В реляционных базах данных контроль целостности, как правило, ограничивается синтаксическими и структурными ограничениями (ключи, типы данных, ограничения ссылочной целостности) и не охватывает семантический уровень предметной области.

Таким образом, онтологическая модель данных и язык DL-Query обеспечивают принципиально иной уровень работы с юридической информацией по сравнению с SQL. Если SQL ориентирован на эффективную обработку явно заданных данных, то DL-Query позволяет извлекать знания с учётом семантики предметной области и логического вывода. Эти различия создают основу для их сопоставления по ряду критериев, включая модель данных, тип обрабатываемых знаний, работу с неполной информацией и возможности автоматического вывода [5].

3. Сравнительный анализ SQL и DL-Query при обработке юридических данных

Для сопоставления возможностей SQL и DL-Query целесообразно рассматривать их не только как разные синтаксисы запросов, но как инструменты, опирающиеся на различные модели данных и различные предпосылки интерпретации знаний. SQL ориентирован на выборку и манипулирование явно сохранёнными данными в рамках реляционной модели, тогда как DL-Query использует онтологическую модель и результаты логического вывода, опираясь на аксиомы и свойства отношений. Это различие принципиально влияет на набор решаемых задач и на архитектурные решения правовых информационных систем. Сравнительные характеристики SQL и DL-Query представлены в табл. 1.

Таблица 1

Сравнительная характеристика SQL и DL-Query

Критерий	**SQL (реляционные БД)**	**DL-Query (онтология OWL)**
Модель данных	Таблицы, атрибуты, ключи, связи (JOIN)	Классы, индивиды, свойства, аксиомы
Тип обрабатываемых знаний	Явные факты (то, что записано в таблицах)	Явные + выводимые (неявные) факты
Предпосылка о полноте данных	Замкнутый мир: отсутствие факта трактуется как «ложь»	Открытый мир: отсутствие факта не означает «ложь»
Целостность	Структурная/ссылочная (ключи, ограничения, типы данных)	Семантическая (аксиомы, ограничения, проверка reasoner)
Семантические свойства отношений	Не поддерживаются декларативно (только через код/запросы)	Декларативно задаются (транзитивность, инверсивность, симметрия)
Работа с иерархиями понятий	Требует отдельной схемы (таблицы и рекурсивные запросы)	Встроена (subClassOf), учитывается в выводе
Извлечение «скрытых знаний»	Только через явное хранение или бизнес-логику приложения	Встроено: reasoner выводит новые факты из аксиом
Типичные области применения	Хранение, поиск, отчёты, транзакции	Классификация, семантический поиск, анализ неполных данных
Стоимость сопровождения при изменении правил	Переписывание запросов/процедур/кода	Изменение аксиом/ограничений онтологии, пересчёт вывода

Таким образом, SQL и DL-Query решают разные классы задач, и в юридической предметной области, насыщенной правилами и исключениями, преимущества онтологического подхода проявляются именно там, где требуется не просто поиск записей, а интерпретация и вывод [6].

4. Сопоставление возможностей на типовых классах запросов

Класс 1. Запросы к явно хранимым фактам (справочные запросы). Если задача сводится к получению информации, прямо зафиксированной

в хранилище (например, перечень объектов, даты, реквизиты, номера статей, сроки), SQL обеспечивает простую реализацию и высокую эффективность. Аналогичные результаты достижимы и в онтологической модели, однако в таком случае DL-Query не даёт принципиального выигрыша: он будет извлекать индивиды или классы, уже явно описанные в онтологии. Следовательно, для задач «каталожного» поиска реляционный подход остаётся рациональным выбором.

Класс 2. Запросы к иерархиям и классификация объектов. В правовой сфере значительная часть информации естественно представляется в виде иерархий: классы объектов интеллектуальной собственности, виды прав, признаки охраноспособности. В онтологии иерархия классов является базовой конструкцией (отношение подкласс – надкласс), поэтому DL-Query позволяет извлекать объекты с учётом наследования автоматически. В реляционной модели иерархии требуют дополнительного проектирования (например, таблицы классификаторов, рекурсивные связи), а запросы к ним – рекурсивных либо специализированных процедур. Это делает SQL-решение более трудоёмким, особенно при усложнении структуры понятий.

Как показано в табл. 2, DL-Query позволяет извлекать не только явно заданные, но и логически выводимые знания, что принципиально отличает его от SQL.

Таблица 2

Сравнение типового запроса в SQL и DL-Query

Критерий сравнения	**SQL (реляционная база данных)**	**DL-Query (онтологическая модель данных)**
Цель запроса	Получение объектов, охраняемых законом	Получение объектов, охраняемых законом
Формулировка запроса	SELECT object FROM Objects WHERE protected = true;	Объект and охраняется_законом some owl:Thing
Источник результата	Значение явно заданного атрибута protected	Явные факты и результаты логического вывода
Учет иерархий понятий	Не учитывается автоматически; требует дополнительных соединений таблиц	Учитывается автоматически через иерархию классов
Учет транзитивных связей	Не поддерживается декларативно; требуется рекурсивная логика	Поддерживается за счёт транзитивных свойств

Работа с неполными данными	Ограничена предпосылкой замкнутого мира	Поддерживается предпосылкой открытого мира
Необходимость дополнитель ной бизнес-логики	Требуется для вывода производных фактов	Не требуется; вывод выполняется reasoner
Возможность извлечения скрытых знаний	Отсутствует	Поддерживается

Класс 3. Запросы, требующие учёта семантических свойств отношений. Для юридической предметной области характерны отношения, имеющие свойства транзитивности и инверсивности. В разработанной онтологии такие свойства заданы для ряда отношений (например, охрана законом, регистрация, «не охраняется законом»), что позволяет автоматически расширять граф связей и получать выводимые зависимости [7]. В SQL подобные эффекты не поддерживаются декларативно: транзитивность приходится реализовывать рекурсивными запросами или заранее рассчитывать транзитивное замыкание и хранить его отдельно; инверсивность – поддерживать дублированием данных или дополнительными запросами. Это увеличивает риск несогласованности и повышает стоимость сопровождения.

Класс 4. Запросы на выявление выводимых (неявных) знаний. Наиболее существенное различие проявляется в задачах извлечения скрытых знаний, когда результат должен быть получен не прямой выборкой, а логическим выводом. В рамках онтологии такие задачи решаются через reasoner: на основе аксиом, ограничений и свойств отношений система выводит новые утверждения о принадлежности объектов к классам и о наличии связей между индивидами. В рамках проведенного исследования показаны примеры, когда при задании ограниченного набора фактов выводятся дополнительные сведения: например, при описании коммерческого обозначения организации автоматически устанавливаются связанные свойства и нормативные привязки; при описании объекта «База данных» выводятся охрана законом и срок исключительных прав; а при указании нормы, исключающей охраноспособность, система выводит перечень сущностей, «не охраняемых законом» в соответствии с заданными правилами.

Для SQL аналогичные результаты не являются «естественной» операцией: они требуют либо хранения всех выводимых атрибутов в явном виде (что ведёт к избыточности и необходимости синхронизации), либо реализации логики вывода в прикладном слое (правила, цепочки проверок, экспертные модули). Тем самым DL-Query демонстрирует

преимущество в задачах, где ценность представляют не только факты, но и их интерпретация.

Таким образом, как показано на рисунке 1, в отличие от SQL, ориентированного на выборку явно заданных данных, DL-Query использует механизм логического вывода, позволяющий получать выводимые знания.

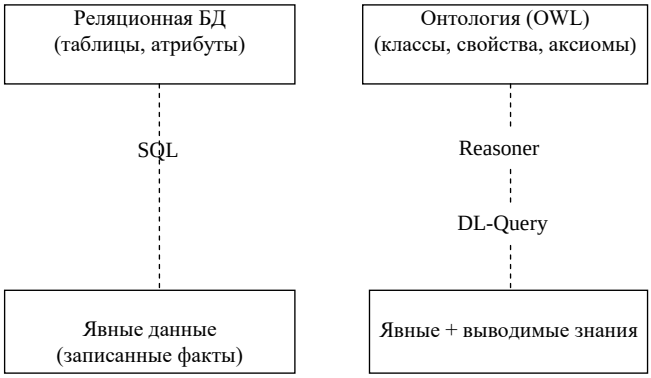


Рис. 1. Концептуальное различие обработки данных при использовании SQL и DL-Query

Исходя из вышеизложенного, SQL и DL-Query являются инструментами для разных уровней работы с данными. SQL наиболее эффективен в задачах хранения и обработки явно заданной информации, где требуется высокая скорость и транзакционность. DL-Query, опирающийся на онтологическую модель и логический вывод, оказывается более результативным при анализе юридической предметной области, включающей сложные связи, иерархии и необходимость извлечения неявных знаний при неполной информации. Следовательно, в правовых информационных системах перспективным является гибридный подход, при котором реляционная база данных используется как средство надёжного хранения и поиска, а онтология – как семантический слой для классификации и поддержки принятия решений.

5. Области применимости SQL и DL-Query и гибридный подход к построению правовых ИС

Результаты сравнительного анализа показывают, что SQL и DL-Query не являются взаимозаменяемыми инструментами, а ориентированы на решение различных классов задач в рамках

информационных систем. Использование исключительно реляционного или исключительно онтологического подхода в правовых информационных системах не всегда является оптимальным решением, особенно при необходимости сочетания высокой производительности хранения данных и интеллектуального анализа предметной области.

Такие возможности востребованы в системах поддержки принятия решений, экспертных и рекомендательных правовых системах, а также в интеллектуальных справочно-правовых сервисах [8].

С учётом выявленных особенностей перспективным представляется гибридный подход к построению правовых информационных систем, при котором реляционная база данных используется в качестве слоя хранения и первичного доступа к данным, а онтология выполняет роль семантического слоя, обеспечивающего интерпретацию, классификацию и логический вывод. В рамках такой архитектуры SQL и DL-Query не конкурируют, а дополняют друг друга: SQL обеспечивает эффективную работу с фактами, а DL-Query — извлечение знаний и семантический анализ. Подобный подход позволяет повысить гибкость системы, упростить сопровождение при изменении нормативной базы и сократить объём прикладной бизнес-логики.

Заключение

В работе выполнен сравнительный анализ языков запросов SQL и DL-Query в контексте обработки данных юридической предметной области на примере прав интеллектуальной собственности. Показано, что различия между данными инструментами обусловлены не только синтаксисом запросов, но и фундаментальными отличиями используемых моделей данных и предпосылок интерпретации информации.

Установлено, что SQL, опирающийся на реляционную модель и предпосылку замкнутого мира, эффективен при работе с явно заданными фактами и хорошо подходит для задач хранения, поиска и отчётности. Однако он ограничен в возможностях семантического анализа и не поддерживает логический вывод без переноса существенной части правил на уровень прикладной логики. В противоположность этому, DL-Query, основанный на онтологической модели и дескриптивной логике, позволяет извлекать как явные, так и выводимые знания, работать с неполной информацией и учитывать сложные семантические свойства отношений.

На примере типовых классов запросов показано, что DL-Query поддерживает задачи автоматической классификации и выявления скрытых зависимостей, которые в реляционных базах данных требуют дополнительных архитектурных решений и усложняют сопровождение

системы. В результате обоснована целесообразность использования гибридной архитектуры правовых информационных систем, сочетающей преимущества реляционных баз данных и онтологического подхода. Полученные выводы могут быть использованы при проектировании интеллектуальных правовых ИС и систем поддержки принятия решений, ориентированных на обработку сложных и семантически насыщенных предметных областей.

Список литературы

1. Калиниченко, Л. А. Эффективная поддержка баз данных с онтологическими зависимостями: реляционные языки вместо дескриптивных логик / Л. А. Калиниченко // Программирование. – 2012. – Т. 38. – № 6. – С. 45-62.
2. Борисова, А. А. Использование онтологии и баз данных в сфере трудового права / А. А. Борисова // Информатика: проблемы, методы, технологии : Материалы XXIV Междун. научно-практической конф. им. Э.К. Алгазинова (Воронеж, 14–15 февраля 2024 г.). – Воронеж, 2024. – С. 835-840.
3. Лапшин, В. А. Онтологии в информационных системах / В. А. Лапшин. – М. : Научный мир, 2020. – 224 с.
4. Понкин, И. В. Цифровые онтологии права и цифровое правовое пространство / И. В. Понкин, А. И. Редькина // Пермский юридический альманах. – 2019. – № 2. – С. 24–37.
5. Сычёв, О. А. Метод разработки интеллектуальных тренажёров на основе онтологии предметной области / О. А. Сычев, Н. А. Пенской, Г. В. Терехов // Онтологическое проектирование. – Т. 15. – 2025. – № 1. – С. 67-81.
6. Борисова, А. А. Разработка и реализация онтологии в сфере интеллектуальной собственности / А. А. Борисова, В. В. Гаршина // Труды молодых учёных факультета компьютерных наук ВГУ : Сб. ст. – Воронеж, 2024. – С. 323-328.
7. Вахитов, А. Р. Преимущества дескриптивной логики при обработке знаний / А. Р. Вахитов, В. Б. Новосельцев // Известия Томского политехнического университета. Инжиниринг георесурсов. –2008. – № 5. – С. 73-76.
8. Валькман, Ю. Р. Проектирование онтологии для правовой предметной области на основе текстового контента с использованием нечеткой логики / Ю. Р. Валькман, Е. А. Хала // Онтология проектирования. – 2014. – № 2(12). – С. 19–37.

Внедрение BI-системы в малый бизнес: принципы, архитектура и практические рекомендации

Н. А. Варданян, e-mail: varnarek@mail.ru

ФГБОУ ВО «Санкт-Петербургский государственный экономический университет»

Аннотация. *В статье рассматриваются ключевые принципы, архитектурные решения и практические рекомендации по внедрению BI-систем в малом бизнесе.*

Ключевые слова: *цифровая трансформация, малый бизнес, интеграция систем, отчетность, облачная аналитика, бизнес-процессы, BI-системы, облачные технологии.*

Введение

Современный малый бизнес сталкивается с увеличением объёма данных и необходимостью повышения операционной эффективности. Однако не каждая IT-платформа, на которой работает определенный бизнес способна к структуризации и визуализации определенных данных для анализа текущего положения бизнеса. С увеличением объёма продаж и закупок увеличиваются и данные, требуемые для тщательного анализа. Для решения проблем с визуализацией данных существуют множество BI-систем, которые позволяют визуализировать данные.

BI-система это решение, которое обеспечивает сбор, хранение, анализ и визуализацию данных компании. Для малого бизнеса из-за ограниченного бюджета ключевыми задачами являются простота внедрения, низкая стоимость владения и высокая полезность для ежедневных операций. Далее будут описаны основные принципы построения BI-системы, архитектура решения, сценарии использования BI-систем, а также будут выделены практические рекомендации для компаний по использованию BI-систем.

1. Основные принципы построения BI-системы для малого бизнеса

Перед внедрением любого ПО в любом случае необходимо изучить IT-инфраструктуру предприятия для понимания того какая BI-система будет эффективнее для определенного стека. Но тем не менее все сводится к пяти принципам, которые будут описаны далее.

© Варданян Н. А., 2026

Принцип модульности: система должна состоять из взаимозависимых модулей, таких как:

1. модуль сбора данных – данный модуль отвечает за получение данных из базы данных компании;
2. модуль хранения данных – данный модуль является хранилищем всей собранной информации с помощью модуля сбора данных;
3. модуль аналитики и визуализации – данный модуль отвечает за визуализацию и трансформацию данных из модуля хранения данных.

Принцип экономичности и масштабируемости: решение должно быть в доступной ценовой категории и легко адаптироваться к росту бизнеса. Для малого бизнеса самые подходящие платформы будут облачные BI-системы (SaaS).

Принцип интероперабельности: BI-система должна иметь возможность безошибочной интеграции с IT-структурой предприятия через готовые коннекторы. Это очень важный принцип, который позволит легко извлекать данные из системы и направлять их напрямую в BI-систему для последующей обработки данных.

Принцип безопасности данных: несмотря на малый масштаб бизнеса, в любом случае необходимо защитить персональные данные клиентов, а также данные о поставщиках и закупочных ценах.

Принцип единого источника данных: аналитика данных должна строиться на основе единой базы данных компании, так как если будут указаны несколько источников данных, то результаты работы могут дублироваться и будет выведена ложная информация.

2. Архитектурные решения для BI-системы в малом бизнесе

Как уже было описано выше, для малого бизнеса стоит использовать облачную архитектуру для снижает затраты на инфраструктуру и обеспечения высокой доступности системы. Необходимо выделить следующие компоненты:

1. слой сбора данных: функция ETL (Extract, Transform, Load) для автоматического сбора данных;
2. слой хранения данных: облачные базы данных (PostgreSQL, Google BigQuery);
3. слой аналитики и визуализации данных: BI-системы (Power BI, Tableau, Yandex DataLens);
4. слой интеграции систем: API-шлюзы для подключения внешних сервисов.

Для успешного внедрения в первую очередь необходимо настроить базу данных компании, выбрать определенные метрики, которые нас интересуют и с помощью функции ETL извлечь из БД компании в BI-систему для последующей трансформации данных в дашборды. Очень

важно чтобы система имела возможность интеграции, так как самая важная часть процесса визуализации данных является автоматическое извлечение данных из БД.

3. Сценарии использования BI в малом бизнесе

С помощью анализа данных можно анализировать рекламные кабинеты из социальных сетей, что позволит в дальнейшем оптимизировать бюджет на рекламу и выявить прибыльную аудиторию. В данном сценарии дашборды могут отражать такие метрики как стоимость лида, конверсию и LTV клиента.

С точки зрения финансового управления автоматизированное формирование отчетов по движению товаров могут позволить компании контролировать дебиторскую задолженность и оперативно реагировать на изменения финансовых показателях.

Автоматическое формирование отчётов о движении денежных средств, прибыльности по продуктам или услугам, структуре затрат. BI-система помогает прогнозировать кассовые разрывы, контролировать дебиторскую задолженность и оперативно реагировать на изменения финансовых показателей.

Анализ воронки продаж, определение этапов с наибольшим оттоком клиентов, сегментация базы по покупательскому поведению позволит персонализировать коммуникацию, повысить средний чек и снизить процент закупки неликвидного товара.

4. Управление внедрением BI-системы

Далее необходимо описать этапы внедрения:

1. Анализ потребностей и IT-инфраструктуры предприятия для определения дальнейшего выбора BI-системы.
2. Стадия планирования внедрения BI-системы.
3. Выбор платформы, который наиболее подходит под IT-инфраструктуру предприятия, при этом сравнение BI-инструментов будет происходить по критериям стоимости, удобства и интеграции.
4. Пробный запуск дашборда в демоверсии BI-системы с несколькими отчетами.
5. Обучение сотрудников работе с системой.
6. Подключение новых источников данных и расширение функционала.
7. Регулярный мониторинг, обновление отчётов, адаптация под меняющиеся требования.

5. Риски внедрения и способы их минимизации

Может возникнуть сопротивление сотрудников из-за нежелания менять привычные процессы (например, работу в Excel). Решением данной проблемы является вовлечение команды с ранних этапов и последующая демонстрация практической пользы системы.

Недостаточное качество данных может стать очень серьёзным риском так как разрозненные, неактуальные или противоречивые данные могут привести к катастрофическим последствиям. Данной ситуации можно избежать если грамотно настроить связь данных между системой компании и BI-системой. Так же регулярная проверка источников данных может свести риски неправильных данных к минимуму.

Несвоевременная выгрузка тоже может оказать негативный эффект на деятельность компании поэтому необходимо настроить автоматическую выгрузку данных из базы данных компании в BI-систему.

6. Практические рекомендации

В первую очередь предприятию надо определиться с необходимыми отчетами, которые будут включены в определённый дашборд. Для этого рекомендуется использование готовых шаблонов для отчетов. Данный подход резко ускорит создание аналитических панелей, избавляя от разработки с чистого листа и экономя десятки часов работы.

Так же необходимо вовлечь конечных пользователей системы, которые будут работать с отчетами. Их прямое участие позволит избежать недопониманий, так как связь с конечными пользователями гарантирует практичность и полезность итоговых результатов. При этом желательно обучить сотрудника, который сможет решить вопросы пользователей, следить за обновлениями и метриками.

Заключение

Внедрение BI-системы позволяет повысить прозрачность бизнеса, ускорить принятие решений и повысить конкурентоспособность. В будущем BI-системы будут использовать технологии искусственного интеллекта. для прогнозирования и автоматизации аналитики, что делает их внедрение ещё более актуальным уже сегодня.

Список литературы

1. ГОСТ Р ИСО/МЭК 24668-2022. Информационные технологии. Аналитика больших данных. Общие требования. — Введ. 2023-01-01. — М.: Рос. ин-т стандартизации, 2022. — 32 с.

2. ГОСТ Р 59344-2021. Системная инженерия. Защита информации при анализе бизнес-процессов. — Введ. 2022-04-01. — М.: Рос. ин-т стандартизации, 2021. — 20 с.

3. ГОСТ Р 59992-2022. Системная инженерия. Жизненный цикл информационных систем. — Введ. 2023-07-01. — М.: Рос. ин-т стандартизации, 2022. — 28 с.

4. Хрусталева, Е.Ю. 1С:Аналитика. BI-система в «1С:Предприятии 8». Изд. 2-е. — М.: 1С-Паблишинг, 2022. — 450 с.

5. Демиденко, А. Аналитика 360: Big Data и BI-системы, которые меняют игру. — М.: Эксмо, 2023. — 256 с.

6. Ковалёвы, С., Ковалёва, В. Настольная книга аналитика: моделирование бизнес-процессов. — СПб.: Питер, 2021. — 384 с.

7. Андерсон, К. Аналитическая культура: от данных к действиям в бизнесе / пер. с англ. — М.: Манн, Иванов и Фербер, 2020. — 320 с.

8. Марр, Б. Ключевые инструменты бизнес-аналитики: 67 методов / пер. с англ. — М.: Эксмо, 2022. — 288 с.

9. Бизнес-аналитика. Введение в обработку и анализ данных: учеб. пособие. — М.: Юрайт, 2025. — 240 с.

Архитектура данных корпоративных систем кадрового электронного документооборота

А. А. Гатауллина, e-mail: alinagataullinaa@yandex.ru

ФГБОУ ВО «Санкт-Петербургский государственный экономический университет»

Аннотация. *В статье приведен анализ архитектуры данных как фундаментального компонента современных корпоративных систем кадрового электронного документооборота. Исследована многоуровневая организация данных, принципы моделирования предметной области и стратегии обеспечения целостности, безопасности и масштабируемости.*

Ключевые слова: *кадровый электронный документооборот, архитектура данных, цифровая трансформация, бизнес-процессы, проектирование базы данных.*

Введение

В современном мире цифровизация бизнес-процессов становится неотъемлемой частью успешного функционирования коммерческих организаций. В условиях стремительного развития информационных технологий и растущей конкуренции компаниям важно отслеживать технологические тренды и соответствовать им. Ключевым трендом цифровой трансформации в сфере управления персоналом является создание целостной и надежной системы кадрового электронного документооборота (КЭДО).

КЭДО представляет собой комплексную систему, которая обеспечивает ведение полного жизненного цикла бизнес-процесса документооборота в компании в цифровом виде. Эффективность такой системы определяется качеством ее архитектурного фундамента – безопасной, целостной и масштабируемой архитектуры данной.

Архитектура данных является важным элементом, преобразующим кадровый электронный документооборот из простого инструмента хранения документов в элемент цифровой экосистемы компании, трансформирующий HR-процессы. Помимо операционной эффективности и скорости обработки запросов она должна обеспечивать соответствие строгим юридическим требованиям, защиту персональных данных, а также интегрируемость с другими корпоративными системами.

© Гатауллина А.А., 2026

1. Многоуровневая модель архитектуры данных кадрового электронного документооборота

Архитектура данных может быть представлена в виде трехуровневой модели, на каждом уровне которой решаются задачи управления данными на разных уровнях абстракции.

Концептуальный уровень. На этом уровне определяется, что система должна знать и какие правила соблюдать. Результатом является концептуальная модель в виде диаграммы «Сущность-связь» (ERD).

Основными сущностями системы являются сотрудник, документ, подразделение и должность, а также маршрут согласования. Сотрудник выступает центральным объектом и содержит атрибуты идентификации, персональные и трудовые данные. Документ представляет собой кадровый документ и включает атрибуты его жизненного цикла, такие как дата, версия и статус. Подразделение и должность отражают элементы организационной структуры компании. Маршрут согласования описывает рабочий процесс, связывает документы с работниками и фиксирует результаты прохождения этапов согласования.

На концептуальном уровне формулируются ключевые бизнес-правила, такие как уникальность табельного номера в рамках компании, обязательное наличие автора у документа, запрет на исключение должности, на которой числятся работники, из штатного расписания.

Логический уровень. Этот уровень описывает, как данные будут структурированы в конкретной модели данных. В большинстве систем КЭДО используется реляционная модель данных, поскольку она соответствует требованиям к структурированному хранению данных, надежна и обеспечивает целостность данных.

Основными задачами логического уровня являются преобразование диаграммы «Сущность–связь» в схему отношений, при котором сущности преобразуются в таблицы, атрибуты — в столбцы с определёнными типами данных, а связи между сущностями реализуются с помощью внешних ключей. Также на данном уровне выполняется нормализация структуры базы данных, направленная на устранение дублирования информации путём приведения таблиц к нормальным формам, как правило к третьей нормальной форме (3НФ). Кроме того, на логическом уровне устанавливаются правила целостности данных: реляционные базы данных позволяют задавать их с помощью механизмов PRIMARY KEY, FOREIGN KEY, UNIQUE, CHECK и NOT NULL.

Физический уровень определяет, где и в каком виде хранятся данные. Принятые на этом уровне решения напрямую влияют на производительность, отказоустойчивость и доступность системы. Он включает выбор системы управления базами данных, который осуществляется между классическими СУБД, такими как Oracle, PostgreSQL и MS SQL Server, и облачными решениями. Также на физическом уровне разрабатываются стратегии индексирования, предусматривающие создание индексов для ускорения поиска по номеру, дате, автору или статусу при работе с большими объёмами кадровых документов.

Дополнительно на физическом уровне применяется сегментация данных, заключающаяся в разделении крупных таблиц на сегменты, чаще всего по диапазонам дат, что позволяет ускорить запросы к актуальным данным и упростить процессы архивирования. Важным аспектом также являются политики резервного копирования и хранения, которые определяют сроки хранения оперативных и архивных данных, а также частоту и тип резервных копий в соответствии с требованиями к восстановлению информации.

Трехуровневая модель служит каркасом для проектирования данных, который дает возможность разделить ответственность между бизнес-аналитиками, архитекторами и администраторами. На концептуальном уровне определяются сущности и правила предметной области, на логическом – они преобразуются в строгую схему таблиц и ограничений, а на физическом – происходит адаптация схемы под конкретную технологическую платформу с учетом всех требований.

2. Специфика моделирования данных для кадрового электронного документооборота

Документ в КЭДО – динамический объект с предопределенным жизненным циклом (Черновик → На согласовании → Утвержден → Архив). Модель данных должна поддерживать следующие аспекты:

– версионность – хранение всех версий документа с возможностью сравнения и возвращения к одной из них;
– метаданные процесса – фиксацию полной истории согласований: кто и когда согласовывал, результат согласования и комментарии.

Также организация данных должна быть произведена таким образом, чтобы электронный документооборот отвечал всем юридическим требованиям. Это реализуется через следующие принципы:

1. Гарантия достоверности и неизменности. Каждый электронный документ должен быть однозначно связан с его автором а его содержание после фиксации исключало любые несанкционированные изменения. Это может быть достигнуто интеграцией механизмов квалифицированной электронной подписи с хранением соответствующих сертификатов и криптографических хэшей документов. Для формирования юридически значимой цепочки событий необходимо создание журнала аудита, записи в котором должны быть защищены от внесения правок.

2. Ориентация на долговременное и регламентированное хранение. Архитектура должна учитывать требования трудового законодательства и учетной политики, устанавливающие сроки хранения кадровых документов, достигающие 75 лет. Это означает, что модель данных должна предполагать миграцию информации между носителями и обеспечивать читаемость в долгосрочной перспективе.

При проектировании архитектуры данных корпоративной системы КЭДО необходимо учитывать обязательные требования Федерального закона №152-ФЗ «О персональных данных». Соблюдение правил, установленных этим нормативным актом, определяет ключевые принципы организации хранения информации:

– разграничение прав доступа – механизмы контроля прав должны быть реализованы на уровне самой базы данных;

– минимизация данных – хранение только необходимых персональных данных;

– реализация прав субъектов персональных данных – создание четких процедур для выполнения сотрудниками запросов на предоставление, исправление и удаление своих персональных данных.

Проектирование архитектуры данных для системы КЭДО представляет собой комплексную задачу, которая требует баланса между операционными, юридическими и регуляторными аспектами. Однако создание корректной структуры хранения информации является лишь необходимым условием. Достаточным условием надежности системы выступает реализация стратегии поддержания целостности и непротиворечивости данных в процессе их постоянного изменения.

3. Стратегии обеспечения целостности и непротиворечивости данных

Надежность системы КЭДО напрямую зависит от того, насколько успешно удается поддерживать корректность и согласованность хранимой информации. Для этого применяются три взаимодополняющие стратегии, соответствующие разным уровням сложности бизнес-правил и архитектурных требований.

1. Декларативная целостность.

Данный подход предполагает описание правил непосредственно в схеме базы данных с использованием встроенных механизмов СУБД. Это наиболее эффективный и производительный способ, поскольку контроль осуществляется на системном уровне.

В рамках декларированной целостности можно выделить два типа ограничений, гарантирующих корректность данных – ссылочную и пользовательскую целостности.

Ссылочная целостность обеспечивается ограничениями внешнего ключа, например, невозможно создать приказ о премировании, ссылающийся на несуществующего работника, или удалить запись о подразделении, в котором числятся работники.

Пользовательская целостность задается с помощью механизмов проверочных ограничений и условий уникальности, например, гарантия того, что дата окончания отпуска будет строго позже даты его начала, или обеспечение уникальности табельного номера в рамках компании.

2. Процедурная целостность.

Для сложных бизнес-правил, которые нельзя описать ограничениями в схеме базы данных, используются процедурные механизмы.

Основные инструменты данного подхода:

– транзакции, гарантирующие атомарность и согласованность операций (например, при утверждении приказа о приеме выполняется последовательность действий: создание записи о приказе, обновление статуса сотрудника, фиксация события в истории изменений);

– триггеры, блоки кода, автоматически выполняемые при определенных событиях в базе данных (например, триггер для проверки корректности цепочки согласующих перед вставкой документа).

3. Непротиворечивость в распределенной среде.

Для компаний с территориально удаленными филиалами или сложным ИТ-ландшафтом, где данные дублируются между несколькими системами, актуальна задача поддержания их общей согласованности. Для ее решения применяются несколько подходов:

– назначение единого источника истины (MDM), при котором для каждого типа данных выбирается одна главная система, которая является авторитетным источником информации, все остальные системы синхронизируются с этим источником;

– асинхронная репликация, которая используется для обеспечения скорости работы в удаленных офисах, где прямой доступ к центральной базе данных может быть медленным (изменения из центральной базы

передаются на реплики с некоторой задержкой, что в конечном счете создает модель согласованности);

– шаблон «Сага» (Saga), применяемый для управления длительными бизнес-процессами, которые затрагивают несколько независимых систем (вместо одной сложной транзакции процесс разбивается на последовательность более простых транзакций в каждой системе, если один из этапов завершился неудачей, запускаются специальные компенсирующие транзакции).

Стратегии обеспечения целостности и непротиворечивости данных формируют внутренний каркас надежности системы КЭДО.

Создание внутренней устойчивости системы – лишь часть задачи. В реальной корпоративной среде система также должна безопасно взаимодействовать с другими компонентами ИТ-ландшафта.

4. Интеграционный аспект и безопасность

Система КЭДО не может существовать в полной изоляции, она всегда работает в интеграции с другими системами, такими как бухгалтерские программы, корпоративные порталы, средства аналитики. Поэтому архитектура данных системы должна предусматривать безопасные способы обмена информации с внешними программами. Она должна включать следующие аспекты:

– интерфейсы интеграции (API) – специальные программные интерфейсы (RESTful или SOLAP API) для безопасного обмена данными с системами расчета заработной платы, бухгалтерии, корпоративными порталами;

– многоуровневая модель безопасности – от защиты на уровне сети и СУБД до криптографического шифрования чувствительных полей в базе данных (например, паспортных данных) и обязательной фиксации всех действий;

– архитектурный подход «Шлюз базы данных» (Database Gateway) – изоляция прямого доступа к базе данных, обмен происходит через специализированный сервисный слой, что усиливает контроль и безопасность.

Интеграционные механизмы и многоуровневая безопасность являются обязательными элементами архитектуры данных системы КЭДО. Без их проектирования система потеряет свою практическую значимость и создаст риски утечки важной кадровой информации.

Заключение

Архитектура данных для корпоративной системы кадрового электронного документооборота является сложной, многоаспектной концепцией, успех которой определяется сбалансированностью решений

на всех уровнях – концептуальном, логическом и физическом. Она должна быть ориентирована не только на хранение и выборку, но в первую очередь на соответствие юридическим требованиям, полную корректность и надежную защиту, поскольку от этого напрямую зависит работа всей организации.

Проектирование архитектуры системы КЭДО требует системного подхода, при котором моделирование предметной области, выбор технологий обеспечения целостности, стратегии защиты информации и планы интеграции с другими системами разрабатываются как взаимосвязанные элементы единого целого. Такая архитектура станет стратегической инвестицией, которая создаст устойчивый и надежный фундамент для цифровой трансформации управления персоналом.

В будущем развитие архитектур данных для систем КЭДО может быть связано с формализацией законодательных требований в виде готовых шаблонов проектирования и создания инструментов для автоматической проверки соответствия архитектуры нормативным актам, что повышает актуальность формирования методологически строгих подходов уже на текущем этапе.

Список литературы

1. Федеральный закон от 06.04.2011 N 63-ФЗ (ред. от 24.04.2023) «Об электронной подписи». URL: https://www.consultant.ru/document/cons_doc_LAW_112701/ (дата обращения: 06.10.2025). – Текст: электронный.

2. Федеральный закон от 27.07.2006 N 149-ФЗ (ред. от 02.07.2021) «Об информации, информационных технологиях и о защите информации». – URL: https://www.consultant.ru/document/cons_doc_LAW_61798/ (дата обращения: 08.10.2025). – Текст: электронный.

3. Приказ Роструда от 19.10.2021 № 258 «Об утверждении форм электронных документов, связанных с трудовой деятельностью». – Текст: электронный // КонсультантПлюс: [сайт]. – URL: https://www.consultant.ru/document/cons_doc_LAW_401260/ (дата обращения: 16.10.2023). – Текст: электронный.

4. ГОСТ Р 55681-2013/ISO/TR 26122:2008 Информация и документация. Анализ процессов работы с точки зрения управления документами (утв. и введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 31 октября 2013 г. № 1303-ст) // Информация и документация: сб. ГОСТов – М: ФГУП «Стандартинформ», 2020. – 20 с. URL:

https://docs.cntd.ru/document/1200142871 (дата обращения: 07.10.2025). – Текст: электронный.

5. ГОСТ Р ИСО/МЭК 27001-2022. Информационная безопасность, кибербезопасность и защита приватности. Системы менеджмента информационной безопасности. Требования: утвержден и введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 29.09.2022 № 831-ст: дата введения 2023-04-01. – URL: https://docs.cntd.ru/document/1200181890 (дата обращения: 16.10.2023). – Текст: электронный.

6. Соколов, Р.В., Проектирование и эксплуатация информационных систем: учеб. пособие / Р.В. Соколов, И.Л. Андреевский. – СПб.: СПбГЭУ, 2023. – 348 с.

7. Сотникова, С.И. Управление персоналом организации: современные технологии: учебник / С.И. Сотникова, Е.В. Маслов, Н.Н. Абакумова [и др.]; под науч. ред. С.И. Сотниковой. – М.: ИНФРА-М, 2019. — 513 с.

Корпоративные информационные системы нового поколения: как микросервисы и гибридные базы данных ломают монолит

С. Ю. Гордеев, e-mail: sergord26@mail.ru

ФГБОУ ВО «Санкт-Петербургский государственный экономический университет»

Аннотация. *В данной статье рассматриваются принципы и практики построения корпоративных информационных систем нового поколения на основе микросервисной архитектуры и гибридных подходов к управлению данными. Анализируется переход от традиционных монолитных приложений к распределённым микросервисным архитектурам, изучаются преимущества и вызовы внедрения полиперсистентных решений для работы с разнородными типами данных. Рассматриваются практические аспекты миграции существующих систем, управления жизненным циклом микросервисов и обеспечения надёжности в распределённых окружениях.*

Ключевые слова: *микросервисная архитектура, монолитные приложения, гибридные базы данных, цифровая трансформация, масштабируемость, распределённые системы, контейнеризация, облачные технологии, NoSQL, полиперсистентность.*

Введение

Современные организации всех размеров стоят перед вызовом цифровой трансформации, которая требует не только обновления технологий, но и полного переосмысления архитектуры информационных систем. На протяжении последних двух десятилетий монолитная архитектура была стандартом разработки корпоративного ПО, однако быстро растущие требования к масштабируемости, производительности, надёжности и скорости разработки делают этот подход всё менее пригодным.

Традиционная монолитная архитектура, где все компоненты системы плотно интегрированы в единое приложение, создаёт целый ряд проблем. Увеличение кодовой базы приводит к усложнению развёртывания, затрудняет тестирование, усложняет командную разработку и делает систему хрупкой перед лицом необходимых изменений. Кроме того, масштабирование монолитного приложения

© Гордеев С. Ю., 2026

требует развёртывания всей системы целиком, что неэффективно с точки зрения использования ресурсов и затрат.

На смену монолиту приходит микросервисная архитектура – подход, при котором приложение разбивается на набор небольших, слабосвязанных, независимо развёртываемых сервисов. Каждый сервис отвечает за конкретную бизнес-функцию, использует собственное хранилище данных и взаимодействует с другими сервисами через чётко определённые интерфейсы. Параллельно с развитием архитектуры приложений происходит революция в области управления данными. Если ранее корпоративные системы опирались на единое централизованное хранилище на основе реляционной базы данных, то сегодня всё больше компаний переходят к гибридным подходам, использующим различные типы баз данных в соответствии с потребностями каждого микросервиса. Этот подход известен как полиперсистентность и открывает новые возможности для оптимизации производительности и функциональности систем.

Целью настоящей статьи является анализ принципов построения корпоративных информационных систем нового поколения, которые используют микросервисную архитектуру в сочетании с гибридными подходами к хранению данных. В статье рассматриваются как технические аспекты, так и практические вызовы, связанные с внедрением таких систем в реальной бизнес-среде.

1. От монолита к микросервисам: историческая эволюция архитектур

Монолитная архитектура долгое время была доминирующим подходом к разработке корпоративного ПО. В этой модели все компоненты системы – интерфейс пользователя, бизнес-логика, доступ к данным и внешние интеграции – объединяются в единое приложение, обычно упакованное в один или несколько исполняемых файлов.

Основные характеристики монолитной архитектуры:

– все части системы тесно интегрированы, что делает модификацию или удаление отдельного компонента сложной операцией, так как может потребоваться изменение других частей системы;

– весь код хранится в одном репозитории, что упрощает разработку на ранних стадиях, но со временем приводит к сложности управления и конфликтам при коллаборативной разработке;

– любое изменение в системе требует пересборки и переразвёртывания всего приложения, что увеличивает риск сбоев и время простоя;

– обычно используется одна большая реляционная база данных, которая служит хранилищем для всех данных системы;

– при необходимости обработать большую нагрузку единственным вариантом является развёртывание дополнительных копий всего приложения, что неэффективно, если проблема заключается в производительности конкретного компонента.

Монолитная архитектура хорошо работает для небольших приложений и команд, но по мере роста системы её недостатки становятся всё более очевидными.

Несколько факторов подтолкнули индустрию к переосмыслению архитектурных подходов:

– современные компании должны быстро адаптироваться к изменяющимся рыночным условиям, что требует быстрой разработки и развёртывания новых функций;

– объёмы данных, которые обрабатывают современные системы, растут экспоненциально. Единая база данных становится узким местом как с точки зрения производительности, так и с точки зрения масштабируемости;

– облачные платформы предоставили гибкую инфраструктуру, позволяющую развёртывать приложения на основе микросервисов с минимальными капитальными затратами;

– по мере роста организаций разработка становится более специализированной, и микросервисная архитектура позволяет разным командам работать независимо над своими сервисами.

Микросервисная архитектура основана на нескольких ключевых принципах:

– микросервисы взаимодействуют через чётко определённые интерфейсы, что позволяет минимизировать зависимости между ними;

– все компоненты, необходимые для реализации конкретной бизнес-функции, должны находиться в одном сервисе;

– каждый микросервис должен быть способен функционировать самостоятельно, с собственным хранилищем данных и логикой обработки ошибок;

– отдельные микросервисы могут масштабироваться независимо от других на основе их конкретных потребностей в ресурсах;

– система должна быть спроектирована так, чтобы отказ одного микросервиса не привёл к отказу всей системы.

2. Гибридные базы данных и полиперсистентность

Одним из ключевых аспектов микросервисной архитектуры является переход от единой централизованной базы данных к гибридному подходу, при котором каждый микросервис может использовать оптимальный для его потребностей тип базы данных. Этот подход известен как полиперсистентность.

В традиционной корпоративной системе все данные хранились в одной большой реляционной базе данных. Это обеспечивало согласованность данных, но создавало узкие места при масштабировании и требовало использования одного стиля моделирования данных для всех типов данных.

Полиперсистентность позволяет разработчикам выбирать наиболее подходящий для конкретной задачи тип базы данных:

– реляционные базы данных используются для данных со сложными связями и требованиями к согласованности, таких как финансовые данные, информация о заказах, учетные данные;

– документоориентированные БД подходят для работы с неструктурированными или слабоструктурированными данными, такими как пользовательские профили, конфигурации, логи событий;

– хранилища ключ-значение типа системы обеспечивают экстремально быстрый доступ к данным и часто используются для кэширования, сессий пользователей и рейтинговых систем;

– графовые базы данных соответствуют для работы с высокосвязанными данными, такими как социальные сети, системы рекомендаций, организационные структуры;

– хранилища временных рядов оптимизированы для хранения и анализа метрик, логов и событий с временными метками;

– поисковые индексы используются для полнотекстового поиска и аналитики больших объёмов неструктурированного текста.

Использование гибридных баз данных открывает ряд преимуществ:

– каждая база данных может быть оптимизирована под конкретные паттерны доступа, что приводит к лучшей производительности;

– различные типы данных могут масштабироваться независимо;

– различные БД имеют разные подходы к схеме данных;

– специализированные БД часто дешевле в эксплуатации для своих конкретных вариантов использования, чем универсальное решение;

– использование новых типов БД позволяет компаниям внедрять инновационные функции, которые было бы сложно реализовать с единственной реляционной БД.

3. Архитектурные решения в микросервисных системах

В микросервисной архитектуре интерфейсы между сервисами являются ключевым элементом конструкции. Поэтому многие организации принимают API-первый подход, при котором разработка API предшествует разработке самих сервисов.

При API-первом подходе команда сначала определяет контракт API – что будут принимать сервисы на вход и что они будут возвращать на

выход. Это позволяет разработчикам других сервисов сразу начинать работу, не дожидаясь завершения разработки целевого сервиса.

REST API стал де-факто стандартом для синхронной коммуникации между микросервисами. REST основан на HTTP и использует стандартные HTTP-методы (GET, POST, PUT, DELETE) для выполнения операций. REST прост в реализации и понимании, и его легко тестировать.

Однако для сложных систем, требующих асинхронной коммуникации, часто используются message brokers, такие как RabbitMQ, Kafka или AWS SQS. Message brokers позволяют сервисам взаимодействовать асинхронно через очереди сообщений, что обеспечивает лучшую развязку и устойчивость к отказам.

При наличии множества микросервисов возникает вопрос о том, как клиентам взаимодействовать с системой. Один подход – дать прямой доступ к каждому микросервису, но это усложняет жизнь клиентам, которые должны знать об адресах всех сервисов.

API-шлюз (API Gateway) решает эту проблему, действуя как единая точка входа для всех клиентов. API-шлюз маршрутизирует запросы к соответствующим микросервисам, может выполнять аутентификацию, авторизацию, логирование, трансформацию данных и другие функции.

Популярные реализации API-шлюзов включают Kong, AWS API Gateway, Azure API Management и другие.

Обнаружение сервисов (Service Discovery) – важный компонент микросервисной архитектуры. Когда микросервисы динамически развёртываются и запускаются в облаке или на кластере контейнеров, их адреса неизвестны заранее. Service Discovery решает эту проблему, позволяя сервисам регистрировать свой адрес при запуске и помогая другим сервисам находить друг друга.

4. Практические вызовы и решения

В монолитной системе с единой базой данных транзакции управляются СУБД, которая гарантирует свойства ACID (атомарность, согласованность, изолированность, долговечность).

В микросервисной системе, где каждый сервис имеет собственное хранилище данных, управление распределёнными транзакциями становится намного сложнее. Если операция требует изменения данных в нескольких микросервисах, возникает риск того, что некоторые изменения будут выполнены, а другие нет, оставляя систему в несогласованном состоянии.

Несколько подходов решают эту проблему:

Saga Pattern. Вместо традиционной ACID-транзакции используется набор локальных транзакций, координируемых сагой. Если одна

локальная транзакция не удаётся, сага выполняет компенсирующие транзакции в других сервисах для отката изменений.

Event Sourcing. Вместо хранения текущего состояния сущности хранится последовательность событий, которые привели к этому состоянию. Это обеспечивает полный аудит-лог и позволяет восстановить состояние системы в любой момент времени.

CQRS (Command Query Responsibility Segregation). Разделение операций чтения и записи, используя разные модели данных для каждой операции. Это позволяет оптимизировать каждую модель под её конкретное назначение.

5. Миграция с монолита на микросервисы

Переход с монолитной архитектуры на микросервисы – это сложный и длительный процесс. Несколько стратегий помогают снизить риск и упростить миграцию:

– стратегия Strangler Fig позволяет вместо полной переписи кода одновременно, новые функции реализуются на микросервисах, а старый монолит постепенно вытесняется;

– разделение по бизнес-доменам дает возможность выделить ключевые бизнес-функции и извлекают их из монолита в отдельные микросервисы, начиная с менее критичных функций;

– API-фасад создает новый API поверх старой системы, который скрывает её сложность. Это может быть промежуточный этап перед полной миграцией на микросервисы.

Заключение

Переход от монолитной архитектуры к микросервисам в сочетании с гибридными подходами к хранению данных представляет собой

фундаментальный сдвиг в том, как строятся корпоративные информационные системы. Этот переход позволяет организациям быть более гибкими, масштабируемыми и инновационными.

Однако, несмотря на множество преимуществ микросервисной архитектуры, её внедрение требует значительных инвестиций в инструменты, обучение и переосмысление процессов разработки. Микросервисы – это не серебряная пуля, и они создают собственный набор вызовов в распределённых системах, управлении данными и операциях.

Ключ к успешному внедрению микросервисов заключается в постепенной миграции. Не следует переводить всю систему на микросервисы одновременно. Инвестиции в инструменты и процессы. Убедитесь, что ваши команды имеют правильные инструменты для развёртывания, мониторинга и управления микросервисами. Культурные

изменения. Микросервисная архитектура требует изменений в том, как организована работа команд. Команды должны быть более самостоятельными и ответственными.

Измерение и итерация. Регулярно измеряйте метрики системы и вносите улучшения на основе полученных данных. Компании, которые успешно внедрили микросервисную архитектуру с гибридными подходами к данным, получают значительные преимущества в скорости разработки, масштабируемости и способности адаптироваться к изменяющимся требованиям бизнеса. Однако путь трансформации требует тщательного планирования, инвестиций и терпения.

По мере развития технологий и появления новых инструментов, микросервисная архитектура будет продолжать эволюционировать. Организации, которые понимают принципы микросервисной архитектуры и гибридного управления данными, будут лучше позиционированы для успеха в быстро меняющейся цифровой экономике. Переход от монолита к микросервисам – это не просто технический вопрос, а стратегический выбор, который влияет на всю организацию и её способность конкурировать в современном мире.

Список литературы

1. ГОСТ Р ИСО/МЭК 27001–2021. Информационная безопасность. Системы менеджмента ИБ. Требования: дата введения 2022-01-01. — М.: Рос. ин-т стандартизации, 2022. — 28 с.

2. Худойкулов Б.Р. О новых поколениях корпоративных информационных систем в XXI веке / Chronos: естественные и технические науки. 2021. Т. 6. № 5 (38). С. 5-9.

3. Попова К. Проектирование базы данных в контексте гибридной архитектуры SaaS-приложения / В сборнике: Фундаментальная и прикладная наука. Сборник научных трудов по материалам XXXI Международной научно-практической конференции, С. 25-33

4. Морозова О.И., Семенихина А.В. Внедрение и использование IT-технологий в системе менеджмента организации // Использование информационных технологий в различных сферах деятельности. - Гомель: Белорусский торгово-экономический университет потребительской кооперации, 2024. - с. 104-108.

5. Миклушевский В.В., Прокошкин А.С., Красильников И.О., Туманов В.Е. Инновации в управлении вузом: новые решения для корпоративной информационной системы / Университетское управление: практика и анализ. 2006. № 6. С. 16-24.

6. Назаров, Д.М. BI-технологии и корпоративные информационные системы в оптимизации бизнес-процессов / Д.М.

Назаров, С.В. Бегичева, Д.А. Азаров. — Екатеринбург: Уральский гос. экон. ун-т, 2018. — 121 с.

7. Соколов, Р.В. Проектирование и эксплуатация информационных систем: учеб. пособие / Р.В. Соколов, И.Л. Андреевский. — СПб.: СПбГЭУ, 2023. — 348 с.

8. Свод знаний по управлению бизнес-процессами: BPM CBOK 4.0 / Джек Х., Дэн М. и др. — М.: Альпина Паблишер, 2023. — 503 с.

Принципы построения и архитектура единой информационной платформы для цифровой трансформации организаций в сфере недвижимости

В. А. Косыгин, e-mail: vovakosygin@gmail.com

ФГБОУ ВО «Санкт-Петербургский государственный экономический университет»

Аннотация. *В данной статье рассматриваются принципы построения и архитектура единой информационной платформы для цифровой трансформации в сфере коммерческой недвижимости.*

Ключевые слова: *цифровая трансформация, единая информационная платформа, CRM, ERP, СЭДО, интеграция систем, управление недвижимостью, бизнес-процессы, микросервисная архитектура, облачные технологии.*

Введение

Современные организации в сфере коммерческой недвижимости сталкиваются с необходимостью цифровой трансформации, обусловленной ростом конкуренции, увеличением объемов данных и потребностью в повышении операционной эффективности. Это особенно актуально для организаций, которые представляют из себя филиальную структуру, где разрозненность систем и процессов приводит к снижению управляемости и росту издержек.

Единая информационная платформа представляет собой комплексное программно-техническое решение, обеспечивающее консолидацию данных, автоматизацию процессов и создание единого информационного пространства для организации. Эффективность такой платформы определяется правильным выбором архитектурных решений и следованием фундаментальным принципам построения корпоративных информационных систем.

1. Фундаментальные принципы построения единой информационной платформы

Принцип модульности и декомпозиции: Модульность является одним из ключевых принципов построения современных информационных платформ. Система должна быть разделена на функционально независимые модули, каждый из которых решает определенный набор задач.

© Косыгин В. А., 2026

Основные модули платформы для организаций в сфере коммерческой недвижимости включают модуль управления проектами для планирования, бюджетирования и контроля сроков выполнения работ; модуль управления взаимоотношениями с клиентами (CRM) для ведения клиентской базы, управления продажами и автоматизации маркетинга; модуль управления ресурсами предприятия (ERP) для финансового учета, управления персоналом и закупками; модуль ЭДО для управления электронным документооборотом и согласованием; модуль аналитики и отчетности для бизнес-аналитики, формирования управленческих отчетов и прогнозирования; модуль управления недвижимостью для эксплуатации объектов, управления арендными отношениями, контроля платежей, планирования и учета ремонтов.

Модули могут представлять из себя как отдельные системы, так и системы внутри систем. Все модули должны иметь возможность интеграции между собой, так как этого и требует суть единой информационной платформы.

Модульная архитектура обеспечивает гибкость системы, позволяя добавлять новые функциональные блоки без необходимости перестройки всей готовой платформы.

Принцип масштабируемости: Масштабируемость определяет способность системы адаптироваться к растущим объемом данных организации, увеличению числа пользователей системы и расширению функциональности при необходимости.

Масштабируемость может быть представлена двумя разными типами:

Вертикальное масштабирование – это увеличение мощности существующих серверов. Подход кажется простым с точки зрения реализации, но несет под собой будущие ограничения.

Горизонтальное масштабирование – это добавление дополнительных вычислительных узлов для распределения нагрузки по ним. Данный подход обеспечивает неограниченные возможности расширения и повышает надежность всей системы.

Исходя из изложенных выше двух типов масштабируемости можно сделать вывод, что для обеспечения масштабируемости системы лучше всего использовать распределение нагрузки между серверами, асинхронную обработку данных, разделение данных между хранилищами.

Принцип интероперабельности: Интероперабельность представляет из себя способность различных информационных систем эффективно обмениваться данными и использовать их совместно. В контексте единой информационной платформы интероперабельность важна для

интеграции систем внутри платформы, интеграции платформы с системами государственных информационных ресурсов, внешними финансовыми сервисами, специализированным программным обеспечением для проектирования и инженерии.

Чтобы обеспечить интероперабельность платформы нужно реализовать несколько уровней: технологический уровень предполагает использование стандартизированных протоколов обмена данными между системами (REST API, SOAP), форматов данных (JSON, XML), интеграционных шин (ESB); семантический уровень – разработка единых справочников, классификаторов, обеспечивающих однозначную интерпретацию данных; организационный уровень – согласование бизнес-процессов и регламентов взаимодействия между различными системами и организациями.

Принцип открытости и стандартизации: Платформа должна использовать открытые стандарты и общепринятые технологии для обеспечения долговечности решения, упрощения интеграции между системами.

К ключевым стандартам в области КИС можно отнести TOGAF, ArchiMate, BPMN, ISO 27001, ISO 9001.

Принцип безопасности и конфиденциальности: Информационная безопасность является критически важным аспектом корпоративных платформ, особенно при работе с персональными данными клиентов и конфиденциальной бизнес-информацией. Платформа должна соответствовать принципам конфиденциальности, целостности и доступности.

Для обеспечения безопасности платформы нужно реализовать принцип наименьших привилегий, при котором пользователям предоставляются только те права, которые относятся к их должностным обязанностям; принцип многофакторной аутентификации с использованием нескольких независимых методов проверки подлинности пользователей; принцип шифрования данных для защиты информации при передаче (протоколы TLS/SSL) и хранении (протокол AES-256); принцип сегментации сети путем разделения инфраструктуры на изолированные зоны с контролем доступа между ними; принцип резервного копирования для обеспечения возможности восстановления данных при сбоях в работе платформы.

Принцип единого источника данных: Наличие централизованного хранилища данных, содержащего актуальную и согласованную информацию для исключения дублирования и противоречивости данных, упрощения процессов аналитики и принятия решений, снижения затрат на хранение и обслуживание данных.

Данный принцип реализуется через создание корпоративного хранилища данных, интегрированного со всеми функциональными модулями платформы.

2. Архитектурные решения единой информационной платформы

Многоуровневая и микросервисная архитектура: Единая информационная платформа строится по многоуровневому принципу, который включает уровни представления, бизнес-логики, доступа к данным и хранения. Это обеспечивает четкое разделение ответственности в системе.

На смену монолитным приложениям приходит микросервисная архитектура. Система делится на набор небольших независимых сервисов, каждый из которых отвечает за свою бизнес-функцию. Это дает ряд преимуществ: независимость разработки и развертывания системы; гибкое масштабирование системы (можно распределять нагрузку на те сервисы, которые используют больше ресурсов); повышенная отказоустойчивость системы.

Для того, чтобы управлять микросервисами нужно использовать API-шлюзы, сервисы обнаружения и механизмы обеспечения устойчивости. Микросервисы обмениваются между собой данными через API, поэтому необходимо иметь стандартизированный подход к их разработке, версионности и управлению. Для документирования API могут использоваться стандарты OpenAPI/Swagger, которые облегчают взаимодействие между командами разработчиков.

Роль облачных технологий и интеграционной шины: Использование облачной инфраструктуры (IaaS, PaaS) обеспечивает эластичность платформы, снижает капитальные затраты и повышает доступность систем. Для сложных сред с большим количеством разнородных систем стоит применять интеграционную шину предприятия (ESB), которая выступает центральным узлом для трансформации и маршрутизации данных между всеми подключенными системами.

Во время построения архитектуры единой информационной платформы важно понимать, что, как и в любом крупном проекте по цифровой трансформации можно столкнуться со сложностями в виде организационного сопротивления, если сотрудники будут сопротивляться изменениям привычных рабочих процессов. Также присутствуют аспекты высокой первоначальной стоимости платформы и риска перегруженности платформы функциональностью. Чтобы избежать организационных проблем стоит проводить работу по изменению корпоративной культуры в компании, в будущем необходимо проводить комплексное обучение персонала по работе на новой платформе. Для урегулирования вопроса высокой первоначальной

стоимости можно использовать метод приоритезации критически важной функциональности, а только потом переходить к реализации второстепенной, либо растягивать проекты по внедрению функциональности по времени, что снизит нагрузку на бюджетирование. Со стороны риска перегруженности функциональностью рекомендуется использовать agile-подход, запуская сначала минимально жизнеспособный продукт и постепенно наращивая функционал платформы.

3. Технологический стек единой информационной платформы

Технологический стек платформы подбирается с учетом специфики каждой конкретной организации. Типичный технологический стек должен включать в себя несколько уровней представления: системы управления базами данных (PostgreSQL, MongoDB); серверные платформы (Apache для обработки потоков данных, Python, Java для разработки бизнес-логики); инструменты мониторинга (Prometheus, Grafana); системы непрерывной интеграции (Jenkins); централизованное хранилище логов.

4. Управление жизненным циклом единой информационной платформы

Внедрение и развитие единой информационной платформы включает в себя следующие этапы:

1. Стратегическое планирование – формирование видения, определение целей и задач платформы, разработка ИТ-стратегии организации.
2. Проектирование архитектуры – разработка целевой архитектуры, выбор технологий, определение интеграционных точек.
3. Разработка и тестирование платформы – итеративный подбор или разработка функциональных модулей, проведение тестирования.
4. Развертывание и переход – поэтапное внедрение платформы, перенос данных из существующих систем, обучение пользователей.
5. Эксплуатация и поддержка – мониторинг работоспособности платформы, устранение инцидентов, регулярное обновление компонентов.
6. Развитие и модернизация – добавление новой функциональности, улучшение производительности платформы, адаптация к изменяющимся требованиям бизнеса.

Для управления изменениями архитектуры платформы целесообразно использовать методологию TOGAF ADM.

Чтобы внедрение платформы было успешным, желательно делать акцент на управление изменениями. Необходимо создать видение ИТ-стратегии, которое будет четко доноситься до сотрудников организации, также важно работать с внутренними или внешними консультантами, которые будут выступать амбассадорами проекта.

Многие программы цифровой трансформации терпят неудачу как раз из-за этого аспекта, поэтому план по управлению изменениями должен быть разработан еще до начала внедрения системы.

5. Ключевые показатели эффективности единой информационной платформы

Метрики эффективности будут представлены тремя разными типами.

Технические показатели включают доступность системы – процент времени, в течение которого система доступна пользователям (целевое значение 99.5% и выше); время отклика системы – среднее время обработки запросов пользователей; пропускную способность системы – количество транзакций, обрабатываемых системой в единицу времени.

Важно отметить, что для организаций в сфере коммерческой недвижимости необходимо отслеживать время, в течение которого система может быть недоступна (RTO), а также максимальный объем данных, которые организация может потерять (RPO). Для критических ситуаций должен быть разработан план восстановления системы с использованием резервных копий, либо переключением на резервные системы.

Бизнес-показатели включают сокращение времени выполнения бизнес-процессов; снижение операционных затрат на ИТ-инфраструктуру; рост удовлетворенности пользователей; окупаемость инвестиций в ИТ (ROI).

Показатели качества данных включают полноту данных – процент заполненных обязательных полей; актуальность данных – частота обновления информации; согласованность данных – отсутствие противоречий между различными источниками информации.

6. Практические рекомендации по внедрению единой информационной платформы

На основе анализа успешных проектов цифровой трансформации организаций в сфере коммерческой недвижимости можно утверждать, что успех внедрения зависит от поэтапности внедрения, вовлечения всех

заинтересованных сторон в процесс внедрения, документирования архитектуры платформы, обучения персонала, мониторинга и аналитики.

Сначала организации необходимо запустить тестовый проект на одном направлении бизнеса, тщательно протестировать все функции, собрать обратную связь от пользователей и решить выявленные проблемы до полного развертывания платформы.

Заключение

Построение единой информационной платформы для цифровой трансформации организаций в сфере коммерческой недвижимости требует системного подхода, основанного на следовании фундаментальным принципам архитектуры корпоративных систем. Ключевыми факторами успеха являются модульность, масштабируемость, интероперабельность, безопасность и стандартизация.

Применение современных архитектурных методов, таких как микросервисная архитектура, использование облачных технологий и интеграционных шин, позволяет создать гибкую и адаптивную платформу, способную эффективно поддерживать бизнес-процессы.

Цифровая трансформация в сфере коммерческой недвижимости – это непрерывный процесс совершенствования, требующий постоянного мониторинга эффективности, адаптации к изменяющимся условиям рынка и внедрения новых технологических решений. Правильно спроектированная и реализованная информационная платформа обеспечивает конкурентные преимущества и устойчивое развитие бизнеса в этой сфере.

В будущем единые информационные платформы в сфере коммерческой недвижимости, вероятнее всего, шагнут в сторону интеграции с системами интернета вещей (IoT), искусственным интеллектом и цифровыми двойниками, что увеличивает актуальность внедрения таких платформ на сегодняшний день.

Список литературы

1. ГОСТ Р ИСО/МЭК 27001–2021. Информационная безопасность. Системы менеджмента ИБ. Требования: дата введения 2022-01-01. — М.: Рос. ин-т стандартизации, 2022. — 28 с.

2. Бобровников, А.Э. Введение в управление проектами внедрения ERP-систем / А.Э. Бобровников. — 2-е изд. — 1С-Паблишинг, 2025. — 320 с.

3. Ильин, И.В. Цифровая трансформация как фактор формирования архитектуры и ИТ-архитектуры предприятия / И.В. Ильин, А.И. Лёвина, А.С. Дубгорн // Научный журнал НИУ ИТМО.

Серия: Экономика и экологический менеджмент. — 2019. — № 3. — С. 50—55.

4. Морозова О.И., Семенихина А.В. Внедрение и использование IT-технологий в системе менеджмента организации // Использование информационных технологий в различных сферах деятельности. - Гомель: Белорусский торгово-экономический университет потребительской кооперации, 2024. - с. 104-108.

5. Киреев В. Г., Мальцев С. В. и др. Филиальная сеть. Развитие и управление. - Питер, 2007. - 352 с.

6. Назаров, Д.М. BI-технологии и корпоративные информационные системы в оптимизации бизнес-процессов / Д.М. Назаров, С.В. Бегичева, Д.А. Азаров. — Екатеринбург: Уральский гос. экон. ун-т, 2018. — 121 с.

7. Чурин, А. Как внедрить CRM. Опыт проектов amoCRM и Битрикс24 / А. Чурин. — СПб.: Питер, 2023. — 336 с.

8. Соколов, Р.В. Проектирование и эксплуатация информационных систем: учеб. пособие / Р.В. Соколов, И.Л. Андреевский. — СПб.: СПбГЭУ, 2023. — 348 с.

9. Свод знаний по управлению бизнес-процессами: BPM CBOK 4.0 / Джек Х., Дэн М. и др. — М.: Альпина Паблишер, 2023. — 503 с.

10. Талонов А.В. и др. Управление недвижимостью. - Юрайт, 2025. - 412 с.

Сравнительный анализ подходов к сбору и формализации требований к программному обеспечению

Д. Д. Скибин, email: d.sckibin2017@yandex.ru

Воронежский государственный университет

Аннотация. *В статье рассматривается проблема оптимизации производительности реляционных баз данных путем перехода от 128-битных случайных идентификаторов (UUID v4) к последовательным идентификаторам (BIGINT). Предложен и реализован программный комплекс, осуществляющий автоматизированную миграцию данных, включающий модули анализа зависимостей на основе теории графов, пакетной трансформации данных и гибридного маппинга идентификаторов. Особое внимание уделено алгоритму топологической сортировки для соблюдения целостности и методу синхронизации данных в режиме реального времени (Memory-Filtered Delta Sync). Приведены результаты экспериментального внедрения системы на массиве в 1 000 000 записей, подтверждающие сокращение объема индексов на 43,2% и повышение скорости операций объединения.*

Ключевые слова: *UUID, BIGINT, B-tree индексы, миграция баз данных, топологическая сортировка, PostgreSQL, ссылочная целостность, ETL.*

Введение

Автоматизация процесса трансформации идентификаторов в крупномасштабных реляционных базах данных

В современных распределенных системах тип данных UUID является стандартом реализации механизма уникальности записей, облегчая масштабирование и предотвращая конфликты при слиянии данных. Однако при значительных объемах данных проявляются фундаментальные недостатки UUID, обусловленные спецификой его обработки в реляционных СУБД.

Основная проблема заключается в несоответствии псевдослучайного типа UUID v4 механизмам оптимизации B-tree индексов. Низкая локальность данных и увеличение размера ключа до 128 бит приводят к «налогу на хранение» и раздуванию вторичных индексов. Случайное распределение значений провоцирует частые разделения страниц, снижая коэффициент их заполнения до 50-60%. Такая фрагментация вызывает деградацию кэша и рост операций

© Скибин Д. Д., 2026

Такая фрагментация вызывает деградацию кэша и рост операций случайного ввода-вывода, что замедляет выполнение операций объединения более чем в два раза.

Переход на последовательные идентификаторы типа BIGINT позволяет восстановить эффективность индексации и оптимизировать использование дискового пространства. При этом ручная трансформация идентификаторов в эксплуатируемых системах со сложной иерархией связей сопряжена с рисками нарушения ссылочной целостности.

Целью работы является проектирование программного комплекса и исследование процесса трансформации базы данных. В рамках исследования решаются задачи динамической интроспекции метаданных СУБД, применения теории графов для реализации механизма топологической сортировки таблиц в зависимом положении, а также реализации механизмов высокопроизводительной пакетной миграции и инкрементальной синхронизации данных в реальном времени.

1. Анализ проблемы и теоретические предпосылки

Эффективность B-tree индексов PostgreSQL напрямую зависит от локальности данных. В отличие от последовательных ключей, идентификаторы, основанные на псевдослучайных механизмах, вызывают интенсивную фрагментацию и снижение коэффициента заполнения страниц. Это удваивает объем занимаемой памяти индексом, но и вызывает эффект усиления записи, нагружая память хост-устройства. При деградации локальности резко возрастает количество ошибочных срабатываний в буфере кэша из-за вынужденного перехода к операциям случайного ввода-вывода.

Процесс трансформации осложняется соблюдением ссылочной целостности, что моделируется задача топологической сортировки ориентированного графа. В целях уменьшения влияния на структуру целевой схемы маппинг идентификаторов выносится в изолированный сервисный слой с доступом за константное время O(1). В условиях крупномасштабных миграций надёжность миграций обеспечивается внедрением пакетной обработки с контрольными точками, что предотвращает избыточное логирование транзакций и позволяет системе корректно обрабатывать циклические зависимости.

Системным фактором является влияние случайных идентификаторов на объем журналов предзаписи (Write-Ahead Log, WAL). Вставка UUID в произвольные страницы индекса вынуждает СУБД фиксировать изменения целых страниц, это приводит к лавинообразному увеличению объема журналов предзаписи. В

распределенных архитектурах это приводит к снижению производительности из-за ограничения пропускной способности диска и увеличивает замедление репликации между узлами кластера. Переход от случайно генерируемых идентификаторов к последовательным ключам позволяет снизить издержки хранения и улучшает стабильность и масштабируемость инфраструктуры за счет стабилизации и систематизации потока записи и снижения нагрузки на механизмы создания резервных копий состояния.

2. Архитектура и реализация программного комплекса

Набор программных решений реализует модульный ETL-цикл, обеспечивающий интроспекцию и трансформацию динамических конфигураций данных. Программный модуль MetadataReader выполняет анализ системных каталогов information_schema, на основе которых DependencyResolver формирует граф зависимостей. Для обеспечения порядка миграции и соблюдения целостности применяется алгоритм топологической сортировки, позволяющий последовательно обрабатывать таблицы от независимых справочников к связанным сущностям.

Для высокопроизводительной замены идентификаторов разработан сервис, использующий двухслойную архитектуру. Уровень L1 реализован в быстрой оперативной памяти (куче) на базе ConcurrentHashMap с ограничением для контроля нагрузки на сборщик мусора. Уровень L2 представляет собой таблицу в целевой СУБД с составным B-tree индексом по полям table_name и old_uuid. Подобная модель гарантирует эффективную обработку при объемах данных, значительно превышающих доступную память.

Класс DataMigrator, реализующий модуль миграции, минимизирует сетевые издержки за счет пакетной обработки с активацией параметра драйвера reWriteBatchedInserts. Данная настройка позволяет СУБД преобразовать в многострочные команды поток инструкций, снижая нагрузку на транзакционный лог. Интеграция механизма возврата идентификаторов обеспечивает получение новых ключей, а также актуализацию маппинга без избыточных запросов, что позволяет достичь пропускной способности до 900 записей в секунду.

3. Методика инкрементальной синхронизации данных

Для поддержания консистентности данных в условиях эксплуатации базы данных, без прерывания в системе разработан механизм Memory-Filtered Delta Sync. В отличие от традиционных методов дедупликации на уровне СУБД (например, ON CONFLICT или SELECT EXISTS), вызывающих избыточную нагрузку потоков ввода и выводов, а также

сетевые задержки, предложенный подход переносит логику фильтрации на прикладной уровень.

Алгоритм базируется на использовании структуры HashSet в короткосрочной быстрой памяти, в которую загружаются все ранее мигрировавшие идентификаторы из реестра маппинга. Использование хеш-таблицы обеспечивают сложность поиска O(1), позволяя мгновенно верифицировать входящий поток данных. Только новые записи, отсутствующие в памяти, подвергаются трансформации и вставке, что снижает количество транзакций в целевой базе данных.

Экспериментальная оценка при обработке данных, на базе основных таблиц по 1 000 000 строк в каждой показала, что время верификации таблицы составляет ~16 мс. Это подтверждает эффективность алгоритма для актуализации хранилищ данных в режиме реального времени и реализации миграции с малым временем простоя.

4. Экспериментальное исследование

Апробация системы проводилась в изолированной среде на базе СУБД PostgreSQL 15 с использованием имитационной модели E-commerce (20 таблиц/сущностей). Для обеспечения репрезентативности объем выборки составил 1 000 000 записей для каждой ключевой сущности системы.

Сбор метрик осуществлялся программным модулем ResultCollector посредством системных функций pg_indexes_size. В результате выполнения ряда экспериментов подтверждено сокращение среднего объема индексов, реализованных структурой B-tree, на 43,2% с 37,8 МБ для UUID до 21,4 МБ для BIGINT. Данный эффект обусловлен как двукратным уменьшением веса ключа, так и повышением плотности заполнения страниц индекса за счет последовательного характера вставок, что минимизирует фрагментацию структуры.

Сравнительный анализ временных затрат выявил устойчивое преимущество целочисленных идентификаторов в ресурсоемких операциях. Наиболее выраженная динамика зафиксирована в операциях объединения таблиц: использование BIGINT обеспечивает ускорение процесса в 2,2 раза (1,239 с против 2,777 с для UUID v4) за счет лучшей локальности данных в кэше и снижения нагрузки на CPU при сравнении компактных ключей.

Заключение

В ходе работы спроектирован и реализован программный комплекс для автоматизированной трансформации идентификаторов в реляционных СУБД. Экспериментально подтверждена гипотеза о «налоге на хранение» при использовании UUID переход

на последовательные ключи BIGINT обеспечил экономию дискового пространства для индексов на 43,2% (на выборке в 1 млн записей), что существенно повысило эффективность использования оперативной памяти и дисковой подсистемы.

Высокая производительность системы достигнута за счет применения топологической сортировки графов зависимостей и внедрения гибридного сервиса маппинга. Это позволило корректно обрабатывать структуры из 20+ взаимосвязанных таблиц со скоростью до 900 записей в секунду. Реализованный алгоритм «Memory-Filtered Delta Sync» доказал свою эффективность, обеспечив актуализацию данных в режиме реального времени без значительной нагрузки на исходную СУБД.

Список литературы

1. Leach P., Mealling M., Salz R. A Universally Unique Identifier (UUID) URN Namespace // RFC 4122. – 2005.
2. Peabody B., Davis K. New UUID Formats // Internet-Draft. – 2021.
3. Silberschatz A., Korth H. F., Sudarshan S. Database System Concepts. – 6th ed. – McGraw-Hill, 2011. – 1376 p.
4. Garcia-Molina H., Ullman J. D., Widom J. Database Systems: The Complete Book. – 2nd ed. – Upper Saddle River : Pearson Prentice Hall, 2009. – 1200 p.
5. Cormen T. H. Introduction to Algorithms / T. H. Cormen [и др.]. – 3rd ed. – Cambridge : MIT Press, 2009. – 1292 p.
6. Домбровская Г., Новиков Б., Бейликова А. Оптимизация запросов в PostgreSQL. – М. : ДМК Пресс, 2022. – 278 с.
7. Скибин Д. Д. Сравнительный анализ типов данных, используемых для идентификации записей в современных базах данных // Материалы конференции ВГУ. – 2025.
8. PostgreSQL Documentation. Indexes [Электронный ресурс]. – Режим доступа: https://www.postgresql.org/docs/current/indexes.html (дата обращения: 12.01.2026)

Обеспечение связности компонентов электронных документов на основе бинарного дерева Меркла

Д. И. Тали, к.т.н., email: dimatali@mail.ru

Краснодарское высшее военное училище

Аннотация. *Выполнен анализ системных свойств электронного документа, что позволяет сформировать его обобщенную модель и обеспечить интегративную целостность, достигаемую за счет связности его компонентов (контента и метаданных). Предложенное решение позволит повысить уровень защищенности электронных документов, обрабатываемых автоматизированными информационными системами юридически значимого электронного документооборота, в условиях преднамеренных и непреднамеренных воздействий уполномоченных пользователей (инсайдеров).*

Ключевые слова: *автоматизированная информационная система юридически значимого электронного документооборота, контент, метаданные, управление электронным документом, целостность электронного документа.*

Введение

Цифровая трансформация общества приводит к повсеместному использованию электронных документов (ЭлД) в различных процессах управления, в том числе и юридически значимых. Однако, существующее положение дел не позволяет произвести качественную модернизацию автоматизированных информационных систем юридически значимого электронного документооборота (АИС ЮЗ ЭД), ввиду устаревшей нормативно-правовой и терминологической базы не поспевающей за темпами цифрового развития [1-3]. Причиной чего является отождествление традиционного «бумажного» документа с электронным, несмотря на техническую природу происхождения последнего. Следствием данного процесса стало формирование структуры ЭлД, представляющей собой совокупность контента и его метаданных, что указывает на дихотомию исследуемого объекта, и двусмысленность в формулировке целевого предназначения документных систем. Вместе с тем, нарастающие темпы цифровизации заставляют системы данного типа совершенствоваться и развиваться до уровня систем обслуживания информации, характеризующихся централизацией обслуживаемых процессов и масштабируемостью обрабатываемой информации. В

© Тали Д. И., 2026

сложившихся условиях основной задачей становится обеспечение целостности не столько ЭлД, сколько самой системы (инфраструктуры), обслуживающей циркулирующую в ней информацию [4]. Первым шагом на пути к решению этой задачи является обеспечение интегративной целостности электронных документов, достигаемой за счет связности их компонентов (контента и метаданных) [5].

1. Анализ системных свойств электронного документа

Анализ трудов специалистов в области теории систем и системного анализа позволяет интерпретировать понятие «система» (от греч. Systema – целое, составленное из частей; соединение), как множество элементов, находящихся в отношениях и связях друг с другом, образующих определенную целостность, единство [6-9].

Следует оговорить относительность этих понятий. Так, элемент данной системы из-за иерархической структуры мира сам оказывается системой со своими элементами. Фиксация системы делит мир на две части – систему и среду. При этом подчеркивают большую силу связей элементов внутри системы, чем с элементами среды, что и привело к возникновению дуализма ЭлД.

То есть ЭлД, с технической точки зрения, являющийся элементом АИС ЭД как объект воздействия, с позиции системного анализа может быть рассмотрен как система из двух элементов – контента и метаданных. Причем их функциональная зависимость очевидна – без создания контента невозможно формирование метаданных, в свою очередь без метаданных невозможно управление контентом, в то время как в среде АИС ЭД эти элементы имеют разную степень защиты.

В связи с чем, в целях формирования ЭлД, как целостного объекта, предлагается принцип интегративности, базирующийся на фундаментальном свойстве систем, которое может быть сформулировано следующим образом: функции системы не есть простая сумма функций составляющих ее частей. Следовательно, нельзя получить адекватного представления о системе, изучая ее отдельные компоненты.

Целостность такого объекта проявляется в двух аспектах: структурном и процессуальном, другими словами целостность – это пространственная связность структурных элементов и временная согласованность их существования.

Пространственная связность будет заключаться в наличии устойчивых и относительно мощных связей между структурными компонентами объекта, чего возможно добиться путем разработки механизма связности контента ЭлД и его метаданных.

2. Обеспечение интегративной целостности электронного документа

Разработанное решение реализуется следующим образом (рис. 1).

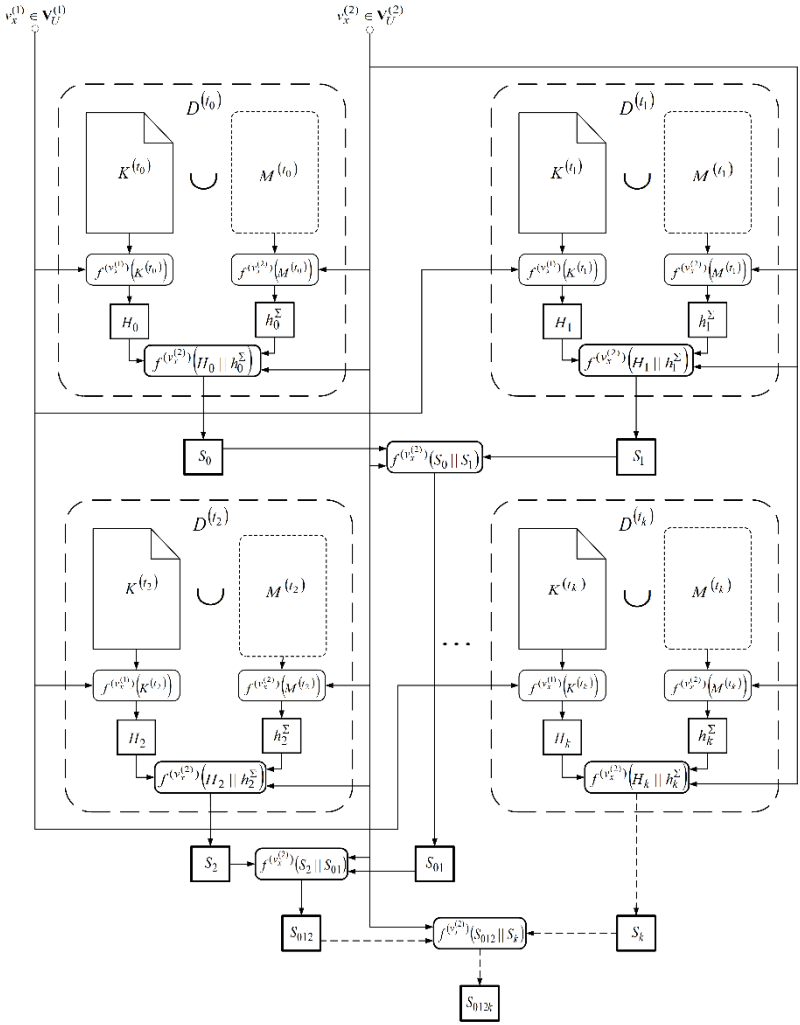


Рис. 1. Функциональная модель формирования ЭлД

Множество ключей $\mathbf{V}_U$ разбивается на два подмножества: $\mathbf{V}_U \in \left\{\mathbf{V}_U^{(1)}, \mathbf{V}_U^{(2)}\right\}$ – ключи формирования ЭлД, и $\mathbf{V}_U^* \in \left\{\mathbf{V}_U^{*(1)}, \mathbf{V}_U^{*(2)}\right\}$ – ключи проверки целостности ЭлД. В рассматриваемом варианте реализации данного способа ключи $\mathbf{V}_U^{(1)}$ являются ключами исполнителя (автора) электронного документа, ключи $\mathbf{V}_U^{(2)}$ – внутренними системными ключами.

В момент формирования t_0 ЭлД в памяти системы обработки данных образуется информационный блок $D^{(t_0)} = \left\{K^{(t_0)}, M^{(t_0)}\right\}$, представляющий собой электронный документ, создаваемый подмножеством контента $K^{(t_0)}$ и подмножеством записей метаданных $M^{(t_0)} \in \{m_{01}, m_{02}, m_{03}, \ldots, m_{0n}\}$, над которыми одновременно производятся операции криптографического преобразования: $H_0 = f^{(v_x^{(1)})}\left(K^{(t_0)}\right)$ на ключах $v_x^{(1)} \in \mathbf{V}_U^{(1)}$, $h_{01} = f^{(v_x^{(2)})}(m_{01})$, $h_{02} = f^{(v_x^{(2)})}(m_{02})$, $h_{03} = f^{(v_x^{(2)})}(m_{03})$, …, $h_{0n} = f^{(v_x^{(2)})}(m_{0n})$ на ключах $v_x^{(2)} \in \mathbf{V}_U^{(2)}$.

Над вычисленными значениями сигнатур $h_{01}, h_{02}, h_{03}, \ldots, h_{0n}$ записей статических метаданных производятся двухместные операции конкатенации: $h_{01} \,||\, h_{02}$, $h_{03} \,||\, h_{0n}$, над результатами которых выполняются операции криптографического преобразования на ключах $v_x^{(2)} \in \mathbf{V}_U^{(2)}$: $h_{0102} = f^{(v_x^{(2)})}(h_{01} \,||\, h_{02})$, $h_{030n} = f^{(v_x^{(2)})}(h_{03} \,||\, h_{0n})$.

После чего над вновь вычисленными сигнатурами последовательно выполняются операции конкатенации и криптографического преобразования на ключах $v_x^{(2)} \in \mathbf{V}_U^{(2)}$: $h_0^{\Sigma} = f^{(v_x^{(2)})}(h_{0102} \,||\, h_{030n})$, результатом которой является формирование корневой сигнатуры h_0^{Σ} записей метаданных.

Затем операции конкатенации и криптографического преобразования на ключах $v_x^{(2)} \in \mathbf{V}_U^{(2)}$ производятся над значением сигнатуры H_0 контента и корневой сигнатурой h_0^{Σ} записей метаданных, в целях получения общей корневой сигнатуры ЭлД, созданного в момент времени t_0: $S_0 = f^{(v_x^{(2)})}\left(H_0 \,||\, h_0^{\Sigma}\right)$.

После произведенных операций ЭлД вводится в систему, причем во время его жизненного цикла $t_1, t_2, \ldots, t_k$ при выполнении записи и редактирования повторно производятся вышеописанные операции.

При этом общая корневая сигнатура S_1 ЭлД, вычисленная в момент времени t_1, конкатенирует с общей корневой сигнатурой S_0 ЭлД, вычисленной в момент времени t_0, с одновременным криптографическим преобразованием на ключах $v_x^{(2)} \in \mathbf{V}_U^{(2)}$, тем самым образуя сигнатуру жизненного цикла S_{01} ЭлД: $S_{01} = f^{(v_x^{(2)})}(S_0 \,||\, S_1)$.

По мере изменения ЭлД операции повторяются в аналогичном порядке в соответствии с моментами времени $t_3, t_4, \ldots, t_k$ его жизненного цикла (рис. 1) [10].

Контент $K^{(t_i)}$ и вычисленные значения его сигнатур H_i, записи метаданных $m_{i1}, m_{i2}, m_{i3}, \ldots, m_{in}$, их сигнатуры $h_{i1}, h_{i2}, h_{i3}, \ldots, h_{in}$, общие сигнатуры h_{i1i2}, h_{i3in} записей метаданных, корневые сигнатуры h_i^{Σ} записей метаданных, общие корневые сигнатуры S_i ЭлД, а также корневые сигнатуры S_{in} жизненного цикла ЭлД сохраняются в памяти системы обработки данных и используются для проверки целостности ЭлД в моменты времени его жизненного цикла $t_0, t_1, t_2, \ldots, t_k$.

Контроль целостности ЭлД осуществляется на основе извлечения из памяти системы обработки данных контента $K^{*(t_i)}$ и вычисленных значений его сигнатур H_i^*, записей метаданных $m_{i1}^*, m_{i2}^*, m_{i3}^*, \ldots, m_{in}^*$ и их сигнатур $h_{i1}^*, h_{i2}^*, h_{i3}^*, \ldots, h_{in}^*$, общих сигнатур h_{i1i2}^*, h_{i3in}^* записей метаданных, корневых сигнатур $h_i^{\Sigma *}$ записей метаданных, общих корневых сигнатур S_i^* ЭлД, а также корневых сигнатур S_{in}^* жизненного цикла ЭлД, прошедших процедуру хранения и подлежащих контролю целостности.

После чего выполняются повторные операции криптографического преобразования над извлеченным контентом $K^{*(t_i)}$ и записями метаданных $m_{i1}^*, m_{i2}^*, m_{i3}^*, \ldots, m_{in}^*$. В результате чего вновь вычисленные сигнатуры: H_i^{**}; $h_{i1}^{**}, h_{i2}^{**}, h_{i3}^{**}, \ldots, h_{in}^{**}$; $h_{i1i2}^{**}, h_{i3in}^{**}$; $h_i^{\Sigma **}$; S_i^{**}; S_{in}^{**}

попарно сравниваются с ранее извлеченными: H_i^*; $h_{i1}^*, h_{i2}^*, h_{i3}^*, \ldots, h_{in}^*$; h_{i1i2}^*, h_{i3in}^*; $h_i^{\Sigma*}$; S_i^*; S_{in}^*.

Заключение об отсутствии нарушения целостности ЭлД делается при выполнении равенст $f^{(v_x^{(2)})}\left(m_{ij}\right)$ в: $S_{in}^{**} = S_{in}^*$; $S_i^{**} = S_i^*$; $H_i^{**} = H_i^*$; $h_i^{\Sigma**} = h_i^{\Sigma*}$; $h_{i1i2}^{**} = h_{i1i2}^*, h_{i3in}^{**} = h_{i3in}^*$; $h_{i1}^{**} = h_{i1}^*, h_{i2}^{**} = h_{i2}^*, h_{i3}^{**} = h_{i3}^*, \ldots, h_{in}^{**} = h_{in}^*$, в противном случае делается заключение о нарушении целостности для соответствующих номеров сигнатур.

В качестве криптографического преобразования , используется либо ключевая хеш-функция, при этом $\mathbf{V}_U = \mathbf{V}_U^*$, либо электронная подпись, при этом $\mathbf{V}_U \neq \mathbf{V}_U^*$ [10].

3. Обобщенная модель электронного документа

На основании вышеизложенного обобщенную модель ЭлД можно представить в графическом виде (рис.2), где ядром модели является ЭлД, как информационный объект, с присущей для него функцией юридической значимости, обеспечиваемой механизмом ЭП, а оболочкой сервисные функции, реализованные посредством представления его метаданных в виде юридически значимой базы метаданных, манипуляции над которой осуществляются посредством системы управления [5].

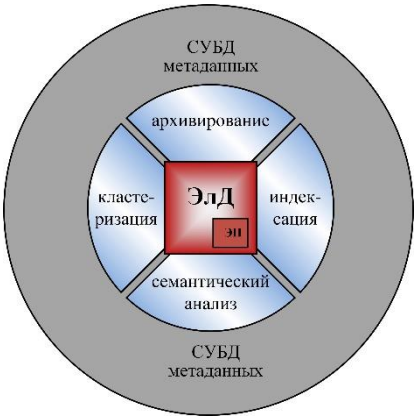


Рис. 2. Обобщенная модель ЭлД в АИС ЮЗ ЭД

Реализовать подобную модель возможно на основе использования графовых баз данных, преимуществом которых является прирост производительности при работе со взаимосвязанными данными, по сравнению с реляционными базами, в условиях увеличения объема хранимых данных.

В таком случае для построения новой структуры ЭлД будет использоваться мультиграф. При этом такая структура ЭлД, ввиду постоянного изменения в процессе жизненного цикла, преобразуется в гиперграф, то есть граф, в котором между собой связаны более двух вершин, где вершинами являются контент с его метаданными, а ребрами действия совершаемые над этими вершинами (рис. 3).

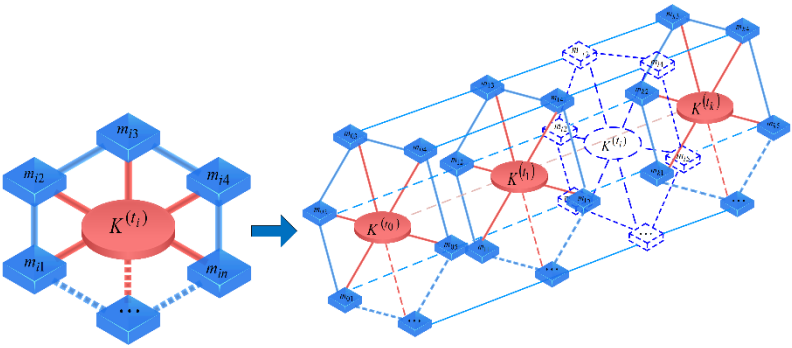


Рис. 3. Графовая модель ЭлД

Заключение

Предлагаемые решения могут быть использованы в автоматизированных информационных системах юридически значимого электронного документооборота, позволяя повысить уровень защищенности обрабатываемых электронных документов, а также реализовать предъявляемые требования по обеспечению связности его компонентов.

Список литературы

1. Шваб, К. Четвертая промышленная революция / К. Шваб – М: Эксмо, 2021. – 208 с

2. Ларин, М. В. Электронные документы: теоретические аспекты / М. В. Ларин // Самарский архивист. – 2021. – № 2. – С. 3-9.

3. Бороздина, А. Г. К вопросу о нормативном регулировании структурных компонентов электронных документов / А. Г. Бороздина // Вестник ВНИИДАД. – 2019. – № 2. – С. 54-60.

4. Ларин, М. В. Управление документами в цифровой экономике: организация, регламентация, реализация / М. В. Ларин [и др.] – М.: РГГУ, 2021. – 242 с.

5. Тали, Д. И. Принцип целостности и интегративности в формировании электронного документа / Д. И. Тали // Правовая информатика. – 2022. – № 3. – С. 72-83. – DOI: 10.21681/1994-1404-2022-3-72-83

6. Волкова, В. Н. Управление устойчивым развитием социально-экономической системы в условиях внедрения эмерджентных технологий / В. Н. Волкова, В. В. Ланкин // В сборнике: Технологическая перспектива в рамках евразийского пространства: новые рынки и точки экономического роста. Материалы 4-ой Международной научной конференции. – 2018. – С. 315-320.

7. Калинин, В. Н. Теория системных исследований как основа фундаментальной общенаучной подготовки молодых ученых / В. Н. Калинин, Б. В. Соколов // Стратегические приоритеты. – 2017. – № 2 (14). – С. 127-137.

8. Новиков, Д. А. Классификации систем управления / Д. А. Новиков // Проблемы управления. – 2019. – № 4. – С. 27-42.

9. Игнатьев, М. Б. Адаптационные возможности сложных систем / М. Б. Игнатьев, П. И. Макин // Международная конференция по мягким вычислениям и измерениям. – 2017. – Т. 2. – С. 307-310.

10. Способ обеспечения интегративной целостности электронного документа [Текст] : пат. 2812304 Российская Федерация : МПК G06F 21/64, G06F 40/194 / Тали Д. И. и др. ; заявитель и патентообладатель Краснодарское высшее военное училище. – № 2022124983 ; заявл. 22.09.2022 ; опубл. 29.01.2024, Бюл. № 4 – 32 с.

Секция 5
Прикладное моделирование и E-business

Разработка системы для автоматизации ключевых процессов дополнительного образования

А. И. Бригадина, email: brigadinaai@gmail.com
Е. А. Копытина, email: zhemkaterina@yandex.ru

Воронежский государственный университет

Аннотация. *В статье рассматривается разработка системы для автоматизации процессов дополнительного образования. Отмечается актуальность внедрения прикладных информационных решений в условиях роста объёма административных данных и необходимости упрощения документооборота в образовательных организациях*

Ключевые слова: *автоматизация процессов, дополнительное образование, информационная система, веб-приложение, обработка данных, электронные таблицы, документооборот, генерация документов.*

Введение

В условиях цифровизации автоматизация рабочих процессов становится одной из ключевых задач для организаций. Рост объёма данных и сопроводительной документации приводит к тому, что ручная обработка информации затрудняет выполнение административных функций и усложняет управление процессами. Несмотря на наличие большого числа программных решений, многие из них не учитывают специфику конкретных организаций или обладают избыточной функциональностью, что особенно актуально экономики для небольших структур, как показано в [1-7].

Сфера образования, включая дополнительное образование, также сталкивается с данными проблемами. Подразделения работают с электронными таблицами, отчётами, приказами и рассылками, что нередко приводит к разрозненности данных и дублированию операций. Дополнительное образование характеризуется высокой динамикой и требует оперативной обработки информации.

В связи с этим актуальной является разработка системы, обеспечивающей автоматизацию ключевых процессов дополнительного образования, включая формирование базы данных на основе файлов .xlsx, подготовку отчётности, рассылку уведомлений и генерацию приказов

© Бригадина А. И., Копытина Е. А., 2026

1. Постановка задачи

Данная статья посвящена системе для автоматизации процессов дополнительного образования, обеспечивающей обработку данных, формирование документации и поддержку организационных операций. Разрабатываемая система должна быть ориентирована на использование в условиях небольшой образовательной структуры и учитывать специфику работы с табличными данными и отчётностью.

Для достижения поставленной цели в работе предусматривается решение следующих задач:

- реализация системы в виде веб-приложения с разделением клиентской и серверной частей;
- поддержка загрузки и обработки файлов формата `.xlsx` с проверкой структуры и корректности данных;
- использование централизованного хранилища данных с возможностью выполнения выборок и обновлений;
- реализация взаимодействия клиентской и серверной частей посредством REST-ориентированного API;
- обеспечение генерации документов формата `.docx` на основе шаблонов;
- поддержка автоматизированной почтовой рассылки уведомлений;
- обеспечение возможности расширения функциональности без изменения базовой архитектуры системы;
- использование современных средств разработки, обеспечивающих производительность и стабильность работы приложения.

2. Средства реализации программного продукта

Для реализации серверной части системы использован язык программирования Python, обеспечивающий широкие возможности для работы с данными и интеграции с внешними библиотеками. В качестве серверного фреймворка выбран FastAPI, позволяющий создавать REST-ориентированные веб-сервисы с поддержкой асинхронной обработки запросов и автоматической генерацией документации API.

Реляционная модель лежит в основе работы с базой данных, что обеспечивает целостность данных и возможность выполнения сложных запросов. Для обработки файлов формата .xlsx применяются специализированные библиотеки Python, реализующие валидацию структуры данных, проверку корректности заполнения полей и дедупликацию записей. Генерация документов формата .docx осуществляется с использованием библиотеки python-docx, что позволяет

формировать приказы и служебные записки на основе шаблонов с плейсхолдерами и единым стилевым оформлением.

Клиентская часть системы реализована с использованием библиотеки React и сборщика Vite. Веб-интерфейс обеспечивает выполнение операций добавления, редактирования и просмотра данных, запуск процессов импорта, рассылки и генерации документов. Взаимодействие между клиентом и сервером осуществляется через HTTP-запросы с использованием библиотеки Axios.

Для упрощения развертывания и обеспечения воспроизводимости окружения используется контейнеризация на основе Docker и Docker Compose.

3. Реализация логики программного продукта

Логика работы программного продукта построена на последовательном взаимодействии отдельных функциональных модулей, каждый из которых отвечает за определённый этап обработки данных и выполнения операций. Такое разделение позволяет упорядочить внутренние процессы системы и обеспечить согласованную работу всех её компонентов.

Исходной точкой функционирования системы является загрузка и обработка данных. Пользователь взаимодействует с интерфейсом приложения, загружая исходные файлы с информацией о слушателях и образовательных программах. После этого данные проходят этап проверки и структурирования, в ходе которого они приводятся к единому формату и подготавливаются для дальнейшего использования. Подготовленные сведения сохраняются во внутреннем хранилище системы, что обеспечивает их централизованное хранение и повторное использование.

На основе хранимых данных осуществляется работа модуля отчётности и статистики. Он использует актуальную информацию для формирования сводных таблиц и аналитических показателей, отражающих текущее состояние образовательных программ. Полученные результаты могут использоваться как для внутреннего контроля, так и для подготовки официальной отчётной документации.

Отдельным элементом логики системы является модуль генерации документов. Он обеспечивает автоматическое создание приказов и иных документов на основе заранее подготовленных шаблонов. При формировании документа необходимые данные подставляются автоматически, что исключает ручное заполнение и снижает вероятность ошибок. Сформированные документы доступны для просмотра, редактирования и дальнейшего использования.

Взаимодействие с участниками образовательных программ реализуется через модуль почтовых рассылок. Он использует информацию из базы данных и позволяет формировать сообщения, адресованные отдельным группам слушателей или всем участникам сразу. Рассылки могут инициироваться как вручную, так и в рамках выполнения других операций, например, после формирования приказов или обновления данных. Интерфейс системы представлен на рис.1 и рис.2.

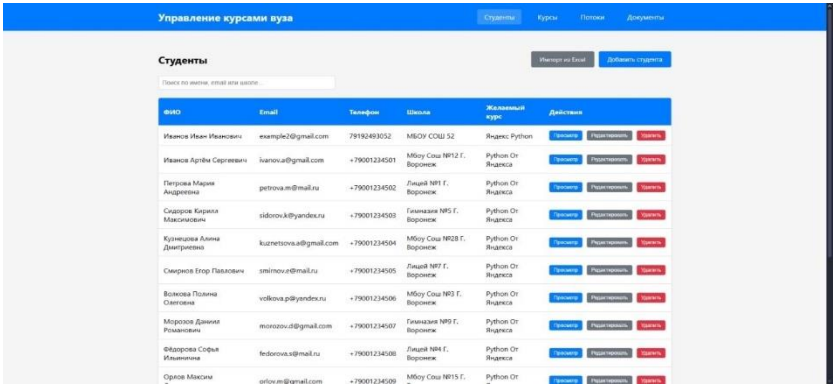


Рис. 1. Работа с контингентом

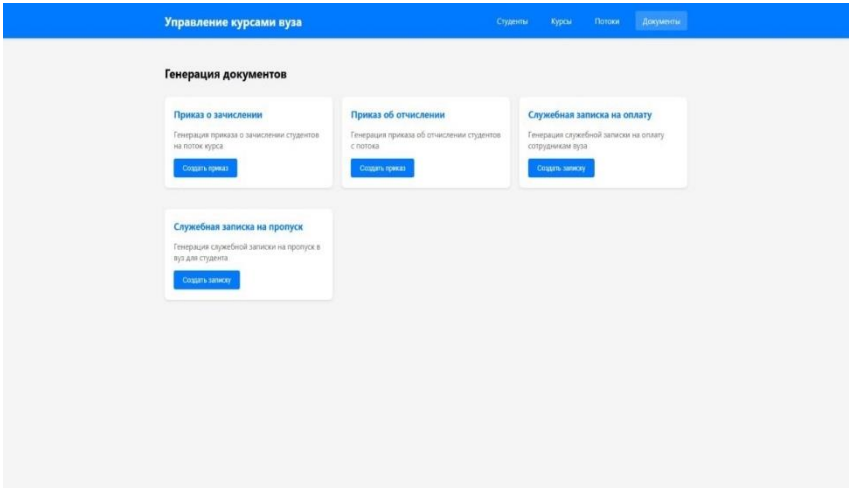


Рис. 2. Просмотр шаблонов документов и их генерация

В итоге логика работы программного продукта направлена на упрощение организационных процессов дополнительного образования и обеспечение целостности всех этапов обработки данных.

Заключение

Система, описанная в данной статье, представляет собой работоспособное программное решение, ориентированное на автоматизацию ключевых процессов дополнительного образования. Архитектура веб-сервиса и выбранные средства реализации обеспечивают возможность дальнейшего развития и адаптации системы под изменяющиеся требования.

К перспективным направлениям развития можно отнести расширение набора формируемых документов, реализацию системы разграничения прав доступа пользователей, интеграцию с внешними образовательными платформами, а также добавление аналитических модулей для углубленной обработки статистических данных. Дополнительным направлением доработки является внедрение механизмов журналирования действий пользователей и резервного копирования данных.

Результаты работы могут быть использованы в практической деятельности подразделений дополнительного образования, а также послужить основой для дальнейших исследований и разработок в области прикладных информационных систем для автоматизации административных процессов.

Список литературы

1. Бокова, О. А. Опытно-экспериментальная деятельность учреждения дополнительного образования как основа реализации программ дополнительного образования / О. А. Бокова, Д. Н. Рыбин, А. Г. Степанищев // Мир науки, культуры, образования. – 2020. – № 6(85). – С. 316-320. – DOI 10.24412/1991-5500-2020-685-316-320. – EDN YXMKTX.

2. Свидетельство о государственной регистрации программы для ЭВМ № 2025664443 Российская Федерация. Программа для автоматизации работы отдела интеллектуальной собственности : заявл. 20.05.2025 : опубл. 04.06.2025 / П. О. Филонов, Е. А. Копытина ; заявитель федеральное государственное бюджетное образовательное учреждение высшего образования «Воронежский государственный университет». – EDN HFQPRS.

3. Кисарин, А. С. Проблемы инноваций в дополнительном образовании в условиях цифровизации образования / А. С. Кисарин // Заметки ученого. – 2021. – № 6-1. – С. 159-162. – EDN USZYAL.

4. Свидетельство о государственной регистрации программы для ЭВМ № 2025664444 Российская Федерация. Программа для генерации пакета документов, сопутствующих процессу ГИА : заявл. 20.05.2025 : опубл. 04.06.2025 / С. И. Бутовецкая, Е. А. Копытина ; заявитель федеральное государственное бюджетное образовательное учреждение высшего образования «Воронежский государственный университет». – EDN BKYQIQ.

5. Смирнов, А. И. Инновации в дошкольном, общем и дополнительном образовании: от системного штампа - к эффективной деятельности / А. И. Смирнов // Наука и образование. Вестник Хабаровского края. – 2023. – № 1. – С. 40-45. – EDN PVNBJL.

6. Свидетельство о государственной регистрации программы для ЭВМ № 2024661693 Российская Федерация. Программа для проведения экспертизы работ проектных конкурсов : № 2024660532 : заявл. 13.05.2024 : опубл. 21.05.2024 / Я. В. Солодовникова, Е. А. Копытина ; заявитель федеральное государственное бюджетное образовательное учреждение высшего образования «Воронежский государственный университет». – EDN HDOJTC.

7. Бордюжа, Е. О. Разработка информационной системы, автоматизирующей генерацию шаблонов документов, сопутствующих прохождению учебных практик ФКН ФГБОУ ВО "ВГУ" / Е. О. Бордюжа, Е. А. Копытина, Н. С. Богданова // Труды молодых учёных факультета компьютерных наук ВГУ : Сборник статей. – Воронеж : Воронежский государственный университет, 2024. – С. 18-25. – EDN ITOFPW.

Реализация нейронечеткой модели для цифровой обработки изображений

В. А. Великодный, email: isnt1321@gmail.com
Е. А. Сирота, email: atoris@list.ru

Воронежский государственный университет

Аннотация. *В статье предложен новый подход к повышению робастности классификации изображений за счёт интеграции нечёткой логики в архитектуру свёрточной нейронной сети (CNN). Вместо традиционного полносвязного классификатора реализован гибридный нейронечёткий классификатор (ANFIS), обучаемый совместно со свёрточным энкодером end-to-end. Теоретически обоснованы преимущества подхода: способность явно моделировать неопределённость и потенциальная интерпретируемость решений через лингвистические правила. Экспериментально на медицинских данных (рентгенограммы лёгких) и стандартных наборах (CIFAR-10) доказано, что предложенная архитектура сохраняет высокую точность на чётких данных и обеспечивает значительный прирост точности (более 20%) на зашумлённых и неоднозначных изображениях по сравнению с классической CNN.*

Ключевые слова: *классификация изображений, нечеткая логика, нечеткие множества, компьютерное зрение, нейро-нечеткие системы, гибридные интеллектуальные системы.*

Введение

Современные задачи компьютерного зрения, такие как классификация изображений, лежат в основе множества критически важных приложений – от автономных транспортных средств и медицинской диагностики до систем видеонаблюдения и дистанционного мониторинга Земли. Несмотря на впечатляющие успехи, достигнутые с помощью методов глубокого обучения и традиционных алгоритмов машинного обучения, эти подходы часто сталкиваются с фундаментальной проблемой – неспособностью адекватно обрабатывать присущую визуальным данным неопределенность, нечёткость и субъективность интерпретации. Визуальные данные по своей природе часто амбивалентны: объекты могут быть частично скрыты, освещение – изменчивым, а границы между классами – размытыми (например, стадии заболевания, переходные погодные условия, смешанные ландшафты).

© Великодный В. А., Сирота Е. А., 2026

Традиционные "чёткие" классификаторы, оперирующие бинарной логикой, зачастую оказываются излишне жёсткими в таких условиях, что может приводить к потере информации на этапе принятия решений и снижению общей робастности системы. Эта проблема актуализирует поиск методов, способных работать с неточностью не как с шумом, а как с содержательной характеристикой данных. В частности, методы нечёткой кластеризации демонстрируют эффективность для анализа сложных цветовых распределений и сегментации изображений с нечёткими границами [1].

В этой связи особый интерес представляет аппарат нечёткой логики, созданный именно для моделирования рассуждений в условиях неточности и приближённости. В отличие от бинарной логики, нечёткая логика оперирует степенями принадлежности элемента к множеству (от 0 до 1), что позволяет естественным образом формализовать лингвистические понятия ("тёмное", "шероховатое", "похожее на опухоль") и экспертные знания. Интеграция нечёткой логики в конвейер классификации изображений открывает путь к созданию более гибких, интерпретируемых и устойчивых к помехам систем. Перспективность данного направления подтверждается исследованиями, в которых нечёткие системы используются для построения классификаторов, имитирующих процесс принятия решений человеком-экспертом, что особенно востребовано в сложных предметных областях, таких как медицина или дистанционное зондирование [2]. Более того, гибридные подходы, сочетающие глубокие нейронные сети для извлечения признаков и нечёткие системы для интерпретируемого принятия решений, показывают потенциал для повышения как точности, так и прозрачности моделей компьютерного зрения [3].

Таким образом, целью данного исследования является реализация и экспериментальная оценка гибридной архитектуры нейронной сети для классификации изображений, в которой традиционный полносвязный классификатор заменяется адаптивной нейронечёткой системой вывода (ANFIS) с целью повышения робастности и интерпретируемости модели при работе с зашумлёнными данными и в условиях неопределённости.

1. Разработка модели

Современные системы компьютерного зрения достигли значительных успехов благодаря парадигме глубокого обучения, в особенности – свёрточным нейронным сетям (Convolutional Neural Networks, CNN). Их стандартная архитектура разделена на два ключевых модуля: глубокий свёрточный энкодер для иерархического извлечения абстрактных признаков и полносвязный (FC) классификатор с функцией активации softmax для финального отнесения к одному из классов.

Однако эта, ставшая классической, архитектура обладает фундаментальными ограничениями, критичными для работы в условиях неопределённости:

1. Дихотомичность решений: Функция softmax, нормирующая выходы в распределение вероятностей, где сумма равна 1, вынуждает модель всегда выбирать доминирующий класс, даже если входное изображение является зашумлённым, атипичным или представляет собой смесь классов. Это приводит к «чрезмерно уверенным» и часто ошибочным предсказаниям на граничных случаях.

2. Непрозрачность («чёрный ящик»): Процесс преобразования высокоуровневого вектора признаков в метку класса внутри полносвязных слоев неинтерпретируем. Невозможно понять, какая комбинация извлечённых признаков и в какой мере повлияла на итоговое решение, что недопустимо в критических областях, таких как медицина или безопасность.

3. Неспособность явно моделировать неопределённость: Традиционная CNN не отделяет понятие «уверенности в классе» от общей «неопределённости модели». Например, низкий максимальный выход softmax может означать как близость к границе решения, так и попадание в область, не представленную в обучающих данных (out-of-distribution sample). Система не может явно сказать: «Это похоже на класс A или B, но я не уверен» или «Это не похоже ни на один из известных мне классов».

Эти ограничения актуализируют поиск гибридных подходов, способных сочетать мощность глубокого обучения с формализмом, приспособленным для работы с неточностью и обеспечивающим элементы интерпретируемости.

В основе нашего подхода лежит принцип замены стандартного полносвязного классификатора CNN на адаптивную нейронечёткую систему вывода (Adaptive Neuro-Fuzzy Inference System – ANFIS), интегрированную в единый конвейер сквозного обучения. Данная архитектура получила название нейронечёткой свёрточной сети (Fuzzy-CNN).

2. Детальное описание архитектуры и алгоритма работы

1. Глубокое извлечение признаков (CNN-энкодер):

– Исходное изображение проходит через последовательность свёрточных, пулинговых и нормализационных слоёв.

– Задача этого модуля - преобразовать пиксельное представление в компактный вектор $X = [x_1, x_2, ..., x_n]$ семантически значимых

высокоуровневых признаков. Этот этап отвечает за «восприятие» и абстракцию.

2. Фаззификация входных признаков:

– Каждый i -й чёткий признак x_i из вектора X поступает на вход нечёткого классификатора.

– Для каждого признака предопределяется набор лингвистических переменных (термов), например: «Низкий», «Средний», «Высокий». Каждому терму T_{ij} сопоставлена своя функция принадлежности (ФП) $\mu_{ij}(x_i)$, обычно гауссовского вида:

$$\mu_{ij}(x_i) = \exp(-((x_i - c_{ij})^2 / (2\sigma_{ij}^2))) \quad (1)$$

где c_{ij} – центр, а σ_{ij} – ширина ФП.

– Слой фаззификации преобразует каждый чёткий вход x_i в вектор степеней принадлежности ко всем его термам: $[\mu_{i1}(x_i), \mu_{i2}(x_i), ..., \mu_{im}(x_i)]$. Таким образом, чёткий вектор признаков X отображается в матрицу степеней принадлежности, переводя задачу в пространство лингвистических понятий.

3. Формирование и агрегация нечётких правил:

– На основе всех возможных комбинаций лингвистических термов для разных признаков автоматически генерируется набор нечётких продукционных правил R_k.

– Пример правила: ЕСЛИ (Признак$_1$ = Высокий) И (Признак$_2$ = Низкий) ... ТО (Класс $= C_p$ с весом w_{kp}).

– Слой нечёткого вывода: для каждого правила R_k вычисляется степень его активации (firing strength) F_k как произведение (или минимум) степеней принадлежности соответствующих термов-антецедентов. Далее проводится **агрегация выводов**. Для каждого выходного класса C_p аккумулируется взвешенный вклад всех правил, которые указывают на этот класс в своей заключительной части:

$$S_p = \sum_k (F_k * w_{kp}) \quad (2)$$

где w_{kp} — настраиваемый вес связи правила R_k с классом C_p.

4. Дефаззификация и принятие решения:

– Полученные агрегированные значения S_p для каждого класса уже являются чёткими числами, но они не нормированы.

– В нашем подходе в качестве финального выхода используется вектор степеней принадлежности объекта к каждому классу

$Y = [y_1, y_2, ..., y_p]$, где $y_p = S_p$ (или может быть нормализована относительно суммы для удобства интерпретации, но не как в softmax).

– Ключевое отличие: Сумма компонент вектора Y не обязана быть равна 1. Если все y_p близки к нулю, это является прямым индикатором высокой неопределённости модели - объект не соответствует ни одному классу.

– Финальное решение может приниматься по принципу максимума:

$$Class = \arg\max(Y) \quad (3)$$

3. Уникальность подхода

Наша архитектура представляет собой качественный синтез двух парадигм, преодолевая ограничения каждой из них в отдельности.

Уникальные особенности предлагаемого нейронечёткого классификатора:

– Сквозное обучение нечёткой системы на глубоких признаках: в отличие от многих работ, где нечёткая система применяется к сырым данным или признакам низкого порядка [1], наш классификатор обучается на признаках, извлечённых глубоким CNN. Это означает, что лингвистические термы («Высокий», «Средний») описывают не яркость пикселя, а семантические абстракции (например, «степень выраженности текстуры» или «активность определённого фильтра»), что значительно повышает осмысленность генерируемых правил.

– Единый градиентный оптимизационный контур: все параметры системы – и веса свёрточного энкодера, и параметры функций принадлежности (центры c_{ij}, ширины σ_{ij}), и веса правил w_{kp} – обучаются совместно методом обратного распространения ошибки (backpropagation). Это обеспечивает глубокую адаптацию нечёткого классификатора к специфике решаемой задачи и данных, что отражено в исследованиях по созданию гибридных нейронечётких моделей для анализа изображений [2].

– Явное моделирование уверенности и неопределённости: Выход модели – это не вероятности, а степени принадлежности. Низкое значение максимальной степени принадлежности является прямым и интерпретируемым сигналом о неуверенности модели, что позволяет реализовать механизмы отказа от классификации или передачи контроля эксперту.

– Потенциал для интерпретируемости (Explainable AI, XAI): Активированные для конкретного изображения правила могут быть проанализированы. Можно определить, какие высокоуровневые

признаки (в их лингвистической форме) оказали решающее влияние на отнесение к классу. Это сближает модель с процессом принятия решений экспертом-человеком, рассуждающим в категориях «если объект имеет сильно выраженную характеристику A и слабо выраженную характеристику B, то он, вероятно, относится к классу C» [3-5].

4. Численный эксперимент

Для валидации предлагаемого подхода и оценки его практической эффективности была проведена серия численных экспериментов на разнородных наборах данных. Основная цель – сравнить производительность классической свёрточной нейронной сети (CNN) с полносвязным классификатором и гибридной архитектуры, в которой полносвязные слои заменены на нейронечёткий классификатор (Fuzzy-CNN). Ключевым метрикой сравнения выбрана точность классификации (Accuracy) на тестовой выборке.

Условия эксперимента:

– Для обеспечения чистоты сравнения архитектуры CNN-энкодеров (количество свёрточных слоёв, их параметры) в классической CNN и в Fuzzy-CNN были идентичными.

– Изменялась только часть классификатора: FC-слои vs. ANFIS-слой.

– Обучение проводилось с использованием оптимизатора Adam, функция потерь – категориальная кросс-энтропия (для Fuzzy-CNN с предварительной нормализацией выходов для совместимости).

– Каждый эксперимент повторялся несколько раз для учёта стохастичности, в таблице представлены усреднённые результаты.

Результаты экспериментов сведены в Таблицу ниже.

Таблица

Сравнение точности (Accuracy) классической CNN и Fuzzy-CNN на различных наборах данных.

N	Датасет	Задача	Точность обычной CNN	Точность CNN + Fuzzy Layer	Разница
1	Синтезированные снимки новообразований	Определение наличия новообразования и ее вида	1.0000	1.0000	0%
2	Рентген легких	Определение наличии пневмонии	0.4054	0.6250	22%
3	Cifar-10	Классификация животных и транспорта	0.9433	0.9439	0.0006%

Заключение

Проведённое исследование было посвящено разработке и экспериментальной валидации нового подхода к классификации изображений, основанного на глубокой интеграции аппарата нечёткой логики в архитектуру свёрточных нейронных сетей. В рамках работы была предложена гибридная модель – нейронечёткая свёрточная сеть (Fuzzy-CNN), в которой стандартный полносвязный классификатор заменён адаптивной нейронечёткой системой вывода (ANFIS), обучаемой совместно со свёрточным энкодером методом обратного распространения ошибки.

Основные теоретические и практические результаты работы заключаются в следующем:

1. Разработана и описана новая архитектура, принципиальным отличием которой является сквозное обучение нечёткого классификатора на высокоуровневых семантических признаках, извлечённых глубоким CNN. Это позволяет системе оперировать не абстрактными числовыми векторами, а лингвистическими понятиями, формируя решения на основе интерпретируемых правил вида «ЕСЛИ-ТО». Данный подход напрямую адресует ключевые недостатки классических CNN: непрозрачность («чёрный ящик») и неспособность явно моделировать неопределённость.

2. Экспериментально доказана эффективность и практическая значимость предложенного подхода. Серия численных экспериментов на разнородных наборах данных показала, что нейронечёткая архитектура не уступает по точности классической CNN на стандартных задачах (CIFAR-10) и идеальных данных, при этом демонстрируя качественный прорыв в условиях высокой неопределённости. На реальных медицинских данных (рентгенограммы лёгких) прирост точности составил более 21%, что подтверждает гипотезу о повышенной робастности Fuzzy-CNN к зашумлённости, артефактам и размытым межклассовым границам.

3. Определены ключевые преимущества нейронечёткого классификатора, вытекающие из его архитектуры:

– Явное моделирование уверенности и неопределённости: Выход модели в виде вектора ненормированных степеней принадлежности позволяет чётко отделить случаи низкой уверенности модели, открывая возможность для реализации механизмов отказа от классификации или передачи решения эксперту.

– Потенциал для интерпретируемости (XAI): Анализ активированных для конкретного изображения нечётких правил даёт возможность понять, какие высокоуровневые признаки (в

лингвистической форме) обусловили принятие решения, приближая логику работы модели к рассуждениям человека-эксперта.

– Гибридная природа: Модель сочетает в себе мощь глубокого обучения в части извлечения сложных признаков и гибкость, адаптивность нечёткой логики в части принятия решений.

Таким образом, работа вносит вклад в развитие методов объяснимого искусственного интеллекта (XAI) и робастного компьютерного зрения. Предложенная гибридная архитектура Fuzzy-CNN представляет собой не просто модификацию, а новый класс моделей, которые за счёт синергии глубокого обучения и нечёткой логики становятся более адекватными для работы с несовершенными данными реального мира. Полученные результаты открывают перспективы для успешного применения подобных систем в критически важных и чувствительных к надёжности областях: медицинской диагностике, дистанционном зондировании, автономных системах и других сферах, где требуется не только высокая точность, но и способность модели осознавать границы своей компетенции и объяснять свои выводы.

Направления будущих исследований могут включать: разработку более эффективных методов инициализации и регуляризации нечёткого классификатора для предотвращения переобучения; исследование автоматического формирования компактных и релевантных наборов правил; а также применение предложенного подхода к задачам семантической сегментации и детекции объектов, где неопределённость часто проявляется на уровне границ регионов.

Список литературы

1. Применение нечёткой кластеризации для детального анализа цветных изображений. – [Электронный ресурс]. – Режим доступа: https://cyberleninka.ru/article/n/primenenie-nechetkoy-klasterizatsii-dlya-detalnogo-analiza-tsvetnyh-izobrazheniy/viewer

2. Бондаренко, Б.А. Нечёткие классификаторы в задачах анализа изображений [Электронный ресурс]: Дис. ... канд. техн. наук. – Курск: ЮЗГУ, 2020. – Режим доступа: https://swsu.ru/upload/iblock/102/7xa644phx6ns20nr9nrqjc62bq9olc6m/Avtoreferat-BondarenkoBA.pdf

3. Использование нечёткой логики в задачах компьютерного зрения. – [Электронный ресурс]. – Труды Международной научной конференции «Системный анализ и моделирование» (СПб., 2021). – Режим доступа: https://scm.etu.ru/assets/files/2021/scm21/papers/336-339.pdf

4. Анисимова, Э. С. НЕЙРО-НЕЧЕТКИЕ СЕТИ. – [Электронный ресурс]. – Экономика и социум. 2015. №3-3 (16). URL: https://cyberleninka.ru/article/n/neyro-nechetkie-seti

5. Анисимова, Э. С. НЕЧЁТКАЯ ЛОГИКА: ПРЕДПОСЫЛКИ ВОЗНИКНОВЕНИЯ И ОСНОВНЫЕ ПРИНЦИПЫ. – [Электронный ресурс]. – Экономика и социум. 2015. №2-5 (15). – Режим доступа: https://cyberleninka.ru/article/n/nechyotkaya-logika-predposylki-vozniknoveniya-i-osnovnye-printsipy

Об одном подходе применения технологий общего назначения в рамках проекта «Экономика данных»

А. А. Зацаринный, email: alex250451@mail.ru
В. И. Меденников, email: dommed@mail.ru

ФИЦ ИУ РАН

Аннотация. *В статье дан анализ национального проекта «Экономика данных» с целью проецирования его на производственные отрасли страны на основе основных принципов цифровой экономики в виде некоторых цифровых стандартов. Показано, что реализация их в виде единой цифровой платформы управления экономикой приведет экономику данных к технологиям общего назначения подобно в свое время железным дорогам, электричеству.*

Ключевые слова: ***Ключевые слова:*** *экономика данных, цифровая экономика, цифровая платформа управления, цифровой стандарт, технологии общего назначения.*

Введение

В 2025 г. цифровизация страны после завершения проекта «Цифровая экономика» была продолжена в рамках национального проекта «Экономика данных» со сроком завершения в 2030г. Замена названий проектов обусловлена возрастающим значением данных как одного из ключевых ресурсов интенсивной цифровизации страны [1]. Это особенно важно для искусственного интеллекта, развитие которого требует интеграции огромного количества достаточно структурированной и надежной информации. Такой подход коррелируется с международным опытом в сфере цифровой трансформации, включая развитие экономики данных в направлениях формирования удобного рынка данных, формирования механизмов регулирования оборота данных, создания комфортных условий для интенсивного вовлечения их в социально-экономическую деятельность с наибольшей эффективностью использования, создания индустрии сбора, хранения, обмена данными и еще ряда целевых показателей [2].

Однако анализ состава подпроектов нацпроекта «Экономика данных», перечня мероприятий для его выполнения, а также целевых показателей показывает, что основные акценты в нацпроекте (также как и в предыдущем проекте «Цифровая экономика») сделаны не на цифровизации реального производства, а в основном на цифровизации

© Зацаринный А. А., Меденников В. И., 2026

государственных услуг, статистики, социальной сферы; особое внимание уделено инфраструктуре доступа к информационно-телекоммуникационной сети интернет путем создания низкоорбитальной спутниковой группировки.

Вместе с тем, при этом не учитываются системные подходы к управлению информацией на основе интеграционных механизмов сбора разрозненных данных и алгоритмов в единую структурированную облачную среду, которые предложены в [3-5] в виде комплементарного изменения технологий и организации управления экономическими агентами. Отметим, что подобные интеграционные механизмы частично реализованы в рамках программ цифровой экономики в некоторых странах и привели к появлению ряда широко используемых терминов (цифровая платформа (ЦП), цифровой двойник (ЦД), экосистема (ЭС) и других).

Более того, интеграционные подходы и системные решения по автоматизированному управлению экономикой страны были предложены несколько десятилетий назад академиком В.М. Глушковым в рамках известного проекта общегосударственной автоматизированной системы учёта и обработки информации (ОГАС), но не были одобрены к реализации руководством СССР [6, 7].

Очевидно, что целевая направленность нацпроекта «Экономика данных» должна учитывать как известные фундаментальные научно-технические наработки, так и реальные особенности современного развития экономики России, требующие комплексного учета огромных объемов накопленных данных.

В связи с изложенным, в статье рассматриваются некоторые аспекты технологий общего назначения применительно к направлениям, реализуемым в нацпроекте «Экономика данных».

1. Понятие технологий общего назначения

Массовое внедрение в экономику научно-технических разработок в конце прошлого века привело к понятию технологий общего назначения (ТОН), идею которой выдвинул П. Дэвид [8]. Развитие ТОН как базовых привело к появлению на их основе целого ряда прикладных технологий. В [8] приведены основные отличительные черты ТОН:

- наличие значительного технологического потенциала для развития и внедрения на новые отрасли человеческой деятельности;
- большое разнообразие технологий, видов бизнеса, в которых они могут быть использованы;
- проявление высокого уровня комплементарности с возникшими на их основе новыми технологиями в производстве;

- пропорциональное изменение комплементарных активов в виде материально-технических ресурсов, систем и структур управления, человеческого капитала.

Вместе с тем, представляется необходимым выделить еще одну отличительную черту ТОН – экономическую, которая и явилась главным мотиватором их появления. Согласно определению технологического разрыва возникшая новая технология должна существенно превышать результативность действующего аналога [9]. Так, в [3] рассматривается цифровой инструмент, эффективно реализующий основные принципы ЦЭ, в виде единой цифровой платформы управления (ЦПУ) экономикой, включающей облачные подплатформы сбора и хранения первичной учетной информации, технологических БД и баз знаний.

В статье показано, каким образом ЦПУ в свете проводимых исследований можно рассматривать как ТОН в рамках процессов цифровизации страны.

2. Основные признаки технологий общего назначения у ЦПУ

Рассмотрим нарастающие тенденции проявления признаков ТОН у ЦПУ на современном этапе развития цифровой экономики в свете проекта «Экономика данных».

Во-первых, наличие значительного технологического потенциала ТОН для развития и внедрения ЦПУ на ряде отраслей экономики подтверждается универсальностью блока сбора и хранения первичной учетной информации для ряда производственных отраслей на примере АПК, логистики, фармацевтики, строительства с последующей интеграцией отраслевых подплатформ в единую цифровую платформу страны; это наглядно демонстрируется в ряде работ, например [10]. При этом важно, что при переходе к ЦПУ может быть достигнута наибольшая эффективность цифровизации производства, особенно в управлении им, поскольку информация становится доступной для предприятий всех размеров, а не только для отдельных наиболее крупных из них.

Во-вторых, второй признак ТОН у ЦПУ (большое разнообразие технологий и видов бизнеса) характеризуется следующим:

– Единая ЦПУ производством при переходе на нее всех отраслей позволяет осуществлять сбор и хранение информации о всех совершенных технологических операциях, включая также данные интернета вещей, по единому цифровому стандарту. Другими словами, единая ЦПУ является технологической основой реализации идеи прослеживаемости продукции, под которой подразумевается цифровая технология оперативного информирования ответственных ведомств, конечных пользователей, поставщиков всех комплектующих, а также всех участников процесса о характеристиках изготовителя и параметрах

выпускаемых товаров [11, 12]. Сейчас такие технологии активно применяются в сфере услуг (прослеживаемость товаров широкого потребления, продуктов питания, лекарств и т.п.).

– Единая ЦПУ производством явилась основой формирования системы управления логистической деятельностью как связующее звено всех участников цепочки создания добавленной стоимости, что особенно актуально в условиях введенных против России санкций. В этом случае использование технологий ЦПУ создает возможность перехода к прямым продажам, при которой производитель прослеживает конечного потребителя, объем и структуру его спроса и исключает из цепочки множество посредников, на которых сейчас приходится до 80% стоимости в розничной цене товара [12].

– Единая ЦПУ (как показано в [3]) позволяет создать условия для эффективного использования искусственного интеллекта (ИИ), включая возможность построения необходимой информационной архитектуры на основе онтологического моделирования [13], позволяющего представить знания предметной области в структурированном для машинной обработки виде. ЦПУ определяет рациональную схему интеграции как информации, так и алгоритмов принятия решения, в том числе и ИИ. Такая интеграция на основе ЦПУ будет способствовать повышению эффективности производства продукции, отвечающей необходимым ценовым, качественным и экологическим требованиям.

– Существенное снижение рыночной стоимости цифровых технологий наряду с появлением более совершенных и дешевых электронно-оптических мониторинговых средств, дающих технологическую возможность установки их на различных подвижных и стационарных машинах, аппаратах, привело к ситуации почти полного контроля каждой технологической операции производственных процессов и связанной с ними средой с точным анализом и прогнозом последствий всех действий. Данное явление получило название точного (прецизионного) производства. В сельском хозяйстве наиболее известной такой технологией стало точное земледелие (ТЗ), основанное на мониторинге, учете и возможности обратного воздействия на максимально возможное число сельскохозяйственных процессов с комплексом взаимосвязанных с ними материалов, техники, оборудования, живых организмов [14]. При этом ТЗ, как и любое другое прецизионное производство, подразумевает необходимость учета значительного количества разнородных данных на полях с учетом антропогенной, природной пространственно-временной неоднородности каждого из них, которые необходимо количественно учитывать с использованием интеграционных технологий на основе онтологического

моделирования предметных областей не только в растениеводстве, но и в ряде смежных отраслей. Однако, ведомственная разобщенность, сопровождаемая фрагментарным, позадачным подходом к проектированию еще информационных систем (ИС), перенесенным и на системы цифровой экономики, приводит к малоэффективной реализации интеграционных проектов [15].

– В настоящее время в России создается система распределенных ситуационных центров, которая направлена на обеспечение информационно-аналитической и интеллектуальной обработки информационных потоков из различных источников для постоянного повышения качества управленческих решений. В качестве методологического и инструментального подхода, обеспечивающего интеграцию различных подсистем и цифровых технологий, могла бы обеспечить единая ЦПУ с использованием единого регламента взаимодействия [16].

– В работе [15] показано, что на базе ЦПУ успешно могут быть решены онтологические проблемы интеграции данных дистанционного зондирования земли (ДЗЗ) в следующих отраслях: в сельском хозяйстве, в картографии, в метеорологии и климатологии, в гидрологии, в лесном хозяйстве, в охране окружающей среды, при чрезвычайных ситуациях.

– Наконец, в работе [17] рассматривается концепция трансформации продуктовой модели общественного развития в сервисную с единых позиций на базе разработанной ЦПУ экономикой страны. Концепция предусматривает разработку цифрового инструмента формирования в организациях, в отдельных случаях у физических лиц, экосистемной модели потребления, которая отражает возросший социальный заказ на более эффективное использование убывающих природных ресурсов, а также на снижение экологических нагрузок на окружающую среду.

В-третьих, к третьему и четвертому признаку ТОН у ЦПУ (высокий уровень комплементарности и пропорциональное изменение комплементарных активов в виде материально-технических ресурсов) отнесем следующие.

Основные принципы успешности реализации цифровой экономики (ЦЭ) как драйвер разработки ЦПУ получили подтверждение в работах [18-19], в которых показано, что экономическая эффективность цифровой трансформации бизнеса обусловлена инвестициями не только в ИКТ, но также в такие комплементарные активы, как усовершенствование управления, в человеческий ресурс.

Для внедрения новых цифровых технологий в экономику требуются изменения в системы управления на всех уровнях страны, включая формирование единого понятийного поля и семантического пространства

путем разработки стандартов на соответствующие системы управления [20]. Более тщательные исследования компании Economist Intelligence Unit с целью проверки теории комплементарности подтвердили очень важные для стратегии цифровизации России выводы в части форм и степени влияния ИКТ, соответственно, ЦЭ на экономические показатели [18]. Таким образом, применение новой ТОН требует изменения бизнес-моделей, стратегий и организационных структур подобно появлению электричества, породившего во многих отраслях бизнес-модели массового производства. Далее, в работе [3] показано также, что в пространстве проектирования цифровых стандартов ЦПУ данные, алгоритмы решения функциональных управленческих задач и инструментарий автоматизации в виде программного обеспечения, находятся в комплементарных связях с сильным синергетическим эффектом. Так, алгоритмы приводят к росту эффективности использования информации, так же, как и структурированные, все более интегрированные данные совершенствуют методы их обработки. Особое внимание при реализации проекта «Экономика данных» должно быть уделено очень актуальному для России комплементарному активу – научно-образовательному, в виде ЦП информационных научно-образовательных ресурсов (ЦП ИНОР) с интеграцией ее с ЦПУ производством [21].

В-четвертых, рассмотрим экономическую составляющую, мотивирующую разработку ЦПУ. Как уже рассмотрено выше, цифровизация потребовала сбора и использования огромного массива данных, а также совершенствования алгоритмов их обработки, значительного увеличения функциональных управленческих задач. Интеграция этих комплементарных активов для конкретного объекта (предприятия), исходя из некоторых целевых показателей на основе выбранных критериев их формирования, породила концепцию цифровых двойников (ЦД), стоимость разработки которых имеет экспоненциальную зависимость от объемов информации и задач, критериев формирования ЦД, показателей и прочих ограничений. Например, в [22] финансовые затраты формирования ЦД, имеющего десятки тысяч объектов, оцениваются в сумме 100$ миллиардов. Из первых опытных образцов таких больших ЦД следует важный практический вывод о невозможности разработки индивидуальных ЦД для каждого предприятия в большинстве отраслей, поэтому выходом является разработка типового единого ЦД для большого их числа. Однако для этого должны быть созданы единые понятийное, информационное и алгоритмическое пространства на основе онтологического моделирования предметных областей не только в одной, но и в ряде

смежных отраслей. Данным критериям как раз удовлетворяет единая ЦПУ экономикой, которая может рассматриваться в настоящее время как прообраз ЦД предприятий, поскольку, например для сельского хозяйства, перечень задач ЦПУ является полным и типовым для большинства хозяйств с моделями, агрегирующими все знания, применяемые при проектировании севооборотов, разработке технологических карт, годовых планов и анализе их выполнения [23].

Заключение

В работе представлены предложения по созданию цифрового инструмента в виде ЦПУ экономикой страны, который позволит повысить эффективность внедрения экосистемного точного земледелия, прослеживаемости продукции, искусственного интеллекта, интеграции с технологиями ДЗЗ, интеграции с единой ЦП логистики, методологического и инструментального подхода к развитию ситуационных центров, формирования экосистемной модели потребления общества, усовершенствования управления на всех уровнях со значительной экономической эффективностью цифровой трансформации бизнеса и т.д.

Реализация предлагаемого подхода требует проведения исследовательских работ по онтологическому моделированию экономики страны для устранения существующих межотраслевых противоречий в понимании процесса цифровизации отраслей.

Список литературы

1. Основные показатели и мероприятия национального проекта «Экономика данных и цифровая трансформация государства» [Электронный ресурс] : – Режим доступа : http://government.ru/info/54314/

2. Экономика данных: мировые подходы к управлению [Электронный ресурс] : – Режим доступа : https://issek.hse.ru/news/865612618.html

3. Зацаринный, А. А. Интеграция приложений искусственного интеллекта в единую цифровую платформу АПК / А. А. Зацаринный, В. И. Меденников, А. Н. Райков // Информационное общество. – 2023. – № 1. – С. 127-138.

4. Алексеева, Н. А. Экономические и управленческие проблемы землеустройства и землепользования в регионе / Н. А. Алексеева, А. К. Осипов, В. И. Меденников и др. // Ижевск: Шелест. – 2022. – 225 с.

5. Меденников, В. И. Математическая модель формирования цифровых платформ управления экономикой страны / В. И. Меденников // Цифровая экономика. – 2019. – №1(5). – С. 25-35.

6. Глушков, В. М. Макроэкономические модели и принципы построения ОГАС / В. М. Глушков. М. : Статистика, 1975. – 160 с.

7. Меденников, В. И. Актуальность проекта ОГАС в цифровую эпоху для эффективности управления экономикой / В. И. Меденников // Финансовый менеджмент. – 2023. – № 4-2. – С. 250-257.

8. David, P. The Dynamo and the Computer: An Historical Perspective on the Modern Productivity Paradox / P. David // The American Economic Review. – 1990. – Vol. 80. – No. 2, Papers and Proceedings of the Hundred and Second Annual Meeting of the American Economic Association. – P.355-361.

9. Фостер, Р. Обновление производства: атакующие выигрывают: Пер. с англ. / Р. Фостер; Общ. ред. и вступ. статья В. И. Данилова-Данильяна. – М. : Прогресс, 1987. – 270 с.

10. Меденников, В. И. Цифровые стандарты - основа интеграции цифровых платформ АПК и других отраслей / В. И. Меденников, Ю. И. Микулец // Вестник Московского гуманитарно-экономического института. – 2021. – № 1. – С. 208-226.

11. Gontar, N. N., Galina A. Y. Traceability as an element of the system of state control over the goods turnover / N. N. Gontar, A. Y. Galina // European Scientific e-Journal. – 2022. – 4(19). – 19-28.

12. Микулец, Ю. И. Единый подход к прослеживаемости продукции на основе математической модели цифровой платформы управления логистикой / Ю. И. Микулец, В. И. Меденников // Вестник Московского гуманитарно-экономического института. – 2023. – № 1. – С. 69-83.

13. ИИ более продуктивен, когда он основан на онтологии [Электронный ресурс] : – Режим доступа : https://ontograph.ru/ontologiya-klyuch-k-raskrytiyu-vozmozhnostey-ii/

14. Walter, A. Smart farming is key to developing sustainable agriculture / A. Walter, R. Huber, N. Buchmann // Proc Nat Acad Sci USA. – 2017. – 114(24):6148–6150.

15. Medennikov, V. I. The role of agrarian science in transforming methods of using the Earth remote sensing data into publicly available technology / Medennikov V. I., Y. A. Flerov // CEUR Workshop Proceedings, Novosibirsk, August 24–27, 2021. Novosibirsk. 2021. 180-193.

16. Ilyin, N. Distributed situational centres system of cutting edge development / N. Ilyin, G. Malinetsky, K. Kolin, A. Zatsarinny, et al. // Proceedings of 10th International Conference on Management of Large-Scale System Development (MLSD). 2017.

17. Зацаринный, А. А. О подходах к внедрению сервисной модели общественных отношений на основе цифровых технологий / А. А.

Зацаринный, В. И. Меденников // Информатика: проблемы, методы, технологии : Материалы XXV Международной научно-практической конференции им. Э.К. Алгазинова, Воронеж, 12–13 февраля 2025 года. – Воронеж: Воронежский государственный университет. – 2025. – С. 756-763.

18. Brynjolfsson, E. Intangible Assets: Computers and Organizational Capital / E. Brynjolfsson, L. Hitt, S. Yang // Brookings Papers on Economic Activity. – 2002. – Vol.2. – No.1.

19. Milgrom, P. The Economics of Modern Manufacturing: Technology, Strategy and Organization / P. Milgrom, J. Roberts // American Economic Review. – 1990, – Vol. 80. – № 3. – P. 511–528.

20. Агеев, А. И. Насколько Россия подготовлена к вызовам XXI века / А. И. Агеев // НГ-ЭНЕРГИЯ от 16.01.2019.

21. Зацаринный, А. А. Теория и практика цифровой трансформации общества в рамках приоритетов научно-технологического развития России / А.А. Зацаринный, К.К. Колин // В сборнике «Проектирование будущего. Проблемы цифровой реальности». – 2019. – № 1 (2). – С. 29-39.

22. Боровков, А. И. Цифровые двойники и цифровая трансформация предприятий ОПК / А. И. Боровков, Ю. А. Рябов, К. В. Кукушкин, В. М. Марусева, В. Ю. Кулемин // Оборонная техника. – 2018. – № 1. – С. 6–23.

23. Меденников, В. И. Экономическая целесообразность разработки единого цифрового двойника предприятия на примере сельского хозяйства / В. И. Меденников // Тенденции развития интернет и цифровой экономики : Труды VIII Международной научно-практической конференции, Симферополь-Алушта, 29–31 мая 2025 года. – Симферополь: ИП Зуева Т.В. – 2025. – С. 31-34.

Результаты вычислительного эксперимента по оценке эффективности применения модели управлении зонами покрытия в спутниковой IoT-инфраструктуре

Д. А. Иванов, email: 1142230113@rudn.ru [1]
В. В. Алексеев, email: vvalex1961@yandex.ru [2]

[1]Российский университет дружбы народов им. Патриса Лумумбы,
[2]Тамбовский государственный технический университет

Аннотация. *В условиях ограниченности вычислительных ресурсов для оперативного перерасчета и корректировки зон покрытия гибридной спутниковой инфраструктуры Интернета вещей в реальном времени предложена модель управления зонами покрытия, основанная на интеграции подсистемы поддержки принятия решения с модулями кластеризации телеметрических терминалов и прогнозирования нагрузки, обеспечивающая динамическую реконфигурацию зон покрытия. Апробация модели выполнена методом имитационного моделирования для сценариев геостационарных, низкоорбитальных и гибридных спутниковых систем, что демонстрирует уменьшение задержек передачи, снижение коэффициента перегрузки и увеличение площади устойчивого покрытия при высокой точности прогнозирования нагрузки. Полученные результаты подтверждают эффективность применения модели поддержки принятия решения по управлению зонами покрытия в спутниковых гибридных системах интернета вещей, позволяя интегрировать ее в систему управления для обеспечения непрерывности и оперативности принятия решений в условиях изменяющейся нагрузки и инфраструктурных трансформаций.*

Ключевые слова: *Интернет вещей, модель управления, подсистема поддержки принятия решения, зоны покрытия..*

Введение

Технология Интернета вещей (IoT) обеспечивает обмен телеметрией между распределенным сенсорными терминалами. Для сценариев вне зон устойчивого наземного покрытия ключевым элементом становится спутниковая связь, обеспечивающая независимость от географической инфраструктуры [1]. Спутниковые системы являются фактически единственным технологическим базисом, обеспечивающим работу технологии IoT на транспорте, добыче ресурсов, мониторинге экологии и

© Иванов Д. А., Алексеев В. В., 2026

критической инфраструктуре, где задержки передачи должны быть минимальны, а пропускная способность высокой.

В настоящее время спутниковая инфраструктура IoT-решений реализуется на основе трёх архитектур: геостационарная (GEO), низкоорбитальная (LEO) и гибридная, объединяющая первую и вторую.

GEO-системы обеспечивают широкую площадь стабильного покрытия, но характеризуются высокой задержкой передачи информации, порядка 500-600 м/с, и запуск космических аппаратов на геостационарную орбиту требует значительных финансовых вложений. Задержка сигнала до 600 м/с может быть критичной для приложений Интернета вещей, требующих минимальных задержек.

LEO-системы имеют низкую задержку передачи от 20 до 50 м/с, обеспечивают более быструю двустороннюю передачу данных, что критично для таких задач, как управление беспилотными транспортными средствами, но нуждается в сложных механизмах динамического распределения нагрузки на сеть в условиях ограниченности ресурсов [2].

Гибридные конфигурации предполагают совместную эксплуатацию обоих сегментов с их преимуществами для повышения качественных характеристик исследуемой системы.

С позиции системного анализа важной задачей, которая возлагается на гибридную спутниковую IoT-инфраструктуру, включающую подсистему маршрутизации и распределённые телеметрические терминалы, является обеспечение многокритериального управления сетевыми ресурсами в условиях неопределённости, требующее интеграции алгоритмов кластеризации и прогнозирования нагрузки на сеть [3]. При этом существующие схемы управления в спутниковых сетях преимущественно статичны и не учитывают пространственно-временную изменчивость распределения терминалов, что проявляется в виде кластерных нагрузок, характерных для задач мониторинга, логистики и управления объектами [4, 5].

Гибридная спутниковая IoT-инфраструктура, при ограниченных вычислительных ресурсах требует оперативного перерасчета и корректировки зон покрытия. Для этого предложена модель управления зонами покрытия на основе интеграции подсистемы поддержки принятия решения, включающей модули кластеризации телеметрических терминалов и модулей прогнозирования нагрузки. Она обеспечит динамическую реконфигурацию зон покрытия.

1. Архитектурные особенности модели

Модель управления зонами покрытия основана на интеграции подсистемы поддержки принятия решения с модулем кластеризации

телеметрических терминалов и модулем прогнозирования нагрузки сети, что обеспечивает динамическую реконфигурацию зон покрытия.

Для анализа оценки эффективности модели необходимо обратиться к ее расположению в архитектуре, которая включает пять уровней:

1 – организационный,

2 – поддержка принятия решения (ППР),

3 – наземный,

4 – орбитальный (GEO/LEO-инфраструктура),

5 – терминальный.

Организационный уровень задает приоритеты, определяемые требуемыми задачами заказчика, поставившего в контур оборудования IoT. ППР осуществляет сбор, анализ и прогнозирование телеметрии, включая модель модуль кластеризации телеметрических терминалов и модуль прогнозирования нагрузки. Наземные станции выполняют маршрутизацию трафика. Орбитальные сегменты обеспечивают физическую ретрансляцию сигналов. Терминальный уровень формирует распределённые точки спроса, генерируя пространственно-временную нагрузку на сеть (суда, станции мониторинга, нефтегазовые платформы и ледовый станции) [6].

Перечисленные подсистемы и взаимоотношения между ними представлены в обобщенной структурной схеме гибридной спутниковой IoT-инфраструктуры на рис. 1. Стрелками изображены более важные и частые зависимости между подсистемами.

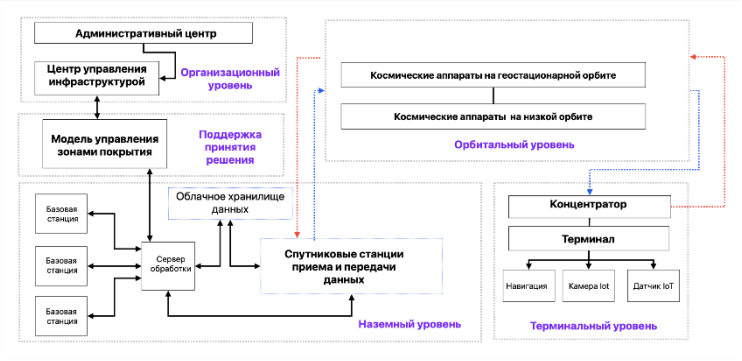


Рис. 1. Структурная схема гибридной спутниковой IoT-инфраструктуры

Информационные потоки в системе проходят полный цикл: телеметрия - ППР - прогноз - решения - реконфигурация - контроль качественных показателей. Управление в системе спутникового

интернета вещей зависит от основных функциональных подсистем и определенных взаимосвязей между ними. Управление спутниковым Интернетом вещей организовано через комплексную систему, включающую несколько уровней и подсистем, которые взаимодействуют между собой для обеспечения непрерывности и качества связи [7].

2. Экспериментальная апробация эффективности модели

Имитационное моделирования в спутниковой гибридной IoT-инфраструктуре проведено с учетом распределения потоков информации и корректировки зон покрытия Z(t). У каждой зоны обслуживания покрытия необходимо определить конфигурацию C(t), обеспечивающую непрерывность и оперативность принятия решений в условиях изменяющейся нагрузки и инфраструктурных трансформаций. Пусть конфигурация C(t) задаёт распределение зон покрытия Zi(t) для i=1…N.

Уравнения отдельных устройств СУ имеют вид:

$$F(C) = [f1, f2, f3, f4],$$

- f_1 – средняя задержка пакетов в системе (средняя время ожидания обслуживания);
- f_2 – площадь покрытия (доля территории, охваченной зонами обслуживания);
- f_3 – коэффициент перегрузки (отношение пиковой нагрузки к ёмкости зоны покрытия);
- f_4 – устойчивость системы (измеряемая, например, дисперсией отклоненйи от целевых показателей системы).

Искомая оптимальная конфигурация C*(t) выбирается методом градиентного спуска [8], путем максимизации скалярного критерия свёртки:

$$J(C) = w_1 *(1- f_{1, норм}) + w_2 * f_{2, норм} + w_3 *(1- f_{3, норм}) + w_4 * f_{4, норм},$$

- $f_{1, норм}$ – нормированные значения критериев (обычно по минимаксному приниципу или Z-нормализации на множестве возможных конфигураций),
- $w_{i,}$ – весовые критерии, заданные на множестве $[w_1, w_2, w_3, w_4]$.

Такой выбор отражает приоритет устойчивости и покрытия над минимизацией задержек и перегрузок.

Кластеризация телеметрических терминалов выполняется алгоритмом DBSCAN (Density-Based Spatial Clustering of Application with

Noise) для выявления плотностных кластеров IoT-терминалов на основе их пространственного положения и текущей неравномерной нагрузки [9].

Распределённые IoT-терминалы (суда, станции мониторинга температур, движущиеся платформы) образуют неравномерное поле нагрузки $\rho(x,y)$ [10].

Для дальнейшего распределения ресурсов [11] и устойчивого управления зоной покрытия требуется выявление кластеров с повышенной нагрузкой. Эффективно распределить ресурсы между кластерами в предлагаемой модели управления зонами покрытия позволяет алгоритм DBSCAN, применяемый к координатам IoT-терминалов в географической системе, минимизируя перегрузки внутри кластера.

Прогнозирование нагрузки выполняется на основе технологий долгой краткосрочной памяти LSTM (Long Short-Term Memory) и используется для предотвращения реактивной перенастройки конфигурации C(t) [12].

Алгоритм LSTM на основе рекуррентных нейронных сетей обучается на исторических данных о нагрузке в кластере, позициях IoT-терминалов и внешних негативных факторах и обеспечивает динамичную реконфигурацию на, прогнозируя распределение нагрузки в течение периода времени ΔT (например, 30 минут) [13].

Алгоритм принимает на вход временные ряды объёма трафика и количества активных IoT-терминалов по кластерам, а такжедополнительные контекстные признаки (время суток, погодные условия, навигационная обстановка). Для каждого кластера k формируется модель:

$$T_k(t+\Delta T) = LSTM_k\,(T_k\,(t - L+1), \ldots, T_k(t)),$$

- $T_k(t)$ – измеренный трафик в момент t,
- L – длина окна передачи.

Обучение производится по историческим данным и данным имитационного моделирования. Критерием качества служит среднеквадратичная ошибка прогноза RMSE, а также метрика MAPE.

Экспериментальная модель управления зонами покрытия включает алгоритмы DBSCAN и LSTM в блоке системы поддержки принятия решения, где для каждого допустимого варианта конфигурации C(t) рассчитывается значение J(Cj) [14].

Оптимальная конфигурация C*(t) определена найденными параметрами конфигурации и скорректирована набором экспертным правил для учёта внешних воздействующих негативных факторов [14],

таких как запрет на отключение критических зон, приоритет движения транспортных средств и т.п.

Апробация выполнена методом имитационного моделирования для сценариев работы в спутниковой гибридной IoT-инфраструктуре, демонстрируя значительное снижение задержек передачи информации, уменьшение коэффициента перегрузки и повышение площади устойчивого покрытия при высокой точности прогнозирования нагрузки. Таким образом, система поддержки принятия решения является замкнутым контуром принятия решений, который приведен на рис. 2.



Рис. 2. Контур принятия решения в гибридной спутниковой IoT-инфраструктуре

Регион имитационного моделирования охватывает участок Северного морского пути с распределенными судоходными маршрутами, береговыми станциями и ледовыми беспилотные средствами наблюдения. Плотность IoT-терминалов варьируется от 5 до 20 устройств на 100 км^2 с пиковыми значения в районах портов и промышленных объектов.

Оценка эффективности применения модели при управлении зонами покрытия в спутниковой IoT-инфраструктуре проведена для сравнения на нескольких сценариях:

S1 – GEO-система без интеграции модели ППР;

S2 – LEO-система без интеграции модели ППР;

S3 – Гибридная система с интеграцией модели ППР.

Для сравнения сценариев используются следующие показатели:

τcp – средняя задеркжа передачи информационных пакетов,

τmax – максимальная задеркжа передачи информационных пакетов,

Psucc – доля, успешо доставленных, информационных пакетов,

Kload – коэффициент перегрузок зон покрытия,

Pcov – площадь устойчивого покрытия,

QoS – интегральный показатель устойчивости.

Результаты имитационного моделирования представлены в таблице ниже.

Таблица

Сравнение результатов для сценариев S1, S2, S3

№ п/п	Характеристика	Сценарий S1	Сценарий S2	Сценарий S3
1	Точность прогнозируемой нагрузки, %	–	-	98
2	Средняя задержка, мс	500	120	50
3	Коэффициент перегрузки	0.15	0.25	0.08
4	Покрытие, %	85	95	97

Полученные результаты по сценарию S3 подтверждают высокую эффективность применения модели управления зонами покрытия в гибридных спутниковых IoT-инфраструктурах.

Сценарий S3 демонстрирует снижение задержки более, чем в 10 раз по сравнению с S1 и в 2,4 раза по сравнению с S2.

Низкий коэффициент перегрузок 0,08 обусловлен кластеризацией телеметрических терминалов и распределением ресурсов системы. Точность прогноза нагрузки 98% достигается за счет применения алгоритма LSTM, что соответствует известным результатам прогнозирования трафика [13]. Увеличение зоны покрытия до 97% подтверждает устойчивость работы предложенной модели.

Таким образом, результаты вычислительного эксперимента позволяют интегрировать модель управлении зонами покрытия в спутниковую IoT-инфраструктуру для обеспечения непрерывности и оперативности принятия решения в условиях изменяющейся нагрузки, снижая среднюю задержку, уменьшая перегрузка в зонах покрытия и увеличивая площадь устойчивого покрытия.

Заключение

Системный анализ результатов показывает, что предложенная модель обеспечивает устойчивость работы IoT-инфраструктуры в условиях неопределенности за счет комбинации организационного управления и модели поддержки принятия решения по управлению зонами покрытия в спутниковых гибридных системах Интернета вещей.

Многокритериальная постановка решаемой задачи позволяет адаптировать конкретные критерии под конкретные классы внутри исследуемой системы, оставаясь в рамках единой методологии.

Представлена структурная схема гибридной спутниковой IoT-инфраструктуры (рис. 1). Разработана модель и предложена ее интеграция в качестве системы поддержки принятия решения при управлении зонами покрытия для обеспечения многокритериального управления сетевыми ресурсами в условиях неопределённости.

Отражен интересующий подход по достижению искомых показателей за счет синтеза алгоритмов DBSCAN и LSTM для кластеризации нагрузки и предиктивного управления ресурсами.

С позиции системного анализы экспериментального апробирована модель поддержки принятия решения по управлению зонами покрытия в спутниковых гибридных системах интернета вещей для арктического региона Российской Федерации. Полученные результаты подтверждают эффективность применения модели в условиях изменяющейся нагрузки и инфраструктурных трансформаций с целью повышения устойчивости и качества обслуживания.

Дальнейшее развитие модели может быть использовано при проектировании и модернизации государственных спутников систем связи для нужд Северного морского пути, мониторинга ледовой обстановки и экстренных служб, а также при уточнении нормативно-методических рекомендации по управлению зонами покрытия в спутниковых IoT-сетях.

Список литературы

1. ГОСТ Р 59670-2021. Системы спутниковой связи. Общие требования. – М. : Стандартинформ, 2021. – 24 с.

2. Ань, Н. Д. Система "Марафон IoT" и новый рынок спутникового IoT / Н. Д. Ань, Ершов А. // Технологии и средства связи. – 2021. – С. 75–80.

3. Васильев, К. К. Теория электрической связи : учеб. пособие / К. К. Васильев [и др.] ; под общ. ред. К. К. Васильева. – Ульяновск : УлГТУ, 2008. – 452 с.

4. Рожкова, А. А. Формирование оптимальной зоны покрытия в сотовых системах связи / А. А. Рожкова // Современные наукоёмкие технологии. – 2014. – № 5-2. – С. 54–54.

5. Алексеев, В. В. Планирование, мониторинг и анализ использования техники на базе ГИС и спутниковой навигации / В. В. Алексеев, А. С. Полянский, Д. С. Ленков // Цифровизация АПК : сб. ст. – Тамбов : ТГТУ, 2020. – С. 32–35.

6. Прокофьева, О. Перспективы спутниковых технологий на рынке M2M/IoT / О. Прокофьева // Технологии и средства связи. – 2017. – С. 44–47.

7. Гадисова, А. М. Обзор современных систем управления сетями связи и их применения / А. М. Гадисова // Молодой ученый. – 2022. – № 6 (401). – С. 9-11.

8. Rappaport, T. Wireless Communications: Principles and Practice / T. Rappaport. – 3rd ed. – Prentice Hall, 2023. – 752 p.

9. Haykin, S. Modern Wireless Communications / S. Haykin, M. Moher. – Pearson, 2022. – 845 p.

10. Русанова, Т. В. Спутниковые системы связи в индустрии интернета вещей / Т. В. Русанова, А. Ю. Кочанов, А. П. Денискин, В. В. Золотухин // Инновационные технологии…. – 2020. – С. 222–223.

11. Покаместов, Д. А. Расчёт зон обслуживания кластера… / Д. А. Покаместов // Вестник СибCUTI. – 2017. – № 3. – С. 45–52.

12. Haykin, S. Neural Networks and Learning Machines / S. Haykin. – 3rd ed. – Pearson, 2021. – 937 p.

13. Иванова, Е. Системы принятия решений в IoT / Е. Иванова // Вестник связи. – 2025. – № 2. – С. 34–42.

14. Алексеев, В. В. Анализ и синтез модульных сетевых информационных систем в интересах повышения эффективности целенаправленных процессов / В. В. Алексеев, Ю. Ю. Громов [и др.]. – Тамбов; М.; СПб.; Баку; Вена; Гамбург : Изд-во МИНЦ «Нобелистика», 2012. – 130 с.

Система хранения и организации значимой для управления изменениями в SAP S/4 HANA корпоративной информации

И. В. Илларионов, email: igor.illarionov@gmail.com
Н. О. Гришаев, email: zebulos@yandex.ru

Воронежский государственный университет

Аннотация. *рассматривается задача централизации и структурирования корпоративной информации, используемой при управлении изменениями и инцидентами в системе SAP S/4 HANA. Актуальность работы обусловлена высокой фрагментацией данных между различными информационными системами, что приводит к росту трудозатрат, снижению прозрачности процессов и увеличению числа ошибок. Предложено решение в виде корпоративного веб-портала, обеспечивающего единое хранилище данных, поддержку процессов Change Management и Incident Management, а также прогнозирование нагрузки на службу поддержки.*

Ключевые слова: *SAP S/4 HANA, управление изменениями, инцидент-менеджмент, корпоративный портал, тикет-система, ARIMA.*

Введение

В крупных компаниях работают тысячи, десятки или даже сотни тысяч сотрудников, функционируют сотни различных бизнес-процессов. Для управления этими процессами используется целый комплекс сложных и объемных программных продуктов, для работы с которыми сотрудники проходят специального обучение. При этом постоянно возникают потребности в каких-либо изменениях, а также учете различного рода инцидентов. К сожалению, зачастую такой учет ведется в разных системах, поэтому администраторы вынуждены постоянно переключаться между различными порталами для выполнения учетных операций и администрирования заявок.

В работе предлагается решение проблемы отсутствия структурированной единой базы данных, содержащей всю необходимую информацию по внедряемому в SAP изменению или инциденту. Система позволяет решить проблему фрагментации данных, необходимых для решения одной бизнес-задачи.

© Илларионов И. В., Гришаев Н. О., 2026

Таким образом, не будет необходимости работать в множестве несвязных программных систем, вся необходимая информация по учету и обработке изменений и инцидентов будет находится в единой системе.

1. Обоснование потребности в автоматизации

В рамках исследования предметной области была составлена модель «Как есть» процесса управления изменениями и инцидентами.

Спецификации разработок расположены на «Microsoft SharePoint», база знаний разбросана между статьями в разделе Knowledge Base в «Service Now» и презентациями из папки на сервере, а транспорты (объект из системы «SAP»), относящиеся к какому-либо изменению/проекту, и вовсе не имеют никакой связи с внешним миром, нигде не указываются и при необходимости ищутся в «SAP»-системе.

Документация по Job'ам (задачам) расположена на отдельном портале, как уже было выше описано.

В существующем способе управления изменениями и инцидентами выделены следующие недостатки:

- большая децентрализация источников данных, что приводит к утомительному поиску информации;
- повышенная трудоемкость операций обработки тикетов;
- большое число ошибок обработки информации;
- невозможность эффективного построения отчетов, например, для того, чтобы оптимизировать загрузку службы техподдержки.

Для их минимизации предлагается разработка и внедрение корпоративного портала, агрегирующего всю информацию. В предлагаемом процессе после делегирования заявки основная информация о ней заносится консультантом в корпоративную систему. Далее происходит поиск информации в этой единой базе, и регистрация операций различного типа. В результате вся информация аккумулируется в системе, что ускоряет процесс ее поиска и обработки. Работа ведется в едином интерфейсе портала.

Появляется возможность прогнозирования тикетов по категориям, что позволяет прогнозировать сезонную нагрузку консультантов и бюджет, выделяемый на привлечение специалистов из консалтинга.

В проекте требуется разработать корпоративную базу знаний, совмещенной с тикет-системой, формирующей каталог изменений в системе.

Необходимо разработать веб-портал, который решает проблему отсутствия структурированной единой базы данных, содержащей всю необходимую информацию по внедряемому в SAP изменению (Change Management) или инциденту (Incident Management).

Система должна позволять решать проблему фрагментации данных, необходимых для решения одной бизнес-задачи.

Изначально поэтапное планирование и отражение изменений осуществляется на портале ServiceNow, спецификации разработок размещены на Microsoft SharePoint, а база знаний распределена между статьями в разделе Knowledge Base на ServiceNow и презентациями, хранящимися в папке на сервере.



Рис. 1. Функции корпоративного портала по управлению изменениями

Транспорта (объекты из системы SAP), связанные с изменениями или проектами, не имеют внешней связи и нигде не указываются, их приходится искать непосредственно в SAP-системе. Документация по Job'ам (еще один объект из SAP) находится на отдельном портале.

Система должна реализовывать набор функций (рис. 1), которые можно разделить на две группы. К основным относятся:

– регистрация тикетов;
– поиск информации в едином хранилище, базе данных и документах.
– учет операций с тикетами;
– прогнозирование числа тикетов по категориям и типам.

Обеспечивающие функции нужны для поддержания работы основных. Они должны включать:

– прикрепление файлов к тикетам;

– фильтрация тикетов;
– поддержка личных кабинетов пользователей;
– поддержка пользовательского интерфейса.

Наиболее привлекательным вариантом для разработки тикет-системы выглядит PHP. Программирование на стороне клиента ведется в JavaScript. Среда разработки Visual Studio Code. СУБД для разработки тикет-системы управления изменениями – MySQL. Поскольку заявки слабо подвержены сезонности в качестве модели для прогнозирования типов заявок на изменения и инциденты использована ARIMA.

2. Проектирование и реализация системы

Составлена диаграмма прецедентов на языке моделирования UML (рис. 2).

Выделены две роли: начальник и сотрудник службы техподдержки.

У начальника полные права, включая делегирование тикетов сотрудникам. А у сотрудников есть возможность работы с делегированными им лично тикетами.

Управление тикетами включает их фильтрацию по параметрам с возможность выгрузки в формате CSV, редактирование отдельных тикетов, просмотр истории операций с тикетами, прогнозирование числа тикетов по их типам. Также у тикетов регистрируются дополнительные параметры, формируя единую базу данных (это важнейшая функция проектируемой ИС).

Детализированная архитектура системы (рис. 3).

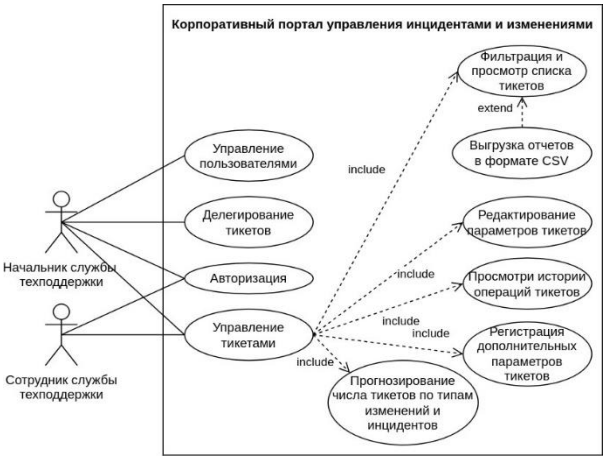


Рис. 2. UML диаграмма вариантов использования

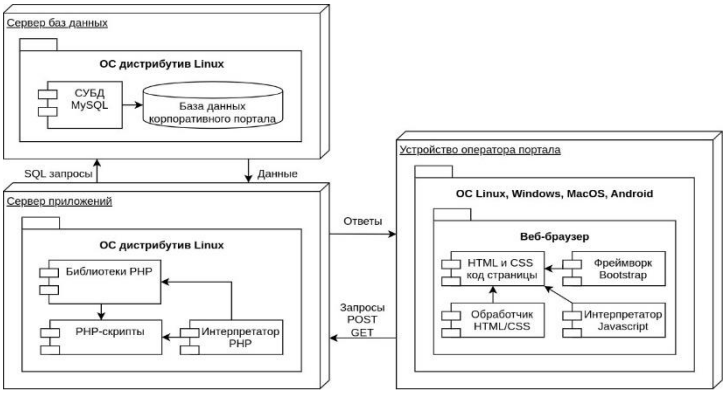


Рис. 3. UML диаграмма компонентов ИС

Устройство оператора портала может работать на любой ОС. Главное требование – это наличие современного веб-браузера, способного отображать HTML5 и CSS3, а также обрабатывать сценарии JavaScript. Взаимодействие между уровнями трехзвенной архитектуры выполняется посредством запросов и ответов.

UML диаграмма последовательностей процесса управления тикетами (рис. 4).

Взаимодействие идет посредством механизма «Запрос-ответ». Запрашивается форма от сервера, на ней устанавливаются параметры, затем происходит применение действия. То есть это классическая модель клиент-серверного взаимодействия.

Если обратится к диаграмме компонентов, то устройство пользователя является клиентом для сервера приложений, но и сам сервер приложений выступает клиентом для сервера баз данных, запрашивая у него информацию или записывая данные в БД.

Проектирование базы данных начинается с модели «Сущность-связь». Далее составляется логическая модель, которая описывающая предметную область. Она детализирует ранее составленную модель «Сущность-связь» (рис. 5). Разработка велась в среде ERWin Data Modeler.



Рис. 4. UML диаграмма последовательностей процесса управления тикетами

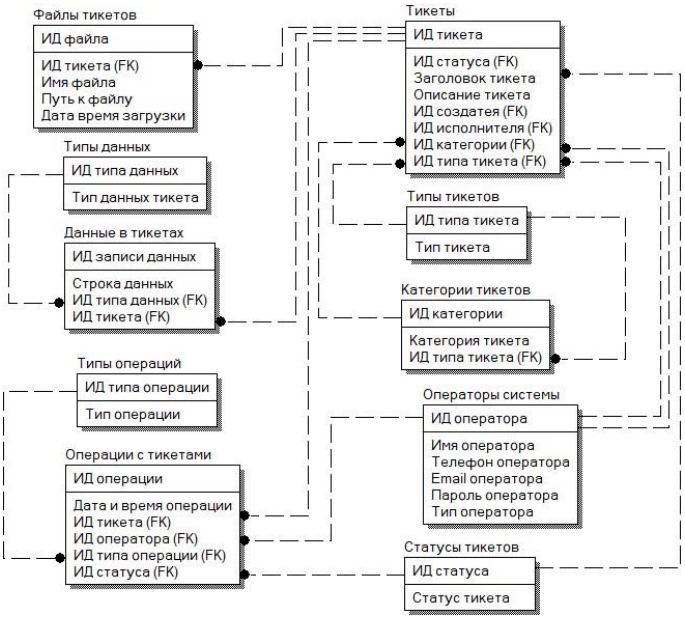


Рис. 5. Логическая модель базы данных

Тип тикета может быть либо изменением, либо инцидентом. Категория тикета отображает его разбиение на подвиды. Также у тикетов по мере их обработки меняются статусы, которые регистрируются в таблице «Операции с тикетами».

Информационная база по тикетам формируется в таблице «Данные в тикетах». То есть по мере обработки тикета к нему можно прикреплять документы, сообщения, транспорты, задачи, менять статусы. Вся эта история обработки будет храниться в БД, и можно будет в любой момент просмотреть все действия поэтапно.

Теперь можно перейти к физической модели (рис. 6).

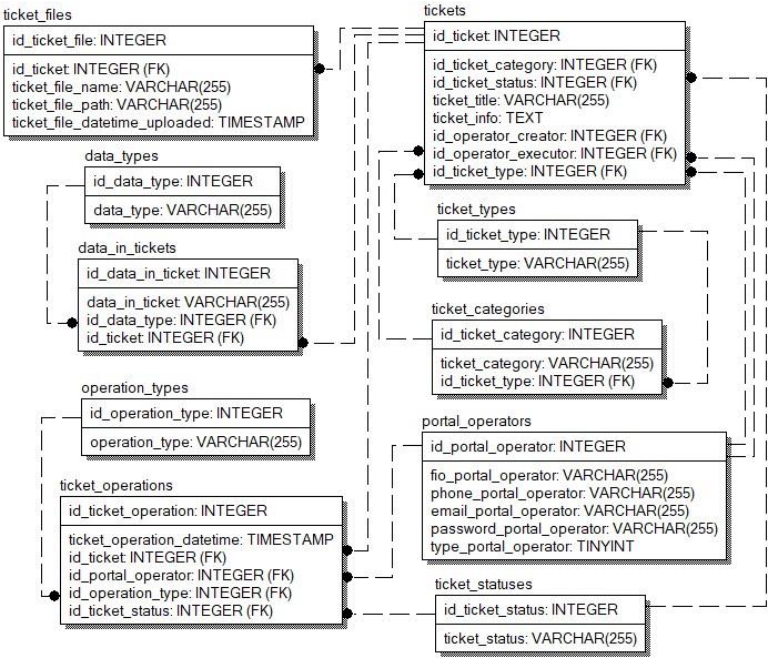


Рис. 6. Физическая модель данных

3. Дерево диалога и макеты форм интерфейса

Составлено дерево диалога корпоративного портала, показывающее доступные пользовательские формы интерфейса (рис. 7).

После страницы авторизации в зависимости от типа оператора происходит загрузка личного кабинета. Прогнозирование числа тикетов по типам и категориям доступно начальнику службы техподдержки.

На странице прогнозирования система автоматически по методу ARIMA считает число тикетов каждого типа на будущий месяц (группировка идет по месяцам, так удобней для планирования ресурсов предприятия).



Рис. 7. Дерево диалога портала с тикет-системой

4. Структурная схема

Составлена структурная схема модулей тикет-системы (рис. 8).

Модули страниц включают модуль шапки и футер, которые задают общий шаблон. Модуль шапки подключает модуль глобальных настроек, общих функций, а также модули интерфейсного фреймворка Bootstrap. В целом все названия модулей информативно описывают их назначение, поэтому в более детальном описании ну нуждаются.

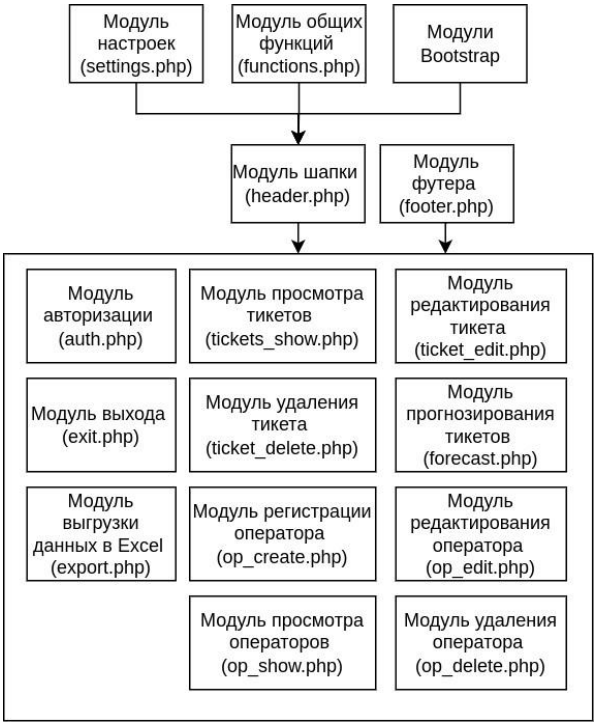


Рис. 8. Структурная схема модулей тикет-системы

Эффективность проекта обусловлена (таблица):

– снижением трудозатрат в результате внедрения проекта;
– снижением финансовых затрат.

Экономия на одном тикете составляет 1,63 часа. При количестве тикетов порядка 500 в год экономия трудозатрат составит 815 часов в год. На одном тикете происходит экономия 603,1 руб. При числе тикетов 500 в год экономия составит 301550 руб..

Срок окупаемости проекта будет равен: 263648,64 / 301550 = 0,87 года.

Таблица

Показатели эффективности проекта

	Оценка 1 часа времени сотрудника службы техподдержки, руб./час.	Время на ручное выполнение, часов	Время на выполнение с программой, часов	Экономия, часов	Стоимость процесса до автоматизации, руб.	Стоимость процесса после автоматизации, руб.	Экономия, руб.
Регистрация тикетов	370	0,15	0,01	0,14	55,5	3,7	51,8
Поиск информации	370	0,5	0,2	0,3	185	74	111
Учет операций с тикетами	370	1,2	0,5	0,7	444	185	259
Прогнозирование числа тикетов по типам и категориям	370	0,5	0,01	0,49	185	3,7	181,3
ИТОГО	1480	2,35	0,72	1,63	869,5	266,4	603,1

Заключение

В работе решена задача централизации и структурирования корпоративной информации, используемой при управлении изменениями и инцидентами в системе SAP S/4 HANA. В настоящее время имеет место высокая фрагментация данных между различными информационными системами, что приводит к росту трудозатрат, снижению прозрачности процессов и увеличению числа ошибок. Реализовано решение в виде корпоративного веб-портала, обеспечивающего единое хранилище данных, поддержку процессов Change Management и Incident Management и прогнозирование нагрузки на службу поддержки

Список литературы

1. Система SAP S/4HANA [Электронный ресурс]. – Режим доступа: https://www.sap.com/central-asia-aucasus/products/erp/s4hana.html

2. Основы и стандарты ITIL. Библиотека ITIL. Все, что нужно знать об ITIL [Электронный ресурс]. – Режим доступа: https://otus.ru/nest/post/694/
3. Среда разработки PHP Expert Editor [Электронный ресурс]. – Режим доступа: https://www.phpexperteditor.com/ru/
4. Среда разработки WebStorm [Электронный ресурс]. – Режим доступа: https://www.jetbrains.com/ru-ru/webstorm/
5. СУБД Oracle [Электронный ресурс]. – Режим доступа: https://www.oracle.com/cis/database/

Извлечение транспортных расходов из общего объема выручки в SAP S/4 HANA

И. В. Илларионов, email: igor.illarionov@gmail.com
В. В. Каплий, email: vladk.03@mail.ru

Воронежский государственный университет

Аннотация. *В работе рассматриваются подходы к извлечению транспортных расходов в SAP S/4 HANA, проводится анализ существующих решений и практик, предлагаются методики, позволяющие автоматизировать данный процесс. Особое внимание уделено практическому применению разработанных решений на примере конкретной компании, что позволяет оценить их эффективность в реальных условиях.*

Ключевые слова: *информационная система, информационные технологии, ERP-системы, транспортные расходы, выручка, SAP S/4 HANA, SAP GRC PC.*

Введение

Реализация автоматической контрольной процедуры по извлечению транспортных расходов из общего объема выручки в SAP S/4 HANA — многоступенчатый процесс, включающий в себя определение требований, выбор методов контроля, подбор программных средств, тестирование, запуск в продуктивной системе заказчика, а также последующий сбор данных для анализа эффективности внедренной процедуры.

Внедрение автоматических контрольных процедур обладает рядом значительных преимуществ. Оно позволяет существенно повысить точность и прозрачность учета, снизить количество ошибок, автоматизировать рутинные операции, а также повысить уровень доверия к финансовой отчетности организации со стороны как внутренних, так и внешних пользователей.

Среди ключевых преимуществ автоматических контрольных процедур можно выделить:

– Автоматизация процессов. Контрольная процедура по извлечению транспортных расходов автоматизирует ранее ручные операции, связанные с распределением выручки. Это минимизирует влияние человеческого фактора и повышает

© Илларионов И. В., Каплий В. В., 2026

согласованность отражения расходов на транспортировку в отчетности.
- Повышение точности. АКП позволяют обрабатывать большие объемы данных в реальном времени, мгновенно выявляя расхождения и ошибки, возникающие при некорректном распределении транспортных расходов, и информируя сотрудников о найденных несоответствиях.
- Снижение рисков. Автоматическая проверка правильности отражения транспортных расходов позволяет значительно снизить риски ошибок, искажений и возможных претензий со стороны контролирующих органов, а также внутренних аудиторов.
- Оперативный доступ к информации. АКП дают возможность ответственным лицам получать информацию о расхождениях и нарушениях немедленно, что ускоряет принятие управленческих решений и повышает эффективность взаимодействия между подразделениями.

1. Автоматизация процессов в SAP S/4 HANA

Автоматизация процессов является одной из ключевых концепций в архитектуре SAP S/4HANA. Система предлагает широкие возможности для минимизации ручного труда, сокращения количества ошибок, повышения прозрачности операций и ускорения обработки бизнес-событий в режиме реального времени. Это особенно актуально для задач, связанных с выделением и контролем транспортных расходов, поскольку такие затраты формируются на основе множества разнотипных документов – от заказов и отгрузок до счетов и платежей.

SAP S/4HANA поддерживает автоматизацию на всех этапах работы с финансовыми и логистическими данными. Среди основных механизмов автоматизации можно выделить:

- Автоматическое формирование бухгалтерских проводок при создании или изменении логистических документов (например, при поставке товара автоматически формируется проводка, отражающая выручку и расходы на транспортировку);
- Механизмы автоматического расчета и распределения затрат, в том числе с использованием шаблонов и правил оценки (например, транспортные расходы могут быть распределены по заказам или направлениям пропорционально объему отгрузки);
- Использование правил учета (account determination), которые позволяют системе самостоятельно определять нужные бухгалтерские счета для отражения выручки и затрат, включая счета для транспортных расходов;

– Автоматическая интеграция между модулями FI, CO, SD, MM и TM, благодаря которой данные о логистических операциях (например, маршруты, грузоперевозки, поставки) автоматически передаются в финансовый учет и становятся частью общего аналитического массива;
– Настраиваемые рабочие процессы (workflow), обеспечивающие согласование, утверждение и контроль бизнес-процессов без участия пользователя, включая автоматическую проверку затрат на транспорт при закрытии периода или формировании отчетности;

Интеллектуальные алгоритмы обработки данных (например, на базе SAP Machine Learning и SAP Business Rules Framework), которые позволяют предсказывать транспортные расходы, выявлять отклонения и автоматически инициировать корректирующие действия.

В задаче извлечения транспортных расходов из выручки автоматизация играет определяющую роль, поскольку позволяет:

– Исключить влияние человеческого фактора при выделении соответствующих затрат;
– Повысить точность и скорость расчетов;
– Формировать отчетность в режиме реального времени;
– Создавать репозитории данных, пригодные для последующего анализа и контроля.

Дополнительную ценность приносит использование SAP Fiori-приложений, которые позволяют в интуитивно понятной форме отслеживать ключевые показатели (KPI), управлять аналитикой и проводить drill-down анализ транспортных расходов на любом уровне детализации – от глобальной выручки до конкретного клиента или вида продукции.

Таким образом, SAP S/4HANA обеспечивает высокий уровень автоматизации, необходимый для системного подхода к идентификации и анализу транспортных расходов в составе выручки. Это не только упрощает учет и контроль, но и открывает возможности для оптимизации затрат, повышения рентабельности и стратегического управления логистическими направлениями деятельности компании.

2. Анализ существующих подходов к извлечению транспортных расходов

Модуль FI (Financial Accounting) в SAP S/4HANA включает в себя различные встроенные инструменты для учета и анализа затрат, включая транспортные расходы. Эти инструменты позволяют фиксировать и классифицировать затраты в разрезе счетов и бизнес-операций, что является основой для последующего анализа. Однако для полноценного

использования возможностей модуля важно учитывать особенности каждого из них.

Стандартные проводки и аналитика по счетам затрат обеспечивают отражение транспортных расходов на отдельных счетах бухгалтерского учета, что позволяет выделять их из общей суммы затрат. Однако если транспортные расходы изначально не отделены от стоимости товаров или услуг в первичных документах, то точное их извлечение становится проблематичным и требует дополнительной детализации.

Финансовые отчеты, формируемые через Report Painter/Writer и стандартные формы SAP, предоставляют агрегированные данные о суммарных транспортных расходах по счетам, периодам и организационным единицам. При этом они не всегда позволяют провести многомерный анализ, например, по клиентам или проектам, что ограничивает возможности глубокой детализации.

Автоматизация проводок на основании первичных документов, таких как счета поставщиков транспортных услуг, упрощает и стандартизирует учет транспортных расходов. При отсутствии правильной настройки правил и шаблонов проводок есть риск некорректной классификации или распределения затрат, что снижает точность учета.

Использование дополнительных аналитических признаков (Business Area, Functional Area, Profit Center) дает возможность более точной классификации затрат, в том числе транспортных, и их анализа в разрезе организационных единиц и направлений деятельности. Однако для этого необходимо строгое и последовательное заполнение этих признаков при вводе первичных документов – в противном случае данные окажутся неполными и не позволят провести достоверный анализ.

Интеграция FI с системами внешней отчетности и консолидации позволяет использовать данные о транспортных расходах в итоговых отчетах для управленческого и внешнего аудита. В такой схеме обычно доступны только сводные суммы расходов, что не дает возможности детально проанализировать структуру транспортных затрат без дополнительных инструментов.

Таким образом, модуль FI SAP S/4HANA предоставляет необходимые базовые возможности для учета и анализа транспортных расходов. Для получения точной и детальной аналитики требуется правильная настройка учетных процедур и тесная интеграция с другими модулями системы.

3. Определение требований к алгоритму

Первым шагом при разработке механизма автоматического извлечения транспортных расходов в SAP S/4HANA является формулировка требований к данному процессу. На этом этапе определяется, какие задачи должен решать механизм, какие данные подлежат обработке, а также устанавливаются ключевые параметры и характеристики реализации.

Ключевыми задачами на данном этапе являются:

- Определение области применения механизма: необходимо определить, какие транзакции и объекты учета в SAP S/4HANA будут включены в обработку. В частности, устанавливаются типы бухгалтерских документов в модуле FI, содержащие транспортные компоненты затрат, а также специфика аналитик (например, cost elements, document types, delivery-related postings).
- Определение критериев корректного извлечения: формируются правила, по которым можно считать процесс извлечения выполненным корректно. Например, это может быть соответствие суммы транспортных расходов между первичными учетными данными и результатами выборки, либо точность распределения по видам продукции или заказам.
- Определение методов валидации результатов: необходимо обозначить подходы к тестированию механизма извлечения. Это может включать выборку реальных данных для сравнительного анализа, использование контрольных выборок или настройку тестовых сценариев с заранее известными значениями.
- Определение частоты выполнения процедуры: определяется, как часто должен запускаться механизм извлечения – в рамках ежедневной отчетности, по окончании отчетного периода или в режиме реального времени (при использовании CDS Views и embedded analytics).
- Формализация параметров извлечения: уточняются ключевые параметры процедуры, включая перечень транспортных расходов (например, расходы на логистику, доставку, перемещение), правила фильтрации и отбора данных, а также возможные действия по обработке исключений или отклонений от нормы.

Результатом данного этапа является создание документа требований, в котором фиксируются цели, границы, логика обработки и технические характеристики процедуры извлечения транспортных расходов. Этот документ служит основой для дальнейшей реализации

логики в системе SAP S/4HANA, включая настройку трансформаций, потоков данных и хранилищ в рамках архитектуры SAP BW/4HANA или Embedded Analytics.

4. Определение методов контроля

После того как требования к извлечению транспортных расходов из структуры выручки были определены, следующим этапом является выбор и формализация методов контроля, которые будут применяться в рамках автоматической процедуры. Каждый метод должен быть логически привязан к конкретному требованию, а также учитывать специфику учета и отражения транспортных затрат в модуле FI SAP S/4HANA.

В SAP S/4HANA для реализации автоматического контроля целесообразно использовать следующие методы:

- Проверка полноты данных. Целью этого метода является убедиться, что все релевантные финансовые документы, содержащие транспортные расходы, корректно загружены и присутствуют в системе. Применяется сверка значений по определённым G/L-счетам, document types и cost elements в модуле FI, а также кросс-сравнение с логистическими источниками (например, Delivery или Shipment) при наличии связи через Accounting Interface.
- Проверка точности данных. Метод применяется для проверки правильности отнесения сумм к транспортным расходам. Это включает в себя анализ правильного использования account assignment objects (например, cost center, internal order), корректность сумм и валют, а также соответствие расчетов нормативам логистики. Часто реализуется через сопоставление сумм с расчетной базой (расходы на доставку по контракту, тарифные ставки).
- Проверка соответствия правилам учета. Этот контроль направлен на соответствие данных внутри FI установленным бухгалтерским политикам компании. Применяются правила, например, что транспортные расходы не могут быть отнесены на доходные счета, либо что они должны использовать только определённые cost elements или profit center. Проверка может производиться средствами validation rules или BRF+ логики.
- Проверка дублирующих проводок. В рамках обработки большого объема финансовых документов возможны дублирующие записи. Метод сравнивает документы с идентичными суммами, датами и ссылками на объекты логистики. При этом особое внимание уделяется документам,

загруженным из внешних систем через интерфейсы (например, при использовании middleware или IDoc).

– Проверка на наличие аномалий. Здесь осуществляется контроль на предмет нетипичных значений. Примеры: сумма транспортных расходов превышает допустимое процентное значение от общей суммы выручки; резкий рост затрат в отчетном периоде; отсутствие транспортных расходов при наличии отгрузки. Используются методы статистического анализа, а также построение baseline-моделей, отклонения от которых сигнализируют об ошибке или необходимости верификации.

Каждый из этих методов контроля может быть реализован в виде отчетов на базе CDS Views, аналитических queries, либо через batch jobs с уведомлениями по результатам проверок. Совокупность подходов позволяет обеспечить автоматизированный и системный подход к контролю извлечения транспортных расходов, повысив надежность отчетности и точность аналитики.

При выборе программных средств реализации автоматической процедуры контроля транспортных расходов важно учитывать как специфику бизнес-требований, так и техническую совместимость с архитектурой SAP S/4HANA. На практике наилучшие результаты достигаются при комбинированном использовании SAP BIS для интеллектуального анализа транзакций, SAP Process Control для постоянного мониторинга данных в FI, а также SAP Audit Management и SAP GRC для контроля соответствия, аудита и управления рисками.

5. Тестирование алгоритма

После реализации автоматической контрольной процедуры, направленной на проверку корректности извлечения транспортных расходов в модуле FI SAP S/4HANA, проводится этап функционального и технического тестирования. Его цель – убедиться, что разработанный механизм корректно определяет транспортные расходы в первичных документах и соблюдает заданные условия контроля.

Подготовка тестовой выборки. На этом этапе подбираются бухгалтерские документы, содержащие записи транспортных расходов. При этом важно обеспечить охват разных сценариев: корректные проводки, расходы, ошибочно отнесённые на неподходящие счета, случаи отсутствия обязательных аналитик (например, объекта учета, поставщика, направления движения средств). Выборка составляется на основе данных в таблицах BSEG, ACDOCA, а также с учетом специфики проводок по транспортным услугам (например, тип документа, комбинация счетов и налоговых кодов).

Контрольная процедура загружается в QA-среду через Change Request. Далее она запускается с использованием разработанных правил и фильтров, реализованных, например, через CDS Views, ABAP-правила или SAP Business Integrity Screening. Контроль выполняется по заранее определенным параметрам – допустимые счета учета, наличие ссылок на контрагента, правильность кодов операций.

После ее выполнения формируются результаты в виде логов, отчетов или таблиц нарушений. Система должна корректно определять документы, в которых:

- транспортные расходы отсутствуют, несмотря на наличие признаков (например, объект закупки или договор с перевозчиком);
- расходы отражены на некорректных счетах бухгалтерского учета;
- нарушена структура аналитик (например, не указан соответствующий поставщик, не задан центр затрат);
- дублируются расходы на одну и ту же перевозку.

Результаты процедуры сравниваются с эталонными значениями. Анализ проводится как на уровне контента (правильность логики отбора), так и на уровне технической реализации (проверка фильтрации, объединения данных, вычислений). При необходимости производится трассировка логики извлечения данных по ключевым полям (например, KOSTL, LIFNR, HKONT).

На основе анализа вносятся изменения в логику процедуры – корректируются условия фильтрации, уточняются правила сопоставления данных, вносятся изменения в источники (например, уточняются признаки транспортных расходов в FI-документах). После этого процедура повторно запускается на том же или расширенном наборе данных.

Финальный запуск демонстрирует устойчивую работу алгоритма: корректное выявление транспортных расходов и фиксирование всех случаев, отклоняющихся от заданных критериев. После утверждения заказчиком, процедура переносится в продуктивную систему для постоянного мониторинга.

Тестирование играет ключевую роль в обеспечении качества автоматической контрольной процедуры. Корректное извлечение транспортных расходов важно как для соблюдения бухгалтерской точности, так и для последующего анализа затрат в системах управленческого учета.

6. Запуск алгоритма в продуктивной системе заказчика

Этап запуска автоматической контрольной процедуры по извлечению транспортных расходов в модуле FI SAP S/4HANA является критическим, так как он переводит разработанную логику из тестовой среды в реальную продуктивную систему заказчика. Целью этого этапа является обеспечение стабильной и корректной работы процедуры в условиях действующей бизнес-инфраструктуры, с соблюдением всех технических, организационных и регламентных требований.

Перед установкой проводится проверка корректности всех настроек и параметров процедуры, включая используемые источники данных (например, ACDOCA, BSEG), трансформации, фильтры и расписания запуска. Особое внимание уделяется соответствию корпоративным требованиям информационной безопасности, ролям доступа и разрешениям пользователей, задействованных в контроле транспортных расходов.

Для минимизации рисков возможных сбоев или некорректной работы процедуры в новой среде создается бэкап продуктивной системы. Это позволяет при необходимости откатить изменения и восстановить систему до первоначального состояния.

Процедура устанавливается через Change Request, импортируется в продуктивную систему и интегрируется в существующий ландшафт SAP. На этом этапе подключаются необходимые объекты: CDS Views, правила фильтрации (например, по счетам учета HKONT, поставщикам LIFNR), используемые InfoProvider'ы (например, ADSO), а также расписание через SAP Job Scheduler.

После установки проводится функциональное тестирование на ограниченной выборке реальных данных. Проверяется корректность извлечения расходов, определение некорректных проводок (например, случаи, когда транспортные расходы отнесены не на те счета или без обязательных аналитик), а также формирование результатов в требуемом формате.

После валидации процедуры на ограниченной выборке производится настройка периодичности ее запуска – например, ежедневный или еженедельный запуск в фоновом режиме. Это позволяет встроить контрольную процедуру в регулярный цикл финансовой отчетности и оперативного контроля затрат.

Контрольная процедура повторно тестируется в продуктивной среде с полной выборкой данных. Оценивается ее производительность, полнота извлечения информации и точность обнаружения отклонений. При необходимости вносятся точечные корректировки.

Проводится обучение ключевых пользователей по интерпретации результатов работы процедуры, взаимодействию с отчетами и действиям при выявлении отклонений. Обучение охватывает как базовых пользователей, работающих с отчетностью, так и специалистов контроллинга и финансового отдела, ответственных за дальнейшую обработку обнаруженных случаев.

После всех проверок, обучений и корректировок процедура переводится в рабочий режим. С этого момента она выполняется в продуктивной системе по заданному расписанию и служит инструментом автоматизированного контроля корректности отражения транспортных расходов в бухгалтерских документах.

Этап продуктивного запуска завершает цикл внедрения автоматической контрольной процедуры и открывает этап эксплуатационного мониторинга. С этого момента заказчик получает регулярный инструмент анализа и контроля, который повышает качество данных и снижает финансовые риски, связанные с некорректным учетом транспортных затрат.

7. Мониторинг и сбор данных

Конечной целью извлечения транспортных расходов из общего объема выручки в SAP S/4 HANA является обеспечение корректности и прозрачности финансовых данных для анализа и отчетности. Однако даже после настройки автоматических процедур распределения транспортных расходов, могут возникать ситуации, требующие мониторинга и сбора данных для анализа и последующей оптимизации процесса.

В SAP S/4 HANA для мониторинга и сбора данных по выделению транспортных расходов из выручки используются различные инструменты и технологии. Основные из них:

SAP предоставляет набор инструментов для контроля состояния системы и выявления возможных проблем. С их помощью можно отслеживать корректность и своевременность выделения транспортных расходов из общего объема выручки, а также собирать данные для дальнейшего анализа.

В SAP S/4 HANA реализован мощный механизм построения отчетов, позволяющий формировать отчеты на основании данных по распределению транспортных расходов. С помощью этих отчетов можно контролировать результаты выделения расходов, оценивать их влияние на финансовые показатели и выявлять отклонения.

Для углубленного анализа и визуализации данных по транспортным расходам могут использоваться решения, такие как SAP Analytics Cloud или SAP Lumira. Эти инструменты позволяют проводить анализ в

реальном времени и создавать дашборды, помогающие принимать обоснованные бизнес-решения.

Для оперативного реагирования на ошибки или несоответствия при выделении транспортных расходов применяются системы автоматических уведомлений. Они могут отправлять сообщения на электронную почту или мобильные устройства ответственных сотрудников.

В целом процесс мониторинга и сбора данных по выделению транспортных расходов из выручки в SAP S/4 HANA включает следующие этапы:

- Определение ключевых показателей мониторинга. Необходимо определить метрики, которые будут отслеживаться, например, доля транспортных расходов в общей выручке, количество ошибок распределения, время выполнения процедуры.
- Настройка системы мониторинга. На этом этапе настраивается отслеживание выбранных показателей, включая конфигурацию уведомлений, мониторинг производительности и корректности расчетов.
- Анализ собранных данных. После запуска процедуры и сбора данных проводится их анализ для выявления проблемных зон и определения возможностей для улучшения процесса выделения транспортных расходов.
- Отчетность. На основании собранной информации формируются отчеты для заинтересованных лиц, содержащие статистику, результаты мониторинга и рекомендации по оптимизации.
- Оптимизация процесса. На основании анализа и отчетов принимаются меры по совершенствованию процедуры выделения транспортных расходов, включая корректировку настроек, улучшение алгоритмов и повышение эффективности системы в целом.

8. Тестирование

При подготовке к запуску автоматической процедуры извлечения транспортных расходов из общего объема выручки было проведено функциональное и интеграционное тестирование. В рамках тестирования был смоделирован запуск автоматической процедуры на реальных операционных данных компании.

Целью тестирования являлась проверка корректности алгоритмов извлечения транспортных расходов, точности отбора условий (например, FRA1, ZFRT и других) и правильности их распределения в отдельную статью учета. Тестирование позволило убедиться, что процедура

работает в соответствии с заданными правилами, корректно выделяет транспортные компоненты из общей выручки и формирует отчетность без отклонений.

Автоматическая контрольная процедура “Извлечение транспортных расходов из общего объема выручки” была развернута и успешно запущена в продуктивной системе SAP S/4 HANA заказчика.

С помощью планировщика контрольных процедур был настроен ежемесячный запуск данной процедуры (каждое 1-е число месяца). Выбранная периодичность обеспечивает регулярную проверку выручки, полученной за прошедший месяц, позволяет оперативно выявлять некорректное включение транспортных расходов в состав доходов, при этом не перегружает систему ежедневными проверками. При необходимости руководитель отдела или ответственный сотрудник может инициировать запуск процедуры в ручном режиме.

После запуска автоматической контрольной процедуры был налажен механизм мониторинга и сбора данных. Основной задачей мониторинга являлось определение, как изменилась работа подразделений после внедрения АКП, а также оценка эффективности внедренного решения.

9. Результаты реализации алгоритма

Оценка эффективности и анализ результатов внедрения автоматических контрольных процедур в SAP S/4 HANA является одним из ключевых этапов реализации данного решения. Для оценки эффективности и анализа результатов применяются различные методы и инструменты.

Одним из основных методов оценки является сравнение показателей до и после внедрения автоматических контрольных процедур. На основании данных, представленных в таблице 6, можно сделать вывод о том, что общее количество ошибок, связанных с некорректным извлечением транспортных расходов из выручки, после внедрения автоматической контрольной процедуры 1 марта 2025 года существенно сократилось. Также было отмечено снижение трудозатрат на их устранение (рис. 1, рис. 2).

Значительное снижение временных затрат по сравнению с сокращением общего числа ошибок объясняется спецификой работы автоматических контрольных процедур. АКП идентифицируют ошибку и указывают на её источник, избавляя сотрудников от необходимости самостоятельно искать причину, что ускоряет процесс устранения и позволяет перераспределить ресурсы на другие задачи.

На рис. 3 представлено распределение ошибок по видам. Анализ позволяет отметить снижение доли технических ошибок на 8,73%. Увеличение доли прочих ошибок связано с общим сокращением

количества ошибок и, соответственно, изменением долевого распределения.

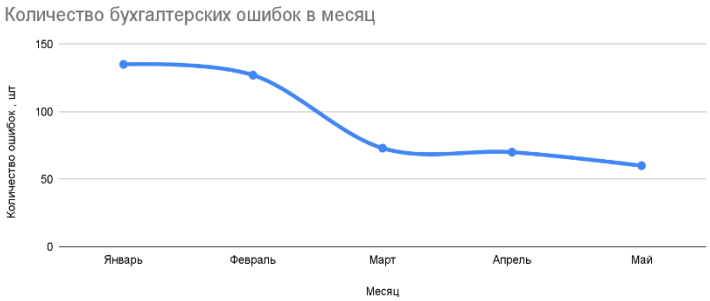


Рис. 1. Снижении количества ошибок после введения АКП

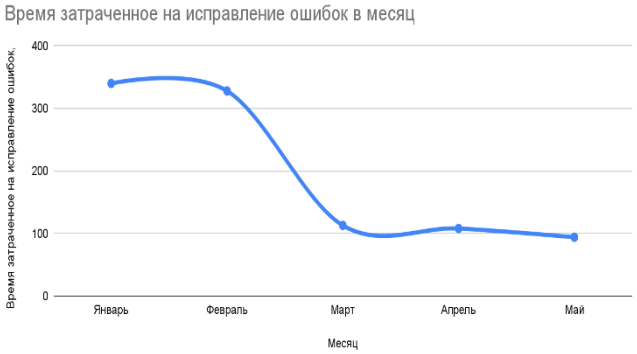


Рис. 2. Уменьшение временных затрат на исправление ошибок



Рис. 3. Распределение ошибок по видам

Таким образом, можно с уверенностью утверждать, что внедрение автоматической контрольной процедуры, направленной на корректное извлечение транспортных расходов из состава выручки, существенно повышает прозрачность финансовой отчетности, снижает риск ошибок и повышает операционную эффективность. Помимо количественного сокращения ошибок, АКП позволяет значительно сократить время на их поиск и исправление, тем самым повышая общую производительность бухгалтерского и финансового подразделений.

Список источников

1. Бранд, Х. SAP R/3 и SAP S/4HANA. Управление бизнесом от процессов до систем / Х. Бранд, Л. Моних, Р. Хинкельманн. – 1-е. – Вальдорф: SAP Press, 2016 – 250 с.
2. Фролов, А. SAP S/4HANA: функциональные возможности и управление проектами / А. Фролов, А. Буянов – 1-е. – М.: SAP Press, 2018 – 248 с.
3. Хоппер, Р. SAP S/4HANA. Управление компанией. Настройка и использование / Р. Хоппер – 1-е. – Вальдорф: SAP Press, 2020 – 185 с.
4. Крылова, Г. Программирование на ABAP для SAP S/4HANA. / Г. Крылова, А. Чекалюк – 1-е – М.: SAP® PRESS, 2020 420 с.
5. Броун, Р. SAP HANA для разработчиков. / Р. Броун – Вальдорфе: SAP Press, 2019 – 180 с.
6. Патель, М. Руководство по сертификации SAP S/4 HANA в финансовом учете. / М. Патель, Н. Неру – М.: Издательство ABC, 2018. – 300 с.

Разработка алгоритма автоматизированного определения приоритета заявок технической поддержки на основе метода анализа иерархий (МАИ) в системе SAP Service Desk

И. В. Илларионов, email: igor.illarionov@gmail.com [1]
Д. Р. Крупчак, email: krupchak.dim9@gmail.com [1]

[1] Воронежский государственный университет

Аннотация. *разработан алгоритм автоматизированного определения приоритета инцидента за счёт замены компоненты Impact на модель, построенную с использованием метода анализа иерархий (МАИ, AHP), решение интегрировано в SAP Service Desk посредством Z-класса. Предложен аналитический блок для оценки влияния новой модели на показатели SLA.*

Ключевые слова: *SAP Service Desk, метод анализа иерархий.*

Введение

Эффективная работа службы технической поддержки является ключевым фактором обеспечения стабильности информационных систем и непрерывности бизнес-процессов. В условиях цифровизации объём обращений пользователей возрастает, что требует автоматизации и повышения качества решений о приоритетах инцидентов. В стандартной функциональности SAP Service Desk приоритет определяется по формуле Priority = f(Impact, Urgency), где Impact отражает масштаб воздействия на бизнес, а Urgency – срочность решения. Однако практическая эксплуатация показывает, что стандартные шкалы Impact часто оказываются излишне грубыми и не учитывают множество факторов, влияющих на истинную критичность инцидента.

Цель настоящей работы – разработать алгоритм автоматизированного определения приоритета инцидента за счёт замены компоненты Impact на модель, построенную с использованием метода анализа иерархий (МАИ, AHP), и интегрировать решение в SAP Service Desk посредством Z-класса. Кроме того, предлагается аналитический блок для оценки влияния новой модели на показатели SLA.

1. Общая постановка задачи

Стандартная матрица SAP Impact × Urgency имеет ограниченное число состояний и выглядит, в упрощённой форме, следующим образом (табл. 1):

© Илларионов И. В., Крупчак Д. Р., 2026

Таблица 1

Классическая матрица SAP Service Desk для определения приоритета

	Low	Medium	High
Low	Priority 4	Priority 3	Priority 2
Medium	Priority 3	Priority 2	Priority 1
High	Priority 2	Priority 1	Priority 1

Такая матрица удобна, но слишком упрощённая: Impact принимает всего три состояния и не учитывает множественные транспортирующие факторы (масштаб, важность процесса, финансовые риски и т. п.). Наша задача – заменить упрощённый Impact на интегральный показатель, рассчитываемый по принципам МАИ, чтобы приоритизация отражала экспертные знания и реальные характеристики инцидента.

2. Общая архитектура приложения

Решение интегрируется в стандартный процесс обработки инцидентов SAP Solution Manager / SAP Service Desk.

Первым этапом является автоматический сбор информации о заявке. Система анализирует объект инцидента, извлекая из него данные, необходимые для дальнейшего применения МАИ. Это могут быть сведения о бизнес-процессе, в рамках которого произошла ошибка, характеристики подразделения-инициатора, информация о масштабе влияния, признаке повторяемости, а также о наличии или отсутствии временного решения, позволяющего продолжить работу.

– Сам модуль включает в себя следующие параметры:
– тип инцидента,
– затронутый бизнес-процесс,
– количество пользователей, которых коснулась проблема,
– наличие обходного решения,
– критичность подразделения-инициатора,
– прочие критерии, определяемые аналитической моделью.

Задача модуля заключается не только в том, чтобы собрать данные, но и привести их в стандартизированный вид, удобный для последующего расчёта. Каждому параметру назначаются числовые значения или категориальные оценки, которые используются в расчетной модели.

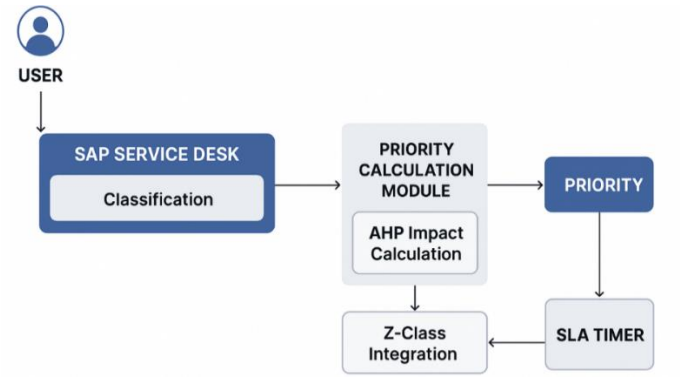


Рис. 1. Архитектура модуля автоматизированного расчёта приоритета в SAP Service Desk.

Ядром всего решения является модуль расчёта интегральной величины MAI Impact. Он применяет метод анализа иерархий, который позволяет определить относительную важность различных факторов. Сначала формируется набор критериев, которые организация считает значимыми при оценке влияния инцидента. Для каждого критерия предусмотрена собственная шкала, отражающая степень его проявления в конкретной заявке.

Далее используется матрица парных сравнений критериев, сформированная на этапе настройки. Она задаёт относительную важность критериев друг относительно друга. Формально матрица фиксируется на уровне кастомизации SAP, но может адаптироваться при изменении внутренних процессов.

После этого проводится вычисление весов критериев – значений, отражающих степень их влияния на итоговую оценку. Метод анализа иерархий требует также проверки согласованности экспертных оценок; если матрица содержит значимые противоречия, система может использовать резервный набор весов.

В заключение модуль объединяет реальные значения критериев с их весами и вычисляет итоговый показатель MAI Impact. Полученная величина нормируется к стандартным шкалам SAP, что позволяет бесшовно интегрировать её в существующие процессы.

После получения MAI Impact система использует привычную формулу SAP Priority = f(Impact, Urgency), заменяя лишь один из её компонентов. Параметр Urgency рассчитывается стандартными

средствами SAP и не требует изменений. Такой подход позволяет внедрить новый механизм без вмешательства в основную структуру системы поддержки и сохранить совместимость с существующими отчётами и SLA-политиками.

Расчёт и применение нового значения выполняется на уровне кастомного ABAP-класса. Этот класс отвечает за организацию всех вычислительных этапов: извлечение данных, вызов модуля МАИ, получение результата, обновление приоритета и логирование результатов. Класс подгружается при создании или изменении инцидента, что обеспечивает автоматичность и прозрачность процесса.

Кроме того, Z-класс может выполнять проверку корректности данных, вызывать резервные алгоритмы, а также предоставлять сервисные методы для отладки или анализа.

Чтобы оценить эффективность новой модели, архитектура решения включает аналитический модуль, который собирает статистику по времени реакции, количеству нарушенных SLA, распределению заявок по приоритетам до и после внедрения. Это позволяет определить, улучшилась ли точность и справедливость приоритизации, и выявить области, где МАИ оказывает наиболее заметное влияние.

3. Разбор отдельного модуля: модуль вычисления MAI Impact

Это ядро решения: здесь задаются критерии, формируется матрица парных сравнений, рассчитываются веса и итоговый Impact Score.

На этапе определения модели организация выбирает критерии, которые наилучшим образом описывают влияние инцидента. Каждый критерий должен быть измеримым и связанным с фактическими атрибутами заявки. Задача модуля – корректно преобразовать эти данные в числовой показатель, входящий в расчетную формулу.

В качестве примера используются четыре измеримых критерия:

K1 – Масштаб инцидента (доля/количество затронутых пользователей);

K2 – Важность бизнес-процесса (оценка влияния на ключевые процессы);

K3 – Критичность подразделения (уровень SLA/приоритетности подразделения);

K4 – Наличие обходного решения (наличие временной альтернативы снижает влияние).

Каждый критерий имеет свою шкалу измерения и переводится в нормированное значение $x_i \in [0,1]$

Метод Саати (AHP) использует шкалу 1–9 для парных сравнений: число a_{ij} показывает, насколько критерий i важнее критерия j. В исходной матрице A диагональные элементы равны 1, а $a_{ij} = 1/a_{ij}$

Пример матрицы парных сравнений A (4x4) (табл. 2):

Таблица 2

Матрица парных сравнений

	K1	K2	K3	K4
K1	1	3	5	7
K2	1/3	1	2	4
K3	1/5	0,5	1	3
K4	1/7	0,25	1/3	1

Эта матрица отражает экспертное мнение: K1 заметно важнее K2–K4, K2 значимо важнее K3 и K4 и т.д.

Для установления относительной значимости критериев используется матрица парных сравнений. Эксперты определяют, насколько один критерий важнее другого, выбирая оценки по шкале Саати. Из этих данных формируется квадратная матрица, в которой каждая ячейка отражает предпочтительность одного критерия над другим. Матрица сохраняется в таблицах SAP и используется во всех дальнейших вычислениях.

Метод анализа иерархий основывается на подходе, предложенном Томасом Саати. Для определения относительной важности критериев используется шкала Саати (1–9), где каждому числу соответствует степень превосходства одного критерия над другим. Оценки заносятся в матрицу парных сравнений размерности N×N, в которой элемент a_{ij} отражает предпочтительность критерия i перед j.

Если критерий i важнее j в k раз, то:

$a_{ij} = k$

$a_{ji} = 1/k$

Заполнение такой матрицы является основой дальнейшего расчёта весов критериев.

После формирования матрицы модуль рассчитывает веса критериев. Это делается путём вычисления собственного вектора матрицы, который отражает распределение значимости факторов. Веса нормируются таким образом, чтобы их сумма равнялась единице, и затем применяются в формуле расчёта интегральной оценки.

Шаг 1 – суммы столбцов

Вычисляем сумму каждого столбца $S_j=\sum a_{ij}$

s1=1+1/3+1/5+1/7≈1.67619048

s2=3+1+1/2+1/4=4.75

s3=5+2+1+1/3≈8.33333333

s4=7+4+3+1=15

Шаг 2 – нормализация матрицы (делим каждый элемент на сумму столбца) (табл. 3).

Таблица 3

Нормализованная таблица

	K1	K2	K3	K4
K1	0,5966	0,6316	0,600	0,4667
K2	0,1989	0,2105	0,2400	0,2667
K3	0,1193	0,1053	0,1200	0,2000
K4	0,0852	0,0526	0,0400	0,0667

Шаг 3 – среднее по строкам → векторы весов (не нормированные)

Вычисляем среднее значение по каждой строке нормированной матрицы:

$w'_1 = \text{mean}(0.5966, 0.6316, 0.6000, 0.4667) \approx 0.573709$

$w'_2 \approx 0.229014$

$w'_3 \approx 0.136145$

$w'_4 \approx 0.061131$

Эти значения уже нормированы (их сумма ≈ 1). Для контроля можно нормировать явным делением на сумму – в данном примере сумма ≈ 1.000000 (табл. 4).

Таблица 4

Итоговые веса (округлённо)

Критерий	Вес (Wi)
K1	0.5737
K2	0.2290
K3	0.1361
K4	0.0611
Сумма	1.0000

Эти веса показывают относительную важность критериев: K1 наиболее значим, затем K2, K3 и K4.

Важным свойством метода анализа иерархий является требование согласованности экспертных оценок. Если матрица содержит противоречия, весовая модель будет ненадёжной. Поэтому модуль вычисляет коэффициент согласованности и сравнивает его с допустимым порогом. В случае его превышения система может применить резервный набор весов или уведомить администратора.

Полученные веса показывают, что в конкретной экспертной оценке масштаб инцидента (K1) доминирует (≈57%), затем важность процесса, критичность подразделения и, наконец, наличие обходного решения. Эти веса используются при вычислении итогового Impact Score для каждой конкретной заявки.

Итоговое значение MAI Impact определяется как сумма произведений веса каждого критерия и его фактической оценки в рассматриваемой заявке. После получения результата он нормируется в диапазон значений SAP и интегрируется в существующую таблицу определения приоритета. Таким образом, итоговая величина учитывает как количественные параметры инцидента, так и экспертные знания, заложенные в модель.

Теперь применим рассчитанные веса к реальной заявке (табл. 5). Для каждого критерия вычислено нормированное значение $x_i \in [0,1]$.

Таблица 5

Веса реальной заявки

Критерий	Вес (Wi)	Фактическое значение (Xi)	Произведение Wi * Xi
K1 – Масштаб инцидента	0,5737	0,80	0,4590
K2 – Важность процесса	0,2290	1,00	0,2290
K3 – Критичность подразделения	0,1361	0,50	0,0680
K4 – Наличие обходного решения	0,0611	0,20	0,0122
Итоговый показатель (Impact Score)	-	-	0,7682

При нормировании SAP, например:
0.0–0.25→Low,
0.26–0.50→Medium,
0.51–0.75→High,
0.75 → Very High.

Impact Score = 0.7982 → Impact = Very High.

4. Интеграция в SAP: Z-класс и бизнес-логика

Реализация производится через кастомный ABAP-класс (например, ZCL_PRIORITY_CALCULATOR), который вызывается при создании/обновлении инцидента. Основные методы:

- calculate_impact(iv_incident_id) – собирает атрибуты заявки, нормирует значения XiX_iXi, вызывает расчёт по AHP и возвращает Impact Score и категорию Impact.
- calculate_priority(iv_impact, iv_urgency) – сопоставляет Impact и Urgency по таблице customizing и записывает Priority.
- log_calculation(iv_incident_id, it_details) – сохраняет метаданные расчёта (матрицы, веса, Impact Score) в Z-таблицы для аудита и аналитики.
- fallback_strategy( ) – в случае недоступности данных возвращает стандартный запуск SAP Impact или заранее установленные веса.

Архитектурно Z-класс взаимодействует с SAP-таблицами инцидентов (CRMD_ORDER и др.), Z-таблицами с экспертными оценками и таблицами аналитики.

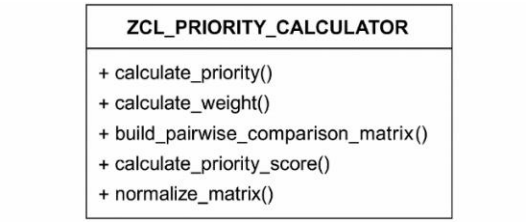


Рис. 2. UML-схема ключевых методов ZCL_PRIORITY_CALCULATOR

5. Модуль аналитики и оценка влияния на SLA

Для оценки пользы решения собирается статистика «до/после» внедрения:

- среднее время реакции, среднее время решения по приоритетам;
- процент нарушений SLA по приоритетам;
- доля перераспределённых приоритетов (сколько заявок получили другой приоритет после внедрения МАИ);
- качество классификации (по отклику бизнес-пользователей).

Ожидаемый эффект: снижение доли нарушений SLA за счёт более объективной приоритизации и корректного направленного распределения ресурсов.

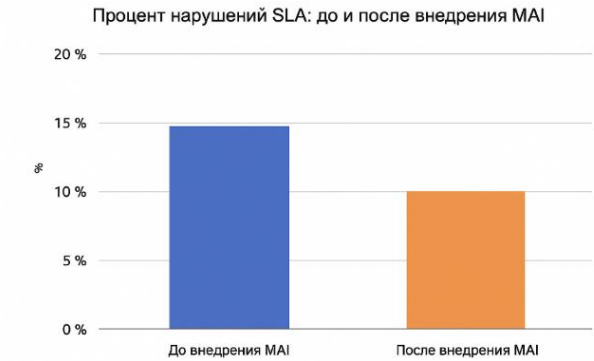


Рис. 3. График «Процент нарушений SLA: до и после внедрения MAI» – вставляется после анализа статистики

Заключение

В работе предложена и подробно описана методика замены стандартного компонента Impact в SAP Service Desk на интегральный показатель, рассчитанный методом анализа иерархий (МАИ). Представленная архитектура позволяет без глубокой модификации ядра SAP интегрировать новый алгоритм через Z-класс и сохранить стандартную логику Priority = f(Impact, Urgency). Приведённый численный пример демонстрирует, как формируются веса критериев, проверяется согласованность оценок и применяется результат к реальной заявке.

Преимущества подхода: формализация экспертных знаний, многофакторность оценки, воспроизводимость результатов, улучшение распределения ресурсов и потенциальное снижение нарушений SLA. Рекомендуется дополнительно провести пилотное внедрение, собрать статистику и, при необходимости, скорректировать набор критериев и экспертные оценки.

Список литературы

1. SAP Service Desk – Управление инцидентами [Электронный ресурс] : сайт. – Режим доступа: https://help.sap.com
2. ITIL v3 Incident Management Overview [Электронный ресурс] : сайт. – Режим доступа: https://wiki.en.it-processmaps.com
3. SAP Priority Calculation: Impact & Urgency Model [Электронный ресурс] : сайт. – Режим доступа: https://blogs.sap.com
4. Thomas L. Saaty. Decision Making with the Analytic Hierarchy Process [Электронный ресурс]: сайт. – Режим доступа: https://www.researchgate.net/publication/22882020

Разработка системы автоматизированного сопровождения психологического консультирования на основе IOS-приложения с функцией аналитики

С. С. Лавлинская, email: lavlinskaas@gmail.com
И.Ф. Астахова, email: astachova@list.ru
Е. А. Копытина, email: zhemkaterina@yandex.ru

Воронежский государственный университет

Аннотация. *В статье представлена реализация мобильного приложения для iOS, предназначенного для автоматизации рабочих процессов консультирующего психолога. Цель разработки — интеграция в единую цифровую среду функций управления клиентами, проведения стандартизированной психодиагностики, интеллектуального формирования терапевтических рекомендаций и аналитической оценки динамики состояния. Обоснованы методологический базис системы, основанный на доказательном подходе и использовании валидного диагностического инструментария, а также архитектурные и технологические решения (SwiftUI, Firebase, MVVM). Особое внимание уделено вопросам информационной безопасности и соответствия этическим стандартам психологической практики. Результатом работы является комплексно спроектированная система, способная стать основой для нового стандарта цифрового инструментария в психологическом консультировании.*

Ключевые слова: *автоматизация психологической практики, мобильное приложение, iOS, психодиагностика, рекомендательная система, облачные технологии, доказательный подход, аналитика данных, информационная безопасность.*

Введение

Цифровая трансформация охватывает различные сферы профессиональной деятельности, включая психологическое консультирование, как показано в [1-6]. Несмотря на рост числа онлайн-платформ для подбора специалистов, большинство психологов, особенно в частной практике или государственных учреждениях, по-прежнему сталкиваются с необходимостью ручного управления рабочими процессами. Это включает ведение бумажных или электронных журналов записей, проведение тестирования с использованием бумажных бланков, ручной подсчёт баллов, интерпретацию результатов,

© Лавлинская С. С., Астахова И. Ф., Копытина Е. А., 2026

хранение архивов и подбор терапевтических материалов. Подобный подход не только неэффективно расходует время специалиста (до 30% рабочего времени по оценкам экспертов), но и повышает риск ошибок, затрудняет стандартизацию диагностики и объективную оценку динамики состояния клиента.

Анализ существующих решений выявил их фрагментарность. Приложения для управления практикой (Calendly, SimplyBook.me) эффективно автоматизируют планирование, но лишены диагностических и терапевтических функций. Специализированные платформы для тестирования (Psychometrica) предлагают богатый арсенал методик, но, как правило, не интегрированы в рабочий цикл и не имеют полноценной мобильной адаптации. Рекомендательные системы и базы знаний предоставляют контент, но не персонализируют его под конкретный диагностический профиль. Таким образом, существует объективная потребность в создании целостного мобильного решения, интегрирующего административные, диагностические, терапевтические и аналитические функции в единый рабочий поток психолога.

1. Постановка задачи

Целью исследования является проектирование, разработка и обоснование архитектуры мобильной системы автоматизированного сопровождения психологического консультирования, выполняющей роль персонального цифрового ассистента специалиста.

Перечислим функциональные требования к приложению:

1. Управление практикой и клиентами: управление клиентской базой.

– Создание, просмотр, редактирование и архивирование профилей клиентов с хранением контактных данных, демографической информации и анамнеза.
– Календарь и планирование сессий: Функция записи на прием с возможностью создания, редактирования и отмены сессий. Интеграция с системными календарями (Apple Calendar) и настройка различных типов встреч.
– Система уведомлений: Автоматическая отправка напоминаний психологу и клиенту (с его согласия) о предстоящей сессии.

2. Психодиагностический модуль:

– Библиотека валидных тестов: Наличие встроенных стандартизированных методик (Шкала депрессии Бека, Шкала тревожности Спилбергера-Ханина, MBI) в адаптированном для мобильного использования формате.

– Проведение тестирования: Интуитивный пошаговый процесс предъявления вопросов теста клиенту непосредственно с устройства психолога.
– Автоматическая обработка результатов: Мгновенный подсчет сырых и стандартизированных баллов с автоматической интерпретацией согласно ключам методик.
– Хранение истории тестирования: Привязка всех результатов к профилю клиента и сохранение в хронологическом порядке для отслеживания динамики.
3. Интеллектуальный модуль рекомендаций:
– Формирование персонализированных рекомендаций: Автоматическая генерация списка терапевтических ресурсов (статьи, упражнения, литература) на основе результатов последнего теста.
– Ранжирование по релевантности: Сортировка предлагаемых материалов на основе алгоритмической оценки соответствия диагностическим маркерам клиента.
– Централизованная база ресурсов: Структурированное хранилище материалов с тегами, связанными с психологическими конструктами (тревожность, депрессия, стресс).
4. Аналитический модуль:
– Визуализация динамики состояния: Построение графиков изменения ключевых показателей по шкалам для выбранного клиента за любой временной период.
– Сравнительная агрегированная аналитика: Просмотр обобщенной статистики по группе клиентов в обезличенном виде для выявления общих тенденций и эффективности методик.
– Профессиональные заметки: Возможность добавлять текстовые комментарии к сессиям или результатам тестов.
5. Системные функции:
– Безопасная аутентификация: Вход в систему через защищенный механизм (Firebase Authentication).
– Работа в офлайн-режиме: Возможность просматривать данные, вводить информацию о сессиях и результатах тестов без подключения к интернету с последующей синхронизацией.
– Сквозная синхронизация данных: Автоматическое и надежное обновление информации на всех устройствах психолога.

2. Реализация проекта: методология, архитектура и ключевые модули

Система строится на принципах доказательной практики (evidence-based practice). В её ядро включены проверенные, валидные и надежные психодиагностические методики:

- Шкала депрессии Бека (BDI): для оценки выраженности депрессивной симптоматики.
- Шкала тревожности Спилбергера-Ханина: для дифференциации личностной и ситуативной тревожности.
- Опросник выгорания MBI: для диагностики профессионального выгорания по трехфакторной модели.
- Операционализация конструктов обеспечивает перевод сырых баллов в стандартизированные профили («умеренная депрессия»), что служит основой для алгоритмического формирования рекомендаций.

Клиентское приложение для iOS разрабатывается на языке Swift с использованием фреймворка SwiftUI, что обеспечивает современный, отзывчивый интерфейс и кроссплатформенную поддержку (iPhone, iPad). Архитектура MVVM (Model-View-ViewModel) применяется для четкого разделения логики данных, бизнес-логики и представления, что повышает тестируемость и сопровождаемость кода.

Серверная часть базируется на облачной платформе Google Firebase, что обеспечивает быстрый старт и масштабируемость, как показано в [7]. Используются следующие сервисы:

- Cloud Firestore (NoSQL): для гибкого и структурированного хранения данных о клиентах, сессиях, результатах тестов.
- Cloud Functions: для реализации серверной бизнес-логики (алгоритмы рекомендаций, сложные аналитические расчеты) в бессерверной модели.
- Firebase Authentication: для безопасной аутентификации и авторизации пользователей.

Общая архитектура системы является событийно-ориентированной (Event-Driven), где ключевые действия (завершение теста, добавление заметки) инициируют каскад событий для обработки данных и обновления интерфейсов.

Стоит отметить ключевые модули приложения:

- Модуль управления практикой: интегрированный календарь для планирования сессий, клиентская база с историей взаимодействий, система уведомлений.

– Модуль психодиагностики: цифровое проведение выбранных тестов с автоматическим подсчетом баллов, интерпретацией и сохранением результатов в профиль клиента.
– Интеллектуальный модуль рекомендаций: является системой поддержки принятия решений (Decision Support System). Алгоритм, основанный на взвешивании по методу TF-IDF (Term Frequency — Inverse Document Frequency), устанавливает связи между диагностическими маркерами клиента и базой терапевтических ресурсов (статьи, упражнения, литература), обеспечивая персонализированный ранжированный вывод рекомендаций.
– Модуль аналитики: предоставляет визуализацию динамики психологических показателей клиента в виде графиков. Включает функции сравнительного анализа (в агрегированном, обезличенном виде) для выявления групповых закономерностей и оценки эффективности методик.

Для формализации функциональных требований к системе была разработана диаграмма вариантов использования (рис. 1), которая показывает ключевые сценарии взаимодействия психолога с приложением.

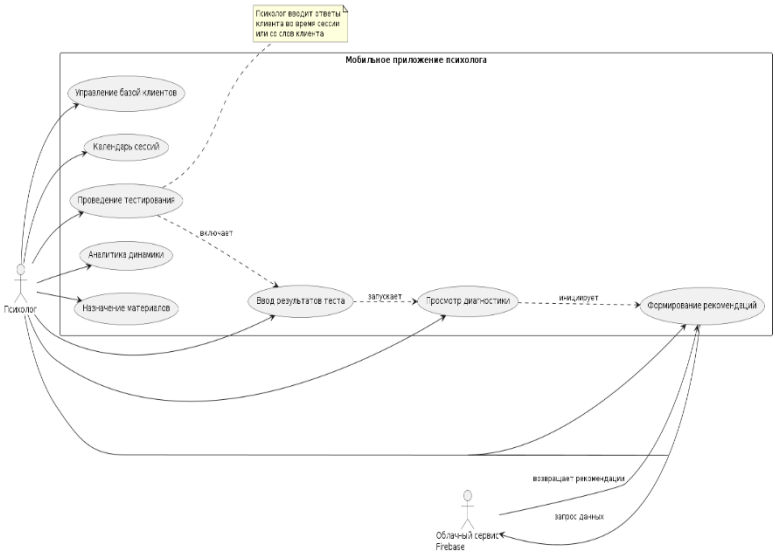


Рис. 1. Диаграмма вариантов использования мобильного приложения для психолога

Как видно из диаграммы, система поддерживает полный цикл работы специалиста: от управления клиентской базой и планирования сессий до проведения диагностики и анализа результатов.

Для обеспечения психолога инструментами отслеживания динамики состояния клиентов был разработан специализированный модуль аналитики (рис. 2).

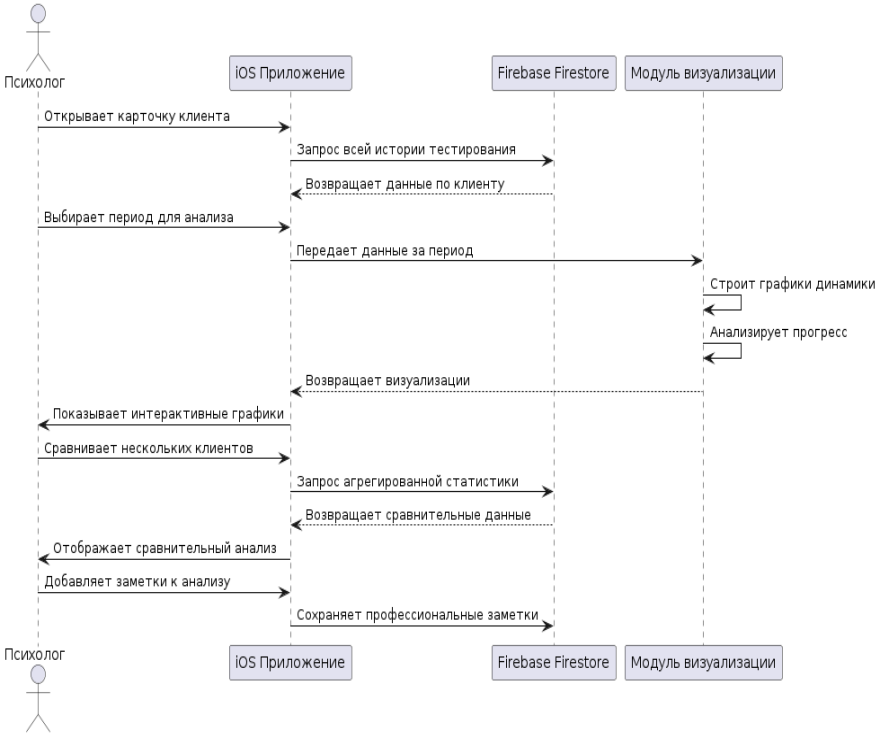


Рис. 2. Диаграмма последовательности процесса аналитики для психолога

Разработка ведется с учетом строгих этических норм и требований к защите персональных данных (ФЗ-152):

- Конфиденциальность: End-to-end шифрование, fine-grained доступ на уровне документов в Firestore.
- Профессиональный контроль: система — помощник, а не замена психологу. Окончательная интерпретация и решение остаются за специалистом.

– Прозрачность: психолог имеет возможность проследить логику формирования рекомендаций.
– Принцип «не навреди»: заложен в методологическую основу и алгоритмические решения.

Заключение

В ходе исследования был спроектирован и разработан комплексный прототип мобильной системы, направленной на автоматизацию рутинных задач психолога-консультанта. Предложенное решение отличается целостностью, объединяя в едином интерфейсе ранее разрозненные функции управления, диагностики, терапии и анализа. Методологическая строгость, обеспечиваемая использованием валидного инструментария и доказательного подхода, сочетается с современными технологическими решениями (SwiftUI, Firebase), что гарантирует как профессиональную состоятельность, так и техническую реализуемость проекта.

Внедрение подобной системы способно привести к значительной оптимизации рабочего времени специалиста, повышению стандартизации и качества услуг, а также формированию объективной доказательной базы эффективности терапевтических интервенций.

– В качестве направления дальнейшего исследования и развития стоит отметить:
– Расширение банка диагностических методик и библиотеки терапевтических ресурсов.
– Внедрение элементов предиктивной аналитики и машинного обучения для прогнозирования динамики состояния.
– Разработка клиентского модуля приложения для выполнения домашних заданий и самоотчетов.
– Интеграция с wearable-устройствами для получения объективных психофизиологических данных.
– Проведение пилотных исследований для количественной оценки эффективности системы в реальной практике.

Разработанный проект представляет собой значимый шаг в формировании нового цифрового стандарта для психологической практики, отвечающего вызовам современности.

Список литературы

1. Юргенсон, Е. А. Перспективы развития в психологии: новаторские подходы и инновации / Е. А. Юргенсон, С. А. Флейшман, А. А. Смирная // Мир человека : Материалы ежегодной Всероссийской конференции, Красноярск, 25 апреля 2024 года. – Красноярск: Федеральное государственное бюджетное образовательное учреждение

высшего образования Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева, 2024. – С. 144-150. – EDN EOWIIM.

2. Конычева, Ю. Д. Инновации в психологии труда в условиях цифровой трансформации / Ю. Д. Конычева // Проблемы теории и практики управления. – 2025. – № 7. – С. 113-122. – EDN UQCXAS.

3. Свидетельство о государственной регистрации программы для ЭВМ № 2025664443 Российская Федерация. Программа для автоматизации работы отдела интеллектуальной собственности : заявл. 20.05.2025 : опубл. 04.06.2025 / П. О. Филонов, Е. А. Копытина ; заявитель федеральное государственное бюджетное образовательное учреждение высшего образования «Воронежский государственный университет». – EDN HFQPRS.

4. Белоусова, А. Представление об инновациях и инновационных процессах в психологии и смежных науках / А. Белоусова // Вестник непрерывного образования. – 2016. – № 1. – С. 82-84. – EDN WDCHZZ.

5. Свидетельство о государственной регистрации программы для ЭВМ № 2025616922 Российская Федерация. Программа для автоматизации процесса документооборота медицинских организаций : заявл. 19.03.2025 : опубл. 21.03.2025 / А. А. Барышев, Е. А. Копытина ; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования "Воронежский государственный технический университет". – EDN RPSEGV.

6. Свидетельство о государственной регистрации программы для ЭВМ № 2024661693 Российская Федерация. Программа для проведения экспертизы работ проектных конкурсов : № 2024660532 : заявл. 13.05.2024 : опубл. 21.05.2024 / Я. В. Солодовникова, Е. А. Копытина ; заявитель федеральное государственное бюджетное образовательное учреждение высшего образования «Воронежский государственный университет». – EDN HDOJTC.

7. Firebase for Apple platforms [Электронный ресурс] // Firebase Documentation / Google. – URL: https://firebase.google.com/docs/ios/setup (дата обращения: 14.11.2025)

Математические модели анализа качества образовательного процесса на основе цепей Маркова

А. Е. Лоскутова, email: anyaa02@yandex.ru [1]
Н. А. Алейникова, email: balbashovan@mail.ru [2]

[1] Муниципальное бюджетное общеобразовательное учреждение «Лицей № 1»
[2] ФГБОУ ВО «Воронежский государственный университет», г. Воронеж, Россия

Аннотация. *В данной работе рассматриваются примеры применения моделей и алгоритмов скрытых марковских цепей в образовательном процессе. Аппарат скрытых марковских цепей разработан для марковских процессов, состояния которых непосредственно не наблюдаемы. Уровень усвоения разделов образовательной программы учащимся является качественной характеристикой. Его категории характеризуются множеством показателей, таких как отметки, посещаемость, активность учащегося на уроке, регулярность и качество выполненных им домашних работ и т.д. Модели скрытых марковских цепей позволяют по наблюдаемой последовательности показателей, например отметок, классифицировать последовательность состояний уровня усвоения материала.*

Ключевые слова: *скрытые марковские цепи, уровень усвоения разделов дисциплины, алгоритм Витерби, алгоритм прямого-обратного хода (Forward-Backward), алгоритм Баума-Велша (EM-алгоритм).*

Введение

Разработка математического аппарата для создания алгоритмов автоматического управления качеством образовательного процесса в школе является актуальной задачей. В данной работе рассматривается задача управления качеством усвоения разделов образовательной программы учащимися в школе, которая осуществляется, в том числе, на основе текущего и периодического контроля успеваемости. С помощью текущего контроля определяется степень достижения планируемых результатов общеобразовательной программы в процессе ее освоения учащимися – текущая успеваемость.

Периодический контроль направлен на определение промежуточной успеваемости и производится по определенному разделу учебного курса

© Лоскутова А. Е., Алейникова Н. А., 2026

в запланированные преподавателем или образовательной организацией моменты времени. Его цель – проверка качества усвоения части пройденного материала и корректировка в случае необходимости траектории успеваемости обучающегося.

Здесь необходимо сделать следующее замечание. Степень усвоения материала каждым учащимся, вообще говоря, непосредственно не наблюдаема. Она представляет собой качественную характеристику, измеренную в категориальной шкале. Возможные ее значения (состояния) зависят от способа соотнесения цели и результата обучения, обеспечения объективности оценки учебных достижений [1]. Разрабатываются критерии, описывающие соответствие между состояниями степени усвоения и различными характеристиками учащихся. Один из возможных способов косвенной оценки степени усвоения материала – с помощью выставления и анализа отметок в ходе текущего контроля. Обычно текущий контроль в школе осуществляется поурочно или по пройденным темам, причем порядок, форма и периодичность его проведения определяется учителем. Форма проведения текущего контроля письменная, устная или комбинированная. В то же время проведение текущего контроля в таких формах не всегда обеспечивает объективность оценки степени усвоения, так как в этом случае, текущий контроль проводится в виде выборочного контроля. Так же колеблемость и динамика в определенных пределах отметок учащегося необязательно означает смену состояний степени усвоения. Например, при отличном уровне усвоения знаний, учащийся мог быть невнимательным и получил плохую отметку. Или, наоборот, при удовлетворительном уровне усвоения, учащемуся повезло и ему достался вопрос, на который он знал ответ.

Задача состоит в определении наиболее вероятных состояний степени усвоения пройденного материала на основе наблюдаемых отметок учащегося с целью поддержки принятия решений учителем при мониторинге динамики процесса обучения конкретного ученика.

Для описания динамики состояний степени усвоения учебного материала в работе предлагается использовать модели скрытых марковских цепей (A Hidden Markov model, HMM). HMM были разработаны для формализации марковского процесса, когда не известно, в каком состоянии находится система (состояния скрыты) на каждом шаге, но каждое состояние может с некоторой вероятностью произвести событие, которое можно наблюдать [2-3].

В качестве скрытых состояний возьмем значения качественного показателя степени усвоения материала, измеренного в категориальной шкале. События, которые производятся скрытыми состояниями,

представляют собой отметки, полученные в ходе текущего контроля учащимся.

1. Основные понятия теории HMM

Дискретная марковская цепь задается множеством состояний, которые обозначим как

$$S = \left(S_i\right)_{i=\overline{1,n}}$$

и матрицей переходных вероятностей

$$M = \left(p_{ij}\right)_{i,\,j=\overline{1,n}},$$

каждый элемент которой соответствует условной вероятности перехода системы из состояния S_i в состояние S_j за один шаг, то есть

$$p_{ij} = P\left(S_j / S_i\right).$$

Вероятность перехода в каждое из состояний зависит только от текущего состояния вне зависимости от того, какими были предшествующие состояния.

Скрытая марковская цепь [3-6] представляет собой модификацию марковской цепи для случая, когда состояния системы непосредственно не наблюдаемы (скрыты). Вместо этого имеется некоторая последовательность или временной ряд наблюдаемых величин

$$Y = \left(y_1, ..., y_T\right),$$

в каждый момент времени принимающих одно значение из конечного множества

$$O = \left(o_1, ..., o_m\right),$$

называемого алфавитом.

Последовательность (4) порождается последовательностью скрытых состояний (1). Кроме матрицы (2) модель HMM содержит матрицу эмиссии

$$B = \left(b_{ij}\right)_{i=\overline{1,n};\,j=\overline{1,m}} = \left(b_i\left(o_j\right)\right)_{i=\overline{1,n};\,j=\overline{1,m}},$$

каждый элемент которой b_{ij} – вероятность наблюдения $o_j \in O$, при условии, что система находилась в состоянии S_i.

Модель HMM задается в виде кортежа

$$\mu = \left(S, O, M, B, \pi\right)$$

или более кратко

$$\mu = \left(M, B, \pi\right),$$

где $\pi = (\pi_i)_{i=\overline{1,n}}$ – начальное распределение вероятностей состояний (1).

Обозначим через x_t – скрытое состояние системы на шаге t из множества возможных состояний (1). Последовательность $X = (x_t)_{t=\overline{1,T}}$ не известна.

При моделировании процесса усвоения знаний учащимся – уровень знаний в каждый момент времени непосредственно не наблюдаем и его градации можно описывать с помощью скрытых состояний (1). Наблюдаемые величины (4) соответствуют отметкам, которые выставляются в ходе текущего контроля.

Перечислим задачи контроля качества усвоения знаний учащегося, которые можно решить с помощью HMM. Первая задача состоит в определении по наблюдаемым отметкам наиболее вероятной последовательности состояний уровня усвоения знаний за определенный промежуток времени, например, за один период периодического контроля. Эту задачу можно решить с помощью алгоритма Витерби.

Вторая задача заключается в определении вероятностей всех состояний на каждом шаге текущего или периодического контроля. По сути это задача классификации состояний на каждом шаге. В теории HMM это производится с помощью алгоритма прямого-обратного хода (Forward-Backward). Решение данной задачи позволит выполнить более глубокий анализ процесса обучения учащегося.

Третья задача – в переоценке матриц M и B с учетом наблюдаемых отметок. Это можно сделать с помощью алгоритма Баума-Велша (EM-алгоритма). С помощью EM-алгоритма можно уточнить параметры модели (8) для каждого ученика с учетом его индивидуальных особенностей.

Одна из проблем при использовании HMM в рамках рассматриваемых задач – начальная инициализация параметров модели (8). Для задания матриц (2) и (6) можно использовать экспертный подход. Например, учитель указывает вероятности перехода из состояния в состояние для каждого ученика. С помощью EM-алгоритма эти вероятности потом уточняются. Матрицу (6) можно получить на основе статистического подхода, оценив элементы этой матрицы относительными частотами той или иной отметки при определенном уровне знаний. Вектор $\pi = (\pi_i)_{i=\overline{1,n}}$ учитель может задать по результатам входного контроля. Далее рассмотрим каждую задачу по отдельности. С помощью алгоритма Витерби по последовательности отметок (4) определяется наиболее вероятная последовательность состояний.

Алгоритм Витерби реализует задачу динамического программирования [7].

Обозначим через

$$\delta_j(t) = \max_{x_1,\ldots,x_{t-1}} P(x_1,\ldots,x_{t-1}, y_1,\ldots,y_{t-1}, x_t = j, y_t)$$

максимальную вероятность, которая будет достигнута, при условии, что на шаге t наблюдается состояние S_j.

В алгоритме Витерби сначала осуществляется прямой проход, при котором на каждом шаге максимизируется накопленная вероятность $P(x_1,\ldots,x_t, y_1,\ldots,y_t, x_{t+1} = j, y_{t+1})$,то есть находится

$$\delta_j(t+1) = \max_{x_{t+1}} \delta_j(t) a_{ij} b_j(y_{t+1}).$$

Аргумент, при котором был достигнут максимум (10), запоминается в функции

$$\psi_j(t+1) = \arg\max_{x_{t+1}} \delta_j(t) a_{ij} b_j(y_{t+1}).$$

Затем осуществляется обратный проход, при котором восстанавливается наиболее вероятная последовательность состояний

$$\hat{x}_T = \arg\max_{x_T} \delta_j(T),$$

$$\hat{x}_t = \psi_{\hat{x}_{t+1}}(t+1).$$

Рассмотрим теперь алгоритм прямого-обратного хода [8] с помощью которого можно классифицировать скрытые состояния, то есть определять распределения их вероятностей. Обозначим через

$$\alpha_t(i) = P(y_1, y_2,\ldots, y_t, x_t = S_i/\mu) = P(Y, x_t = S_i/\mu)$$

вероятность при принятой модели μ увидеть наблюдаемые значения $y_1, y_2,\ldots, y_t$ до момента времени t включительно и при этом на шаге t – состояние S_i (α-вероятности).

α-вероятности определяются по рекуррентной формуле

$$\alpha_1(i) = \pi_i b_i(y_1),$$

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{n} \alpha_t(i) p_{ij}\right] \cdot b_j(y_{t+1}).$$

Далее обозначим через

$$\beta_t(i) = P(y_{t+1}, y_{t+2},\ldots, y_T / x_t = S_i, \mu)$$

вероятность (β -вероятность) увидеть после шага t последовательность $y_{t+1}, y_{t+2}, ..., y_T$, в условиях модели μ, а также при условии состояния S_i на шаге t. $\beta_t(i)$ рассчитываются по формулам

$$\beta_T(i) = 1,$$

$$\beta_t(i) = \sum_{j=1}^{n} p_{ij} b_j(y_{t+1}) \beta_{t+1}(j).$$

Алгоритм прямого-обратного хода состоит из следующих этапов:

1. Прямой проход заключается в вычислении α -вероятностей по формулам (15)-(16).
2. Вычисляется вероятность последовательности (4) в условиях μ

$$P(Y/\mu) = P(y_1, ..., y_T/\mu) = \sum_{i=1}^{n} \alpha_T(i).$$

3. Обратный проход заключается в вычислении β -вероятностей по формулам (18)-(19).
4. Определяются вероятности нахождения в состоянии S_i на каждом шаге

$$\gamma_t(i) = P(x_t = i/Y, \mu) = \frac{\alpha_t(i)\beta_t(i)}{P(Y/\mu)},$$

а также вероятности перехода из состояния в состояние

$$\xi_t(i, j) = P(x_t = i, x_{t+1} = j/Y, \mu) = \frac{\alpha_t(i)\, p_{ij} b_j(y_{t+1}) \beta_{t+1}(j)}{P(Y/\mu)}.$$

Если вероятностное распределение состояний уровня усвоения учебного материала на каком-то шаге начинает существенно отличаться от предыдущего, например, повышаются вероятности неудовлетворительного уровня, то это говорит о вероятном негативном изменении в процессе обучения учащегося и требует внимания учителя.

Последний алгоритм, рассматриваемый в работе, – алгоритм Баума-Велша [9] позволяет обучить модель (8) на основе наблюдаемой последовательности отметок (4). Возможно, что матрицы в (8) были идентифицированы учителем не совсем верно для данного учащегося и расходятся с наблюдаемыми отметками. Алгоритм Баума-Велша основан на оптимизации функции правдоподобия и включает следующие шаги:

Е-шаг: Для текущих $\mu^{(k)} = \left(M^{(k)}, B^{(k)}, \pi^{(k)}\right)$ (где начальные параметры модели $\mu^{(0)} = \left(M^{(0)}, B^{(0)}, \pi^{(0)}\right)$ заданы учителем) вычисляются: $\alpha_t(i)$, $\beta_t(i)$ с помощью алгоритма Forward-Backward; $\gamma_t(i)$ – по формуле (21); $\xi_t(i, j)$ – по формуле (22)

М-шаг: пересчет параметров

$$\pi_i^{(k+1)} = \gamma_1(i),$$

$$a_{ij}^{(k+1)} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)},$$

$$b_i^{(k+1)}(k) = \frac{\sum_{t:\, y_t = k} \gamma_t(i)}{\sum_{t=1}^{T-1} \gamma_t(i)}.$$

Критерий остановки:

$$\left|\log P\left(Y/\mu^{(k+1)}\right) - \log P\left(Y/\mu^{(k)}\right)\right| < \varepsilon$$

2. Применение алгоритмов HMM к анализу уровня усвоения учебного материала

Рассмотрим пример применения алгоритма прямого-обратного хода для классификации уровня усвоения знаний на этапах периодического контроля.

Предположим, имеется четыре состояния уровня успеваемости учащегося $S = (\text{Плохое, Удовлетворительное, Хорошее, Отличное})$.

Оценивание происходит по 4-х балльной шкале. Отметки являются элементами множества $O = (2, 3, 4, 5)$. Предположим, что матрица переходных вероятностей для конкретного учащегося была оценена следующим образом

$$M = \begin{pmatrix} 0.2 & 0.5 & 0.29 & 0.01 \\ 0.2 & 0.5 & 0.2 & 0.1 \\ 0.05 & 0.25 & 0.4 & 0.3 \\ 0.01 & 0.29 & 0.35 & 0.35 \end{pmatrix}.$$

Матрица эмиссии

$$B = \begin{pmatrix} 0.3 & 0.4 & 0.25 & 0.05 \\ 0.3 & 0.5 & 0.15 & 0.05 \\ 0.05 & 0.2 & 0.6 & 0.15 \\ 0.01 & 0.09 & 0.4 & 0.5 \end{pmatrix}.$$

Начальное распределение вероятностей состояний $\pi = (0, 0, 0.2, 0.8)$. Наблюдаемая последовательность отметок $Y = (5,5,4,5,4,4,2,5,4,5,4,3)$.

Рассчитаем с помощью алгоритма прямого-обратного хода вероятности состояний в каждый момент времени. Вероятности состояний, рассчитанные по формуле (21) приведены в таблице ниже.

Таблица

Определение вероятностей состояний

t	y_t	$\gamma_t(S_1)$	$\gamma_t(S_2)$	$\gamma_t(S_3)$	$\gamma_t(S_4)$	S
1	5	0	0	0	1	S_4
2	5	0,00001	0,000002	0,001063	0,998934	S_4
3	4	0,000012	0,000045	0,023059	0,976884	S_4
4	5	0,000002	0,000009	0,009658	0,990332	S_4
5	4	0,000005	0,000021	0,028819	0,971155	S_4
6	4	0,000004	0,000019	0,044279	0,955698	S_4
7	2	0,000553	0,001805	0,031585	0,966057	S_4
8	5	0,000039	0,000168	0,003548	0,996245	S_4
9	4	0,000013	0,000053	0,035005	0,964929	S_4
10	5	0,000007	0,000027	0,031862	0,968105	S_4
11	4	0,000003	0,000011	0,038617	0,961369	S_4
12	3	0,000015	0,00004	0,721374	0,278537	S_3

Вероятность последовательности

$$P(Y/\mu) \approx 2.79 \times 10^{-11}.$$

Логарифм функции правдоподобия

$$\log P(Y/\mu) \approx -24.42.$$

Следовательно, модель, заданная учителем, плохо объясняет данные наблюдения. На протяжении почти всей последовательности (кроме

последнего наблюдения) наиболее вероятным является состояние S_4 с вероятностью > 95%.

Применим алгоритм Баума-Велша для корректировки параметров модели. В результате получим скорректированные параметры модели:

матрица переходных вероятностей

$$M^* = \begin{pmatrix} 0.1425 & 0.5219 & 0.2849 & 0.0507 \\ 0.1874 & 0.5052 & 0.1974 & 0.11 \\ 0.0428 & 0.2478 & 0.4061 & 0.3033 \\ 0.0008 & 0.0975 & 0.3526 & 0.5491 \end{pmatrix};$$

матрица эмиссии

$$B^* = \begin{pmatrix} 0.3012 & 0.3985 & 0.2503 & 0.0500 \\ 0.2987 & 0.5012 & 0.1498 & 0.0503 \\ 0.0485 & 0.1992 & 0.6015 & 0.1508 \\ 0.0098 & 0.0893 & 0.4012 & 0.4997 \end{pmatrix};$$

вектор начального распределения вероятностей

$$\pi^* = (0,0003,\ 0,0002,\ 0,1995,\ 0,8000).$$

При этом значение логарифма правдоподобия

$$\log P\left(Y/\mu^*\right) \approx -24.37\ .$$

изменилось не значительно, как и все параметры модели.

На основании матрицы переходных вероятностей (27) можно вычислить предельные вероятности состояний

$$\pi_f^* = (0.0487,\ 0.1423,\ 0.2958,\ 0,5132)\ .$$

При этом предельные вероятности состояний при исходной матрице были следующими

$$\pi_f = (0.1170,\ 0.3830,\ 0.3006,\ 0,1994)\ .$$

То есть скорректированная модель дает более оптимистичный прогноз. Вероятность того, что учащийся большую часть времени в дальнейшем будет находиться в состоянии S_4 гораздо больше, чем при исходной модели (исходных ожиданиях учителя).

Заключение

Таким образом, в работе показаны примеры применения математических моделей и алгоритмов скрытых марковских цепей к анализу уровня усвоения разделов образовательной программы учащимся. Рассмотренные алгоритмы позволяют по последовательности

отметок, полученных учащимся в ходе текущего контроля, классифицировать последовательность состояний уровня усвоения им пройденного материала, определить наиболее вероятную последовательность состояний, а также скорректировать матрицу переходных вероятностей и эмиссии, отражающих первоначальное представление учителя о возможностях ученика. При этом учитывается, что краткосрочные колебания отметок не обязательно означают смену состояний уровня усвоения. Приведенные алгоритмы позволяют учителю более объективно оценивать процесс обучения учащегося.

Список литературы

1. Шентурк, С. Об уровне усвоения учебного материала / С. Шентурк // Теория и практика современной науки. – 2016. – № 5 (11). – С. 1061-1066.
2. Кельберт, М. Я. Вероятность и статистика в примерах и задачах. Т. II: Марковские цепи как отправная точка теории случайных процессов и их приложения / М. Я. Кельберт, Ю. М. Сухов. – Москва: МЦНМО, 2009. – 551 с.
3. Rabiner, L. R. A tutorial on hidden Markov models and selected applications in speech recognition / L. R. Rabiner // Proceedings of the IEEE. – 1989. – Vol. 77. – no. 2. – P. 257-286.
4. Elliott, R. J. Hidden Markov Models: Estimation and Control / R. J. Elliott, L. Aggoun, J. B. Moore. – New York : Springer-Verlag, 1994. – 357 p.
5. McCallum, A. Introduction to Natural Language Processing : Lecture 7: Hidden Markov Models. The Baum-Welch Algorithm. – 2004. – 21 p.
6. Mor, B. A systematic review of hidden markov models and their applications / B. Mor, S. Garhwal, A. Kumar // Archives of Computational Methods in Engineering. – 2021. – Vol. 28. – no. 3. – P. 1429-1448. DOI: 10.1007/s11831-020-09422-4
7. Viterbi, A. J. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm / A. J. Viterbi // IEEE Transactions on Information Theory. – 1967. – Vol. 13. – no. 2. – P. 260-269.
8. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains / Baum L. E. [et al.] // The Annals of Mathematical Statistics. 1970. Vol. 41, no. 1. P. 164-171.
9. Baum, L. E. An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology / L. E. Baum // Bulletin of the American Mathematical Society. – 1967. – Vol. 73. – no. 3. – P. 360-363.

Сравнительный анализ решения задачи нечеткой регрессии с использованием классических подходов и W-алгебры

М. Г. Матвеев, email: mgmatveev@yandex.ru [1]
Е. Н. Свиридова, email: sviridova.e.n@mail.ru [2]
Н. С. Богданова, email: bogdanova2077@yandex.ru [1]

[1] Воронежский государственный университет
[2] Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н. Е. Жуковского и Ю. А. Гагарина» (г. Воронеж)

Аннотация. *В представлении и обработке данных все более широкое применение находят нечеткие числа и методы нечетких вычислений. Это делает актуальной задачу выбора наиболее эффективных методов реализации нечетких вычислений и построения нечетких зависимостей. В статье приводится сравнительный анализ известных классических методов и метода, использующего W-алгебру, для решения задачи нечеткой линейной регрессии как с четкими, так и с нечеткими параметрами. Представлены основные свойства W-алгебры нечетких треугольных LR-чисел и методы решения задачи построения линейной регрессии с использованием W-алгебры. На примерах решения задачи построения линейной регрессии различными методами показаны преимущества метода с использованием W-алгебры, что обосновывает эффективность и перспективность его использования.*

Ключевые слова: *нечеткая линейная регрессия, методы оценивания четких и нечетких параметров, W-алгебра, нечеткие метрики качества моделирования.*

Введение

Задачи регрессии занимают важное место в теории и практике принятия решений, обеспечивая оценку зависимости одних величин от других на основе анализа соответствующих статистических данных. Обычно методы классического регрессионного анализа предполагают, что эти данные по своей природе точные (четкие), но их наблюдаемые значения содержат аддитивную случайную ошибку, как правило, с известным распределением. Однако, во многих прикладных областях, например, экономике, медицине, социологии, анализируемые данные могут иметь размытый, интервальный характер, обусловленный экспертными оценками, не имеющими точных значений. В этих случаях

© Матвеев М. Г., Свиридова Е. Н., Богданова Н. С., 2026

оценка зависимости исследуемых величин производится с использованием нечеткой регрессии [1], например, нечеткой линейной регрессии (НЛР)

$$\tilde{y} = \tilde{a}_0 + \tilde{a}_1\tilde{x}_1 + ... + \tilde{a}_n\tilde{x}_n,$$

где $\tilde{x}_i$, $i = 1;...;n$ – нечеткие регрессоры (факторы) модели; $\tilde{a}_i$, $i = 0;1;...;n$ – нечеткие параметры модели.

Целью нашей работы является сравнительный анализ статистических методов оценки четких и нечетких параметров модели (1), основанных на известных классических подходах, с методами, основанными на использовании W-алгебры. Для сравнения были использованы результаты исследований, опубликованные в работах [2–8]. Сравнительный анализ проводился по общепринятым для регрессии метрикам, адаптированным к нечетким числам. Кроме того, было проведено сравнение классических методов и методов W-алгебры по вычислительной сложности алгоритмов.

1. Используемые модели и методы

Разнообразие классических методов решения задачи нечеткой регрессии не позволяет привести соответствующие модели и методы. С ними можно ознакомиться по приведенным публикациям. В этом разделе даны основные сведения об операциях W-алгебры и приводятся модели, методы, метрики сравнения нечетких чисел и метрики качества, адаптированные к нечетким числам.

В W-алгебре множество нечетких чисел представимо в форме треугольных или трапецеидальных LR-чисел, модифицированных в соответствии со следующими определениями.

Определение 1. Автономной компонентой (левой или правой) треугольного w-числа называется параметрическая запись функции принадлежности нечеткого числа в виде

$$x^L(\alpha) = x^L(0) + \alpha(x^L(1) - x^L(0)),$$
$$x^R(\alpha) = x^R(0) + \alpha(x^R(1) - x^R(0)),$$

где $x^L(1) = x^R(1) = x(1) \in \mathbb{R}$, $x^L(0), x^R(0) \in \mathbb{R}$, $\alpha \in [0,1]$.

Определение 2. w-числом (или нечётким двухкомпонентным треугольным числом) называется нечёткое число $\tilde{x}$, у которого левая и правая границы задаются функциями $x^L(\alpha)$, $x^R(\alpha)$, причем $x^L(1) = x^R(1) = x(1)$. Общая точка $x(1)$ является модой треугольного w-

числа, значения $x^L(0)$, $x^R(0)$ определяют соответственно левый и правый коэффициенты нечёткости (параметры разброса) числа $\tilde{x}$.

Пусть $x(\alpha) = x(0) + \alpha(x(1) - x(0))$, $y(\alpha) = y(0) + \alpha(y(1) - y(0))$ — компоненты (левые или правые) двух w-чисел. Арифметические операции в W-алгебре определяются независимо покомпонентно.

Сложение:

$$x(\alpha) \oplus y(\alpha) = x(0) + y(0) + \alpha\Big(\big(x(1) + y(1)\big) - \big(x(0) + y(0)\big)\Big).$$

Вычитание:

$$x(\alpha) \ominus y(\alpha) = x(0) - y(0) + \alpha\Big(\big(x(1) - y(1)\big) - \big(x(0) - y(0)\big)\Big).$$

Умножение:

$$x(\alpha) \odot y(\alpha) = x(0) \cdot y(0) + \alpha\Big(\big(x(1) \cdot y(1)\big) - \big(x(0) \cdot y(0)\big)\Big).$$

Модель нечеткого регрессионного анализа имеет вид

$$\tilde{y} = \tilde{a}_0 \oplus \tilde{a}_1 \odot \tilde{x}_1 \oplus ... \oplus \tilde{a}_n \odot \tilde{x}_n,$$

где $\tilde{y}$ – нечёткое значение зависимой переменной; $\tilde{x}_1, ..., \tilde{x}_n$ – нечёткие значения регрессоров; $\tilde{a}_0, \tilde{a}_1, ..., \tilde{a}_n$ – нечёткие в общем случае параметры модели, подлежащие оценке; $\oplus$ и $\odot$ – это операторы сложения и умножения (3), (5) соответственно, определенные в рамках W-алгебры.

Для получения оценок НЛР с четкими параметрами [9] предлагается использовать метрический подход с минимизацией расстояния между наблюдаемыми (или заданными экспертами) и модельными нечеткими значениями. В W-алгебре нечеткие числа являются выпуклыми компактами, поэтому расстояние между наблюдаемым значением нечеткого числа $\tilde{y}(\alpha)$ и соответствующим модельным значением $\tilde{y}_m(\alpha)$ можно ввести как агрегированное расстояние между соответствующими α-срезами, вычисленное отдельно для левых и правых компонент [9]:

$$d^L(\tilde{y}, \tilde{y}_m) = \left| \int_0^1 \left(y^L(\alpha) - y_m^L(\alpha) \right) d\alpha \right|,$$

$$d^R(\tilde{y}, \tilde{y}_m) = \left| \int_0^1 \left(y^R(\alpha) - y_m^R(\alpha) \right) d\alpha \right|.$$

Нетрудно убедиться, что все необходимые аксиомы меры в данном случае соблюдены.

Полное расстояние определяется как сумма расстояний (7) и (8):

$$d(\tilde{y}, \tilde{y}_m) = d^L(\tilde{y}, \tilde{y}_m) + d^R(\tilde{y}, \tilde{y}_m),$$

Утверждение [9]. Численное значение агрегированного расстояния $d(\tilde{y}, \tilde{y}_m)$ между α-срезами двух w-чисел эквивалентно расстоянию между этими числами при $\alpha = 0,5$.

Тогда критерий минимизации расстояния будет представлен следующей функцией

$$I\left(a_0; a_1; ...; a_n\right) = \sum_{i=1}^{n}\left(\tilde{y}_i\left(0,5\right) - \tilde{y}_i^m\left(0,5\right)\right)^2 \to \min .$$

Критерий (10) в случае НЛР четкими параметрами будет иметь вид

$$I(a_0; a_1; ...; a_n) = \sum_{i=1}^{n}\left(y_i^L(0,5) - a_0 - a_1 x_{1i}^L(0,5) - ... - a_n x_{ni}^L(0,5)\right)^2 +$$

$$+\sum_{i=1}^{n}\left(y_i^R(0,5) - a_0 - a_1 x_{1i}^R(0,5) - ... - a_n x_{ni}^R(0,5)\right)^2 \to \min,$$

а задача оценки параметров $a_0, a_1, ..., a_n$ сводится к методу наименьших квадратов.

Для нечетких параметров, не ограничивая общности, мы будем предполагать, что нечёткие входные данные $\tilde{x}_{ij}$, нечёткие выходные данные $\tilde{y}_i$, и нечёткие параметры $\tilde{a}_p$, $p = 0,1,...,n$, $j = 1,...,n$, $i = 1,...,k$ заданы как левые или правые компоненты w-числа.

В рамках W-алгебры представим параметры и данные в виде:

$$\tilde{y}_i = y_0^i + \left(y_1^i - y_0^i\right)\alpha, \quad \tilde{x}_{ij} = x_0^{i,j} + \left(x_1^{i,j} - x_0^{i,j}\right)\alpha,$$

$$\tilde{a}_p = a_0^p + \left(a_1^p - a_0^p\right)\alpha,$$

где $\alpha \in [0,1]$, $y_0^i = y_i\left(0\right)$, $y_1^i = y_i\left(1\right)$, $x_0^{i,j} = x_{ij}\left(0\right)$, $x_1^{i,j} = x_{ij}\left(1\right)$, $a_0^p = a_p\left(0\right)$, $a_1^p = a_p\left(1\right)$ – вещественные числа, $i = 1,...,k$, $j = 1,...,n$, $p = 0,1,...,n$.

Тогда критерий для оценки нечетких параметров можно записать в виде

$$P = P\left(\tilde{a}_0, \tilde{a}_1, ..., \tilde{a}_n\right) = \sum_{i=1}^{k}\left(\tilde{y}_i \ominus \tilde{y}_i^m\right)^2 = \sum_{i=1}^{k} P_i,$$

$$P_i = \left(y_0^i - \mathbf{a_0}^T \mathbf{x_0^i}\right)^2 +$$

$$+\alpha\left[\left(y_1^i - \mathbf{a}_1^T \mathbf{x}_1^{\mathbf{i}}\right)^2 - \left(y_0^i - \mathbf{a_0}^T \mathbf{x_0^i}\right)^2\right],$$

где $\mathbf{a_0} = \left(a_0^0, a_0^1, ..., a_0^n\right)^T$, $\mathbf{a}_1 = \left(a_1^0, a_1^1, ..., a_1^n\right)^T$, $\mathbf{x_0^i} = \left(1, x_0^{i,1}, ..., x_0^{i,n}\right)^T$, $\mathbf{x}_1^{\mathbf{i}} = \left(1, x_1^{i,1}, ..., x_1^{i,n}\right)^T$.

Выражение (12) можно переписать в виде:

$$P_i = (1-\alpha)\left(y_0^i - \mathbf{a_0}^T \mathbf{x_0^i}\right)^2 + \alpha\left(y_1^i - \mathbf{a}_1^T \mathbf{x}_1^{\mathbf{i}}\right)^2.$$

Поскольку $\alpha \in [0,1]$, оба коэффициента $1-\alpha$ и α неотрицательны, минимум P достигается, когда минимизируется каждая из сумм по отдельности: $\sum_{i=1}^{k}\left(y_0^i - \mathbf{a_0}^T \mathbf{x_0^i}\right)^2$ минимизируется по $\mathbf{a_0}$, $\sum_{i=1}^{k}\left(y_1^i - \mathbf{a}_1^T \mathbf{x}_1^{\mathbf{i}}\right)^2$ минимизируется по $\mathbf{a}_1$. Решения в матричной форме имеют вид:

$$\mathbf{a_0} = \left(\mathbf{X_0}^T \mathbf{X_0}\right)^{-1} \mathbf{X_0}^T \mathbf{y_0},\ \mathbf{a}_1 = \left(\mathbf{X}_1^T \mathbf{X}_1\right)^{-1} \mathbf{X}_1^T \mathbf{y}_1,$$

где $\mathbf{X}_0$ — матрица размера $k \times (n+1)$, в которой строка i равна $\left(1, x_0^{i,1}, ..., x_0^{i,n}\right)$; $\mathbf{X}_1$ — матрица размера $k \times (n+1)$, в которой строка i равна $\left(1, x_1^{i,1}, ..., x_1^{i,n}\right)$; $\mathbf{y}_0 = \left(y_0^1, ..., y_0^k\right)^T$, $\mathbf{y}_1 = \left(y_1^1, ..., y_1^k\right)^T$.

2. Метрики качества моделирования

Качество полученных оценок параметров напрямую определить невозможно. Но можно получить косвенные оценки качества найденных параметров посредством метрик качества полученной модели НЛР. Некоторые такие нечеткие метрики, в том числе аналоги статистических метрик, известны. Например, в работах [10,11] используются нечеткие расширения коэффициента детерминации (R^2), MSE и MAE, а также средняя ширина доверительной полосы.

Нечеткий R^2 как аналог классического коэффициента детерминации можно определить через меру расстояния между наблюдаемыми и модельными нечеткими значениями

$$R^2_{fuzzy} = 1 - \frac{\sum_{i=1}^{n} d(\tilde{y}_i, \tilde{y}_i^m)^2}{\sum_{i=1}^{n} d(\tilde{y}_i, \tilde{y}_{cpeд}^m)^2},$$

где $\tilde{y}_{cpeд}^m$ — среднее нечеткое значение, центр данных; $d(\cdot,\cdot)$ – метрика расстояния между двумя нечеткими числами.

Средняя квадратичная ошибка и средняя абсолютная ошибка определяются соответственно как MSE = $\frac{1}{n}\sum_{i=1}^{n}(\tilde{y}_i - \tilde{y}_i^{M})^2$, MAE = $\frac{1}{n}\sum_{i=1}^{n}\left|\tilde{y}_i - \tilde{y}_i^{M}\right|$, $\tilde{y}_i, \tilde{y}_i^{M}$ - наблюдаемые и модельные (предсказанные) нечеткие выходные значения.

Средняя ширина доверительной полосы предсказания, численно равная среднему сумм левых и правых спредов модельных результатов, рассчитывается по формуле AWCB=$\frac{1}{n}\sum_{i=1}^{n}|\, y_i^{RM}(\alpha=0) - y_i^{LM}(\alpha=0)\,|$. Убывание этой метрики интерпретируется как снижение нечеткости результата моделирования.

3. Сравнительный анализ различных методов нечеткой регрессии по метрикам качества моделирования

Сравнительный анализ проводился для двух видов нечеткой регрессии: с четкими и нечеткими параметрами и был выполнен на двух синтетических выборках, состав которых отражен в табл. 1 и 2.

Таблица 1

Синтетический набор данных [2]

$(x^L; x^m; x^R)$	$(y^L; y; y^R)$
(1;1;1)	(6,2;8;9,8)
(2;2;2)	(4,2;6,4;8,6)
(3;3;3)	(6,9;9,5;12,1)
(4;4;4)	(10,9;13,5;16,1)
(5;5;5)	(10,6;13;15,4)

Таблица 2

Синтетический набор данных [5,6]

$(x^L; x^m; x^R)$	$(y^L; y; y^R)$
(1,5;2;2,5)	(3,5;4;4,5)
(3;3,5;4)	(5;5,5; 6)
(4,5;5,5;6,5)	(6,5;7,5;8,5)
(6,5;7;7,5)	(6;6,5;7)

Окончание табл.2

$(x^L; x^m; x^R)$	$(y^L; y; y^R)$
(8;8,5;9)	(8;8,5;9)
(9,5;10,5;11,5)	(7;8;9)
(10,5;11;11,5)	(10;10,5;11)
(12;12,5;13)	(9;9,5;10)

Оценка четких параметров НЛР проводилась методом на основе W-алгебры с последующим сравнением с методами [5,3]. Результаты вычислений для оценки четких параметров НЛР приведены в табл. 3.

Таблица 3

Оцененные параметры моделей и метрики качества моделирования по данным таблицы 2 для методов [5,3] и W-алгебры

Метод	**Параметр** a_0	**Параметр** a_1	**Метрики качества**	
			MSE	R^2_{fuzzy}
Kao, Chyu [5]	3,565	0,522	2,5728	0,793
Bargiela [3]	3,4467	0,5360	2,5850	0,791
Метод с W-алгеброй	3,5724	0,5193	2,5715	0,794

Сравнительный анализ с известными методами, такими как подходы Kao и Chyu [5], а также Bargiela и соавторов [3], показал, что W-алгебра демонстрирует сопоставимое или лучшее качество моделей при существенно меньших вычислительных затратах. Это делает её перспективным инструментом для анализа данных в условиях неопределенности.

Оценка нечетких параметров НЛР проводилась методом на основе W-алгебры с последующим сравнением с методами [2,4,6,7,8]. Результаты вычислений для оценки нечетких параметров НЛР приведены в табл. 4, 5.

Таблица 4

Оцененные параметры моделей и метрики качества моделирования по данным таблицы 1 для методов [2,4] и W-алгебры

Метод	**Параметр** $\tilde{a}_0$	**Параметр** $\tilde{a}_1$	**Метрики качества**		
			MAE	**MSE**	**AWCB**
Tanaka [2]	(0; 3,85; 7,7)	(2,1; 2,1; 2,1)	2,78	5,51	7,00

Окончание табл.4

Метод	**Параметр** $\tilde{a}_0$	**Параметр** $\tilde{a}_1$	**Метрики качества**		
			MAE	**MSE**	**AWCB**
Savic, Pedrycz [4]	(0,92; 4,95; 8,98)	(1,64; 1,71; 1,78)	2,73	5,55	8,48
Метод с W-алгеброй			2,44	3,69	4,64

Таблица 5

Оцененные параметры моделей и метрики качества моделирования по данным таблицы 2 для методов [6,7,8] и W-алгебры

Метод	**Параметр**	**Параметр** $\tilde{a}_1$	**Метрики качества**		
			MAE	**MSE**	**AWCB**
Sakawa, Yano [6]	(2,94; 3,37; 3,79)	(0,45; 0,56; 0,67)	1,58	1,93	3,23
Hojati [7]	(3,00; 3,41; 3,82)	(0,50; 0,52; 0,54)	1,46	1,37	1,77
Вельдяксов, Шведов [8]	(3,27; 3,57; 3,87)	0,52	1,46	1,29	1,25
Метод с W-алгеброй	(3,28; 3,57; 3,84)	(0,52; 0,52; 0,52)	1,46	1,29	1,25

В таблицах 4 и 5, как и при оценке четких параметров, видно, что применение W-алгебры обеспечивает сопоставимую или повышенную точность моделей при существенном сокращении вычислительных ресурсов, что также определяет перспективность метода, основанного на использовании W-алгебры.

Заключение

В рамках исследования продемонстрированы ключевые преимущества W-алгебры при построении моделей нечёткой линейной регрессии с нечёткими параметрами.

Упрощение вычислительных процессов. Подход на основе W-алгебры трансформирует задачу нечёткой регрессии с нечеткими параметрами в задачи обычной линейной регрессии, что радикально снижает вычислительную нагрузку по сравнению с техниками, оперирующими множественными α-срезами. Данное преимущество становится особенно значимым при обработке многомерных данных.

Расширенные возможности работы с данными. Метод не накладывает ограничений на симметричность или неотрицательность разбросов нечётких чисел, что существенно увеличивает его адаптивность к разн α ородным данным.

Повышенная надёжность результатов. В отличие от -срезов, склонных к генерации неустойчивых решений, подход обеспечивает робастность результатов через использование классических уравнений линейной регрессии.

Практическая реализуемость. Алгоритм легко адаптируется к существующим вычислительным платформам, что упрощает его интеграцию в реальные системы.

Таким образом, W-алгебра демонстрирует себя как перспективный инструмент для нечёткого регрессионного анализа, сочетающий методологическую простоту с вычислительной эффективностью и устойчивостью результатов. Перспективными направлениями для дальнейших исследований являются адаптация метода для нелинейных зависимостей и работа со сложными структурами нечётких данных.

Список литературы

1. Zadeh, L. A. Fuzzy Sets / L. A. Zadeh // Information and Control. – 1965. – Vol. 8, № 3. – P. 338–353.

2. Tanaka, H. Possibilistic Linear Regression Analysis for Fuzzy Data / H. Tanaka, I. Hayashi, J. Watada // European Journal of Operational Research. – 1989. – Vol. 40. – P. 389–396.

3. Bargiela, A. Multiple regression with fuzzy data / A. Bargiela, W. Pedrycz, T. Nakashima // Fuzzy Sets and Systems. – 2007. – Vol. 158. – P. 2169–2188.

4. Savic, D. A. Evaluation of fuzzy linear regression models / D. A. Savic, W. Pedrycz // Fuzzy Sets and Systems. – 1991. – Vol. 39. – P. 51–63.

5. Kao, C. Least-Squares Estimates in Fuzzy Regression Analysis / C. Kao, C. L. Chyu // European Journal of Operational Research. – 2003. – Vol. 148. – P. 426–435.

6. Sakawa, M. Multiobjective fuzzy linear regression analysis for fuzzy input-output data / M. Sakawa, H. Yano // Fuzzy Sets and Systems. – 1992. – Vol. 47. – P. 173–181.

7. Hojati, M. A simple method for computation of fuzzy linear regression / M. Hojati, C. R. Bector, K. Smimou // European Journal of Operational Research. – 2005. – Vol. 166, № 1. – P. 172–184.

8. Вельдяксов, В. Н. О методе наименьших квадратов с нечеткими данными / В. Н. Вельдяксов, А. С. Шведов // Экономический журнал ВШЭ. – 2014. – № 2. – С. 328–344.

9. Матвеев, М. Г. W-алгебра в задачах нечеткой линейной регрессии с четкими параметрами / М. Г. Матвеев, Е. Н. Свиридова // Вестник Воронежского государственного университета. Серия «Системный анализ и информационные технологии». – 2025. – №4. – С. 132–144.

10. Kelkinnama, M. Fuzzy least absolute regression / M. Kelkinnama, S. M. Taheri // 4th International IEEE Conference Intelligent Systems. – 2008. – Vol. 2. – P. 3433–3457.

11. Chachi, J. Multiple fuzzy regression model for fuzzy input-output data / J. Chachi, S. M. Taheri // Iranian Journal of Fuzzy Systems. – 2016. – Vol. 13, № 4. – P. 63–78.

Подходы к восстановлению информации с использованием методов нечисловой статистики

В. В. Меньших, email: menshikh@list.ru
В. Р. Гришаева, email: vrkvaleria.2002@mail.ru

Воронежский институт МВД России

Аннотация. *В данной работе рассматриваются переход от количественных к качественным показателям, а также использование методов нечисловой статистики для работы с неполными статистическими данными.*

Ключевые слова: *нечисловая статистика, методы нечисловой статистики, правовая статистика, принятие решений, ранжирование, неполные статистические данные.*

Введение

В современной прикладной статистике при анализе данных возникают ситуации, когда исходная информация о наблюдаемых объектах или явлениях является неполной, неоднородной или структурно ограниченной. Поэтому актуальной является задача обработки таких данных [1].

Примером такой неполной статистической информации являются данные правовой статистики о субъектах Российской Федерации с наименьшими и наибольшими темпами прироста различных показателей преступности, которые доступны не в полном объеме (табл. 1) [2, 3].

Таблица 1

Данные правовой статистики о состоянии преступности на территории некоторых субъектах Центрального федерального округа Российской Федерации за январь 2025 года

	п1, %	п2, %	п3, %	п4, %	п5, %	п6, %	п7, %
Владимирская область	-78,3		164			-32,9	
Калужская область	500			-52,2		-26,5	
Костромская область	-73,1		93	200	75		

© Меньших В. В., Гришаева В. Р., 2026

Окончание табл. 1

	п1, %	п2, %	п3, %	п4, %	п5, %	п6, %	п7, %
Липецкая область	-76,9			-50	-61,5		30,5
Орловская область	850	-35,9		72,7		48,6	
Ярославская область	100			107,7		41,9	28,8

где п1 – темпы прироста числа зарегистрированных преступлений, связанных с незаконным оборотом оружия; п2 – темпы прироста числа преступления, совершенные с использованием информационно-телекоммуникационных технологий или в сфере компьютерной информации; п3 – темпы прироста числа преступления, связанные с незаконным оборотом наркотических средств, психотропных веществ или их аналогов, сильнодействующих веществ, новых потенциально опасных психоактивных веществ; п4 – темпы прироста числа преступные посягательства и материальной ущерб от преступлений; п5 – темпы прироста числа преступления, совершенные на дорогах и трассах вне населенных пунктов; п6 – темпы прироста числа преступления, совершенные на улицах, площадях, в парках, скверах; п7 – темпы прироста числа преступления, совершенные в общественных местах.

Значения показателей преступности для остальных субъектов являются промежуточными.

Вместе с тем следует отметить, что для эффективного принятия решений и статистических оценок необходимо иметь полную статистическую панель для всех объектов исследования [4].

2. Переход от количественных к качественным показателям

Как следует из [4-6] для принятия эффективных решений необходимо иметь данные правовой статистики в полном объеме, а в случае, если статистическая панель является неполной, то необходимо осуществить переход от количественных показателей к качественным.

Количественные показатели – это числовые величины, имеющие измеримый физический, экономический или статистический смысл. Они используются для ответа на вопросы, связанные с частотой, интенсивностью, объёмом или пропорцией (например: «сколько?», «как часто?», «какой процент?»). Однако интерпретация таких данных требует учёта контекста, поскольку числовое значение не отражает субъективного восприятия явления [7].

Примером количественных показателей являются данные правовой статистики, а именно статистические данные о состоянии преступности

на территории субъектов Центрального федерального округа Российской Федерации за январь 2025 года (табл. 1).

Качественные показатели фиксируют нечисловые характеристики объекта исследования и получаются посредством экспертных оценок, наблюдений, интервью или анализа текстов. Они отвечают на вопросы, направленные на выявление причин, механизмов, свойств или субъективного опыта (например: «почему?», «как именно?», «какой?»). Такая формулировка отражает субъективную оценку, недоступную прямому количественному измерению [7].

Вопрос перехода от количественных к качественным показателям при неполной статистической панели затрагивает важную методологическую проблему на стыке когнитивной психологии и статистической теории.

Из работы [8] было определено, что способность человека воспринимать информацию в пределах 7 ± 2. Данные пределы называются числом Миллера, которое предполагает собой знаменитую эмпирическую оценку ограниченной ёмкости кратковременной (рабочей) памяти. В данном случае используются понятия из теории информации (в частности, биты), но подчёркивается, что человеческое восприятие не сводится к простому подсчёту битов, так как важна структура и смысл.

Переход от количественных к качественным показателям обусловлен не столько отказом от больших значений измерений, сколько определением границ применимости количественных методов и стремлением к более обширному пониманию неполной исходной статистической информации.

Во многих случаях можно сказать, что переход от количественных показателей к качественным – это не отказ от строгости и точности исходных статистических данных, а восстановление неполноты информации или расширение понимания того, что считать «данными» и «информацией». Он необходим в случаях, когда:

1. при работе со статистической панелью ряд данных полностью или частично неизвестен;
2. значения ряда показателей представлены большими значениями, работа с которыми не всегда является удобной и эффективной;
3. цель перехода является не измерение, а возможность восстановления информации при неполной статистической панели данных.

Таким образом, можно сделать вывод, что при переходе от количественных к качественным показателям необходимо все данные привести не только к наименьшим значениям возможными методами, но

и систематизировать их на определенные классы, объединив по какому-то признаку.

3. Методы нечисловой статистики

Часто затрагиваемый вопрос при работе с нечисловой статистикой – это неполнота данных, суть которой заключается в том, что не для всех пар исследуемых объектов имеется информация об их сравнении или упорядочивании [9]. Поэтому могут возникнуть ряд проблем при работе с такой статистикой, которые можно решить с помощью следующий методов [9]:

1. метод ранжирования;
2. метод парных сравнений;
3. метод экспертных оценок;
4. метод бинарных отношений.

Рассмотрим некоторые из данных методов. Начнем с метода ранжирования. Ранжирование — это упорядочение объектов в соответствии с убыванием (или возрастанием) степени выраженности некоторого признака [10, 11]. При этом сам признак может быть не измерен количественно; достаточно лишь возможности сравнения любых двух объектов по этому признаку. Таким образом, ранжирование происходит путем реализации измерений в порядковой шкале [12].

Рассмотрим ранжирование нечисловой статистики на примере неполных данных о некоторых видах преступления, совершенных на территории некоторых субъектов Центрального федерального округа Российской Федерации в январе 2025 года (табл. 1).

Исходя из табл. 1 невозможно определить, в каком из субъектов Российской Федерации криминогенная обстановка лучше или хуже. Именно поэтому необходимо произвести ранжирование данных по всей территории Российской Федерации для полного понимания криминогенной обстановки [13, 14].

Для этого необходимо произвести деление на ранги (1):

$$k = \frac{\beta - \alpha}{s},$$

где k – шаг ранга, β – нижняя граница, α – верхняя граница, s – число рангов.

Учитывая, что всех градаций может быть 7 ± 2 [9], можно использовать ранжирование и получить относительно точные качественные оценки на основе неполной статистической информации. Используя формулу (1), получим следующие 7 рангов и их границы для января 2025 года (табл. 2).

Таблица 2

Границы рангов

k	135,7	43,8	44,0	42,9	53,1	19,0	4,2
β	-100	-73	-58,3	-100	-71,4	-80	3,4
верхняя граница 1 ранга	35,7	-29,2	-14,3	-57,1	-18,3	-61,0	7,6
верхняя граница 2 ранга	171,4	14,5	29,8	-14,3	34,7	-41,9	11,7
верхняя граница 3 ранга	307,1	58,3	73,8	28,6	87,8	-22,9	15,9
верхняя граница 4 ранга	442,9	102,0	117,9	71,4	140,8	-3,8	20,1
верхняя граница 5 ранга	578,6	145,8	161,9	114,3	193,9	15,2	24,3
верхняя граница 6 ранга	714,3	189,5	206,0	157,1	246,9	34,3	28,4
α	850,0	233,3	250,0	200,0	300,0	53,3	32,6

Определим принадлежность каждого субъекта Центрального федерального органа Российской Федерации к своему рангу (табл. 3).

Заметим, что регионы с наименьшими темпами прироста зарегистрированных преступлений имеют чаще всего 1-3 ранги, а регионы с наибольшими – 5-7. Следовательно, можно сделать предположение, что остальные субъекты Российской Федерации относятся к 4 рангу.

Таким образом с помощью метода ранжирования данных мы осуществили переход от количественных показателей к качественным, а также смогли восстановить всю информацию статистической панели для дальнейшей работы (табл. 4).

Таблица 3

Ранги областей Центрального федерального округа Российской Федерации

	п1	п2	п3	п4	п5	п6	п7
Владимирская область	1 ранг		6 ранг			3 ранг	
Калужская область	5 ранг			2 ранг		3 ранг	
Костромская область	1 ранг		4 ранг	7 ранг	3 ранг		
Липецкая область	1 ранг			2 ранг	1 ранг		7 ранг
Орловская область	7 ранг	1 ранг		5 ранг		1 ранг	
Ярославская область	2 ранг			5 ранг		7 ранг	7 ранг

Таблица 4

Восстановленная информация о рангах всех областей Центрального федерального округа Российской Федерации

	п1	п2	п3	п4	п5	п6	п7
Владимирская область	1 ранг	4 ранг	6 ранг	4 ранг	4 ранг	3 ранг	4 ранг
Калужская область	5 ранг	4 ранг	4 ранг	2 ранг	4 ранг	3 ранг	4 ранг
Костромская область	1 ранг	4 ранг	4 ранг	7 ранг	3 ранг	4 ранг	4 ранг
Липецкая область	1 ранг	4 ранг	4 ранг	2 ранг	1 ранг	4 ранг	7 ранг

Окончание табл. 4

	п1	п2	п3	п4	п5	п6	п7
Орловская область	7 ранг	1 ранг	4 ранг	5 ранг	4 ранг	1 ранг	4 ранг
Ярославская область	2 ранг	4 ранг	4 ранг	5 ранг	4 ранг	7 ранг	7 ранг

Перейдем к рассмотрению следующего метода нечисловой статистики для работы с неполными статистическими данными. Метод парных сравнений – это метод, позволяющий последовательно сравнить все элементы статистической панели [10]. При неполноте статистических данных прямо рассматривается ситуация, когда эксперт не может (или не хочет) сравнить некоторые пары объектов, следовательно, возникает неполное бинарное отношение.

Рассмотрим на том же примере (табл. 1). Произведем попарное сравнение Владимирской и Костромской областей. Заметим, что значения показателя п1 больше в Костромской области, а п3 – в Владимирской

области. Значения показателей п4-п6 вообще не предоставляется сделать возможным, так как из-за неполноты статистической панели нельзя реализовать сравнение данных показателей. Следовательно, данный метод применим не ко всем типам неполных статистических панелей, и с помощью метода попарных сравнений невозможно осуществить полное восстановление информации при переходе от количественных к качественным показателям.

Следующий рассмотренный метод экспертных оценок предполагает, что ряд независимых экспертов дает оценку по определенным показателям [10]. Данный метод является субъективным, поскольку эксперты опираются не только на методы, но и на свое мнение. В качестве оценки показателей эксперт может выбрать сам метод работы с нечисловой статистикой. Поэтому метод экспертных оценок может в себя включать как метод ранжирования, так и метод попарных сравнений.

Заключение

Таким образом, использование методов нечисловой статистики, особенно метода ранжирования, позволяет восстановить информацию статистической панели, осуществить сравнение исследуемых объектов и повысить эффективность принятия решений в условиях дефицита количественных данных. Это особенно ценно для правоохранительных органов, где своевременная оценка криминогенной ситуации напрямую влияет на эффективность профилактики и борьбы с преступностью.

Список литературы

1. Морозова, В. О. Существующие подходы к описанию данных правовой статистики / В. О. Морозова, В. Р. Кучеряева // Охрана, безопасность, связь. – 2025. – № 10-3. – С. 20-26.

2. Портал правовой статистики : официальный сайт. – Москва. – Обновляется в течение суток. – URL: http://crimestat.ru/ (дата обращения: 13.01.2026). – Текст : электронный.

3. Сведения о состоянии преступности в Российской Федерации за 2012–2025 годы: аналитический отчет ФКУ ГИАЦ МВД России : официальный сайт. – Москва. – Обновляется в течение суток. – URL : https://xn--b1aew.xn--p1ai/reports/item/21933965/ (дата обращения: 13.01.2026). – Текст : электронный.

4. Морозова, В. О. Модели и алгоритмы оценки и прогнозирования показателей, используемых при осуществлении аналитической работы сотрудниками правоохранительных органов : диссертация на соискание ученой степени кандидата технических наук / Морозова Валерия Олеговна. – Воронеж, 2023. – 194 с.

5. Меньших, В. В. Метод декомпозиции статистических данных при прогнозировании показателей преступности / В. В. Меньших, В. О. Морозова, В. Р. Кучеряева // Криминологический журнал. – 2025. – № 2. – С. 251-259.

6. Меньших, В. В. Сравнительный анализ статистических методов исследования и прогнозирования данных правовой статистики (на примере данных о преступлениях в сфере информационных технологий) / В. В. Меньших, В. О. Морозова, В. Р. Кучеряева // Вестник Воронежского института МВД России. – 2024. – № 4. – С. 42-51.

7. Айвазян, С. А. Прикладная статистика и основы эконометрики / С. А. Айвазян, В. С. Мхитарян. – Москва : Юнити, 1998. – 1006 с.

8. Миллер, Дж. А. Магическое число семь плюс или минус два. О некоторых пределах нашей способности перерабатывать информацию / Дж. А. Миллер. – Москва : Инженерная психология. – 1964. – С. 564-581.

9. Орлов, А. И. Нечисловая статистика / А. И. Орлов. – Москва : МЗ – Пресс, 2004. – 513 с.

10. Орлов, А. И. Искусственный интеллект : нечисловая статистика : учебник / А. И. Орлов. – Москва : Ай Пи Ар Медиа, 2022. – 446 с.

11. Ларичев, О. И. Теория и методы принятия решений, а также Хроника событий в Волшебных Странах : Учебник / О. И. Ларичев. – Москва : Логос, 2000. – 296 с.

12. Большев, Л. Н. Таблицы математической статистики / Л. Н. Большев, Н. В. Смирнов. – Москва : Наука, 1983. – 416 с.

13. Амелина, А. В. Модель и численный метод учета латентных значений показателей правовой статистики при оценке криминогенной ситуации / А. В. Амелина, Т. В. Меньших // Вестник Воронежского института МВД России. – 2025. – № 3. – С. 56-63.

14. Меньших, А. В. Моделирование частичной неопределенности и неполноты данных при принятии управленческих решений / А. В. Меньших, Т. В. Меньших // Вестник Воронежского института МВД России. – 2023. – № 2. – С. 132-137.

Задачи распределения ресурсов элементов мультиагентных систем в условиях «мягкого» управления

В. В. Меньших, email: menshikh@list.ru
А. Н. Никитенко, email: nikitenkoan27@mail.ru

Воронежский институт МВД России

Аннотация. *В данной работе рассматривается задача распределения ресурсов для агентов мультиагентных систем при их функционировании в условиях «мягкого» управления, предполагающего возможность частичного расходования ресурсов одних агентов в интересах выполнения задач других агентов. Введены понятия функций индивидуальной и коллективной полезностей функционирования агентов. Приведён пример поиска параметров «мягкого» управления, приводящий к максимизации функции полезности мультиагентной системы в целом.*

Ключевые слова: *мультиагентные системы, «мягкое» управление, распределение ресурсов, индивидуальная функция полезности функционирования агента, коллективная функция полезности функционирования агента.*

Введение

Современная парадигма развития мультиагентных систем (МАС) [1-3] характеризуется переходом от анализа поведения изолированных субъектов к исследованию сложных процессов их коллективного взаимодействия [4-6].

В МАС эффективность функционирования системы определяется не только индивидуальными показателями, но и качеством возникающих связей между агентами, их способностью к взаимопомощи и синергии. При рассмотрении отдельных агентов системы их функционирование направлено на оптимизацию функции полезности решения назначенных для выполнения этим агентам задач, которую далее будем называть индивидуальной функцией полезности. При исследовании системы в целом следует учесть вклад других агентов в полезность функционирования данного агента, которую далее будем называть коллективной функцией полезности.

Таким образом, возникает необходимость решения задачи распределения ресурсов агентов МАС так, чтобы обеспечить и выполнение собственных задач этих агентов, и их участие в решении задач других агентов.

© Меньших В. В., Никитенко А. Н., 2026

Для решения данной задачи представляется перспективным использование концепции «мягкого» управления [7-9], предполагающей отказ от прямого директивного распределения ресурсов в пользу создания такой информационной и стимулирующей среды, в которой агенты в процессе самоорганизации [4, 5] приходят к распределению ресурсов, максимизирующему их коллективную полезность.

1. Концепция «мягкого» управления

При функционировании агентов МАС в условиях «мягкого» управления существует возможность оказания агентами помощи при решении задач системы. «Мягкое» управление предполагает, что отдельные агенты, так называемые доноры, могут в определённые моменты времени частично либо полностью (в форме эксплуатации) функционировать с отдачей собственного ресурса, а тем самым в некоторых случаях со снижением своей полезности при решении отдельных задач других агентов — реципиентов. При этом один и тот же агент в рамках функционирования МАС может выступать как в роли донора при решении одних задач, так и в роли реципиента при решении других задач. Следовательно, при функционировании в рамках принципа «мягкого» управления может возрастать как индивидуальная полезность функционирования отдельных агентов, так и полезность функционирования МАС в целом.

В современной теории интеллектуальных систем, в том числе МАС, принято выделять три основных направления, различающиеся способом организации взаимодействия агентов и степенью централизации управления: распределенный искусственный интеллект (РИИ), децентрализованный искусственный интеллект (ДИИ) и системы искусственной жизни [10, 11]. В рамках научных интересов авторов в данном исследовании будут рассмотрены только первые два направления. При организации РИИ особое внимание уделяется глобальным задачам системы, то есть агенты решают задачи в интересах полезности функционирования всей системы. При ДИИ агент действует рационально, исходя из своих собственных интересов. Целью исследования в рамках «мягкого» управления является поиск компромиссного решения между ДИИ И РИИ, то есть функционированием агентов в рамках глобальных целей и локальных целей агентов.

2. Описание модели

Обратимся к формальному описанию исследуемой модели.

Пусть имеется некая среда, в которой функционирует МАС $M = \left\{m_1, ..., m_{|M|}\right\}$, агенты которой решают задачи из множества задач $W = \left\{w^1, ..., w^{|W|}\right\}$. При этом не все задачи, которые потенциально может выполнить агент m_i, могут быть ему назначены на рассматриваемом этапе функционирования МАС.

Введём обозначения:

β_i — важность агента m_i;

$$f_i^k = \begin{cases} 1, \text{ если задача } w^k \text{ может быть выполнена агентом } m_i \\ 0, \text{ если иначе} \end{cases};$$

$$g_i^k = \begin{cases} 1, \text{ если задача } w^k \text{ должна быть выполнена агентом } m_i \\ 0, \text{ если иначе} \end{cases}.$$

Примем, что

$\forall i, j, i \neq j \;\; g_i^k \cdot g_j^k = 0$, т. е. множества назначенных на выполнение агентам задач не пересекаются;

$\exists i, j, i \neq j \;\; f_i^k \cdot g_j^k \neq 0$, т. е. найдутся агенты, которые потенциально могут выполнить задачи, назначенные другим агентам.

Каждый агент имеет для выполнения задач ресурс r_i, который он может потратить как на выполнение назначенных ему задач, так и задач в рамках помощи другим агентам.

Обозначим:

r_i^k — ресурс, выделяемый агентом m_i для решения задачи w^k;

$r_i^{k\,\min}$ — минимальный ресурс, который может быть выделен агенту m_i на выполнения задачи w^k;

$r_i^{k\,\max}$ — максимальный ресурс, который может быть выделен агенту m_i на выполнения задачи w^k;

$q_i^k(r_i^k)$ —индивидуальная полезность выполнения задачи w^k агентом m_i при выделении ресурса r_i^k;

α_i^k — степень влияния выполнения задачи w^k на полезность функционирования агента m_i;

ρ_i — доля ресурса, оставшаяся после выделения минимальных объёмов ресурса для решения назначенных агенту-донору задач:

$$\rho_i = r_i - \sum_{k=1}^{|W|} \left(r_i^{k\,\min} \cdot g_i^k \right).$$

Введём ограничение на выделяемый ресурс:

$$r_i^{k\,\min} \le r_i^k \le r_i^{k\,\max}.$$

В рамках «мягкого» управления могут быть определены параметры $\nu_{i,j}^k$, составляющие трёхмерную матрицу

$$\mathrm{N} = \left(\nu_{i,j}^k \right)_{i,\,j=\overline{1,|M|}}^{k=\overline{1,|W|}}$$

и характеризующие долю участия агента-донора m_i в решении задачи w^k агента-реципиента m_j. Данный параметр численно определяет долю ресурса, затрачиваемого агентом-донором m_i на решение этой задачи.

При этом должны выполняться следующие ограничения:

$$\nu_{i,j}^k \in \left[\nu_{i,j}^{k\,\min}; \nu_{i,j}^{k\,\max} \right] \subseteq [0;1].$$

Перейдём к определению «коллективной» полезности функционирования агентов в условиях «мягкого» управления.

Вклад агента m_i в полезность функционирования агента m_j при решении задачи w^k определяется как

$$\alpha_j^k \cdot q_i^k (\nu_{i,j}^k \cdot \rho_i) \cdot f_i^k \cdot g_j^k.$$

Вклад всех агентов МАС в полезность функционирования агента m_j при решении задачи w^k определяется как

$$l_j^k \left(R^k, \mathrm{N} \right) = \sum_{i=1}^{|M|} \left(\alpha_j^k \cdot q_i^k (\nu_{i,j}^k \cdot \rho_i) \cdot f_i^k \cdot g_j^k \right) = \alpha_j^k \cdot g_j^k \sum_{i=1}^{|M|} \left(q_i^k (\nu_{i,j}^k \cdot \rho_i) \cdot f_i^k \right)$$

где $R^k = \left(r_1^k, r_2^k, ..., r_{|M|}^k \right)$.

Тогда коллективная полезность функционирования агента m_j при решении всех задач является суммой полезности, возникающей в результате самостоятельного выполнения назначенных ему задач при выделении минимальных объёмов ресурса, и полезности, возникающей в результате выполнение назначенных ему задач как самостоятельно, так и другими агентами в рамках реализации концепции «мягкого» управления, при распределении оставшейся части ресурса ρ :

$$L_j\left(R,\mathrm{N}\right)=\sum_{k=1}^{|W|}l_j^k\left(R^k\right)+\sum_{k=1}^{|W|}\alpha_j^k\cdot g_j^k\cdot q_j^k\left(r_j^{k\min}\right).$$

Таким образом, необходимо найти:

$$\left(R^*,\mathrm{N}^*\right)=Arg\max_{\mathrm{N},R}\sum_{j=1}^{|M|}\left(\beta_j\cdot L_j\left(R,\mathrm{N}\right)\right).$$

Необходимо осуществить поиск параметров $\nu_{i,j}^k$, обеспечивающих оптимальное распределение ресурсов агентов с целью повышения полезностей функционирования агентов.

3. Численный пример

Рассмотрим численный пример. Пусть на некоторой территории функционирует МАС, состоящая из двух агентов-БПЛА $M=\left\{m_1,m_2\right\}$, решающих задачи обеспечения безопасности при возникновении чрезвычайной ситуации [12, 13]. Множество задач агентов содержит две задачи — $W=\left\{w^1,w^2\right\}$.

Обозначим матрицы

– потенциально выполнимых задач агентами МАС

$$F=\begin{array}{c}\\ m_1\\ m_2\end{array}\begin{array}{c}\begin{array}{cc}w^1 & w^2\end{array}\\ \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\end{array};$$

– назначенных на выполнение задач агентам МАС

$$G=\begin{array}{c}\\ m_1\\ m_2\end{array}\begin{array}{c}\begin{array}{cc}w^1 & w^2\end{array}\\ \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\end{array}.$$

Из матриц F и G следует, что агент m_1 имеет возможность выполнить задачу w^2, выделив оставшуюся после выполнения собственной задачи часть ресурса, т. е. может возникнуть ситуация, представленная на рисунке.

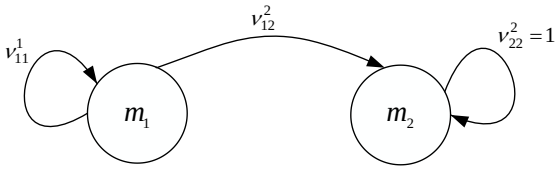


Рисунок. Граф взаимодействия агентов

Исходные данные заданы следующим образом:

Таблица

Исходные данные

m_1	m_2
$r_1 = 10$;	$r_2 = 5$;
$\alpha_1^1 = 1$;	$\alpha_2^1 = 1$;
$r_1^{1\min} = 4$;	$r_2^{2\min} = 2$;
$r_1^{1\max} = 10$;	$r_2^{2\max} = 5$;
$\rho_1 = 6$;	$\rho_2 = 3$;
$q_1^1 = 4r$;	$q_2^2 = 12r$;
$q_1^2 = 6r$;	$\beta_2 = 1$.
$\beta_1 = 1$.	

Тогда:

1. $l_1^1 = q_1^1 \cdot \left(\nu_{1,1}^1 \cdot 6\right) = 24 \cdot \nu_{1,1}^1$;
2. $l_2^2 = q_1^2 \cdot \left(\nu_{1,2}^2 \cdot 6\right) + q_2^2 \cdot \left(\nu_{2,2}^2 \cdot 3\right) = 36 \cdot \nu_{1,2}^2 + 36 \cdot \nu_{2,2}^2$.

При решении задачи (3), (1)-(2) $\left(R^*, \mathrm{N}^*\right) = Arg \max_{\mathrm{N},R} \left(\beta_1 \cdot L_1 + \beta_2 \cdot L_2\right)$ необходимо найти:

$$\begin{cases} 24 \cdot \nu_{1,1}^1 + 36 \cdot \nu_{1,2}^2 + 36 \cdot \nu_{2,2}^2 \to \max \\ \nu_{1,1}^1 + \nu_{1,2}^2 = 1 \\ \nu_{2,2}^2 = 1 \end{cases}.$$

Результатом решения задачи являются следующие значения:
$\nu_{1,1}^1 = 0$, $\nu_{1,2}^2 = 1$, $r_1^1 = 4$, $r_1^2 = 6$ и $r_2^2 = 5$.

Заключение

В данной работе рассмотрена задача распределения ресурсов в мультиагентных системах в условиях «мягкого» управления. Основное внимание уделено проблеме согласования индивидуальных интересов агентов с коллективными целями системы, что особенно актуально в контексте современных исследований по децентрализованному и распределённому искусственному интеллекту. Ключевым результатом работы является введение и формализация понятий индивидуальной и коллективной полезности функционирования агентов, что позволяет

количественно оценивать эффективность как отдельных агентов, так и системы в целом. Предложенная модель «мягкого» управления обеспечивает гибкий механизм распределения ресурсов между агентами-донорами и агентами-реципиентами, способствуя достижению синергетического эффекта. Таким образом, использование концепции «мягкого» управления открывает новые возможности для проектирования адаптивных, устойчивых и эффективных мультиагентных систем, способных функционировать условиях меняющейся среды. Дальнейшие исследования на расширение модели для более сложных сценариев взаимодействия агентов.

Список литературы

1. Рассел, С. Искусственный интеллект: современный подход : в 2 т. – Т. 1 : Решение проблем: знания и рассуждения / С. Рассел, П. Норвиг ; пер. с англ. – 4-е изд. – Санкт-Петербург : Диалектика, 2021. – 704 с.
2. Wooldridge, M. An Introduction to Multiagent Systems / M. Wooldridge. – Chichester : John Wiley & Sons, Ltd, 2002. – 492 p.
3. Словохотов, Ю. Л. Распределенный интеллект мультиагентных систем. Ч. 2. Коллективный интеллект социальных систем / Ю. Л. Словохотов, Д. А. Новиков // Проблемы управления. – 2023. – № 6. – С. 3–21.
4. Городецкий, В. И. Самоорганизация и многоагентные системы. I. Модели многоагентной самоорганизации / В. И. Городецкий // Известия РАН. Теория и системы управления. – 2012. – № 2. – С. 92–120.
5. Городецкий, В. И. Самоорганизация и многоагентные системы. II. Приложения и технология разработки / В. И. Городецкий // Известия РАН. Теория и системы управления. – 2012. – № 3. – С. 102–123.
6. Gorodetsky, V. Multi agent Peer-to-Peer Intrusion Detection / V. Gorodetsky, O. Karsaev, V. Samoylov, S. Serebryakov // Communications in Computer and Information Science. – 2007. – Vol. 1. – P. 260–271.
7. Меньших, В. В. Модель оценки эффективности функционирования мультиагентной системы в условиях использования «мягкого» управления / В. В. Меньших, А. Н. Никитенко // Вестник Воронежского института МВД России. – 2025. – № 1. – С. 40–48.
8. Menshikh, V. V. Balance on Nash in "Soft" Control Problems for Systems of the Hierarchal Type / V. V. Menshikh, D. E. Orlova // Bulletin of the South Ural State University. Series: Mathematical Modelling, Programming and Computer Software. – 2019. – Vol. 12, No. 1. – P. 144–149.
9. Никитенко, А. Н. Компьютерное моделирование "мягкого" управления в мультиагентных системах / А. Н. Никитенко // Моделирование систем критического применения. – 2025. – № 2. – С. 27–34.

10. Городецкий, В. И. Децентрализация, самоорганизация и эмерджентный интеллект - цифровой взрыв умных технологий / В. И. Городецкий, О. Н. Граничин, П. О. Скобелев // Материалы общих заседаний 15-й Мультиконференции по проблемам управления : Материалы конференции, Санкт-Петербург, 04–06 октября 2022 года. – Санкт-Петербург: Концерн «Центральный научно-исследовательский институт «Электроприбор», 2022. – С. 40–54.

11. Сохова, З. Б. Модель самоорганизации автономных агентов в децентрализованной среде / З. Б. Сохова, В. Г. Редько // Проблемы управления. – 2021. – № 2. – С. 42–51.

12. Многоагентная система для разрешения конфликтных ситуаций в воздушном пространстве / А. В. Малышев, В. И. Городецкий, О. В. Карсаев [и др.] // Труды СПИИРАН. – 2006. – Т. 1, № 3. – С. 88–104.

13. Ямалов, И. У. Моделирование процессов управления и принятия решений в условиях чрезвычайных ситуаций / И. У. Ямалов. – Москва : Лаборатория знаний, 2020. – 291 с.

Модифицированный алгоритм кумулятивных сумм (Adaptive CUSUM) для обнаружения разладок в временных рядах метеоданных

М. С. Петренко, email: mariapetrenko003@gmail.com
Е. А. Сирота, email: sirota_ea@sc.vsu.ru

Воронежский государственный университет

Аннотация. *Статья посвящена задаче выявления точек разладки в временных рядах на примере метеоданных. В ней представлен усовершенствованный вариант алгоритма Adaptive CUSUM с динамическим порогом, учитывающим вариабельность и шумы в данных, что обеспечивает более высокую точность детекции по сравнению с традиционным CUSUM. Выполнен сравнительный анализ производительности Adaptive CUSUM и классического CUSUM на основе реальных данных о температуре воздуха в Москве за 2024 год. Доказано превосходство Adaptive CUSUM в повышении качества прогнозирования временных рядов.*

Ключевые слова: *обнаружение разладок, временные ряды, AdaptiveCUSUM, CUSUM, метеорологические данные, адаптивные алгоритмы, прогнозирование, анализ временных рядов, сезонность, волатильность, точность прогноза, статистические методы, обработка данных, моделирование.*

Введение

В наше время анализ временных рядов (ВР) приобретает первостепенное значение в множестве сфер, таких как метеорология, экономика, медицина, промышленность и другие [1]. Одна из центральных проблем в работе с ВР — выявление моментов разладки, когда резко меняются ключевые характеристики ряда, включая среднее значение, дисперсию, тренд или сезонность. Такие разладки провоцируют внешние события (например, климатические сдвиги, финансовые кризисы, катастрофы природы) или внутренние факторы (например, неисправности в системах сбора данных).

При наличии структурных сдвигов возникают сложности:

1. Изменяется модель параметров системы, что при игнорировании приводит к ненадёжным оценкам и снижению прогнозной точности.
2. Усложняется классификация типа ряда.

© Петренко М. С., Сирота Е. А., 2026

3. Метеоданные отличаются повышенной изменчивостью, выраженной сезонностью и шумами, требуя сложных методов для детекции разладок.

Классические алгоритмы вроде CUSUM (Cumulative Sum Control Chart) ограничены жёсткими порогами и слабой адаптацией к динамике данных, что провоцирует ложные тревоги или упущения реальных изменений.

Поэтому востребована разработка продвинутых методов, гибко реагирующих на структурные трансформации ВР с учётом особенностей метеоданных.

В статье вводится доработанная версия CUSUM под названием AdaptiveCUSUM.

Цель исследования — создание и внедрение улучшенного алгоритма AdaptiveCUSUM для поиска разладок в ВР на примере метеоданных.

2. Классический алгоритм CUSUM

CUSUM основывается на накоплении отклонений текущих значений от ожидаемого среднего. Алгоритм фиксирует разладку, когда накопленная сумма отклонений превышает заданный порог.

Пусть ВР x_t - последовательность наблюдений $x_1, ..., x_N$, где каждое наблюдение x_t соответствует значению процесса в момент времени *t*.

До момента разладки t_a значения временного ряда распределены вокруг среднего μ_0 с некоторым стандартным отклонением σ. После разладки среднее значение процесса изменяется на μ_1 [2].

Таким образом, рассматриваются следующие гипотезы:

6. Гипотеза H_0 (отсутствие изменений): среднее значение временного ряда принимает значение μ_0 .

7. Гипотеза H_1 (наличие изменений): среднее значение временного ряда принимает значение μ_1 [3].

Алгоритм использует “онлайн”-подход, то есть обнаружение изменений происходит постепенно, для каждого нового наблюдения x_t вычисляются верхняя кумулятивная сумма S_t^+ и нижняя кумулятивная сумма S_t^- по формулам (1), (2) [3, 9]:

$$S_t^+ = \max(0,\ S_{t-1}^+ + (y_t - \mu_0) - k), \quad (1)$$

$$S_t^- = \min(0,\ S_{t-1}^- - (y_t - \mu_0) - k), \quad (2)$$

где $S_0^+ = 0$, $S_0^- = 0$, k - параметр дрейфа, определяется по формуле (3) как половина разницы между μ_0 и μ_1:

$$k = \frac{|\mu_1 - \mu_0|}{2}. \quad (3)$$

Если выполняется соотношение (4):

$$S_t^+ > h, \quad (4)$$

среднее значение временного ряда увеличилось[3].

Если выполняется соотношение (5):

$$S_t^- > h, \quad (5)$$

среднее значение временного ряда уменьшилось. В этих случаях принимается гипотеза H_1 и фиксируется разладка.

Пороговое значение h определяется эмпирически. Как правило, для данных с низким уровнем шума $h \in [4,6]$. Если данные сильно зашумлены $h \in [6,10]$, если ложные срабатывания недопустимы, $h \in [8,12]$ [9].

Как отмечалось ранее, классический алгоритм CUSUM обладает рядом ограничений:

1. Высокая зависимость от подбора параметров порога h и дрейфа k.
2. Трудности настройки под нестационарные ряды: CUSUM рассчитан на стационарность (постоянство среднего и дисперсии), поэтому при наличии трендов или сезонности возникают ложные тревоги.
3. Эффективность только для структурных сдвигов среднего. медленные или мелкие изменения часто остаются незамеченными.
4. Чувствительность к выбросам, которые провоцируют ложные срабатывания без настоящей разладки.

3. Адаптивный алгоритм CUSUM

Авторы предлагают модификацию классического алгоритма CUSUM под названием AdaptiveCUSUM, которая объединяет проверенные статистические методы: анализ волатильности данных, обработку выбросов через z-оценки, коэффициент чувствительности к сдвигам и t-тест для проверки значимости изменений среднего. Эти приёмы, описанные в источниках, применяются для динамической настройки адаптивного порога.

В отличие от статичного порога h в традиционном CUSUM, ориентированного на стационарную дисперсию и гомоскедастичность,

предложенный адаптивный порог корректируется с учётом гетероскедастичности, сезонных пиков волатильности и аномальных значений, типичных для метеоданных.

Таким образом, адаптивный порог является функцией локальной волатильности и экстремальности наблюдений, сохраняя целевое значение средней задержки между ложными тревогами ARL_0 в среднем по ряду, и задается формулой (6):

$$h_t = f(z_t, \sigma_t, \beta) = h_0 + \beta z_t \sigma_t, \tag{6}$$

где h_0 - базовый порог, который обеспечивает минимальный уровень накопления доказательств при стабильной волантильности σ_t. Соответствует классической настройке CUSUM на $ARL_0 \approx 370$. Значение определяется эмпирически и лежит в диапазоне [4, 5] [9].

σ_t - локальная волатильность, вычисляемая по скользящему окну w (7), (8):

$$\sigma_t = \sqrt{\frac{1}{w} \sum_{i=t-w+1}^{t} (x_i - \bar{x}_t)^2}, \tag{7}$$

$$\bar{x}_t = \frac{1}{w} \sum_{i=t-w+1}^{t} x_i. \tag{8}$$

Локальная волантильность масштабирует порог пропорционально текущему уровню шума.

z_t - Z-оценка экстремальности (9):

$$z_t = \frac{x_t - \bar{x}_t}{\sigma_t}. \tag{9}$$

Значения $z_t > F^{-1}(1-0.005) = 1.96$ интерпретируются как выбросы и значение порога мгновенно повышается, блокируя ложную тревогу от единичного события [7]. $F^{-1}(1-\alpha)$ - квантиль нормального распределения, $\alpha = 0.005$ - уровень значимости.

β - Коэффициент чувствительности, который балансирует адаптивность:

При росте β сильнее реакция на изменения волатильности. Как следствие, меньше ложных тревог, но увеличится задержка обнаружения.

При уменьшении β наблюдается более слабее адаптация, быстрее обнаружение разладки, но больше ложных срабатываний.

На первом этапе данного алгоритма для ВР, состоящего из наблюдений $x_1, ..., x_w$, где w- размер скользящего окна, рассчитываются начальные значения среднего, стандартного отклонения по формулам (10), (11) [6]:

$$\mu_0 = \frac{1}{w}\sum_{i=1}^{w} x_i, \tag{10}$$

$$\sigma_0 = \sqrt{\frac{1}{w}\sum_{i=1}^{w}(x_i - \mu_0)^2}. \tag{11}$$

А так же определяются начальные значения верхних и нижних кумулятивных сумм $S_0^+ = 0$, $S_0^- = 0$ [4].

На втором этапе для каждого значения $t = w+1, ..., N$ ВР $x_{w+1}, ..., x_N$ рассчитывается локальная статистика окна по формулам (12), (13):

$$\sigma_t = \sqrt{\frac{1}{w}\sum_{i=t-w+1}^{t}(x_i - \bar{x}_t)^2}, \tag{12}$$

$$\bar{x}_t = \frac{1}{w}\sum_{i=t-w+1}^{t} x_i. \tag{13}$$

Расчёт z-оценки производится по формуле (9). Затем рассчитывается адаптивный порог (6) и, аналогично классическому алгоритму CUSUM, происходит обновление кумулятивных сумм (1), (2) и проверка превышения порогового значения (4), (5).

На следующем этапе используется t-критерий Стьюдента для проверки статистической значимости сдвига среднего после превышения порогового значения по формулам (14)-(16) [7]:

$$s_t = \sigma_t\sqrt{\frac{w-1}{w}}, \tag{14}$$

$$t = \frac{\bar{x}_t - \mu_0}{\frac{s_t}{\sqrt{w}}}, \tag{15}$$

$$t_{крит.} = t_{w-1}(1-\frac{\alpha}{2}). \tag{16}$$

Если значение $t > t_{крит}$, фиксируется момент разладки.

По окончании алгоритм возвращает массив точек разладки, значения верхних и нижних кумулятивных сумм и информацию о сегментах.

4. Численный эксперимент. Сравнение результатов прогнозирования ВР моделей, построенных на базе алгоритма Хольта-Винтерса

Сравнительный анализ обсуждаемых алгоритмов проводился на примере модельных данных ежедневных показателей температуры воздуха в г. Москва за период 1.01.2024–1.01.2025. Данные взяты с сайта http://meteo.ru/data/temperature-precipitation/. В рамках эксперимента были реализованы классический алгоритм CUSUM и его модификация AdaptiveCUSUM, построена модель Хольта-Винтерса, которая рассматривалась в работе [8].

На первом этапе проводилось исследование ВР и определение моментов разладки с помощью адаптивного алгоритма CUSUM, результаты работы которого отображены на рис.1.

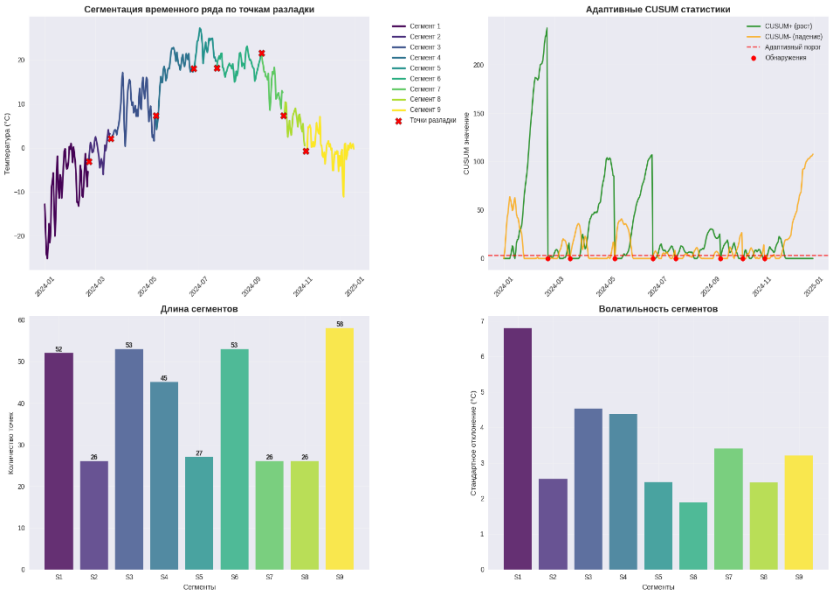


Рис. 1. Исследование ВР и определение моментов разладки с помощью адаптивного алгоритма CUSUM

Следующий этап — построение прогноза модельных данных и данных, прошедших проверку на моменты разладки. Прогнозирование

выполнялось на базе модели Хольта-Винтерса [8]. Так же была проведена оценка качества прогноза построенных моделей.

Пример прогнозирования значений временного ряда с учётом точек разладки представлен на рис. 2.

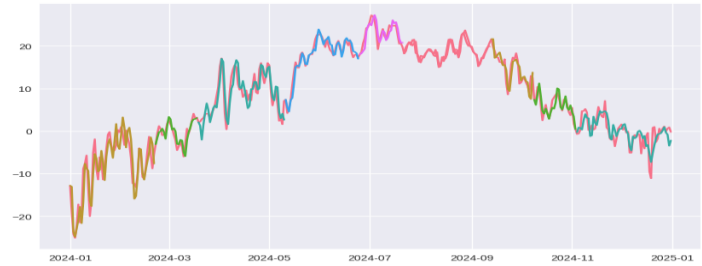


Рис. 2. Пример прогнозирования значений временного ряда с учётом точек разладки

Пример прогнозирования значений временного ряда с помощью модели Хольта-Винтерса без учёта точек разладки представлен на рис. 3.

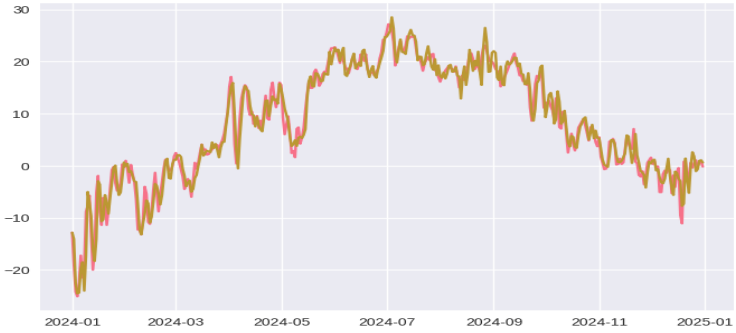


Рис. 3. Пример прогнозирования значений временного ряда с помощью модели Хольта-Винтерса без учёта точек разладки

Показатели функций потерь MAE, MSE, MAPE для оценки адекватности моделей с применением адаптивного алгоритма CUSUM и без него в задаче прогнозирования временного ряда приведены в таблице.

Из таблицы видно, что наиболее качественный прогноз был построен моделью с учетом точек разладки, полученных с помощью AdaptiveCUSUM.

Таблица

Показатели ошибок

	MSE	MAE	MAPE
AdaptiveCUSUM	3.72	1.24	0.63
Классический CUSUM	3.8	1.55	0.7
Прогнозирование без детекции разладок	4.61	1.52	0.8

Заключение

Полученные результаты численного эксперимента однозначно подтверждают необходимость предварительной детекции разладок при прогнозировании нестационарных временных рядов метеоданных. Игнорирование структурных сдвигов приводит к систематическим ошибкам параметрических моделей: MSE исходного ряда достигает 4.61 против 3.72 у AdaptiveCUSUM. Неучтённые разладки искажают оценки тренда, сезонности и уровня шума при построении прогноза ВР, что особенно критично при изменяющемся климате, где частота аномалий растёт.

Модифицированный алгоритм AdaptiveCUSUM оптимален для гетероскедастичных данных по следующим причинам:

Учёт волатильности: порог автоматически растёт на шумных участках, блокируя сезонный шум;

Т-тест дополнительно фильтрует шум, повышая интерпретируемость разладок;

Также сравнение метрик из таблицы 1 демонстрирует превосходство: MAPE 0.63%, (AdaptiveCUSUM) против 0.8% (прогнозирование без предобработки).

Список литературы

1. Гребенюк, Е. А. Методы анализа нестационарных временных рядов с неявными изменениями свойств / Е. А. Гребенюк // Автомат. и телемех. – 2005. – №12. – С. 3–30.

2. Гребенюк, Е. А. Анализ и оперативная диагностика систем, описываемых нестационарными случайными процессами / Е. А. Гребенюк // Пробл. управл. – 2003. – №4. – С. 23–29.

3. Гребенюк, Е. А. Мониторинг нестационарных процессов: анализ и исследование изменений свойств стационарности / Е. А. Гребенюк // Пробл. управл. – 2004. – №3. – С. 15–20.

4. Гребенюк, Е. А. Обнаружение изменений свойств нестационарных случайных процессов / Е. А. Гребенюк // Автомат. и телемех. – 2005. – №12. – С. 3–30.

5. Матвеев, М. Г. Моделирование динамики атмосферных процессов на основе анализа многомерных временных рядов / М. Г. Матвеев [и др.] // Воздушно-космические силы. Теория и практика. Электронное периодическое издание. – 2017. – №1 (1). – С. 98–110.

6. Ross, G. J. Adaptive CUSUM Charts for Monitoring a Process Mean / Ross, G. J., Adams, N. K., Woodall, W. H. // Journal of Quality Technology – 2011. –№43(1). – С. 1 –14.

7. Sparks, R. Adaptive Charts for Process Monitoring / R. Sparks // Journal of Quality Technology – 2000. –№32(3). – С. 262 –271.

8. Петренко, М. С. Анализ и исследование некоторых методов прогнозирования нестационарных временных рядов / М. С. Петренко, Е. А. Сирота // Труды молодых учёных факультета компьютерных наук ВГУ. Выпуск 5 / под общ. ред. Д. Н. Борисова ; Воронежский государственный университет. - Воронеж, 2025. - 719 с. - С. 222-229.

9. Granjon P. The CUSUM algorithm a small review/ P. Granjon //GIPSA-lab – 2014.

Создание сервиса автоматизации поиска работы и стажировок для студентов

М.С. Полумеев, email: polumeevm@gmail.com
Е. А. Копытина, email: zhemkaterina@yandex.ru

Воронежский государственный университет

Аннотация. *В данной статье рассматривается разработка сервиса автоматизации поиска работы и стажировок для студентов, ориентированного на упрощение процесса подбора вакансий, анализа требований работодателей и отклика на предложения.*

Ключевые слова: *автоматизация поиска работы, стажировки, студенты, веб-сервис, информационная система, карьерное развитие, агрегатор вакансий.*

Введение

Для студентов и выпускников высших учебных заведений поиск первой работы или стажировки является наиболее актуальной и сложной задачей. Изучение современных исследований и практик показывает, что значительная часть студентов сталкивается с трудностями в поиске релевантных вакансий, отсутствием структурированной информации, а также необходимостью вручную отслеживать большое количество сайтов и платформ с предложениями о трудоустройстве.

В связи с этим задача автоматизации поиска работы и стажировок для студентов является актуальной и востребованной, как показано в [1-6]. Создание специализированного сервиса, ориентированного именно на молодые кадры, позволит не только упростить процесс поиска, но и повысить качество подбора вакансий за счёт учета индивидуальных характеристик пользователя.

Целью данной работы является разработка сервиса автоматизации поиска работы и стажировок для студентов, создающий централизованный доступ к вакансиям, интеллектуальную фильтрацию и поддержку процесса отклика.

1. Постановка задачи

В рамках разработки сервиса автоматизации поиска работы и стажировок для студентов была сформулирована следующая задача: создать веб-приложение, способное подбирать вакансии из различных

© Полумеев М. С., Копытина Е. А., 2026

источников, анализировать профиль пользователя и предоставлять персонализированные рекомендации.

К разрабатываемому сервису были предъявлены следующие требования:

- реализовать систему регистрации и аутентификации пользователей;
- предоставить возможность заполнения и редактирования профиля студента (образование, навыки, интересы, опыт);
- разработать механизм автоматического сбора и обновления вакансий и стажировок из внешних источников;
- реализовать фильтрацию и сортировку вакансий по различным критериям (направление, формат занятости, город, уровень подготовки);
- обеспечить формирование персонализированных рекомендаций на основе профиля студента;
- предоставить возможность сохранения избранных вакансий и отслеживания статуса откликов.

2. Средства реализации программного продукта

Для реализации сервиса автоматизации поиска работы и стажировок была выбрана клиент-серверная архитектура. Основные средства и технологии разработки включают:

- Python - высокоуровневый язык программирования, используемый для реализации серверной логики приложения, отличающийся простотой синтаксиса, широкой экосистемой библиотек и удобством разработки веб-сервисов;
- Django - обеспечивает встроенные механизмы работы с базами данных, аутентификации пользователей и административной панели;
- Django Rest Framework - набор инструментов для создания RESTful веб-сервисов на основе Django, упрощающий реализацию API и обмен данными между клиентом и сервером;
- PostgreSQL - объектно-реляционная система управления базами данных для хранения информации о пользователях, вакансиях, профилях студентов и истории откликов;
- Встроенный ORM Django - технологии объектно-реляционного отображения, обеспечивающие удобное взаимодействие приложения с базой данных без необходимости написания сложных SQL-запросов;
- Встроенные механизмы аутентификации Django - для реализации безопасной системы регистрации, аутентификации и авторизации пользователей;

- Requests, BeautifulSoup, Selenium - библиотеки для автоматического сбора данных о вакансиях и стажировках из внешних источников (парсинг сайтов, работа с API сторонних сервисов);
- HTML, CSS - языки разметки и стилизации пользовательского интерфейса;
- JavaScript - для реализации интерактивных элементов клиентской части;
- Celery и Redis - для организации фоновых задач, таких как периодический сбор и обновление вакансий, обработка больших объемов данных и отправка уведомлений пользователям.

Взаимодействие клиента и сервера осуществляется посредством HTTP-запросов типов GET, POST, PUT и DELETE с передачей данных в формате JSON, что обеспечивает удобную интеграцию между различными компонентами системы и возможность масштабирования сервиса.

3. Реализация логики программного продукта

Функционально разрабатываемый сервис состоит из следующих модулей: модуль управления пользователями, модуль обработки вакансий, модуль рекомендаций и модуль взаимодействия с внешними источниками данных.

После регистрации и авторизации в систему пользователю предлагается заполнить профиль, указав направление обучения, уровень образования, профессиональные интересы и набор навыков. Указанные данные сохраняются в базе данных и используются для дальнейшего анализа.

Модуль обработки вакансий выполняет автоматический сбор информации о стажировках и вакансиях из внешних источников. Полученные данные нормализуются, очищаются и сохраняются в базе данных сервиса.

На основе информации из профиля пользователя и параметров вакансий формируются персонализированные рекомендации, учитывающие факторы соответствия навыков требованиям работодателя, уровень подготовки студента и формат занятости.

Пользователь может просматривать список рекомендованных вакансий, применять дополнительные фильтры, сохранять интересные предложения и отслеживать историю откликов.

Заключение

В данной статье был рассмотрен процесс разработки сервиса автоматизации поиска работы и стажировок для студентов. Реализация

предложенного варианта способствует существенному сокращению затрат времени студентов на поиск подходящих вакансий, повышение качества подбора вариантов рабочих мест и упрощение взаимодействия с работодателями.

Разработанный сервис соответствует поставленным требованиям и может быть использован в качестве основы для дальнейшего развития. В перспективе планируется расширение функциональности за счёт внедрения интеллектуальных алгоритмов анализа резюме, интеграции с карьерными центрами вузов и использования методов машинного обучения для повышения точности рекомендаций.

Список литературы

1. Джаббаров, Р. Р. Разработка серверного приложения для сбора рыночных исследований / Р. Р. Джаббаров, Е. А. Копытина // Труды молодых ученых факультета компьютерных наук ВГУ : Сборник научных трудов. – Воронеж : Воронежский государственный университет, 2025. – С. 61-67. – EDN UFKOCV.

2. Габидуллин, Т. Р. Разработка web-приложения "БСК.работа" для поиска работы студентам на АО БСК / Т. Р. Габидуллин // Малоотходные, ресурсосберегающие химические технологии и экологическая безопасность : Сборник материалов VI Международной молодежной научно-практической конференции, Стерлитамак, 28 сентября 2023 года. – Стерлитамак: Уфимский государственный нефтяной технический университет, 2023. – С. 481-484. – EDN ROLXAE.

3. Быков, М. Н. Разработка системы аутсорсинга краткосрочных работ на предприятиях / М. Н. Быков, Е. А. Копытина // Труды молодых учёных факультета компьютерных наук ВГУ : Сборник статей. – Воронеж : Воронежский государственный университет, 2024. – С. 42-50. – EDN VLZHBI.

4. Шарипов, Б. Р. Автоматизированная система поиска работы и найма персонала / Б. Р. Шарипов // Международная молодежная научная конференция "XXII Туполевские чтения (школа молодых ученых)" : МАТЕРИАЛЫ КОНФЕРЕНЦИИ СБОРНИК ДОКЛАДОВ, Казань, 19–21 октября 2015 года / Российский фонд фундаментальных исследований, Казанский национальный исследовательский технический университет им. АН. Туполева-КАИ (КНИТУ-КАИ). Том IV. – Казань: Издательство "Фолиант", 2015. – С. 375-377. – EDN VAZDDR.

5. Копытина, Е. А. Разработка веб-сервиса для формирования горячего коммерческого предложения ООО «ЧерноземАгромаш» / Е. А. Копытина, Ф. А. Десятириков, А. В. Копытин // Информатика: проблемы, методы, технологии : Материалы XXII Международной научно-практической конференции им. Э.К. Алгазинова, Воронеж, 10–12

февраля 2022 года / Под редакцией Д.Н. Борисова. – Воронеж: Общество с ограниченной ответственностью "Вэлборн", 2022. – С. 990-996. – EDN SXFFFI.

6. Абрамцова, А. С. Разработка web-приложения для поиска работы и публикации вакансий / А. С. Абрамцова // Цифровой регион: опыт, компетенции, проекты : Сборник статей VII Международной научно-практической конференции, посвященной 95-летнему Юбилею Брянского государственного инженерно-технологического университета, Брянск, 29–30 мая 2025 года. – Брянск: Брянский государственный инженерно-технологический университет, ООО "Новый проект", 2025. – С. 19-23. – EDN IODKOD.

Разработка веб-приложения для создания индивидуальных планов развития сотрудников IT-компаний

П. Н. Попова, email: popovapolina2133@gmail.com
Е. А. Копытина, email: zhemkaterina@yandex.ru

Воронежский государственный университет

Аннотация. *В данной статье описывается разработка информационной системы для создания и ведения индивидуальных планов развития сотрудников IT-компаний с возможностью отслеживания прогресса их выполнения.*

Ключевые слова: *ментор, индивидуальный план развития (ИПР), канбан-доска, задача.*

Введение

Развитие профессиональных компетенций сотрудников является одним из ключевых факторов успешной деятельности IT-компаний. В условиях быстрого обновления технологий и высокой конкуренции на рынке труда особую роль играет системный подход к обучению и развитию специалистов. Для этого в компаниях используются индивидуальные планы развития сотрудников, позволяющие учитывать текущий уровень знаний, цели специалиста и требования проектов.

В настоящее время индивидуальные планы развития часто ведутся с использованием онлайн-таблиц. Такой подход усложняет отслеживание прогресса выполнения целей и повышает риск потери данных. Кроме того, отсутствие единого инструмента делает процесс взаимодействия между сотрудниками и менторами менее прозрачным, как показано в [1-6].

Данная статья посвящена разработке веб-приложения для создания индивидуальных планов развития сотрудников IT-компаний, ориентированного на работу сотрудников и менторов. Реализация такого приложения позволит повысить эффективность процесса развития персонала и сделать его более структурированным и прозрачным.

1. Постановка задачи

Постановка задачи в рамках разработки веб-приложения для создания индивидуальных планов развития сотрудников IT-компаний заключается в необходимости разработки информационной системы, предназначенной для формирования, хранения и сопровождения

© Попова П. Н., Копытина Е. А., 2026

индивидуальных планов развития сотрудников. К разрабатываемому веб-приложению были выдвинуты следующие требования:

- Реализовать систему аутентификации для двух ролей пользователей (ментор и сотрудник);
- Предоставить менторам возможность создавать индивидуальные планы развития для своих подчиненных с автоматической генерацией уникального кода доступа;
- Разработать функционал управления задачами в рамках ИПР с указанием названия, описания, приоритета, срока выполнения и связанных навыков;
- Реализовать визуализацию задач в формате канбан-доски с тремя статусами выполнения (к выполнению, в работе, готово);
- Предоставить сотрудникам возможность отслеживать свой прогресс и самостоятельно изменять статус задач

2. Средства реализации программного продукта

Основной код веб-приложения написан на объектно-ориентированном языке Python с использованием следующих технологий:

- FastAPI – веб-фреймворк для создания REST API и веб-приложений на Python. FastAPI обеспечивает высокую производительность, автоматическую генерацию документации API и встроенную валидацию данных.
- PostgreSQL – объектно-реляционная система управления базами данных (СУБД) с открытым исходным кодом.
- SQLAlchemy – библиотека для языка Python, которая связывает объектно-ориентированное программирование с реляционной базой данных, сопоставляя типы данных Python с типами данных SQL.
- Uvicorn – ASGI сервер для запуска веб-приложений на Python, обеспечивающий асинхронную обработку запросов.
- HTML5 – язык гипертекстовой разметки документов. CSS3 – формальный язык для описания стиля
- JavaScript – язык программирования для реализации интерактивности на стороне клиента.

Разработанное веб-приложение имеет клиент-серверную архитектуру. Клиент и сервер взаимодействуют посредством обмена сообщений типа GET, POST, DELETE, PATCH протокола HTTP.

3. Реализация логики программного продукта

Система управления индивидуальными планами развития может быть условно разделена на три части: создание ИПР с выбором задач из каталога, управление задачами на канбан-доске и отслеживание прогресса выполнения.

На рис. 1 представлена страница создания индивидуального плана развития. Ментор заполняет форму с данными сотрудника: ФИО и email. После заполнения формы система переходит к этапу выбора задач.

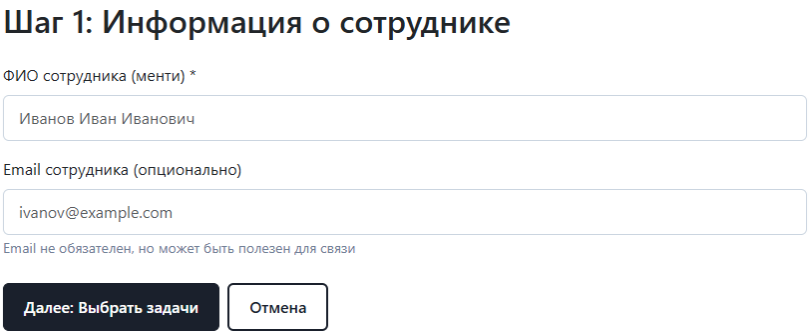


Рис. 1. Форма для заполнения данных о сотруднике

На рис. 2 представлен каталог готовых задач, сгруппированных по категориям навыков. Система содержит базу знаний с задачами для развития различных компетенций. Каждый таск включает название навыка, уровень сложности, описание и критерии приемки.

Ментор выбирает необходимые задачи из каталога, отмечая их галочками. После выбора система создает ИПР и автоматически генерирует уникальный код доступа для сотрудника.

Перед реализацией логики веб-приложения были определены основные функции, необходимые для работы с индивидуальными планами развития сотрудников IT-компаний. В результате в системе реализованы следующие компоненты:

- База данных с задачами, сгруппированными по категориям навыков;
- Форма создания индивидуального плана развития с двухэтапным процессом заполнения;
- Канбан-доска для визуализации задач и управления процессом их выполнения;
- Механизм автоматической генерации кода доступа для сотрудников.

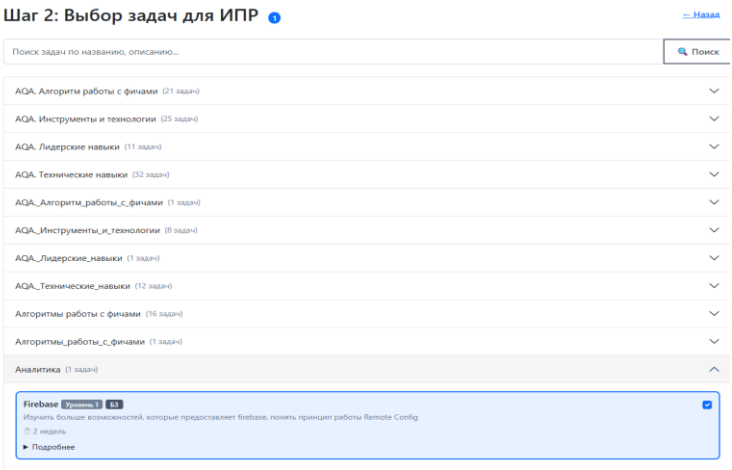


Рис. 2. Выбор задач для ИПР

При разработке системы особое внимание уделялось удобству работы для различных ролей пользователей. На рис. 3 представлена канбан-доска с задачами сотрудника.

На канбан-доске задачи распределены по трём колонкам: «К выполнению», «В работе» и «Готово». Пользователь может перемещать карточки задач между колонками, изменяя их текущий статус. Каждая карточка содержит основную информацию о задаче: название, приоритет, срок выполнения и связанные навыки.

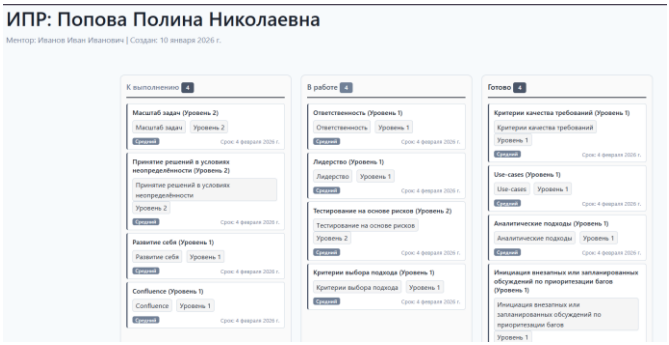


Рис. 3. Канбан-доска с задачами ИПР сотрудника

Фрагмент карточки задачи с подробной информацией представлен на рис. 4. При выборе карточки открывается модальное окно с описанием задачи, критериями выполнения и возможностью добавления комментариев.

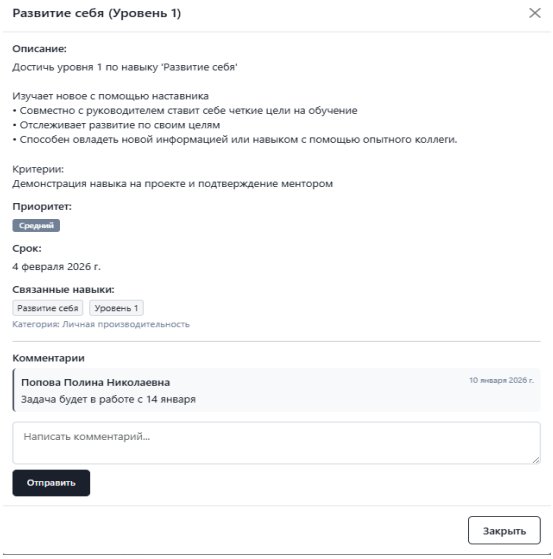


Рис. 4. Детальная карточка задачи

Также в системе реализована аутентификация пользователей и профиль, содержимое которого зависит от роли пользователя. Для сотрудника на странице профиля отображается информация о нем. Для ментора страница профиля содержит список сотрудников с активными планами развития.

Данные хранятся реляционной базе данных PostgreSQL. После выполнения всех задач индивидуальный план развития считается завершённым и закрывается.

Заключение

В данной статье была рассмотрена реализация веб-приложения для создания и управления индивидуальными планами развития сотрудников IT-компаний. Разработанные функциональные возможности соответствуют всем поставленным требованиям. В будущем планируется расширение функциональности данного программного продукта: добавление возможности для сотрудника выполнять несколько ИПР

одновременно, реализация архива выполненных планов развития с возможностью просмотра истории, внедрение системы уведомлений о приближающихся сроках задач на электронную почту.

Список литературы

1. Манукян, С. Ш. Автоматизация взаимоотношений с клиентами в SEO- компании / С. Ш. Манукян, Е. А. Копытина // Труды молодых ученых факультета компьютерных наук ВГУ. Том Выпуск 2. – Воронеж : ООО «ВЭЛБОРН», 2022. – С. 108-113. – EDN PUNHZU.

2. Баранова, Ю. К. Формирование индивидуального плана развития сотрудника как технология управления профессиональным развитием персонала организации / Ю. К. Баранова // Экономика, управление, общество: история и современность : Материалы XXI Всероссийской научно-практической конференции, Хабаровск, 19 декабря 2023 года. – Хабаровск: Российская академия народного хозяйства и государственной службы при Президенте РФ, 2024. – С. 19-23. – EDN CFXSEQ.

3. Погорелов, Р. И. Адаптация тиражируемого решения от фирмы 1С (1С: Управление нашей фирмой) и разработка приложения на мобильной платформе 1С / Р. И. Погорелов, А. В. Копытин, Е. А. Копытина // Труды молодых учёных факультета компьютерных наук ВГУ : Сборник статей / Под редакцией Д.Н. Борисова. Том Выпуск 1. – Воронеж : Общество с ограниченной ответственностью "Вэлборн", 2021. – С. 121-125. – EDN EOZHFH.

4. Вырышева, А. Д. Индивидуальный план развития как инструмент профессионального развития сотрудников / А. Д. Вырышева // Экономическая наука - основополагающая дисциплина всеобщего развития : Сборник материалов V Международной научно-практической конференции, Кемерово, 26 декабря 2017 года. – Кемерово: Общество с ограниченной ответственностью "Западно-Сибирский научный центр", 2017. – С. 113-116. – EDN YPEIPB.

5. Копытина, Е. А. Определение тенденции развития строительной организации на основе прогнозирования временных рядов / Е. А. Копытина, Н. А. Петрикеева, Д. М. Чудинов // Информационные технологии в строительных, социальных и экономических системах. – 2020. – № 1(19). – С. 87-91. – EDN KYQXLC.

6. Саенко, Е. О. Способы разработки и эффективность индивидуальных планов развития сотрудников в рамках стратегического управления кадровым потенциалом на предприятии / Е. О. Саенко, Е. А. Сеньшова // Вестник молодых ученых Санкт-Петербургского государственного университета технологии и дизайна. – 2024. – № 1. – С. 514-519. – EDN SMWQLO.

Секция 6
Интеллектуальные информационные системы, компьютерная лингвистика, технологии информационного поиска

Разработка и внедрение предметно-ориентированного GPT-ассистента для поддержки математического обучения

А. В. Абрамян, email: avabramyan@sfedu.ru[1]

[1]Южный федеральный университет,
Институт математики, механики и компьютерных наук им. И.И.Воровича

Аннотация. *В статье представлен опыт разработки предметно-ориентированного GPT-ассистента для поддержки курса «Непрерывная математика». Ассистент реализован на основе Custom GPT и использует авторскую базу знаний, содержащую строго структурированные лекционные материалы. В работе подробно рассмотрена методология подготовки базы знаний и настройка поведения большой языковой модели для обеспечения дословного извлечения определений, теорем и доказательств без их переформулирования. Рассматриваются типовые ошибки, возникающие при использовании LLM в математическом контексте, и инженерные решения по их устранению. Обсуждаются результаты практического использования ассистента и ограничения предложенного подхода.*

Ключевые слова: *GPT-ассистент; искусственный интеллект в образовании; математическое образование; база знаний; RAG; непрерывная математика.*

Введение

В последние годы большие языковые модели получили широкое распространение в качестве основы интеллектуальных систем различного назначения. Их возможности активно исследуются и в образовательном контексте, где такие модели рассматриваются как инструмент поддержки учебного процесса, автоматизации консультаций и расширения доступа к учебным материалам. Применение интеллектуальных систем в образовании опирается на общие принципы теории искусственного интеллекта и многоагентных систем, а также на современные подходы к проектированию интеллектуальных программных средств [1].

Одновременно с этим использование больших языковых моделей в образовательных приложениях выявило ряд принципиальных ограничений. В частности, для дисциплин, требующих строгой формализации и точности изложения, характерны проблемы, связанные

© Абрамян А. В., 2026

с произвольным переформулированием определений, нарушением логической структуры доказательств и использованием знаний, не относящихся к конкретному учебному курсу. Эти особенности существенно ограничивают возможность прямого применения универсальных языковых моделей в математическом образовании без дополнительного архитектурного и методического контроля [2].

Одним из направлений решения указанных проблем является разработка предметно-ориентированных ассистентов, в которых поведение языковой модели ограничивается заранее заданным корпусом учебных материалов. Подобный подход предполагает использование механизмов извлечения релевантной информации из базы знаний и формирование ответов на основе этого извлечения, что позволяет снизить влияние априорных знаний модели и повысить воспроизводимость формулировок. Такие архитектурные решения развиваются, в частности, в рамках подходов retrieval-augmented generation [3].

В данной работе рассматривается опыт разработки и внедрения предметно-ориентированного GPT-ассистента «Константин», предназначенного для поддержки изучения теоретического материала математического курса на начальном этапе обучения в вузе. Ассистент интегрирован в реальный учебный процесс и ориентирован на работу исключительно с авторской базой знаний, содержащей лекционные материалы преподавателя. Целью работы является описание архитектуры ассистента, методологии подготовки базы знаний и анализа типовых ошибок, возникающих при использовании больших языковых моделей в математическом контексте.

1. Контекст внедрения и образовательная среда

GPT-ассистент «Константин» был внедрён в учебный процесс в октябре 2025 года в Институте математики, механики и компьютерных наук им. И. И. Воровича Южного федерального университета. Ассистент использовался в рамках обязательного курса, читаемого студентам первого курса направления «Фундаментальная информатика и информационные технологии». Курс реализуется в течение двух семестров и является одним из базовых элементов математической подготовки студентов данного направления.

Содержательно курс представляет собой сжатое изложение основных разделов математического анализа, дополненное элементами теории дифференциальных уравнений. Учебный материал характеризуется высокой плотностью теоретических понятий и строгими требованиями к точности формулировок и доказательств. Аудиторная работа организована в формате еженедельных занятий, включающих лекции и практические занятия, при этом значительная часть усвоения

теоретического материала переносится в самостоятельную работу студентов.

Общее число обучающихся на направлении составляет около 120 человек. Для целевой аудитории характерен неоднородный уровень предварительной математической подготовки, что на начальном этапе обучения приводит к затруднениям при работе с формальными определениями, теоремами и доказательствами, а также при самостоятельном повторении лекционного материала.

GPT-ассистент «Константин» был задуман и реализован как вспомогательный инструмент консультационной поддержки, дополняющий традиционные формы обучения. Его использование не предполагает подмены преподавателя или автоматизации оценки знаний и ориентировано исключительно на работу с теоретическим материалом курса.

2. Архитектура GPT-ассистента «Константин»

GPT-ассистент «Константин» реализован в формате Custom GPT и предназначен для диалогового взаимодействия со студентами в рамках учебного курса. Архитектура ассистента ориентирована на решение задачи консультационной поддержки при изучении теоретического материала и построена с учётом необходимости строгого контроля содержания ответов, формулировок и используемых источников информации.

Принципиальным архитектурным решением является жёсткое разграничение ролей языковой модели и базы знаний. Языковая модель используется исключительно как механизм интерпретации пользовательского запроса и извлечения релевантных фрагментов информации, тогда как все содержательные элементы ответа формируются на основе авторской базы знаний, содержащей лекционные материалы курса. Ассистенту запрещено использовать априорные знания модели, а также привлекать внешние источники информации, не включённые в базу знаний.

Используемый подход концептуально близок к архитектурам retrieval-augmented generation, в которых генерация ответа осуществляется на основе извлечения релевантных фрагментов из заданного корпуса документов [3]. Применение данного подхода позволяет существенно снизить риск появления альтернативных формулировок и некорректных интерпретаций, характерных для универсальных языковых моделей, и обеспечить воспроизводимость содержания в соответствии с авторским изложением курса.

Ключевым элементом архитектуры является системный промпт, задающий правила поведения ассистента. В промпте зафиксированы

ограничения на использование информации, требования к стилю и структуре ответов, а также правила отказа от ответа в случаях, когда запрашиваемая информация отсутствует в базе знаний. Кроме того, промпт регламентирует взаимодействие ассистента с пользователем в нестандартных ситуациях, включая эмоционально окрашенные или неконструктивные запросы, что повышает устойчивость поведения системы в условиях реального использования.

База знаний ассистента представляет собой набор файлов, содержащих лекционные материалы курса, оформленные в формате Markdown с использованием математической разметки LaTeX. Материалы структурированы по темам и разделам курса и включают определения, формулировки теорем и доказательства, оформленные с явным выделением логических блоков. Такая организация базы знаний обеспечивает корректное извлечение информации и минимизирует вероятность частичного или искажённого воспроизведения содержания.

Общие характеристики и ограничения используемой языковой модели соответствуют возможностям современных больших языковых моделей и описаны в официальных материалах разработчика [4, 5]. В рамках данной работы параметры модели не модифицировались, а управление поведением ассистента осуществлялось за счёт архитектурных решений, системного промпта и структуры базы знаний.

В процессе разработки ассистента использовался вспомогательный LLM-ассистент, выполнявший функции методического и инженерного консультанта. Его применение позволило выявлять причины некорректных ответов и вносить целенаправленные изменения в системный промпт и структуру базы знаний, обеспечив итеративное уточнение архитектуры ассистента.

3. Методология подготовки базы знаний

Корректная работа предметно-ориентированного GPT-ассистента в значительной степени определяется качеством и структурой базы знаний, используемой в качестве единственного источника содержания ответов. В процессе разработки ассистента база знаний рассматривалась не как вспомогательное хранилище информации, а как центральный элемент архитектуры ассистента, от которого напрямую зависит воспроизводимость формулировок и устойчивость поведения системы.

В основу базы знаний были положены авторские лекционные материалы курса, используемые в аудиторной работе. Исходные материалы были переработаны и приведены к единому формату с целью устранения неоднозначностей, допустимых в устном изложении, но критичных при автоматической обработке текста. Все материалы оформлялись в формате Markdown с использованием математической

разметки LaTeX, что позволило сохранить формальную точность математических выражений и обеспечить корректное извлечение формул языковой моделью.

Структурирование базы знаний осуществлялось в соответствии с логикой курса и включало явное выделение ключевых логических блоков: определений, формулировок теорем и доказательств. Особое внимание уделялось оформлению доказательств, поскольку в ходе тестирования было выявлено, что протяжённые тексты без внутренней структуры нередко воспроизводятся неполностью. Для устранения этой проблемы доказательства оформлялись с явным выделением логических шагов, что обеспечило последовательное и полное извлечение текста при ответах ассистента.

Отдельным аспектом методологии стала работа с терминологией и наименованиями разделов. Использование общеупотребительных названий математических утверждений без дополнительной конкретизации в ряде случаев приводило к извлечению альтернативных формулировок, известных языковой модели из других курсов и источников. Для предотвращения подобных ситуаций соответствующие утверждения выносились в самостоятельные разделы с уточнёнными заголовками, однозначно соотносящимися с авторским изложением курса.

Формирование базы знаний носило итеративный характер. После каждого этапа подготовки материалов проводилось тестирование ассистента на наборе типовых вопросов, соответствующих программе курса и экзаменационным требованиям. Обнаруженные некорректные ответы анализировались, после чего вносились изменения либо в структуру базы знаний, либо в формат представления материала. Существенную роль в этом процессе сыграло использование вспомогательного LLM-ассистента, выполнявшего функции методического и инженерного консультанта и позволившего выявлять расхождения между интуитивным восприятием текста человеком и его интерпретацией языковой моделью.

В результате была сформирована база знаний, адаптированная для использования в составе предметно-ориентированного GPT-ассистента и обеспечивающая воспроизводимость формулировок и доказательств в соответствии с авторским лекционным материалом при сохранении строгих ограничений на поведение языковой модели.

4. Типовые ошибки и их исправление

В процессе разработки и эксплуатации GPT-ассистента «Константин» был выявлен ряд типовых ошибок, обусловленных особенностями работы больших языковых моделей в условиях строгих

предметных ограничений. Анализ этих ошибок позволил уточнить архитектурные решения, методологию подготовки базы знаний и правила поведения ассистента.

Одной из наиболее значимых проблем стала склонность языковой модели к использованию альтернативных формулировок математических утверждений, известных ей из других источников. Это проявлялось в ситуациях, когда запрос пользователя содержал общеупотребительное название утверждения, допускающее несколько различных формулировок в рамках разных курсов. В таких случаях ассистент мог воспроизводить формулировку, не совпадающую с авторским изложением курса. Для устранения данной ошибки соответствующие утверждения были вынесены в самостоятельные разделы базы знаний с уточнёнными заголовками, однозначно соотносящимися с используемым в курсе содержанием.

Другой характерной проблемой стало неполное воспроизведение доказательств значительного объёма. При отсутствии явной внутренней структуры ассистент в ряде случаев возвращал лишь часть доказательства, опуская отдельные логические шаги. Анализ показал, что данная ошибка связана с особенностями извлечения протяжённых текстовых фрагментов. В качестве решения была принята практика пошагового оформления доказательств с явным выделением каждого логического этапа, что обеспечило последовательное и полное воспроизведение доказательств в ответах ассистента.

Отдельную группу составили ошибки, связанные с поиском релевантного фрагмента в базе знаний. В частности, при неточной или существенно переформулированной формулировке запроса ассистент мог возвращать частичный или нерелевантный ответ, несмотря на наличие необходимой информации в базе знаний. Для снижения вероятности подобных ситуаций была уточнена структура файлов базы знаний и усилена роль заголовков, отражающих ключевые понятия и утверждения курса.

Также в ходе тестирования были выявлены ошибки поведения ассистента, не связанные напрямую с содержанием учебного материала. К ним относились попытки ссылаться на внутренние элементы базы знаний, выводить названия файлов или предлагать пользователю скачать материалы. Данные проблемы были устранены путём корректировки системного промпта и введения дополнительных ограничений на допустимые формы ответа.

Наконец, было установлено, что эмоционально окрашенные или провокационные запросы могут приводить к нестабильному поведению модели и выходу за рамки заданной роли. Для повышения устойчивости

ассистента в подобных ситуациях в системный промпт были включены инструкции, регламентирующие стиль ответа и порядок реакции на неконструктивные запросы. Это позволило обеспечить предсказуемость поведения ассистента при сохранении его основной функциональности.

В совокупности выявленные ошибки и способы их устранения показали, что корректная работа предметно-ориентированного GPT-ассистента в математическом контексте требует не только наличия качественной базы знаний, но и целенаправленной инженерной настройки, основанной на анализе реального поведения системы и итеративном уточнении принятых решений.

5. Практика использования ассистента

GPT-ассистент «Константин» использовался в учебном процессе после завершения основного этапа настройки и тестирования. За рассматриваемый период было зафиксировано более 700 запросов к ассистенту, из которых свыше 100 были сформулированы преподавателем в ходе проверки корректности ответов и уточнения поведения системы. Остальные запросы поступали от студентов и отражали их реальную практику взаимодействия с ассистентом при самостоятельной работе с теоретическим материалом.

Показано, что применение больших языковых моделей в математическом образовании требует целенаправленных архитектурных и методических решений. Жёсткая привязка ответов к авторской базе знаний, структурирование учебных материалов и итеративная настройка системных инструкций позволили обеспечить воспроизводимость формулировок определений, теорем и доказательств без использования априорных знаний языковой модели.

Особенно заметной оказалась востребованность ассистента в периоды подготовки к контрольным мероприятиям и экзамену. В эти периоды наблюдался рост числа запросов, направленных на систематизацию теоретического материала и повторение ключевых утверждений курса. Ассистент позволял студентам оперативно получать доступ к необходимым формулировкам без необходимости поиска соответствующих фрагментов в конспектах, что способствовало более структурированной работе с теорией.

Одновременно практика использования выявила и характерные ограничения ассистента, связанные с формулировкой запросов. В случаях, когда студент существенно переформулировывал вопрос по сравнению с принятым в курсе изложением или использовал неточные обозначения, ассистент мог не находить релевантный фрагмент в базе знаний и отказываться от ответа. Такое поведение рассматривалось как

ожидаемое и соответствующее архитектурным ограничениям системы, а не как ошибка в её работе.

В целом опыт использования GPT-ассистента «Константин» показал, что предметно-ориентированный ассистент, разработанный с учётом особенностей математического контента и ограничений языковых моделей, может быть эффективно интегрирован в учебный процесс в качестве инструмента поддержки самостоятельной работы студентов с теоретическим материалом.

6. Ограничения и риски использования

Несмотря на положительный опыт внедрения GPT-ассистента «Константин» в учебный процесс, его использование сопряжено с рядом ограничений и потенциальных рисков, которые необходимо учитывать при интерпретации полученных результатов и при возможном расширении области применения предложенного подхода.

Одним из ключевых ограничений является зависимость корректности работы ассистента от формулировки пользовательского запроса. Архитектура ассистента предполагает сопоставление запроса с фрагментами базы знаний, сформированной на основе авторских лекционных материалов. В случаях, когда студент существенно переформулирует вопрос по сравнению с принятым в курсе изложением или использует неточные обозначения, ассистент может не находить релевантный фрагмент и отказываться от ответа. Данное поведение является следствием сознательно введённых ограничений и направлено на предотвращение появления некорректных или произвольных формулировок.

Существенным ограничением является также высокая зависимость качества ответов от структуры и формата базы знаний. Предложенный подход требует строгой организации учебных материалов, явного выделения логических блоков и однозначных заголовков. Использование неструктурированных или слабо формализованных текстов может привести к частичному извлечению информации или к нарушению целостности воспроизводимых фрагментов. В связи с этим перенос архитектуры ассистента на другие курсы требует сопоставимого уровня методической подготовки материалов.

К числу ограничений относится и отсутствие гарантии полноты ответа даже при наличии релевантной информации в базе знаний. В отдельных ситуациях ассистент может воспроизводить лишь часть необходимого фрагмента или отказываться от ответа, что обусловлено особенностями обработки текста языковой моделью и механизмами извлечения информации. Подобные случаи рассматриваются как

ожидаемые и требуют дополнительного контроля со стороны преподавателя на этапе настройки системы.

Отдельную группу рисков составляют ограничения, связанные с использованием внешней инфраструктуры. Работа ассистента зависит от доступности сервисов, предоставляющих доступ к языковой модели, а также от качества интернет-соединения. В периоды высокой нагрузки на серверы возможны задержки в ответах или увеличение числа отказов, что снижает предсказуемость работы системы. Общие характеристики и ограничения используемых языковых моделей описаны в официальных материалах разработчика [4, 5].

Наконец, принципиальным ограничением является функциональная роль ассистента в образовательном процессе. GPT-ассистент не предназначен для оценки знаний, проверки решений или формирования индивидуальных образовательных траекторий и не может рассматриваться как замена преподавателя. Его использование предполагает сохранение ведущей роли преподавателя в отборе и интерпретации учебного материала, а также в контроле образовательных результатов.

Перечисленные ограничения и риски отражают текущие возможности применения больших языковых моделей в математическом образовании и подчёркивают необходимость осознанного и контролируемого подхода к их внедрению в учебный процесс.

Заключение

В данной работе представлен опыт разработки и внедрения предметно-ориентированного GPT-ассистента «Константин», предназначенного для поддержки изучения теоретического материала математического курса на начальном этапе обучения в вузе. Ассистент был интегрирован в реальный учебный процесс и использовался студентами первого курса в качестве инструмента консультационной поддержки при самостоятельной работе с лекционными материалами.

Показано, что применение больших языковых моделей в математическом образовании требует целенаправленных архитектурных и методических решений. Ключевым фактором устойчивой работы ассистента является жёсткая привязка ответов к авторской базе знаний и отказ от использования априорных знаний языковой модели. Реализованная архитектура, основанная на извлечении релевантных фрагментов из заранее подготовленного корпуса учебных материалов, позволила обеспечить воспроизводимость формулировок определений, теорем и доказательств в соответствии с авторским изложением курса.

Существенную роль в достижении устойчивого поведения ассистента сыграла методология подготовки базы знаний, включающая

структурирование учебных материалов, явное выделение логических блоков и пошаговое оформление доказательств. Анализ типовых ошибок, возникавших в процессе настройки и эксплуатации ассистента, показал, что многие проблемы могут быть устранены за счёт корректировки структуры базы знаний и системных инструкций без изменения параметров используемой языковой модели.

Практика использования ассистента продемонстрировала его востребованность как инструмента доступа к теоретическому материалу и навигации по курсу, особенно в периоды подготовки к контрольным мероприятиям и экзамену. При этом выявленные ограничения подтверждают, что подобные ассистенты не могут рассматриваться как замена преподавателя и должны использоваться исключительно в роли вспомогательного средства обучения, дополняющего традиционные формы работы.

Представленный опыт показывает, что предметно-ориентированные GPT-ассистенты могут быть эффективно интегрированы в учебный процесс при условии осознанного управления их архитектурой и содержанием. Описанный подход может быть адаптирован для других математических дисциплин и курсов, требующих строгого следования формальным определениям и доказательствам, при сохранении ключевых принципов подготовки базы знаний и настройки поведения языковой модели.

Список литературы

1. Бессмертный, И. А. Искусственный интеллект. Введение в многоагентные системы : учебник для вузов / И. А. Бессмертный. – М. : Издательство Юрайт, 2025. – 148 с.

2. Luckin, R. Intelligence Unleashed: An Argument for AI in Education / R. Luckin, W. Holmes, M. Griffiths, L. B. Forcier. – London : Pearson Education, 2016. – 210 p.

3. Lewis, P. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks / P. Lewis [и др.] // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 9459–9474.

4. OpenAI. GPT-4 Technical Report [Электронный ресурс]. – 2023. – Режим доступа: https://arxiv.org/abs/2303.08774

5. OpenAI. Introducing GPT-5 [Электронный ресурс]. – 2025. – Режим доступа: https://openai.com/ru-RU/index/introducing-gpt-5/

Создание виртуальных ассистентов закрытого контура и их применение в образовательной, научной, и повседневной деятельности военного вуза

С. А. Алейников, e-mail: alfredus@mail.ru[1]
Е. Н. Карлова, e-mail: ekaterina-n-karlova@yandex.ru[1]

[1]Военный учебно-научный центр военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина»

Аннотация. *В данной статье рассматриваются ключевые аспекты создания виртуальных ассистентов закрытого контура (ВАЗК) на базе больших языковых моделей (LLM), функционирующих в изолированной локально-вычислительной сети (ЛВС) военного вуза, без доступа к сети Интернет. Особое внимание уделяется потенциалу их применения в подготовке офицерских кадров, анализе больших образовательных данных и обеспечении оперативной поддержки принятия решений. Подчеркивается важность обеспечения конфиденциальности, надежности и адаптации моделей к специфике военных задач.*

Ключевые слова: *военное образование, искусственный интеллект, большие языковые модели, анализ образовательных данных.*

Введение

В условиях цифровизации образования и роста требований к безопасности информации в военной сфере, разработка ВАЗК становится актуальной задачей. Одним из преимуществ цифровой трансформации образования является возможность анализа обширных массивов данных о различных аспектах образовательного процесса. Практика использования технологий искусственного интеллекта (ТИИ) в этих целях открывает новые перспективы для оценки результативности обучения, структурирования и анализа полученных данных при помощи языковых моделей.

Использовать ТИИ в сфере военного образования необходимо исходя из тех направлений, где по объективным причинам можно обеспечить качественный прирост эффективности в подготовке офицерских кадров.

В современных условиях, когда военные учреждения сталкиваются с ростом угроз кибербезопасности и необходимости обработки

© Алейников С. А., Карлова Е. Н., 2026

чувствительной информации, традиционные облачные сервисы, основанные на доступе в сеть Интернет, становятся менее привлекательными. Это побуждает к разработке альтернативных решений, которые могут работать локально, использовать LLM и данные, хранящиеся в защищенных системах.

Основным источником формирования цифровых данных об образовании является специальное программное обеспечение (СПО), разрабатываемое как военными вузами, так и поставляемое им централизовано. При этом, использование монолитной архитектуры проектирования СПО обеспечивает единый подход в эксплуатации продукта на начальных этапах, однако с учётом развития технических решений и индивидуальных особенностей организации деятельности вуза, усложняет масштабирование и внесение правок в программное обеспечение, так как любая правка требует пересборки и развёртывания всего приложения целиком. На текущий момент, использование таких программ для анализа образовательных данных осложнено ограничением доступа к сведениям об активности пользователей (логам) и полноте баз данных для формирования выборок и срезов интересующей информации. Это связано как с требованиями информационной безопасности, так и с техническими сложностями выгрузки структурированных данных в связи с отсутствием или ограниченностью функциональных возможностей встроенных инструментов.

Для оперативного выполнения частных задач с использованием ТИИ, предлагается использовать микросервисную архитектуру [1] закрытых информационных систем на основе веб-технологий, как альтернативу существующим подходам или в дополнение к ним.

Благодаря веб-технологиям данные для включения в обработку LLM могут быть получены одновременно из нескольких источников в рамках изолированной локально-вычислительной сети (ЛВС) военного вуза, где для сбора, хранения и обработки цифровой информации используются разнотиповые базы данных. Сформированный кластер больших данных необходим как для полноты анализа закономерностей образовательного процесса, так и для прогнозирования и классификации методами машинного обучения.

Фреймворки на базе веб-технологий неплохо себя показали при анализе образовательных данных. Разработанная в Военно-воздушной академии программа для углубленного анализа отзывов на выпускников «МАРС» [2], ставшая лауреатом конкурса «Лучший электронный образовательный ресурс» в рамках тематической выставки «Военное образование – на службе Отечеству» МВТФ «Армия-2024», использует алгоритмы машинного обучения для анализа успеваемости обучающихся

[3]. Нейронная сеть Brain.js, интегрированная в программу, с высокой точностью предсказывает значение оценки общего уровня подготовленности, которую выпускник получит в войсках, а также рекомендует наиболее подходящий тип воинской должности для каждого выпускника в зависимости от его достижений во время обучения в вузе. На основе оперативного прогноза успешности профессиональной деятельности могут быть выработаны меры поддержки выпускников, попадающих в зону риска получения низких оценок [4].

Необходимо отметить, что в военных вузах разрабатывается собственная система критериев и показателей оценки качества образования. Согласно этим критериям оцениваются достижения обучающихся на соответствие результатов обучения заявленным целям и компетенциям, а также выполняется оценка уровня профессиональной подготовки выпускников по месту их распределения. Одним из критериев является удовлетворённость обучением и обратная связь. Показатели основаны на опросах обучающихся, преподавателей, выпускников и их непосредственных командиров. Информация о результатах данных опросов необходима для успешного прохождения процедуры аккредитации и, согласно приказам Минобрнауки 2013 года №462,1324 должна включаться в ежегодный отчет о самообследовании.

Ручной анализ текста опросов и отзывов, написанных в свободной форме, требует значительных временных затрат, поэтому здесь на помощь приходят модели глубокого обучения LLM, включающие в себя мощные нейронные сети, способные эффективно обрабатывать и структурировать большие объемы данных военной специфики. ВАЗК является программной реализацией такого помощника.

1. Создание ВАЗК. Подготовка ИИ-сервера к работе

На сегодняшний день существует не так много open-source инструментов для использования возможностей LLM в изолированной программной среде военного вуза. Одним из активно развивающихся проектов в этой области является Ollama. Инструмент не имеет графического веб-интерфейса, однако уже позволяет запускать и использовать многие известные языковые модели (например, Gemma, Qwen, LLaMA, Mistral, DeepSeek, GPT-OSS, YandexGPT и другие) на локальном компьютере или сервере. Ollama обеспечивает высокую степень конфиденциальности, так как данные не отправляются в облако, а обработка происходит локально. Программа поддерживает работу с различными архитектурами процессоров (включая GPU и CPU) и упрощает процесс загрузки, настройки и использования моделей, что делает её популярной среди разработчиков, исследователей и пользователей, которым нужны мощные языковые модели без

зависимости от облачных сервисов. ВАЗК будет использовать доступные команды Ollama для обращения к подключенным языковым моделям (рисунок 1).

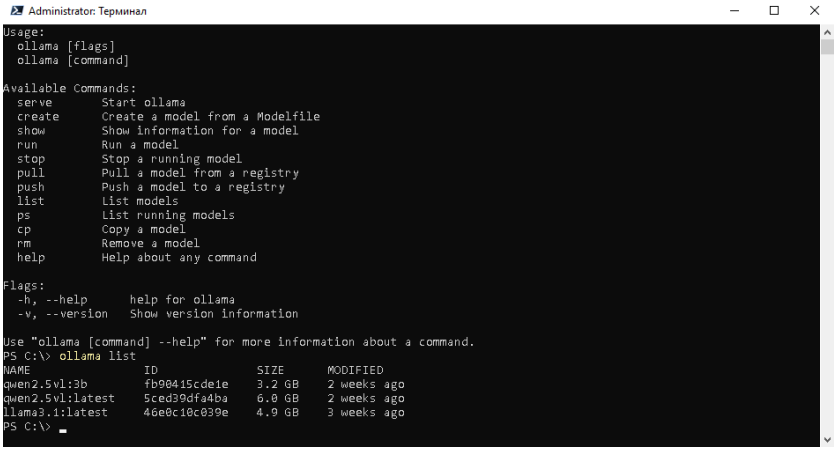


Рис. 1. Доступные команды для работы с LLM через Ollama

Для обращения к ИИ-серверу из ЛВС необходимо установить Ollama и провести первоначальную настройку программы и переменных окружения ОС Windows в PowerShell. Рассмотрим в качестве примера параметры, которые разрешают обращение к серверу всем пользователям локальной сети:

1. Прослушивание входящих соединений на всех доступных сетевых интерфейсах:
2. $env:OLLAMA_HOST="0.0.0.0"
3. setx OLLAMA_HOST "0.0.0.0" /M
4. Разрешение доступа ко всем серверам, где хранятся LLM:
5. $env:OLLAMA_ORIGINS="*"
6. setx OLLAMA_ORIGINS "*" /M
7. Изменение максимальной длины контекста:
8. $env:OLLAMA_CONTEXT_LENGTH="2048"
9. setx OLLAMA_CONTEXT_LENGTH "2048" /M
10. Использование нового механизма обработки моделей:
11. $env:OLLAMA_NEW_ENGINE="true"
12. setx OLLAMA_NEW_ENGINE "true" /M
13. Указание пути к расположению моделей:
14. $env:OLLAMA_MODELS = "C:\ollama\models"
15. setx OLLAMA_MODELS "C:\ollama\models" /M

Подробнее о параметрах настройки Ollama можно почитать в технической документации [5].

Также, необходимо создать в брандмауре правило для разрешения порта 11434, т.к. по умолчанию он используется программой для обработки промптов в локальной сети.

После внесения всех изменений – перезапустить ПК.

Для понимания технических требований к аппаратной части ИИ-сервера рассмотрим основные характеристики LLM:

1. Размер модели (количество параметров). Определяет сложность модели, её способность запоминать и обобщать информацию. Большее количество параметров обычно означает лучшую производительность, но требует больших вычислительных ресурсов.

2. Архитектура. Основная используемая архитектура – Transformer (encoder-decoder, decoder only). Существуют также Reformer, Longformer и другие. Включают механизмы внимания, которые позволяют модели учитывать контекст при обработке текста.

3. Данные для обучения. Модели обучаются на огромных объёмах текстовых данных (книги, статьи, веб-страницы, код). Качество и разнообразие этих данных сильно влияют на производительность модели. Используются датасеты Common Crawl, C4, Pile, RedPajama и другие.

4. Контекстное окно (context). Определяет максимальную длину текста, которую модель может учитывать при обработке. Измеряется в количестве токенов (единиц текста). Большие контекстные окна помогают модели лучше понять контекст запроса, но требуют больше вычислительных ресурсов.

5. Способности и задачи. Модели могут быть специализированы под конкретные задачи, такие как генерация текста (стихи, сценарии, код), перевод текста между языками, ответы на вопросы из разных областей, классификация, суммаризация текста, анализ тональности и другие.

6. Производительность. Метрики производительности напрямую зависят от выполняемых задач. Например:

- Perplexity – мера неопределённости модели;
- Bleu – для машинного перевода;
- Rouge – для суммаризации текста;
- Accuracy – точность модели;
- Recall – полнота ответа;
- F1 – для классификации текста.

– Эффективность. Характеризуется скоростью обработки данных (tokens per second), потреблением памяти (RAM, GPU) и в целом требованиям к вычислительным ресурсам.

– Доступность и лицензия. Здесь подразумеваются лицензионные ограничения на использование и распространение языковых моделей, является ли исходный код открытым (open-source) или закрытым (closed-source), а также доступность модели для использования в своих проектах через открытый программный интерфейс (API).

– Проведём оценку нагрузки на выделенный сервер, используемый для работы с LLM под управлением Ollama в локальной сети, с учётом временных затрат на обработку промптов.

– Расчёт вычислительной нагрузки.

Общее количество FLOPs для обработки N запросов вычисляется по формуле:

$$F_{total} = N \cdot T_{gen} \cdot F \cdot P \cdot 10^9 ,$$

– N – количество пользователей в локальной сети, одновременно отправляющих запросы;
– T_{gen} – время генерации ответа на один запрос (в секундах);
– F – количество флот-операций (FLOPs), необходимых для обработки одного токена;
– P – количество параметров модели.

В формуле (1) учитывается коэффициент перевода миллиардов параметров в абсолютные значения.

– Расчёт нагрузки на память.

Объём памяти, необходимый для хранения моделей, вычисляется по формуле:

$$M_{type} = \sum_{i=1}^{k} S_i ,$$

– S_i – размер файла i-й модели;
– k – количество одновременно используемых моделей.

Объём памяти, необходимый для буферизации запросов, вычисляется по формуле:

$$M_{buffer} = N \cdot T_{gen} \cdot \bar{t} ,$$

– t – среднее количество токенов на запрос.

Общая память, необходимая ИИ-серверу, вычисляется по формуле:

$$M_{total} = M_{type} + M_{buffer} ,$$

7. Условия для выбора технических характеристик сервера.

Пропускная способность GPU вычисляется по формуле:

$$R_{server} \geq \frac{F_{total}}{T_{gen}},$$

– R_{server} – реальная пропускная способность сервера (FLOPs/сек).

Требуемый объём видеопамяти вычисляется по формуле:

$$M_{server} \geq M_{total} \cdot \eta^{-1},$$

– M_{server} – объём видеопамяти сервера (в гигабайтах);
– ŋ – коэффициент эффективности использования видеопамяти ($0 < \eta < 1$).

При единовременном использовании на сервере сразу нескольких моделей – нагрузка возрастает, её расчёт будет иметь следующий вид:

$$F_{total}^{общ} = \sum_{i=1}^{n} F_{total}^{i},$$

При этом объём необходимой памяти также суммируется и вычисляется по формуле:

$$M_{total}^{общ} = \sum_{i=1}^{n} M_{type}^{i} + M_{buffer},$$

Таким образом, технические характеристики сервера напрямую зависят от числа единовременно работающих с моделями пользователей (N), размерам развёрнутых на сервере моделей (M_{total}) и организации доступа к ним. Время на генерирование ответа (T_{gen}) зависит от пропускной способности GPU (R_{server}).

При последовательном выполнении промпт-запросов (при помощи очередей), или обработке запроса небольшим количеством пользователей, достаточно иметь выделенный сервер со следующими техническими характеристиками:

1. Материнская плата: MSI B650 или аналог;
2. Процессор: Intel Core i7-13700K или AMD Ryzen 7700;
3. Оперативная память (RAM): 16Гб и выше;
4. Видеокарта: GeForce RTX 4070 Ti (16 Гб VRAM) и выше;
5. Жёсткий диск: SSD M2 1Tb;
6. Блок питания: DeepCool PL750D или аналог;
7. Система охлаждения: DeepCool LT720 или аналог;
8. Корпус ATX;

2. Архитектура и компоненты ВАЗК

Типичная архитектура ВАЗК включает следующие компоненты:

1. Модель LLM – основной вычислительный элемент, отвечающий за обработку запросов и генерацию ответов. Может быть обучена заранее или дообучена на данных военной специфики.

2. База знаний – локальная база данных, содержащая информацию, необходимую для функционирования ВАЗК, включая учебные материалы, инструкции, протоколы и другие релевантные данные.

3. Модуль обработки естественного языка (NLP) – обеспечивает понимание запросов пользователя и формирование ответов на естественном языке.

4. Графический интерфейс пользователя (GUI) – Предоставляет возможность взаимодействия пользователя с ВАЗК (текстовый чат, голосовой интерфейс и т.п.)

5. Модуль безопасности – обеспечивает защиту данных и ограничение доступа к ресурсам.

Обращение к модели из нашего приложения будет происходить по сети. ИИ-сервер получает промпт-запрос с помощью асинхронных функций inLLM и genLLM (рисунок 2).

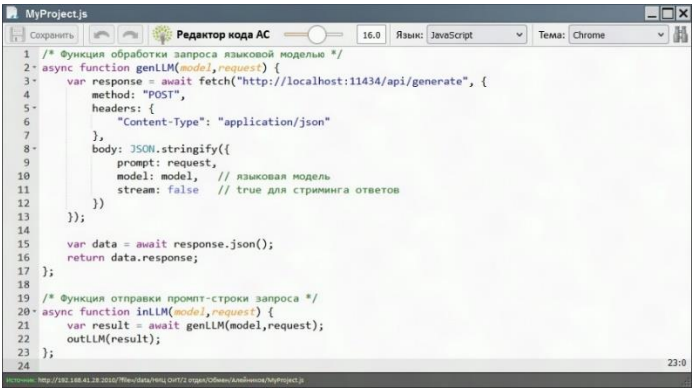


Рис. 2. Пример программной реализации кода для отправки запроса LLM

В качестве параметра «model» (строка 20) в асинхронную функцию inLLM передаётся название подключенной к Ollama модели (например, qwen2.5vl:3b). В параметре «request» содержится промпт-запрос, например – «Проведи статистический анализ недостатков профессиональной деятельности военных специалистов по отзывам их командиров (далее идут недостатки из отзывов через запятую)». Асинхронная функция genLLM (строка 2) возвращает ответ языковой

модели, который в дальнейшем можно использовать в своём приложении (вывести на экран, сформировать документ и т.п.). При помощи вызова функции outLLM (строка 22) с параметром result ответ передаётся в ВАЗК.

При небольших технических возможностях сервера необходимо продумать реализацию очереди запросов для параллельной работы пользователей с программой при пиковых нагрузках. А также, реализовать хранение контекста диалога для лучшего понимания сущности запросов нейросетью. На основе созданного механизма отслеживания контекста в ВУНЦ ВВС «ВВА» ведутся исследования по наращиванию знаний LLM в области военного образования. Для этих целей выбран подход связывания контекста с базой данных ключевых слов и описаний.

Методика обучения LLM через расширение контекста запроса (In Context Learning, ICL) – это подход, при котором модель учится выполнять новые задачи на основе примеров и инструкций, предоставленных в тексте запроса. Ключевая особенность – отсутствие обновления весов (параметров) модели во время обработки промпта. Вместо дообучения модель использует уже существующие знания, адаптируя их к задаче с помощью демонстраций. В нашем случае источником демонстраций будет база знаний ВАЗК, которая представляет собой базу данных с ключевыми словами и описаниями их контекста (рисунок 3).

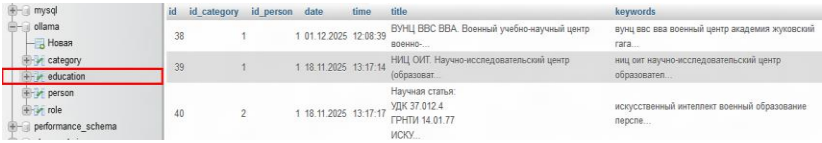


Рис. 3. Пример реализации базы знаний для обучения LLM

В ICL такой подход обучения LLM называется Few-shot Learning (обучение на нескольких примерах). Он позволит обучить нейросети военной специфике, улучшить точность и релевантность ответов. Для поиска и сравнения ключевых слов в запросе (приведение к существительным в единственном числе именительном падеже), в ВАЗК необходимо интегрировать морфологический словарь. Данный механизм станет крепкой основой наполнения базы знаний данными военного образования.

Заключение

Создание и использование ВАЗК на базе больших языковых моделей является перспективным направлением развития военного образования. Несмотря на возникающие трудности, преимущества, такие как персонализация обучения, повышение эффективности учебного процесса и снижение зависимости от внешних ресурсов, делают разработку и внедрение таких систем крайне важными и востребованными. Дальнейшие исследования в области оптимизации моделей, создания специализированных датасетов и разработки эффективных мер безопасности позволят раскрыть полный потенциал ВАЗК и значительно улучшить подготовку военных специалистов.

Список литературы

1. Алейников, С.А., Толстых, А.В., Мельников, А.В., Жукова, С.И. Структура специального программного обеспечения с микросервисной системной архитектурой для распределённых автоматизированных систем // Электронное периодическое издание «Воздушно-космические силы. Теория и практика» – Воронеж: ВУНЦ ВВС «ВВА». – 2024 – №31. – С. 64-76.
2. Алейников, С.А., Карлова, Е.Н., Гавриленко А.В. – «Информационная система поддержки принятия решений на основе анализа отзывов на выпускников из войск» / Свидетельство о государственной регистрации программы для ЭВМ №2024612307 от 30.01.2024.
3. Боронин, О.А., Карлова, Е.Н., Алейников, С.А. Доказательное развитие образования // Научно-методический журнал Министерства Обороны Российской Федерации «Вестник военного образования» – Москва – 2026 – №57. – С. 62-65.
4. Алейников, С.А., Карлова, Е.Н. Перспективы применения машинного обучения в программных модулях на микросервисной архитектуре для анализа образовательных данных // Материалы VII Международной научно-практической конференции «Динамика развития системы военного образования» (14 марта 2025 г.) – Омск: Изд-во ОмГТУ, 2025. С. 260-264.
5. Ollama’s documentation. [Электронный ресурс]. Режим доступа: https://docs.ollama.com (дата обращения: 21.01.2026).

Разработка системы выявления галлюцинаций в ответах больших языковых моделей

Е. А. Бабкина, email: bkatya5577@gmail.com [1]
В. В. Гаршина, email: garshina.veronika@gmail.com [1]

[1] Воронежский государственный университет

Аннотация. *Статья посвящена разработке системы выявления галлюцинаций в ответах больших языковых моделей (LLM). Описаны средства реализации, архитектура системы, ключевые моменты разработки. Приведены примеры использования системы на практике, наглядно иллюстрирующие результаты работы.*

Ключевые слова: *большие языковые модели, LLM, галлюцинации, качество генерации текста, достоверность ответов, самопроверка, внешние источники данных.*

Введение

В настоящее время большие языковые модели являются неотъемлемой частью активно развивающейся сферы искусственного интеллекта. Они применяются в различных областях: от создания развлекательного контента до поддержки научных исследований. Однако, несмотря на их успех, одной из основных проблем языковых моделей является феномен «галлюцинаций». Под этим термином понимается генерация информации, которая выглядит правдоподобно, но не соответствует фактическим данным или логически не вытекает из контекста запроса пользователя.

В предыдущей статье [1] были подробно описаны типы и причины галлюцинаций. Такие ошибки больших языковых моделей могут приводить к дезинформации пользователей и вызывать риски в критически важных областях, таких как медицина, юриспруденция, образование [2].

Актуальность разработки методов выявления галлюцинаций вызвана необходимостью повышения надежности и безопасности использования языковых моделей. Разработка комплексной системы, способной оценивать вероятность наличия галлюцинаций как на основе внутренних свойств модели, так и с использованием внешних источников информации, является важной задачей современных специалистов в области искусственного интеллекта. Ее решение станет важным шагом на

© Бабкина Е. А., Гаршина В. В., 2026

пути к безопасному внедрению больших языковых моделей во многие сферы жизни общества.

1. Средства реализации

Система для выявления галлюцинаций реализована на языке программирования Python, так как он имеет поддержку большого количества библиотек для машинного обучения и обработки естественного языка. К его преимуществам можно также отнести динамическую типизацию, простоту, доступность документации. Используемые в работе библиотеки:

- Transformers (Hugging Face) – основная библиотека для работы с предобученными большими языковыми моделями. Используется при инициализации моделей и токенизаторов, а также для генерации текстовых ответов.
- PyTorch – библиотека для тензорных вычислений и работы с глубокими нейронными сетями [3].
- Sentence-Transformers – специализированная библиотека для получения векторных представлений текстовых предложений (эмбеддингов). Используется для оценки смысловой близости текстов в модуле внешней проверки.
- Scikit-learn – библиотека для машинного обучения. В системе применяется для оценки согласованности ответов.
- Wikipedia-API – клиент для работы с API энциклопедии Wikipedia. В системе служит внешним источником фактических данных для проверки достоверности сгенерированных ответов.
- Accelerate – библиотека, позволяющая упростить процесс обучения. Используется для оптимизации работы модели.

В качестве платформы для разработки и тестирования системы был выбран Google Colab – облачная среда выполнения, которая предоставляет бесплатный доступ к вычислительным ресурсам. Google Colab позволяет проводить эксперименты с моделями среднего размера (например, FLAN-T5-base), без необходимости иметь специализированное локальное оборудование.

2. Архитектура системы

Архитектура разработанной системы организована в виде последовательности взаимосвязанных этапов:

- Этап базовой генерации;
- Этап внутренних проверок;
- Этап внешней верификации;
- Этап агрегации оценок.

Такая структура обеспечивает гибкость: в будущем можно будет легко добавлять новые методы проверки или модифицировать существующие.

Работу системы можно изобразить в виде схемы, представленной на рисунке 1.

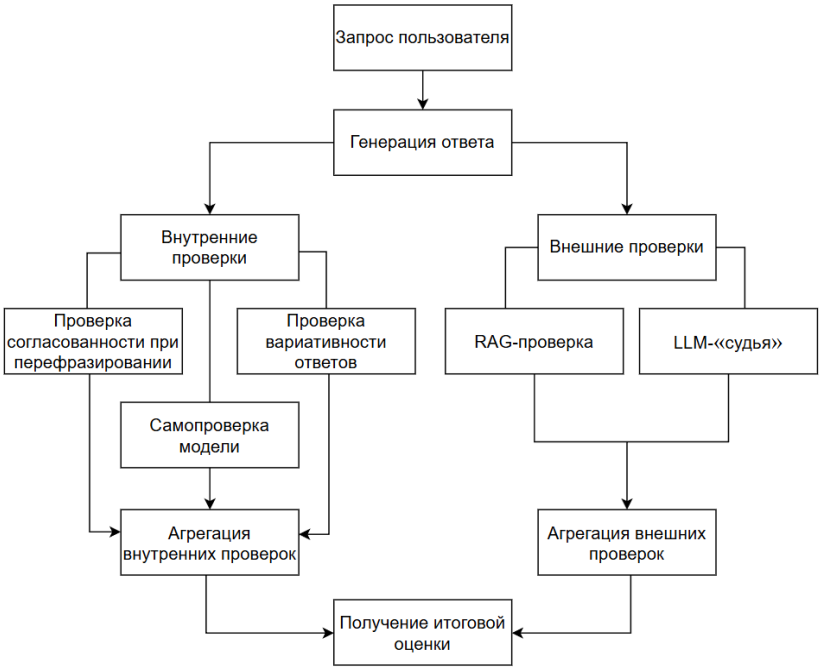


Рис. 1. Схема работы системы выявления галлюцинаций в ответе языковой модели

Входной точкой системы является запрос пользователя. На первом этапе основная языковая модель (например, FLAN-T5-base) генерирует ответ на поставленный вопрос. Этот ответ модели становится центральным объектом для всех последующих проверок и проходит через конвейер обработки. Система спроектирована с учетом возможности замены базовой модели на любую другую из семейства трансформеров. Это обеспечивает свойство системы адаптироваться под различные задачи и вычислительные ресурсы.

Слой внутренних проверок направлен на анализ результата генерации без обращения к внешним источникам информации. Он включает три основных метода:

– Проверка согласованности при перефразировании пользовательского запроса;
– Самопроверка модели;
– Проверка вариативности ответов без изменения запроса.

В рамках первого метода система формирует несколько перефразированных версий исходного запроса и генерирует ответы на каждую из них. Разброс в содержании ответов на семантически эквивалентные вопросы указывает на неустойчивость модели и является индикатором потенциальной галлюцинации. Для оценки близости ответов используется метод TF-IDF векторизации [4] с последующим расчетом сходства между полученными векторами. Чем ниже среднее попарное сходство ответов, тем выше оценка риска галлюцинации.

В рамках второго метода модели задается прямой вопрос о правильности ее собственного ответа («Is this answer correct? YES or NO»). Ответ модели об ошибке (ответ «NO») является сигналом о возможной недостоверности сгенерированной информации.

Третий метод включает в себя следующее: модель несколько раз генерирует ответ на один и тот же запрос (исходных запрос пользователя). Параметр температуры, используемый при настройке генерации, вносит элемент стохастичности [5]. Низкая согласованность между полученными вариантами указывает на высокую неопределенность модели и может служить признаком галлюцинации.

Слой внешней проверки предназначен для сопоставления ответа модели с независимыми источниками информации. Он состоит из двух методов:

– Проверка на основе извлечения данных из внешней базы знаний;
– Проверка с помощью LLM-«судьи».

Первый метод использует Wikipedia API для поиска соответствующих теме запроса статей. С помощью модели эмбеддингов (all-MiniLM-L6-v2) находится наиболее семантически близкий документ, который затем сравнивается с первоначальным ответом модели. Этот метод реализует концепцию Retrieval-Augmented Generation (RAG) в обратном направлении: вместо использования внешних данных для генерации, они применяются для проверки уже сгенерированного контента [6]. Значительное расхождение между ответом модели и найденной информацией повышает вероятность галлюцинации.

В рамках второго подхода отдельный экземпляр языковой модели (например, FLAN-T5-large, как более точный аналог FLAN-T5-base) выступает в роли независимого «судьи», оценивая пару «вопрос-ответ» и вынося вердикт о корректности ответа. Модель-«судья» отвечает на вопрос «You are a fact-checker LLM. Is this answer correct? YES or NO».

Заключительным этапом является агрегация полученных частных оценок. Каждая описанная выше проверка возвращает числовую оценку. Она может принимать значения от 0 до 1, где более высокое значение означает большую вероятность галлюцинации. Каждая проверка имеет вес, отражающий субъективную оценку важности данного типа проверки. Система рассчитывает взвешенную оценку отдельно для группы внутренних и внешних проверок. Итоговая вероятность галлюцинации представляет собой линейную комбинацию этих двух оценок. Отдельно заданный коэффициент регулирует баланс доверия между внутренней проверкой модели и внешней верификацией. Такой подход обеспечивает целостную оценку вероятности галлюцинаций в ответе языковой модели.

3. Тестирование работы системы

Тестирование системы проводилось на языковой модели FLAN-T5-base с использованием серии вопросов:

– «Who invented the telephone?» (Кто изобрел телефон?);
– «What is the capital of Australia?» (Какова столица Австралии?);
– «In which year did Albert Einstein win the Nobel Prize?» (В каком году Альберт Эйнштейн получил Нобелевскую премию?).

Тестирование системы подтвердило ее способность выявлять галлюцинации в ответах языковой модели. На запрос об изобретателе телефона модель дала ошибочный ответ «teddington». Все внутренние проверки показали высокий риск галлюцинации: самопроверка вернула результат «NO» (модель выявила ошибку в своем ответе), а проверки согласованности и вариативности при повторных генерациях выявили разброс в ответах модели («charles bell», «William McKinley», «lincoln johnson», «alexander nixon»). Внешние проверки также подтвердили ошибку: LLM-«судья» вынес вердикт «NO» (значит, ответ модели является некорректным). Итоговая вероятность галлюцинации составила 0.916.

Аналогичные результаты получены для запроса о столице Австралии. Модель ответила «melbourne» вместо Канберры. Метод самопроверки дал результат «NO», а множественные генерации дали противоречивые варианты («adelaide», «sydney», «balmain», «georgia»). Внешняя проверка через Wikipedia API извлекла корректные данные о Канберре. Финальная оценка вероятности галлюцинации – 0.901.

Наиболее показательным стал тест с годом присуждения Нобелевской премии Эйнштейну. Модель ошибочно указала 1908 год, но проявила ложную уверенность в правильности своего ответа: самопроверка вернула «YES». Однако другие проверки выявили внутреннюю несогласованность (варианты «1971», «1915», «1957», «1994» при повторных запусках генерации) и несоответствие внешним данным. LLM-«судья» ошибочно согласился с ответом модели. Итоговая вероятность галлюцинации составила 0.59, что ниже предыдущих случаев, но все равно указывает на значительный риск ошибки в ответе модели.

Отчет системы при анализе ответа модели на вопрос «What is the capital of Australia?» представлен на рисунке 2.

```
QUESTION: What is the capital of Australia?

--- MAIN LLM ANSWER ---
melbourne

--- SELF CRITIQUE ---
NO

--- PARAPHRASE ANSWERS ---
[1] adelaide
[2] capital of australia
[3] nps
[4] sydney

--- VARIANCE ANSWERS ---
[1] melbourne
[2] adelaide
[3] balmain
[4] melbourne
[5] georgia

--- EXTERNAL RAG ---
Canberra (  KAN-brə; Ngunawal: Kanbarra) is the capital city of Australia and
the only substantial settlement of the Australian Capital Territory. Founded
following the federation of the colonies of Australia as the seat of government
for the new nation, a compromise between Melbourne and Sydney, it is Australia's
largest inland city, and the eighth-largest Australian city by population. The
city is located at the northern tip of the Australian Alps, the country's
highest mountain range. As of J

--- LLM JUDGE ---
NO

--- SCORES ---
Internal score: 0.895
External score: 0.906
FINAL hallucination probability: 0.901
```

Рис. 2. Отчет системы о результатах анализа ответа модели на наличие галлюцинаций

Заключение

В данной статье описана разработанная система выявления галлюцинаций в ответах больших языковых моделей, основанная на сочетании внутренних и внешних методов проверки. Предложенный подход позволяет комплексно оценивать надежность ответов моделей. Результаты тестирования показали, что система способна эффективно находить как очевидные ошибки, так и менее явные случаи галлюцинаций. Использование многоуровневой верификации повышает практическую ценность полученной оценки. Разработанная система является перспективным инструментом для повышения безопасности применения языковых моделей в различных областях.

Список литературы

1. Бабкина, Е. А. Галлюцинации в больших языковых моделях (LLM) и методы их компенсации / Е. А. Бабкина, В. В. Гаршина // Сборник научных трудов XXV Международная конференции Информатика: проблемы, методология, технологии. – Воронеж: Издательский дом ВГУ, 2025. – С. 910-915.
2. Шадеркина, В. А. ChatGPT в медицине: возможности и ограничения / В. А. Шадеркина // Российский журнал телемедицины и электронного здравоохранения. – 2024. – Т. 10, № 1. – С. 33-43.
3. Paszke, A. PyTorch: An Imperative Style, High-Performance Deep Learning Library / A. Paszke, S. Gross, F. Massa [et al.] // 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). – 2019. – P. 12.
4. Senbel, S. Fast and Memory-Efficient TFIDF Calculation for Text Analysis of Large Datasets / S. Senbel, In H. Fujita, A. Selamat [et al.] // Advances and Trends in Artificial Intelligence: Artificial Intelligence Practices. – 2021. – P. 557-563.
5. Kabakus, A. T. Evaluating the Impact of Temperature and Instruction Strategies on Hallucination in Large Language Models / A. T. Kabakuş; GU J Sci, Part A. – 2026. – Vol. 13, No. 1. – P. 13-32.
6. Ma, X. Query Rewriting for Retrieval-Augmented Large Language Models / X. Ma, Y. Gong, P. He [et al.] // Empirical Methods in Natural Language Processing. – 2023. – P. 5303-5315.

Применение мультимодальной большой языковой модели для первичной аналитической интерпретации визуальных материалов, публикуемых хакерскими группировками

А. А. Барков, email: a.brkv_fstec@inbox.ru[1]

[1]ФАУ «ГНИИИ ПТЗИ ФСТЭК России»

Аннотация. *Рассмотрен подход к применению мультимодальной большой языковой модели для первичной аналитической интерпретации визуальных материалов, публикуемых хакерскими группировками. На основе экспериментальных сценариев показана возможность распознавания действий на видеозаписях и скриншотах, а также классификации характера наблюдаемой активности. Определены преимущества и ограничения использования данного подхода в аналитической деятельности.*

Ключевые слова: *мультимодальная языковая модель, хакерские группировки, информационная безопасность.*

Введение

В последние годы хакерские группировки всё чаще используют визуальные материалы – видеозаписи и скриншоты – в качестве средства демонстрации своей деятельности, публикуя их в открытых информационных каналах [1–2]. Подобные материалы могут содержать демонстрацию пользовательского доступа, фрагменты корпоративной ИТ-инфраструктуры, результаты нарушения работоспособности систем или элементы управления промышленными объектами.

Для аналитических подразделений существенную сложность представляет оперативная интерпретация подобных материалов и определение характера наблюдаемой активности: является ли она реальной вредоносной либо носит имитационный (демонстрационный) характер. Ручной анализ визуальных данных требует значительных временных затрат и зачастую привлечения профильных специалистов, что снижает оперативность аналитической обработки информации.

В этой связи актуальной является задача применения мультимодальных больших языковых моделей в качестве инструмента первичной аналитической интерпретации визуальных материалов. При этом модель рассматривается не как замена эксперта, а как средство преобразования визуальной информации в структурированное текстовое

© Барков А. А., 2026

описание и предварительную аналитическую гипотезу, позволяющую аналитику принять решение о дальнейшем направлении анализа.

1. Объект, предмет, цель и задачи исследования

Объект исследования – наблюдаемая активность, зафиксированная в видеозаписях и статических изображениях, публикуемых хакерскими группировками в открытых информационных каналах.

Предмет исследования – методы применения мультимодальных больших языковых моделей для интерпретации визуально наблюдаемой активности и классификации её характера.

Цель исследования – оценка применимости мультимодальных больших языковых моделей для автоматизации первичной аналитической интерпретации визуальных материалов и классификации наблюдаемой активности как реальной вредоносной либо имитационной.

В рамках исследования были поставлены и решены следующие задачи:

Анализ интерпретации пользовательской активности в графической среде операционной системы.

Оценка распознавания и понимания команд терминала в операционной системе на базе ядра Linux.

Интерпретация статических изображений корпоративной ИТ-инфраструктуры, опубликованных в открытых источниках.

Анализ видеоматериалов промышленной SCADA-системы с точки зрения характера наблюдаемой активности.

2. Методика исследования

На подготовительном этапе исследования был выполнен предварительный анализ нескольких мультимодальных больших языковых моделей [4–6], поддерживающих обработку визуальной информации. Анализ осуществлялся путём применения моделей к типовым аналитическим сценариям, отражающим практику первичной обработки визуальных материалов аналитиком, и был направлен на проверку функциональной пригодности моделей для формирования осмысленных и логически связных текстовых интерпретаций визуального контекста. Проведение количественной или качественной оценки, построение метрик, а также формирование рейтинга моделей в рамках данного этапа не предусматривались. По результатам предварительного анализа в качестве объекта настоящего исследования была выбрана модель Gemini 3 Pro Preview [3], показавшая наиболее стабильное и интерпретируемое поведение при решении задач первичной аналитической интерпретации визуальных материалов.

Рассмотрение в рамках исследования одной мультимодальной модели обусловлено заданными условиями эксперимента и направлено на оценку принципиальной применимости данного класса моделей для первичной аналитической интерпретации визуальных материалов без проведения сравнительного анализа различных реализаций.

Основной этап исследования был реализован в диалоговом режиме с применением стандартизированных текстовых запросов, ориентированных на получение структурированного описания происходящего на экране и аналитической интерпретации наблюдаемой активности.

Аналитическая интерпретация, формируемая моделью, рассматривалась как предварительная и вспомогательная и предназначалась для поддержки аналитика на этапе первичной обработки визуальных материалов. Функциональная роль модели заключалась в преобразовании визуально наблюдаемой активности, зафиксированной на видеозаписях и статических изображениях, в структурированное текстовое представление, а также в формировании предварительных аналитических гипотез о характере и контексте наблюдаемых действий. Использование модели на данном этапе направлено на снижение временных затрат и когнитивной нагрузки аналитика при первичном ознакомлении с материалами и принятии решения о целесообразности их дальнейшего углублённого анализа.

Проверка предметной, инженерной или технологической корректности выводов модели осознанно не осуществлялась и не входила в задачи исследования, поскольку требует профильной экспертизы в соответствующих предметных областях и, как правило, предполагает привлечение дополнительных источников данных, выходящих за рамки визуального контекста. Указанное ограничение является методологическим решением и отражает реальные условия аналитической деятельности при работе с визуальными материалами из открытых информационных источников, в рамках которых первичная аналитическая интерпретация предшествует экспертной верификации и специализированному техническому анализу.

Качество аналитики модели оценивалось по совокупности критериев, включающих визуально-семантическую согласованность интерпретации с наблюдаемым контекстом, логическую непротиворечивость описания, отсутствие явных «галлюцинаций», а также практическую применимость полученного текстового представления для аналитика при принятии решения о дальнейшем анализе материала.

3. Экспериментальные задачи

3.1 Объект, предмет, цель и задачи исследования

Цель задачи. Оценка способности мультимодальной большой языковой модели распознавать визуально наблюдаемую пользовательскую активность в графической среде операционной системы, корректно интерпретировать последовательность действий (навигация по файловой системе, открытие документов, просмотр содержимого) и формировать текстовое описание, пригодное для первичной аналитической обработки.

Исходные данные. Видеозапись экрана, представленная на рис.1, демонстрирующая работу пользователя в графической среде ОС Windows: переход по каталогам, использование встроенных механизмов поиска, открытие и просмотр документов различных форматов (PDF, текстовый редактор, электронные таблицы). Наблюдаемая активность представлена без ввода команд в командной строке и без использования специализированных средств проведения атак.

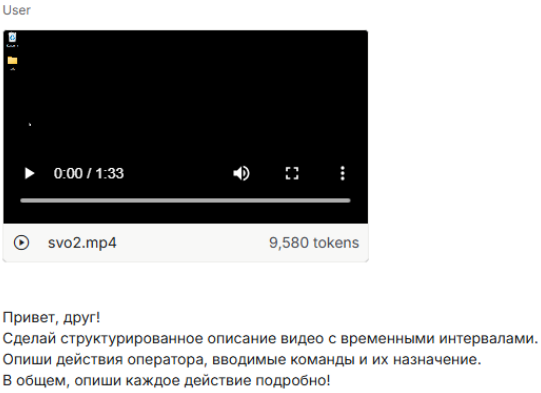


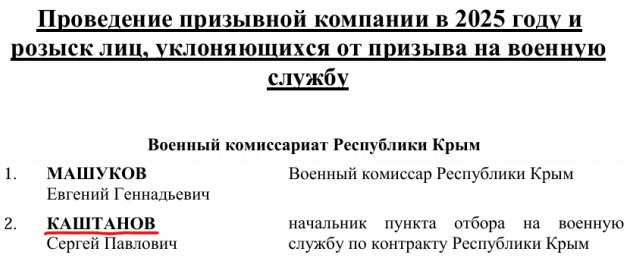
Рис. 1. Прикреплённое видео для анализа и запрос

Интерпретация модели. Модель сформировала структурированное описание происходящего на экране с временной привязкой, выделив основные этапы навигации и открытия файлов. Интерпретация включала распознавание типов открываемых документов и общего контекста демонстрируемого доступа.

Классификация активности. Поскольку на видеозаписи отсутствуют признаки технического воздействия (эксплуатация уязвимостей, выполнение команд, использование специализированных инструментов атак), наблюдаемая активность интерпретируется как демонстрация

пользовательского доступа в графической среде. В рамках задач настоящей работы подобный сценарий относится к имитационной (демонстрационной) активности, поскольку сам по себе не является свидетельством реализации вредоносных действий.

Неточности и вывод. Выявлены локальные неточности, связанные с частным распознаванием отдельных элементов (например, фрагментов текста/именований, представленных на рис. 2 и 3) и интерпретацией некоторых действий пользователя.

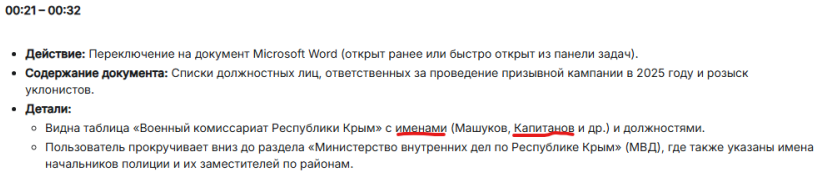


Проведение призывной компании в 2025 году и розыск лиц, уклоняющихся от призыва на военную службу

Военный комиссариат Республики Крым

1.	**МАШУКОВ** Евгений Геннадьевич	Военный комиссар Республики Крым
2.	**КАШТАНОВ** Сергей Павлович	начальник пункта отбора на военную службу по контракту Республики Крым

Рис. 2. Фамилия в документе на видео



00:21 – 00:32

- **Действие:** Переключение на документ Microsoft Word (открыт ранее или быстро открыт из панели задач).
- **Содержание документа:** Списки должностных лиц, ответственных за проведение призывной кампании в 2025 году и розыск уклонистов.
- **Детали:**
  - Видна таблица «Военный комиссариат Республики Крым» с именами (Машуков, Капитанов и др.) и должностями.
  - Пользователь прокручивает вниз до раздела «Министерство внутренних дел по Республике Крым» (МВД), где также указаны имена начальников полиции и их заместителей по районам.

Рис. 3. Фамилия в ответе нейросети

Данные неточности не приводили к искажению общего смысла происходящего и не влияли на итоговую интерпретацию сценария как демонстрационного. Задача показала способность модели формировать семантически согласованное текстовое описание пользовательской активности в графической среде.

3.2 Оценка распознавания и понимая команд терминала в операционной системе на базе ядра Linux

Цель задачи. Оценка способности мультимодальной большой языковой модели распознавать команды, выполняемые в терминале операционной системы на базе ядра Linux, корректно интерпретировать выполняемые действия и формировать предварительную аналитическую интерпретацию, позволяющую классифицировать характер активности как реальный вредоносный либо имитационный (демонстрационный).

Исходные данные. Видеозапись экрана, представленная на рис. 4, демонстрирующая выполнение команд в терминале операционной системы Linux и взаимодействие между двумя вычислительными средами (предположительно в виртуальном окружении). На экране также присутствуют элементы интерфейса установленного средства защиты, позволяющие учитывать контекст тестирования.

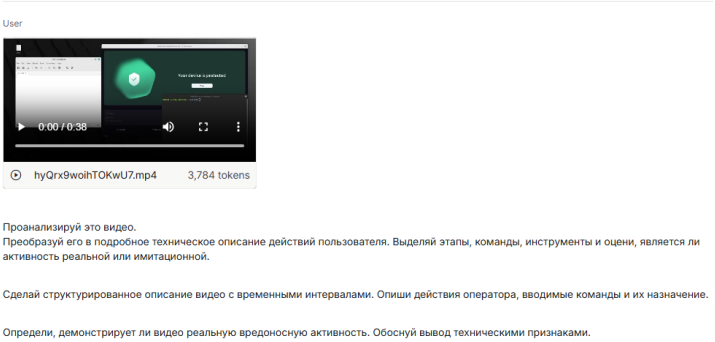


Рис. 4. Прикреплённое видео, содержащее выполняемые пользователем команды в терминале и запрос к нейросети

Интерпретация модели. Модель корректно распознала характер выполняемых команд, их назначение и функциональную роль в рамках сценария удалённого доступа. Существенным результатом является визуально-семантическое распознавание контекста: модель связала отображаемые окна терминала, последовательность действий пользователя и наблюдаемые результаты (установление соединения и выполнение базовых операций подтверждения доступа) в единый логически непротиворечивый сценарий.

Классификация активности. На основании совокупности признаков (демонстрационный контекст, признаки контролируемой/лабораторной среды, отсутствие элементов скрытности и эксплуатационного этапа, характер действий пользователя) активность была классифицирована как имитационная (демонстрационная). При этом модель сформировала аналитическую гипотезу о сценарии удалённого доступа, применимую для принятия решения о дальнейшей экспертной эскалации.

Неточности и вывод. Зафиксирована локальная неточность (рис. 5 и 6) в определении частных параметров интерфейса защитного средства (уровня/версии), не влияющая на интерпретацию последовательности действий и итоговую классификацию активности.

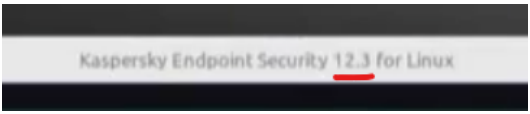


Рис. 5. Версия Kaspersky Endpoint Security для Linux на видео

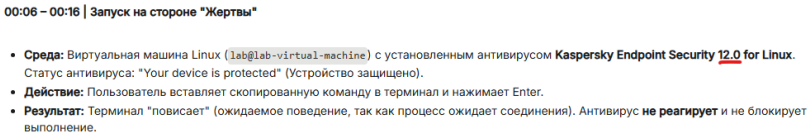


Рис. 6. Версия Kaspersky Endpoint Security для Linux, распознанная нейросетью

Задача подтверждает способность модели понимать происходящее на экране, интерпретировать терминальные действия и различать демонстрационные и потенциально вредоносные сценарии на основании контекстных признаков.

3.3 Интерпретация скриншотов, опубликованных в открытых источниках

Цель задачи. Оценка способности мультимодальной большой языковой модели распознавать и интерпретировать статические изображения, опубликованные в открытых источниках и позиционируемые как подтверждение осуществлённой активности хакерской группировки, представленной на рис. 7, а также классифицировать характер наблюдаемой активности как реальный вредоносный либо имитационный на основе визуально-семантических признаков.

Исходные данные. Набор скриншотов, опубликованных в открытом информационном канале и содержащих элементы интерфейсов корпоративной ИТ-инфраструктуры (административные консоли, файловые структуры серверных систем, окна командной оболочки), а также визуальные признаки нарушения штатного состояния данных и системных объектов (рис. 8 – 11). Анализ выполнялся без дополнительного направляющего запроса, что позволило оценить самостоятельность интерпретации визуальной информации.

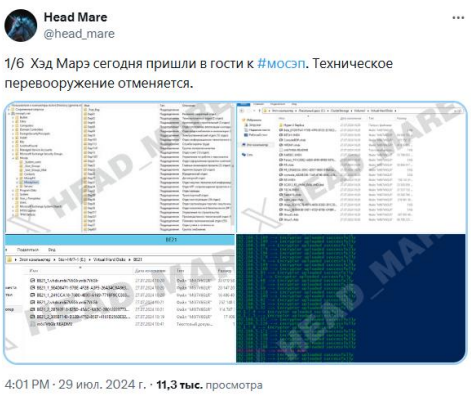


Рис. 7. Сообщение хакерской группировки Head Mare в социальной сети X (Twitter)

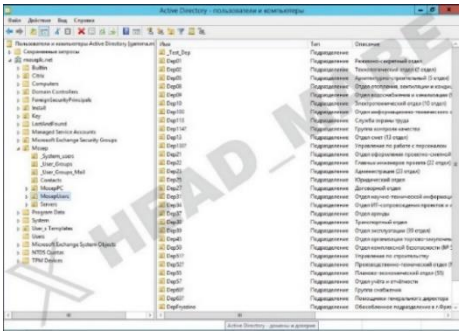


Рис. 8. Скриншот Active Directory

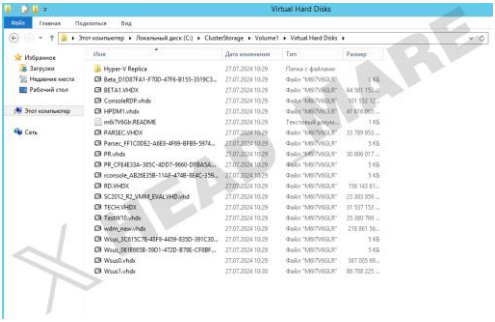


Рис. 9. Результат работы ВПО типа шифровальщик

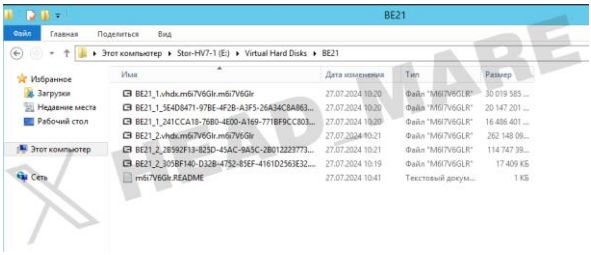


Рис. 10. Результат работы ВПО типа шифровальщик

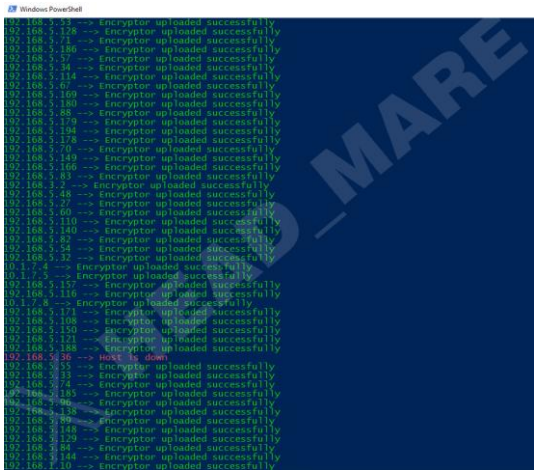


Рис. 11. Демонстрация загрузки ВПО

Интерпретация модели. Модель сформировала целостную интерпретацию набора изображений, корректно идентифицировав типы отображаемых интерфейсов и назначение элементов инфраструктуры. Важным аспектом является согласование отдельных изображений в единый сценарий: модель связала признаки административного доступа, изменения состояния файлов и выполнение действий в командной оболочке как компоненты одного процесса.

Классификация активности. На основании совокупности признаков (масштабность и системность изменений, типовые последствия вредоносного воздействия, согласованность признаков между изображениями, отсутствие явных индикаторов лабораторного стенда) наблюдаемая активность была классифицирована как соответствующая сценарию реальной вредоносной активности. При этом подчёркивается, что в рамках настоящей работы не выполняется атрибуция и не делаются

юридически значимые выводы; классификация относится к характеру наблюдаемой активности, отражённой в визуальных материалах.

Неточности и вывод. Существенных неточностей интерпретации визуального содержимого выявлено не было. Задача демонстрирует способность модели интерпретировать статические изображения инфраструктурного контекста и формировать предварительную классификацию характера активности на основе визуально-семантических признаков.

3.4 Классификация активности в промышленной SCADA-системе

Цель задачи. Оценка способности мультимодальной большой языковой модели распознавать и интерпретировать визуально наблюдаемую активность оператора в промышленной SCADA-системе и классифицировать её как эксплуатационную либо потенциально деструктивную (вредоносную/имитационную) на основании контекста действий пользователя и структуры интерфейса.

Исходные данные. Видеозапись экрана, демонстрирующая работу оператора в SCADA-системе: навигацию по мнемосхемам, мониторинг параметров инженерных подсистем и использование штатных элементов управления. В запросе, представленном на рис.12, к модели использовалось ограничение на генерацию неподтверждённых деталей (требование интерпретировать только наблюдаемое на экране).



Рис. 12. Запись работы инженерной SCADA-системы и запрос к модели

Интерпретация модели. Модель корректно распознала тип интерфейса как SCADA/HMI, интерпретировала назначение мнемосхем и связала действия пользователя с типовыми сценариями мониторинга и подтверждения настроек. Существенным является визуально-семантическое распознавание: модель выделила нормальные и

потенциально аномальные состояния параметров, не выходя за рамки наблюдаемых данных, и при этом избежала генерации утверждений о факте кибератаки при отсутствии соответствующих признаков.

Классификация активности. На основании того, что все действия выполнялись через штатный интерфейс, отсутствовали признаки несанкционированного воздействия (ввод команд, использование внешних инструментов, обход механизмов управления), а наблюдаемое взаимодействие соответствовало эксплуатационному сценарию, активность была классифицирована как эксплуатационная. Наличие пограничных/аномальных параметров интерпретируется как эксплуатационный контекст, который сам по себе не является достаточным основанием для отнесения сценария к вредоносной активности без фиксации целенаправленных деструктивных действий.

Неточности и вывод. Существенных неточностей в интерпретации визуального содержимого не выявлено. Отдельно подчёркивается, что в рамках данной задачи не выполнялась проверка технологической корректности режимов и уставок, поскольку это требует профильной инженерной экспертизы и выходит за рамки настоящего исследования. Задача демонстрирует, что модель может быть использована для первичной интерпретации промышленного интерфейса и предварительной классификации характера наблюдаемой активности в условиях ограниченной предметной экспертизы аналитика.

4. Основные результаты

Результаты, полученные в ходе проведённого исследования, демонстрируют, что мультимодальная большая языковая модель способна эффективно выполнять первичную аналитическую интерпретацию визуальных материалов, публикуемых хакерскими группировками в открытых информационных каналах. Во всех рассмотренных экспериментальных задачах модель корректно распознавала тип визуальной среды и формировала семантически согласованное текстовое описание наблюдаемых действий, опираясь исключительно на визуально представленный контекст.

Принципиально важным результатом является способность модели различать реальную вредоносную активность и имитационные (демонстрационные) сценарии без привлечения внешних источников информации и без явного указания аналитика на характер происходящего. Классификация активности формировалась на основе совокупного анализа визуальных признаков, последовательности действий и контекста среды выполнения, что соответствует реальным условиям аналитической работы при обработке материалов из открытых источников.

Следует отметить, что в отдельных задачах модель допускала локальные неточности в интерпретации частных деталей (например, элементов интерфейса или характеристик программного обеспечения). Однако данные неточности не нарушали логическую целостность интерпретации и не приводили к искажению итоговой классификации характера наблюдаемой активности. Это позволяет сделать вывод о том, что для задач первичной аналитической обработки ключевым фактором является семантическая согласованность и контекстная адекватность интерпретации, а не абсолютная точность отдельных деталей.

Отдельного внимания заслуживают результаты, полученные в сценариях, требующих предметной экспертизы (в частности, при анализе промышленной SCADA-системы). В данных условиях модель продемонстрировала способность формировать осмысленную предварительную интерпретацию происходящего, не выходя за рамки визуально наблюдаемых данных и избегая генерации неподтверждённых утверждений о факте кибератаки. Это подтверждает целесообразность использования модели в качестве инструмента поддержки аналитика в ситуациях, когда профильная экспертиза отсутствует или доступна не на первичном этапе анализа.

Таким образом, полученные результаты позволяют рассматривать мультимодальные большие языковые модели как эффективный инструмент снижения когнитивной нагрузки аналитика, ускорения первичной обработки визуальных материалов и формирования предварительных аналитических гипотез. При этом модель не заменяет эксперта, а выступает связующим звеном между визуальным материалом и последующими этапами экспертного анализа.

Заключение

В рамках настоящей работы проведено экспериментальное исследование применимости мультимодальных больших языковых моделей для автоматизации первичной аналитической интерпретации визуальных материалов, публикуемых хакерскими группировками в открытых информационных каналах. Основное внимание было уделено задаче распознавания и понимания действий, зафиксированных на видеозаписях и скриншотах, а также классификации характера наблюдаемой активности как реальной вредоносной либо имитационной.

В ходе исследования на примере четырёх независимых экспериментальных задач показано, что мультимодальная большая языковая модель Gemini 3 Pro Preview способна корректно интерпретировать визуально наблюдаемую активность в различных средах – от графического интерфейса операционной системы и командной строки до корпоративной ИТ-инфраструктуры и

промышленной SCADA-системы. Модель формирует осмысленные и логически согласованные текстовые описания происходящего, пригодные для использования аналитиком при первичной обработке информации.

Полученные результаты подтверждают, что применение мультимодальных больших языковых моделей позволяет существенно сократить временные затраты аналитика на первичную интерпретацию визуальных материалов и повысить эффективность принятия решений о дальнейшем направлении анализа. При этом подчёркивается, что аналитическая интерпретация, формируемая моделью, носит вспомогательный характер и не заменяет экспертную оценку, а лишь служит инструментом предварительного анализа и поддержки принятия решений.

Ограничения проведённого исследования связаны с отсутствием проверки предметной корректности аналитических выводов модели, что обусловлено необходимостью профильной экспертизы в соответствующих предметных областях. Данное ограничение рассматривается как осознанное методологическое решение, отражающее реальные условия аналитической деятельности при работе с материалами из открытых источников.

В целом результаты работы подтверждают целесообразность использования мультимодальных больших языковых моделей в аналитических подразделениях в качестве элемента систем поддержки аналитической деятельности в области информационной безопасности. Перспективными направлениями дальнейших исследований являются расширение набора анализируемых сценариев, формализация критериев оценки аналитической интерпретации и интеграция подобных моделей в автоматизированные аналитические платформы.

Список литературы

1. Ресурс Telegram [Электронный ресурс]. – Режим доступа: https://t.me/kiberkorpus/566

2. Ресурс X (Twitter) [Электронный ресурс]. – Режим доступа: https://x.com/head_mare/status/1817908298831515838

3. Мультимодальная большая языковая модель Gemini 3 Pro Preview [Электронный ресурс]. – Режим доступа: https://aistudio.google.com/prompts/new_chat?model=gemini-3-pro-preview

4. Мультимодальная большая языковая модель ChatGPT 5.2 [Электронный ресурс]. – Режим доступа: https://chatgpt.com/

5. Мультимодальная большая языковая модель Claude Sonnet 4.5 [Электронный ресурс]. – Режим доступа: https://claude.ai

6. Мультимодальная большая языковая модель DeepSeek-V3.2 [Электронный ресурс]. – Режим доступа: https://www.deepseek.com/en/

Верифицируемая система поддержки принятия решений на основе больших языковых моделей

В. И. Борисенко, email: borisenko.vi.661@suitd.ru[1]
В. И. Пименов, email: v_pim@mail.ru [1]

[1] Санкт-Петербургский государственный университет промышленных технологий и дизайна

Аннотация. *В работе представлена концепция применения больших языковых моделей для системного моделирования бизнес-задач. Детализирована архитектура интеграции больших языковых моделей в процесс принятия решений, которая охватывает этапы предобработки данных, генерации, верификации и адаптации результатов. В качестве основы для повышения когнитивной устойчивости бизнес-аналитики предложен комбинированный подход, использующий технологии RAG и ReAct.*

Ключевые слова: *большая языковая модель, RAG, ReAct, система поддержки принятия решений, когнитивный агент*

Введение

Активное внедрение больших языковых моделей (large language models, LLM) в системы поддержки принятия решений (СППР), открывает новые возможности для бизнес-аналитики [1]. Однако такие системы в корпоративной среде сталкивается с недостаточной надёжностью и интерпретируемостью выводов, генерируемых LLM.

Проблема «галлюцинаций» LLM, то есть генерации правдоподобной, но фактически некорректной или неподтверждённой информации, остаётся критическим вызовом. Таким образом, возникает противоречие между креативным и контекстно-зависимым потенциалом LLM и жёсткими требованиями бизнес-среды к обоснованности и проверяемости принимаемых решений. Отсутствие прозрачных механизмов объяснения «чёрного ящика» LLM может снижать доверие конечных пользователей, таких как менеджеров и аналитиков, и, тем самым, препятствовать полномасштабному внедрению.

1. Анализ существующих подходов к обеспечению достоверности и интерпретируемости LLM

Повышение доверия к выводам LLM является активной областью исследований, что породило множество методов, которые условно можно

© Борисенко В. И., Пименов В. И., 2026

разделить на два класса: архитектурные решения, дополняющие модель внешними модулями, и методологии инструкций к LLM, структурирующие её внутренние рассуждения.

1.1. Архитектурные подходы

Архитектура, интегрирующая подход генерации, дополненной поиском (retrieval-augmented generation, RAG), стала стандартом для снижения «галлюцинаций» за счёт привязки генерации к внешним документам. Как отмечается в работе [2], RAG эволюционировал от «наивных» архитектур к сложным системам с интегрированными механизмами ранжирования, что позволило повысить точность ответов в задачах, требующих фактологических знаний.

Основное преимущество RAG – возможность актуализации знаний системы без дорогостоящего переобучения модели, а также обеспечение связанности выводов до исходных документов.

Однако критический анализ показывает, что RAG в его текущем виде не позволяет решить проблему «галлюцинаций» LLM. Во-первых, его эффективность жёстко зависит от качества и актуальности релевантного корпуса данных. Если база знаний неполна, устарела или содержит противоречащую информацию, генеративная часть модели может либо усилить существующие ошибки, либо, пытаясь их разрешить, создать новую непроверенную информацию. Во-вторых, механизм извлечения часто оптимизирован под поверхностную лексическую схожесть, а не под глубокую семантическую релевантность, что приводит к пропуску контекстуально важных, но лексически отличных документов. В-третьих, существует фундаментальная проблема интеграции извлечённого контекста с внутренними знаниями модели. LLM может проигнорировать предоставленные доказательства в пользу собственных, потенциально ошибочных, знаний, особенно если контекст противоречит её сильным внутренним представлениям. Это подтверждается исследованиями, показывающими, что даже при наличии релевантных извлечений модели могут генерировать утверждения, не следующие напрямую из предоставленных источников.

Таким образом, RAG обеспечивает необходимую, но недостаточную основу для доверия. Он решает проблему доступа к знаниям, но не гарантирует их корректного использования и интерпретации моделью в рамках конкретной бизнес-логики.

1.2. Методологии структурированных рассуждений

Для улучшения внутренней прозрачности и логической последовательности выводов были предложены методы структурированных рассуждений. Метод цепочки рассуждений (chain-of-

thoughts, CoT) требует от LLM генерации промежуточных логических шагов перед финальным ответом. В работе [3] подчёркивает, что CoT не только улучшает выполнение сложных задач, но и делает процесс принятия решений модели более интерпретируемым для человека. Однако CoT по своей сути является «чёрным ящиком внутри чёрного ящика». Он делает видимым ход мыслей модели, но не предоставляет механизмов проверки истинности каждого из этих шагов. Модель может сгенерировать безупречно выглядящую логическую цепочку, основанную на ложной предпосылке или «галлюцинированном» факте.

Парадигма с синергией размышления и действия (reasoning and acting, ReAct) представляет собой шаг вперёд, объединяя генерацию рассуждений с выполнением действий, таких как запросы к внешним API или базам данных. Это позволяет модели итеративно уточнять свои гипотезы на основе поступающей извне информации, что критически важно для бизнес-аналитики, где решение часто требует проверки предположений против реальных данных. В исследовании [4] демонстрирует, что ReAct снижает ошибки, вызванные «галлюцинациями», в задачах проверки фактов, поскольку модель может запросить внешний источник для подтверждения. Однако и у ReAct есть существенные ограничения для корпоративного применения. Во-первых, он вводит значительные задержки из-за необходимости многократных циклов «рассуждение-действие-ожидание ответа». Во-вторых, надёжность всего процесса становится зависимой от надёжности и доступности внешних сервисов (CRM, ERP, BI-систем). В-третьих, ReAct не решает проблему семантической интеграции. Модель получает сырые данные, например, числовой показатель из отчёта, но её способность корректно интерпретировать этот показатель в контексте конкретного бизнес-процесса остаётся на её усмотрение и может быть ошибочной.

1.3. Специализированные методы верификации

Современные подходы к проверке входных данных, как описано в работе [5], включают извлечение утверждений из текста LLM, сбор доказательств из доверенных источников (внешние базы знаний, онтологии) и оценку согласованности между утверждением и доказательствами. Эти методы эффективны для проверки конкретных фактов (дат, имен, статистических данных), но сталкиваются с трудностями при верификации сложных, контекстно-зависимых бизнес-утверждений или прогнозов, истинность которых невозможно установить по текущим данным.

Семантические методы, такие как оценка смысловой схожести, используются для проверки согласованности сгенерированного текста с предоставленным контекстом или бизнес-правилами. В работе [6],

показывают, что современные LLM могут эффективно выполнять задачи оценки смысловой схожести, особенно когда для оценки требуется фоновое знание о мире. В контексте СППР это может быть использовано для проверки, соответствуют ли рекомендации модели корпоративным наборам правил или историческим прецедентам.

Главный недостаток этих методов заключается в их фрагментарность и реактивность. Проверка входных данных работает постфактум, проверяя уже сгенерированный текст, а не предотвращая ошибки в процессе генерации. Семантический анализ часто опирается на общие метрики схожести (косинусное расстояние эмбеддингов), которые плохо адаптируются к специфике предметной области и могут не улавливать важные для бизнеса логические связи или нарушения правил.

1.4. Системные ограничения и пробелы

Проведённый анализ позволяет выявить ключевые системные пробелы, препятствующие построению доверенных LLM-ориентированных СППР:

− Отсутствие интеграции с доменными онтологиями и бизнес-логикой. Большинство рассмотренных методов работают на уровне текста или общих знаний. Они не имеют встроенного механизма для проверки соответствия выводов модели формальным бизнес-правилам, процессным моделям или отраслевым онтологиям. Рекомендация может быть фактологически верной и логически последовательной, но при этом нарушать внутренний регламент или быть невыполнимой в рамках существующих процессов.

– Фрагментарность и отсутствие сквозного контроля. Существующие подходы (RAG, проверка входных данных, семантический анализ) часто применяются изолированно или последовательно, но не в рамках единого сквозного контура верификации. Это приводит к ситуации, когда ошибка, пропущенная на одном этапе (например, RAG предоставил нерелевантный документ), не может быть исправлена на последующих.

– Слабая адаптивность к динамике предметной области. Бизнес-среда характеризуется постоянными изменениями: появляются новые продукты, меняются регуляторные требования, оптимизируются процессы. Традиционные методы верификации требуют ручного обновления правил и корпусов данных, что делает систему инерционной и отстающей от реальности.

– Недостаточная интерпретируемость для конечного пользователя. Даже такие методы, как CoT и ReAct, генерируют объяснения на естественном языке, которые могут быть многословными, избыточными или сложными для восприятия бизнес-аналитиком. Отсутствует

стандартизированный способ представления обоснований решений, увязанных с конкретными источниками данных, этапами рассуждений и бизнес-метриками.

Таким образом, существует необходимость в гибридном подходе, который бы интегрировал сильные стороны рассмотренных технологий – способность RAG работать с внешними знаниями, итеративность ReAct, структурированность CoT, потенциал методов проверки входных данных – в единую архитектуру, дополненную механизмами семантической привязки к доменной онтологии и бизнес-процессам. Такой подход должен обеспечить не разовую проверку, а сквозной, многоуровневый контроль достоверности и генерацию целевых объяснений на всех этапах работы LLM в цикле поддержки принятия решений.

2. Предлагаемая архитектура системы решения бизнес-задач на основе больших языковых моделей

Основываясь на выявленных в предыдущем разделе системных ограничениях существующих методов, предлагается целостная архитектура модели, предназначенной для применения LLM в задачах бизнес-аналитики, прогнозирования и принятия решений. Модель описывает взаимодействие внутренних компонентов модели друг с другом и с внешними системами предприятия, формируя движение от данных к практическим действиям.

2.1. Структура системы и внешняя среда

Модель описывает предприятие как открытую динамическую среду. В неё интегрированы внутренние корпоративные системы (CRM, ERP, BI, журналы событий) и внешние потоки данных (социальные сети, новости, аналитика рынка). Эти внешние источники формируют поток сигналов, поступающих на вход интеллектуального модуля, построенного на основе LLM.

Функциональная архитектура системы моделирования реализована в виде комплекса взаимосвязанных модулей, к которым относятся:

- модуль предобработки;
- модуль LLM-генерации гипотез;
- модуль верификации и оценки;
- модуль коррекции и адаптации;
- модуль визуализации и принятия решений.

На рисунке схематически представлена архитектура системы, иллюстрирующая сквозной процесс от получения исходных данных до формирования управленческих решений, а также ключевые функциональные модули и схему их взаимодействия.

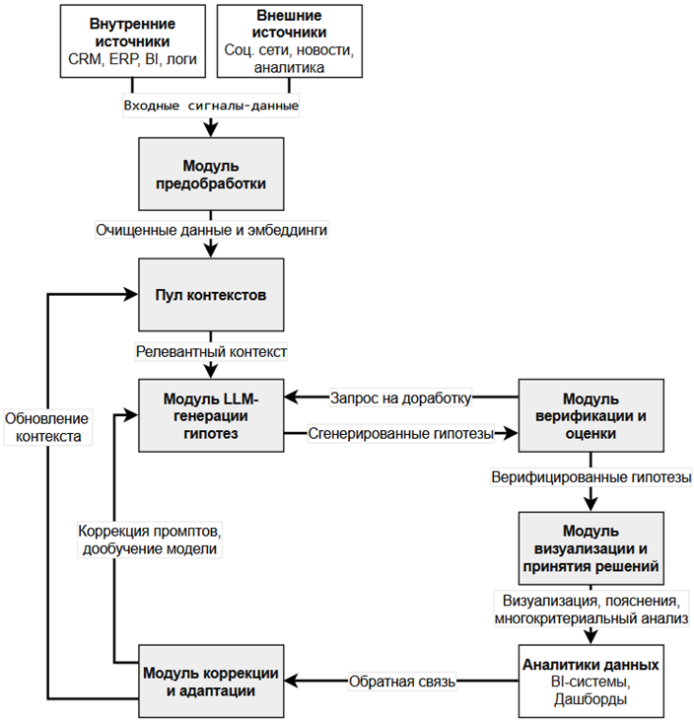


Рисунок. Концептуальная модель решения бизнес-задач на основе LLM

2.2. Модуль предобработки

Модуль предобработки отвечает за очистку, разбиение текста на более мелкие части (токенизация), нормализацию, извлечение признаков и преобразование текстов в векторные представления (эмбеддинги) с последующей кластеризацией:

$$D = \{d_{int} \cup d_{ext}\} \rightarrow [\text{Tokenization}] \rightarrow T = \{t_1, t_2, \ldots, t_n\}$$

$$T \rightarrow [\text{Normalization}] \rightarrow T' = \{t_1', t_2', \ldots, t_m'\}$$

$$T' \rightarrow [\text{Vectorization}] \rightarrow E = \{e_1, e_2, \ldots, e_m\},\ e_i \rightarrow R^d$$

$$E \rightarrow [\text{Clasterization}] \rightarrow C = \{C_1, C_2, \ldots, C_k\}$$

где D – исходные данные (внутренние d_{int} и внешние d_{ext}), T – токенизированные данные, T' – нормализованные токены, E –

эмбеддинги в d-мерном пространстве, C – кластеры семантически близких элементов.

2.3. Пул контекстов

Кластеризованные эмбеддинги используются для создания структурированного пула контекстов. Для целей верификации каждому элементу в пуле (документу или его фрагменту) помимо семантического вектора присваиваются метаданные, например, источник, временная метка, а также предварительно вычисленный тег доверия. Тег доверия инициализируется на основе источника данных, например, внутренняя транзакция из ERP имеет высокий тег, новостная статья – средний, пользовательский отзыв в социальной сети – низкий. На основе частоты успешных ссылок и обратной связи тег доверия может динамически корректироваться в модуле коррекции и адаптации. Таким образом, на вход модулю генерации гипотез поступает не просто набор релевантных текстовых фрагментов, а семантически структурированный, ранжированный по доверию контекст.

2.4. Модуль LLM-генерации гипотез

На этапе генерации LLM используют в качестве когнитивного агента системы. LLM принимает подготовленные входные данные, состоящие из инструкции и релевантного контекста, извлечённого из пула, и на их основе формирует гипотезу. Работа модуля строится по принципу ReAct, где каждое действие «думай» (reasoning) и «действуй» (act) сопровождается автоматической внутренней проверкой. Получив задачу, он не генерирует единый поток CoT, а разбивает его на элементарные проверяемые шаги. Перед тем как принять это утверждение как часть рассуждения, обязательно инициируется запрос к контекстному пулу и внешним источникам, например, к мониторингу рынка, для проверки на соответствие сгенерированного шага с извлечённым фактом. Только шаг, прошедший эту проверку, становится частью финальной цепочки рассуждений. Шаги, не прошедшие проверку, либо переформулируются, либо помечаются как «неподтверждённые предположения».

Процесс построения гипотез средствами LLM можно формализовать следующим способом:

$$G = \text{LLM}(P \oplus C_r) \rightarrow H = \{h_1, h_2, ..., h_n\}$$

где P – инструкция к LLM (входной запрос), C_r – релевантный контекст из пула эмбеддингов, $\oplus$ – операция конкатенации, LLM() – функция вызова LLM, H – множество сгенерированных гипотез.

2.5. Модуль верификации и оценки

После генерации гипотезы проходят этап оценки качества и достоверности. В качестве примера этапы можно представить в виде двухуровневой системы контроля.

На первом уровне проверяется синтаксико-семантическая согласованность. Данный этап направлен на оценку внутренней непротиворечивости сгенерированного текста и его соответствия онтологии предметной области. Нарушение онтологических правил, например, предложение изменить несуществующий атрибут процесса, ведёт к отбраковке или коррекции гипотезы.

На втором уровне проверяется фактологическая обоснованность с использованием подходов RAG и методов проверки входных данных. Для каждого ключевого утверждения в сценарии система автоматически извлекает цитаты-обоснования из контекстного пула, использованного при генерации, с учётом тега доверия к источнику. Затем выполняется перекрёстный запрос к доверенным внешним источникам, не задействованным в первоначальном контексте, таким как актуальные регуляторные базы данных или индексы фондового рынка, что позволяет выявить расхождения с авторитетной информацией. По итогам проверки каждому утверждению присваивается количественная оценка достоверности, вычисляемая как функция от релевантности, свежести и уровня доверия к источникам-обоснованиям.

На третьем уровне проверяется логико-процессная выполнимость и соответствие. На данном этапе сгенерированный сценарий конвертируется в набор входных параметров для цифрового двойника соответствующего бизнес-процесса, после чего запускается симуляция гипотезы. Она оценивает возможные последствия реализации сценария по ключевым метрикам. Сценарий, приводящий к сбою симуляции или выходу метрик за допустимые границы, отправляется на доработку. Дополнительно система проверяет, не нарушает ли рекомендация установленных регламентов, выявляя конфликты путём сопоставления с эталонными моделями процессов.

2.6. Модуль визуализации и принятия решений

Выходом модифицированной системы является не только гипотеза, прошедшая многоуровневый контроль, но и структурированное объяснение, автоматически формируемое на основе данных, собранных на всех этапах работы. Структура этого объяснения включает несколько ключевых компонентов. Во-первых, это очищенная цепочку проверенных шагов. Во-вторых, список ключевых утверждений с привязанными к ним источниками из контекстного пула и соответствующими оценками достоверности. В-третьих, онтологическая карта, которая визуализирует

связи ключевых сущностей гипотезы в рамках доменного графа. Наконец, в объяснение интегрируются результаты симуляции, полученные при запуске сценария через цифрового двойника.

Данное структурированное объяснение интегрируется в модуль визуализации и принятия решений. Конечный пользователь, такой как менеджер или аналитик, видит не только итоговую рекомендацию, но и интерактивную панель, позволяющую детально раскрыть любой её аспект. Это даёт возможность просмотреть исходные данные, проверить логику рассуждений, оценить прочность аргументации и спрогнозированные последствия. Таким образом, система трансформирует LLM из простого источника ответов в инструмент для совместного анализа, где окончательное решение принимается человеком на основе максимально полной, проверенной и интерпретируемой информации.

2.7. Модуль коррекции и адаптации

Модуль коррекции и адаптации аккумулирует обратную связь по всем отклонённым и принятым сценариям. Если пользователь отвергает рекомендованный системой вариант, это фиксируется как отрицательная обратная связь. Система анализирует, на каком уровне контроля не было выявлено противоречий, которое оказалось значимым для человека. Эта информация используется для корректировки онтологии (добавление новых правил или связей), настройки порогов срабатывания проверок на разных уровнях и обновления тегов доверия источников в контекстном пуле.

Заключение

Таким образом, архитектура замыкает петлю постоянного улучшения, где каждый цикл принятия решения способствует уточнению и специализации самой системы верификации, повышая её адекватность конкретной бизнес-среде. Данный подход позволяет преодолеть выявленные ранее ограничения фрагментарности, слабой адаптивности и недостаточной интерпретируемости, создавая основу для доверенного и эффективного симбиоза человеческого интеллекта и когнитивных способностей больших языковых моделей в корпоративной практике.

Список литературы

1. Борисенко, В. И. Генерации синтетических наборов данных с использованием больших языковых моделей для решения бизнес-задач / В. И. Борисенко, В. И. Пименов // ПРОМТЕХДИЗАЙН. Естественные и технические науки : Сборник статей всероссийской научной

конференции молодых ученых с международным участием. Часть 3 (СПбГУПТД, 1-27 апреля 2025 г.). – Санкт-Петербург, 2025. – С. 40-42.

2. Retrieval-Augmented Generation for Large Language Models: A Survey [Электронный ресурс] / Y. Gao [et al.]. – 2024. – Режим доступа : https://arxiv.org/abs/2312.10997v5

3. Towards Better Chain-of-Thought Prompting Strategies: A Survey [Электронный ресурс] / Z. Yu [et al.]. – 2023. – Режим доступа : https://arxiv.org/abs/2310.04959

4. ReAct: Synergizing Reasoning and Acting in Language Models [Электронный ресурс] / S. Yao [et al.]. – 2023. – Режим доступа : https://arxiv.org/abs/2210.03629v3

5. Factuality of Large Language Models: A Survey [Электронный ресурс] / Y. Wang [et al.]. – 2024. – Режим доступа : https://arxiv.org/abs/2402.02420v3

6. Text Encoders Lack Knowledge: Leveraging Generative LLMs for Domain-Specific Semantic Textual Similarity [Электронный ресурс] / J. Gatto [et al.]. – 2023. – Режим доступа : https://arxiv.org/abs/2309.06541

Применение методов машинного обучения для прогнозирования негативных факторов в системах распределения энергии

М.А. Гришкин, email: mbabul686868@mail.ru[1]

[1]Тамбовский государственный технический университет

Аннотация. *В статье рассматривается актуальная проблема повышения стабильности и безопасности систем распределения энергии за счет внедрения методов машинного обучения. В мире постепенной деградации инфраструктуры, увеличения доли нестабильного производства из возобновляемых источников энергии и усложнения управления сетью традиционные методы мониторинга и диагностики становятся неэффективными. Исследование сосредоточено на возможности использования алгоритмов машинного обучения для решения двух основных задач: прогнозирования вероятности возникновения негативных факторов (перегрузки, отказа устройства, колебаний мощности) и их классификации практически в режиме реального времени. Результаты исследования показывают, что интеллектуальные системы, основанные на машинном обучении, могут стать основой для перехода от гибкого и адаптивного управления рисками, что имеет решающее значение для обеспечения надежности и эффективности современных энергетических систем*

Ключевые слова: *машинное обучение, негативные факторы, теплоснабжение, распределение энергии, интеллектуальные системы*

Введение

Современные энергетические системы сталкиваются с проблемой перехода от моделей реактивного управления к моделям интеллектуального прогнозирования. Эта трансформация особенно сложна в контексте системного кризиса, распространения энергии и, в особенности, ее нагрева. Создается технологическая проблема, проявляющаяся в крайнем износе инфраструктуры и хронически высоких потерях, усугубляется экономической нестабильностью и организационными проблемами, создавая среду с постоянно высоким операционным риском. При традиционном подходе в прошлом практически не уделялось внимания рутинному планированию и реагированию на большие аварийные ситуации, которое зарекомендовало себя как комплексное в области информационных технологий. В этих

© Гришкин М.А., 2026

случаях приоритетной задачей является не только разработка новых инструментов, но и изменение акцента на прогнозирование появления негативных факторов и предотвращение их последствий.

Один из способов сделать это — использовать инструменты машинного обучения. Его главным преимуществом в этом контексте является способность выявлять сложные закономерности и корреляции, которые часто служат предвестниками чрезвычайных ситуаций, а также способность анализировать большое количество разнообразных оперативных данных. Это позволяет нам перейти от управления регулирования к управлению реальной ситуации и предотвращения, что необходимо для решения изнурительных проблем, связанных с работой в необъяснимых и часто опасных режимах. Таким образом, применение машинного обучения для прогнозирования негативных факторов является не просто техническим усовершенствованием, но и становится необходимым механизмом обеспечения базовой промышленной безопасности и стабильности систем электроснабжения в условиях ограниченных ресурсов и высокой неопределенности. Данная статья посвящена исследованию возможностей этих методов, анализу основных классов решаемых задач и рассмотрению способов их интеграции в существующие технические и управленческие схемы отрасли.

1. Негативные факторы

Актуальность разработки и внедрения методов машинного обучения и классификации негативных факторов в отечественных системах распределения энергии обусловлена глубоким системным кризисом, охватившим, в частности, сферу теплоснабжения. Кризис носит комплексный характер, где технологическая деградация, экономическая неустойчивость и организационные просчеты взаимно усиливают друг друга, создавая среду с хронически высоким уровнем операционных рисков.

Фундаментальной предпосылкой выступает критическое состояние материально-технической базы [1-3]. Инфраструктура объектов теплоснабжения характеризуется экстремальной степенью износа: по различным оценкам, около 29 % тепловых сетей требуют срочной замены, а 22 % уже перешли в категорию полностью ветхих. Существующие темпы реконструкции, не превышающие 1.9 – 2.0 % в год, оказываются вдвое ниже минимально необходимых для стабилизации ситуации 4 – 5 %, что свидетельствует не о развитии, а о медленной физической деградации активов. Это напрямую отражается на эффективности: фактические потери энергии в сетях, официально оцениваемые в 12 – 13 %, по экспертным данным могут достигать 20 – 30 %. Параллельно низкий коэффициент использования установленной

мощности генерирующего оборудования (в среднем около 10 – 11 %) указывает на его постоянную работу в нерасчетных и заведомо неэффективных режимах, что лишь ускоряет износ оборудования и повышает вероятность его отказов [1-3]. Таким образом, объективная реальность эксплуатации данных объектов радикально расходится с нормативными требованиями к надежности и безопасности [4].

Экономическая модель отрасли не способствует разрешению данного противоречия. Сектор теплоснабжения демонстрирует устойчивую убыточность, что исключает возможность формирования внутренних источников для капитальной модернизации. Годовой объем инвестиций остается на порядок ниже экспертных оценок необходимых вложений, а с учетом инфляции реальный их рост отсутствует. Действующая затратная модель тарифообразования, в структуре которой амортизационная составляющая не превышает 7 %, закрепляет этот порочный круг, лишая отрасль финансовых механизмов для технологического обновления. В условиях хронического недофинансирования поддержание работоспособности системы неизбежно происходит за счет отложенного ремонта и нарастания объемов износа [3].

Наложение организационно-управленческих и кадровых проблем на ветхую инфраструктуру и дефицит ресурсов создает накопительный эффект, многократно умножающий риски. Отрасль сталкивается с перманентным дефицитом квалифицированного инженерно-технического персонала [1-3], способного эксплуатировать сложные объекты в нештатных условиях. Широкое распространение устаревших, слабо автоматизированных систем контроля и управления не позволяет перейти от реагирования на аварии к их профилактике. Унаследованная от предыдущих эпох неоптимальная структура генерирующих и сетевых мощностей приводит к их постоянной работе в аварийно-опасных режимах, предъявляя завышенные требования как к оборудованию, так и к персоналу.

Логическим следствием этих проблем (технической, экономической, организационной) является устойчиво высокий уровень аварийности. Статистика указывает, что наиболее серьезные инциденты концентрируются в зонах повышенной технологической опасности – машинных и котельных отделениях, где ключевым фактором риска выступают системы топливоподачи. При этом внешние экономические шоки, такие как резкий рост цен на традиционные энергоносители, могут провоцировать попытки компенсировать убытки за счет использования более дешевых, но менее изученных видов топлива или сознательного

нарушения регламентов эксплуатации, что дополнительно повышает вероятность катастрофических событий.

Таким образом, современное состояние отрасли представляет собой нестабильный механизм, где традиционные управленческие подходы исчерпали свой потенциал. Интеллектуальные системы на основе методов машинного обучения и анализа данных способны решать задачи адаптивного управления в условиях энергоёмких процессов [5], что подтверждается исследованиями в области применения интеллектуальных методов анализа данных для повышения качества решений в управлении ресурсами информационных систем [5]. Разработка и внедрение нейросетевых технологий для прогнозирования и классификации негативных факторов становится необходимым механизмом для обеспечения базовой промышленной безопасности.

2. Методы машинного обучения

Одной из важнейших задач является прогнозирование вероятности выхода оборудования из строя и аварийных ситуаций. Для решения этой задачи используются алгоритмы классификации, такие как градиентный бустинг (CatBoost) и случайный лес, обученные на основе исторических данных о воздействии, температурных условиях, вибрациях и прошлых событиях [6], которые предварительно должны быть обработаны, выполняя следующие задачи: сжатие, поиск, восстановление, преобразование [7]. Эти модели способны точно оценить риск возникновения того или иного вида неисправности (например, в системах подачи топлива или в районах с максимальным износом тепловых сетей) за несколько часов или даже дней до ее возможного возникновения. Повторяющиеся нейронные сети, особенно с долговременной памятью (LSTM), особенно эффективны при анализе временных рядов датчиков, которые регистрируют долгосрочные зависимости и данные и прогнозируют процесс деградации устройства [8]. Это позволяет планировать профилактические меры, минимизировать незапланированные простои и сократить количество отложенных ремонтов.

В условиях работы значительной части генераторного и сетевого оборудования в нерасчетных, заведомо неэффективных режимах становится очень важным выявлять аномалии и нештатные ситуации в режиме реального времени. Именно здесь на первый план выходят неконтролируемые методы обучения, такие как изолирующий лес и автоэнкодеры. Эти алгоритмы, обученные на основе данных о работе системы, способны обнаруживать даже небольшие отклонения в показаниях тысяч датчиков, которые могут быть неочевидны транспортному персоналу [9]. Выявленные аномалии, будь то

нетипичный перепад давления, увеличение потерь в определенном сегменте сети или неоптимальное состояние сгорания, служат ранним сигналом о развитии проблемы, что позволяет нам перейти от решения последствий к предохранению.

Для снижения фактических потерь энергии, которые в некоторых случаях достигают 20 – 30 %, и оптимизации нагрузки на сеть используются методы прогнозирования спроса и адаптивного управления нагрузкой. Модели, основанные на ансамблях деревьев решений или глубоких нейронных сетях и множестве факторов (время суток, день недели, погодные условия, календарь событий), позволяют получать точные прогнозы потребления на разные периоды времени [8]. Основываясь на этих прогнозах, системы, основанные на обучении с подкреплением, могут автоматически рекомендовать оптимальные конфигурации потоков энергии, схемы аварийного питания или параметры управления, принимая во внимание текущее состояние сети. Это помогает увеличить среднюю загрузку установленной мощности и снизить пиковую нагрузку, приводящую к ускоренному износу.

Экономическая модель отрасли, характеризующаяся недостатком инвестиций, требует наиболее целесообразного и эффективного распределения ограниченных средств на модернизацию. В связи с этим методы машинного обучения используются для создания приоритетных программ плановой профилактики и замены оборудования. Благодаря объединению моделей вероятности ошибок прогнозирования с данными о критичности системы, количестве недочетов, сроке службы и стоимости замены, алгоритмам ранжирования. Это позволяет переключить фокус инвестиций с затрат на реагирование в чрезвычайных ситуациях на целенаправленные меры, обеспечивая тем самым максимальный прирост надежности в единицу времени.

Наконец, методы интерпретируемого машинного обучения (интерпретации прогнозов (SHAP), прогноз сложной модели (LIME)) и кластеры (K-средних (K-means), с шумом на основе плотности (DBSCAN)) используются для анализа и классификации инцидентов, которые уже произошли из-за нехватки персонала и поддержки персонала при эксплуатации сложных объектов [10]. Эти инструменты не только автоматически классифицируют аварии по типу и причине, но и объясняют, какие факторы (технологические, эксплуатационные и внешние) внесли наибольший вклад в возникновение чрезвычайной ситуации [11-12]. Результаты проведенных анализов становятся основой для корректировки нормативных актов, разработки целевых программ обучения сотрудников и совершенствования операционных процессов.

Заключение

Машинное обучение способно решать важные оперативные и стратегические задачи, от прогнозирования неисправностей оборудования и обнаружения аномалий в режиме реального времени до оптимизации нагрузки и планирования ремонта. Использование алгоритмов классификации, нейронных рекуррентных нейронных сетей, неконтролируемых методов обучения и обучения с подкреплением формирует интеллектуальный цикл, способный анализировать большие объемы данных, выявлять скрытые зависимости и предлагать обоснованные решения в условиях высокой неопределенности.

Внедрение таких систем не только помогает сократить количество аварий и сократить потери энергии, но и создает основу для рационального распределения ограниченных финансовых ресурсов и смещает акцент с управления стихийными бедствиями на предотвращение. Это особенно важно в ситуациях, когда темпы физического уничтожения активов превышают темпы их традиционного использования.

Таким образом, интеграция методов машинного обучения в системы распределения энергии представляет собой стратегический шаг для создания устойчивой, адаптируемой и безопасной энергетической инфраструктуры. Дальнейшие исследования должны быть направлены на разработку гибридных моделей, сочетающих физические принципы систем с возможностями искусственного интеллекта, а также стандартизированных платформ для создания их реализации в условиях существующих технологических и организационных ограничений. Только так можно преодолеть системный кризис и перейти к интеллектуальной и устойчивой модели энергоснабжения будущего.

Список литературы

1. Есаян, Л.Н. Современные проблемы теплоснабжения городов и рациональные пути их решения // Актуальные исследования. – 2021. – №28. – С. 13.
2. Сазонов, В.С., Савельев, Д.В. Анализ причин возникновения аварий на теплоэлектростанциях // Меридиан. – 2022. – С. 1.
3. Терентьева, А.С. Анализ основных проблем централизованного теплоснабжения в России на современном этапе // Научные труды: Институт народнохозяйственного прогнозирования РАН. – 2020. – С. 253-273.
4. Приказ "Минтруда России от 17.12.2020 N 924н "Об утверждении Правил по охране труда при эксплуатации объектов теплоснабжения и теплопотребляющих установок"" от 17.12.2020 № 924

// Официальный интернет-портал правовой информации. – 2022 г. – с изм. и допол. в ред. от 29.12.2020.

5. Шабанов, К.Б., Алексеев, В.В. Применение методов интеллектуального анализа данных для повышения качества принятия решений при управлении ресурсами информационной медиасистемы // Вестник Тамбовского Государственного Технического Университета. – 2021. – №1. – С. 14-19.

6. Айдарова, А.Ш. Применение алгоритма CatBoost для прогнозирования долгосрочной прибыльности клиентов в e-commerce // Актуальные исследования. – 2024. – №40

7. Обухов, А. Д., Патутин, К. И., Назарова, А. О. Алгоритмы обработки данных в автоматических системах управления на основе компьютерного зрения // Вестник Тамбовского государственного технического университета. – 2022. – №4. – С. 573-585.

8. Xinlin, T., Weiwei, C. Production scheduling problem under peak power constraint // IEEE Sustainable Power and Energy Conference (iSPEC). – 2020. – С. 2083-2088.

9. Фомина, Е.Е. Использование алгоритма Random Forest для обработки социально-экономических данных // Вестник ПНИПУ. – 2022. – №1. – С. 142-150.

10. Краснянский, М. Н., Обухов А. Д., Соломатина, Е. М., Воякина, А. А. Сравнительный анализ методов машинного обучения для решения задачи классификации документов научно-образовательного учреждения // Вестник ВГУ. – 2018. – №3. – С. 173-182.

11. Осипова, Ю.А., Лавров, Д.Н. Применение кластерного анализа методом К-средних для классификации текстов научной направленности // Математические структуры и моделирование. – 2017. – №3. – С. 108-121.

12. Гринева, Н.В., Иванов, И.Д., Аблязина, Н.Х. Интерпретируемая модель машинного обучения для задачи геоаналитики: моделирования размещения торговых точек розничной сети // Инновации и инвестиции. – 2023. – №6. – С. 327-331.

Структура системы управления коллаборативным роботом-манипулятором для свободного манипулирования объектами внешней среды

К. А. Калушев, email: konstantin.kalushev@gmail.com[1]

[1]Национальный исследовательский ядерный университет «МИФИ»

Аннотация. *Ключевой задачей пути воплощения в реальность концепции универсального сервисного работа общего назначения выступает создание системы управления, позволяющей машине свободно манипулировать объектами открытого мира. Рассматривается структура системы управления коллаборативным роботом-манипулятором для произвольного взаимодействия с объектами внешнего мира. Предложенная в работе система состоит из пяти основных модулей и пяти вспомогательных, создаваемых как посредством дообучения нейросетевых моделей, так и на основании алгоритмической архитектуры. Новизна предложенного подхода заключается во включение в состав системы алгоритмического модуля планирования движений робота и модуля контроля выполнения задания с двумя уровнями контроля – за счёт использования датчиков очувствления схвата робота и семантического оценки рабочей сцены с использованием большой визуально-языковой модели.*

Ключевые слова: *Коллаборативный робот, робот-манипулятор, система управления, Vision-Language-Action Model (VLAM), манипулирование объектами.*

Введение

Человечество всегда мечтало о создании универсального сервисного робота общего назначения (англ. – general purpose service robot). Данная концепция подразумевает машину, способную выполнить любое поручение человека на уровне бытовых задач (в реальном окружении обычной среды деятельности человека, например, на кухне). Для этого универсальный сервисный робот должен быть наделен своего рода «бытовым» общим искусственным интеллектом. Такая машина должна интегрировать целый ряд умений, как это описано, например, в [**Ошибка! Источник ссылки не найден.**], включая навигацию, комплексное восприятие окружающего пространства, манипулирование объектами и интерактивное взаимодействие с оператором. Правилами известных робототехнических соревнований RoboCup@Home,

© Калушев К. А., 2026

предусматривается отдельный блок для роботов общего назначения [1], целью которого является демонстрация возможности понять и выполнить команды оператора, что требует от машины широкого спектра различных навыков.

Следует отметить, что в отечественной литературе представлено понятие «интерактивный коллаборативный робот» (ИКо-робот) [3], под которым понимается кобот, самостоятельно выполняющий различные задачи по указаниям оператора, причем ИКо-робот сохраняет активное взаимодействие с оператором в процессе выполнения задач. Понятие «интерактивный коллаборативный робот» относится к форме взаимодействия человека и робота, которая приобретает характер речевого диалога [4]. Вместе с тем, при всей видимой схожести, концепция сервисного робота общего назначения сущностно отличается от ИКо-робота именно дженерализмом и шириной спектра решаемых задач, в то время как содержание понятия ИКо-робот подразумевает скорее подход к взаимодействию с человеком на базе диалога, хотя бы и для решения какой-то одной отдельно взятой задачи.

Ключевым навыком сервисного робота общего назначения выступает свободное манипулирование объектами внешней среды (англ. – open-world robotic manipulation), осуществляемое как по прямым командам оператора, так и в результате пошагового выполнения задачи высокого уровня после ее декомпозиции в субзадачи. Применение традиционной коллаборативной робототехники принципиально ограничено для этого самой концепцией существующих систем управления – программированием движения робота по принципу «от точки к точке».

Навык же свободного манипулирования объектами внешнего мира, в свою очередь, является сложным, состоящим из ряда субзадач, к которым можно отнести анализ речи оператора, декомпозицию сложных задач в примитивы действий, оценку состояния рабочей области манипулятора, прогноз деформируемости объектов при взаимодействии, планирование движений и т.д. Данные умения робота далеко выходят за рамки построения традиционных систем управления, основанных на принципе «от точки к точке».

Для решения задачи свободного манипулирования объектами внешней среды требуется создание фреймворка сквозной архитектуры, охватывающего все субзадачи манипулирования. Настоящая работа посвящена разработке такого целостного фреймфорка, ориентированного на практическое решение задач свободного манипулирования. Предложен состав модулей фреймворка, описаны их функциональные связи и принципы взаимодействия. В связи с тем, что имеющиеся на

настоящий момент аналогичные фреймворки не уделяют должного внимания контролю исполнения примитива действия, разработан двухуровневый модуль исполнения задания.

1. Постановка задачи манипулирования произвольными объектами

Задача свободного манипулирования объектами заключается в осуществлении роботом физического взаимодействия с объектами рабочей области (сцены) в неструктурированной внешней среде с целью выполнения команды оператора. Данная задача противостоит задаче манипулирования объектами в закрытой среде.

В общем виде задача произвольного манипулирования объектами может быть сформулирована следующим образом. Пусть имеется инструкция от оператора l, представление рабочей области в начальный момент времени O_0 (такое представление может быть сделано на основе снимков RGB или RGB-D, либо иметь описание на языке PDDL) и информация о текущем пространственном положении звеньев и рабочего органа робота r_0 (встречается также термин проприоцепция робота).

Для выполнения следующего действия (например, захвата объекта) необходимо определить пространственное положение звеньев робота и рабочего органа r_1, такое что

$$r_1 = f\left(l, O_0, r_0\right)$$

В более широком смысле необходимо определить ведущее к достижению цели, поставленной в l, множество {r1, r2,…, rN}, где N – количество шагов по выполнению задачи, а каждый последующий шаг можно представить как

$$r_i = f\left(l, O_{i-l}, r_{i-1}\right)$$

При этом существует целевое состояние рабочей области ОТ, получаемое на основе функции преобразования Ф

$$\Phi : l \to O_T$$

Такое что

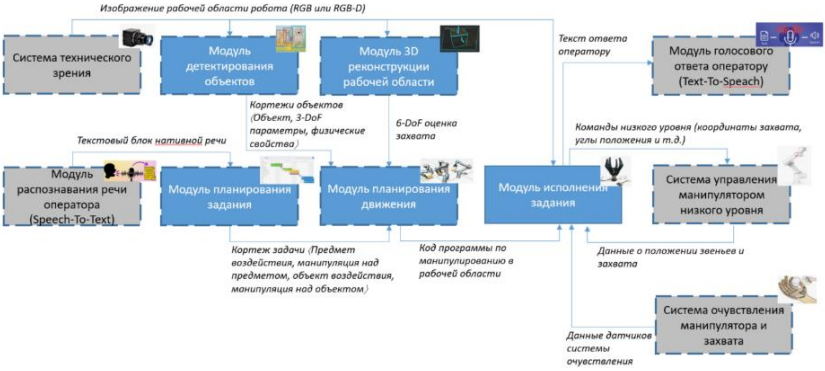
$$O_T \approx O_N$$

Иными словами, задачи свободного манипулирования объектами рабочей области робота-манипулятора сводится к достижению такого O_N, которое будет максимально приближено к O_T.

2. Предлагаемая структура системы управления коллаборативным роботом-манипулятором

Система управления роботом-манипулятором для целей свободного манипулирования объектами внешней среды является комплексной, что подразумевает создание фреймворка, состоящего из набора модулей. В литературе выделяют два подхода к его архитектуре: 1-ый основанный на использовании нейросетевой модели «видение-язык-действие» (англ. – Vision-Language-Action Model – VLAM) в качестве ядра фреймворка и 2-ой использующий сквозную цепочку модулей.

Предлагается организовать систему управления коллаборативным роботом-манипулятором, нацеленным на решение задачи свободного манипулирования объектами внешней среды, по модульному принципу. На рис. 1 приводится структура системы управления и иллюстрируется порядок взаимодействия модулей.



синим цветом – основные модули,
серым цветом – вспомогательные модули

Рис. 1. Структура системы управления коллаборативным роботом-манипулятором

Ниже приводится описание и функциональное назначение отдельных модулей:

1. Система технического зрения

Строится на основе RGB или RGB-D камеры в положении eye-in-hand. Возможно использование двух камер: одной в положении eye-in-hand, а второй в положении над рабочей сценой (стационарно).

2. Модуль распознавания речи оператора

Вход модуля: Аудио файл, содержащий команду оператора l.

Решаемые задачи: Трансформация аудио информации в текстовый блок.

Метод реализации: Известные решения, такие как модуль Whisper [5].

Выход модуля: Текстовый блок, содержащий представление l.

3. Модуль детектирования объектов

Вход модуля: RGB или RGB-D изображение(я) рабочей области O_i.

Решаемые задачи: Детектирование всех объектов, находящихся в рабочей области, сегментация изображения с целью точного определения границ и форм, оценка 3-DoF позиции выбранных объектов (x, y, θ), определение физических свойств объектов с целью коррекции характера последующего взаимодействия с захватом манипулятора.

Метод реализации: Дообучение известных решений, таких как SAM [6] для сегментации, Grounding DINO [7] или YOLO-World [8] для open-vocabulary object detection.

Выход модуля: Кортежи объектов <Объект, 3-DoF параметры, физические свойства>.

4. Модуль планирования задания

Вход модуля: Текстовый блок, содержащий задание l.

Решаемые задачи: Выделение из текстового блока непосредственно задания низкого уровня в том виде, в котором планировщик сможет его в дальнейшем обработать.

Обоснованно применять иерархический подход к планированию заданий, выполняемых машиной (например, как это продемонстрированно в работе) [9].

Метод реализации: Дообучение языковой модели, например Qwen3 [10] или Gemma3 [11] с помощью методов LoRA and DoRA как это сделано, например в [12].

Выход модуля: Кортеж задачи «Предмет воздействия, манипуляция над предметом, объект воздействия, манипуляция над объектом».

5. Модуль планирования движения

Вход модуля: Кортежи объектов, кортеж задачи, по запросу 6-DoF оценка захвата объекта

Решаемые задачи: Решает две задачи: (а) оценивает возможность (англ. – affordance) реализации l, исходя из Oi и (б) генерирует код программы на языке низкого уровня, понимаемого системой управления роботом (поставляемой изготовителем робота). Алгоритм работы модуля проиллюстрирован на рисунке 2.

Метод реализации: Алгоритмически реализованный модуль, использование LLM для генерации непосредственно машинного кода.

Выход модуля: Код программы по манипулированию, написанный на языке программирования низкого уровня (например, KUKA Robot Language).

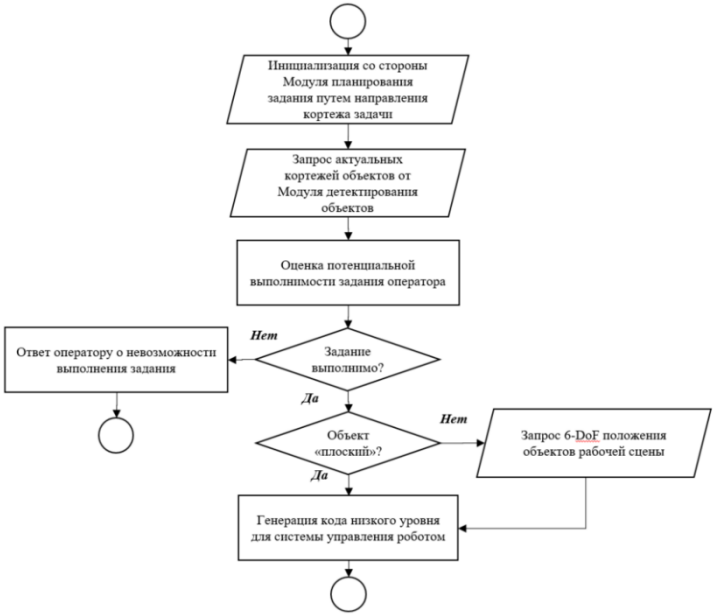


Рис. 2. Алгоритм работы Модуля планирования движения

6. Модуль 3D реконструкции рабочей области

Вход модуля: Изображения рабочей области робота от RGB или RGB-D камеры.

Решаемые задачи: Построение облака точек на основании множества RGB или RGB-D изображений рабочей области робота, определение 6-DoF позиции для осуществления захвата объекта интереса.

Метод реализации: Дообученные под конкретные наборы предметов 6-DoF модули, такие как Contact-GraspNet [13] или AnyGrasp [14].

Выход модуля: 6-DoF позиция (x, y, z, ϕ, θ, ψ) объекта интереса, модель рабочей области в виде облака точек

7. Модуль исполнения задания

Вход модуля: Код программы по манипулированию, написанный на языке программирования низкого уровня, данные о положении захвата и звеньев робота, данные от системы очувствления захвата робота.

Решаемые задачи: Реализует построчное исполнение кода программы путем направления команд системе управления роботом низкого уровня. Последующий контроль исполнения, реализованным в два контрольных слоя с помощью: (1) датчиков очувствления непосредственно робота и (2) итоговой оценки LLM соответствия конечного состояния рабочей сцены ON ожидаемому OT. Алгоритм работы модуля проиллюстрирован на Рис. 2.

Метод реализации: Алгоритмически реализованный модуль, дообучение LLM.

Выход модуля: Построчные команды низкого уровня, текст ответа оператору для поддержания интерактивного диалога.



Рис. 3. Алгоритм работы Модуля исполнения задания

8. Модуль голосового ответа оператору

Вход модуля: Текст ответа оператору для поддержания интерактивного диалога.

Решаемые задачи: Трансформация текстового блока в аудио информацию.

Метод реализации: Известные решения, такие как модуль Whisper.

Выход модуля: Аудио информация.

9. Система очувствления манипулятора и захвата

Вход модуля: Запросы от Модуля исполнения задания.

Решаемые задачи: Опрос тензорезистивных и силомоментных датчиков очувствления захвата робота.

Метод реализации: Алгоритмический модуль, основанный на библиотеках для обработки сигналов датчиков очувствления.

Выход модуля: Данные датчиков очувствления захвата робота.

10. Система управления манипулятором низкого уровня

Вход модуля: Построчные команды низкого уровня.

Решаемые задачи: Преобразование команды низкого уровня в движения звеньев робота и захвата.

Метод реализации: Предоставляется производителем робота (например, KUKA Robot Language).

Выход модуля: Данные о положении захвата и звеньев робота.

Выводы

Концепция сервисного робота общего назначения, выполняющего произвольные задачи оператора бытового характера, давно являлась ориентиром и мотиватором для практических исследований. Ключевым шагом на пути к реализации данной концепции выступает задача манипулирования объектами открытого мира, то есть осуществление произвольных манипуляций с любыми объектами интереса, находящимися в поле технического зрения робота (с учётом ограничений, налагаемых конструкцией захвата робота). Принципиальный подход к решению данной задачи заключается в создании фреймворков, решающих отдельные подзадачи (интерпретация команд оператора, распознавания объектов рабочей сцены и т.д.) в рамках единой архитектуры. Подобные фреймформи можно разделить на две большие группы: 1 опирающиеся на визуально-языковую модель действий (VLAM) и 2, состоящие из множества относительно независимых модулей.

Несмотря на концептуальные преимущества подхода с использованием VLAM, представляется, что на настоящий момент подход с использованием множества независимых модулей представляется более актуальным с точки зрения практического использования. Рассмотрена конкретная архитектура системы управления коллаборативным роботом на основе такого подхода с ядром из пяти основных модулей. Научная новизна предложенной сквозной архитектуры заключается в 1 использовании большой языковой модели (Large Language Model – LLM) для генерации кортежей задач, поставленных оператором, 2 использовании Модуля исполнения задания, осуществляющего покомандный контроль за действиями робота на основании датчиков очувствления и 3 использовании большой визуально-

языковой модели (VLM) для оценки соответствия конечного состояния рабочей сцены инструкции оператора.

Список литературы

1. Desiderata for Planning Systems in General-Purpose Service Robots. [Электронный ресурс]. – Режим доступа: https://arxiv.org/pdf/1907.02300
2. RoboCup@Home Rules & Regulations. [Электронный ресурс]. – Режим доступа: https://robocupathome.github.io/RuleBook/rulebook/master.pdf
3. Ющенко, А.С. Коллаборативная робототехника и человеческий фактор // Актуальные проблемы психологии труда, инженерной психологии и эргономики / Под редакцией А.А. Обознова, А.Л. Журавлева. – Москва: Институт психологии РАН, 2020. – С. 83-103. – DOI 10.38098/ergo.2020.007. – EDN ZAQAVN.
4. Ющенко, А. С. Интерактивная коллаборативная робототехника // XIV Всероссийская мультиконференция по проблемам управления МКПУ-2021: материалы XIV мультиконференции: в 4 т., Дивноморское, Геленджик, 27 сентября – 02 2021 года. Том 1. – Ростов-на-Дону: Южный федеральный университет, 2021. – С. 156-158. – EDN LVUPZU.
5. Haz, A.L. A Study of Audio-to-Text Conversion Software Using Whispers Model / A. L. Haz, E. D. Fajrianti, N. Funabiki and S. Sukaridhoto // 2023 Sixth International Conference on Vocational Education and Electrical Engineering (ICVEE). – Surabaya, Indonesia, 2023. – P. 268-273.
6. Kirillov, A. Segment Anything / A. Kirillov // IEEE/CVF International Conference on Computer Vision (ICCV). – Paris, France, 2023. – P. 3992-4003.
7. Shilong, L. Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection / L. Shilong, Z. Zhaoyang, R. Tianhe, L. Feng, Z. Hao, Y. Jie, J. Qing, L. Chunyuan, Y. Jianwei, S. Hang, Z. Jun, Z. Lei // In Computer Vision – ECCV 2024: 18th European Conference. – 2024. – Part XLVII. – P. 38-55.
8. Cheng, T. YOLO-World: Real-Time Open-Vocabulary Object Detection / T. Cheng, L. Song, Y. Ge, W. Liu, X. Wang and Y. Shan // IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – Seattle, WA, USA, 2024. – P. 16901-16911.
9. Song, C. H. LLM-Planner: Few-Shot Grounded Planning for Embodied Agents with Large Language Models / C. H. Song, B. M. Sadler, J. Wu, W. -L. Chao, C. Washington and Y. Su // IEEE/CVF International Conference on Computer Vision (ICCV). – Paris, France, 2023. – P. 2986-2997.

10. Qwen3 Technical Report. [Электронный ресурс]. – Режим доступа: https://doi.org/10.48550/arXiv.2505.09388

11. Gemma 3 Technical Report. [Электронный ресурс]. – Режим доступа: https://doi.org/10.48550/arXiv.2503.19786

12. Tereshchenko, V. The use of LLM to identify goals and objectives from the text of degree thesis using fine-tuning and prompt-engineering methods / V. Tereshchenko, V. Martsinkevich, N. Aminov and A. Dukhanov // International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA). – Antalya, Turkiye, 2025. – P. 1-5.

13. Sundermeyer, M. Contact-GraspNet: Efficient 6-DoF Grasp Generation in Cluttered Scenes / M. Sundermeyer, A. Mousavian, R. Triebel, D. Fox // IEEE International Conference on Robotics and Automation (ICRA). – Xi'an, China, 2021. P. 13438-13444.

14. Fang, H. S. AnyGrasp: Robust and Efficient Grasp Perception in Spatial and Temporal Domains / H. S. Fang, W. Chenxi, F. Hongjie, G. Minghao // IEEE Transactions on Robotics. – 2023. – Vol. 39, – № 5, – С. 3929-3945.

Применение машинного обучения для классификации кибератак

Ю. А. Крыжановская, email: jak @mail.ru[8]

[1]Воронежский государственный университет

Аннотация. *Рассмотрена реализация модели Classifier Chain (Random Forest) для многометочной классификации кибератак и произведена оценка качества полученной модели с применением различных метрик на открытом наборе данных SR-BH 2020.*

Ключевые слова: *Classifier Chain, Random Forest, классификация, оценка качества, набор данных SR-BH 2020.*

Введение

Информационные технологии в настоящий момент применяются практически во всех сферах человеческой деятельности, включая такие, как медицина, финансы, производство и многие другие, в которых случайное или, тем более, преднамеренное разглашение, искажение, подмена информации может иметь самые серьезные последствия. В связи с этим сложно переоценить важность обеспечения безопасности и защиты данных. Кроме того, необходимо уметь оценивать риски, определять возможные пути и способы воздействия на чувствительную информацию, определять факт проникновения, чтобы противодействовать этому. Также следует учесть, что многие организации для решения своих задач применяют веб-приложения, и достаточно малое число систем обработки информации не предполагают взаимодействия через сети, которые могут быть использованы, в том числе, для реализации атак. Для обнаружения атак могут применяться различные подходы и средства, а в последнее время в этой области широкое распространение получило машинное обучение. Так, в [1] анализировалось использование гладкого частного леса и логистической модели деревьев для обнаружения атак. В работе [2] предложен метод классификации, основанный на предварительно обученной языковой модели ELECTRA, с использованием метрик precision, recall, F1-score. Машинное обучение с использованием методов k-ближайших соседей, случайного леса, дерева решений применительно к обеспечению безопасности цепочек поставок рассматривается в [3]. В данной работе производительность моделей исследуется с использованием метрик accuracy, precision, recall, F1-score, коэффициента корреляции Мэтьюса

© Крыжановская Ю.А., 2026

(MCC), размера модели (Mem), времени обучения (TT) и времени прогнозирования (PT). В статье [4] производительность алгоритма случайного леса (RF), многослойного персептрона (MLP) и сети долговременной кратковременной памяти (LSTM) оценивается с применением метрик accuracy и F1-score.

В данной работе для обнаружения атак предлагается реализация на python модели Classifier Chain [5] с базовым классификатором Random Forest, а для оценки качества модели предполагается использовать метрики accuracy, F1-score, Hamming loss, ROC AUC, jaccard similarity, informedness, markedness.

1. Набор данных

В качестве набора данных для модели использовался SR-BH 2020 [2, 6], представляющий собой многоклассовый набор данных. Общий объем составляет 907814 записей, содержащих информацию о веб-запросах, каждому из них присвоена классификация CAPEC (Common Attack Pattern Enumeration and Classification) [7]. Большая часть запросов – 525195 (57,8%) – относятся к «нормальным» запросам, а 382619 (42,2%) запросов являются атаками. Для набора данных определено 24 признака, включающие сетевые параметры, http-заголовки, ответ сервера, и 13 различных типов атак. В качестве особенностей данного набора стоит отметить то, что один запрос может относиться к разным видам атак (многометочность), что классы не сбалансированы по размеру, а в данных имеются пропуски, в частности, NaN встречается 616419 раз. Эти особенности обуславливают необходимость предварительной обработки данных, которая осуществлялась в несколько этапов:

– Переименование (листинг 1);

Листинг 1

Переименование

```
d=d.rename(index=str,columns=['000 – Normal': 'normal']))
d=d.rename(index=str,columns=['272 – Protocol Manipulation':
'protocol manipulation']))
d=d.rename(index=str,columns=['242 – Code Injection': 'code
injection']))
d=d.rename(index=str,columns=['242 – OS Command Injection':
'os command injection']))
```

– Удаление строк с значениями NaN и замена пропусков в response_content_length на 0 (листинг 2);

Листинг 2

Замена пропусков на 0

```
d['response_content_length’].fillna(0,inplace=True)
d['response_content_length’]=
d['response_content_length’].astype(int)
d.isnull().sum()
```

– Кодирование категориальных признаков (CountEncoder из category_encoders);
– Визуализация данных (распределение response_content_length: гистограмма, KDE-график, круговая диаграмма; круговые диаграммы для request_referer и request_http_method).

На рис. 1 и рис. 2 приведены, соответственно, гистограмма и KDE-график распределения response_content_length .

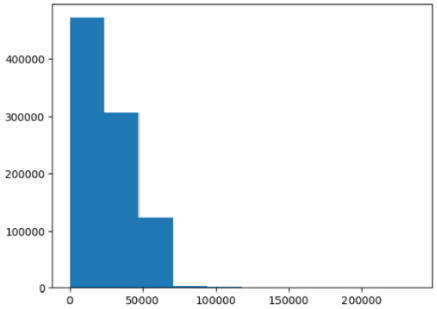


Рис. 1. Распределение response_content_length: гистограмма

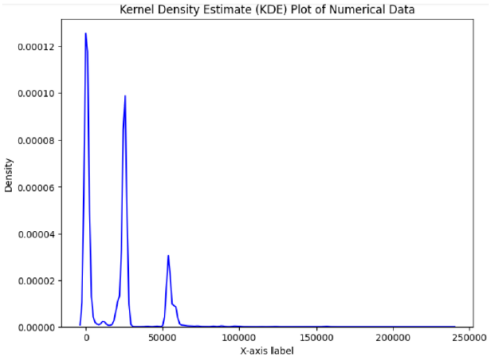


Рис. 2. Распределение response_content_length: KDE-график

После предобработки данные разделяются на обучающую и тестовую части в пропорции 80% и 20%., что представлено на рис. 3.

```
X = d4[['request_http_method','request_http_request','request_referer','request_user_agent
y = d4[['normal','protocol manipulation','code injection','os command injection','path tr

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
base_classifier = RandomForestClassifier()
classifier_chain = ClassifierChain(base_classifier, order='random')
```

Рис. 3. Разделение 80/20

2. Классификация. Модель и оценка качества

Для решения задачи классификации применялась модель Classifier Chain с базовым алгоритмом Random Forest. При выборе модели была учтена зависимость между метками, а также то, что данный алгоритм хорошо работает с категориальными данными и устойчив к переобучению.

В процессе решения задачи использовались три основные метрики классификации, четыре детализированных метрики и две дополнительных, полученные результаты представлены в табл.1-3.

Таблица 1

Основные метрики классификации

Метрика	Значение	Интерпретация
Accuracy	0.8507	Модель правильно предсказывает ~85.07% всех примеров.
F1-score (micro)	0.8520	Гармоническое среднее precision и recall (микро-усреднение). Хороший баланс.
Hamming Loss	0.0230	Очень низкая доля ошибочно предсказанных меток (2.3%). Чем ближе к 0, тем лучше.

Таблица 2

Детализированные метрики

Метрика	Значение	Интерпретация
Precision (micro)	0.8534	Из всех предсказанных положительных классов 85.34% верны.
Recall (micro)	0.8505	Модель находит 85.05% всех реальных положительных примеров.
Micro-Averaged ROC AUC	0.9191	Отличная способность разделять классы (AUC > 0.9 – высокая точность).
Micro-Averaged Jaccard	0.7421	Умеренное сходство между предсказанными и истинными метками (макс. = 1).

Таблица 3

Дополнительные метрики

Метрика	Значение	Интерпретация
Informedness	0.7799	Модель хорошо избегает ложных срабатываний и пропусков (макс. = 1).
Markedness	0.7520	Высокая согласованность предсказаний с реальными данными (макс. = 1).

Графически анализ результатов на различных метриках представлен на рис. 4.

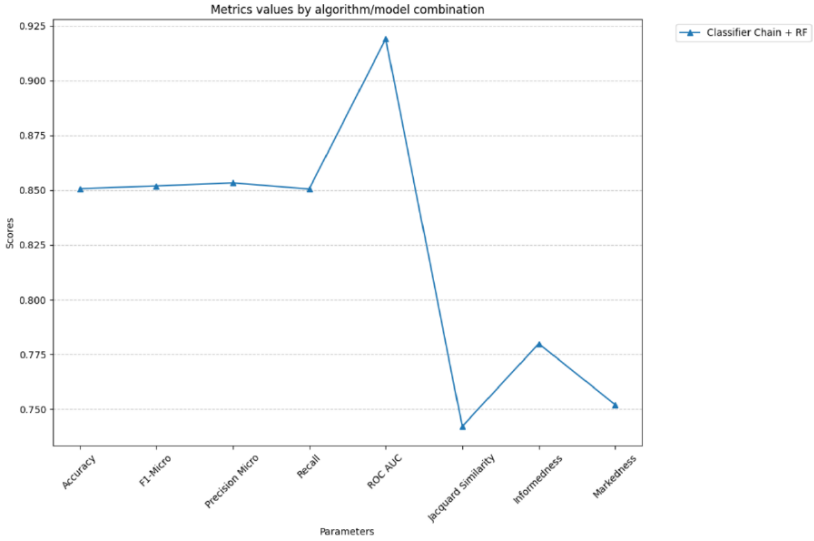


Рис. 4. Анализ на различных метриках

Заключение

В данной статье рассмотрено решение задачи классификации кибератак с использованием модели Classifier Chain, релизованной на языке python. Для обучения и тестирования модели был выбран открытый набор данных SR-BH 2020, данные из набора были предварительно обработаны (удаление дубликатов, кодирование категориальных признаков). Построенное решение было оценено с использованием различных метрик. Анализ полученных результатов показал, что полученная реализация демонстрирует хороший баланс, высокую согласованность предсказаний с реальными данными и низкую долю

ошибочно предсказанных меток, что говорит о высокой эффективности выбранного подхода при решении задачи многометочной классификации.

Список литературы

1. Kareem, M. I. Machine Learning-based Detection of Web Attacks Using Logistic Model Trees / M. I. Kareem, G. Bahaa. Abdulhussein, D. A. Mohammed, A. Z. K. Matloob and H. K. Alwan // International Conference on Computer Science and Software Engineering (CSASE) (Duhok, Iraq, 2025). – Duhok, 2025. – pp. 252-257, doi: 10.1109/CSASE63707.2025.11053993
2. Gu, F. Web Attack Classification Method Based on Pre-Trained Language Model ELECTRA / F. Gu, Y. Sun, T. Ji, Y. Zhang, C. Yang and T. Wang // 11th IEEE International Conference on Privacy Computing and Data Security (PCDS) (Hakodate, Japan, 2025). – Hakodate, 2025. – pp. 363-371. doi: 10.1109/PCDS65695.2025.00056
3. A Comparative Study of Lightweight Machine Learning Techniques for Cyber-Attacks Detection in Blockchain-Enabled Industrial Supply Chain / S. Ismail, [etc] // IEEE Access. 2024. – vol. 12. – pp. 102481-102491. doi: 10.1109/ACCESS.2024.3432454
4. Intelligent Cyber Attack Detection and Classification for Network-Based Intrusion Detection Systems / N. Oliveira, [etc.] // Applied Sciences. – 2021. – № 11(4). – pp. 1674. doi: 10.3390/app11041674
5. Classifier Chains: A Review and Perspectives / J. Read, [etc.] // Journal of Artificial Intelligence Research. – 2021. – №70. – pp. 683-718.
6. HarvardDataverse, “SR-BH 2020 multi-label dataset” [Электронный ресурс]. – Режим доступа : https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/OGOIXX
7. MITRECorporation, “CAPEC – Common Attack Pattern Enumeration and Classification (CAPEC TM)” [Электронный ресурс]. – Режим доступа : https://capec.mitre.org/index.html

Анализ современных моделей синтеза речи в задачах межъязыковой адаптации

Н. С. Кутепов, email: kutepov.nikita.bmstu@mail.ru
К. Н. Ткаченко, email: kutepov.nikita.bmstu@mail.ru

Российский университет дружбы народов имени Патриса Лумумбы

Аннотация. *Рассмотрены различные архитектуры моделей синтеза, используемые в области машинного обучения и искусственного интеллекта.*

Ключевые слова: *Синтез речи, нейросетевые модели, межъязыковая адаптация, TTS.*

Введение

В условиях глобализации цифровых технологий мультиязычные системы искусственного интеллекта приобретают критическое значение для обеспечения равного доступа к сервисам в лингвистически разнообразных регионах. Особую актуальность это приобретает в контексте синтеза речи, где эффективная межъязыковая адаптация моделей позволяет преодолевать барьеры между языковыми сообществами. Современные исследования демонстрируют растущий спрос на механизмы переноса знаний о фонетике и просодии с высокоресурсных на низкоресурсные языки. Данная проблема становится ключевой для разработки инклюзивных голосовых интерфейсов следующего поколения.

Несмотря на значительный прогресс в области речевого синтеза, отсутствие систематизированной классификации современных моделей создает существенные препятствия для исследователей и разработчиков. Это особенно ощутимо при работе с языками, имеющими ограниченные корпусы для обучения, где традиционные подходы требуют непропорционально высоких вычислительных и временных затрат. Существующий разрыв между теоретическими возможностями моделей и их практическим применением замедляет внедрение технологий в реальные системы. Таким образом, задача структурирования знаний приобретает первостепенное значение для оптимизации процессов адаптации.

© Кутепов Н. С., Ткаченко К. Н., 2026

1. Обзор современных архитектур моделей синтеза речи

Архитектура Tacotron представляет собой последовательную модель типа encoder-decoder с механизмом внимания. Она преобразует входной текст в спектрограммы через поэтапное прогнозирование акустических признаков. Ключевой особенностью является использование location-sensitive attention, обеспечивающего точное выравнивание текста и аудио. Данный подход позволяет генерировать естественную речь с выраженной просодией, но требует значительных вычислительных ресурсов.

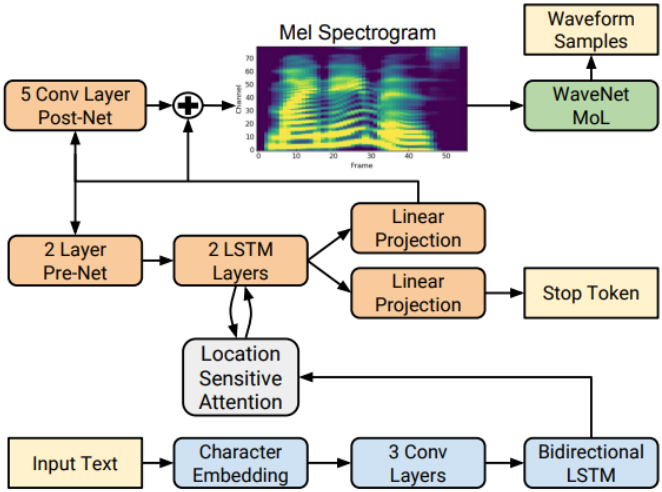


Рис. 1. Структурная схема модели Tacotron2

В отличие от Tacotron, модель FastSpeech использует неавторегрессивный подход, что значительно ускоряет процесс синтеза. Она применяет фонемные выравнивания и длительности для параллельной генерации. Структурно система состоит из блока преобразования текста в скрытые представления и модуля генерации спектрограмм. Такая архитектура обеспечивает стабильность вывода и контроль над скоростью речи.

Модель VITS интегрирует преимущества вариационных автоэнкодеров и диффузионных подходов, объединяя текстозависимую и текстонезависимую генерацию. Она использует стохастические процессы для моделирования вариативности речевых паттернов. Современное развитие архитектуры демонстрирует улучшенное качество синтеза за счет совместной оптимизации всех компонентов системы.

Сравнительная оценка архитектур по критериям естественности и выразительности показывает, что системы на базе глубокого обучения демонстрируют выдающиеся результаты. Tacotron обеспечивает высокую выразительность за счет авторегрессивного механизма, но уступает в скорости. FastSpeech достигает сопоставимого качества с улучшенной эффективностью, а VITS сочетает естественность звучания с гибкостью генерации. Эти характеристики делают рассмотренные модели базовыми решениями для задач межъязыковой адаптации.

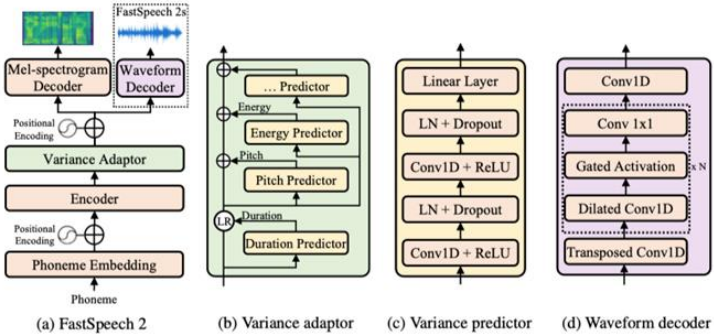


Рис. 2. Структурная схема модели FastSpeech2

Перенос лингвистических знаний в современных моделях синтеза речи базируется на переиспользовании скрытых представлений, полученных при предобучении на больших корпусах данных. Многослойные энкодеры и декодеры формируют абстрактные признаки, кодирующие фонетические и субфонетические характеристики, которые могут быть релевантны для нескольких языков. Общие латентные пространства позволяют моделям сохранять информацию о артикуляции и акустических паттернах, что снижает необходимость в обширных целевых данных. Это создает предпосылки для эффективной межъязыковой адаптации при ограниченных ресурсах целевого языка. Практические подходы к использованию скрытых представлений включают заморозку ранних слоев, дообучение верхних слоев и внедрение параметрически эффективных адаптеров. Адаптеры и слойные трансферные блоки позволяют сохранить универсальные фонетические представления и одновременно вводить язык-специфические корректировки без полного переобучения модели. Дополнительные техники, такие как регуляризация и совместное обучение на смешанных языковых корпусах, способствуют сохранению переносимых знаний и предотвращают катастрофическое забывание. В итоге комбинирование

общих представлений с небольшими языковыми параметрами обеспечивает баланс между универсальностью и спецификой. Роль предобученных скрытых представлений проявляется в их способности обеспечивать исходную структуру для генерации просодии и артикуляции в новых языках. При переносе важно учитывать различия в фонологической системе и фонетическом инвентаре целевого языка, чтобы адаптация скрытых представлений не вводила систематических ошибок. Методы нормализации и привязки признаков к языковым метаданным помогают корректировать универсальные представления под локальные особенности. Такие механизмы повышают стабильность и качество синтеза при межъязыковой адаптации.

Адаптация фонетических характеристик для низкоресурсных языковых пар часто основана на введении единой фонетической репрезентации, например, через универсальные алфавиты или артикуляторные признаки. Преобразование графем в фонемы с последующей нормализацией и аппроксимацией недостающих фонем позволяет расширить покрытие моделируемых звуков без значительных затрат данных. Дополнительные методы включают синтетическую генерацию тренировочных примеров и использование межъязыковых словарей соответствий, что уменьшает разрыв между источником и целевым языком. Это обеспечивает более корректную артикуляцию при ограниченных ресурсах целевой языковой пары. Просодическая адаптация направлена на перенос интонационных шаблонов, длительности и динамики речи с учетом языковых контекстов и выражательных особенностей. Техники включают обучение условных просодических эмбеддингов, использование языковых или диалектных кодов и многозадачное обучение, где просодия моделируется как отдельная, но связанная задача. Также применяются методы разделения факторов, такие как адверсариальное обучение для отделения говорящих и языковых характеристик, что способствует лучшей генерализации просодии. В совокупности эти стратегии позволяют моделям корректно отражать ритм и интонацию целевого языка при минимальном объеме адаптационных данных.

2. Анализ подходов к межъязыковой адаптации: fine-tuning и zero-shot learning

Fine-tuning предполагает архитектурные модификации для адаптации предобученных моделей к новым языкам. Ключевым механизмом является настройка слоёв акцентации и внедрение адаптивных слоёв (adapter layers), которые позволяют сохранять исходные лингвистические представления. Заморозка весов базовых слоёв минимизирует катастрофическое забывание при дообучении на

ограниченных данных. Эти изменения обеспечивают гибкую подстройку модели без полной перетренировки. Эффективность fine-tuning напрямую зависит от качества данных целевого языка. Обязательным требованием выступает наличие точных выравниваний (alignments) между фонетическими и акустическими представлениями. Процедуры фонемного отображения (phoneme mapping) решают проблему различий в фонетических инвентарях языков. Формат данных должен соответствовать исходному представлению, используемому при предварительном обучении модели.

Zero-shot learning устраняет необходимость в целевых языковых данных через использование общих многоязычных эмбеддингов. Условные языковые токены и встраивания говорящего позволяют генерировать речь на непредставленных в обучении языках. Ключевую роль играет предварительное мультилингвальное обучение на разнородных корпусах, формирующее обобщённые фонетические представления. Факторы обобщаемости включают архитектурную ёмкость моделей и степень лингвистического разнообразия тренировочных данных.

Выбор метода адаптации для низкоресурсных языков определяется чувствительностью к объёму данных и вычислительной стоимостью. Fine-tuning демонстрирует превосходство при наличии хотя бы минимальных размеченных данных, обеспечивая фонетическую достоверность. Zero-shot подходы эффективны при полном отсутствии целевых данных, но требуют устойчивости к шумам в разметке мультиязычных корпусов. Критическим параметром остаётся баланс между качеством синтеза и ресурсными затратами на адаптацию.

Практическая реализация fine-tuning для Tacotron начинается с подготовки фонетических транскрипций целевого языка и выравниваний временных характеристик. Ключевым этапом является частичное дообучение декодера и attention-механизма на ограниченном корпусе данных, что позволяет сохранить лингвистические особенности исходной модели. Для предотвращения деградации качества синтеза на базовом языке применяются методы регуляризации, такие как заморозка начальных слоев энкодера и использование адаптивного коэффициента обучения.

Zero-shot адаптация Tacotron достигается за счет интеграции языковых маркеров в скрытые представления и комбинации speaker embeddings с лингвистическими признаками. Мультилингвальные расширения модели используют смешанный тренировочный корпус из нескольких языков, что формирует общее фонетическое пространство. Техники метаподстройки (meta-learning) позволяют генерировать речь на

языках, не представленных в обучающих данных, путем экстраполяции языковых зависимостей.

Адаптация FastSpeech к новым языкам фокусируется на модификации предсказателя длительностей и генерации мел-спектрограмм с учетом фонетических особенностей целевого языка. Параллельная архитектура модели обеспечивает стабильность адаптационного обучения и сокращает время обработки по сравнению с последовательными системами. Выбор вокодера существенно влияет на качество синтеза, причем модели на основе диффузионных процессов демонстрируют лучшую адаптивность к мультилингвальным сценариям.

Эмпирические наблюдения выявили различия в интеллигибельности (CER 5.8% для Tacotron vs 4.2% для FastSpeech) и натуралистичности синтезированной речи при межъязыковой адаптации. Типичные ошибки включают некорректную передачу просодии и артикуляцию несвойственных фонем, особенно при отсутствии родственных языков в обучающих данных. Для достижения приемлемых результатов требуется не менее 4 часов аннотированных записей, причем качество данных критически влияет на устойчивость адаптации.

Заключение

Проведённый анализ современных архитектур синтеза речи, включая Tacotron, FastSpeech и VITS, подтвердил их высокую эффективность для задач межъязыковой адаптации. End-to-end подходы на базе глубокого обучения демонстрируют способность генерации естественной речи с сохранением просодических характеристик. Ключевую роль играют механизмы transfer learning, позволяющие переносить знания о фонетических и интонационных паттернах между языками при минимальной доработке моделей.

Сравнительная оценка методов адаптации выявила дифференцированную область применения fine-tuning и zero-shot learning. Fine-tuning показывает превосходные результаты при наличии даже ограниченных данных целевого языка, обеспечивая высокое качество синтеза. Zero-shot подходы остаются перспективным направлением для абсолютно низкоресурсных сценариев, хотя требуют дальнейшей оптимизации для достижения сопоставимой естественности речи.

Разработанная классификация моделей по критериям ресурсоёмкости, качества синтеза и скорости адаптации устраняет проблему несопоставимости подходов в существующих исследованиях. Предложенная система категоризации предоставляет исследователям структурированный инструмент для выбора оптимальных решений.

Классификация учитывает компромиссы между вычислительной сложностью и лингвистической универсальностью моделей.

Перспективы практического применения связаны с развитием few-shot обучения и созданием универсальных кросс-лингвистических представлений. Это особенно актуально для образовательных платформ и голосовых ассистентов в регионах с лингвистическим разнообразием. Оптимизация данных направлений позволит расширить доступность речевых технологий для носителей низкоресурсных языков.

Список литературы

1. Андреев, П.К Генеративные модели для улучшения речи: диссертация. – Москва, Национальный исследовательский университет «Высшая школа экономики», 2024.

2. Бакуменко Я.Д., Таджибова А.Н. Современные проблемы и решения в области машинного перевода // Успехи кибернетики. – 2025. – №2. – С. 47–59.

3. Бондаренко, В.И., Елисеев В.О., Ермоленко Т.В. Анализ эффективности глубоких языковых моделей для задачи определения тональности русскоязычных текстов // Проблемы искусственного интеллекта. – 2024. – №1. – С. 51–55.

4. Залевская, А.А. Общенаучное и узкоспециальное значение термина // Вестник ТвГУ. Серия "Филология". – 2017. – №2. – С. 7–13.

5. Израилова, Э.С. Особенности машинного обучения средствами CNN в рамках синтеза речи // Вестник ГГНТУ. Технические науки. – 2019. – №2. – С. 29–33.

Разработка системы хранения и подготовки данных для машинного перевода русского жестового языка

В. В. Ландовский, email: landovskij@corp.nstu.ru[1]
М. Г. Гриф, email: grifmg@mail.ru [1]
О.О. Королькова, email: ookorol@mail.ru[1]

[1]Новосибирский государственный технический университет

Аннотация. *В статье рассматривается разработка системы управления набором данных для распознавания русского жестового языка. Исходя из особенностей алгоритмов распознавания выделены ключевые сущности, предложена концептуальная модель системы. Разработана база данных с использованием СУБД PostgreSQL. Интерфейс пользователя реализован на языке C# в среде Microsoft Visual Studio.*

Ключевые слова: *распознавание жестов, русский жестовый язык, база данных.*

Введение

Одна из ключевых проблем машинного перевода русского жестового языка (РЖЯ) заключается в том, что в настоящее время существует ограниченное количество доступных датасетов для РЖЯ, а сбор данных для машинного перевода жестов требует участия людей, способных грамотно исполнять жесты и объяснять их значения на звуковом языке. Уже имеются разработки, способные распознавать жесты [1-4], ряд из которых применяется в различных мультимедиасистемах и бытовых системах автоматизации. Данное направление активно развивают зарубежные и отечественные компании [2, 5]. Несмотря на существование технологий распознавания жестов, отсутствие структурированных синтаксических наборов данных ограничивает точность перевода, что определяет актуальность проектов в данной области.

Сферой научных интересов авторов статьи является разработка системы распознавания русского жестового языка, включая создание машинно-ориентированного датасета синтаксических конструкций русского жестового языка с многоуровневой аннотацией [6], способного помочь решить проблемы коммуникации людей с ограниченными возможностями здоровья по слуху со слышащими собеседниками. В отличие от существующих электронных словарей РЖЯ, предлагаемая

© Ландовский В.В., Гриф М.Г., Королькова О.О., 2026

система содержит не только лексические единицы, но и модели синтаксических связей, а также контекстуальные варианты интерпретации. Такой датасет помогает понять, как формируются сложные предложения, каким образом связаны между собой части фразы и какую роль играют различные жесты в передаче смысла.

1. Ключевые факторы, влияющие на аннотирование данных

Предметом изучения лексики РЖЯ является жест как основная единица русского жестового языка. В составе жеста выделяют мануальный компонент, включающий конфигурацию кисти, ориентацию рук в пространстве относительно друг друга и корпуса говорящего, локализацию и движение, и немануальный компонент, включающий движения головы и тела, взгляд, выражение лица и артикуляцию, называемую маусингом. Особенностью жеста является одновременное выполнение всех его элементов, в отличие от линейного следования звуков в устной речи.

Основной единицей синтаксиса РЖЯ является высказывание. В высказывании выделяют главные составляющие – подлежащее, дополнение и предикат, аналогичные подлежащему, прямому дополнению и сказуемому в предложениях звукового языка. Часто для их описания в лингвистической литературе используются термины «субъект», «объект» и «предикат». Говоря о порядке жестов в высказывании, имеют в виду порядок следования этих главных составляющих. Исследователями русского жестового языка было установлено существование базового порядка жестов в высказываниях и нескольких вариантов следования языковых единиц в них. Порядок следования жестов позволяет передать значения подлежащего, дополнения и предиката и зависит от частеречной принадлежности языковых единиц [7]. Кроме того, в РЖЯ важны позиции обстоятельства и определения, которые зависят от их семантического типа.

В составе высказываний выделяют именную группу, под которой понимают жест или последовательность жестов, которые представляют собой референт или указывают на него в непрямом смысле [8]. Е. В. Прозоровой было установлено, что вершиной полной именной группы является субстантивный или местоименный жест.

Кроме того, в составе высказываний могут выделяться объединения жестов, названные Г. Л. Зайцевой синтагмами [9]. Дефектолог проанализировала высказывания, включающие синтагмы с одновременным исполнением двух жестов, и пришла к выводу, что они позволяют выражать пространственные, атрибутивные и причинно-следственные отношения.

Для обеспечения лингвистической полноты и пригодности для анализа аннотация каждой фразы включает как формальные, так и содержательные параметры. Структура схемы разметки представлена в табл.

Таблица 1

Структура схемы разметки фразы

Параметр	Описание	Пример
Номер жеста	Порядковый номер жеста внутри фразы	«1»
Значение жеста	Уточнение, если значение жеста отличается от стандартного или зависит от синтаксической позиции	
Начало жеста (сек.)	Временные метки, определяющие границы исполнения жеста на видео	«0.12»
Конец жеста (сек.)		«1.10»
Часть речи	Грамматическая категория жеста	«сущ.», «гл.», «мест.» и др.
Синтаксическая роль	Функция жеста в структуре фразы	«именная группа»
Композиционное отношение	Описание связей между жестами внутри фразы	«1» – главная составляющая (субъект, обьект, предикат), «0» – остальные жесты (обстоятельства, определения)

Особенности синтаксиса РЖЯ необходимо учитывать при аннотировании высказывания. Процесс аннотирования включает разметку высказывания/фразы. Важным этапом разметки является сегментация высказывания – процесс, при котором для каждого жеста во фразе на жестовом языке фиксируются моменты начала и конца его выполнения на временной шкале. В отличие от покадровой разметки, такой подход позволяет выделять целостные жестовые единицы для дальнейшего анализа и автоматического распознавания.

2. Проектирование базы данных

Анализ предметной области позволил выделить основные сущности и отношения между ними. Схема логической модели представлена на рис. 1.

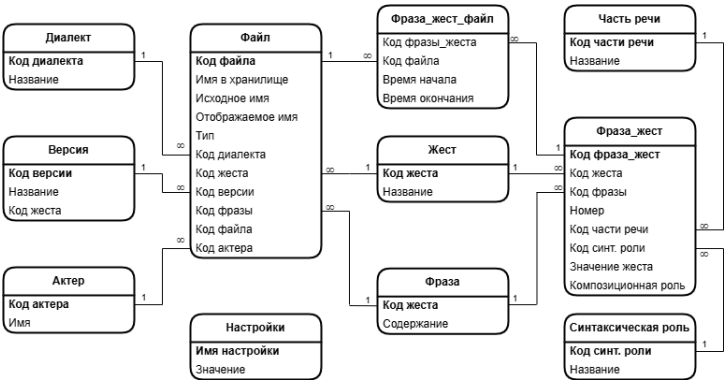


Рис. 1. Логическая модель данных

Связь между жестом и фразой реализована через дополнительную таблицу «Фраза_жест». Записи этой таблицы содержат параметры жеста, входящего в состав фразы. Поскольку для фразы может быть записано несколько видеофайлов, наборов временных меток должно быть также несколько, они находятся в таблице «Фраза_жест_файл». Каждая запись в этой таблице связывает вхождение жеста во фразу и определяет его временные метки для конкретного файла.

В качестве СУБД выбрана PostgreSQL [10] – свободная объектно-реляционная система управления базами данных с открытым кодом. Она существует в реализациях для множества UNIX-подобных платформ, а также для Microsoft Windows. PostgreSQL отличается высокой производительностью, не накладывает ограничений на размер базы данных, поддерживает параллельные вычисления. Выбранная система отлично справляется с большими объёмами данных, что позволяет масштабировать её без существенных изменений архитектуры и обеспечивает высокую скорость выборок и обработки информации.

Хранение файлов реализовано отдельно от СУБД. В базе данных хранится лишь информация о загруженных файлах. В настоящее время таких файлов более 200 тысяч. Содержимое файлов копируется в специальный каталог, на которой развёрнута СУБД. Путь к этому каталогу задаётся в настройках системы. Файл помещается в подкаталог, соответствующий уникальному номеру фразы в базе данных.

3. Разработка программного обеспечения

Интерфейс пользователя разработан с учётом типовых требований к пользовательским интерфейсам и реализован на языке C# в среде Microsoft Visual Studio [11].

Существует три режима работы с программой, переключение между которыми осуществляется в верхней части окна: работа с жестами, добавление файлов и редактирование фраз. Первые два режима позволяют создавать новые жесты, редактировать свойства существующих жестов и загружать файлы. Для каждого жеста можно добавить видеофайлы и файлы разметки. В третьем режиме работы осуществляется добавление и редактирование фраз. Внешний вид окна программы представлен на рис. 2. В списке, расположенном в центре окна, отображаются жесты, входящие во фразу. Чтобы пополнить состав фразы из левого списка нужно выбрать предварительно добавленный жест, в контексте текущей фразы указать его значение, часть речи, синтаксическую и композиционную роль, и после этого нажать на кнопку с двумя стрелками вправо, обозначающую перемещение. Для изменения последовательности предусмотрены кнопки со стрелками вверх и вниз.

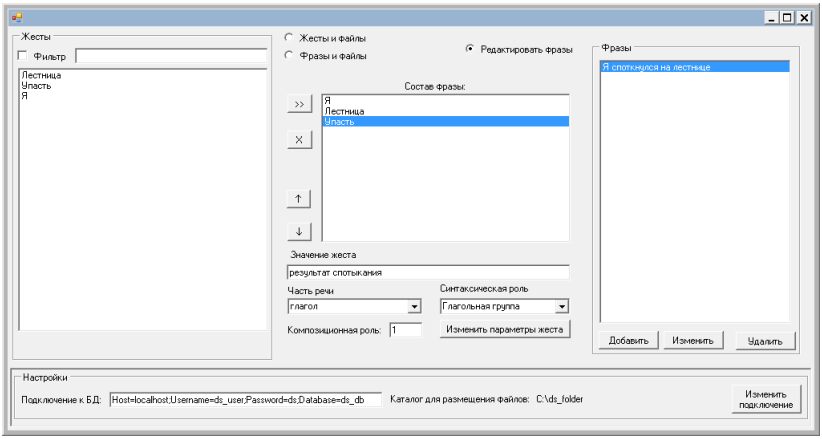


Рис. 2. Главное окно программы

После выбора файла будет доступно окно, в котором перечислены все жесты, входящие во фразу, напротив каждого из них необходимо проставить временные границы начала и конца жеста в секундах. Данное окно показано на рис. 3.

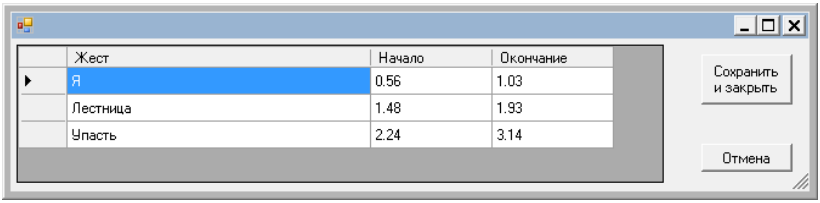


Рис. 3. Окно редактирования временных меток

Заключение

Результатом реализации первого этапа проекта стала разработка системы хранения, наполнения и редактирования набора данных русского жестового языка. На следующем этапе предполагается осуществить создание веб-версии пользовательского интерфейса для обеспечения удобного многопользовательского режима работы на удалённом сервере.

Список литературы

1. Oudah, M. Hand Gesture Recognition Based on Computer Vision: A Review of Techniques / M. Oudah, A. Al-Naji, J. Chahl // Journal of Imaging. 2020. – Vol. 6. – Iss. 8. – P. 73-102.

2. Slovo: Russian Sign Language Dataset / A. Kapitanov [et al.] // Computer Vision Systems. Springer Nature Switzerland. – 2023. – P. 63-73.

3. MediaPipe Hands: On-device Real-time Hand Tracking [Электронный ресурс] / F. Zhang [et al.] // CVPR Workshop on Computer Vision for Augmented and Virtual Reality, Seattle, WA, USA. – 2020. – Режим доступа: https://arxiv.org/abs/2006.10214

4. Online Detection and Classification of Dynamic Hand Gestures with Recurrent 3D Convolutional Neural Networks / P. Molchanov [et al.] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – P. 4207-4215.

5. Kapitanov, A. HaGRID – HAnd Gesture Recognition Image Dataset [Электронный ресурс] / A. Kapitanov, A. Makhlyarchuk, K. Kvanchiani // IEEE/CVF Winter Conference on Applications of Computer Vision. – 2022. – Режим доступа : https://arxiv.org/abs/2206.08219

6. Хнюнин, М. В. Система распознавания жестов в реальном времени с использованием MediaPipe / М. В. Хнюнин, М. Г. Гриф // Информатика: Проблемы, Методы, Технологии. Материалы XXIII Международной научно-практической конференции им. Э. К. Алгазинова. – Воронеж, 2023. – С. 1097-1102.

7. Введение в лингвистику жестовых языков. Русский жестовый язык : [учебник / С. И. Буркова и др. ; ред. С. И. Буркова, В. И. Киммельман]. – Новосибирск : Изд-во НГТУ, 2019. – 354 с.

8. Прозорова, Е. В. Референциальные характеристики именных групп в российском (русском) жестовом языке. Дипломная работа [Электронный ресурс] / Е. В. Прозорова. – М., 2006. – Режим доступа: http://signlang.ru/

9. Зайцева, Г. Л. Жестовая речь. Дактилология: Учебник для студентов высших учебных заведений / Г. Л. Зайцева. – М. : Гуманит. изд. центр ВЛАДОС, 2000. – 92 с.

10. Рогов, Е. В. PostgreSQL изнутри / Е. В. Рогов. – М.: ДМК Пресс, 2022. – 660 с.

11. C# 5.0 и платформа .NET 4.5 для профессионалов / К. Нейгел [и др.]; пер. с англ. Ю. Н. Артеменко. – М.: Диалектика, 2014. – 1440 с.

Принципы построения онтологической модели организационно-технической системы для задачи мониторинга состояния её элементов на основе семантического анализа

Д. А. Михин, email: dimamikhin200184@yandex.ru
Н. С. Толстошеин, email: Nik.tolstoshein@yandex.ru

Тамбовский Государственный Технический Университет

Аннотация. *Рассмотрены основные принципы построения онтологической модели организационно-технической системы для задачи мониторинга состояния её элементов на основе семантического анализа текстовых источников. Обоснована необходимость онтологии как семантического каркаса, обеспечивающего унификацию разнородных сведений и их привязку к элементам системы при оценке состояния. Предложена структура онтологии, ориентированная на получение объяснимых результатов мониторинга и поддержку решений о корректирующих действиях и модернизации состава ОТС.*

Ключевые слова: *системный анализ, семантический анализ организационно-техническая система, онтология, мониторинг состояния.*

Введение

Организационно-техническая система (ОТС) представляет собой совокупность технических средств, персонала, программных компонентов, объединённых общей целью и связями взаимодействия. Для таких систем характерны многоуровневая структура, изменяемость состава и режимов функционирования, а также ограничения по ресурсам, безопасности и регламентам. В системном анализе это означает необходимость фиксировать не только состав элементов, но и их функции, зависимости, условия применения и критерии эффективности работы системы в целом.

При эксплуатации ОТС мониторинг состояния элементов редко может быть построен исключительно на числовых показателях, получаемых в процессе работы системы. Значимая часть используемых сведений о состоянии и перспективах модернизации элементов системы выражена в текстах: актах технического обслуживания, дефектных ведомостях, регламентах, инструкциях производителя, научных статьях, пособиях, и так далее. Эти источники неоднородны: присутствуют

© Михин Д.А., Толстошеин Н.С., 2026

синонимия, аббревиатуры, различия в стиле и полноте описания. Следовательно, для обработки текстовых источников требуется модуль семантического анализа [1], представляющий собой совокупность методов, позволяющих извлекать из текста упоминания объектов, признаков, событий и отношений, а затем использовать их в оценке состояния и системе поддержки принятия решений о модернизации или изменении состава ОТС.

Чтобы перейти от разрозненных сведений к формализованной постановке задачи мониторинга, анализ должен завершаться согласованной моделью предметной области: что считать элементом ОТС, какие состояния компонентов различать, какие признаки и источники допустимы в качестве основания вывода, как учитывать временную динамику, зависимости, и так далее. Для составления модели предметной области используется онтологическая модель, то есть концептуализированное представление информации о какой-либо области реальности, представленное в электронном виде [2]. В контексте мониторинга ОТС онтология становится семантическим каркасом, связывающим показания, тексты и выводы в единую базу знаний.

Задача мониторинга в организационно-технических системах не сводится к фиксации отклонения параметра. Как правило, требуется интерпретация контекста: какой элемент затронут, какие функции он поддерживает, насколько критична деградация, какие зависимости с другими компонентами существуют и какие корректирующие действия могут быть использованы. Именно поэтому на этапе системного анализа важны не только перечни элементов, но и описание их связей, ролей и причинно-следственных зависимостей.

Онтологическая модель выполняет в мониторинге несколько взаимодополняющих функций. Во-первых, она фиксирует структуру системы в терминах «состав – структура – функции», обеспечивая единые определения и классификацию элементов. Во-вторых, она задаёт типизацию состояний и событий, позволяя описывать изменения во времени. В-третьих, онтология позволяет связать наблюдаемые симптомы и измеряемые показатели с элементами и их состояниями, что фактически формирует основу для интерпретации диагностических данных.

Системы мониторинга часто используются специалистами, которые принимают последующие решения, следовательно, вывод о состоянии элемента должен быть обоснован: указаны применённые правила, поддерживающие показатели и конкретные фрагменты документов. Онтология позволяет хранить и предъявлять такую доказательную цепочку в единообразном виде.

1. Принципы построения онтологической модели

Принципы построения онтологии для мониторинга ОТС должны отражать специфику предметной области: большое число компонентов, разнородность источников, динамическая конфигурация и необходимость объяснимых выводов. Важно подчеркнуть, что для данной задачи онтология проектируется не как абстрактная модель знаний, а как часть архитектурного модуля СППР, поддерживающий семантический анализ как структурированных, так и неструктурированных данных [3], а также формирование состояния.

Ниже приведены принципы построения онтологической модели для задачи мониторинга состояния элементов ОТС:

1. Принцип определённости границ и уровня детализации. Границы ОТС и глубина декомпозиции элементов выбираются исходя из целей мониторинга и доступности данных. Системный анализ предполагает выделение уровня, на котором принимаются управленческие решения: например, мониторинг может вестись на уровне подсистем, узлов или конкретных компонентов.
2. Принцип связности с деревом целей и критериями эффективности. Онтология должна поддерживать не только описание объектов, но и связь их состояния с целями функционирования ОТС. На практике это означает наличие понятий для функций/ролей элементов и критериев (качество выполнения работ, время выполнения, надёжность), через которые интерпретируется значимость отклонений. Такой подход превращает онтологию из «словаря терминов» в инструмент системной формализации.
3. Принцип разделения сущностей «объект – состояние – событие – действие». В системном анализе важна временная динамика: элемент (объект) существует на протяжении жизненного цикла, его состояние меняется, события фиксируют факты эксплуатации и обслуживания, а действия отражают вмешательства. Чёткое разделение этих сущностей устраняет логические ошибки и обеспечивает корректное накопление истории.
4. Принцип двуслойности: концептуальный слой + лексико-семантический слой. Семантический анализ работает с реальными формулировками, поэтому помимо понятий предметной области необходим слой терминов: синонимы, аббревиатуры, варианты написания, типовые шаблоны дефектов. Этот слой связывает «как пишут в документах» с «как

это представлено в онтологии», снижая чувствительность мониторинга к вариативности текста.

5. Принцип представления неопределённости и качества извлечения. Результаты извлечения сущностей и отношений из текста часто имеют вероятность/уверенность. Онтология должна позволять хранить оценку уверенности сопоставления, источник и контекст, а также различать гипотезу и подтверждённый факт.
6. Принцип доказательной базы и трассируемости [4]. Каждый вывод о состоянии должен быть объясним: какие показатели, события или фрагменты документов его поддержали. Для этого вводится сущность «свидетельство/основание» и связи, позволяющие построить цепочку объяснения «элемент → симптом/показатель → свидетельство → вывод».
7. Принцип расширяемости и сопровождаемости. ОТС развиваются: появляются новые компоненты и методики, поэтому онтология проектируется модульно, как ядро (универсальные понятия мониторинга) и расширения по подсистемам. С точки зрения системного анализа это обеспечивает устойчивость модели при изменениях и позволяет масштабировать мониторинг без пересмотра всей схемы.
8. Принцип интерпретируемости и объяснимости результатов мониторинга. Система должна формировать выводы, которые можно проверить и обсуждать. Онтология должна поддерживать представление «цепочки» объяснения: элемент → наблюдение/симптом → показатель/событие → основание в документе → вывод о состоянии. Это делает мониторинг пригодным для эксплуатации в организационной среде, где решения должны быть обоснованы для ЛПР, решая проблемы срочности, квалификации и информативности [5].
9. Принцип учёта жизненного цикла и конфигурации элемента. Интерпретация состояния зависит от того, когда элемент введён в эксплуатацию, какие версии и настройки использует, какая история обслуживания и замен накоплена. Поэтому онтология должна позволять описывать версии, конфигурации, интервалы обслуживания и изменения состава системы во времени. В рамках системного анализа данный принцип соответствует учёту изменений структуры и условий функционирования [6].

2. Рекомендуемая структура онтологии для задачи мониторинга состояния элементов ОТС

Перечисленные принципы задают требования к содержанию онтологии, а следующим шагом является переход к структуре модели, в рамках которого необходимо определиться с тем, какие классы и отношения являются обязательными для мониторинга и какие данные должны быть представлены в базе знаний. Ниже приводится ядро онтологии, достаточное для формализации состояния элементов ОТС и их связки с результатами семантического анализа документов.

Рекомендуемая структура онтологии должна напрямую обслуживать контур мониторинга с семантическим анализом, где измеримые данные и тексты преобразуются в единые факты, на основе которых формируются состояния. Для удобства предлагается выделить следующие элементы онтологической модели:

- ядро ОТС, описывающее элементы и функции;
- ядро мониторинга, описывающее состояния и признаки;
- лексико-семантический слой, описывающий документы термины и результаты извлечения;
- слой доказательности, фиксирующий основания и объяснения.

Такая декомпозиция сохраняет системную связь между структурой ОТС и интерпретацией наблюдений.

Ядро онтологии должно включать следующие классы:

- Элемент (Component) – объект контроля, являющийся функциональным элементом организационно-технической системы, например: технический узел, программный модуль.
- Функция/роль (Function/Role) – назначение элемента, позволяет интерпретировать критичность отклонений.
- Показатель/параметр (Metric/Parameter) – измеряемая характеристика, которая используется для формализации признаков состояния.
- Симптом/проблема (Symptom/Issue) – наблюдаемое проявление дефекта или деградации компонента ОТС.
- Состояние (State) – типизированная оценка состояния элемента с временной привязкой.
- Событие (Event) – инцидент, отказ, обслуживание, обновление, для фиксирования фактов жизненного цикла и контекста изменения состояния элементов ОТС.
- Действие (Action) – используется для указания взаимодействия с компонентами системы, таких как диагностика, настройка, ремонт, замена, корректировка методики.

– Ограничение (Constraint) – используется для внедрения условий и запретов, таких как совместимость, регламент, безопасность, ресурсные лимиты.
– Основание/свидетельство (Evidence) – ссылка на документ, фрагмент текста, запись журнала или значение показателя, поддерживающее вывод и являющиеся обоснованием для модернизации элемента или структуры ОТС.

Для формирования целостной картины состояния важны отношения между классами. На уровне структуры достаточно выделить типовые связи, обеспечивающие как аналитическую пригодность, так и объяснимость связей:

– структурные связи: partOf (часть-целое), belongsTo (принадлежность к подсистеме), dependsOn/interactsWith (зависимость/взаимодействие);
– связи мониторинга: hasState (элемент находится в состоянии), indicatedBy (состояние определяется показателем), relatesTo (причина относится к элементу);
– временные связи: causedBy/precededBy (логическая/временная связность событий), changedBy (состояние изменено действием/событием);
– связи объяснимости: evidenceSupports (основание поддерживает вывод), extractedFrom (основание получено из источника), limitedBy (действие ограничено условиями).

В результате рекомендуемая структура онтологии обеспечивает замыкание цикла «семантический анализ → факты → состояние → объяснение». Именно это отличает онтологическую модель мониторинга ОТС от абстрактной онтологии предметной области: она содержит сущности и связи, непосредственно поддерживающие извлечение и интерпретацию текстовых свидетельств, что направлено на решение задачи мониторинга состояния элементов ОТС. Дополнительно в рамках повышения качества принятия решений целесообразно применение метода нечеткой логики, что позволит снизить процент ошибок [7].

Заключение

Онтологическая модель ОТС для мониторинга состояния элементов является не только средством представления знаний, но и результатом системного анализа предметной области. Она фиксирует согласованные определения элементов, функций, состояний и событий, обеспечивает связь показателей и текстовых свидетельств, а также поддерживает трассируемость и объяснимость выводов. Сформулированные принципы и предложенная структура ядра онтологии позволяют перейти от разрозненных описаний состояния к формализованной модели, которая

затем используется при формировании требований к системе мониторинга и последующей поддержке принятия решений о корректирующих действиях и модернизации.

Литература

1. Михин, Д. А. Перспективы использования семантического анализа в задачах мониторинга состояния и модернизации организационно-технических систем / Д. А. Михин, Н. С. Толстошеин // Проблемы техногенной безопасности и устойчивого развития : материалы двадцатой межвузовской научной студенческой конф. Ассоциации «Объединенный университет им. В. И. Вернандского» (Тамбов, 16 октября 2025 г.). – Тамбов, 2025. – С. 44-47
2. Горшков, С. В. Введение в онтологическое моделирование / С. В. Горшков – 2-е изд., перераб. и доп. – Екатеринбург. : ООО «ТриниДата», 2018. – 150 с.
3. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика : учебник / Е. И. Большакова [и др.] – М : МИЭМ, 2011. – 272 с.
4. Бычков, М. Ю. Основные подходы к обеспечению трассируемости требований / М. Ю. Бычков, С. В. Котлинской // E-Scio. – Автономная некоммерческая организация «Информационная Мордовия», 2020. – Т. 11. – С. 91-99.
5. Сыроид, А. В. Структура информационной системы принятия управленческих решений / А. В. Сыроид // Труды ТГТУ: сб. научных статей молодых ученых и студентов. – Тамбов: Тамб. гос. техн. ун-т, 2008. – В. 21. – С. 173-176.
6. Тырсин, А. Н. Теория систем и системный анализ : учебное пособие / А. Н. Тырсин ; УрСЭИ АТиСО. – Челябинск, 2002. – 128 с.
7. Шабанов, К. Б. Применение методов интеллекутального анализа данных для повышения качества принятия решений при управлении ресурсами информационной медиасистемы / К. Б. Шабанов, В. В. Алексеев // Вестник Тамбовского государственного технического университета. – Тамбов, 2021. – Т. 27. – С. 14-20.

Формирование информационного критерия классификации входных данных для большой языковой модели

К. И. Пашигорев [1], e-mail: kpashigorev@yandex.ru
В. А. Головской [2], e-mail: golovskoy_va@mail.ru

[1] Высшая школа экономики
[2] Краснодарское высшее военное училище

Аннотация. *Представлены промежуточные результаты исследований по созданию алгоритма поиска критерия, используемого большой языковой моделью при сокращении данных, необходимых ей для формирования новости из технических логов. Рассматривается итеративное применение модели для классификации данных, которые в дальнейшем будут поданы ей на вход. Сформулированы гипотезы о содержании искомого критерия классификации, учитывающем информативность ансамблей исходных данных. Проведено сравнение результатов работы модели с классификацией экспертами. Приведены результаты эксперимента, проведенного в рабочих условиях компании.*

Ключевые слова: *Энтропия Шеннона, информационная энтропия, количество информации, модель, промпт, машинное обучение, большие языковые модели, LLM.*

Введение

Успехи в исследованиях больших языковых моделей (LLM) индуцировали рост их популярности [3. 2] при использовании в различных областях бизнеса, науки, техники и развлечений. В рамках исследований на стыке науки и бизнеса авторами предложена система [3. 3], которая будет получать из смежных систем ИТ-ландшафта организации технические данные, представленные логами, и преобразовывать их с помощью LLM в новость – человекочитаемый текст, посвященный описанию дата-продукта (ДП), под которым понимается имеющий конечное описание информационный продукт – программа, приложение и т.д. Далее каждая новость должна будет рекомендоваться конкретным корпоративным пользователям из множества на основании их профессиональных интересов. На рис. 1 представлена компонентная диаграмма проектируемой системы [3. 4], построенная первым автором на основании известных [3. 5] принципов и подходов, и методологии [3. 6].

Авторами проводятся дальнейшие исследования по оптимизации ресурсов системы, представленной на рис. 1.

© Пашигорев К.И., Головской В.А., 2026

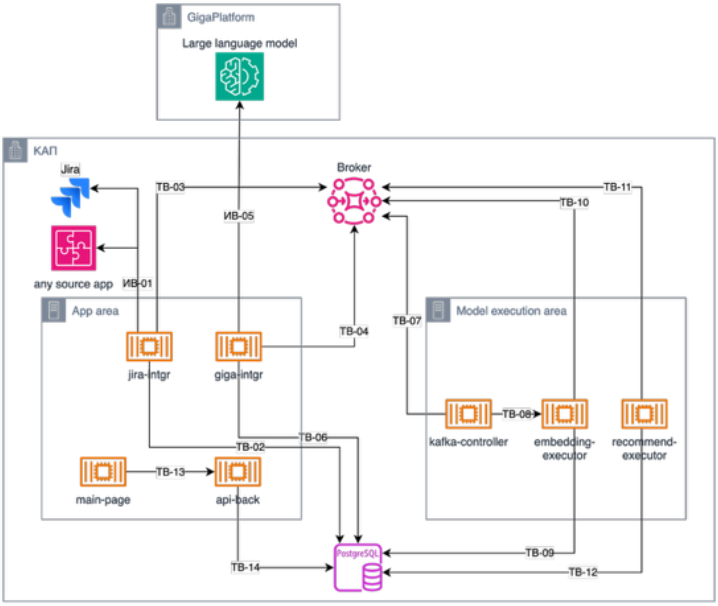


Рис. 1. Компонентная диаграмма системы

1. Постановка задачи

Одним из направлений оптимизации является итеративное использование LLM, формирующей новость, для предварительного сокращения количества используемых ею для формирования новости ключей и их значений. Здесь и далее под ключом подразумевается элемент структуры данных dict языка программирования python, где каждому ключу в структуре данных соответствует определенное значение. На основании такого массива элементов «ключ:значение» $\left\{(key:value)_j\right\}$ приложение формирует новость для пользователя. Таким образом, предмет настоящего исследования может быть сформулирован как исследование нетривиального семантического свойства (НСС) θ_{KD} из множества НСС LLM $\Theta_{LLM}=\left\{\theta_i\right\}$, $i=\overline{1,N_I}$, реализующего сокращение ключей перед формированием из них новости. Объект исследования – формирование новости приложением из состава системы. Здесь и далее под приложением будет пониматься отдельный

микросервис, взаимодействующий с LLM, а под системой – совокупность всех микросервисов, брокеров и баз данных, вне зависимости от их назначения. Цель исследования – определение критерия KD сокращения количества ключей проприетарной LLM при проявлении НСС θ_{KD}.

Особенность исследования состоит в том, что используемая LLM предназначена и обучена для решения задач, отличающихся от предлагаемой авторами, т.е. при её построении и обучении была задача сформировать подмножество требуемых НСС $\Theta` \subset \Theta_{LLM}$, $\Theta` = \{\theta_r\}$, $\theta_{KD} \notin \Theta`$. Исследование реализуется в русле актуального направления [3. 7] повышения интерпретируемости систем искусственного интеллекта.

Ключи, полученные при реальном функционировании системы, представлены в таблице. В целях исследования вопросов интерпретируемости ключи были экспертно классифицированы на две группы: влияющие на качество новости и не влияющие, что отражено в графе 4 таблицы. Первая группа образовала экспертный массив ключей (ЭМК), представляющий собой подмножество исходного, полного массива ключей $EMK \subset MK$, $MK = \{(key : value)_j\}$, $j = \overline{1,42}$. И как следствие массив из исключенных ключей $DMK \subset MK \,\&\, DMK \not\subset EMK$, $DMK \bigcup EMK = MK$, $DMK \bigcap EMK = \varnothing$. Перед LLM также была поставлена задача формирования сокращенного массива ключей (СМК) $CMK \subset MK$ путем удаления из MK не влияющих на качество новости ключей, что с теоретико-информационных позиций может быть записано выражением

$$A_{LLM}^{\theta_{KD}}(MK, prompt_l) = CMK,$$

где $A_{LLM}^{\theta_{KD}}$ – запись алгоритма функционирования LLM, акцентирующая внимание на анализируемом НСС θ_{KD}; $prompt_l$ – промпт, побуждающий LLM к формированию СМК, $l = \overline{1,2}$.

Промпты $prompt_l$, как и метрики качества новости, являются предметом другого этапа исследования, и не приведены ввиду ограничений на объем статьи.

Ставится задача построить алгоритм поиска критерия KD, используемого LLM для удаления ключей при формировании ею CMK из MK. Данная задача может рассматриваться как композиция задачи определения критерия классификации ключей и задачи поиска объекта во множестве, сформированном классификационными признаками. Обе

подзадачи рассматриваются традиционно [3. 7, 3. 8] с позиций теории информации.

1.

Экспертный массив ключей

j	Ключ словаря (key_j)	Тип данных	Включен в ЭМК
1	Функциональный блок трайба	Строка	Да
2	Трайб	Строка	Да
3	Название ДП	Строка	Да
4	ДП	Строка	Да
5	Категория в каталоге	Строка	Да
6	Тип ДП	Строка	Да
7	Кластер	Строка	Да
8	Платформа схемы	Строка	Да
9	Схема базы данных	Строка	Да
10	Количество тэгов	Строка	
11	Имя тэга	Строка	
12	Табличный номер Менеджеров данных ДП	Строка	
13	Менеджеры данных ДП	Строка	Да
14	Uuid	UUID	
15	Ревизия структуры	Булево	
16	Внешний текстовый код	Строка	
17	Последний Статус	Строка	Да
18	Статус ДП	Строка	Да
19	Дата назначения МД	Дата	Да
20	Время изменения s2t с ФМ витрины / реплики	Дата	Да
21	Время снятия признака окончания разметки витрины / реплики	Дата	Да
22	Время окончания разметки витрины / реплики	Дата	Да
23	Время создания ДП	Дата	Да
24	Время активации ДП	Дата	Да
25	Дата появления ДП в метаданных	Дата	Да
26	Дата последнего изменения ДП	Дата	Да
27	Время появления ДП в каталоге	Дата	Да
28	Время публикации ДП в каталоге	Дата	Да
29	Скорость публикации ДП (в рабочих днях)	Число	Да
30	Длительность работы МД с ДП (в рабочих днях)	Число	Да
31	Флаг публикации	Булево	
32	Флаг скорости публикации версии ДП	Булево	
33	Имя системы-приемника	Строка	Да
34	Площадка распространения	Строка	Да
35	Статус ДП «не Архив»	Строка	Да
36	is_active_dp	Булево	
37	is_active_schema	Булево	
38	Заполнена внешняя сущность	Строка	Да

j	Ключ словаря (key_j)	Тип данных	Включен в ЭМК
39	Флаг «Данные метаданных ДП»	Булево	
40	Флаг «Данные внешних метаданных»	Булево	
41	Флаг «Данные каталога»	Булево	
42	SLA	Булево	

Сформировано 4 гипотезы о содержании критерия KD. Общим для всех гипотез является Высказывание «Классификационный критерий – определенное значение энтропии ключей для формирования СМК», являющееся предположением об использовании LLM для формирования новости ключей, обладающих определенными информационными характеристиками, что согласуется как с классическими положениями теории информации, так и с современными подходами к машинному обучению [3. 8, 3. 9]. В контексте текущей задачи энтропия является количественной оценкой разнообразия значений конкретного ключа, вследствие чего сделано общее для всех гипотез предположение об удалении LLM из MK малоинформативных ключей.

Гипотеза №1. Высказывание. Допущение: взаимосвязь между ключами отсутствует или не играет роли.

Гипотеза №2. Высказывание. Допущение: взаимосвязь между ключами присутствует и играет роль.

Гипотеза №3. Высказывание. Допущение: взаимосвязь между ключами присутствует в виде простых причинно-следственных связей и играет роль.

Гипотеза №4. Высказывание. Допущение: взаимосвязь между ключами присутствует в виде сложных причинно-следственных связей и играет роль.

В работе представлены только результаты проверки Гипотезы №1, для чего спланирован и проведен эксперимент в реальных условиях функционирования корпоративной системы.

2. Описание эксперимента

План эксперимента содержит приведенные ниже этапы, предполагает исследование LLM с позиций черного ящика.

1. Трансформировать датасет в словарь dict.
2. Вычислить энтропию для каждого ключа в исходном словаре dict, для чего по результатам обработки 17863 ДП рассчитать значение информационной энтропии для каждого key_j следующим образом:

$$H_j = -\sum_{m=1}^{17863} p(value_{j,m}) \log_2 \left(p(value_{j,m}) \right),$$

где $value_{j,m}$ – m-е значение $value$ ключа key_j; $p(value_{j,m})$ – частотность m-го значения $value_{j,m}$ ключа key_j. Необходимо отметить, что $value_{j,m}$ принимает одно из M^j уникальных значений key_j, и в общем случае $M^j \neq M^{j+1}$, $j = \overline{1,41}$.

3. Сформировать ЭМК из словарей dict.
4. Подготовить промпты $prompt_l$, $l = \overline{1,2}$, с инструкцией по сокращению ключей и вариациями:

– с инструкцией о рассуждении при каждом удалении ключа;
– без инструкций о рассуждении при каждом удалении ключей.

5. Получить на выходе LLM согласно **Ошибка! Источник ссылки не найден.** четыре варианта СМК: $CMK_{1a}, CMK_{1b}, CMK_{2a}, CMK_{2b}$.
6. Вычислить мощности N_{LP} подмножеств ключей $LP \subset MK$, $LP = DMK \bigcap CMK$, добавленных LLM при построении $CMK_{1a}, CMK_{1b}, CMK_{2a}, CMK_{2b}$:

$$N_{LP}^{1a} = \left|LP_{1a}\right|, key_j^{'} \in LP_{1a} \Leftrightarrow key_j^{'} \notin EMK \wedge key_j^{'} \in CMK_{1a} \wedge key_j^{'} \in DMK_{1a},$$
$$N_{LP}^{1b} = \left|LP_{1b}\right|, key_j^{'} \in LP_{1b} \Leftrightarrow key_j^{'} \notin EMK \wedge key_j^{'} \in CMK_{1b} \wedge key_j^{'} \in DMK_{1b},$$
$$N_{LP}^{2a} = \left|LP_{2a}\right|, key_j^{'} \in LP_{2a} \Leftrightarrow key_j^{'} \notin EMK \wedge key_j^{'} \in CMK_{2a} \wedge key_j^{'} \in DMK_{2a},$$
$$N_{LP}^{2b} = \left|LP_{2b}\right|, key_j^{'} \in LP_{2b} \Leftrightarrow key_j^{'} \notin EMK \wedge key_j^{'} \in CMK_{2b} \wedge key_j^{'} \in DMK_{2b}.$$

7. Сформировать критерий KD.

3. Результаты эксперимента

Полученные в результате расчета значения энтропии приведены на рис. 2, анализ которых позволил сделать вывод о предполагаемой LLM информативности ключей, не включенных в ЭМК:

– Ключи с низкой энтропией: «Флаг публикации», «is_active_schema», «Флаг «Данные метаданных ДП», «Флаг «Данные внешних метаданных», «Флаг «Данные каталога», «SLA»;
– Ключи со средней энтропией: «Количество тэгов», «Имя тэга», «Табличный номер Менеджеров данных ДП», «Ревизия структуры», «Флаг скорости публикации версии ДП», «is_active_dp»;
– Ключи с высокой энтропией: «uuid», «Внешний текстовый код».

На рис. 3-6 представлены значения N_{LP}^{1a}, N_{LP}^{1b}, N_{LP}^{2a}, N_{LP}^{2b}, отложенные по оси абсцисс, и процент случаев, когда были получены эти величины, отложенный по оси ординат.

Анализ рис.3, где в эксперименте использовался промпт №1 без рассуждений, позволяет сделать вывод, что LLM в 92% случаев считает полезными еще 3 ключа, не включенных в ЭМК ($N_{LP}^{1a} = 3$), и в 8% случаев считает полезным 1 ключ ($N_{LP}^{1a} = 1$), не включенный в ЭМК.

Ключ	Энтропия	Уник.знач.	Всего
Внешний текстовый код	13.6911	16674	17863
Дата появления дата-продукта в метаданных	13.6082	16507	17863
Дата последнего изменения ДП	12.5930	13839	17863
Время публикации ДП в каталоге	12.2965	14619	17863
Время изменения s2t с ФМ витрины/реплики	12.2473	9570	17863
uuid	12.0921	6264	17863
Время создания ДП	12.0791	6237	17863
Название ДП	11.9943	6990	17863
Время активации ДП	11.1813	13131	17863
Длительность работы МД с ДП (в раб днях)	11.1635	13028	17863
Время окончания разметки витрины	11.0934	12984	17863
Скорость публикации ДП (в раб днях)	9.9616	11777	17863
Время появления дата-продукта в каталоге	9.5068	3691	17863
Схема базы данных	8.9010	3549	17863
Имя тэга	8.8936	2961	17863
ДП	8.7275	3455	17863
Табличный номер Менеджеров данных ДП	6.8418	916	17863
Менеджеры данных ДП	6.8417	915	17863
Дата назначения МД	6.2158	710	17863
Кластер	4.9816	117	17863
Трайб	4.8875	127	17863
Площадка распространения	3.5864	150	17863
Функциональный блок трайба	3.2564	21	17863
Категория в каталоге	2.9966	17	17863
Ревизия структуры	2.9531	62	17863
Количество тэгов	2.0450	32	17863
Имя системы-приемника	1.6373	157	17863
Флаг скорости публикации версии ДП	1.6085	4	17863
Тип ДП	1.5549	3	17863
is_active schema	1.5058	3	17863
Платформа схемы	1.4653	6	17863
Последний статус	1.3644	7	17863
SLA	1.2022	3	17863
is_active dp	1.1675	3	17863
Заполнено внешняя сущность	1.1636	3	17863
Статус ДП	1.0899	3	17863
Флаг публикации	0.7506	3	17863
Статус ДП "не Архив"	0.6248	2	17863
Флаг «Данные метаданных ДП»	0.2741	2	17863
Флаг «Данные каталога»	0.2427	2	17863
Время снятия признака окончания разметки	0.2000	194	17863
Флаг «Данные внешних метаданных»	0.0142	2	17863

Рис. 2. Результаты расчета энтропии Шеннона

На рис.4 представлен результат с промптом №1 с рассуждением, где в 83% случаев все ключи, вошедшие в СМК, в полном объеме включены в ЭМК ($N_{LP}^{1b} = 0$), но в 17% случаев LLM включает в СМК 1 ключ, не входящий в ЭМК ($N_{LP}^{1b} = 1$). На рис.5 представлен результат с промптом №2 без рассуждения, где в 100% случаев все ключи, вошедшие в СМК, входят в ЕМК ($N_{LP}^{2a} = 0$). На рис.6 представлен результат с промптом №2 с рассуждением, где в 92% случаев LLM включает в СМК 2 ключа, не

входящих в ЕМК ($N_{LP}^{2b}=2$), и в 8% случаев LLM включает в СМК 3 ключа, не входящих в ЕМК ($N_{LP}^{2b}=3$).

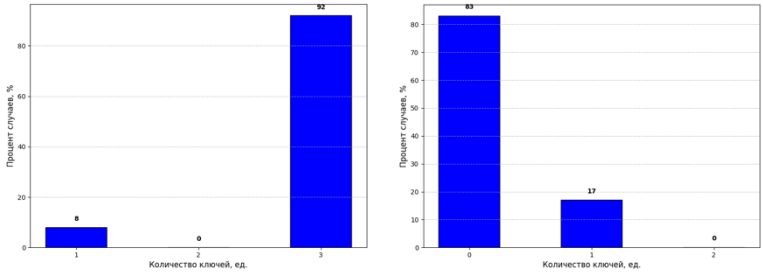


Рис. 3. Величины N_{LP}^{1a}

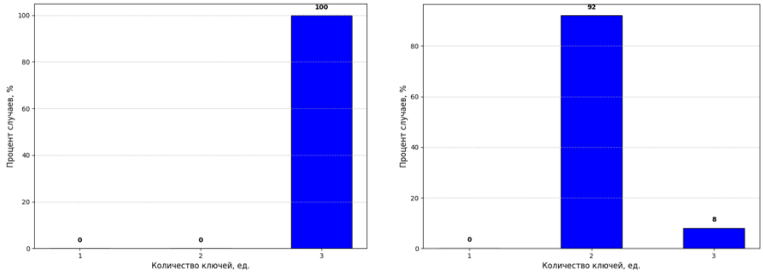


Рис. 4. Величины N_{LP}^{2a}

Обсуждение результатов

Из полученных результатов можно сделать вывод, что сокращение ключей LLM, в целом, соответствует действиям экспертов при определении ЭМК. При этом имеют место отклонения в 1-3 ключа. Данный вывод обуславливает предварительно содержание критерия $KD=\langle S,H_s\rangle$, где S – количество пороговых значений энтропии, H_s – значения энтропии пороговых значений, $s=\overline{1,S}$.

Пороги величин значений энтропии для классификации могут быть определены различными подходами, однако предварительно для рассматриваемого примера они имеют следующие значения:

– $S=2$;

– Нижние 25%: Низкая энтропия, $H_s < 1.39$ бит – это 25-й перцентиль, т.е. 25% энтропий всех ключей находятся ниже этого порога (1.39 бит). В представленном ЭМК это 11 ключей внизу массива, представленного на рис. 2.
– Верхние 25%: Высокая энтропия: $H_s > 10.81$ бит – 75-й перцентиль – т.е. 25 % энтропий всех ключей находятся выше этого порога. В представленном ЭМК это 11 ключей вверху массива на рис.2.

Заключение

Большая языковая модель эффективно справляется с задачей оптимизации используемых ключей и без применения классических алгоритмов формирует список ключей, которые достаточно стабильно попадают в выбранный ЭМК.

В дальнейших исследованиях планируется осуществить проверку гипотез № 2-4 и уточнение критерия KD. При проверке гипотезы № 2 предлагается исследовать взаимные зависимости между ключами и учесть их путем замены **Ошибка! Источник ссылки не найден.** на выражение для вычисления условной энтропии. Также представляет интерес рассмотрение информационного критерия с позиций Колмогоровской сложности, поскольку обеспечивает оценивание не только разнообразия, но и предсказуемости.

Литература

1. Das, B. C. Security and privacy challenges of large language models: A survey / B. C. Das, M. H. Amini, Y. Wu // ACM Computing Surveys. – 2025. – Т. 57. – № 6. – P. 1-39.

2. Пашигорев, К.И. Модель рекомендательной системы на основе технических событий / К. И. Пашигорев, А. О. Резников // Бизнес-информатика. – 2025. – Т. 19. – № 1. – С. 7-21.

3. Пашигорев, К.И. Агентно-ориентированная модель обработки логов, основанная на синтезе методов промптинжиниринга и цепочки рассуждений / К.И.Пашигорев // III Научная конференция факультета компьютерных наук НИУ ВШЭ (Москва, 26-29 октября 2025 г.) Режим доступа: https://cs.hse.ru/mirror/pubs/share/1096375414.pdf

4. Разработка и применение мультиагентных систем в корпоративной среде [Электронный ресурс]. Режим доступа: https://b2b.giga.chat/multi-agent-system

5. Rah! recsys–assistant–human: A human-centered recommendation framework with llm agents / Y. Shu [et al.] // IEEE Transactions on Computational Social Systems. – 2024. – Т. 11. – № 5. – P. 6759-6770.

6. Доверенный искусственный интеллект: вызовы и перспективные решения / Д. Ю. Турдаков [и др.] // Доклады Российской академии наук. Математика, информатика, процессы управления. – 2022. – Т. 508. – № 1. – С. 13-18.

7. Прокаев, А. Н. Принцип максимума энтропии в теории принятия решений / А. Н. Прокаев // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. – 2024. – Т. 20. – № 2. – С. 154-169.

8. Акимова, П.Р. О влиянии энтропии внутри класса на точность и полноту классификации / П.Р. Акимова [и др.] // Информационные процессы. – 2021. – Т. 21. – № 3. – С. 149–161.

9. Khalid, S. A survey of feature selection and feature extraction techniques in machine learning / S. Khalid, T. Khalil, S. Nasreen // 2014 science and information conference // IEEE – 2014. – P. 372-378.

Использование программных инструментов автоматического анализа текстов на русском языке на базе алгоритмических методов и инструментов на основе нейросетевых технологий для решения производственных и научных задач

Е. В. Полицына, email: kathrin.beaver@mail.ru
С. А. Полицын, email: pul_forever@mail.ru

Аннотация. *В докладе представлены результаты анализа особенностей применения инструментов автоматического анализа текстов на русском языке на основе алгоритмических методов и нейросетевых технологий, предложены пути решения существующих проблем использования больших языковых моделей для решения производственных и научных задач автоматизации обработки естественно-языкового текста на основе их совместного использования с алгоритмическими методами.*

Ключевые слова: *Компьютерная лингвистика, алгоритмические методы обработки текста, большие языковые модели*

Введение

Объем обрабатываемых человечеством данных достиг громадных масштабов и маловероятно, что эта тенденция будет меняться в обозримое время. Однако, с этими данными, в первую очередь текстовыми, люди работают очень по-разному. Для решения задач получения ответа на конкретные вопросы, получения представления о чем-либо, классификации поступающих заявок, отзывов, данных со средств автоматизации различных бытовых и технологических процессов и др. необходимо максимально быстро обрабатывать большие объемы данных с приемлемым качеством. Тогда как при решении научно-исследовательских задач лингвистического анализа текстов и исследования естественного языка, анализа и подготовки юридических и технических документов, законотворчества и др. важна максимальная точность и детерминированность результатов.

На протяжении десятков лет развиваются методы автоматического анализа текстов, в том числе на русском языке, среди них всегда были алгоритмические методы, основанные на правилах и закономерностях естественного языка, статистические методы и их комбинации. В первую очередь эти методы получили распространение для исследовательских и научных задач, далее стали использоваться в специализированном

© Полицына Е.В., Полицын С.А., 2026

программном обеспечении для обработки текстов (поисковики, агрегаторы новостей и др.), оставаясь практически не востребованными в бытовой сфере.

Последние годы бурное развитие получили статистические методы, основанные на машинном обучении, большие языковые модели (Large Language Model, LLM) получили распространение не только в научной сфере и разработке крупного промышленного ПО, но и повседневной жизни, существенно упрощая решение многих задач поиска информации, получения систематизированной информации по запросу, перевода, генерации текстов и изображений и т.д.

При этом такой явный прорыв в области автоматизации обработки больших текстов имеет и некоторые особенности: модели гарантируют получение некоторой информации с некоторой точностью, получение достоверной информации с ссылками на авторитетные источники, которую можно использовать при выполнении профессиональных задач, для принятия ответственных решений и для полностью автоматических операций в программных системах, требует от пользователей гораздо больше времени, усилий и знаний в области использования LLM, включая природу формирования результатов, нежели быстрое получение первичного результата.

Исходя из этого, по-прежнему актуальным является комбинирование существующих методов автоматического анализа текста, но уже с новыми мощными инструментами, основанными на статистических подходах.

Программные инструменты на базе алгоритмических методов анализа текстов

Практически с появлением ЭВМ стали развиваться «классические» методы анализа текстовой информации. В задачи алгоритмических методов входит извлечение, классификация и получение какой-либо информации из текста по заранее определённым правилам. Разработка этих правил также часто входит в состав методов, а точность полученного результата напрямую зависит полноты этих правил и алгоритма их применения.

Среди программных инструментов на базе алгоритмических методов можно выделить как узконаправленные, решающие одну определенную задачу, так и более общие, которые могут быть использованы для решения целого ряда задач анализа текста, используя общий подход [1].

Ключевые алгоритмические методы анализа текста, включают в себя следующие [2]:

1. Предобработка текста: методы токенизации, стемминга и лемматизации, разметки текста (POS), синтаксического разбора и др. Даже в больших промышленных системах стадия предобработки текста часто выполняется с помощью инструментов, основанных на алгоритмических методах [3].
2. Статистический анализ и моделирование текста: метод TF-IDF – пример простого, но мощного статистического метода для оценки важности термина; методы тематического моделирования, методы векторного представления текста (Word2Vec) [4].
3. Извлечение информации из текста: методы извлечения именованных сущностей (NER), получения ключевых фраз, методы реферирования, информационный поиск в тексте. Различные алгоритмические методы находят свое применение в широко применяемых инструментах таких как ElasticSearch, Lucene и др.
4. Методы определения тональности текста, классификации текстов по заданным критериям и др.

Современные программные инструменты для использования этих методов обычно включают в себя реализацию сразу нескольких из них в соответствии с выбранной направленностью. Можно выделить следующие: NLTK (Natural Language Toolkit) – фреймворк, который содержит реализации множества классических алгоритмов (стеммер Портера, токенизаторы, NER и т.д.); spaCy – промышленный инструмент, включает в себя методы токенизации, разметки текста, NER; Gensim реализует методы тематического моделирования, различные методы векторизации текста, реализацию алгоритма TextRank; scikit-learn – самый популярный инструмент алгоритмической поддержки создания инструментов на базе машинного обучения для языка Python и др.

Основными преимуществами алгоритмических методов анализа текста, которые позволяют им занимать свою нишу на современном этапе развития технологий являются:

– Интерпретируемость и управляемость. Всегда можно проанализировать, почему, например, документу присвоена та или иная тема или по каким ключевым словам отзыв отнесен к положительным, и настроить параметры алгоритмов, если требуется.

– Эффективность на малых данных. Алгоритмические методы не требуют дополнительной информации для работы и используют только входной текст. Логистическая регрессия на TF-IDF может показать отличные результаты, используя всего несколько сотен размеченных примеров.

– Скорость и ресурсы. Алгоритмические методы работают быстрее и требуют меньше вычислительных ресурсов, чем LLM.

Многие задачи предобработки, первичного поиска документов, базовой классификации или извлечения фактов по шаблону эффективно и экономично решаются алгоритмическими методами, а значит их развитие совместно с нейросетевыми методами является актуальной задачей.

Программные инструменты на основе больших лингвистических моделей

С появлением LLM происходит переход от решения задач с помощью алгоритмов к решению задач с помощью «управляемого» рассуждения, задачей разработчика становится задание направления этого рассуждения с помощью контекста, инструкций и инструментов.

Принципиальное отличие больших языковых моделей от классических NLP-инструментов заключается в их природе: они являются не алгоритмическим решением конкретной задачи, а вычислительной базой для обеспечения выполнения операций над семантическим содержанием текста. Инструменты на базе LLM расширяют круг решаемых лингвистических задач.

В первую очередь это задачи «понимания» и генерации текста по инструкции. Модель выступает как универсальный исполнитель инструкций на естественном языке. В рамках этой задачи обычно рассматриваются следующие подзадачи: анализ тональности с пониманием контекста; извлечение структурированной информации; классификация с нуля по описанию категорий; перефразирование, упрощение, формальное или неформальное преобразование стиля; генерация текста по шаблону (письма, отчеты, креативный контент, код) и др. Ключевые инструменты включают в себя:

– API (OpenAI GPT, Gemini): Хорошо подходят для прототипирования и сред, где не требуется работа с конфиденциальными данными.
– Фреймворки для промпт-инженерии (LangChain, Guidance), которые позволяют оптимизировать запросы, создавать сложные шаблоны с условной логикой [5].

Второй важной задачей, решаемой с помощью инструментов на базе LLM, является ведение диалога на основе знаний из предметной области, а также поддержка принятия решений, LLM становится "ядром", которое «отвечает» на вопросы, «анализирует», «систематизирует» и планирует последовательность действий, используя внешние инструменты (API, код, поиск).

Для решения этой задачи служат инструменты, построенные на архитектуре RAG – Retrieval-Augmented Generation является востребованной для корпоративных решений, т.к. она частично решает проблему «галлюцинаций» и актуальности знаний LLM. Ключевые инструменты на базе этой архитектуры включают в себя: LlamaIndex для индексирования, хранения и извлечения контекста из разнородных данных; LangChain позволяет LLM «решать», когда и какой инструмент использовать для поиска (например, поиск в векторной БД, в обычной БД или через поисковик) [6]; векторные базы данных (Chroma, Weaviate, Qdrant, Pinecone) хранят представления текстов для семантического поиска [7].

Важной новой задачей, которая стала очень актуальной именно в контексте бурного роста применения LLM стала адаптация модели под узкую предметную область. Часто модель нужно "дообучить" на специфических данных предметной области для повышения качества результатов, а также проверить, является ли результат достоверным для принятия решения из-за стохастической природы LLM. Для этого созданы специальные инструменты для оценки вывода модели (RAGAS, TruLens, DeepEval), они автоматически генерируют метрики для RAG-систем (например, метрика соответствия ответа контексту) [8].

Современный уровень развития компьютерной лингвистики трансформирует лингвиста в архитектора сложных когнитивных систем. Требуется комбинация детерминированных алгоритмов (для поиска, проверки) с генеративными возможностями LLM, чтобы создавать приложения, которые не просто обрабатывают текст, а «понимают» и решают задачи в предметной области.

Эксплуатация и обеспечение безопасности использования LLM-решений

При использовании LLM-решений в связи с общей низкой детерминированностью результатов зачастую остро встают вопросы безопасности. Помимо общепринятых мер по обеспечению безопасности программного обеспечения: использования изолированной среды развертывания и использования, контроля сетевой безопасности с использованием специализированных средств, аутентификации и авторизации API, а также настройки средств мониторинга работоспособности, необходима защита от атак, ориентированных на LLM-решения. Основными видами атак и способами защиты от них на основании OWASP Top 10 for Large Language Model Applications 2025 [9] являются:

Prompt Injection. Тип атаки, когда пользовательские запросы изменяют поведение или вывод LLM непредусмотренным образом,

заключается в прямом внедрении команд в запрос через контекст диалога. Защита реализуется путем разделения данных/инструкций, валидации и фильтрации входных данных, ограничения поведения модели и контроля привилегий.

Утечка конфиденциальной информации. Конфиденциальная информация может повлиять как на LLM, так и на контекст ее применения. Возможна утечка персональных данных и раскрытие алгоритмов. Защита возможна также путем валидации входных запросов и контроля привилегий, ограничение источников данных, предварительной очистки данных для обучения модели (удаление, маскирование, шифрование чувствительных данных), а также обучение пользователей безопасному использованию инструментов на базе LLM.

Отравление данных и модели. Отравление данных происходит, когда данные, используемые на подготовительных этапах, используются для введения уязвимостей или искаженных представлений данных. Использование непроверенных данных для обучения также увеличивает риск создания ошибочных выводов. Верификация данных для обучения, проверка источников данных обязательна для защиты от подобных атак, кроме того, необходимо версионирование данных для отслеживания изменений и возможности отката версий при выявлении проблем.

Некорректная обработка выходных данных. Некорректная обработка выходных данных относится к недостаточной проверке, очистке и обработке информации, генерируемой LLM, перед ее передачей другим компонентам и системам. Основные риски заключаются в том, что выходные данные напрямую передаются в другие системы в качестве исполняемых команд или кода приложений (system-shell, SQL-запросы, генерация запросов с подстановкой параметров и др.). Требуется тщательная валидация данных и использование систем мониторинга для выявления аномалий.

Неограниченное потребление ресурсов. Неограниченное потребление описывает риск, когда LLM или сопутствующие сервисы используют вычислительные и сетевые ресурсы без ограничений. Отсутствие ограничений на длину генерируемого текста, время обработки запросов или объем обрабатываемых данных может привести к отказу в обслуживании, чрезмерному расходу ресурсов и даже непредвиденным финансовым затратам, особенно в облачных средах. Для защиты от этого необходимо: управление распределением ресурсов, ограничение времени выполнения и частоты запросов, комплексный мониторинг.

Помимо приведенных атак существует и ряд других, ориентированных на конкретную предметную область или особенности

использования [9-10], поэтому универсальных комплексных средств защиты решений с использованием LLM не существует, что существенно усложняет их использование для решения производственных задач и требует наличия отдельных специалистов в области эксплуатации и безопасности LLM-решений.

Использование программных инструментов анализа текстов на русском языке для решения производственных и научных задач

Результаты сравнения качества работы программных инструментов анализа текстов на русском языке демонстрируют довольно высокий уровень современных инструментов, например, для морфологического анализа на текстах Национального корпуса русского языка и корпуса SynTagRus LLM-решения показывают точность 94-97% [11], алгоритмические решения – 95-99% [12].

Несмотря на повсеместное использование LLM-решений тенденции меняются: по данным Gartner в 2023 году [13] шел рост использования AI-технологий, тогда как по данным 2025 года [14] пошло снижение и начался рост в области развития специализированных решений и общемировых моделей. В таблице приведены результаты сравнения решений на основе алгоритмического подхода и LLM-подхода:

Таблица 2

Результаты сравнения решений

Алгоритмический подход	LLM-подход
Основа	
ингвистические знания экспертов	Большие массивы текстовых данных
Интерпретируемость	
Высокая (прозрачные правила)	Очень низкая ("черный ящик")
Масштабируемость	
Низкая (ручное добавление новых правил и алгоритмов)	Высокая при наличии данных (добавление данных), низкая при отсутствии нужных данных
Гибкость	
Низкая: сложности при наличии неоднозначности	Средняя: работает с неоднозначностью, но результат может быть непредсказуемым
Зависимость от предметной области и задачи	
Средняя: при использовании словарей требуется их асширение, решение прикладных задач строится на базе основных алгоритмов	Высокая: требуются размеченные корпусы екстов для предметной области и решаемой задачи

Затраты на разработку	
Средние: работа лингвистов и программистов	Высокие: вычислительные ресурсы, работа специалистов по подготовке данных и обучению моделей
Безопасность и эксплуатация	
Аналогично любому ПО	Отдельные уязвимости LLM-решений, дополнительно к общим для любого ПО отдельные задачи MLOps и LLMOps
Требования к ресурсам и стоимость использования	
Средние и низкие требования к ресурсам (по сравнению с промышленным ПО)	Высокие требования для локального развертывания, минимум бесплатных облачных решений с ограничениями по производительности (неприменимы для производственных задач)

Наиболее критичными для практического применения программных средств автоматического анализа текста являются интерпретируемость и достоверность результатов, применение для новых предметных областей и задач в них, т.к. при решении производственных задач всегда существуют технические, экономические и репутационные риски, а решаемые задачи отличаются своей уникальностью.

Несмотря на привлекательность скорости и простоту получения первичного решения задачи с использованием LLM, получение достоверного результата может потребовать больше времени или даже привести к невозможности его получения, т.к. модели часто выдают сгенерированный на основе обобщенных данных результат.

Таким образом, инструменты на основе LLM позволяют решать более широкий круг задач в качестве базовых, на основе которых могут строиться более сложные уникальные решения в конкретной предметной области или компании. Однако, с точки зрения интерпретируемости и достоверности результатов возникет большой риск получения недостоверных результатов, что может приводить к критичным последствиям. При этом немаловажным фактором является и стоимость внедрения и эксплуатации решений на основе LLM: стоимость оборудования или использования облачных решений (при наличии возможности это делать с точки зрения обеспечения конфиденциальности используемой информации), а также сроки и объем работ по обучению моделей на подходящих данных, адаптации готовых моделей под нужные задачи и процессы, обеспечение безопасности внедряемого решения и его встраивание в существующие производственные процессы требует существенных финансовых и временных затрат.

Научные задачи также отличаются своей уникальностью, причем часто не только с точки зрения предметной области, но и с точки зрения решаемой задачи, исследуемого явления. Для этого практически всегда требуется разработка новых алгоритмов и новых наборов данных.

Таким образом, производственные и научные задачи объединяет их уникальность, их решение невозможно путем сведения к полностью ранее решенной задач для той же предметной области и использования готового инструмента, а значит всегда требуется разработка новых комплекстных алгоритмов, комбинрование существующих подходов и инструментов. Так, алгоритмические методы могут дополнять решения на основе LLM для проверки достоверности ответов путем поиска в открытых источниках и дополнительного анализа и дополнительной валидации и фильтрации входных данных и итоговых результатов. Инструменты на базе LLM могут использоваться для решения научных задач при предварительно проведенной специфичной ориентированной на задачу векторизации текстов и подготовки модели для этой задач.

Заключение

Инструменты анализа текстов на русском языке достигли промышленной зрелости, что позволяет их активно применять как в научных исследованиях, так и для решения производственных задач.

Основными направлениями развития инструментов на основе LLM остаются создание специализированных моделей для предметных областей для более точного решения узких задач и промышленного применения в разных сферах, инструментов на основе алгоритмического подхода – уточнение правил анализа текста и создание новых алгоритмов для автоматического выявления новых языковых особенностей [15].

Оптимизация производительности существующих решений актуальна для инструментов обоих направлений и их совместного использования: усложнение алгоритмов анализа приводит к увеличению расхода ресурсов и увеличению времени обработки, использование LLM предполагает высокие требования к инфраструктуре и ее дальнейшей эксплуатации.

Таким образом, ключевыми проблемами практического развития программных инструментов автоматического анализа текста становятся высокие технические и финансовые требования к переходу от их применения для исследований и пилотных проектов к массовому промышленному использованию для решения производственных задач, а также оптимальный выбор и комбинированное использование существующих инструментов для решения уникальных производственных и научных задач.

Список литературы

1. Полицына, Е. В. Алгоритмы автоматизации анализа текста на русском языке для решения прикладных задач с применением фреймворка TAWT / Е. В. Полицына, С. А. Полицын, А. С. Поречный // Программные продукты и системы. – 2021. – № 2. – С. 257-268.
2. Jurafsky, D., Martin, J. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, with Language Models. (2026), https://web.stanford.edu/~jurafsky/slp3/, Online manuscript released January 6, 2026
3. Bird, S. Natural Language Processing with Python / S. Bird, E. Klein, E. Loper. – O'Reilly Media, Inc., 2009. – 504 с. – ISBN 978-0-5965-164-99
4. Manning, C. Foundations of Statistical Natural Language Processing / Christoph Manning, Hinrich Schütze. – Cambridge: The MIT Press, 1999. – 718 с. – ISBN 978-0-26213-360-9
5. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models / Wei и др. / Advances in Neural Information Processing Systems – 2022.
6. ReAct: Synergizing Reasoning and Acting in Language Models / Shunyu Yao et al. / Proceedings at International Conference on Learning Representations – 2023.
7. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks / Lewis P. et al. /arXiv.2005.11401, – 2020.
8. Holistic Evaluation of Language Models / Bommasani, R., Liang, P., Lee, T. / Annals Of The New York Academy Of Sciences, – 2023.- №5
9. OWASP TOP 10 for LLMs 2023 [Электронный ресурс] – Режим доступа: https://owasp.org/www-project-top-10-for-large-language-model-applications/
10. Внимание, программирует ИИ! [Электронный ресурс] – Режим доступа: https://www.kaspersky.ru/blog/vibe-coding-2025-risks/40659
11. Russian SuperGLUE benchmark [Электронный ресурс] – Режим доступа: https://russiansuperglue.com/
12. Полицына, Е.В. Анализ качества работы и расширение возможностей инструментов морфологического анализа текстов на русском языке / Е.В. Полицына, С. А. Полицын, А.С. Поречный, А.Н. Рыкунов // Вестник ВГУ. Серия: Системный анализ и информационные технологии. – 2023. – № 2. – С. 171-180
13. What’s New in AI from the 2023 Gartner Hype Cycle [Электронный ресурс] – Режим доступа:

https://www.gartner.com/en/articles/what-s-new-in-artificial-intelligence-from-the-2023-gartner-hype-cycle

14. The 2025 Hype Cycle for Artificial Intelligence Goes Beyond GenAI [Электронный ресурс] – Режим доступа: https://www.gartner.com/en/articles/hype-cycle-for-artificial-intelligence

15. Развитие инструментов фреймворка TAWT для работы с текстами на русском языке / Е. В. Полицына, С. А. Полицын, А. С. Поречный, А. Н. Рыкунов // Информатика: проблемы, методы, технологии : Материалы XXV Международной научно-практической конференции им. Э.К. Алгазинова, Воронеж – Воронеж: ВГУ, 2025. – С. 964-973.

Использование статических анализаторов в генерации программного кода на основе больших языковых моделей

Е. С. Скопин, email: eternalpulsar90@gmail.com

ФГБОУ ВО «Вятский государственный университет»

Аннотация. *В данной работе рассматривается метод улучшения генерации программного кода большими языковыми моделями основанный на выборке с отклонением на базе статического анализатора кода в качестве внешнего верификатора. Подход не требует дополнительного обучения модели и позволяет повысить долю корректных решений для моделей малого размера.*

Ключевые слова: *машинное обучение, нейронные сети, большие языковые модели, генерация кода, статические анализаторы.*

Введение

Современный этап развития информационных технологий характеризуется стремительным проникновением методов искусственного интеллекта (ИИ) во все сферы деятельности, и программная инженерия не является исключением. Одним из наиболее прорывных направлений на стыке ИИ и разработки программного обеспечения стала генерация кода с использованием больших языковых моделей [1]. LLM активно применяются для автоматической генерации программного кода, однако, как правило, повышение корректности и качества сгенерированных решений достигается за счёт увеличения размера моделей или применения методов дополнительного обучения. Эти подходы требуют значительных вычислительных ресурсов. В данной работе рассматривается простой метод улучшения на этапе генерации с использованием моделей с малым количеством параметров (< 8млрд.), не требующий большого количества ресурсов.

1. Данные

В качестве тестового набора данных в данной работе использовался набор MBPP (Mostly Basic Programming Problems) [2]. Это коллекция задач по программированию базового уровня. Датасет включает в себя 974 задачи, коллекция разделена на несколько отдельных наборов:

- тестовый (500 задач);
- обучающий (374 задач);
- валидационный (90 задач);

© Скопин Е.С., 2026

– few-shot набор – это 10 задач предназначеных для использования в few-shot инструкциях;

Каждая задача включает текстовое описание, готовый код и набор тестов для проверки корректности решения. Данный набор данных широко применяется для оценки качества генерации программного кода большими языковыми моделями и позволяет проводить воспроизводимые сравнения различных методов генерации. В рамках данной работы тесты из набора MBPP использовались для проверки функциональной корректности сгенерированных решений.

2. Метод выборки с отклонением

Для каждого входного запроса модель выполняет несколько независимых стохастических генераций, после чего полученные варианты последовательно проверяются на корректность. Генерация выполняется с использованием стохастического сэмплирования. В каждом шаге используется генерация с Top-p сэмплированием (ядерная выборка) [3], а также с сэмплированием с температурой. Для каждого запроса генерируется не более N вариантов.

Сгенерированный моделью ответ может содержать вспомогательный текст или форматирование в виде markdown-блоков. Перед проверкой корректности из ответа извлекается программный код, приоритетно из блоков, помеченных как «```python». В случае отсутствия таких блоков используется весь сгенерированный ответ. Данная процедура позволяет унифицировать входные данные для последующей проверки. В качестве внешнего верификатора используется статический анализатор кода ruff. Для каждого сгенерированного варианта код сохраняется во временный файл и проверяется с использованием стандартного режима статического анализа. Вариант считается прошедшим проверку только в случае отсутствия диагностических сообщений и ненулевого кода возврата анализатора.

Процесс генерации продолжается до получения первого варианта, успешно прошедшего статическую проверку, либо до достижения максимального числа попыток. В случае если ни один из сгенерированных вариантов не удовлетворяет критерию, в качестве результата используется последний сгенерированный код.

Оценка функциональной корректности сгенерированных решений выполняется отдельно с использованием тестов из набора данных MBPP. Для каждого задания применяется стандартная процедура выполнения тестов, предоставленных в датасете.

3. Эксперименты

Эксперименты проводились с использованием библиотеки Hugging Face Transformers [4] для работы с большими языковыми моделями и выполнения генерации. Для загрузки и предобработки набора данных MBPP использовалась библиотека datasets [5]. Функциональная корректность сгенерированных решений оценивалась с применением стандартной процедуры выполнения тестов, реализованной в библиотеке human-eval [6]. В качестве статического анализатора кода использовался инструмент ruff [7]. В работе использовались авторегрессионные языковые модели семейства Qwen3 [8]:

- Qwen/Qwen3-0.6B;
- Qwen/Qwen3-1.7B;
- Qwen/Qwen3-4B;
- Qwen/Qwen3-8B-FP8;

Все модели использовались в режиме инференса без дополнительного обучения или адаптации параметров.

В качестве основной метрики использовалась доля функционально корректных решений, успешно прошедших тесты из набора данных. Для каждой модели сравнивались результаты базовой генерации и генерации с применением выборки с отклонением. По каждому методу было проведено 20 экспериментов. В таблице представлены средние значения доли функционально корректных решений, а также минимальные и максимальные значения.

Таблица. Результаты генерации кода на наборе MBPP.

Модель	**Параметры (млрд.)**	**Базовая генерация**	**Выборка с отклонением**
Qwen3-0.6B	0.6	27.3 (26.0-28.4)	**27.9** (25.2-29.8)
Qwen3-1.7B	1.7	43.1 (41.0-45.6)	**43.9** (42.0-45.4)
Qwen3-4B	4.0	56.4 (54.7-57.9)	**56.8** (55.5 – 58.3)
Qwen3-8B	8.2	**61.3** (59.9-62.9)	60.8 (60.1-61.9)

В проведённых экспериментах использовались следующие параметры генерации:

- Температура = 0.6.
- top-p = 0.95.
- Максимальная длина последовательности генерации = 1024 токена.

– Максимальное количество попыток генерации = 5.

Для моделей меньшего размера наблюдается стабильное улучшение качества генерации. В частности, для моделей порядка 0.6B–4B параметров применение статического анализатора позволяет увеличить долю функционально корректных решений по сравнению с базовой генерацией. Однако для более крупных моделей значимый эффект отсутствует.

Заключение

В работе был рассмотрен метод для улучшения генерации программного кода на основе больших языковых моделей при помощи статического анализатора.

Предоставлены результаты экспериментов в сравнении с традиционными подходами генерации на основе языковых моделей. Результаты показали, что простой подход выборки с отклонением на основе статического анализатора позволяют получить более корректную генерацию кода на языковых моделях с небольшим количеством параметров, порядка 0.6-4 млрд.. Метод не показал значимого улучшения с более крупными моделями. Это может объясняться изначально высокой способностью более крупных моделей к генерации синтаксически корректного кода, а также возможной избыточностью внешней проверки для моделей такого масштаба.

Таким образом, предложенный подход демонстрирует потенциал использования статических анализаторов как простого и эффективного инструмента повышения надёжности генерации кода малыми языковыми моделями, что может способствовать их более широкому внедрению в практику разработки программного обеспечения.

Перспективы дальнейших исследований могут включать:

1. Интеграцию семантических верификаторов и динамического тестирования;
2. Адаптацию метода для других языков программирования.

Список литературы

1. A Survey on Large Language Models for Code Generation [Текст] / J. Jiang, F. Wang, J. Shen [и др.] – 2024.
2. Program synthesis with large language models [Текст] / A. Odena, C. Sutton, D.M. Dohan, E. Jiang [и др.] – 2021.
3. The Curious Case of Neural Text Degeneration [Текст] / A. Holtzman, J. Buys, L. Du, [и др.] – 2020
4. Transformers: State-of-the-Art Natural Language Processing. [Электронный ресурс] – Режим доступа: https://github.com/huggingface/transformers (дата обращения 24.01.2026).

5. Datasets: A Community Library for Natural Language Processing. Datasets [Электронный ресурс] – Режим доступа: https://github.com/huggingface/datasets (дата обращения 24.01.2026).
6. HumanEval: Hand-Written Evaluation Set [Электронный ресурс] – Режим доступа: https://github.com/openai/human-eval (дата обращения 24.01.2026).
7. Ruff: An extremely fast Python linter and code formatter, written in Rust. [Электронный ресурс] – Режим доступа: https://github.com/astral-sh/ruff (дата обращения 24.01.2026).
8. Qwen3 Technical Report / A. Yang, A. Li, B. Yang, [и др.] arXiv:2505.09388 [cs]. – arXiv, 2025. – URL: http://arxiv.org/abs/2505.09388 (дата обращения: 24.01.2026).

Возможности применения технологии Retrieval Augmented Generation в системах поддержки принятия решений при разработке информационных систем

К. В. Смотрова, e-mail: smotrova-kristina@mail.ru
В. В. Гаршина, e-mail: garshina@cs.vsu.ru

Воронежский государственный университет

Аннотация. *В работе рассматриваются возможности применения Retrieval Augmented Generation (RAG) в системах поддержки принятия решений (СППР), используемых в процессе разработки информационных систем. Предложена практико-ориентированная архитектура RAG-СППР, которая объединяет корпоративную базу знаний (требования, архитектурные решения, регламенты, постмортемы инцидентов) и генеративную модель, формируя обоснованные рекомендации со ссылками на источники. Описаны ключевые этапы конвейера (подготовка данных, индексация, гибридный поиск, ранжирование, контроль достоверности), а также подход к оценке качества ответов и риски внедрения. Результаты могут быть использованы для построения помощников архитектора/аналитика, ускоряющих согласование решений и повышающих прозрачность разработки.*

Ключевые слова: *Retrieval Augmented Generation, RAG, система поддержки принятия решений, СППР, разработка информационных систем, база знаний, корпоративные данные, LLM.*

Введение

Разработка информационных систем (ИС) в организациях сопровождается постоянным принятием решений: выбор архитектурного стиля и технологий, приоритизация требований, оценка рисков изменений, реагирование на инциденты, согласование нефункциональных требований (безопасность, производительность, надежность). Источники знаний для таких решений обычно распределены между трекером задач, репозиториями, регламентами, журналами инцидентов и перепиской команд. В результате возрастает время поиска релевантной информации, снижается повторное использование накопленного опыта, а решения фиксируются неполно либо без явных обоснований.

© Смотрова К.В., Гаршина В. В., 2026

Технология Retrieval Augmented Generation (RAG) объединяет информационный поиск по внешнему хранилищу знаний и генерацию ответа языковой моделью. Это позволяет формировать рекомендации на основе актуальных документов и одновременно предоставлять источники, на которые опирается вывод. Подход широко обсуждается как способ повысить фактичность и обновляемость генерации по сравнению с использованием только «памяти» модели.

1. RAG как компонент СППР в разработке ИС

Классические СППР ориентированы на поддержку человека-эксперта за счет предоставления данных, моделей и интерфейса анализа альтернатив. Для разработки ИС характерно, что решение часто требует сопоставления нескольких классов знаний:

– нормативные требования и политики (например, безопасность и комплаенс);

– внутренние архитектурные решения (ADR);

– опыт эксплуатации (инциденты, SRE-отчеты, runbook);

– контекст текущей задачи (требования, ограничения сроков и ресурсов).

RAG-СППР можно рассматривать как надстройку над корпоративной базой знаний, которая:

2. извлекает релевантные фрагменты документов;

– синтезирует ответ в удобном виде для принятия решения;

– прикладывает ссылки на первоисточники и фиксирует «след решения» для последующего аудита.

Важная особенность – обновляемость: при изменении регламентов или накоплении новых ADR нет необходимости переобучать модель – достаточно обновить индекс знаний. Этот принцип напрямую следует из идеи сочетания параметрической и непараметрической памяти в RAG-подходе. [1]

2. Практико-ориентированная архитектура RAG-СППР

Эффективная RAG-система на практике реализуется как многоступенчатый конвейер, начинающийся со сбора и нормализации данных. В качестве источников выступают разнородные корпоративные артефакты: документация, требования и задачи из трекеров, содержимое репозиториев (включая README, архитектурные решения и changelog), эксплуатационные материалы вроде постмортемов и ранбуков, а также шаблоны решений и проверочные чек-листы. Все документы приводятся к единому формату и обогащаются метаданными – указанием системы или модуля, ответственной команды, даты создания, типа документа и уровня доступа.

Далее следует этап сегментации, при котором текст разбивается на фрагменты умеренного размера (обычно 300–800 токенов) с небольшим перекрытием соседних частей. [2] При этом сохраняется исходная иерархия документа: заголовки, структура разделов и ссылки на первоисточник. Особенно важно для инженерных документов уважать их внутреннюю структуру – не разрывать таблицы, списки и кодовые блоки, а также избегать смешения различных по смыслу частей в одном фрагменте.

Каждый полученный фрагмент проходит векторизацию: для него вычисляется эмбеддинг и сохраняется в векторном хранилище. В корпоративной среде практичным решением оказывается использование PostgreSQL с расширением pgvector, что позволяет хранить векторные представления и метаданные в единой базе и выполнять фильтрацию непосредственно на уровне СУБД. [4]

Извлечение релевантного контекста строится на гибридном подходе, сочетающем классический ключевой поиск по алгоритму BM25 с векторным поиском по семантической близости. Дополнительную точность обеспечивают фильтры по метаданным – продукту, версии, команде или типу документа. После первичного отбора кандидатов применяется этап реранжирования, в ходе которого отбираются 5–12 наиболее полезных фрагментов для передачи в языковую модель. Такая многоуровневая стратегия извлечения снижает уровень шума и помогает уложиться в ограничения контекстного окна генеративной модели. [2]

На этапе генерации модель получает вопрос пользователя вместе с отфильтрованным контекстом и формирует ответ в структуре, принятой для систем поддержки принятия решений. Ответ включает несколько альтернативных вариантов решения, аргументы «за» и «против» по каждой опции, выявленные риски и ограничения, конкретные рекомендации по следующим шагам (что проверить или с кем согласовать), а также точные цитаты и ссылки на источники – пункты регламентов, архитектурные решения или описания инцидентов. Для интеграции в чат-боты или плагины к системам управления задачами удобно возвращать ответ в структурированном JSON-формате с полями для вариантов, преимуществ, недостатков, рисков и источников.

Завершающим элементом конвейера выступает контроль достоверности: на этапе постобработки система проверяет, что ключевые утверждения ответа подкреплены предоставленными фрагментами контекста. При этом выделяются прямые цитаты, формируется список источников, а в случаях недостатка информации система либо запрашивает у пользователя уточнение (например, версию продукта или дополнительный контекст), либо выдаёт ограниченный ответ с оговоркой

о неполноте данных. Подобные техники верификации знаний и оценки надёжности RAG-ответов активно развиваются как самостоятельное направление в области прикладного применения генеративного ИИ. [3]

3. Сценарий применения: помощник для архитектурных и проектных решений Оценка эффективности RAG-СППР

Рассмотрим типовой сценарий применения RAG-системы поддержки принятия решений: выбор архитектурного подхода при разработке нового функционала. Эксперт задаёт вопрос о целесообразности выделения отдельного микросервиса для реализации функции X или её размещения внутри существующего монолитного приложения. Для формирования взвешенного ответа система анализирует комплекс релевантных корпоративных артефактов:

– архитектурные решения (ADR) по аналогичным кейсам;

– нефункциональные требования: ожидаемая нагрузка, целевые SLA, параметры восстановления (RTO/RPO);

– ограничения инфраструктурной платформы (лимиты на количество сервисов, требования к сетевой топологии);

– данные прошлых инцидентов – включая точки отказа, корневые причины и применённые меры;

– внутренние стандарты команды: критерии выделения границ сервисов, требования к логированию и метрикам.

На основе этого анализа пользователю предлагаются два обоснованных варианта. Первый – выделение самостоятельного микросервиса – обеспечивает независимое масштабирование компонента, изоляцию отказов и автономный цикл релизов. При этом система честно указывает на сопутствующие риски: рост операционной нагрузки на команду, сложности с распределёнными транзакциями и необходимость дополнительных усилий для сквозной наблюдаемости. Второй вариант – реализация функционала внутри монолита – ускоряет поставку, упрощает отладку и трассировку запросов, а также снижает число инфраструктурных зависимостей. Однако система предупреждает о потенциальных проблемах: увеличении внутренней связанности кодовой базы, трудностях с избирательным масштабированием отдельных компонентов и формировании «бутылочных горлышек» в процессе релизов.

Ключевое преимущество подхода – полная прослеживаемость аргументации: каждый довод сопровождается прямыми ссылками на первоисточники (конкретные ADR, отчёты об инцидентах, внутренние гайды). [2] Такая прозрачность не только ускоряет согласование решения в архитектурном комитете или на ревью, но и служит надёжной

профилактикой повторения ошибок, уже зафиксированных в истории эксплуатации системы.

4. Оценка эффективности RAG-СППР

Для практического внедрения RAG-системы поддержки принятия решений критически важно оценивать не только стилистическое качество сгенерированного текста, но и его реальную пользу в процессе принятия обоснованных решений. Эффективный подход предполагает применение сбалансированного набора метрик, охватывающего несколько измерений качества.

Качество извлечения релевантного контекста измеряется классическими информационно-поисковой метрикой – Recall@K – на контрольном наборе вопросов, что позволяет объективно оценить, насколько полно и точно система находит нужные документы и фрагменты. Параллельно оценивается достоверность (faithfulness) ответов: рассчитывается доля утверждений, прямо подтверждённых источниками, а также выявляется процент «битых» или нерелевантных ссылок, которые могут ввести пользователя в заблуждение. [6]

Важным компонентом оценки остаётся экспертная пользовательская обратная связь: специалисты выставляют баллы по таким критериям, как практическая полезность рекомендаций, полнота анализа альтернатив и прозрачность аргументации с привязкой к источникам. Дополняет картину мониторинг операционных показателей – задержки ответа, стоимости обработки запроса и стабильности работы системы после обновлений базы знаний.

Анализ практики внедрения и обзоров современных RAG-решений показывает, что значимый вклад в итоговое качество вносит не только выбор генеративной модели, но и тщательная подготовка данных: продуманная структура чанков с учётом семантических границ, насыщенность метаданными для точной фильтрации, а также корректно выстроенная многоступенчатая стратегия ранжирования контекста перед передачей в модель. [2] Именно эти инженерные аспекты часто определяют разницу между теоретически возможным и практически применимым результатом.

Заключение

RAG является перспективной технологией для повышения практической ценности СППР в разработке информационных систем, так как объединяет доступ к корпоративной базе знаний и генерацию рекомендаций, обеспечивая обновляемость и прослеживаемость выводов. Предложенная архитектура RAG-СППР ориентирована на реальные артефакты разработки ИС (требования, ADR, регламенты,

эксплуатационные отчеты) и позволяет формировать структурированные ответы, поддерживающие выбор альтернатив.

К направлениям дальнейшей работы относятся: интеграция с графами знаний для учета связей между сущностями (компоненты–инциденты–решения), автоматическая оценка достоверности и персонализация рекомендаций в зависимости от роли пользователя (аналитик, архитектор, SRE). Подходы «RAG + knowledge graph» также рассматриваются в исследованиях СППР как способ усилить связность знаний и объяснимость. [5]

Литература

1. Lewis, P., Perez E., Piktus A. et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks [Электронный ресурс]. – 2020. – URL: https://arxiv.org/abs/2005.11401. – Дата обращения: 10.01.2026.
2. Gao, Y., et al. Retrieval-Augmented Generation for Large Language Models: A Survey [Электронный ресурс]. – 2023. – URL: https://arxiv.org/abs/2312.10997. – Дата обращения: 12.01.2026.
3. Zeng, S., et al. Towards Knowledge Checking in Retrieval-augmented Generation: A Representation Perspective [Электронный ресурс]. – 2025. – URL: https://aclanthology.org/2025.naacl-long.151/. – Дата обращения: 11.01.2026.
4. pgvector. pgvector: Open-source vector similarity search for Postgres [Электронный ресурс]. – URL: https://github.com/pgvector/pgvector. – Дата обращения: 10.01.2026.
5. Wang, S,. et al. Construction of intelligent decision support systems through integration of retrieval-augmented generation and knowledge graphs [Электронный ресурс]. – 2025. – URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC12514041/. – Дата обращения: 11.01.2026.
6. Gan, A., Yu, H., Zhang, K. et al. Retrieval Augmented Generation Evaluation in the Era of Large Language Models: A Comprehensive Survey [Электронный ресурс]. – 2025. – URL: https://arxiv.org/abs/2504.14891. – Дата обращения: 12.01.2026.

Использование алгоритмов обучения с подкреплением для мультиагентных систем в задачах управления складом

Я. В. Солодовникова, email: solodovnikovasyv@yandex.ru
В. В. Гаршина, email: garshina.veronika@gmail.com

Воронежский государственный университет

Аннотация. *В данной статье рассматривается практическое применение методов обучения с подкреплением для мультиагентных систем в задаче управления складом. Проведена сравнительная оценка эффективности алгоритмов IDQN, QMix и MAPPO, определены их сильные стороны и ограничения в условиях динамичной среды. Полученные результаты показывают преимущество специализированных алгоритмов MARL и подчеркивают важность выбора подхода в зависимости от требований конкретной системы.*

Ключевые слова: *методы обучения агентов, мультиагентные системы, координация агентов, адаптивные системы, взаимодействие агентов, управление складом.*

Введение

С развитием технологий в современном мире стала актуальной задача управления автоматизированными складами [1]. Это действительно сложный вызов, так как требуется координировать большое число роботов-штабелеров (AVG) в условиях ограниченных ресурсов и динамического изменения окружающей среды. При этом классические системы, основанные на предварительном планировании, не подходят для оперативного реагирования на изменения. Поэтому для такого рода задач представляется перспективным применение методов мультиагентного обучения с подкреплением (MARL). Они позволяют создавать адаптивные системы, в которых агенты самостоятельно обучаются координированным стратегиям.

В предыдущей статье [2] был проведен обзор различных алгоритмов MARL, данное же исследование направлено на практическую реализацию и сравнительный анализ нескольких перспективных алгоритмов: Independent DQN, QMix и MAPPO. Способность алгоритмов оптимизировать ключевые показатели проверяется в задаче управления складом. В данном случае важны такие метрики, как пропускная способность, время выполнения заказа, а также минимизация конфликтов между агентами. Это позволяет определить сильные и слабые стороны

© Солодовникова Я.В., Гаршина В.В., 2026

каждого подхода для практического применения в системах складской логистики.

1. Средства реализации

Средства реализации включают в себя программную платформу и аппаратные ресурсы. Далее рассмотрим каждую из этих составляющих более подробно.

Программная платформа предстпаляет собой язык программирования, библиотеки и среду разработки. В качестве языка программирования был выбран Python 3.9, так как он является стандартным для задач машинного обучения благодаря своей обширной экосистеме библиотек. Для реализации алгоритмов обучения использовалась библиотека PyTorch, обработка данных и вычисления выполнялись с использованием NumPy и Pandas. Для визуализации результатов обучения применялись библиотеки Matplotlib и Seaborn. Весь процесс разработки проходил в среде PyCharm, она была выбрана за свои удобные инструменты отладки, которые позволяют проводить работу с большими проектами. Все это важно при реализации комплексных алгоритмов MARL [3].

Теперь перейдем к аппаратным ресурсам. Эксперименты проводились на выделенной рабочей станции со следующей конфигурацией:

- Процессор: Intel Core i5-8600K (6 ядер, 6 потоков);
- Оперативная память: 32 GB DDR4;
- Графический ускоритель: NVIDIA GeForce GTX 1060 (6 GB VRAM);
- Операционная система: Windows 10 22H2;
- Хранение данных: Samsung SSD 860 EVO M.2 250 GB.

Данная конфигурация обеспечивает необходимую производительность для параллельного обучения нескольких агентов и обработки данных в реальном времени.

2. Архитектура решений и моделирование среды

В качестве задачи, которую необходимо решить, выступает управление роботами-штабелерами на автоматизированном складе. Роботам необходимо находить среди стеллажей на складе необходимые товары и доставлять их на точку упаковки. При этом важно, чтобы задача решалась эффективно в условиях ограничения ресурсов, то есть роботы должны выбирать минимальные маршруты, избегать столкновений со стеллажами и стенами, а также не мешать работе других роботов.

Модель среды представляет собой дискретную сетку фиксированного размера, где в качестве агентов выступают автономные

роботы-штабелеры. Пространство действий каждого агента включает перемещения в четырех направлениях, действия загрузки или разгрузки товара и ожидания для разрешения конфликтов.

Чтобы реализовать данную среду, была разработана модульная симуляция (см. рис. 1). Модуль карты склада (Warehouse Grid Manager) хранит состояние сетки, в которой клетки могут быть свободными, содержать стеллаж или являться станцией приема и отгрузки. Диспетчер заказов (Order Dispatcher) генерирует асинхронный стохастический поток задач, в результате чего возникает динамическая нагрузка. В каждый момент времени система наблюдения (Observation Generator) генерирует агенту локальный вектор признаков, состоящий из информации о состоянии клеток в ограниченном радиусе, наличии груза и статусе ближайших целей.

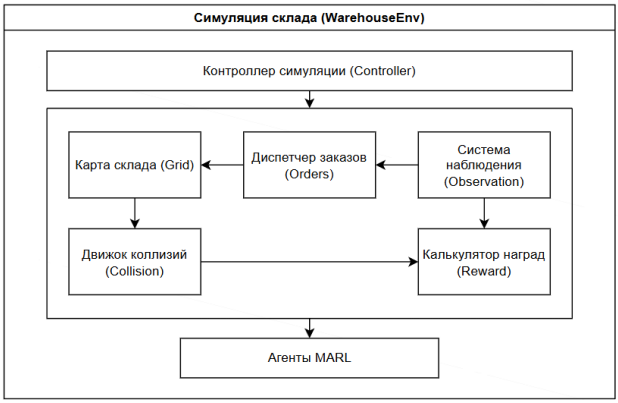


Рис. 1. Архитектура симуляции склада

Функция награды, которая рассчитывается калькулятором наград (Reward Calculator), сконструирована для поощрения эффективности. Система награждения работает следующим образом:

– при успешной доставке товара присваивается большая награда;
– за каждый шаг робот получает небольшой штраф (для стимулирования быстроты);
– при столкновениях назначается существенный штраф.

Движок коллизий (Collision Engine) отслеживает и обрабатывает коллизии между агентами. Все модули управляются контроллером симуляции (Simulation Controller), который управляет временными шагами и переходами между состояниями системы.

Теперь перейдем к алгоритмам обучения, как было сказано раньше, рассматриваются алгоритмы Independent DQN, QMix и MAPPO.

Начнем с алгоритма Deep Q-Network (DQN), он является одним из фундаментальных алгоритмов глубокого обучения с подкреплением. В мультиагентном контексте применяется подход Independent DQN, при котором каждый агент обучает свою собственную DQN-сеть, рассматривая других агентов как часть динамической и нестационарной среды. Это позволяет сохранять простоту и вычислительную эффективность, но приводит к появлению проблемы нестационарности [4]. Архитектура данного алгоритма представлена на рис. 2.

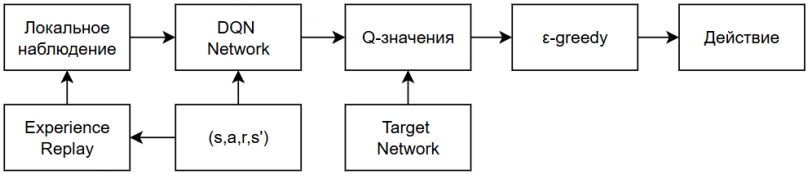


Рис. 2. Архитектура Independent DQN

Далее рассмотрим алгоритм QMix, разработанный специально для кооперативных задач. Он использует архитектуру централизованного обучения с децентрализованным исполнением (CTDE). Каждый агент имеет свою индивидуальную RNN-сеть, которая оценивает локальные Q-значения на основе истории его наблюдений. Затем эти значения агрегируются через смешивающую сеть (mixing network), параметры которой зависят от глобального состояния системы. Ключевое свойство mixing network – монотонность, гарантирующая, что увеличение локальных Q-значений ведет к увеличению общего Q-значения команды [5]. Архитектура данного алгоритма представлена на рис. 3.

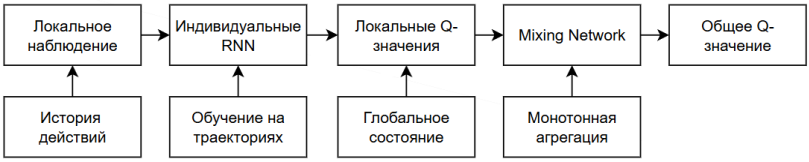


Рис. 3. Архитектура QMix

Завершим рассмотрение алгоритмом MAPPO (Multi-Agent Proximal Policy Optimization), который представляет собой актор-критический подход, адаптированный для мультиагентных систем. Он оптимизирует непосредственно политику агентов и так же использует архитектуру

CTDE. Децентрализованные акторы (actor networks) принимают решения на основе локальных наблюдений, а централизованный критик (critic network) оценивает общую ценность состояния, имея доступ к глобальной информации во время обучения. Для стабилизации используется техника PPO-clip, ограничивающая величину изменения политики на каждом шаге обновления [6]. Архитектура данного алгоритма представлена на рис. 4.

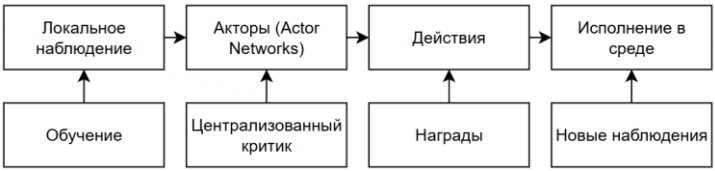


Рис. 4. Архитектура MAPPO

3. Результаты обучения

В результате симуляции и обучения агентов описанными выше алгоритмами были получены графики, отображающие эффективность решения задачи.

Кривые обучения (см. рис. 5) показывают, что QMix достигает стабильной производительности быстрее всех, тогда как MAPPO сходится медленнее, но стабильнее. Independent DQN демонстрирует высокую дисперсию и медленную сходимость.

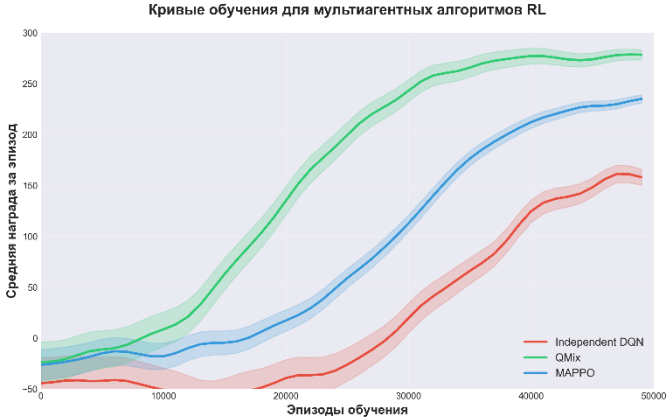


Рис. 5. Кривые обучения для мультиагентных алгоритмов

По финальной производительности (см. рис. 6) QMix лидирует по пропускной способности (27.5 заказов/час) и скорости выполнения заказов (8.2 мин). MAPPO показывает близкие результаты по производительности, но обеспечивает наименьшее количество столкновений (0.9 за эпизод). Independent DQN уступает по всем метрикам.

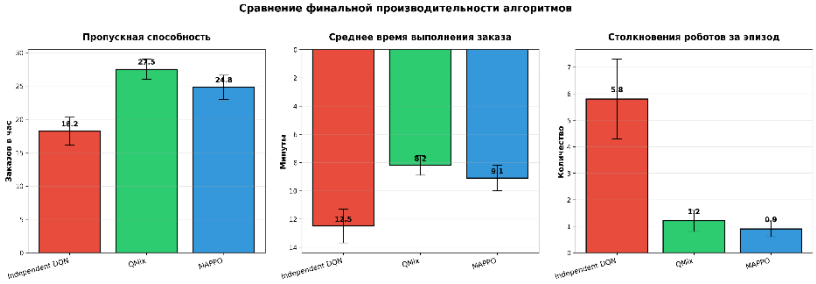


Рис. 6. Сравнений финальной производительности алгоритмов

Исследование масштабируемости (рис. 7) подтверждает, что MAPPO наиболее устойчив к увеличению числа агентов, теряя лишь 6% эффективности при переходе к 20 агентам. QMix теряет 14%, а Independent DQN – 60% производительности.

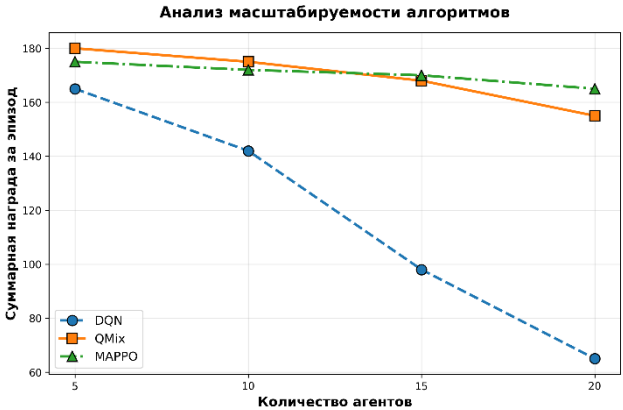


Рис. 7. Анализ масштабируемости алгоритмов

Заключение

В данной статье было проведено исследование по применению мультиагентного обучения с подкреплением для управления роботами-штабелерами на складе. С помощью алгоритмов Independent DQN, QMix и MAPPO получилось решить поставленную задачу и обучить агентов. Сравнительный анализ в свою очередь показал, что QMix демонстрирует наилучшую производительность при фиксированном числе агентов, а MAPPO – лучшую масштабируемость и безопасность. IDQN оказался недостаточно эффективен для задач, требующих координации. Результаты подтверждают практическую применимость MARL в складской логистике и определяют критерии выбора алгоритма для различных сценариев эксплуатации.

Список литературы

1. Kumar, N. V. Development of collision free path planning algorithm for warehouse mobile robot / N. V. Kumar, C. S. Kumar // Procedia Computer Science. – 2018. – Vol. 133. – P. 456–463.
2. Солодовникова, Я. В. Обзор методов и подходов для обучения агентов в мультиагентных системах / Я. В. Солодовникова, В. В. Гаршина // Информатика: проблемы, методология, технологии : материалы XXV Международной научой конференции. – Воронеж : Издательский дом ВГУ, 2025. – С. 995–1000.
3. Ziogas, A. N. Productivity, Portability, Performance: Data-Centric Python / A. N. Ziogas, T. Schneider, T. Ben-Nun, A. Calotoiu, T. De Matteis, J. de Fine Licht, L. Lavarini, T. Hoefler // Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. – New York : ACM, 2021. – P. 1–14.
4. Rashid, T. Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning / T. Rashid, M. Samvelyan, C. Schroeder de Witt, G. Farquhar, J. Foerster, S. Whiteson // Journal of Machine Learning Research. – 2020. – Vol. 21. – P. 1–51.
5. Colom, A. Empirical analysis of exploration strategies in QMIX : master thesis / A. Colom. – Barcelona : Universitat Pompeu Fabra, 2021. – P. 51.
6. Parada, L. Safe and efficient manoeuvring for emergency vehicles in autonomous traffic using multi-agent proximal policy optimisation / L. Parada, E. Candela, L. Marques, P. Angeloudis // Transportation Research Part C: Emerging Technologies. – 2022. – Vol. 135.

Сравнительная характеристика методов машинного обучения в задачах управления БПЛА в рамках эргатической системы

Н. Д. Солопанов, email: nik.solopanov@mail.ru [1]
М. В. Яковлева, email: mariaakovleva14507@gmail.com [1]
В. А. Санько, email: ronisan22@mail.ru[1]

[1] Воронежский государственный университет

Аннотация. *В статье представлены результаты работы, в которой изучалось поведение операторов в условиях ограниченной видимости, а также построение моделей прогнозирования потенциальных команд управления. При анализе результатов экспериментального исследования, включающего сбор данных и последующее обучение моделей машинного обучения, установлено, что наилучшие показатели точности классификации достигаются при использовании алгоритма CatBoost. Значение точности на тестовой выборке составило 0.95, что превышает соответствующие показатели моделей Случайного леса (0.896) и Деревьев решений (0.887). Наименьшее значение точности было получено при использовании модели градиентного бустинга (0.576). Предикторы моделей разбиты на 3 класса – параметры окружающей среды, психологические параметры и параметры истории команд. Модели из исследуемого пула обладающие высокой точностью (деревья решений и случайные леса) характеризуются умеренной зависимостью от психологических параметров – 0.11 и более выраженной зависимостью от параметров окружения – 0.61 и предыдущих команд – 0.28. Модели градиентного бустинга располагаются в области низкой точности – 0.576 и сравнительно равной важности всех трех групп признаков (параметры окружения – 0.39, психологические признаки – 0.26, история команд – 0.35). CatBoost выделяется как единственная модель, сочетающая высокую значимость предыдущих команд – 0.87 с высокой точностью, психологических параметров 0.06 и параметров окружения – 0.07.*

Ключевые слова: *Анализ данных, машинное обучение, классификация, психологические факторы.*

Введение

В связи с увеличением числа беспилотных летательных аппаратов (БПЛА) в гражданских и специализированных областях, появляется необходимость анализа поведения конкретного оператора и операторов в

© Солопанов Н. Д., Яковлева М. В., Санько В. А., 2026

целом. Увеличение сложности выполняемых задач и продолжительности полетов приводит к возрастанию когнитивной нагрузки на оператора, что делает человеческий фактор одним из ключевых ограничений эффективности управления эргатических систем [2, 5]. Несмотря на большое количество операторов высокого с точки зрения точности управления уровня, каждый из них нередко принимает неэффективные решения. Это может происходить из-за накопления усталости, снижения концентрации внимания, эффекта «замыливания» глаза, а также индивидуальными особенностями реагирования оператора на внешние стимулы. При длительном управлении подобные факторы могут существенно снижать точность, скорость и надежность управления. В связи с чем возникает задача разработки системы поддержки принятия решений, способных повышать устойчивость управления. Однако, существующие подходы, в основном, ориентированы на использование только информации об окружающей обстановке, не учитывая индивидуальное поведение и психологические паттерны оператора [4]. Игнорирование индивидуальных различий между операторами, таких как уровень тревожности, особенности внимания, скорость реакции и склонность к определенным стратегиям принятия решений, может существенно ограничивать адаптивность подобных систем и снижать их потенциальную эффективность в задачах управления беспилотными летательными средствами [1]. Таким образом, актуальной является задача анализа и разработки адаптивной модели, способной предсказывать лучшую команду управления с учетом индивидуальных особенностей оператора и условий выполнения задачи. Решение данной задачи позволит повысить качество управления БПЛА, снизить вероятность ошибок и создать основу для интеллектуальных систем поддержки принятия решений, схема такой системы предложена в работе [6]. Целью данной работы является разработка и сравнительный анализ моделей машинного обучения для прогнозирования команд управления БПЛА с учетом психологических характеристик оператора и параметров окружающей среды.

1. Материалы и методы

В формировании набора данных участвовало 33 человек возраста 20-23 лет, из них 18 мужчин и 15 женщин. Все испытуемые психологически и физически здоровы. Управление осуществлялось 2D моделью БПЛА в среде с препятствиями, сформированной на основе лидарной съемки реального леса. Для повышения вариативность применялась аугментация карт путем их поворота относительно центра карты. Каждый оператор прошел 50 раундов управления. Также каждый испытуемый был протестирован с использованием рядя психологических

тестов, в том числе: "Какое полушарие у вас преобладает?", "Шкала тревоги Спилбергера-Ханина, STAI", "Личностный опросник Айзенка, EPI", "Ваш стиль обучения и мышления, SOLAT" , "Мельбурнский опросник принятия решений, MDMQ", "Ведущий орган чувств" [3]. В процессе эксперимента осуществлялся сбор данных о траектории движения, временные интервалы между командами и параметры окружающей среды.

Предикторы моделей были разделены на несколько категорий, первая категория – параметры окружающей среды, включают в себя расстояния до ближайших деревьев, радиусы этих деревьев, среднюю площадь деревьев, процент расстояния до финиша, вторая категория – параметры истории принятия решений, такие как пять предыдущих команд, разница между текущей и предыдущей пользовательской командой, разница во времени между текущей и предыдущей командами. Следующая категория предикторов – пользовательские психологические характеристики, такие как Экстраверсия/интроверсия (из 24), Нейротизм (из 24), Зрительный канал (из 16), Сенсорный канал (из 16), Звуковой канал (из 16), Шкала ситуативной тревожности (из 80), Шкала личностной тревожности (из 80), ШСТ Рестанд 2011 (из 10), ШЛТ Рестанд 2011 (из 10), Темперамент. Категориальные признаки закодированы посредством One-Hot-Encoding. Целевой переменной выступала «лучшая» команда – дискретное направление движения минимизирующее значение углового отклонения от вектора на финиш при условии отсутствия препятствий в конечной точке шага. В результате была сформирована итоговая выборка объемом 1437303 наблюдений и 85 признаков.

2. Результаты и обсуждения

В рамках решения задачи построения модели прогнозирования лучшей команды управления на каждом шаге оператора были исследованы следующие алгоритмы машинного обучения: деревья решения (CART), случайный лес, градиентный бустинг, CatBoost. Из всего набора данных были сформированы 100 подвыборок, которые разбивались на тренировочную и тестовую выборки в следующих пропорциях 75% и 25% соответственно. Каждая модель обучалась на всех 100 подвыборках. Статистическая значимость различий между моделями проверялась с использованием критериев Краскела-Уоллиса и Уилкоксона. В таблице представлены описательные статистики точности за 100 подвыборок: среднее значение (M), стандартная ошибка среднего: (m), квартильные оценки (Q1, Me, Q3), коэффициенты асимметрии (As) и эксцесса (Ex), а также 5-й и 95-й процентили.

Таблица 1

Выборочные оценки метрик

	M ± m*	Q1	Me	Q3	As	Ex	5,00 %	95,00 %
Дерево решений, максимальная глубина 30	0.72 ± 0.0003	0.72	0.72	0.72	-0.20	-0.44	0.72	0.73
Дерево решений, максимальная глубина 50	0.72 ± 0.0004	0.72	0.72	0.73	-0.27	0.12	0.72	0.73
Случайный лес, количество деревьев 100	0.78 ± 0.0002	0.77	0.78	0.78	-0.05	0.17	0.77	0.78
Случайный лес, количество деревьев 200	0.78 ± 0.0002	0.78	0.78	0.78	-0.19	-0.19	0.77	0.78
Градиентный бустинг, количество деревьев 20	0.53 ± 0.0003	0.53	0.53	0.53	0.002	-0.19	0.53	0.54
Градиентный бустинг, количество деревьев 50	0.58 ± 0.0007	0.58	0.58	0.58	-2.26	5.9	0.56	0.59
Catboost	0.95± 0.0001	0.95	0.95	0.95	0.34	0.13	0.95	0.95

Примечание: * – $p < 0.001$ при проверке критерия Краскела – Уоллиса для сравнения показателей всех методов машинного обучения; aa – $p < 0.001$ при проверке лучшего результата со вторым лучшим по критерию Уилкоксона. Была сформулирована нулевая гипотеза об отсутствии статистически значимых отличий между результатами оценки моделей. С использованием парного критерия Уилкоксона получено следующее: лучшие результаты показала модель Catboost, точность на тестовой выборке – 0.95, в сравнении с деревьями решений результаты оказались статистически значимыми ($p<0.001$). Деревья и случайный лес показали близкие результаты, но статистически отличающиеся – 0.886-0.896 ($p<0.001$). Специфичность для всех моделей находится на высоком уровне. Чувствительность лучше всех у модели Случайного леса с количеством деревьев 200, эта модель лучше всех предсказывает класс команды. Модели градиентного бустинга продемонстрировали низкие показатели точности и чувствительность, но высокую специфичность. Между результатами всех моделей имеются статистические отличия, во всех $p<0.001$. По результатам сравнения моделей, лучшей была выбрана модель Catboost, обученная на 100 итерациях.

Важность признаков в используемых моделях машинного обучения представляет собой величину, отражающую относительный вклад каждого признака в процесс формирования решений модели. Значения важности рассчитываются на основе внутренних критериев конкретного алгоритма (уменьшение неоднородности для деревьев решений и случайного леса, а также изменение функции потерь для моделей градиентного бустинга и CatBoost) и нормируются таким образом, что сумма важностей по всем признакам равна единице.

На рисунке 6 показано распределение важности признаков в разбивке на 3 класса: параметры окружения, психологические параметры, предыдущие команды.

Параметры окружающей среды – «радиус ближайшего дерева», «средняя площадь деревьев» и «дистанция до ближайшего дерева» – заметно важны только для классических деревьев решений и частично для случайного леса. Их вклад находится в диапазоне 0.05–0.13. Это отражает склонность простых моделей опираться на реальные видимые параметры.

Для деревьев решений и случайного леса самым важным психологическим признаком был Темперамент – около 0.02, Catboost практически не учитывал этот признак. Для градиентного бустинга самым важным психологическим признаков оказался уровень слухового рецептора – 0.06. Параметры личной и ситуативной тревожность не принимались во внимание моделями, их вклад близок к нулю. Больше всего на психологические признаки опирались модели градиентного бустинга – 26% вклада в результат прогнозирования, за ними идут модели деревьев решений – 12% вклада, затем случайные леса – 10% вклада. Меньше всего психологические факторы учитывала модель Catboost – около 6%.

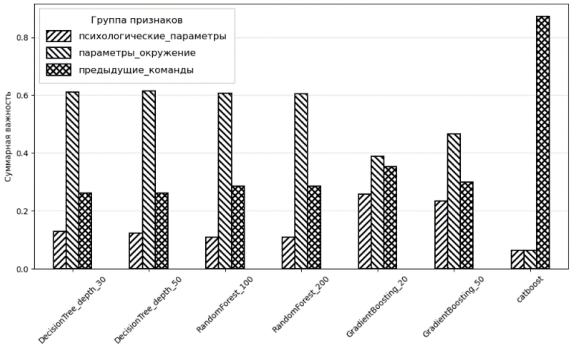


Рисунок. Важность признаков

Заключение

В результате проведенного исследования были собраны и проанализированы данные логов управления операторов, осуществлявших управление моделью беспилотного летательного аппарата в условиях полета в лесной среде, а также данные, характеризующие психологические особенности операторов. На основе полученного экспериментального материала была сформирована обучающая выборка и реализован ряд моделей машинного обучения, предназначенных для прогнозирования лучшей команды управления на каждом шаге оператора. Модель Catboost показала лучший результат на тестовой выборке – точность 0.96, модели деревьев решений и случайного леса показали точность около 0.88, градиентный бустинг показал наихудший результат – 0.57. Для моделей на основе деревьев решений и случайного леса важнейшими предикторами были параметры окружения; у Catboost – предыдущие команды; у градиентного бустинга важность параметров окружения, психологические признаки и предыдущие команды была на близком уровне. Полученные результаты подтверждают целесообразность учета индивидуальных особенностей оператора и истории его действий при построении моделей поддержки управления беспилотными летательными аппаратами. В дальнейшем развитии данной темы представляется перспективной разработка интеллектуальной информационной системы поддержки принятия решений при управлении беспилотными летательными аппаратами. Такая система может адаптировать рекомендации по управлению в зависимости от текущего состояния оператора, что потенциально может позволить повысить надежность управления, снизить вероятность ошибок и уменьшить когнитивную нагрузку на оператора в условиях длительных и сложных полетов.

Авторы выражают глубокую признательность Туровскому Ярославу Александровичу, ведущему научному сотруднику ИПУ РАН, доктору технических наук, кандидату медицинских наук, заведующему лабораторией медицинской кибернетики ВГУ, за научное руководство и поддержку при выполнении исследования.

Список литературы

1. Бирин, Д. А. Проблемы управления группой РТК / Д. А. Бирин, Р. В. Мещеряков, Я. А. Туровский // XVI Всероссийская мультиконференция по проблемам управления (МКПУ-2023): материалы мультиконференции. В 4 т. Т. 2. – Волгоград : Волгоградский государственный технический университет, 2023. – С. 287-290. – EDN PFMELT.

2. Мамонтов, Е. В. Человеческий фактор и безопасность полетов при управлении воздушным движением / Е. В. Мамонтов // Надежность и

качество сложных систем. – 2022. – № 1(37). – С. 41-45. – DOI 10.21685/2307-4205-2022-1-5. – EDN XJMBHT.

3. Навалихина, А. И. Тест на определение ведущей модальности с целью применения в сфере образования. / А.И. Навалихина // Вестник ПНИПУ. Проблемы языкознания и педагогики. – 2022. – № 3 – С. 197-201.

4. Подольский, В. А. Система классификации состояний оператора для повышения качества управления технической системой / В. А. Подольский, А. И. Михальский, Я. А. Туровский // Материалы XIII Международной конференции по прикладной математике и механике в аэрокосмической отрасли (AMMAI'2020) (Алушта, 06–13 сентября 2020 года). – Москва: Московский авиационный институт (национальный исследовательский университет), 2020. – С. 656-658. – EDN VYNLDJ.

5. Рыбалкина, А. Л. Управление рисками связанными с утомляемостью при техническом обслуживании воздушных судов как аспектом человеческого фактора / А. Л. Рыбалкина, А. С. Семенов // XXI век: итоги прошлого и проблемы настоящего плюс. – 2019. – Т. 8, № 1(45). – С. 159-163. – EDN YZAPYD.

6. Туровский, Я. А. Способ повышения надежности и противоаварийной безопасности эргатических систем / Я.А. Туровский, А.С. Суровцев, И.А. Авцинов // Управление развитием крупномасштабных систем (MLSD'2022) : Труды Пятнадцатой международной конференции (Москва, 26–28 сентября 2022 г.) / под общ. ред. С.Н. Васильева, А.Д. Цвиркуна. – Москва: Институт проблем управления им. В.А. Трапезникова РАН, 2022. – С. 975-979. – DOI 10.25728/mlsd.2022.0975. – EDN KFYBAO.

Автоматическая оценка релевантности научных статей тематикам конференции с помощью большой языковой модели

А.В. Сычев, email: sav@cs.vsu.ru

Воронежский государственный университет

Аннотация. *В докладе приведены результаты эксперимента по оценке релевантности текстов научных статей путем сопоставления их текстового содержимого с аннотациями секций научной конференции. Сопоставление выполнялось автоматически с помощью локально развернутой большой языковой моделью QWEN3-14B. Результатом сопоставления являлся вектор оценок релевантости и краткая текстовая аннотация этой оценки. Размерность вектора была равна количеству секций конференции. В докладе представлены также анализ результатов кластеризации сформированных векторов релевантности методом K-средних и t-SNE визуализация выявленных кластеров. Дополнительно был проведен анализ текстовых описаний существующих секций и сформированы рекомендации по их редизайну, полученные с помощью нескольких больших языковых моделей.*

Ключевые слова: *генеративный искусственный интеллект, большие языковые модели, оценка релевантности, кластеризация, t-SNE визуализация, RAG.*

Введение

Ежегодно организаторы научных конференции сталкиваются с ростом числа подаваемых работ и расширением междисциплинарных тематик. На этапе первичного отбора организаторам необходимо быстро и последовательно оценить, насколько статья соответствует заявленным тематическим направлениям конференции. На практике эта задача часто решается вручную – программным комитетом или назначенными рецензентами – что требует значительных трудозатрат и неизбежно приводит к вариативности решений из-за субъективности, неоднозначности критериев и ограничений по времени.

Автоматическую оценку релевантности научных статей тематическим секциям конференции можно рассматривать как частный случай задач информационного поиска и тематической классификации научных текстов. Традиционные подходы включают правила и ключевые

© Сычев А.В., 2026

слова, методы векторизации (TF-IDF, тематическое моделирование), классические модели машинного обучения и нейросетевые эмбеддинги. Однако у этих методов есть ограничения: они часто чувствительны к терминологическим вариациям, хуже обрабатывают междисциплинарные работы и слабо учитывают контекстные и аргументационные связи, важные для корректной интерпретации научного текста.

Появление больших языковых моделей (Large Language Models, LLM) создало предпосылки для более «глубокой» оценки соответствия: LLM способны учитывать контекст, выводить скрытую тему работы из аннотации или полного текста, сопоставлять ее с описаниями секций и формировать обоснование решения.

В то же время применение LLM в процедурах отбора статей порождает новые методические вопросы:

- как формулировать критерии и промпты для устойчивой работы модели;
- как измерять качество релевантности и калибровать уверенность; как минимизировать галлюцинации и обеспечить воспроизводимость; как учитывать справедливость (bias) и прозрачность;
- как интегрировать LLM в рабочий процесс с участием экспертов.

Цель данного доклада – рассмотреть подходы к автоматической оценке релевантности научных статей тематикам конференции с использованием большой языковой модел, сравнить их с традиционными решениями и апробировать на практике конкретный алгоритм оценки релевантности.

Обзор литературы

Рассмотрим существующие подходы к решению задачи оценки релевантности научных статей тематикам конференции.

1. Ранние и «классические» методы. Включают в себя тематическое моделирование и базовую классификацию. TF-IDF в сочетании с линейными классификаторами и SVM долгое время оставались базовой линией для классификации статей по предметным областям. Такие подходы просты и интерпретируемы, но плохо работают на коротких аннотациях, слабее учитывают контекст и синонимию. Тематическое моделирование LDA применялось для выделения тем и группировки документов. Оно полезно для разведочного анализа, но не всегда дает точное сопоставление с заранее заданными треками конференции и чувствительно к настройкам и корпусу текстов [1], [2].

2. Нейросетевые представления текста включают в себя ряд моделей от word2vec до контекстных эмбеддингов. Переход к эмбеддингам улучшил «семантическое» сопоставление документов и описаний тематик. Для научных текстов важны доменные модели (например, SciBERT), так как терминология и стиль научной речи заметно отличаются от новостей и веб-текстов [3], [4], [5].

3. Семантический поиск, например, bi-encoder / cross-encoder, dense retrieval [6, 7, 8, 9]. Для сопоставления «статья ↔ описание секции» часто используют две схемы:

– Bi-encoder (двухбашенная модель). Отдельно кодируются статья и запрос/тема, затем рассчитается близость (по косинусной метрике). Это быстро и удобно для поиска по большому числу треков или при большом корпусе документов.
– Cross-encoder. Объединяет пару текстов и делает более точную оценку соответствия, но работает медленнее. Хорошо подходит для переранжирования. В информационном поиске широко применяется следующий конвейер: поисковая выборка → переранжирование. Сначала выполняется быстрый отбор кандидатов, затем строится точная модель для лучших K кандидатов. Для сценария конференции это выглядит естественно: сначала подобрать несколько наиболее вероятных секций, затем уточнить ранжирование и вероятность «вне тематики».

4. Применение больших языковых моделей (БЯМ) как zero-shot / few-shot классификатора и «оценщика» релевантности [10], [11], [12]. БЯМ позволяют:

– делать *zero-shot* классификацию по описанию секций без обучения на размеченных данных;
– использовать *few-shot* примеры, чтобы подстроить критерии релевантности;
– генерировать обоснования, повышая прозрачность решений для программного комитета.

При этом возникают проблемы: чувствительность к промпту, нестабильность, необходимость калибровки уверенности, риск галлюцинаций (особенно при оценке по неполному тексту), «подгонка» формулировок в описании секции.

5. Retrieval-Augmented Generation (RAG) и использование внешнего контекста [13].

Если секции описаны коротко или неоднозначно, помогает расширение контекста: добавление регламента конференции, тезауруса ключевых терминов, примеров уже принятых статей, описаний областей.

RAG-подход повышает устойчивость и уменьшает вероятность галюцинаций.

Применительно к задаче оценки релевантности научных статей тематикам конференции можно рассматривать такие вопросы как оценка качества, воспроизводимость и риски [14, 15, 16, 17].

Для реальной системы оценки релевантности научных статей важны:

- Метрики: accuracy/F1 (для классификации по секциям), AUROC (для сравнения «в-фокусе против вне-фокуса»), MRR/nDCG (для ранжирования секций), калибровка (ECE).
- Межэкспертное согласие (например, Коэна), чтобы понимать верхнюю границу качества и неоднозначность секций.
- Непредвзятость (fairness) и смещение (bias), особенно если решения влияют на возможность участия авторов в конференции.
- Надежность и детерминизм: фиксирование версии модели, параметра температуры, промптов; оценка устойчивости оценки к перефразированию.

Эксперимент по автоматической оценке релевантности

Для проверки возможности автоматической оценки релевантности научных статей с помощью БЯМ был проведен машинный эксперимент, схема которого показана на рисунке.

Эксперимент заключался в сопоставлении текстового содержимого статьи с описанием секций научной конференции. Данное сопоставление выполнялось с помощью локально развернутой большой языковой модели *Qwen3-14B*. Результат оценивания, возвращаемый языковой моделью, представлял собой вектор вещественных чисел:

$$(r_1, r_2, \ldots, r_N), \quad r_i \in [0,1]$$

где компонента r_i представляла собой оценку релевантности текстового содержимого статьи описанию i-ой секции конференции. Дополнительно языковая модель возвращала краткий текстовый комментарий к оценке релевантности. Количество секций N в эксперименте было установлено равным 9.

Для проведения эксперимента было разработано несколько скриптов на языке Perl, с помощью которых выполнялись все процедуры по оценке статей, кластеризации вектров релевантности и визуализации кластеров.

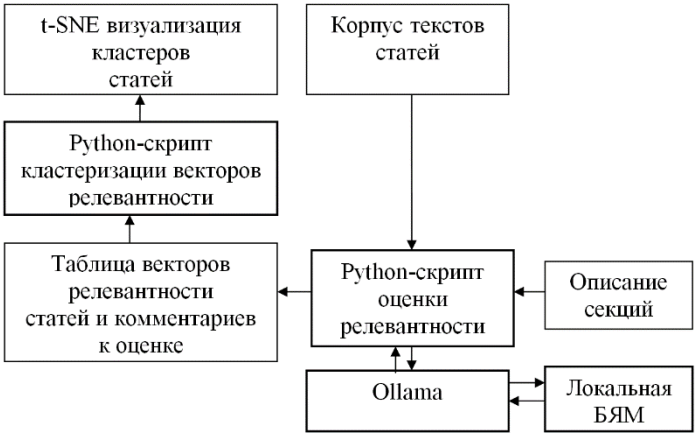


Рис. 1. Схема эксперимента по автоматической оценке релевантности статей

Взаимодействие скрипта с БЯМ выполнялось через платформу для локального запуска больших языковых моделей на компьютере – Ollama. Непосредственно оценка текста статьи в формате PDF выполнялась с помощью техники RAG на основе следующего промпта:

```
Цель: Оценить степень релевантности содержания рассматриваемой
научной статьи секциями научной конференции.

На научной конференции имеются следующие секции:
1 секция. Методы и технологии разработки программных систем.
2 секция. Компьютерное моделирование в фундаментальных и
прикладных исследованиях.
3 секция. Технологии обработки и защиты информации.
4 секция. Информационные системы и базы данных.
5 секция. Прикладное моделирование и E-business.
6 секция. Интеллектуальные информационные системы,
компьютерная лингвистика, технологии
информационного поиска.
7 секция. Суперкомпьютерные вычисления и BigData. Квантовые
информационные системы.
8 секция. САПР и цифровые технологии.
9 секция. Инфокоммуникационные системы, системы
телекоммуникаций, радиоэлектроника, обработка
сигналов.

Для каждой секции сравни текст научной статьи с названием и
списком ключевых тем этой секции, обязательно оцени степень
релевантности статьи для этой секции по числовой шкале:
```

```
минимальное значение для оценки – 0, максимальное значение для
оценки – 1.0

Все девять оценок объедини в один массив, который должен
возвращаться в поле "score".

Верни JSON:
{
  "score": [<0..1>, <0..1>, <0..1>, <0..1>, <0..1>, <0..1>,
<0..1>, <0..1>, <0..1>],
  "feedback": "Пояснение по результатам оценки подобия"
}
```

Описание секций конференции (название секции и перечень направлений) исходно представлялось также в PDF формате и после векторизации добавлялось к промпту.

Характеристики БЯМ Qwen3-14B:

– количество параметров: 14.8 миллиардов;
– количество слоев: 40;
– длина контекста: 32768 токенов.

В эксперименте оценка релевантности для корпуса размером порядка 200 статей заняла по времени около 3 часов. Для эксперимнта был использован компьютер Intel Core i7, 32ГБ DDR5, GeForce RTX 5060 Ti 16ГБ, SSD 1000 ГБ.

Результаты оценки релевантности статей были сведены в таблицу. Ниже для примера приводится фрагмент, содержащий 6 записей.

На основе векторов релевантности статей была проведена кластеризация методом k-средних. Результат кластеризации был визуализирован с помощью инструмента t-SNE. На рисунке 2 выявленные кластеры обозначены контурами.

Визуальный анализ полученного распределения позволяет увидеть несовпадение результатов ручного и автоматического распределения статей по секциям. Так для отдельных секций (например, 5) оба способа распределения дают хорошо перекрывающиеся группировки. Статьи секции номер 2, сформированной вручную, вполне могут быть разделены на три и более части.

Таблица

Результаты оценки релевантности статей

id	section	score	feedback
1	1	[0.2, 0.7, 0.1, 0.1, 0.1, 0.3, 0.1, 0.5, 0.1]	Статья наиболее релевантна для секции 2 (компьютерное моделирование) и 8 (САПР), так как описывает алгоритм для 3D-моделирования и оптимизации данных. Однако темы статьи слабо пересекаются с большинством других секций, за исключением упоминания вычислительных методов в секции 6.
2	1	[0.3, 0.1, 0.2, 0.1, 0.4, 0.1, 0.1, 0.1, 0.1]	Статья слабо соответствует большинству секций, кроме 1 (технологии блокчейн) и 5 (E-business). Основной фокус на анализе рынка NFT и игровых экономик, что частично пересекается с секциями 1 и 5, но не соответствует ключевым темам других секций.
3	1	[0.7, 0.2, 0.2, 0.2, 0.2, 0.1, 0.5, 0.2, 0.1]	Статья наиболее релевантна для секции 1 (методы разработки ПО) и 7 (суперкомпьютерные вычисления), так как описывает интеграцию Unity Job System в ECS-фреймворк для параллельных вычислений. Меньшая релевантность для остальных секций из-за отсутствия тем, связанных с моделированием, базами данных, квантовыми системами или телекоммуникациями.
4	1	[0.2, 0.4, 0.8, 0.1, 0.1, 0.9, 0.1, 0.1, 0.3]	Статья наиболее релевантна для секций 3 (технологии обработки информации) и 6 (интеллектуальные системы), где рассматриваются нейросети и машинное обучение. Умеренная релевантность для секции 2 (моделирование) и 9 (обработка сигналов), но низкая релевантность для остальных секций из-за отсутствия соответствия ключевым тем.
5	1	[0.8, 0.3, 0.2, 0.1, 0.5, 0.3, 0.1, 0.2, 0.1]	Статья наиболее релевантна для секции 1 (методы разработки ПО) и секции 5 (моделирование бизнес-процессов), так как описывает архитектурные подходы к обработке данных с использованием AOP и Java. Меньшая релевантность для остальных секций, так как не затрагивает темы BigData, квантовых систем, телекоммуникаций или обработки сигналов.
6	1	[0.2, 0.3, 0.1, 0.1, 0.1, 0.5, 0.1, 0.1, 0.4]	Статья наиболее релевантна для 6 секции (интеллектуальные системы, машинное обучение) и 9 секции (обработка сигналов). Менее релевантна для остальных секций, так как фокусируется на конкретных алгоритмах распознавания лиц, а не на широком спектре тем секций.

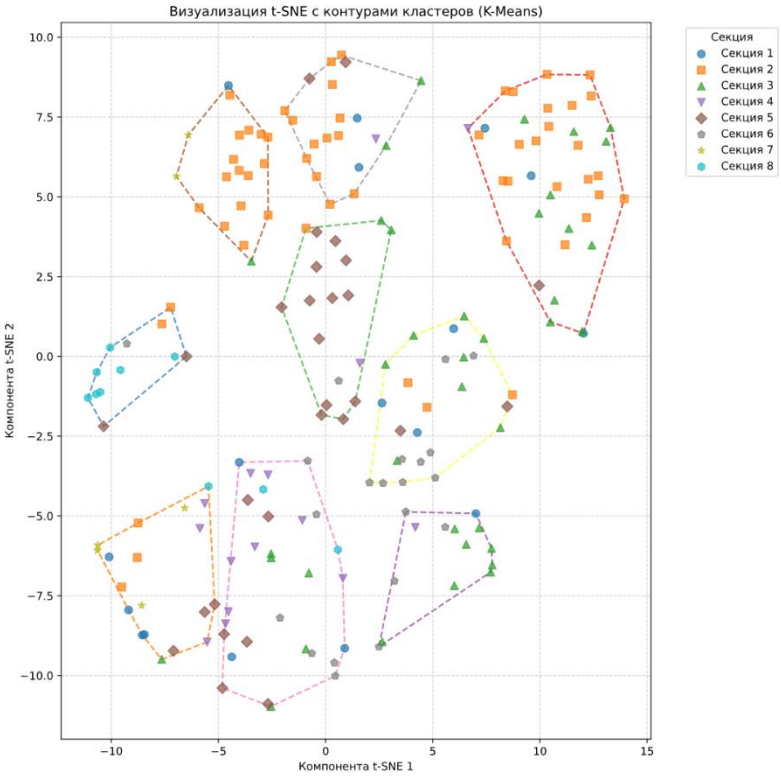


Рис. 2. Визуализация кластеров статей

Для объяснения полученного результата был проведен автоматический анализ текстовых описаний секций конференции, доступных на веб-сайте конференции. С этой целью были использованы четыре БЯМ, которые на основе технологии RAG по промпту выполняли анализ описаний секций и формировали конкретные рекомендации по редизайну существующей структуры секций. Обобщенный результат оценки структуры секций с помощью БЯМ приводится далее.

«Текущая структура отражает историческое наложение тематик, а не логическую декомпозицию современной исследовательской повестки в области информатики и вычислительных наук.

Анализ текущей структуры выявил несколько системных проблем, требующих корректировки секций для повышения логичности и современности конференционной программы:

1. Множественные тематические пересечения. Наиболее критичная проблема – дублирование ключевых направлений в разных секциях.
2. Неравномерность тематической нагрузки. Некоторые секции аккумулируют до 5–6 разнородных направлений. Другие секции остаются узкими и слабонаполненными. Следствие: Дисбаланс загрузки секций, сложность формирования программы, потеря фокуса конференции.
3. Смешение методологических и предметно-ориентированных подходов. В структуре отсутствует единая логика разделения: одни секции построены вокруг методов (моделирование, ИИ, разработка ПО); другие – вокруг прикладных областей (телеком, радиоэлектроника, бизнес). Отсутствует единая классификационная логика, нет ответа на вопрос: секции группируются по методам, по областям применения или по технологическим стекам? Следствие: структура не масштабируется и плохо адаптируется к появлению новых направлений.
4. Устаревшая терминология и отсутствие упоминания современных трендов в описаниях секций.
5. Неоптимальная балансировка фундаментальных и прикладных аспектов.

Ключевой вывод. Текущая структура отражает историческое наложение тематик, а не логическую декомпозицию современной исследовательской повестки в области информатики и вычислительных наук. Требуется либо четкое разграничение границ секций при сохранении их числа, либо стратегическая реорганизация с выделением самостоятельных треков по ИИ, кибербезопасности, управлению данными и цифровой трансформации. Наиболее критичным является выделение информационной безопасности в отдельную секцию и устранение дублирования тематики моделирования».

Можно заметить, что некоторые из упоминавшихся в обобщенном ответе критические проблемы в имеющейся структуре секций конференции визуально отражены на рисунке 2.

Заключение

В настоящем докладе рассмотрена задача автоматической оценки релевантности научных статей тематическим секциям конференции с применением больших языковых моделей. В ходе исследования был реализован и апробирован экспериментальный подход, основанный на локальном развёртывании модели Qwen3-14B и технологии RAG для сопоставления текстового содержания статей с описаниями секций.

Эксперимент показал практическую реализуемость такого подхода: за приемлемое время (примерно 3 часа на корпус из 200 статей) модель формирует векторную оценку релевантности по девяти секциям с кратким текстовым обоснованием. Полученные векторы были подвергнуты кластеризации методом k-средних и визуализированы с помощью t-SNE, что позволило наглядно выявить несоответствия между автоматическим распределением статей и их исходным распределением по секциям, выполненным организаторами. В частности, обнаружено, что для ряда секций (например, секция 2) ручная классификация объединяет тематически разнородные работы, тогда как автоматическая оценка указывает на возможность их дробления.

Дополнительно проведен анализ описаний секций силами нескольких БЯМ, который позволил сформулировать системные недостатки текущей структуры конференции: тематические пересечения, неравномерность нагрузки, смешение методологического и прикладного принципов группировки, устаревшую терминологию. Предложенные рекомендации направлены на устранение дублирования тем, выделение самостоятельных треков (информационная безопасность, ИИ, управление данными) и приведение рубрикатора в соответствие с современной исследовательской повесткой.

Таким образом, экспериментально подтверждено, что большие языковые модели способны выступать не только в роли «оценщика» релевантности, но и в роли аналитического инструмента для редизайна самой структуры конференции. Применение БЯМ в данной задаче повышает прозрачность принятия решений, снижает субъективность первичного отбора и дает организаторам объективную обратную связь для совершенствования тематической классификации.

Дальнейшие исследования было бы целесообразно направить на:

- расширение экспериментальной базы с привлечением экспертных оценок для верификации результатов;
- изучение влияния различных промптов и методов калибровки на устойчивость и воспроизводимость оценок;
- исследование возможности адаптации предложенного подхода к другим научным мероприятиям и областям знания.

Внедрение подобных автоматизированных инструментов в практику работы программных комитетов способно повысить эффективность обработки заявок и качество формирования научной программы конференций в условиях растущего потока междисциплинарных исследований.

Благодарности

Работа выполнена при поддержке Фонда Потанина.

Список литературы

1. Blei, D. M., Ng, A. Y., Jordan, M. I. Latent Dirichlet Allocation. Journal of Machine Learning Research. 3, null (3/1/2003), 993–1022. [Электронный ресурс]. – Режим доступа: https://dl.acm.org/doi/10.5555/944919.944937
2. Joachims, T. (1998). Joachims, Thorsten. (1998). Text Categorization with Support Vector Machines. Machine Learning: ECML-98. ECML 1998. Lecture Notes in Computer Science, vol. 1398. Springer, Berlin, Heidelberg. – [Электронный ресурс]. – Режим доступа: https://doi.org/10.1007/BFb0026683
3. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. arXiv:1301.3781. – [Электронный ресурс]. – Режим доступа: https://arxiv.org/abs/1301.3781
4. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. 10.48550/arXiv.1810.04805. – [Электронный ресурс]. – Режим доступа: https://arxiv.org/pdf/1810.04805
5. Beltagy, I., Lo, K., Cohan, A. SciBERT: A Pretrained Language Model for Scientific Text. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp. 3615–3620.
6. Reimers, N., Gurevych, I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pp. 3982–3992.
7. Nogueira, R., Cho, K. Passage Re-ranking with BERT. arXiv:1901.04085. – [Электронный ресурс]. – Режим доступа: https://arxiv.org/abs/1901.04085
8. Karpukhin, V., Oguz, B., Min, S., et al. (2020). Dense Passage Retrieval for Open-Domain Question Answering. EMNLP.
9. Lin, J., Nogueira, R., Yates, A. Pretrained Transformers for Text Ranking: BERT and Beyond. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Tutorials, pp. 1–4. – [Электронный ресурс]. – Режим доступа: https://aclanthology.org/2021.naacl-tutorials.1.pdf.
10. Brown, T. B., Mann, B., Ryder, N., et al. Language Models are Few-Shot Learners. 10.48550/arXiv.2005.14165. – [Электронный ресурс]. – Режим доступа: https://doi.org/10.48550/ARXIV.2005.14165.

11. Wei, J., Wang, X., Schuurmans, D., et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. In Proceedings of the 36th International Conference on Neural Information Processing Systems (NIPS '22). Curran Associates Inc., Red Hook, NY, USA, Article 1800, 24824–24837.

12. Liu, P., Yuan, W., Fu, J., et al. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. ACM Comput. Surv. 55, 9, Article 195 (September 2023), 35 p. – [Электронный ресурс]. – Режим доступа: https://doi.org/10.1145/3560815

13. Lewis, P., Perez, E., Piktus, A., et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS '20). Curran Associates Inc., Red Hook, NY, USA, Article 793, 9459–9474.

14. Belinkov, Y., Glass, J. Analysis Methods in Neural Language Processing: A Survey. Transactions of the Association for Computational Linguistics. 7. 49-72. – [Электронный ресурс]. – Режим доступа: https://dx.doi.org/10.1162/tacl_a_00254

15. Mitchell, M., Wu, S., Zaldivar, A., et al. Model Cards for Model Reporting. In Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19). Association for Computing Machinery, New York, NY, USA, 220–229. – [Электронный ресурс]. – Режим доступа: https://doi.org/10.1145/3287560.3287596.

16. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21). Association for Computing Machinery, New York, NY, USA, 610–623. – [Электронный ресурс]. – Режим доступа: https://doi.org/10.1145/3442188.3445922.

17. OpenAI. GPT-4 Technical Report. arXiv:2303.08774. – [Электронный ресурс]. – Режим доступа: https://arxiv.org/abs/2303.08774

Модель педагогических измерений на основе мультимодальных генеративных технологий ИИ

А. В. Сычев, email: sav@sc.vsu.ru
А. И. Дымов, email: artem.dymovv4884@gmail.com

Воронежский государственный университет

Аннотация. *Статья посвящена анализу потенциала мультимодальных генеративных технологий искусственного интеллекта (ИИ) для преодоления системного кризиса широко используемых сегодня методов педагогических измерений. Рассматриваются ключевые недостатки существующих подходов к оцениванию: доминирование итоговой оценки, фрагментарность, деконтекстуализация и субъективность. Целью работы является обоснование необходимости и обсуждение конкретных подходов к концептуальному проектированию модели педагогических измерений нового поколения, основанной на интеграции мультимодальных данных (текст, аудио, видео, поведенческие сигналы). Обсуждается опыт апробации платформы с элементами дополненной реальности ИКСАР в качестве мультимодального интерфейса пользователя в перспективной системе управления обучением, построенной на базе генеративных технологий ИИ.*

Ключевые слова: *педагогические измерения, оценивание образовательных результатов, мультимодальный искусственный интеллект, генеративные модели, персонализированное обучение, формирующее оценивание, цифровой профиль обучающегося, этика ИИ, AR-системы.*

Введение

Современное образование переживает системный кризис традиционных методов оценивания, основанных преимущественно на измерении легко формализуемых знаний с помощью стандартизированных тестов и экзаменов. Эти методы, выполняющие функцию административного контроля, все чаще демонстрируют свою ограниченность и несоответствие целям профессиоевльного развития обучающихся. Ключевые противоречия заключаются в следующем:

1. Доминирование итоговой оценки над формирующей. Традиционные методы (главным образом, экзамены и контрольные работы) фиксируют результат в отдельный момент времени, но не дают

© Дымов А. И, Сычев А. В., 2026

понимания самого процесса обучения: как обучающийся пришел к решению, какие когнитивные стратегии использовал, где возникли и почему преодолевались затруднения.

2. Фрагментарность и деконтекстуализация. Оцениванию подвергаются изолированные, часто искусственные учебные задания, а не комплексные способности к применению знаний в реальных, многокомпонентных ситуациях.

3. Субъективность и отсутствие персонализации. Даже при использовании тестов интерпретация результатов и, особенно, экспертные оценки преподавателей подвержены влиянию человеческого фактора.

Эти проблемы формируют насущную необходимость в новой парадигме педагогических измерений, которая должна быть комплексной, динамичной, объективной и персонализированной. Разработка модели педагогических измерений на основе мультимодальных генеративных систем ИИ представляется адекватным ответом на этот вызов, так как такие системы, имитируя человеческую способность к полисенсорному восприятию и контекстному анализу, обеспечивают техническую возможность реализации этой парадигмы оценки образовательного процесса во всей его многогранности.

Кризисная ситуация в области педагогических измерений усугубляется технологической ограниченностью широко используемых сегодня цифровых инструментов оценивания. Доминируют одномодальные системы, которые анализируют лишь один тип данных. Этот фундаментальный недостаток не позволяет улавливать контекстные нюансы и скрытые смыслы. Реальный образовательный процесс насыщен ситуациями, где истинный смысл коммуникации зачастую находится за пределами текста. Например, словесный ответ «все понятно», произнесенный с саркастической интонацией и сопровождающийся определенной мимикой, несет прямо противоположное значение. Отсутствуют как интегративная методология так и технологическая платформа, способные синхронно принимать, анализировать и интерпретировать гетерогенные данные об учебной деятельности.

В данной статье обсуждаются вопросы, связанные с разработкой концептуальной модели педагогических измерений нового поколения, которая интегрирует потенциал мультимодальных генеративных систем искусственного интеллекта. Модель должна быть нацелена на перевод педагогической диагностики из эпизодической констатации результатов в процесс непрерывного, контекстно-обогащенного понимания индивидуального учебного пути обучающегося.

В частности рассматриваются следующие задачи:

- Анализ теоретических основ и методологических вызовов современных педагогических измерений.
- Определение архитектурных и функциональных принципов мультимодальных генеративных систем ИИ, релевантных для образовательного контекста.
- Определение структурно-функциональной модели интеграции и ее верификация на практических кейсах.
- Выявление и систематизиация этических, методологических и технологических вызовов, связанных с внедрением модели.

2. Мультимодальные генеративные системы ИИ: их сущность и архитектура

Мультимодальные генеративные системы искусственного интеллекта – это модели, которые способны принимать, обрабатывать и генерировать информацию, относящуюся к нескольким различным модальностям. К этим модальностям можно отнести текст, изображения, аудио, видео или сенсорные данные. Существенным является также, что работа со всеми этими источниками информации может происходить одновременно. Комбинированное понимание разных модальностей позволяет интеллектуальной системе строить более полные, контекстуально насыщенные и точные ответы. Такая мультисенсорность аналогична обрабаботке человеком информации, поступающей в его мозг через множество органов чувств. В сравнении с моделями, которые работают с одной из модальностей, здесь достигаются более высокие показатели точности анализа информации благодаря перекрестным ссылкам на информацию, поступающую из разных каналов. Благодаря этому противоречия или неопределенности в одном потоке данных могут быть либо подтверждены либо исправлены с использованием информации в других модальностях. Мультимодальная модель способна анализировать не только текст, но и поведение студента [2].

Функциональная структура мультимодальных моделей может включать три основных компонента: модуль ввода данных, модуль обработки и модуль вывода результатов. Модуль ввода данных работает как интерфейс для получения данных системой ИИ, предоставляя ей текст, изображения, аудио, видео и сигналы датчиков. Также он может производить их предварительную обработку, чтобы сделать эти данные пригодными для дальнейшего анализа. Модуль обработки объединяет данные из разных источников, определяет скрытые закономерности в них, выделяет ключевые признаки и формирует единое представление, отражающее смысл всей мультимодальной информации. Модуль вывода формирует прогнозные оценки, рекомендательные решения,

синтезированный контент или практические выводы. Формат представления выходной информации также может быть мультимодальным: текст, изображения, видео, аудио. Это определяется спецификой приложения [4].

3. Интеграция педагогических измерений в мультимодальные системы ИИ

Педагогические измерения – это область педагогики и психометрии, занимающаяся количественной и качественной оценкой образовательных процессов и результатов обучения. Традиционно они включают в себя следующие аспекты:

- академическую успеваемость (баллы за тесты и экзамены, оценки и т.д.);
- когнитивные показатели (уровень понимания, способность к анализу и синтезу, решение проблем);
- психоэмоциональные и поведенческие показатели (вовлеченность, мотивация, уровень стресса, социальная активность).

Основной целью педагогических измерений является получение и предоставление объективной информации о реальном уровне учебных достижений обучающихся [10].

В свою очередь эти измерения и метрики делятся на группы, каждая из которых имеет свои задачи:

- Метрики результата обучения. Эти измерения оценивают непосредственно знания и навыки.
- Диагностические метрики, которые применяются для выявления проблем в обучении.
- Процессуальные метрики необходимы для описания того, как именно студент учится.
- Также выделяют группы вовлеченности, адаптивности, прогресса, качества педагогического контента [9].

Если говорить о педагогических метриках в контексте мультимодальных генеративных систем ИИ, то здесь открываются новые возможности для измерения того, что раньше было невозможно оценить. Например, это могут быть невербальные реакции на сложные темы, соответствие реальных навыков заявленным, определение момента потери внимания и многое другое. Рассмотрение педагогических метрик помогает понять, какие показатели могут быть использованы для анализа и интерпретации учебной деятельности. Однако важно определить, каким образом мультимодальные генеративные системы ИИ могут применять эти метрики на практике и какие педагогические задачи способны решать.

С точки зрения целей мультимодальных систем в педагогике, разработка и применение таких моделей направлены на решение задач, связанных с формированием комплексной и динамической модели обучаемого. Для этого объединяются разнообразные данные: результаты тестирования и контрольных работ (знания), поведенческие характеристики (активность, вовлеченность), эмоциональное состояние (выражение лица, тон голоса) и личные предпочтения в обучении. Такая комплексная картина позволяет не просто оценить уровень усвоения материала, но и понять особенности обучаемого – его сильные и слабые стороны, мотивацию, стрессовые реакции и другие факторы, влияющие на учебный процесс. Кроме того, цифровой профиль динамически обновляется по мере прохождения курса, отражая изменения и прогресс в реальном времени. Мультимодальные системы дают возможность выявлять не только явные пробелы в знаниях, но и причины затруднений – будь то когнитивные сложности, низкая мотивация или эмоциональные барьеры. Как результат – материал преподносится с учетом индивидуальных особенностей: меняется сложность заданий, формат подачи (текст, видео, интерактивные упражнения), темп обучения и другие параметры. Это способствует более эффективному усвоению материала и снижает риск демотивации и выгорания.

Мультимодальные модели позволяют внедрить принцип формирующего оценивания – постоянного мониторинга прогресса обучаемого на протяжении всего образовательного процесса. Это достигается за счет сбора и анализа данных из множества источников и модальностей, позволяя своевременно выявлять затруднения и корректировать обучение. На основе анализа данных и выявленных потребностей каждого обучаемого мультимодальные системы способны автоматически создавать персонализированные учебные материалы и задания. Это могут быть тексты с учетом уровня понимания, интерактивные упражнения, подсказки, рекомендации и обратная связь, которые адаптируются под текущие возможности и состояние студента.

Кроме того, система может подстраивать формат подачи: например, использовать больше визуальных элементов для визуалов или аудиоконтент для аудиалов. Такой подход значительно повышает вовлеченность и эффективность обучения [9].

4. Подходы к разработке модели педагогических измерений на основе мультимодальных генеративных ИИ

Для реализации выше рассмотренных выше педагогических целей одних лишь данных и метрик недостаточно. Важное значение имеет архитектура самих мультимодальных систем и методы их обучения.

Насколько точно и эффективно модель сможет анализировать поведение обучаемого – напрямую зависит от применяемых технологических подходов.

Рассмотри ключевые архитектурные и методологические решения, заложенные в основу мультимодальных ИИ-моделей.

Раннее и позднее слияние

При данном подходе анализ разделяется на два этапа. Раннее слияние подразумевает объединение сырых или низкоуровневых признаков из разных модальностей перед их обработкой моделью. Позднее слияние – это независимая обработка каждой модальности с последующим объединением результатов на уровне предсказаний или решений.

Механизм скрещенного внимания

Это современный подход, при котором модель учится «сопоставлять» фрагменты информации из разных модальностей. Например, в образовательной среде система может сопоставлять текстовое описание «ученик допускает ошибку при решении уравнений» с конкретной областью на изображении тетради или планшета, где эта ошибка допущена. Такой механизм лежит в основе моделей типа CLIP и позволяет ИИ более точно интерпретировать учебные данные и выявлять причины затруднений. Кроме того, к современным подходам для решения перечисленных задач относятся трансформеры с мультимодальной архитектурой, генеративные модели и другие методы.

Архитектурные и методологические решения от раннего и позднего слияния до механизмов скрещенного внимания и генеративных мультимодальных трансформеров определяют не только техническую структуру системы, но и ее функциональные возможности в педагогическом контексте.

Мультимодальные модели не существуют в отрыве от инструментов, которые обеспечивают их разработку, обучение, настройку и последующую оценку качества. Именно поэтому следующий важный аспект – это экосистема инструментов, позволяющая воплощать перечисленные архитектурные идеи в работающие педагогические решения. Она включает в себя как крупные предобученные мультимодальные модели, так и фреймворки для их дообучения, специализированные методы оптимизации, наборы метрик и инфраструктуру для сбора разнообразных обучающих данных [1].

1. Такие системы, как OpenAI GPT-4V, Google Gemini, Microsoft Kosmos и Meta ImageBind, служат фундаментом для понимания и генерации мультимодальной информации. В контексте

педагогических измерений они позволяют анализировать сложные комбинации данных – текстовые ответы обучаемого, изображения выполненных заданий, аудиозаписи чтения вслух или видеозаписи взаимодействия с учебным материалом. Использование таких моделей повышает точность диагностики и обеспечивает согласованность оценивания при работе с разными модальностями.

2. Инструменты PyTorch, TensorFlow и Hugging Face Transformers предоставляют гибкую инфраструктуру для разработки мультимодальных моделей, позволяя интегрировать методы анализа текста, изображений и аудио в единую систему. С точки зрения педагогических измерений эти фреймворки дают возможность строить модели, соблюдающие психометрические принципы, например, учитывать латентные факторы, распределение ошибок, свойства заданий и показатели обучаемых.
3. Методы эффективной адаптации, такие как LoRA, P-Tuning и адаптеры, позволяют настраивать крупные мультимодальные модели под конкретные педагогические задачи: диагностику ошибок письма, оценку выразительности речи, анализ невербальных проявлений вовлеченности. Такой подход повышает конструктивную валидность (то есть соответствие измеряемой сущности) и уменьшает риски систематических ошибок, поскольку модель обучается на данных, непосредственно связанных с образовательной практикой.

Оценивание мультимодальных систем в педагогике должно сочетать технические и педагогические метрики.

Технические показатели включают:

– точность распознавания эмоций;
– показатели BLEU и ROUGE для анализа качества генерации текстовых объяснений и обратной связи;
– F1-оценку для классификации типов ошибок или уровней понимания.

Педагогические метрики, такие как конструктивная, содержательная и прогностическая валидности, позволяют определить, насколько модель действительно измеряет то, что должна измерять, и способна ли ее диагностика предсказать реальные учебные результаты. Это критически важно для интеграции ИИ в образовательные процессы, где каждое решение должно быть обоснованным, прозрачным и соответствовать этическим стандартам [3].

Для сбора данных для мультимодальных систем могут использоваться камеры, микрофоны, системы eye-tracking, сенсоры позы и активности, а также лог-данные цифровых образовательных платформ. Все они служат основой для построения достоверных мультимодальных датасетов. Для педагогических измерений ключевым является не только объем собранных данных, но и их качество: минимизация шума, согласованность между модальностями, защита конфиденциальности и соблюдение этических норм. Корректно собранные данные обеспечивают надежность оценивания и позволяют проводить анализ латентных характеристик обучаемых мотивации, внимания, когнитивной нагрузки [6].

Примерами успешного внедрения технологий мультимодальных генеративных ИИ в учебный процесс для педагогических измерений могут служить следующие кейсы: персонализированный тьютор-бот «Khanmigo» от Khan Academy (на базе GPT-4) и исследовательские прототипы от Stanford и MIT.

Пример 1. Обучающийся решает геометрическую задачу на интерактивном холсте. Система «видит» какие фигуры и линии рисует обучающийся не через камеру, а посредством анализа действий в цифровой среде. Что касается текста, обучающийся печатает или рассуждает вслух. Исходя из этих данных, модель помогает ему и параллельно составляет портрет его успеваемости [5].

Пример 2. Оценка навыков презентации и публичных выступлений в приложениях Yoodli и Orai. Пользователь записывает свое выступление на видео – например, репетицию защиты выпускной квалификационной работы, – а мультимодальная модель параллельно анализирует текст, аудио и визуальные сигналы. По транскрипту речи система оценивает наличие слов-паразитов, повторов, убедительность формулировок и логичность структуры. Аудиоанализ выявляет темп речи, паузы, интонацию, эмоциональную выразительность и количество заполнителей вроде «э-э-э». Видеосигнал позволяет определить качество зрительного контакта с камерой, осанку, жестикуляцию, степень открытости позы и наличие улыбки. На основе всего этого ИИ формирует персонализированный отчет с оценками по категориям и конкретными рекомендациями. Например, система указывает, что выражение «в общем-то» прозвучало N раз и предлагает заменить его осознанными паузами; отмечает, что скорость речи составляет 175 слов в минуту и советует сделать несколько глубоких вдохов, чтобы замедлиться; фиксирует, что взгляд в камеру держался лишь 38% времени и рекомендует завершать каждую мысль короткой фиксацией взгляда.

Кроме того, система может предложить упражнения, направленные на развитие слабых сторон и отработку конкретных навыков.

5. Ограничения и этические вызовы внедрения модели

Несмотря на значительный потенциал мультимодальных систем генеративного искусственного интеллекта в образовательной среде, их использование сопряжено с рядом серьезных проблем и ограничений. Эти трудности затрагивают не только технологическую, но и педагогическую, этическую и психометрическую стороны внедрения.

Одной из наиболее острых проблем является сбор и обработка большого массива биометрических и поведенческих данных учащихся: изображения лица и позы, голос, направления взгляда, реакции в реальном времени, результаты действий в цифровой среде. Они имеют высокую ценность для построения точных моделей, но несут потенциальную угрозу приватности и безопасности обучаемых. Любая утечка подобной информации может иметь серьезные последствия, включая риск идентификации личности, психологического давления или дискриминации. Поэтому внедрение мультимодальных систем требует строгого регулирования, прозрачных процедур хранения данных и обязательного информированного согласия всех участников образовательного процесса.

Следующая проблема – смещение оценок. Обучение моделей на нерепрезентативных или однобоких выборках неизбежно приводит к возникновению алгоритмических смещений. Например, модели, обученные преимущественно на лицах одной этнической группы, хуже распознают эмоции представителей других групп. Аналогично, системы могут неправильно интерпретировать выражения лица у людей с расстройствами аутистического спектра или неверно классифицировать определенные акценты в речи. В контексте педагогических измерений такие ошибки могут привести к неадекватным оценкам, неверным диагностическим выводам и, как следствие, к нарушению равных образовательных возможностей. Проблема смещения требует глубоких исследований, разнообразных обучающих датасетов и постоянной проверки моделей на устойчивость и справедливость.

Современные мультимодальные модели, особенно основанные на трансформерах, обладают высокой сложностью внутренней структуры. Их решения трудно интерпретировать как педагогам, так и обучаемым. В результате возникает проблема доверия. Например, если система выставляет оценку уровню концентрации или предполагает наличие затруднения в понимании темы, педагог должен иметь возможность объяснить это решение. Недостаточная прозрачность ИИ подрывает принятие технологии и затрудняет ее внедрение в реальные

образовательные процессы. Для решения этой проблемы требуется развитие механизмов интерпретируемого ИИ, визуализация внимания модели, а также адаптация психометрических методик к мультимодальным системам.

Мультимодальные модели предъявляют высокие требования к вычислительным ресурсам, качеству входных данных и инфраструктуре. Для корректной работы необходимы стабильное освещение, качественные камеры и микрофоны, точный сбор лог-данных. Малейшие технические сбои шум в аудио, низкое разрешение, нестабильное интернет-соединение могут существенно снизить точность интерпретации модели. Кроме того, создание высококачественных датасетов требует трудоемкой ручной разметки, что делает разработку педагогически ориентированных систем дорогостоящей и долгосрочной задачей. В педагогике крайне важно, чтобы любое измерение соответствовало заявленному конструкту: если модель утверждает, что обучающийся рассеян, необходимо доказать, что анализируемые сигналы действительно отражают сниженное внимание, а не, например, усталость, тревогу или особенности мимики. В мультимодальных системах эта проблема усугубляется множеством источников данных, каждый из которых добавляет собственные источники шума и неопределённости. Поэтому необходимо тщательно проверять модели на содержательную, конструктивную и прогностическую валидности.

В совокупности перечисленные проблемы формируют комплекс вызовов, без решения которых мультимодальные модели не могут быть полноценно интегрированы в образовательную практику. Их осознание и критическое исследование являются необходимым шагом для разработки безопасных, справедливых и педагогически обоснованных технологий мультимодального искусственного интеллекта.

6. Аппаратно-программные пользовательские интерфейсы мультимодальных систем управления обучением

Переход к мультимодальной модели организации учебного процесса и педагогических измерений потребует также использование новых технических средств обучения. Помимо мультимодальных нейросетевых систем обработки информации в рамках новой организации учебного процесса будут также востребованы устройства, обеспечивающие мультимодальный аппаратно-программный интерфейс, обеспечивающий взаимодействие пользователя с системой управления обучением.

Как нам представляется, в качестве одного из русскоязычных инструментов, обеспечивающих требуемый мультимодальный интерфейс, можно рассматривать цифровой инструмент «полевого» сотрудника с элементами дополненной реальности – ИКСАР

(https://iksar.pro/), разработка и поддержка которого осуществляется компанией «ИКСАР Технологии».

В 2025-2026 учебном году на факультете компьютерных нак ВГУ в порядке апробации возможностей платформы ИКСАР магистрантами были выполнены лабораторные работы, в которых им ставилась задача разработки сценария выполнения некоторого задания с последующей реализацией этого сценария в виде процесса, сконструированного в специальном визуальном редакторе и состоящего из последовательности шагов с мультимодальным контентом. Готовый процесс далее загружался на мобильное устройство пользователя (смартфон, AR-очки) и исполнялся с помощью команд управления: голосовых (очки, смартфон), экранно-сенсорных (смартфон), кнопочных (очки). Отображение шагов исполняемого процесса производилось либо на экране смартфона, либо на мини-дисплее AR-очков. Помимо визуального отображения инструктивная информация могла выдаваться пользователю в виде аудиозаписи. Выполнение функций на каждом шаге активировалось одним из способов (голос или сенсорный ввод) с возможностью фиксации действий, выполняемых пользователем, в любых модальностях: текст, фото, аудио, видео. Сохраненная таким образом на мобильном устройстве информация отправлялась затем на облачный сервер ИКСАР, где заносилась в журнал процесса и могла быть использована в дальнейшем для анализа и контроля действий мобильного пользователя. Платформа также позволяет в рамках исполняемого процесса активировать элемент видеоконференцсвязи для консультации с удаленным экспертом.

Дополнительными преимуществами интерфейсной части платформы являются: свободные руки у мобильного пользователя (при использовании AR очков), графические цифровые инструкции, головное управление, видео-конференц связь, поддержка offline, Wi-Fi, 3G, 4G, 5G. Серверная часть платформы предоставляет следующие функциональные возможности: планирование задач, управление ресурсами, управление устройствами, информационная безопасность. Встроенный графический редактор обеспечивает быструю no-code разработку цифрового процесса. Обеспечена поддержка AR очков всех ведущих производителей для промышленного рынка AR: RealWear, Rokid, Vuzix.

Апробация возможностей ИКСАР подтвердила реальные мультимодальные возможности этой платформы, потенциально позволяющие ее использовать в качестве одной из важных компонент системы управления обучением на основе мультимодальных генеративных технологий ИИ. Для реальной интеграции ИКСАР в обучающий контур требуется отработка механизмов ее взаимодействия

(интерфейсов) с мультимодальными большими языковыми моделями и классическими системами управления обучения (например, Moodle).

Заключение

Проведенный анализ показывает, что кризис традиционных педагогических измерений связан не только с их методологической ограниченностью, но и с принципиальной неспособностью одномодальных подходов адекватно отражать сложную, контекстно насыщенную природу образовательного процесса. Современные методы фиксации результатов обучения, концентрируясь преимущественно на измерении академических знаний, игнорируют эмоционально-поведенческие, когнитивно-процессуальные и мотивационные аспекты деятельности обучающегося. В результате педагогическое оценивание оказывается оторванным от реальных процессов понимания и развития.

Интеграция мультимодальных генеративных систем искусственного интеллекта открывает путь к формированию новой парадигмы педагогических измерений – динамичной, контекстуальной и персонализированной. Такие системы способны одновременно анализировать текстовые, аудио-, видео- и поведенческие сигналы, что позволяет перейти от суммарного контроля к непрерывному, формирующему оцениванию. Использование мультимодальных данных делает возможным построение цифрового профиля обучающегося, в котором объединяются академические результаты, когнитивные стратегии, эмоциональные реакции и особенности взаимодействия с учебной средой. Этот профиль может служить основой как для точной диагностики, так и для автоматической генерации адаптивных учебных рекомендаций, способствующих росту автономности и мотивации обучаемого.

Практическая апробация мультимодальной платформы ИКСАР подтвердила возможность поддержки такого подхода в условиях реальной образовательной среды. Использование технологий дополненной реальности и мультимодального взаимодействия (голос, жест, визуальные и аудиосигналы) демонстрирует, что подобные решения могут стать компонентом интегрированной системы управления обучением на базе ИИ, обеспечивая сбор, интерпретацию и визуализацию сложных данных о процессах обучения.

Концептуальная модель педагогических измерений нового поколения призвана сочетать принципы психометрической строгости с возможностями интеллектуального анализа мультимодальных данных. Архитектурные решения, такие как механизм скрещенного внимания, раннее и позднее слияние модальностей, использование генеративных трансформеров, формируют технологическую основу для более

достоверного и комплексного анализа учебной деятельности. При этом важным условием становится соблюдение баланса между алгоритмической точностью и педагогической интерпретируемостью результатов.

Однако внедрение мультимодальных ИИ-систем сопряжено с рядом существенных вызовов: рисками нарушения приватности, возможными алгоритмическими смещениями, непрозрачностью принимаемых решений и высокими требованиями к инфраструктуре и качеству данных. Для их преодоления необходимо формирование нормативно-этических рамок, включающих обеспечение конфиденциальности, независимую экспертизу моделей, разработку стандартов валидности и процедур педагогической интерпретации. Только при совмещении технологической инновационности с этической ответственностью и психометрической обоснованностью возможно создание доверия к новым системам оценивания.

Мультимодальные генеративные технологии ИИ способны стать ключевым инструментом перехода от репродуктивной модели контроля знаний к системе анализа и поддержки образовательного развития личности. Реализация обсуждаемой в статье концепции может качественно изменить саму природу педагогических измерений, превратив их из инструмента фиксации результатов в средство понимания, сопровождения и персонализированного продвижения обучающегося в образовательном пространстве.

Дальнейшие исследования должны быть направлены на разработку стандартизированных методик валидизации таких систем, создание этически устойчивых алгоритмов обработки данных и формирование единой технологической экосистемы, объединяющей педагогическую, ИИ- и AR-инфраструктуру в целостное образовательное пространство будущего.

Список литературы

1. Baltrušaitis, T., Ahuja, C., Morency, L.-P. Multimodal Machine Learning: A Survey and Taxonomy // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2019. – Vol. 41, No. 2. – p. 423–443. – DOI: 10.1109/TPAMI.2018.2798607.

2. Explore AI in your teaching: a one hour professional development experience with Khan Academy // Official Blog of Khan Academy. – 2025 [Электронный ресурс] – Режим доступа: https://blog.khanacademy.org/explore-ai-in-your-teaching-a-one-hour-pd-experience-with-khan-academy/

3. Gao, J., Li, P., Chen, Z., Zhang, J. A Survey on Deep Learning for Multimodal Data Fusion // Neural Computation. – 2020. – Vol. 32, No. 5. – p. 829–864. – DOI: 10.1162/neco_a_01273.

4. Guerrero-Sosa, J.D.T., Romero, F.P., Menéndez-Domínguez V.H., et al. A Comprehensive Review of Multimodal Analysis in Education // Applied Sciences. – 2025. – Vol. 15, No 11. – p. 5896. – DOI: 10.3390/app15115896.

5. Guo, W., Wang J., Wang, S. Deep Multimodal Representation Learning: A Survey // IEEE Access. – 2019. – Vol. 7. – p. 63373–63394. – DOI: 10.1109/ACCESS.2019.2916887.

6. Lee, G., Shi, L., Latif, E., et al. Multimodality of AI for Education: Towards Artificial General Intelligence // arXiv. – 2023. – arXiv:2312.06037 [cs.AI]. – DOI: 10.48550/arXiv.2312.06037.

7. Liu, X. Self supervised Learning: Generative or Contrastive // arXiv. – 2021. – Version 5. – arXiv:2006.08218 [cs.LG]. – DOI: 10.48550/arXiv.2006.08218.

8. Brookhart, S. M., Nitko, A. J., Educational assessment of students / A. J. Nitko, S. M. Brookhart. – 8th ed. – Pearson College Div, 2018. – 560 p.

9. Mangaroska, K., Sharma, K., Gašević, D., Giannakos, M. Multimodal Learning Analytics to Inform Learning Design: Lessons Learned from Computing Education // Journal of Learning Analytics. – 2020. – Vol. 7, No. 3. – p. 79–97. – DOI: 10.18608/jla.2020.73.7.

10. Третьякова, Т. В. Педагогические измерения, их роль в оценивании качества образования // Вестник СВФУ. – 2013. – № 4. – С. 109–114. – URL: https://cyberleninka.ru/article/n/pedagogicheskie-izmereniya-ih-rol-v-otsenivanii-kachestva-obrazovaniya (дата обращения: 29.12.2025)

Применение искусственного интеллекта в системе поддержки принятия решений нефтехимических производств

Н. С. Толстошеин, email: Nik.Tolstoshein@yandex.ru
Д. А. Михин, email: dimamikhin200184@yandex.ru

Тамбовский государственный технический университет

Аннотация. *В исследовании рассмотрена роль искусственного интеллекта в системах поддержки принятия решений для нефтехимической индустрии. Описаны ключевые прикладные направления систем поддержки принятия решений. Разработана общая архитектура систем поддержки принятия решений с применением искусственного интеллекта. Проанализированы возможные риски и пути их снижения.*

Ключевые слова: *Искусственный интеллект, система поддержки принятия решений, нефтехимия.*

Введение

Искусственный интеллект сегодня становится ключевым фактором повышения эффективности и конкурентоспособности нефтехимических предприятий. Сложность производственных процессов, огромное количество сенсорных данных и требование к надежности и безопасности заставляют прибегать к интеллектуальным системам, способным объединять физику процессов, данные и бизнес-цели. Системы поддержки принятия решений (СППР) на базе искусственного интеллекта (ИИ) помогают специалистам принимать более обоснованные решения в реальном времени, оптимизировать продукцию и энергоэкономику, снизить риск аварий и соблюсти регуляторные требования. В нефтехимии задачи СППР часто связаны с управлением непрерывными и пакетными процессами, обработкой больших объемов данных и необходимостью учитывать неопределенности, вариативность сырья и состояние рынка.

1. Тенденции применения искусственного интеллекта в СППР нефтехимических предприятий

СППР – это совокупность моделей, инструментов визуализации и механизмов принятия решений, помогающих людям и автоматическим системам находить оптимальные решения в условиях неопределенности и ограничений [1]. В нефтехимической промышленности такие системы

© Толстошеин Н.С., Михин Д.А., 2026

решают, как правило, задачи разделения, переработки, энергосбережения, обеспечения качества продукции и безопасной эксплуатации. Обычно СППР интегрируются с существующими системами управления предприятием, управлением производственными процессами, системами мониторинга состояния оборудования и системами планирования.

Основные прикладные направления СППР нефтехимических предприятий представлены на рис 1.



Рис. 1. Основные прикладные направления СППР [2]

Искусственный интеллект в системах поддержки принятия решений нефтехимического производства повышает точность прогнозов и оптимизацию технологических режимов, обеспечивает быстрые рекомендации в реальном времени, учитывает неопределенность и позволяет проводить сценарный анализ и поддерживает прогнозное обслуживание.

В настоящий момент уже формируются следующие тенденции [3]:

– Интеграция цифровых двойников для всего цикла жизни завода: от проектирования и реконструкции до эксплуатации и модернизации;
– Совместное обучение моделей между несколькими площадками и компаниями без передачи чувствительных данных;
– ИИ будет усиливать энергоустойчивость и снижение выбросов через интегрированные решения;

– Усиление автоматических корректирующих действий, поддержка оператора, но с удержанием контроля;
– Расширение применения графовых нейронных сетей для сложных сетевых взаимосвязей между узлами процесса.
– Прогнозирование устойчивости цепей поставок в условиях рыночных колебаний и глобальных факторов.
– Более тесная интеграция с регуляторными требованиями и аудиторскими процедурами через объяснимые и прозрачные модели.

2. Общая архитектура СППР с применением ИИ

Общая архитектура СППР с ИИ строится как единый конвейер взаимосвязанных уровней (рис 2). В каждом уровне применяются ИИ-технологии: на источниках – обнаружение аномалий, оценка качества данных и заполнение пропусков; на подготовке – автоматическая очистка, нормализация и генерация признаков; в хранении – оценка надёжности источников и политики доступа; в формировании признаков – автоматическое извлечение релевантных признаков; в моделировании – машинное обучение, и объяснимый ИИ; в управлении моделями – версионирование, развёртывание и мониторинг, контроль дрейфа; в принятии решений – прогнозная аналитика, сценарное моделирование и устойчивые оптимизационные решения; в интеграции – обработка событий; в пользовательском интерфейсе – дашборды, сценарии “что – если” и поясняющие интерфейсы; в исполнении – внедрение рекомендаций; в цифровом двойнике – безопасное тестирование и обучение; в мониторинге – предиктивное обслуживание; в безопасности и соответствии – кибербезопасность, аудит и прослеживаемость данных. Такая архитектура поддерживает гибкость цифровой трансформации, управляемость данными и прозрачность решений.

Цифровой двойник нефтехимического комплекса – это интерактивная, обновляемая цифровая модель реального объекта [4]. Он объединяет физические принципы, данные и управление процессами, позволяя:

– Проводить виртуальные испытания и «что если» сценарии без риска для реального оборудования;
– Подготавливать специалистов к новым режимам без влияния на производство;
– Поддерживать точность прогнозов за счет постоянной калибровки по данным текущей эксплуатации;
– Встроенные цифровые двойники применяются для отдельных модулей (производственных линий, секций процесса) и для целых предприятий.

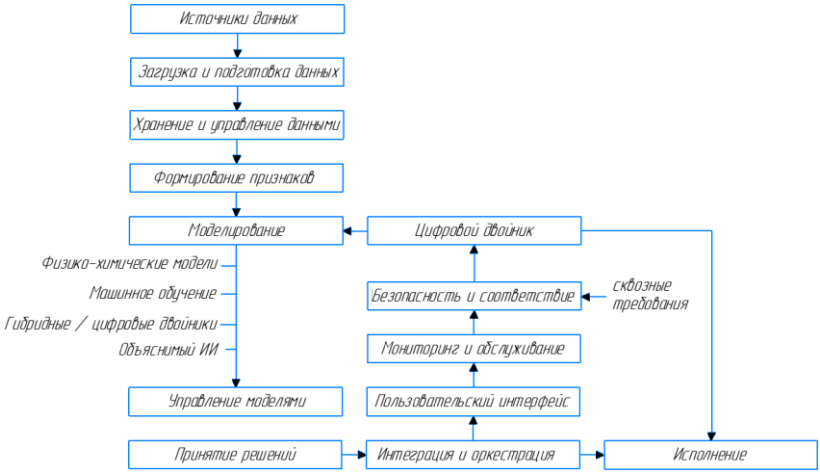


Рис. 2. Общая архитектура СППР с применением ИИ

Но также существуют определенные риски [5]:

– Неполные или шумные данные, пропуски, дрейф сигналов.
– Непрозрачность моделей (кодификация управленческих решений).
– Безопасность и киберугрозы.
– Регуляторные ограничения
– Влияние на рабочие процессы и человеческий фактор.

Чтобы уменьшить эти риски, при внедрении ИИ в СППР важно учитывать некоторые факторы:

Сочетание первопринципных моделей и данных: повышает точность и устойчивость к редким ситуациям, уменьшает зависимость от объема исторических данных.

Прогнозы и решения должны учитывать неопределенности в сырье, условиях реакции и рыночной конъюнктуре. Используются байесовские подходы, методы Монте-Карло и режимы сценариев.

Операторы должны понимать, почему система предлагает конкретное решение. В нефтехимии наличие прозрачности и соблюдение регуляторных требований критично.

СППР должны работать в связке с разными системами управления технологическими процессами и системами планирования. Важна совместимость, безопасность и минимизация воздействия на текущий процесс.

Заключение

Применение технологий искусственного интеллекта в нефтехимии для создания систем поддержки принятия решений открывает широкие возможности для повышения эффективности, надежности и устойчивости производственных процессов. Комбинация гибридных моделей, цифровых двойников, продвинутых методов прогнозирования и оптимизации позволяет не только уменьшить операционные затраты, но и повысить безопасность и качество продукции в условиях изменяющегося рынка и регуляторной среды. Успешная реализация требует системного подхода к данным, архитектуре, управлению изменениями и постоянному мониторингу моделей. В итоге СППР становится не просто инструментом автоматизации, а стратегическим элементом цифровой трансформации нефтехимического предприятия.

Список литературы

1. Николюкин, М.С. Моделирование поддержки принятия решений в системе управления складским комплексом / М. С. Николюкин, А.Д. Обухов, В. В. Алексеев // Правовая информатика. Научный центр правовой информации – 2023. – № 3. – С. 54-64.

2. Трахтенгерц, Э.А. Методы компьютерной поддержки формирования целей и стратегий в нефтегазовой промышленности / Э. А. Трахтенгерц, Ю. П. Степин. – М.: СИНТЕГ, 2007. – 335 с.

3. Егоров, А.А. Искусственный интеллект в нефтегазовой области. Перспективы развития 2025–2034 гг. Зарубежный опыт // Автоматизация и IT в нефтегазовой области. – 2025. – Т. 2. – № 62. – С. 15-27.

4. Кайченов, А.В. Разработка и исследование цифрового двойника лабораторного стенда "автоматическая система регулирования частоты вращения двигателя постоянного тока" / А. В. Кайченов, М. О. Емашев, А. А. Яроцкая // Вестник Тамбовского государственного технического университета. – 2025. – Т. 31. – № 2. – С. 235-244.

5. Глазкова И.А. Алгоритмы оценки эффективности функционирования информационной системы и ее элементов на основе принципа декомпозиции / И. А. Глазкова, М. А. Ивановский, С.С. Куприянова // Вестник Тамбовского государственного технического университета. – 2024. – Т. 30. – № 1. – С. 76-84.

Моделирование расписания занятий на основе генетического алгоритма в системе высшего образования

М. Т. Тухтасинов, e-mail: mumtozali@yahoo.com[1]

Наманганский государственный технический университет, Узбекистан

Аннотация. *В данной статье рассматривается вопрос моделирования автоматического формирования расписания занятий в системе высшего образования. В качестве решения проблемы предлагается использование технологий искусственного интеллекта, в частности, генетического алгоритма. В ходе исследования анализируются основные вопросы процесса формирования расписания занятий – нормальное распределение ресурсов, таких как преподаватели, аудитории и группы, во времени с учетом имеющихся ограничений и критериев.*

Ключевые слова: *расписание занятий, оптимизация, искусственный интеллект, генетический алгоритм, автоматизация, ограничения.*

Введение

Формирование расписания занятий представляет собой сложную комбинаторную задачу оптимизации, относящуюся к классу NP-полных задач. При составлении расписания необходимо учитывать множество критериев и ограничений: доступность аудиторий, загруженность преподавателей, пожелания студентов, технические требования к помещениям и другие факторы.

Традиционно создание расписания занятий выполняется вручную сотрудниками учебных заведений, что требует значительных временных затрат и подвержено человеческим ошибкам. С увеличением числа студентов, преподавателей и учебных дисциплин сложность задачи возрастает экспоненциально, что делает ручное составление расписания крайне трудоемким процессом. В связи с этим автоматизация процесса формирования расписания занятий является актуальной проблемой в современной системе высшего образования.

Для решения данной проблемы в последние десятилетия активно применяются методы искусственного интеллекта и эволюционные алгоритмы [1-5]. Генетические алгоритмы показали высокую эффективность при решении комбинаторных задач оптимизации

© Тухтасинов М.Т., 2026

благодаря способности находить квазиоптимальные решения в приемлемые сроки и избегать застревания в локальных оптимумах.

В данной статье рассматривается применение генетического алгоритма как эффективного метода решения задачи автоматического составления расписания с учетом всех необходимых ограничений и критериев оптимизации.

Математическая формализация задачи и основные определения

В рамках задачи введем следующие математические определения:

$G = \{g_1, g_2, ..., g_n\}$ — множество учебных групп;

$T = \{t_1, t_2, ..., t_n\}$ — множество преподавателей;

$R = \{r_1, r_2, ..., r_n\}$ — множество аудиторий;

$S = \{s_1, s_2, ..., s_n\}$ — множество учебных дисциплин;

$D = \{d_1, d_2, ..., d_n\}$ — множество учебных дней;

$H = \{h_1, h_2, ..., h_n\}$ — множество учебных часов (временных слотов);

$P = D \times H$ — множество временных интервалов (пар).

Каждое занятие (урок) представляется кортежем из пяти элементов:

$$L = \left(g_i, t_j, r_k, s_p\left(d_q, h_l\right)\right),$$

где $g_i \in G$ – учебная группа; $t_j \in T$ – преподаватель; $r_k \in R$ – аудитория; $s_p \in S$ – учебная дисциплина; $\left(d_q, h_l\right) \in P$ — временной интервал (день и час).

Система ограничений

Для формирования качественного расписания занятий необходимо учитывать два типа ограничений: жесткие и мягкие ограничения.

Жесткие ограничения. Жесткие ограничения являются обязательными требованиями, которые должны быть безусловно выполнены. Нарушение хотя бы одного жесткого ограничения делает расписание недопустимым. Введем следующие три основных жестких ограничения:

H_1: Отсутствие конфликтов для групп

Одна учебная группа не может одновременно присутствовать на двух и более занятиях:

$$\sum_{i=1}^{N} \delta\left(g_i = g\right) \cdot \delta\left(p_i = p\right) \le 1, \forall g \in G, \forall p \in P .$$

H_2: Отсутствие конфликтов для преподавателей

Один преподаватель не может одновременно вести два и более занятия:

$$\sum_{i=1}^{N} \delta\left(t_i = t\right) \cdot \delta\left(p_i = p\right) \le 1, \forall t \in T,\ \forall p \in P\ .$$

H_3: Отсутствие конфликтов для аудиторий

Одна аудитория не может одновременно использоваться для двух и более занятий:

$$\sum_{i=1}^{N} \delta\left(r_i = r\right) \cdot \delta\left(p_i = p\right) \le 1, \forall r \in R,\ \forall p \in P\ ,$$

где N – общее количество занятий в расписании; $\delta\left(x = y\right)$ — функция Кронекера, равная 1 при выполнении условия $x = y$ и 0 в противном случае; g_i, t_i, r_i, p_i — соответственно группа, преподаватель, аудитория и временной интервал для i -го занятия.

Мягкие ограничения. Мягкие ограничения представляют собой желательные, но необязательные требования к расписанию. Их нарушение не делает расписание недопустимым, но снижает его качество. Введем следующие три мягких ограничения:

S_1: Предпочтительное размещение лекций в утренние часы

Лекционные занятия рекомендуется проводить в первой половине дня:

$$s_1 = \sum_{i \in L} w_i \cdot h_i\ ,$$

где L – множество лекционных занятий; w_i – весовой коэффициент, отражающий предпочтительность утренних часов (чем раньше час, тем меньше штраф); h_i — номер временного слота (чем больше номер, тем позднее время).

S_2: Последовательность лекций и практических занятий

Практические и лабораторные занятия желательно проводить после соответствующих лекций:

$$S_2 = \sum_{d \in D} \sum_{g \in G} \alpha \cdot \upsilon_{d,g}\ ,$$

где α – коэффициент штрафа за нарушение последовательности; $\upsilon_{d,g}$ — количество случаев, когда практическое занятие по дисциплине проводится раньше лекции для группы g в день d; суммирование производится по всем дням и группам.

S_3: Сбалансированность нагрузки преподавателей

Распределение учебной нагрузки между преподавателями должно быть максимально равномерным:

$$S_3 = \sum_{j=1}^{|T|} \left(L_j - \bar{L} \right)^2 ,$$

где L_j — учебная нагрузка (количество часов) преподавателя t_j; $\bar{L} = \frac{1}{|T|} \sum_{j=1}^{|T|} L_j$ – средняя нагрузка по всем преподавателям; $|T|$ — общее количество преподавателей.

Целевая функция оптимизации. Общая функция качества расписания представляет собой взвешенную сумму мягких ограничений:

$$F = \beta_1 \cdot S_1 + \beta_2 \cdot S_2 + \beta_3 \cdot S_3 \to \min ,$$

где $\beta_1, \beta_2, \beta_3$ —весовые коэффициенты, определяющие относительную важность каждого критерия.

Архитектура генетического алгоритма

В генетическом алгоритме основной единицей представления решения является хромосома. В контексте задачи составления расписания хромосома представляет собой полный вариант расписания занятий. Каждая хромосома содержит исчерпывающее описание всех занятий с указанием всех необходимых параметров.

Хромосома определяется следующим образом:

$$C = \{ L_1, L_2, ..., L_N \} ,$$

где N – общее количество занятий в расписании; L_i — i -е занятие.

Каждое занятие L_i представляет собой кортеж из пяти базовых компонентов:

$$L_i = \begin{pmatrix} group_id_i, teacher_id_i, room_id_i, \\ subject_id_i, time_slot_i \end{pmatrix}$$

где $group_id_i$ — идентификатор учебной группы, для которой проводится занятие; $teacher_id_i$ — идентификатор преподавателя, ведущего занятие; $room_id_i$ — идентификатор аудитории, в которой проводится занятие; $subject_id_i$ — идентификатор учебной дисциплины; $time_slot_i$ — временной интервал проведения занятия (комбинация дня недели и учебного часа).

Таким образом, хромосома С полностью определяет конкретное расписание занятий, указывая для каждого занятия все необходимые параметры: кто, что, где и когда проводит.

Расчет функции приспособленности (фитнес-функции)

Функция приспособленности (fitness function) – это математическая функция, которая оценивает качество каждой хромосомы (варианта расписания) в генетическом алгоритме. Она определяет степень пригодности решения, учитывая как жесткие, так и мягкие ограничения.

Функция приспособленности определяется следующим образом:

$$Fitnes(C) = F_0 - \left(w_1 \cdot P_1 + w_2 \cdot P_2 + w_3 \cdot P_3 + w_4 \cdot P_4 + w_5 \cdot P_5\right)$$

где F_0 – базовое (начальное) значение приспособленности; P_1 – количество конфликтов (нарушений жестких ограничений): -конфликты по группам (одна группа в нескольких местах одновременно), -конфликты по преподавателям (один преподаватель ведет несколько занятий одновременно), -конфликты по аудиториям (одна аудитория используется несколькими группами одновременно); P_2 – количество «окон» в расписании (незанятые временные промежутки между занятиями для групп и преподавателей); P_3 – количество занятий, проводимых в нежелательное время (поздние вечерние часы, последние пары); P_4 – степень дисбаланса учебной нагрузки между преподавателями (отклонение от равномерного распределения); P_5 — количество нарушений последовательности занятий (случаи, когда практические занятия проводятся раньше соответствующих лекций); w_1, w_2, w_3, w_4, w_5 — весовые коэффициенты, определяющие относительную важность каждого критерия штрафования.

Принцип работы: чем выше значение функции Fitnes(C), тем лучше качество расписания. Идеальное расписание имеет минимальное количество штрафов, что приводит к максимальному значению фитнес-функции. Генетический алгоритм стремится максимизировать эту функцию в процессе эволюции популяции решений.

Подсчет конфликтов

Конфликты возникают, когда один и тот же ресурс (группа, преподаватель или аудитория) используется одновременно для нескольких занятий. Общее количество конфликтов рассчитывается по следующей формуле:

$$P_1 = \sum_{p \in P} \begin{pmatrix} \max\left(0, n_g(p) - 1\right) + \max\left(0, n_t(p) - 1\right) + \\ + \max\left(0, n_r(p) - 1\right) \end{pmatrix}$$

где $n_g(p)$ — количество групп, одновременно занятых во временном интервале p; $n_t(p)$ – количество занятий, которые одновременно ведет один преподаватель в интервале p; $n_r(p)$ – количество занятий, одновременно проводимых в одной аудитории в интервале p; P – множество всех временных интервалов.

Каждое превышение единицы означает наличие конфликта соответствующего типа.

Балансировка учебной нагрузки

Дисбаланс учебной нагрузки между преподавателями оценивается через дисперсию распределения нагрузки:

$$P_4 = \sum_{j=1}^{|T|} \left(L_j - \bar{L}\right)^2,$$

где L_j — количество учебных часов (занятий), назначенных преподавателю j; $\bar{L} = \frac{1}{|T|} \sum_{j=1}^{|T|} L_j$ — средняя нагрузка на одного преподавателя; $|T|$ – общее количество преподавателей.

Чем меньше значение P_4, тем более равномерно распределена нагрузка между преподавателями.

Штрафы за временные ограничения

Штраф за «окна» в расписании. «Окна» представляют собой свободные временные интервалы между занятиями в течение учебного дня. Штраф за наличие окон рассчитывается следующим образом:

$$P_2 = \sum_{d \in D} \sum_{g \in G} W_{d,g} + \sum_{d \in D} \sum_{t \in T} W_{d,t},$$

где $W_{d,g}$ — количество окон в расписании группы g в день d; $W_{d,t}$ — количество окон в расписании преподавателя t в день d; суммирование производится по всем дням, группам и преподавателям.

Штраф за занятия в нежелательное время:

$$P_3 = \sum_{i=1}^{N} 1\left(h_i \in H_{late}\right),$$

где $1\left(h_i \in H_{late}\right)$ – индикаторная функция, равная 1, если занятие i проводится в поздние часы (множество H_{late}), и 0 в противном случае; H_{late} – множество нежелательных временных слотов (например, последние пары дня).

Генетические операторы

Генетические операторы – это математические операции, реализующие эволюционный процесс в генетическом алгоритме. Они представляют собой компьютерные модели естественных генетических процессов. Рассмотрим применение основных генетических операторов для решения задачи составления расписания занятий.

Формирование начальной популяции. Начальная популяция – это первое поколение хромосом, создаваемое на начальном этапе работы генетического алгоритма. Она является основой всего эволюционного процесса.

Обычно начальная популяция генерируется случайным образом. Однако в нашем подходе мы формируем начальную популяцию более целенаправленно: занятия распределяются по доступным временным слотам таким образом, чтобы избежать явных конфликтов (одновременного использования одного ресурса). При этом в начальных решениях могут присутствовать другие недостатки, такие как «окна» в расписании, занятия в нежелательное время и неравномерное распределение нагрузки.

Начальная популяция определяется следующим образом:

$$Population_0 = \left\{C_1, C_2, ..., C_{pop_size}\right\},$$

где каждая хромосома представляет собой:

$$C_k = \left\{L_1^k, L_2^k, ..., L_N^k\right\}, k = 1, 2, ..., pop_size,$$

где pop_size — размер популяции (количество особей в поколении); N – общее количество занятий в расписании; L_i^k – i-е занятие в k-й хромосоме.

Оператор селекции (отбора). Цель оператора селекции – отобрать наиболее приспособленные особи для формирования следующего поколения и участия в скрещивании. Существуют различные методы селекции. В данной работе используется метод турнирной селекции.

Турнирная селекция:

$$C_{selected} = \arg \max_{C_i = Tournament} Fitnes\left(C_i\right),$$

где $Tournament = \left\{C_{r_1}, C_{r_2}, ..., C_{r_k}\right\}$ — множество из k хромосом, случайно выбранных из текущей популяции; $C_{selected}$ – победитель турнира (хромосома с максимальным значением функции приспособленности среди участников турнира); k – размер турнира (обычно k=2 или k=3).

Процесс повторяется до тех пор, пока не будет отобрано необходимое количество родительских особей для создания нового поколения.

Оператор кроссовера (скрещивания). Оператор кроссовера объединяет генетический материал двух родительских хромосом для создания одной или нескольких дочерних хромосом. В задаче составления расписания кроссовер играет важную роль в комбинировании успешных распределений занятий из разных решений.

В данной работе применяется метод равномерного кроссовера (uniform crossover):

$$L_i^{child} = \begin{cases} L_i^{parent_1}, \text{ если } rand() < 0.5 \\ L_i^{parent_2}, \text{иначе} \end{cases}$$

где L_i^{child} — i-е занятие в дочерней хромосоме; $L_i^{parent_1}, L_i^{parent_2}$ — соответствующие занятия в первой и второй родительских хромосомах; rand( ) – функция генерации случайного числа в диапазоне [0,1].

Таким образом, каждый ген (занятие) дочерней хромосомы выбирается случайным образом от одного из родителей с равной вероятностью.

Оператор мутации. Мутация – это процесс случайного изменения генов в хромосоме. В контексте задачи составления расписания мутация позволяет исследовать новые варианты решений путем изменения параметров отдельных занятий.

Процесс мутации. Для каждого занятия L_i в хромосоме с вероятностью p_{mut}:

$$L_i^{mutated} = \begin{pmatrix} group_id_i, teacher_id_i, room_id_i, \\ subject_id_i, time_slot_{new} \end{pmatrix}$$

где $time_slot_{new}$ – новый случайно выбранный временной интервал; p_{mut} – вероятность мутации (обычно $p_{mut} = 0.01 - 0.1$).

Также могут применяться другие типы мутаций: изменение аудитории: $room_id_i \to room_id_{new}$; перестановка двух занятий: обмен временными интервалами между двумя случайно выбранными занятиями.

Правильная настройка параметров мутации (вероятность и тип) является одним из ключевых факторов, определяющих эффективность генетического алгоритма при оптимизации расписания занятий.

Заключение

Разработанное решение на основе генетического алгоритма обладает следующими преимуществами:

1. Многокритериальная оптимизация – возможность одновременного учета множества условий и ограничений различной природы;
2. Гибкость и адаптивность – быстрая адаптация к изменяющимся требованиям и различным конфигурациям ограничений образовательного процесса;
3. Поиск глобально оптимального решения – способность избегать застревания в локальных оптимумах благодаря механизмам селекции, кроссовера и мутации.

Результаты исследования демонстрируют, что генетические алгоритмы позволяют эффективно решать задачу автоматического формирования расписания занятий в системе высшего образования. Предложенный подход обеспечивает значительное сокращение времени на составление расписания и повышение его качества по сравнению с ручным методом.

В настоящее время авторским коллективом проводится практическая апробация программного комплекса, разработанного на основе представленных методов и алгоритмов. Параллельно продолжаются научные исследования, направленные на дальнейшее совершенствование системы, включая интеграцию дополнительных критериев оптимизации и повышение производительности алгоритма.

Список литературы

1. Burke, E. K., Petrovic, S. Recent research directions in automated timetabling. European Journal of Operational Research. Vol. 140, issue 2, 2002, p. 266-280.
2. Goldberg, D.E. (1989). Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley. 1989.
3. Астахова, И.Ф., Фирас, А.М. Составление расписания учебных занятий на основе генетического алгоритма // Вестник ВГУ. Серия: Системный анализ и информационные технологии. – 2013. – № 2. – С. 93-99.
4. Конькова, И.С. Использование генетического алгоритма в задаче оптимизации расписания вуза // Вестник ТвГТУ (Тверского

государственного технического университета). – 2012. – Вып. 22. – С. 26-31.

5. To‘xtasinov, M.T., Sadritdinov, N.X. O'zbekiston va rivojlangan davlatlar ta'lim tizimida avtomatlashtirilgan dars jadvalini tuzish tajribalari. “Al-Farg‘oniy avlodlari” elektron ilmiy jurnali, ISSN 2181-4252, Tom: 1, Son: 2, 2025-yil, 57-60 b.

Секция 7
Суперкомпьютерные вычисления и BigData. Квантовые информационные системы

Квантовые алгоритмы поиска в машинном обучении: проблемы, методы и технологии

Р. М. Бозоров, email: ruslan.bozorov2013@ya.ru
А. Ф. Клинских, email: 2000afk@gmail.com

Воронежский государственный университет

Аннотация. *Квантовое машинное обучение (QML) представляет собой одну из наиболее перспективных областей на стыке квантовых вычислений и искусственного интеллекта, предлагая теоретическое ускорение для ряда вычислительных задач. Несмотря на значительный исследовательский интерес, практическая реализация QML сталкивается с комплексными проблемами, связанными как с аппаратной нестабильностью квантовых систем, так и с алгоритмическими и теоретическими ограничениями. В данной работе проводится анализ ключевых проблем в области QML, таких как проблема ввода классических данных, барьеры обучаемости и влияние квантового шума. В качестве одного из методов преодоления этих проблем рассматривается адаптация алгоритма квантового поиска Гровера для задач машинного обучения, обеспечивающая квадратичное ускорение для неструктурированного поиска. Исследуются технологические аспекты, включая разработку устойчивых кубитов и гибридные квантово-классические архитектуры. Работа структурирована для представления на конференции «Информатика: проблемы, методы, технологии» и призвана дать системный обзор текущего состояния и перспектив интеграции квантовых алгоритмов поиска в области ИИ.*

Ключевые слова: *квантовое машинное обучение, алгоритм Гровера, барьерное плато, квантовый шум, квантовая декогеренция.*

Введение

За последнее десятилетие квантовые вычисления превратились из сугубо теоретической области в активно развивающуюся технологию с растущим числом практических демонстраций и алгоритмических предложений. Основная гипотеза заключается в том, что квантовые компьютеры, используя такие явления, как суперпозиция и запутанность, смогут решать определенные классы задач экспоненциально быстрее классических систем. Квантовое машинное обучение стремится

© Бозоров Р. М., Клинских А. Ф., 2026

использовать это преимущество для ускорения задач анализа данных, оптимизации и распознавания образов.

Однако путь от теоретических алгоритмов к их практической реализации сопряжен с фундаментальными трудностями. Проблема декогеренции и шума делает квантовые состояния недолговечными, а сама квантовая информация не может быть скопирована, что усложняет коррекцию ошибок. Кроме того, многие перспективные алгоритмы, такие как алгоритм Шора, для своей работы требуют миллионов стабильных кубитов, что выходит за рамки возможностей современных NISQ-устройств (Noisy Intermediate-Scale Quantum devices).

В этом контексте задача эффективного поиска и оптимизации в больших пространствах данных, характерная для ИИ, становится естественной областью применения квантовых алгоритмов. Алгоритм Гровера, обеспечивающий квадратичное ускорение для задач неструктурированного поиска, представляет особый интерес как относительно простой для понимания и реализации, но мощный инструмент, который может быть использован в качестве подпрограммы в более сложных QML-пайплайнах [1].

1. Проблемы квантового машинного обучения

Переход от классического машинного обучения к квантовому сопряжен с рядом фундаментальных проблем, которые можно разделить на три основные категории.

Классические данные, представленные битовыми строками, должны быть эффективно загружены в квантовый процессор. Этот этап, известный как квантовое кодирование данных (state preparation), сам по себе может стать вычислительным узким местом, сводя на нет потенциальное ускорение от квантового алгоритма. Требуется разработка эффективных схем кодирования, минимизирующих глубину схемы и число необходимых кубитов.

Квантовые нейронные сети (QNN) и другие параметризованные квантовые схемы страдают от так называемого «исчезающего градиента» или «барьерного плато» (barren plateau). При увеличении числа кубит градиенты функции стоимости экспоненциально затухают, делая процесс обучения невозможным. Эта проблема усугубляется присутствием шума в реальных устройствах.

Кубиты чрезвычайно хрупки и подвержены влиянию внешнего шума, что приводит к декогеренции и ошибкам. Современные подходы к коррекции ошибок требуют огромных накладных расходов, когда один логический (стабильный) кубит создается из сотен или тысяч физических. Альтернативные пути, такие как создание топологически защищенных кубитов, находятся на ранней стадии исследования, как,

например, эксперименты с топологическими фазами на сверхпроводниковых процессорах. Кроме того, неспособность к копированию квантовых состояний (теорема о запрете клонирования) накладывает дополнительные ограничения на проектирование алгоритмов.

Основные проблемы квантового машинного обучения представлены в таблице.

Таблица

Проблемы квантового машинного обучения

Категория проблемы	**Конкретная проблема**	**Текущие исследовательские подходы**
Обучаемость	Барьерное плато (Barren Plateaus)	Специальные инициализации, структурированные анзатцы, предварительное обучение
Данные	Эффективное кодирование классических данных	Амплитудное кодирование, кодирование углов поворота, квантовые GAN
Аппаратура	Шум и декогеренция кубитов	Квантовая коррекция ошибок (QEC), топологическая защита, вариационные квантовые алгоритмы (VQE/QAOA) для NISQ
Алгоритмы	Преодоление квадратичного ускорения	Интеграция с классическими методами, гибридные алгоритмы (напр., квантовое усиление)

2. Метод: Алгоритм Гровера в контексте машинного обучения

Алгоритм Гровера — это квантовый алгоритм для неструктурированного поиска, который находит помеченный элемент в базе данных из N элементов за примерно $\sqrt{N}$ операций, обеспечивая квадратичное ускорение по сравнению с классическим алгоритмом, требующим в среднем $N/2$ операций.

Алгоритм работает в итеративном режиме, каждая итерация состоит из двух ключевых шагов:

Применение оракула (Uω): Эта квантовая схема «помечает» искомое состояние, инвертируя его фазу (умножая амплитуду на -1), оставляя остальные состояния неизменными [2].

Применение диффузионного оператора (Us): Этот оператор инвертирует все амплитуды относительно их среднего значения, что приводит к увеличению амплитуды помеченного состояния и уменьшению амплитуд остальных.

Геометрически это можно представить как последовательность отражений в двумерном подпространстве, что приближает состояние системы к искомому.

Алгоритм Гровера может быть использован не как самостоятельный метод, а как мощная подпрограмма в рамках более крупных QML-алгоритмов:

Ускорение поиска решений: в задачах оптимизации, таких как поиск минимальной энергии молекулы или оптимальных параметров модели, перебор вариантов может быть ускорен.

Усиление классических алгоритмов: техника амплитудного усиления, обобщающая алгоритм Гровера, может квадратично улучшать производительность случайных или вероятностных классических алгоритмов [3].

Квантовые базы данных: прямое применение для поиска в специализированных квантовых базах данных, что может быть актуально, например, в квантовой химии для поиска конфигураций молекул.

3. Технологии и перспективы

В эпоху NISQ-устройств наиболее практичным подходом является гибридное вычисление. Классический компьютер управляет процессом обучения, отправляя на квантовый процессор задачи, которые он может решать эффективно, например, вычисление ожидаемых значений или выполнение алгоритма Гровера в качестве подпрограммы. Вариационные квантовые алгоритмы (VQE, QAOA) являются ярким примером такого подхода [4].

Исследования ведутся не только в области программного обеспечения, но и в создании более стабильных аппаратных платформ:

Топологическая защита: эксперименты, подобные проведенным китайской группой Пань Цзяньвэя, показывают возможность создания квантовых состояний, устойчивых к локальным возмущениям, на основе топологических фаз вещества.

Новые кубиты: разработка кубитов на основе майорановских фермионов или других квазичастиц обещает повысить устойчивость к декогеренции.

Заключение

Квантовое машинное обучение находится на раннем, но стремительно развивающемся этапе. Алгоритм Гровера, обеспечивая надежное квадратичное ускорение для задач поиска, является одним из краеугольных камней в этом направлении и естественным кандидатом для интеграции в QML-пайплайны. Однако для реализации этого потенциала необходимо преодолеть комплексные проблемы, связанные с аппаратной реализацией, кодированием данных и обучаемостью моделей [5].

Перспективы лежат в развитии гибридных квантово-классических методов, которые будут использовать сильные стороны обоих типов вычислений в эпоху NISQ, и в параллельной технологической гонке по созданию более масштабируемых и устойчивых квантовых процессоров. Исследования в области квантового поиска и машинного обучения носят междисциплинарный характер, требуя совместной работы специалистов в области информатики, физики и прикладной математики.

Список литературы

1. Quantum Computing in Artificial Intelligence: a Review of Quantum Machine Learning Algorithms / O. F. Olaitan [et al.] // Path of Sience. – 2025. – № 5 – P. 7001–7009.

2. Grover, L.A fast quantum mechanical algorithm for database search / L. Grover // Proceedings of the twenty-eighth annual ACM symposium on Theory of computing. – 1996. – P. 212–219.

3. Integrating artificial intelligence and quantum computing: A systematic literature review of features and applications. / V. C. Pidena [et al.] // International Journal of Cognitive Computing in Engineering. – 2026.– № 7.– P.26–39.

4. A variational eigenvalue solver on a photonic quantum processor. / A. Peruzzo [et al.] // Nature Communications. – 2014. – № 1. – P. 4213.

5. Power of data in quantum machine learning. / H. Y. Huang [et al.] // Nature Communications. – 2021. – № 1. – P. 2631.

Экспериментальное сравнение квантовых классификаторов для бинарной и мультиклассовой задач

Д. А. Ивонин, email: i.daniil1728@gmail.com
А. Ф. Клинских, email: 2000afk@gmail.com

Воронежский государственный университет

Аннотация. *Исследована эффективность различных квантовых алгоритмов классификации на простых и сложных задачах распознавания образов. Рассмотрены три подхода: квантовая машина опорных векторов (Quantum Support Vector Classifier, QSVC), вариационный квантовый классификатор (VQC) и квантовый вариант алгоритма (k)-ближайших соседей (QkNN).*

Ключевые слова: *квантовый машинный интеллект, квантовая машина опорных векторов, вариационный квантовый классификатор, квантовый (k)-ближайших соседей.*

Введение

Современные методы машинного обучения достигли значительных успехов в задачах классификации, однако дальнейшее улучшение производительности ограничивается вычислительными ресурсами классических систем. Квантовые вычисления представляют собой перспективное направление, способное предложить новые способы обработки информации за счёт параллелизма и принципов квантовой механики. Особенно активно развивается область квантового машинного обучения (QML), объединяющая квантовые алгоритмы и задачи интеллектуального анализа данных.

Одной из ключевых задач QML является классификация – определение принадлежности объекта к одному из заранее определённых классов на основе признаков. В условиях квантовых вычислений данная задача решается с использованием специализированных алгоритмов, таких как квантовая машина опорных векторов (QSVM), вариационный квантовый классификатор (VQC) и квантовый аналог метода k-ближайших соседей (QkNN). Каждый из этих подходов использует разные принципы кодирования данных, построения квантовых схем и стратегий обучения, что делает актуальным их сравнительный анализ.

В данной работе проводится экспериментальное сравнение указанных алгоритмов на примерах бинарной и многоклассовой классификации. В качестве наборов данных использованы Iris и модифи-

© Ивонин Д. А., Клинских А. Ф., 2026

цированный MNIST-PCA, подходящие для исследования эффективности квантовых классификаторов в условиях ограниченного числа кубитов. Целью исследования является выявление преимуществ и ограничений каждого подхода, а также оценка их применимости в условиях современных квантовых устройств [1–3].

1. Описание наборов данных

Для проведения экспериментов были выбраны два широко используемых открытых набора данных, предназначенных для решения задач бинарной и многоклассовой классификации.

1. MNIST-PCA (бинарная классификация). Модифицированная выборка из оригинального набора рукописных цифр MNIST была отфильтрована таким образом, чтобы содержать только изображения цифр 3 и 5. Каждое изображение размером 28x28 пикселей было предварительно преобразовано методом главных компонент (PCA) до 5 признаков, нормированных в диапазоне $[0, \pi]$, что делает их пригодными для углового кодирования в квантовых схемах. Итоговый датасет содержит 11 552 обучающих и 1 902 тестовых примера.

2. Iris (многоклассовая классификация). Набор данных содержит 150 объектов, каждый из которых описывается четырьмя числовыми признаками: длина и ширина чашелистика, длина и ширина лепестка. Объекты относятся к одному из трёх классов цветков ириса: Iris setosa, Iris versicolor и Iris virginica. Благодаря своей компактности и линейно разделимой структуре, данный набор широко используется в тестировании алгоритмов квантового машинного обучения.

Оба набора были масштабированы и приведены к виду, совместимому с квантовыми схемами, реализованными в среде Qiskit.

2. Методы квантовой классификации

В данной работе рассмотрены три квантовых алгоритма классификации, реализованных с использованием библиотеки Qiskit: квантовая машина опорных векторов (QSVM), вариационный квантовый классификатор (VQC) и квантовый метод k-ближайших соседей (QkNN). Каждый из них использует собственный способ кодирования признаков и архитектуру квантовой схемы.

1. QSVM (Quantum Support Vector Machine).

QSVM использует квантовую карту признаков $\phi(x)$, преобразующую вектор признаков x в квантовое состояние $|\phi(x)\rangle$. Вычисляется матрица Керна:

$$K(x_i, x_j) = \left| \left\langle \phi(x_i) \middle| \phi(x_j) \right\rangle^2 \right\rangle, \quad (1)$$

где K служит аналогом ядра в классическом SVM. В работе использована схема ZZFeatureMap, реализующая запутывание кубитов через оператор.

$$U_{ZZ}^{(i,j)} = \exp(-i\gamma Z_i Z_j), \quad (2)$$

где Z - оператор Паули-Z, а γ - параметр взаимодействия.

2. VQC (Variational Quantum Classifier).

В VQC входной вектор $x \in \mathbb{R}^n$ кодируется в состояние квантовой системы с помощью вращений:

$$R_y(x_i) = \exp(-i\frac{x_i}{2}Y), i = 1,...,n, \quad (3)$$

что реализует angle encoding. После кодирования к схеме применяется параметризованная вариационная схема $U(\vec{\theta})$, где $\vec{\theta}$ – вектор обучаемых параметров. Классификация осуществляется на основе среднего значения наблюдаемой величины:

$$f(x, \vec{\theta}) = \langle 0 | U^\dagger(\vec{\theta}) Z U(\vec{\theta}) | 0 \rangle \quad (4)$$

Параметры $\vec{\theta}$ оптимизируются с помощью классических градиентных методов.

3. QkNN (Quantum k-Nearest Neighbors).

Метод QkNN реализует идею квантового сравнения состояния $|x\rangle$ с состояниями обучающей выборки $|x_i\rangle$. Вычисляется квантовая метрика близости:

$$d(x, x_i) = 1 - \left| \langle x | x_i \rangle \right|^2, \quad (5)$$

После чего выбираются k ближайших элементов. Для реализации данной процедуры может использоваться квантовая амплитудная амплификация, что в перспективе позволяет достичь квадратичного ускорения по сравнению с классическим kNN [4–6].

3. Экспериментальная постановка

Эксперимент был проведён с использованием симулятора квантовых вычислений Qiskit Aer, обеспечивающего имитацию работы квантовых схем в условиях отсутствия шума. Реализация всех трёх

классификаторов – QSVM, VQC и QkNN – выполнена с использованием библиотеки Qiskit Machine Learning. Исследование направлено на оценку качества классификации в двух задачах: бинарной (модифицированный датасет MNIST-PCA, классы 3 и 5) и многоклассовой (датасет Iris).

Подготовка данных:

Перед подачей на вход квантовых схем все признаки были нормированы в диапазон $[0, \pi]$, обеспечивая совместимость с угловым кодированием (angle encoding), реализуемым через гейты $R_y(x_i)$.

Архитектура схем:

– QSVM: используется схема ZZFeatureMap(depth=2) с двухкубитными гейтами взаимодействия. Измеряется квантовая матрица Керна $K(x_i, x_j)$, затем применяется классический SVM с линейным ядром.

– VQC: реализована трёхслойная вариационная схема с гейтами и CNOT между соседними кубитами. Оптимизация параметров осуществляется методом COBYLA.

– QkNN: для каждого тестового примера рассчитываются оверлапы с квантовыми представлениями обучающей выборки. Классификация происходит по принципу большинства среди $k = 3$ ближайших.

4. Результаты

Сравнительная точность моделей приведена на рис. 1 и рис. 2. Метрики моделей представлены в табл. 1, 2.

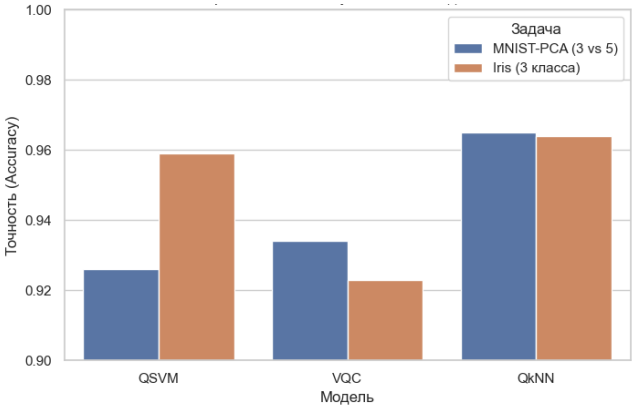


Рис. 1. Точность классификации для QSVM, VQC и QkNN на бинарной и многоклассовой задачах

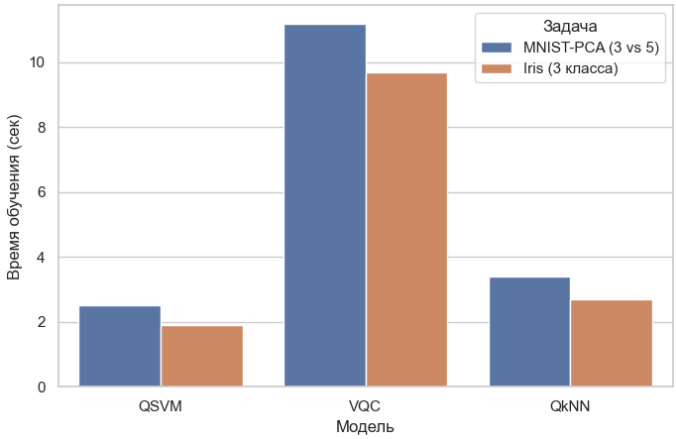


Рис. 2. Время обучения для QSVM, VQC и QkNN на бинарной и многоклассовой задачах

Таблица 1

Метрики моделей для задачи бинарной классификации

Модель	Precision (среднее по классам)	Recall (среднее по классам)	F1-score (среднее по классам)
QSVM	0.925	0.927	0.926
VQC	0.931	0.936	0.933
QkNN	0.967	0.962	0.964

Таблица 2

Метрики моделей для задачи мультиклассовой классификации

Модель	Precision (среднее по классам)	Recall (среднее по классам)	F1-score (среднее по классам)
QSVM	0.960	0.957	0.958
VQC	0.920	0.924	0.922
QkNN	0.965	0.963	0.964

QkNN показал наивысшую точность в обеих задачах. QSVM стабильно работает на малых данных, особенно при чёткой границе классов. VQC демонстрирует хорошую гибкость, но требует тщательной настройки параметров.

Angle encoding показал хорошую применимость для маломерных данных.

ZZFeatureMap эффективно работает в QSVM за счёт двухкубитных взаимодействий.

QkNN оказался наиболее устойчивым и эффективным среди сравниваемых алгоритмов.

Заключение

В данной работе было проведено экспериментальное сравнение трёх квантовых алгоритмов классификации – QSVM, VQC и QkNN – на задачах бинарной и многоклассовой классификации. Анализ показал, что каждый из методов обладает своими сильными и слабыми сторонами в зависимости от характера данных и архитектуры схемы.

QkNN продемонстрировал наилучшие результаты по всем основным метрикам (accuracy, precision, recall, F1-score) как в бинарной, так и в многоклассовой задачах. Его производительность стабильна, а структура алгоритма проста в реализации. Это делает QkNN привлекательным выбором для задач с малым числом признаков и ограниченным числом кубитов.

QSVM показал высокую точность и хорошую устойчивость, особенно в многоклассовой классификации. Использование квантовой карты признаков (ZZFeatureMap) позволяет эффективно извлекать взаимосвязи между признаками. Однако данный подход требует вычисления матрицы Керна, что увеличивает время исполнения при большом объёме данных.

VQC оказался наиболее чувствительным к выбору архитектуры вариационной схемы и параметров оптимизации. Несмотря на это, он обеспечивает гибкость и потенциально может обобщать сложные зависимости. Однако в проведённых экспериментах он уступил остальным моделям по точности и стабильности.

С практической точки зрения, QkNN может быть рекомендован для применения в квантовых классификаторах первого поколения (NISQ-устройства), благодаря своей простоте и высокой точности. QSVM подходит для задач с выраженными границами между классами, а VQC требует более глубокой настройки и может быть эффективен при большем количестве данных и кубитов.

Список литературы

1. Quantum Computing in Artificial Intelligence: a Review of Quantum Machine Learning Algorithms / O. F. Olaitan [et al.] // Path of Sience. – 2025. – № 5 – P. 7001–7009.

2. Quantum machine learning: Classifications, challenges, and solutions / A. Sigov [et al.] // Journal of Industrial Information Integration. – 2024. – № 42.– P. 1–12.

3. Integrating artificial intelligence and quantum computing: A systematic literature review of features and applications. / V. C. Pidena [et al.] // International Journal of Cognitive Computing in Engineering. – 2026.– № 7. – P. 26–39.

4. New directions for artificial intelligence: human, machine, biological, and quantum intelligence. / L. M. Enamoto [et al.] // Frontiers of Information Technology & Electronic Engineering. – 2021. – № 23. – P. 984–990.

5. Mohammadisavadkoohi, E. A Systematic Review on Quantum Machine Learning Applications in Classification / E. Mohammadisavadkoohi, N. Shafiabady, J. Vakilian // IEEE Transactions on Artificial Intelligence. – 2026. – № 7. – P. 4–21.

6. Schuld, M. Quantum computing for pattern classification / M. Schuld, I. Sinayskiya, F. Petruccione // PRICAI 2014: Trends in Artificial Intelligence. – 2014. – Vol. 8862. – P. 208–220.

Разработка загрузочных образов для СнК

А. С. Коваль, email: koval@cs.vsu.ru

Воронежский государственный университет

Аннотация. *Обсуждается создание загрузочного образа для ПЛИС типа СнК, включающего в себя ядро Linux, корневую файловую систему и битстрим. Рассматриваются вопросы взаимодействия программируемой (PL) и процессорной (PS) частей СнК в задачах построения ускорителей вычислений.*

Ключевые слова: *СнК, ПЛИС, ускорители вычислений, embedded Linux, SoC, FPGA, Yocto, BitBake, Buildroot.*

Введение

Эволюция программируемой логики привела к появлению отдельного класса устройств «ПЛИС типа Система-на-Кристалле (СнК)», в которых программируемые блоки логики сочетаются с аппаратными сложнофункциональными (СФ) блоками: обычно процессорами и периферией. Предтечей им были попытки применения софт-процессоров MicroBlaze, PicoBlaze для высокоуровневого управления – формирования системы в одной микросхеме. Софт-процессоры и сейчас применяются и находятся в библиотеках средств разработки, но не могут обеспечить сопоставимую надежность, в сравнении с аппаратными СФ. Одними из первых аппаратных процессоров в ПЛИС были применены PowerPC 405/440. Одно или два аппаратных ядра обеспечили независимость их работы от состояния программируемой логики, тогда как софт-процессоры, особенно в случае репрограммирования частей микросхемы, могли попасть в неопределённые состояния. На данном этапе несколько поколений ПЛИС компаний AMD/Xilinx и Intel/Altera для аппаратных процессорных ядер используют архитектуры ARM v7, v8 (рис.1).

Помимо высокой надежности, СнК-решения с аппаратными процессорами позволяют задействовать обширные наработки ПО, сокращая время на создание сложных систем обработки данных и управления. И на определенном уровне сложности всё чаще применяют ОС как базовую систему управления вычислениями и вводом-выводом. Вопрос о том, в каких именно проектах следует применять ОС, а в каких процессор будет исполнять единую программу-приложение, – дискуссионный, а решение – компромисс между

© Коваль А. С., 2026

многофункциональностью первого и простотой и предсказуемостью второго. Аналогичен и выбор между универсальной ОС GNU/Linux и системой реального времени, например, FreeRTOS. Мы исследовали варианты сборки образов ОС на ядре Linux для популярного семейства ПЛИС AMD/Xilinx Zynq 7Z020xx и Zynq UltraScale+.

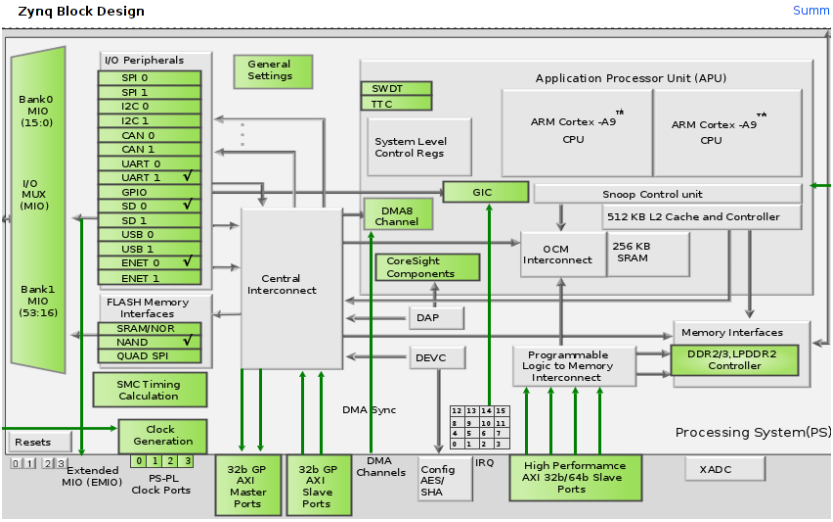


Рис. 1. Процессорная система Zynq в среде разработки Vivado

1. Архитектура СнК Zynq и особенности загрузки ОС Linux

Несмотря на то, что в составе вышеупомянутых микросхем находятся процессоры архитектур ARM (v7 и v8), все-таки требуется инициализация их периферии, самих процессоров, устройств ввода-вывода, загрузки. Также необходим загрузчик первого уровня. Поэтому недостаточно ядра и корневой системы даже только для запуска ОС на процессорах ПЛИС. Кроме того, ПЛИС содержит программируемую логику, для которой формируется конфигурационный битстрим. Все эти компоненты должны быть собраны и переданы в микросхему. Варианты их размещения: NAND, QSPI, eMMC или micro-SD флеш-память. Был выбран вариант micro-SD, как наиболее простой в тестировании, хотя для рабочих проектов удобнее и надежнее NAND, QSPI.

Исследовалось несколько отладочных плат: Alinx AXU2CGB (ПЛИС Zynq Ultrascale+), PYNQ-Z2 (ПЛИС Zynq 7Z020), MYiR Z-turn (ПЛИС Zynq 7Z020) и плата управления от биткоин-майнера Ebit E9 (ПЛИС Zynq 7Z010). На последней сделан акцент, т. к. плата не является отладочной и

ни производитель платы, ни производитель ПЛИС не поставляет для неё готовых сборок или файлов-описателей. Тем не менее, для платы есть электрическая принципиальная схема и достаточная информация. Кроме ПЛИС Zynq 7Z010, на плате есть NAND-память 128MiB, DDR3-память SDRAM 256MiB, microSD слот, сетевой контроллер 10/100Мбит/с компании IC Plus – IP101GA.

Процесс загрузки Zynq начинается с загрузчика нулевого уровня BootROM, код которого выполняется в отдельной внутренней памяти On-Chip Memory (OCM) объемом 256/192КБ. Далее загружается в OCM код загрузчика первого уровня – First State Boot Loader (FSBL). Диаграмма загрузки довольно сложна из-за поддержки т. н. «безопасной» загрузки, с проверкой подлинности образов, шифрованием [1]. В упрощенном небезопасном режиме добавляются традиционный для ПЛИС источник загрузки и отладки: JTAG. Выбор различных источников и режимов загрузки для BootROM выбирается сигналами MIO [8:2]. Сигналы портов MIO подключены к процессорной части ПЛИС и не выделены только под выбор источника – это мультиплексированные порты. Поэтому нельзя, например, задавать постоянный код на MIO [8:2] перемычками на землю, т. к. это приводит к ошибкам передачи данных через задействованные порты. Еще одна особенность (уже самой платы), которую приходится учитывать в проектах, – размещение сетевого PHY-чипа на портах программируемой части микросхемы – EMIO (рис. 1). Эти порты необходимо внутренне маршрутизировать к сетевому контроллеру процессорной части.

2. Сборка загрузочных образов

Сборка загрузочного образа состоит из получения из исходных кодов Linux-ядра, описателя устройств Device Tree Binary, корневой файловой системы, инициализации процессорной части ПЛИС и конфигурации программируемой логики – битстрима.

Минимальная аппаратная часть, описанная в среде проектирования Vivado (рис. 2), содержит процессорный модуль с настройками портов, клоков и других компонентов, сигналы к микросхеме сетевого PHY. В стандартном для ПЛИС применении добавляются модули для программируемой логики, например, нейропроцессор (рис. 3). Итогом работы среды Vivado является файл-описатель оборудования .xsa, включающий в себя и битстрим, и являющийся входным для набора инструментов (tool-chain) сборки ядра и корневой файловой системы [2].

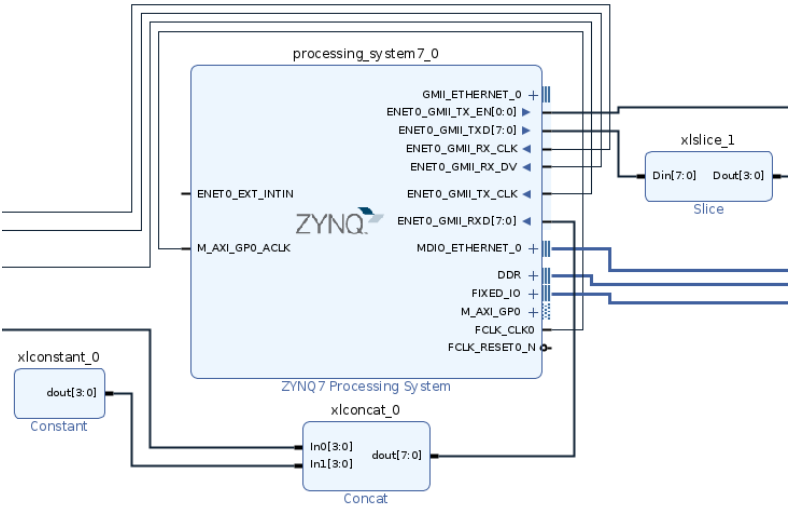


Рис. 2. IP-блоки проекта минимальной системы для загрузки ОС

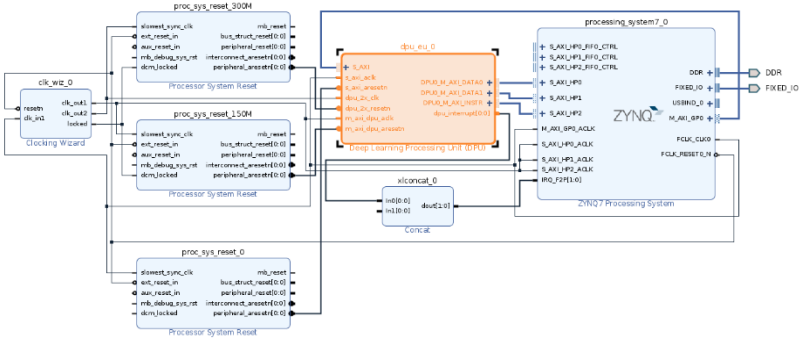


Рис. 3. Проект с нейропроцессорным модулем DPU

Стандартный инструмент сборки среды проектирования построен на основе Petalinux [2], который, в свою очередь, базируется на проекте Linux Foundation с открытым кодом – Yocto Project [3]. На рис. 4 показан конфигуратор Petalinux, в котором задаются основные параметры сборки образа.

Из репозитория AMD/Xilinx и многочисленных зеркал скачиваются и собираются пакеты компонентов ОС в исходных текстах. Офлайн-сборка без постоянного доступа в Интернет или при полном отсутствии доступа на компьютере сборки в Yocto Project требует настройки системы

«сборки рецептов» bitbake и каталога скачиваемых пакетов. Bitbake использует кэш "sstate cache artefacts" (SCA) для определения зависимостей и того, какие именно рецепты нужно скачать и пересобрать.

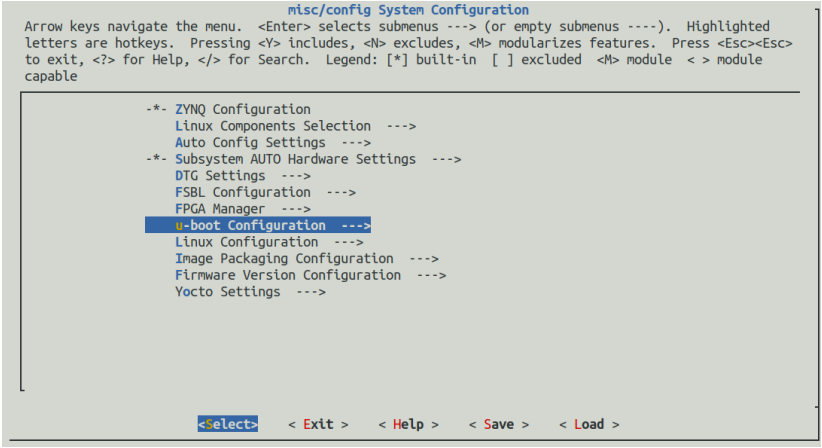


Рис. 4. Конфигуратор Petalinux

Есть по крайней мере два офлайн-варианта [4]:

1. Предварительная загрузка всех архивов пакетов и кэшей sstate, соответствующая версии Petalinux, например, для вер. 2020.2 – файлы sstate_arm_2020.2.tar.gz (9.09 GB), downloads_2020.2.tar.gz. Компания AMD/Xilinx предлагает SCA для трех платформ: ARM, ARM64 и MicroBlaze. Затем – задание локальных путей в конфигураторе petalinux-config.
2. Также можно скачать пакеты с исходными текстами без сборки с помощью команды "bitbake -c fetchall petalinux-image-minimal", что достаточно для последующей сборки проекта.

В зависимости от сценариев разработки, отдельные компоненты, участвующие в загрузке, могут быть неизменны, и их регулярная пересборка нецелесообразна. Например, при наличии уже собранных ядра и корневой системы, в которых не планируются изменения, для загрузки достаточно сформировать файл BOOT.bin, который включает в себя FSBL, Bitstream-файл, U-Boot, и разместить его на FAT-32 разделе microSD-карты [5]. В этом же разделе должны находиться файлы boot.scr (скрипт U-Boot) и образ ОС с ФС – image.ub. Возможно несколько вариантов размещения корневой файловой системы (RAM-диск, EXT4 раздел карты) и ядра. Файл image.ub содержит ядро, RAM-образ корневой ФС и дерево устройств DTB. На стадии сборки выбирается подходящий

вариант. Являясь полноценной сетевой ОС, добавление ПО и даже сборка небольших проектов может проводиться непосредственно на ПЛИС.

Заключение

Загрузочные образы для СнК ПЛИС с ОС Linux позволяют создавать системы сложной обработки данных на «кристалле». Однако сборка образа, состоящего из корневой файловой системы, ядра, битстрима, загрузчика первого уровня и данных инициализации, – длительный процесс с высокой вероятностью дополнительных затрат на исправление ошибок и отладку. Системы автоматизации сборки: Yocto/Petalinux, Buildroot – помогают упорядочить этот процесс, в то же время позволяя регулярно обновлять ОС и её компоненты, приспосабливать системное ПО к конкретным функциям встраиваемой системы.

Список литературы

1. Zynq-7000 All Programmable SoC Software Developers Guide [Электронный ресурс] : документация. – Режим доступа : https://docs.amd.com/v/u/en-US/ug1769-zynq-7000-soc-design-lp
2. PetaLinux Tools Documentation Reference Guide [Электронный ресурс] : документация. – Режим доступа : https://docs.amd.com/r/2020.2-English/ug1144-petalinux-tools-reference-guide
3. Документация Yocto Project [Электронный ресурс] : документация. – Режим доступа : https://wiki.yoctoproject.org
4. AMD Adaptive Computing Wiki (Yocto Builds without an Internet Connection) [Электронный ресурс] : документация ПО. – Режим доступа : https://xilinx-wiki.atlassian.net/wiki
5. Zynq-7000 SoC Technical Reference Manual [Электронный ресурс] : документация. – Режим доступа : https://docs.amd.com/r/en-US/ug585-zynq-7000-SoC-TRM/Zynq-7000-SoC-Technical-Reference-Manual

Квантовые и параллельные вычисления: креативный подход

С. Д. Кургалин, e-mail: kurgalin@bk.ru
С. В. Борзунов, e-mail: borzunov@cs.vsu.ru

ФГБОУ ВО «Воронежский государственный университет»,

Аннотация. *Решение актуальных научно-технических задач требует использования как высокопроизводительных вычислений, так и, в ближайшей перспективе, квантовых вычислений. Подготовка специалистов в области информационных технологий должна осуществляться на высоком научном и методическом уровне, следует широко применять необходимые программно-аппаратные средства и специализированные учебно-методические комплексы, а также обеспечить формирование креативной образовательной среды.*

Ключевые слова: *информационные технологии, параллельные вычисления, квантовые вычисления, высшее образование, креативная образовательная среда.*

Введение

Решение современных задач в области моделирования физических и химических процессов, предсказание свойств новых материалов, проведение экономических расчетов и подготовка прогнозов базируется на высокотехнологичной вычислительной базе. Многие десятилетия парадигма вычислений была ограничена однопроцессорными устройствами, или, как еще говорят, машинами фон Неймана. Неуклонное и быстрое развитие вычислительной мощности, связанное с количественным увеличением характеристик такого рода систем, в частности, частоты процессора, объема доступной оперативной памяти и пропускной способности шин данных, привело к появлению достаточно совершенных вычислительных устройств. Тем не менее, значительное увеличение вычислительной мощности достигается с помощью ресурсов параллелизма, что заключается в совместном применении для решения задачи нескольких одноранговых вычислительных узлов. Таким образом, наиболее мощные компьютеры в настоящее время используют технологии параллельных вычислений, к которым относят также методы программирования графических процессоров, представляющих собой объединение в иерархическую структуру большого количества (порядка 10^3–10^4) индивидуальных вычислительных ядер.

© Кургалин С. Д., Борзунов С. В., 2026

Особенности подготовки специалистов в области информационных технологий

Скорость развития индивидуальных вычислительных ядер, продолжавшегося на протяжении нескольких последних десятилетий, неизбежно снижается в силу достижения физически обусловленных пределов архитектуры системы. Во-первых, размер структурных логических элементов процессора ограничен характерным размером атомов и молекул, поскольку электромагнитные сигналы в операторах инверсии, дизъюнкции, конъюнкции и т. д. формируются на основе электронных динамических процессов. Во-вторых, предел миниатюризации наступает еще даже несколько раньше за счет необходимости диссипации энергии, т. е. отвода тепла от элементов электронной цепи.

Переход к параллельным вычислительным системам позволил направить вектор развития высокопроизводительных вычислений и работы с большими данными таким образом, чтобы физические ограничения на архитектуру не снижали темпа развития. Таким образом, наблюдается монотонный рост производительности мощных вычислительных систем, что позволяет решать всё более масштабные задачи математического моделирования и прогнозирования. Но и характерный размер задач, другими словами – их вычислительная сложность – также неуклонно растёт.

Следует отметить, что существуют классы задач, актуальных для науки, промышленности и экономики, постановка которых приводит к существенным затруднениям в решении даже с использованием предельных по мощности ЭВМ – суперкомпьютеров. К таким задачам относится, например, задача коммивояжера, которая заключается в нахождении гамильтонова цикла минимального веса в полном взвешенном графе. Она возникает, в частности, в области планирования ресурсов, комбинаторной оптимизации и транспортной логистики.

Это приводит к необходимости поиска новых методов вычислений, отличных от известных, классических способов. Направление квантовых вычислений как раз и предлагает другую модель вычислительных операций, базирующуюся на законах микромира – законах квантовой теории. К настоящему моменту считается общепризнанным фактом, что с помощью квантовых алгоритмов можно получить решение значительного количества задач в классической постановке за время, существенно меньшее, чем с помощью суперкомпьютеров. Для демонстрации указанного эффекта используют специальный термин «квантовое превосходство», который обозначает возможность проведения на квантовом компьютере таких вычислений, выполнение

которых на классических вычислительных системах потребовало бы неприемлемо большого времени.

В квантовом компьютере, как предполагается, важны и аппаратная часть, и используемые алгоритмы, причём теория квантовых алгоритмов в настоящее время содержит большое количество теоретических методов и подходов. Полнофункциональная аппаратная часть такого компьютера находится в процессе разработки, это предмет исследований многих ведущих научных и научно-производственных лабораторий.

Квантовые вычисления являются одной из наиболее перспективных технологий [1, 2], что ставит важную задачу для высшей школы: специалисты в области информационных технологий должны обладать компетенциями в области интенсивно развивающихся квантовых компьютеров.

Как можно видеть [1, 2], знакомство с известными квантовыми алгоритмами и принципами их построения, а также изучение методов вычислений на основе законов квантовой теории можно организовать даже без сложной и дорогостоящей аппаратной базы, поскольку возможно использование классических вычислительных ресурсов, с помощью которых организуется эмуляция работы микросистем, но, конечно, ограниченного объема. Как правило, удается эмулировать работу систем из нескольких кубитов или нескольких десятков кубитов.

Компетенции в области высокопроизводительных вычислений достигаются с помощью необходимой аппаратной базы и программных средств [3, 4], а также специальной учебно-методической базы. Это же относится и к квантовым вычислениям. На кафедре цифровых технологий факультета компьютерных наук Воронежского государственного университета подготовка в области супер-компьютерных и квантовых вычислений ведётся уже более 20 лет.

Важнейшим фактором освоения учебных программ, содержащих физико-математические дисциплины, является формирование креативной образовательной среды в вузах [5]. Под креативным подходом будем понимать способ организации учебного процесса, при котором обучающийся на основе накопленных знаний и навыков окажется способным генерировать новые знания. Необходимо стремиться, чтобы такая ситуация возникала как можно раньше в процессе обучения, в идеале – с первого курса. Более того, существуют предложения для введения основ параллельного программирования на уровне средней образовательной школы. В случае квантовых вычислений, видимо, такое раннее начало не приведет формированию нужных компетенций, поскольку как математический формализм квантовой механики, так и фундаментальные физические концепции чрезвычайно трудно было бы

усвоить без физико-математического базиса, с которым обучающие знакомятся на первом-втором курсах (см., например, [6]).

Отметим необходимость практикоориентированности: для креативного подхода характерна стимуляция обсуждения рассматриваемых задач и указание их значимости на уровне научно-технических исследований, промышленности и экономики. Примером подразделения, где уже в течение длительного времени формируется креативный потенциал, является учебно-научная лаборатория медицинской кибернетики кафедры цифровых технологий. В ней успешно ведутся исследовательские работы с непосредственным и активным участием студентов по актуальным направлениям развития медицины и биологии (см., например, направление [7]).

Отметим также, что математической основой для организации ансамбля вычислительных узлов в кластере или суперкомпьютере является теория графов, поскольку в качестве модели большой системы выбирают ориентированный или неориентированный графы, в которых вершины соответствуют индивидуальным вычислительным узлам, а рёбра отражают физические линии связи между узлами. В силу этого теория графов, представляющая собой абстрактный математический аппарат, хорошо применима в области конструирования сложных систем. Кроме того, эта теория позволяет решать и другие практически важные задачи. Например, размещение большого количества структурных элементов процессора в минимальном объеме требует исключения пересечений внутренних линий связи, что в абстрактной формулировке эквивалентно размещению графа на плоскости без пересечений рёбер, т. е. задаче теории планарных графов. Ещё один раздел теории графов, а именно, где рассматриваются их хроматические свойства, позволяет получить расписание в технологических процессах, в том числе расписание распределения ресурсов для пользователей суперкомпьютеров. В итоге, один из разделов дискретной математики, которым является теория графов, позволяет строго сформулировать и привести алгоритмы решения значительного количества задач сугубо практического плана, что обеспечивает использование полученных знаний и результатов в реальной жизни.

Заключение

Подводя итог, можно констатировать, что решение современных научно-технических задач требует использования как высокопроизводительных вычислений, так и перспективных квантовых вычислений. Подготовка современных специалистов в области информационных технологий и, в частности, квантовых вычислений должна осуществляться на высоком научном и методическом уровне,

идти по пути использования учебных курсов, связанных с понятиями классической и квантовой информации, широкого применения необходимых программно-аппаратных средств и специализированных учебно-методических комплексов.

Литература

1. Борзунов, С. В. Основы квантовых вычислений и квантовой теории информации / С. В. Борзунов, С. Д. Кургалин. – Санкт-Петербург : БХВ-Петербург, 2025. – 248 с.

2. Борзунов, С. В. Квантовые вычисления : учебное пособие / С. В. Борзунов, С. Д. Кургалин. – Санкт-Петербург : БХВ-Петербург, 2022. – 144 с.

3. Кургалин, С. Д. Суперкомпьютерные технологии в Воронежском государственном университете / С. Д. Кургалин, С. В. Борзунов // Вестник Воронежского государственного университета. Серия: Проблемы высшего образования. – 2018. – № 3. – С. 183–187.

4. Borzunov, S. V. Using the resources of the Supercomputer of Voronezh State University in learning processes and scientific researches / S. V. Borzunov, S. D. Kurgalin // Суперкомпьютерные дни в России : труды международной конференции (Москва, 24-25 сентября 2018 г.). – Москва, 2018. – С. 972–977.

5. Кургалин, С. Д. Формирование креативной образовательной среды в вузах / С. Д. Кургалин, И. В. Шершень // Вестник Воронежского государственного университета. Серия: Проблемы высшего образования. – 2016. – № 1. – С. 73–78.

6. Kurgalin, S. D. Algebra and Geometry with Python / S. D. Kurgalin, S. V. Borzunov. – Cham : Springer, 2021. – 422 p.

7. Туровский, Я. А. Динамика цепочек локальных максимумов спектров электроэнцефалограмм человека / Я. А. Туровский, С. Д. Кургалин, А. Г. Семенов // Биофизика. – 2014. – Т. 59, № 1. – С. 185–190.

Метод FROST для распределённой оптимизации с отслеживанием градиента на ориентированных графах

К. В. Новиковский, email: Kostya361a@yandex.ru
А. П. Преображенский, email: app@vivt.ru

Воронежский институт высоких технологий

Аннотация. *В работе исследуется задача распределённой оптимизации над ориентированным графом агентов, где каждый узел располагает локальной функцией и стремится совместно минимизировать сумму локальных критериев. Классические методы распределённого градиентного спуска обеспечивают точную сходимость лишь при убывающем шаге, что существенно снижает практическую эффективность. Современные алгоритмы с отслеживанием градиента (gradient tracking) часто опираются на двустохастические матрицы весов или push-sum нормализацию, требующую знания степеней исходящих соединений и увеличивающую коммуникационные накладные расходы. Предлагается алгоритм FROST (Fast Row-Stochastic Optimization with Uncoordinated Step-Sizes), основанный на строково-стохастических матрицах и допускающий некоординированные локальные шаги обновления. Доказана линейная сходимость для L гладких и сильно выпуклых локальных функций при постоянных шагах. Метод не требует push-sum механизма, обладает пониженной сервисной нагрузкой и эффективен на направленных графах. Экспериментальные результаты демонстрируют преимущество по точности по сравнению с классическим DGD и сопоставимую скорость сходимости с push-sum методами при меньших коммуникационных затратах.*

Ключевые слова: *распределённая оптимизация, многоагентные системы, ориентированные графы, линейная сходимость.*

Введение

Распределённая оптимизация рассматривается в сценариях, когда данные и вычислительные ресурсы находятся на нескольких взаимосвязанных узлах сети, а централизованный сервер отсутствует или его использование нежелательно по причинам отказоустойчивости, приватности или ограничений по пропускной способности каналов связи.

© Новиковский К. В., Преображенский А. П., 2026

Каждый агент i наблюдает только собственную локальную функцию $f_i(x),$ зависящую от некоторого общего вектора параметров $x \in R^p$.

Глобальная задача формулируется как минимизация суммарного функционала $F(x) = \sum_{i-1}^{n} f_i(x).$ Коммуникация между агентами осуществляется по ребрам ориентированного графа, через который узлы обмениваются своими локальными оценками и градиентной информацией.

Наиболее простой и широко известный подход — распределённый градиентный спуск (Distributed Gradient Descent, DGD). Он объединяет локальные шаги градиентного спуска с процедурой согласования (consensus) между соседними узлами. Однако для DGD известно, что при постоянном шаге итерации сходятся лишь к окрестности оптимального решения; для достижения точной сходимости требуется убывающая последовательность шагов, что значительно замедляет практическую сходимость и делает выбор шагов нетривиальной задачей.

В последние годы были предложены методы с отслеживанием градиента (gradient tracking), такие как EXTRA, Exact Diffusion, DIGing и их модификации. В них, помимо локальной оценки параметра, поддерживается дополнительная переменная, аппроксимирующая глобальный градиент. Это позволяет компенсировать эффект несогласованности и добиться линейной сходимости при постоянных шагах. Однако большинство подобных алгоритмов предполагают наличие двустохастической матрицы весов, что требует симметричной или, по крайней мере, ненаправленной топологии графа. Для ориентированных сетей построение двустохастичных матриц, как правило, невозможно.

Альтернативный подход основан на push-sum нормализации и использовании столбцово-стохастических матриц. Push-sum позволяет работать на направленных графах, однако требует передачи дополнительных нормирующих коэффициентов и повышает коммуникационные накладные расходы. Более того, такие методы обычно предполагают согласованные шаги обновления у всех агентов.

Цель данной работы заключается в разработке алгоритма распределённой оптимизации, который: функционирует на ориентированных сильно связных графах; использует только строково-стохастические матрицы весов; не требует push-sum нормализации и знания степеней исходящих рёбер; допускает некоординированные (разные) шаги обновления у агентов; обладает линейной скоростью сходимости при постоянных шагах.

Этим требованиям отвечает предлагаемый алгоритм FROST (Fast Row-Stochastic Optimization with Uncoordinated Step-Sizes). В дальнейшем формально описывается постановка задачи, выводится алгоритм, проводится анализ его сходимости и демонстрируются численные результаты.

1. Обзор литературы

Интерес к распределённым методам оптимизации заметно усилился в связи с развитием больших распределённых систем, сенсорных сетей и методов обучения на периферийных устройствах. Классический метод DGD в различных модификациях был проанализирован в ряде работ; показано, что при фиксированном шаге он обеспечивает лишь ограниченную точность, а при убывающем шаге — медленную сходимость к оптимальному решению [1, 2].

Для устранения этого недостатка был предложен класс алгоритмов с коррекцией по градиенту (gradient tracking). Методы EXTRA и Exact Diffusion используют двустохастические матрицы весов, что позволяет сохранять инвариантность среднего значения и корректно учитывать вклад всех локальных градиентов [3, 4]. Аналогичные идеи развиваются в алгоритмах DIGing и Push-DIGing, где достигается линейная сходимость при постоянных шагах обновления [5, 6]. Однако все эти методы требуют либо симметричной топологии, либо специального подбора матриц весов, что для ориентированных графов практически нереализуемо.

Push-sum-подходы расширяют область применимости на directed-графы за счёт введения дополнительной нормы-переменной и использования столбцово-стохастических матриц [7, 8]. Такой механизм позволяет компенсировать неравномерности степеней, но приводит к увеличению объёма служебного трафика и усложнению реализации, особенно в условиях ограниченных ресурсов и возможных потерь сообщений.

В ряде работ рассматривались варианты gradient tracking на ориентированных графах, однако чаще всего они либо сохраняют зависимость от push-sum нормализации, либо предполагают согласованные шаги обновления у всех агентов [9]. В отличие от этого, алгоритм FROST строится на строково-стохастических матрицах и допускает индивидуальный выбор шагов α_i, что повышает гибкость и потенциальную устойчивость к гетерогенности вычислительных ресурсов.

2. Постановка задачи

Рассматривается ориентированный сильно связный граф $G=(V,E)$, где $V=\{1,...,n\}$ множество агентов, а $E\subseteq V\times V$ множество дуг, задающих направление передачи информации. Наличие дуги $(j,i)\in E$ означает, что агент j может передавать сообщения агенту i. Каждый агент i располагает собственной локальной функцией $f_i: R^p \to R,$ которая предполагается μ сильно выпуклой и L гладкой, то есть для любых $x, y \in R^p$ выполняются условия

$$f_i(y) \geq f_i(x) + \langle \Delta f_i(x), y-x\rangle + \frac{\mu}{2}\|y-x\|^2,$$

$$\|\nabla f_i(y) - \nabla f_i(x)\| \leq L\|y-x\|.$$

Глобальная задача оптимизации формулируется как

$$\min_{x\in R^p} F(x) = \sum_{i=1}^{n} f_i(x).$$

Также предполагается, что существует единственное решение $x^{\backslash *} \in R.$

Обмен информацией задаётся матрицей весов $A=\left[a_{ij}\right]_{i,j=1}^{n}$, согласованной с топологией графа $a_{ij}>0$, только если $(j,i)\in E$ или $i=j$ и $a_{ij}=0$, иначе $i=j$. Матрица A строково-стохастична:

$$a_{ij}>0, \sum_{j=1}^{n} a_{ij} = 1.$$

Такая модель, естественно, описывает схемы, где каждый агент формирует выпуклую комбинацию полученных значений от соседей. В отличие от методов, требующих двустохастичности, здесь не предполагается симметрия графа и нет необходимости знать степени исходящих вершин.

Задача состоит в построении алгоритма, реализуемого каждым агентом на основе локальной информации и обмена с соседями, который обеспечивает сходимость всех локальных оценок x_i^k к оптимальному решению $x^{\backslash *}$ с линейной скоростью.

3. Алгоритм FROST

Каждый агент i на итерации k хранит два вектора, а именно $x_i^k \in R^p$ – локальную оценку искомого решения и $y_i^k \in R^p$ – вспомогательную переменную, отслеживающую изменение локального градиента.

Пусть $a_{ij} > 0$ – локальный шаг обновления, выбираемый агентом независимо. Алгоритм FROST задаётся рекуррентной схемой i

$$x_i^{k+1} = \sum_{j=1}^{n} a_{ij} x_i^k - a_i y_i^k \tag{1}$$

$$y_i^{k+1} = \sum_{j=1}^{n} a_{ij} y_i^k + \nabla f_i(x_i^{k+1}) - \nabla f_i(x_i^k) \tag{2}$$

Первая формула сочетает операцию согласования (выпуклое усреднение по строке матрицы y_i^k. Вторая формула реализует механизм gradient tracking: переменная y_i^k корректируется разностью текущего и предыдущего локальных градиентов и одновременно подвергается согласованию с соседями.

Обозначим $x^k = \left[x_1^k, \ldots, x_n^k \right]^T, Y^k = \left[y_1^k, \ldots, y_n^k \right]^T$ и матрицу $\Lambda = diag(a_1, \ldots, a_n)$. Тогда (1)–(2) можно записать в матричном виде

$$X^{k+1} = AX^k - \Lambda Y^k \tag{3}$$

$$X^{k+1} = AX^k - \Lambda Y^k, \tag{4}$$

где $\nabla F(X) = \left[\nabla f_1(x_1), \ldots, \nabla f_n(x_n)) \right]^T$.

Важной особенностью FROST является то, что шаги a_i могут быть различными для разных агентов. Это позволяет учитывать гетерогенность вычислительных возможностей и локальные ограничения по качеству связи, не вводя дополнительной глобальной координации. При $a_i = a$ для всех i алгоритм сохраняет свои свойства, но становится частным случаем с согласованным шагом.

4. Анализ сходимости

Для анализа сходимости используем стандартный аппарат теории матриц Маркова и свойств сильно выпуклых и L гладких функций.

Поскольку A является неприводимой строково-стохастической матрицей, по теореме Перрона–Фробениуса, существует единственный левый собственный вектор $L\pi \in R^n$, такой, что

$$\pi^T A = \pi^T, \pi_i > 0, \sum_{i=1}^{n} \pi_i = 1$$

Вектор π задаёт стационарное распределение марковской цепи с матрицей переходов A. Определим взвешенное среднее локальных оценок и трекеров

$$\bar{x}^k = \sum_{i=1}^{n} \pi_i x_i^k, \bar{y}^k = \sum_{i=1}^{n} \pi_i y_i^k.$$

Введём проектор $\Pi = 1\pi^T$, где 1 – вектор из единиц длины n. Тогда матричные переменные можно разложить на консенсусную и ортогональную компоненты

$$X^k = \Pi X^k + X_\perp^k = 1\bar{x}^k + X_\perp^k,$$

$$Y^k = \Pi Y^k + Y_\perp^k = 1\bar{y}^k + Y_\perp^k,$$

где $$X_\perp^k = (I - \Pi)X^k Y_\perp^k = (I - \Pi)X^k.$$

Сходимость алгоритма разбивается на две подсоставляющие, а именно на убывание ортогональных компонент $X_\perp^k$ и $Y_\perp^k$, что соответствует достижению консенсуса и сходимость усреднённой оценки $\bar{x}$ к оптимальному решению $\bar{x}^k$.

Основной результат можно сформулировать следующим образом.

Теорема. Пусть для всех $i = 1,\ldots,n$ шаги a_i удовлетворяют условию

$$0 < a_i < \frac{2}{L + \mu}.$$

Тогда существует число $\rho \in (0,1)$ константа $C > 0$, такие, что для всех $k \geq 0$ выполняется неравенство

$$\left\| X^k - 1 \otimes x^{\backslash *} \right\|^2 + \left\| Y^k - 1 \otimes \nabla F(x^{\backslash *}) \right\|^2 \leq C\rho^k.$$

Иными словами, локальные оценки x_i^k сходятся к оптимальному решению $x^{\backslash *}$ с линейной скоростью, а переменные y_i^k сходятся к соответствующему значению глобального градиента $\nabla F(x^{\backslash *})$.

В доказательстве теоремы используется функционал Ляпунова вида

$$V^k = \left\|Y_\perp^k\right\|^2 + c\left\|Y_\perp^k\right\|^2 + d\left\|\bar{x}^k - x\right\|^2,$$

где коэффициенты $c > 0$ и $d > 0$ подбираются таким образом, чтобы матрица перехода в расширенном пространстве ошибок обладала спектральным радиусом меньше единицы. Используя свойства матрицы A и неравенства, вытекающие из сильной выпуклости и гладкости, получаем оценку вида

$$V^{k+1} \le pV^k.$$

Таким образом, FROST обеспечивает линейную сходимость при постоянных шагах и не требует ни двустохастичности матрицы весов, ни push-sum нормализации.

5. Экспериментальные результаты

Для подтверждения теоретических результатов была проведена серия численных экспериментов. Рассматривались сети с различным числом агентов n (например, $n = 20, 50, 100$) и типами ориентированных графов, такими как ориентированное кольцо (Directed Ring), где каждый узел получает информацию от одного соседа и передаёт следующему, случайный направленный граф Эрдаша–Реньи с заданной вероятностью появления дуги и случайный сильно связный граф с неравномерной степенью вершин, моделирующий асимметричные коммуникационные структуры.

В качестве тестовых локальных функций использовались квадратичные функции вида

$$f_i(x) = \frac{1}{2}x^T Q_i x + b_i^T x,$$

где симметричные матрицы Q_i выбирались положительно определёнными и удовлетворяющими $\mu I \prec Q_i \prec LI$. Такой выбор позволяет аналитически задать параметры μ и L и легко вычислять оптимальное решение $x^{\backslash *}$.

Шаги a_i выбирались неоднородно по формуле

$$a_i = \frac{\gamma}{\|Q_i\|},$$

где параметр $\gamma \in (0,1)$ подбирался так, чтобы выполнялось условие теоремы о сходимости. Тем самым моделировалась ситуация, когда агенты обладают разными локальными параметрами и не могут согласовать общий шаг.

Для оценки качества сходимости использовались следующие метрики:

$$E_x(k) = \frac{1}{n}\sum_{i=1}^{n}\left\| x_i^k - x^{\backslash *} \right\|,$$

характеризующая среднюю ошибку по всем узлам, и запись, характеризующая точность оценки глобального градиента

$$E_g(k) = \frac{1}{n}\sum_{i=1}^{n}\left\| y_i^k - \nabla f_i(x^{\backslash *}) \right\|.$$

Алгоритм FROST сравнивался с классическим DGD с постоянным шагом, а также с push-sum-модификацией градиентного метода. Во всех рассмотренных конфигурациях наблюдалась геометрическая (линейная) сходимость ошибки для FROST, причём скорость убывания ошибки мало зависела от структуры графа при условии сохранения сильной связности. Типичные значения экспериментально оцениваемого коэффициента сходимости ρ находились в диапазоне от 0.92 до 0.97.

Метод DGD с постоянным шагом демонстрировал быстрое начальное уменьшение ошибки, но далее выходил на плато и не достигал точного решения, что согласуется с теоретическими результатами. Push-sum-метод показывал поведение, близкое к FROST по скорости сходимости, но требовал передачи дополнительных нормирующих переменных, что приводило к увеличению объёма коммуникаций.

Особое внимание уделялось влиянию неоднородности шагов a_i, что подтверждает его пригодность для гетерогенных распределённых систем.

Заключение

В работе предложен и исследован алгоритм FROST (Fast Row-Stochastic Optimization with Uncoordinated Step-Sizes) для решения задач распределённой оптимизации над ориентированными графами. Алгоритм использует строково-стохастические матрицы весов и механизм отслеживания градиента, допускает некоординированные шаги обновления и не требует push-sum нормализации.

Проведённый анализ показал, что при стандартных предположениях о сильной выпуклости и гладкости локальных функций FROST обеспечивает линейную сходимость к точному оптимальному решению при постоянных шагах. Численные эксперименты подтвердили теоретические выводы и продемонстрировали преимущество метода по сравнению с классическим DGD, а также сопоставимую скорость сходимости с push-sum-подходами при меньших коммуникационных затратах.

Перспективными направлениями дальнейших исследований являются: расширение алгоритма на стохастический режим (использование стохастических градиентов), анализ поведения в условиях динамически изменяющихся графов, разработка проксимальных модификаций для негладких задач, а также построение адаптивных стратегий выбора шагов a_i на основе локальной информации агентов.

Список литературы

1. Nedić, A. Distributed subgradient methods for multi-agent optimization / A. Nedić, A. Ozdaglar // IEEE Transactions on Automatic Control. – 2009. – Vol. 54, № 1. – P. 48–61.
2. Nedić, A. Constrained consensus and optimization in multi-agent networks / A. Nedić, A. Ozdaglar, P. A. Parrilo // IEEE Transactions on Automatic Control. – 2010. – Vol. 55, № 4. – P. 922–938.
3. EXTRA: An exact first-order algorithm for decentralized consensus optimization / W. Shi [et al.] // SIAM Journal on Optimization. – 2015. – Vol. 25, № 2. – P. 944–966.
4. Yuan, K. Exact diffusion for distributed optimization and learning – Part I: Algorithm development / K. Yuan, B. Ying, A. H. Sayed // IEEE Transactions on Signal Processing. – 2018. – Vol. 67, № 3. – P. 708–723.
5. Nedić, A. Achieving geometric convergence for distributed optimization over time-varying graphs / A. Nedić, A. Olshevsky, W. Shi // SIAM Journal on Optimization. – 2017. – Vol. 27, № 4. – P. 2597–2633.
6. Nedić, A. Stochastic gradient-push for strongly convex functions on time-varying directed graphs / A. Nedić, A. Olshevsky // IEEE Transactions on Automatic Control. – 2016. – Vol. 61, № 12. – P. 3936–3947.
7. Kempe, D. Gossip-based computation of aggregate information / D. Kempe, A. Dobra, J. Gehrke // Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science. – Cambridge. – 2003. – P. 482–491.
8. Duchi, J. C. Dual averaging for distributed optimization: convergence analysis and network scaling / J. C. Duchi, A. Agarwal, M. J. Wainwright // IEEE Transactions on Automatic Control. – 2012. – Vol. 57, № 3. – P. 592–606.
9. Pu, S. Push–pull gradient methods for distributed optimization in networks / S. Pu, A. Nedić, A. Olshevsky // IEEE Transactions on Automatic Control. – 2021. – Vol. 66, № 1. – P. 1–16.

Особенности использования сопроцессоров Intel Xeon Phi в обучении студентов параллельному программированию

А. Г. Потапов, email: artem2011.ap@inbox.ru

Воронежский государственный университет

Аннотация. *Рассматривается процесс внесения изменений в структуру высокопроизводительного вычислительного кластера Воронежского государственного университета (ВГУ) с целью создания наилучших условий для использования сопроцессоров Intel Xeon Phi в учебном процессе.*

Ключевые слова: *суперкомпьютер, Xeon Phi, образование.*

Введение

Сопроцессоры Intel Xeon Phi дают возможность запускать программы в многопоточном режиме с одновременным использованием 60 ядер и 240 потоков, что значительно превышает среднестатистическое количество ядер на центральных процессорах. Несмотря на невысокую среднюю тактовую частоту (1.24 ГГц), Xeon Phi предоставляют отличный инструмент для отработки студентами методов распараллеливания на системе с общей памятью [1–3].

Xeon Phi имеют два основных режима работы: offload-режим, подразумевающий необходимость модификации программного кода и использование как минимум 1 ядра CPU, и native-режим, в котором сопроцессоры фактически являются самостоятельными вычислительными узлами. Native-режим является более предпочтительным для обучения студентов работе с базовыми конструкциями OpenMP и MPI ввиду отсутствия специфичности системы команд.

В связи с отсутствием технической поддержки со стороны производителя, а также с рядом аппаратно-программных особенностей сопроцессоров, существуют известные трудности для их эксплуатации в составе вычислительного комплекса. Такими особенностями являются отсутствие жёстких дисков и внешних Ethernet-интерфейсов, устаревшая версия 2.6 Linux-ядра и, как следствие, отсутствие поддержки современных версий SLURM.

Рассмотрены способы решения вышеупомянутых трудностей на примере высокопроизводительного кластера факультета компьютерных наук Воронежского государственного университета, включающие

© Потапов А. Г., 2026

использование технологии «external bridging» и одновременное содержание двух версий SLURM.

1. Установление связи с вычислительным полем

Первая задача, которую необходимо выполнить для перевода сопроцессоров в native-режим, состоит в их подключении ко всем вычислительным узлам кластера. Проблема состоит в том, что по умолчанию драйвер MPSS (Manycore Platform Software Stack) открывает доступ сопроцессорам только на хостовой узел, к которому они подключены физически через PCI Express. В связи с отсутствием интерфейсов Ethernet, единственным способом внедрить Xeon Phi в локальную сеть кластера является использование сетевого моста.

Сетевой мост (bridge) фактически является виртуальным сетевым коммутатором (switch), объединяющим несколько устройств во внутренней сети по их MAC-адресам. MPSS поддерживает использование данной технологии, однако для переключения сетевого интерфейса вычислительного узла в режим моста необходимо временно отключить его от сети кластера, что является второй значимой проблемой ввиду монтирования корневого каталога через NFS.

Отключение Ethernet-интерфейса напрямую неизбежно приводит к разрыву соединения с NFS-сервером и прекращению нормального функционирования всей системы. Несмотря на то, что большинство важнейших каталогов содержатся в оперативной памяти при использовании tmpfs, система не способна самостоятельно восстановить доступ к сети без необходимых исполняемых и конфигурационных файлов.

Для решения данной проблемы был написан bash-скрипт, создающий минимально жизнеспособную изолированную среду в папке /tmp/root. В неё помещается командная оболочка и несколько важнейших программ, включая brctl, отвечающую за создание сетевого моста. С помощью команды chroot корень системы временно переносится в выделенную папку, хранящуюся в ОЗУ и не зависящую от NFS-сервера, внутри которой монтируется /proc и выполняются команды по настройке новой сетевой конфигурации. После этого корень системы возвращается в исходную NFS-директорию, а вычислительный узел становится вновь доступным для подключения. Его IP-адрес при этом переносится с eth0 (enp3s0f0) на только что созданный интерфейс br0. Фрагмент вывода команды “ip a” после внесения описанных изменений представлен на рис. 1.

```
2: enp3s0f0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq master br0 state UP qlen 1000
    link/ether 0c:c4:7a:56:f3:ec brd ff:ff:ff:ff:ff:ff
    inet6 fe80::ec4:7aff:fe56:f3ec/64 scope link
       valid_lft forever preferred_lft forever
3: enp3s0f1: <BROADCAST,MULTICAST> mtu 1500 qdisc noop state DOWN qlen 1000
    link/ether 0c:c4:7a:56:f3:ed brd ff:ff:ff:ff:ff:ff
4: mic0: <BROADCAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast master br0 state UNKNOWN qlen 1000
    link/ether 4c:79:ba:73:02:57 brd ff:ff:ff:ff:ff:ff
    inet6 fe80::4e79:baff:fe73:257/64 scope link
       valid_lft forever preferred_lft forever
5: mic1: <BROADCAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast master br0 state UNKNOWN qlen 1000
    link/ether 4c:79:ba:73:02:a9 brd ff:ff:ff:ff:ff:ff
    inet6 fe80::4e79:baff:fe73:2a9/64 scope link
       valid_lft forever preferred_lft forever
6: br0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP qlen 1000
    link/ether 0c:c4:7a:56:f3:ec brd ff:ff:ff:ff:ff:ff
    inet 10.16.255.1/24 brd 10.16.255.255 scope global br0
       valid_lft forever preferred_lft forever
    inet6 fe80::ec4:7aff:fe56:f3ec/64 scope link
       valid_lft forever preferred_lft forever
```

Рис. 1. Структура сетевых интерфейсов вычислительного узла

В результате выполнения скрипта и подключения к br0 виртуальных интерфейсов mic0 и mic1 сопроцессоры Xeon Phi становятся доступными для подключения по SSH с любого узла кластера. Каждому из них присвоено уникальное имя вида node1-mic0, позволяющее однозначно идентифицировать сопроцессоры в локальной сети.

2. Установка SLURM

В 2018 году компания Intel прекратила официальную поддержку сопроцессоров Xeon Phi [4]. В связи с этим дальнейшее обновление их драйверов и, как следствие, программного обеспечения стало невозможным. Сопроцессоры архитектуры Knights Corner, установленные на вычислительном кластере ВГУ, обладают Linux-ядром версии 2.6, которое считается устаревшим. Почти у всех современных программных пакетов отсутствует совместимость с такой версией ядра, в связи с чем требуется установка более старых версий ПО.

Кроме того, MIC-архитектура специфична, для сборки программных пакетов на ней с помощью make требуется внесение изменений в конфигурационные файлы программ. Для этого существует специальная утилита gnu-configize, содержащаяся в MPSS, выполняющая нужные преобразования в автоматическом режиме. В качестве альтернативы возможно использование кросс-компиляции на хостовом узле, однако данный способ на практике оказался менее эффективным.

Поскольку стандартный комплект программ для Xeon Phi является более ограниченным, чем для среднестатистической Linux-системы, для развёртывания на них системы управления очередью исполнения SLURM потребовалась ручная сборка большего количества пакетов (к примеру, hwloc и python2). К тому же, старая версия Linux-ядра не позволила

установить версию SLURM, использующуюся на вычислительных узлах кластера, в связи с чем было принято решение развернуть два параллельно работающих SLURM-сервера (см. рис. 2). Такая реализация стала возможной благодаря созданию нескольких виртуальных машин с использованием Xen-гипервизора, каждая из которых выполняет функцию отдельного служебного узла [5].

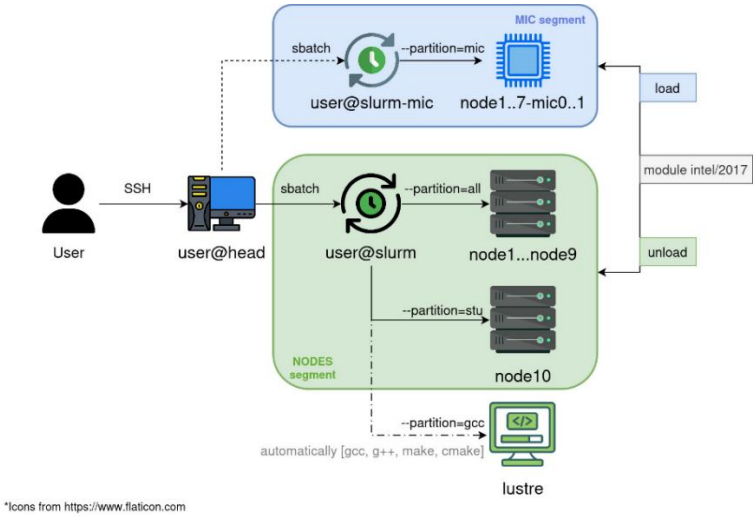


Рис. 2. Структура очередей исполнения на кластере ВГУ

3. Внедрение сопроцессоров в учебный процесс

Для удобства использования Xeon Phi после применения всех вышеописанных изменений был создан модуль intel/2017. В первую очередь, он предоставляет доступ студентам к семейству компиляторов Intel, включая icc, icpc, mpiicc и т. д., поддерживающих ключ компиляции -mmic. Нюанс заключается в ограничении возможности использования данного ключа только теми узлами, на которых установлен MPSS и присутствует соответствующий модуль ядра, что делает непригодным использование для кросс-компиляции головного узла. Поэтому при загрузке модуля intel/2017 выполняется автоматическое перенаправление обращения к icc на один из вычислительных узлов, содержащих в своём составе Intel Xeon Phi.

Следующей важной задачей стало обеспечение доступности сопроцессоров для использования в рамках очереди исполнения с головного узла при минимальном вмешательстве студентов в её работу.

Для этой цели было осуществлено перенаправление основных команд взаимодействия с очередью (squeue, sbatch, srun) на MIC-сегмент при загрузке intel/2017. Если на 10 вычислительном узле (24 ядра) максимальное время выполнения задачи студентами составляет 1 час ввиду высокого спроса на ресурсы, то на 12 сопроцессорах (2880 ядер в сумме) оно было увеличено до 14 дней. Таким образом, появились гораздо более широкие возможности для тестирования технологий параллельного программирования в учебном процессе.

Заключение

В результате проведения серии мероприятий по настройке сопроцессоров Intel Xeon Phi на высокопроизводительном вычислительном кластере ВГУ, включая настройку сетевого моста для установления связи между всеми вычислительными узлами и сопроцессорами, а также организацию двух параллельно работающих серверов управления очередью исполнения SLURM, сопроцессоры Xeon Phi получили возможность работы в native-режиме в качестве самостоятельных вычислительных узлов.

Было существенно упрощено использование сопроцессоров студентами, благодаря чему значительно расширились возможности для изучения особенностей параллельного программирования.

Список литературы

1. Кургалин, С. Д. Суперкомпьютерная система как фактор влияния на образовательный процесс / С. Д. Кургалин, С. В. Борзунов // АППМИМ: сб. тр. Международной научной конференции. – Воронеж, 2024. – С. 4–9.

2. Кургалин, С. Д. Суперкомпьютерные вычисления: практический подход / С. Д. Кургалин, С. В. Борзунов. – СПб. : БХВ, 2019. – 256 с.

3. Гергель, В. П. Теория и практика параллельных вычислений / В. П. Гергель. – М. : Национальный Открытый Университет «ИНТУИТ», 2016. – 501 с.

4. Intel ends the Xeon Phi product line [Электронный ресурс]. – Режим доступа: https://www.networkworld.com/article/966180/intel-ends-the-xeon-phi-product-line.html

5. Потапов, А. Г. Внедрение новой структуры программного комплекса гибридного вычислительного кластера Воронежского государственного университета / А. Г. Потапов // АППМИМ: сб. тр. Международной научной конференции. Секция: Высокопроизводительные вычисления в науке, технике и образовании. – Воронеж, 2026. – С. 22–26.

Библиотеки и инструменты квантового ИИ для Quantum PCA (qPCA)

И. С. Хаустов, email: sparkcolazet@gmail.com
А. Ф. Клинских, email: 2000afk@gmail.com

Воронежский государственный университет

Аннотация. *В работе рассматриваются современные библиотеки и инструменты квантового искусственного интеллекта в контексте реализации алгоритма Quantum Principal Component Analysis (qPCA) – квантового аналога метода главных компонент. Описываются теоретические основы qPCA, включая использование квантового фазового оценивания для извлечения собственных значений и собственных векторов ковариационной матрицы, закодированной в виде квантового состояния.*

Ключевые слова: *Quantum PCA, квантовый ИИ, квантовые вычисления, метод главных компонент, квантовое фазовое оценивание (QPE), ковариационная матрица, квантовые алгоритмы, гибридные квантово-классические методы, Qiskit, квантовое машинное обучение.*

Введение

Quantum Principal Component Analysis (qPCA) – квантовый аналог метода главных компонент, предложенный Ллойдом и соавт. в 2014 г. В классической версии PCA ищет ортогональные главные компоненты данных, обеспечивающие максимальную дисперсию, решая задачу диагонализации ковариационной матрицы. Квантовый алгоритм qPCA оперирует над квантовым состоянием, кодирующим ковариационную матрицу, и извлекает ее наибольшие собственные значения и собственные векторы с потенциальным экспоненциальным ускорением. Идея состоит в использовании квантового параллелизма: фаза собственных состояний определяется через квантовое фазовое оценивание (QPE), что теоретически позволяет получить главные компоненты за логарифмическое число шагов от размерности данных. На практике, однако, реализация qPCA сталкивается с трудностями: требуются эффективная загрузка данных в квантовые амплитуды и многократные копии состояния, а шум в текущих устройствах ограничивает преимущества. Более того, недавние исследования показывают, что экспоненциальный выигрыш qPCA зависит от предположений о подготовке входных состояний – классический алгоритм может достичь схожей сложности при наличии аналогичного доступа к данным. Тем не менее, qPCA остается важным примером

квантового алгоритма для задачи уменьшения размерности и стимулирует разработку гибридных квантово-классических методик (например, вариационных схем PCA).

Квантовый ИИ-фреймворк – это программная платформа для разработки и выполнения квантовых алгоритмов машинного обучения. Ниже мы рассмотрим ключевые современные фреймворки – IBM Qiskit, PennyLane (Xanadu), TensorFlow Quantum (Google), Cirq (Google) – а также другие (Microsoft Q# и Amazon Braket), с акцентом на их возможности для реализации qPCA. Мы сравним их архитектуру, средства для qPCA, примеры кода, ограничения и преимущества, а также интеграцию с квантовыми устройствами. Таблица в конце резюмирует функциональность, производительность и удобство использования этих фреймворков. Также приведем примеры исследований, где qPCA реализован на практике, включая диаграммы квантовых схем и результаты экспериментов [1].

1. Метод главных компонент: классический vs квантовый

Классический PCA вычисляет собственные векторы ковариационной матрицы данных

$$C = \frac{X^T X}{n} \tag{1}$$

X – матрица данных, n – количество строк данных.

Главные компоненты – это собственные векторы с наибольшими собственными значениями (дисперсиями), позволяющие проецировать данные на подпространство меньшей размерности, сохраняя основную информацию. Алгоритмически, PCA требует времени $O\left(n^3\right)$для матрицы, что затрудняет обработку высокоразмерных данных.

Квантовый PCA (qPCA) предполагает наличие квантового состояния ρ, пропорционального ковариационной матрице (1), применения унитарного оператора

$$U = e^{\frac{-i}{\rho}t} \tag{2}$$

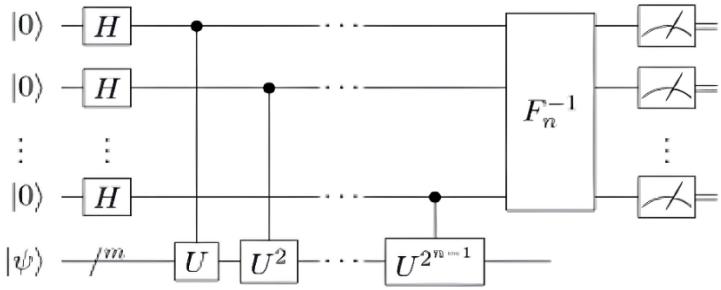
ρ – квантовое состояние.

Использование квантового фазового оценивания (QPE) для извлечения собственных фаз (связанных с собственными значениями λ).

Измерение фазы, коллапсирующее состояние регистра в соответствующий собственный вектор, даёт нам главную компоненту. На выходе qPCA даёт квантовое состояние, представляющее главный компонент, которое можно далее использовать в квантовом алгоритме или измерить для получения классического результата. Теоретически, если ρ

имеет ранг r, qPCA может выявлять первые r компонент за время $O(poly\log(m))$, тогда как классически PCA – $O\left(n^3\right)$ [2].

Квантовое фазовое оценивание (QPE) – ключевой компонент qPCA. QPE использует цепочку из t вспомогательных кубитов и управляемых применений к регистру с состоянием ρ, за которыми следует обратное квантовое Фурье-преобразование, чтобы получить битовую приближенную фазу на вспомогательных кубитах. На схеме ниже показан пример стандартной схемы QPE для унитарного оператора U и одного известного собственного состояния (рис. 1):



Верхние вспомогательные кубиты проходят через цепочку Адамаров и управляемых U^{2^k}, после чего выполняется обратное КФ-преобразование и измерение

Рис. 1. Схема квантового фазового оценивания (QPE) для определения фазы φ унитарного оператора U

В контексте qPCA оператор играет роль «эволюции» под действием ковариационной матрицы. Собственные фазы несут информацию о собственных значениях. После QPE измерение вспомогательного регистра дает оценку; параллельно информационный регистр коллапсирует в соответствующее собственное состояние – главную компоненту. Поскольку U реализуется через разложение $e^{-i\rho t}approx(e^{-i\rho\Delta t})^N$ с помощью дополнительных копий, ресурсные затраты возрастают с точностью оценки и рангом ρ. В ряде работ были предложены оптимизации: резонансный qPCA минимизирует число вспомогательных кубитов, сканируя частоту управления одним кубитом для резонансного возбуждения главной компоненты; гибридные методы объединяют квантовую фазовую оценку с классической пост-обработкой

для повышения устойчивости к шумам и сокращения глубины цепи. Ниже мы рассмотрим, как доступные библиотеки поддерживают реализацию подобных алгоритмов [3].

2. IBM Qiskit

Архитектура и язык: Qiskit – открытый фреймворк на Python от IBM, один из самых популярных в квантовом программировании. Он разделен на модули: Terra (основы описания квантовых цепей, трансляция на уровень аппаратуры), Aer (высокопроизводительные симуляторы состояний и шумовых моделей), Ignis (инструменты смягчения ошибок и калибровки, сейчас интегрированы в Terra) и прикладные пакеты (раньше объединенные в Qiskit Aqua) – например, Qiskit Machine Learning, Qiskit Nature и др., предоставляющие реализации алгоритмов высокого уровня. Qiskit написан на Python с критическими частями на C++ для производительности. Он обеспечивает как низкоуровневый контроль (конструирование схем gate-by-gate, управление импульсами на аппаратном уровне), так и высокоуровневые алгоритмы (готовые классы VQE, QSVM, QAOA и пр.). Важным достоинством Qiskit является прямая интеграция с реальными квантовыми устройствами IBM через облако – пользователь может запустить схему на настоящем процессоре (от 5 до 127 кубитов и более) практически в тех же строках кода, что и на симуляторе. Это сделало Qiskit стандартным выбором для обучения и экспериментов, с обширным сообществом и документацией.

Реализация qPCA: в актуальной версии Qiskit нет специализированного класса qPCA, однако инструменты для построения алгоритма доступны. В ранних версиях Qiskit Aqua содержалась демонстрационная реализация qPCA, основанная на Phase Estimation. Сейчас пользователю предлагается самостоятельно собрать цепь QPE, используя модуль Qiskit Algorithms (classical–quantum algorithms). Например, можно задействовать готовый алгоритм Iterative Phase Estimation (итеративное фазовое оценивание) для оценки собственного значения. Ниже показан фрагмент кода, иллюстрирующий игрушечную реализацию qPCA в Qiskit: мы предполагаем, что имеется оператор (например, полученный через техники экспоненциального приближения с дополнительными кубитами), и применяем к нему фазовое оценивание:

Листинг 1

Фрагмент кода на qiskit

```
from qiskit.algorithms import IterativePhaseEstimation
from qiskit import QuantumCircuit

# Допустим, мы построили унитар U ≈ exp(-i ρ^t)
```

```
qpe = IterativePhaseEstimation(epsilon=0.01)  # алгоритм
фазовой оценки с заданной точностью
prep_circuit = QuantumCircuit(n_qubits)    # схема подготовки
состояния |ψ>, соответствующего ρ
# ... здесь подгружается инициализация данных в prep_circuit
...
result = qpe.estimate(unitary=U,
state_preparation=prep_circuit)
print("Оценка собственного значения:", result.phase)
```

Этот код использует итеративный метод вместо полной схемы QPE, экономя кубиты ценой повторных запусков. Кроме того, в Qiskit есть модуль quantum_info с функциями для работы с плотностными матрицами и эволюцией; им можно воспользоваться для симуляции на небольших размерах. На практике основная сложность – подготовка входного состояния ρ из классических данных. Qiskit предоставляет методы для инициализации произвольных амплитуд (например, Initialize или более оптимальные схемы), но их сложность растет экспоненциально с числом кубитов. Поэтому демонстрации qPCA в Qiskit обычно ограничиваются малыми размерами (2–3 кубита) с простыми, вручную заданными r для подтверждения концепции.

Пример использования: в репозитории Qiskit Tutorials и сторонних источниках есть несколько примеров qPCA. Один из подходов – эмуляция ковариационной матрицы на квантовом регистре: допустим, нужно анализировать квантовое состояние неизвестной 2×2 матрицы плотности. Можно прямым кодом задать ρ как statevector или оператор и провести фазовое оценивание с помощью собственных методов Qiskit (например, используя QPE на основе схемы Китаева). Однако готовых «однострочных» вызовов для qPCA нет – требуется собрать алгоритм вручную из базовых операций. Пользователи Qiskit компенсируют это богатой библиотекой готовых матричных экспонент (например, PauliEvolutionGate для экспоненциального гамильтониана) [4]. Таким образом, Qiskit обеспечивает «конструктор» для qPCA, предоставляя кирпичики, экспоненциальный оператор, фазовое оценивание и измерение.

Ограничения и преимущества: преимущества Qiskit – богатый функционал и поддержка реального железа. Для qPCA это означает, что в Qiskit можно не только промоделировать алгоритм на симуляторе (с учётом шума с помощью Aer), но и попытаться запустить на реальном квантовом процессоре IBM. Например, эксперимент 2021 г. в IBM показал применение QSVM (quantum SVM) на 6-кубитном устройстве для классификации белков, что близко по идеологии к qPCA (квантовое выделение признаков). Прямой запуск qPCA на текущих устройствах

ограничен их шумностью и числом кубитов; однако Qiskit поддерживает динамические схемы и квазиопросы (mid-circuit measurement), что можно использовать для повторного использования кубитов в итеративной фазовой оценке, снижения аппаратных требований. К ограничениям Qiskit можно отнести его ориентацию на gate-модель и IBM-экосистему: работа с непрерывными переменными или отжиговыми системами не покрывается; также интеграция сторонних облачных сервисов требует коннекторов. Тем не менее, Qiskit остается золотым стандартом общего назначения, и новые методы (например, Variational Quantum PCA) могут быть реализованы на нём благодаря модулю оптимизации и автодифференцируемым схемам (через пакет Qiskit Machine Learning).

Интеграция с аппаратурой: Qiskit изначально создавался для IBM Quantum Experience, поэтому он поддерживает прямую отправку схем на суперапроводниковые квантовые компьютеры IBM через API. Для qPCA это значит, что можно, например, провести часть алгоритма на реальном устройстве – скажем, измерить спектр ρ на нескольких эволюционных временах с помощью экспериментальной процедуры. Также Qiskit предоставляет симуляторы: statevector (идеальный, для малых систем), qasm_simulator (имитирует измерения), density_matrix_simulator (для смешанных состояний), с возможностью указать шумовые модели. Квантовые устройства других компаний могут поддерживаться косвенно: некоторые провайдеры (Honeywell, IonQ) позволяли загружать схемы в формате QASM, сгенерированного Qiskit. Кроме того, через Qiskit-метод IBMQ.get_provider() доступны бэкенды, включая эмуляторы на GPU (например, Aer GPU). Таким образом, Qiskit удобно использовать для полнокубитных симуляций qPCA и экспериментов на реальных IBM QPU, хотя из-за требований qPCA такие запуски пока ограничены игрушечными задачами.

3. TensorFlow Quantum (TFQ)

Архитектура и язык: TensorFlow Quantum – библиотека, созданная в сотрудничестве Google и университетов в 2020 г., расширяющая платформу TensorFlow для поддержки квантовых вычислений в моделях машинного обучения. TFQ интегрируется с TensorFlow 2.x (Python), предоставляя слой, позволяющий включать квантовые схемы как часть вычислительного графа TensorFlow. Основой для описания квантовых операций служит библиотека Cirq – схемы строятся на Cirq, а TFQ предоставляет операции (op) для их выполнения и дифференцирования внутри TF. Проще говоря, TFQ позволяет определить квантовую схему, передать ее в модель Keras и обучать параметры схемы совместно с классическими параметрами, используя стандартные инструменты

глубокого обучения (обратное распространение, GPU ускорение и т. п.). TFQ реализован на Python (оболочка) с высокопроизводительными C++ ядрами (для симуляции схем на CPU/GPU) и тесно связан с Cirq: фактически, под капотом, когда мы запускаем квантовую часть модели, TFQ вызывает Cirq-симулятор или API Google Quantum Engine (для реальных запусков, если доступны).

Фокус TFQ – параметризованные квантовые схемы в контексте ML (вариационные квантовые схемы, квантовые слои нейронных сетей). Он предоставляет слой tfq.layers.PQC (Parameterized Quantum Circuit), который можно вставить в Keras-модель как обычный слой. Этот слой принимает на вход кластер квантовых схем (в виде тензора) и выход – ожидания наблюдаемых величин. TFQ автоматически просчитывает градиенты по параметрам схем (методом сдвига параметров или через тензорную аналитику) и поддерживает батчи квантовых данных: можно одновременно подавать на вход несколько схем/наблюдаемых для параллельной обработки, что оптимально задействует ресурсы. Язык схем – Cirq, т. е. пользователь описывает квантовый алгоритм через операции Cirq на кубитах (пример: cirq.H(q0) для Адамара, и т. д.). Это довольно низкоуровневый способ, но TFQ компенсирует, предоставляя готовые функциональные слои [5].

Реализация qPCA: специфического класса/функции для qPCA в TFQ нет – как и PennyLane, этот фреймворк не ориентирован на «алгоритмы без обучения». Тем не менее, qPCA может быть реализован в TFQ различными путями:

Явная QPE-схема в качестве Keras-модели: можно описать всю схему квантового фазового оценивания на Cirq (включая вспомогательные кубиты, контроллы и обратное QFT) и обернуть ее в tfq.layers.PQC. Параметры схемы в данном случае могут отсутствовать (или быть, например, временным параметром t эволюции), то есть мы фактически получим модель без тренировки, выполняющую алгоритм.

Листинг 2

Пример кода на TensorFlow Quantum

```
import cirq, sympy, tensorflow as tf
import tensorflow_quantum as tfq

# кубит и параметр
q = cirq.GridQubit(0,0)
theta = sympy.symbols('theta0')
# схема подготовки ansatz (Ry(theta) на |0>):
circuit = cirq.Circuit(cirq.ry(theta).on(q))
# наблюдаемое (например, rho = 0.8|0><0| + 0.2|1><1| = 0.8*Z +
0.2*I, с точностью до константы)
obs = 0.8 * cirq.Z(q) + 0.2 * cirq.I(q)
```

```
# Превращаем в тензоры для TFQ:
circuit_tensor = tfq.convert_to_tensor([circuit])
obs_tensor = tfq.convert_to_tensor([[obs]])

# Определяем модель: слой PQC возвращает ожидание obs
pq_layer = tfq.layers.PQC(circuit, obs)
model = tf.keras.Sequential([pq_layer])

# Инициализируем параметр (в TensorFlow он внутри слоя)
model.build(input_shape=()) # нет классических входов
# Оптимизация:
optimizer = tf.keras.optimizers.Adam(learning_rate=0.1)
for step in range(100):
    with tf.GradientTape() as tape:
        exp_val = model([])    # ожидание obs для текущего
theta
        loss = -exp_val    # максимизируем ожидание =
минимизируем -ожидание
        grads = tape.gradient(loss, model.trainable_variables)
        optimizer.apply_gradients(zip(grads,
model.trainable_variables))
print("Оптимизированное ожидание rho:", exp_val.numpy())
print("Параметр theta:", model.trainable_variables[0].numpy())
```

Ограничения и преимущества: преимущество TFQ – плотная интеграция с классическим ML-инструментарием Google. Это удобно для исследователей, знакомых с TensorFlow: квантовые элементы вписываются в знакомый pipeline обучения. TFQ эффективен для гибридных нейросетевых архитектур, например, квантовых нейронных сетей, квантового переноса обучения и т. д. В контексте qPCA это менее напрямую полезно, т. к. qPCA – алгоритм без обучаемых параметров. Однако TFQ может блистать в задачах типа квантового обучения особенностям (feature learning), где после выделения главных компонент (набор квантовых состояний) их можно дальше подать на квантовый классификатор – все это оформить единой дифференцируемой моделью.

Ограничения TFQ: он жестко привязан к экосистеме TensorFlow. Пользователю нужно владеть TF2, понимать концепции тензорных вычислений и графов. Это усложняет порог вхождения для тех, кто пришел из квантовой информатики (им проще Qiskit или PennyLane). Кроме того, отладка квантовых схем труднее – нельзя просто «распечатать» состояние кубитов в процессе, т. к. это внутри TF-графа. Обычно приходится писать вспомогательные Cirq-функции для проверки отдельных компонентов схемы вне графа.

С точки зрения qPCA, TFQ не предлагает ничего принципиально нового по сравнению с связкой «Cirq + классический ML», кроме, разве

что, оптимизации на GPU. Симулятор TFQ (Cirq) способен распараллеливать вычисления по сэмплам, но не ускоряет экспоненциально растущие задачи. Поэтому для чисто алгоритмического qPCA – Qiskit или Cirq могут быть даже удобнее. Однако, если задача включает квантовую нейросетевую подстройку для анализа данных – TFQ окажется мощным решением.

Интеграция с аппаратурой: TFQ сам по себе не взаимодействует напрямую с физическим железом. Он ориентирован на симуляцию квантовых моделей на CPU/GPU. Тем не менее, поскольку он построен на Cirq, теоретически можно использовать API Cirq для отправки схемы на Google Quantum Engine (если есть доступ) или другой бэкенд. Например, Cirq поддерживает Google Sycamore (для партнеров) – TFQ-модель можно модифицировать, чтобы выполнять inference (прогон готовой обученной схемы) на реальном процессоре, используя cirq.google.Engine для выполнения. Прямо «из коробки» TFQ не предлагает абстракций для работы с устройствами, что отличает его от Qiskit или PennyLane. Тем не менее, Google активно исследует гибридные вычисления: в 2023 г. анонсирована платформа TensorFlow Quantum Processing Units (QPU) – возможно, в будущем TFQ обретёт прозрачную отправку графов на квантовые ускорители. Пока же TFQ – инструмент в основном для симуляций и прототипирования. Если цель – запустить qPCA на железе IonQ или IBM, TFQ не будет лучшим выбором; лучше прибегнуть к PennyLane или Qiskit, которые имеют прямого доступа к этим устройствам.

4. Выводы

Проведённый обзор современных библиотек и инструментов квантового искусственного интеллекта показал, что, несмотря на раннюю стадию развития квантовых технологий, экосистема программных решений уже позволяет реализовывать квантовые версии классических алгоритмов, включая метод главных компонент. Инструменты, такие как Qiskit, PennyLane, TensorFlow Quantum и Cirq, предоставляют исследователям мощные средства для моделирования квантовых схем, проведения гибридных экспериментов и тестирования алгоритмов машинного обучения в квантовой постановке.

Каждая из рассмотренных библиотек имеет свои особенности: Qiskit отличается глубокой интеграцией с реальными квантовыми устройствами IBM, PennyLane удобен для гибридных моделей с классическими фреймворками, TensorFlow Quantum фокусируется на интеграции с нейронными сетями, а Cirq предоставляет высокоуровневый контроль над квантовыми схемами. Тем не менее, во всех случаях реализация qPCA требует творческого подхода,

комбинирования базовых операций и понимания теоретических принципов квантовых вычислений.

Ключевыми ограничениями остаются шумность современных квантовых устройств, ограниченное количество кубитов и отсутствие эффективных средств загрузки больших объёмов классических данных в квантовую форму. Однако даже при этих ограничениях гибридные подходы уже сегодня позволяют использовать преимущества квантовой обработки в отдельных компонентах моделей.

Таким образом, можно заключить, что библиотеки квантового ИИ, несмотря на технические сложности, открывают широкие перспективы для экспериментов с квантовыми алгоритмами анализа данных. Метод главных компонент, как один из фундаментальных методов в машинном обучении, становится отличной площадкой для демонстрации практической применимости квантовых вычислений. В ближайшие годы ожидается рост числа исследований и разработок, направленных на адаптацию классических алгоритмов к квантовой архитектуре, что постепенно будет приближать нас к эпохе реального квантового преимущества в задачах анализа данных.

Заключение

Квантовые вычисления стремительно развиваются и открывают новые горизонты в области искусственного интеллекта. Метод главных компонент, являясь краеугольным камнем анализа и обработки данных, получил новое развитие в квантовой интерпретации. Реализация qPCA с использованием современных библиотек демонстрирует не только теоретическую осуществимость квантового ускорения, но и практическую ценность гибридных подходов в условиях ограниченных квантовых ресурсов.

Исследование показало, что библиотеки, такие как Qiskit, PennyLane, TensorFlow Quantum и Cirq, уже сегодня предоставляют доступные инструменты для проектирования и тестирования квантовых алгоритмов, включая qPCA. Они становятся важным мостом между теоретическими достижениями и реальными экспериментами, открывая путь к разработке более сложных квантово-классических систем в будущем.

Несмотря на существующие технические барьеры, дальнейшее развитие программных фреймворков и квантового оборудования создаёт прочную основу для трансформации традиционного машинного обучения. В этом контексте метод главных компонент служит важным примером того, как классические аналитические инструменты могут быть переосмыслены и усилены с помощью квантовых технологий.

Список литературы

1. Голышева, Е. Н. Основные подходы к генерации изображений с помощью нейронных сетей / Е. Н. Голышева // Инновационная наука. – 2023. – №11/2.– С. 25–28.

2. Quantum Computing in Artificial Intelligence: a Review of Quantum Machine Learning Algorithms / O. F. Olaitan [et al.] // Path of Sience. – 2025. – № 5 – P. 7001–7009.

3. Grover, L. A fast quantum mechanical algorithm for database search / L. Grover // Proceedings of the twenty-eighth annual ACM symposium on Theory of computing. – 1996. – P. 212–219.

4. Ахромеева, Т. С. Искусственный интеллект как проблема культуры / Т. С. Ахромеева // Обсерватория культуры. – 2020. – № 3. – С. 228–241.

5. Mohammadisavadkoohi, E. A Systematic Review on Quantum Machine Learning Applications in Classification / E. Mohammadisavadkoohi, N. Shafiabady, J. Vakilian // IEEE Transactions on Artificial Intelligence. – 2026. – № 7. – P. 4–21.

6. Schuld, M. Quantum computing for pattern classification / M. Schuld, I. Sinayskiya, F. Petruccione // PRICAI 2014: Trends in Artificial Intelligence. – 2014. –vol. 8862. – P. 208–220.

Секция 8
САПР и цифровые технологии

Стратегическое проектирование

Н. В. Акимов, email: nikita261298@mail.ru[1]
В. В. Сокольников, email: svp_kitp@mail.ru[2]

[1] Воронежский государственный технический университет

Аннотация. *В настоящей статье исследуется эволюция и современное состояние стратегического проектирования как междисциплинарной методологии, интегрирующей принципы системного инжиниринга, управления жизненным циклом и цифрового моделирования. Акцент сделан на анализе технологического стека, включающего системы информационного моделирования (BIM), управления жизненным циклом изделий (PLM), интернета вещей (IoT) и инструменты на основе искусственного интеллекта. Рассматривается трансформация данной области от решения задач детального конструирования к управлению сквозными процессами создания стоимости на протяжении полного жизненного цикла сложных инженерных объектов. Делается вывод о формировании новой парадигмы, в которой стратегическое проектирование выступает ключевым элементом обеспечения устойчивого развития и технологического суверенитета.*

Ключевые слова: *стратегическое проектирование, цифровые двойники, BIM, PLM, генеративное проектирование, системный инжиниринг, жизненный цикл изделия, интероперабельность, прескриптивная аналитика, кибер-физические системы.*

Введение

Современная научно-техническая и производственная сфера претерпевает масштабную трансформацию, обусловленную процессами, объединенными под концепцией «Индустрия 4.0» или Четвертой промышленной революции [1]. Одним из центральных элементов этой трансформации является радикальное переосмысление роли и содержания проектной деятельности. Традиционное проектирование, исторически фокусировавшееся на создании графической и текстовой документации, регламентирующей геометрические параметры, материалы и технологические процессы изготовления отдельных компонентов, демонстрирует свою недостаточность перед лицом усложнения инженерных систем, сокращения инновационных циклов и

© Акимов Н. В., Сокольников В. В., 2026

ужесточения требований к экономической эффективности и экологической устойчивости.

В ответ на эти вызовы формируется новая методологическая дисциплина – стратегическое проектирование. Данная дисциплина может быть определена как систематизированная совокупность принципов, методов и инструментов, направленных на принятие оптимизационных решений на ранних стадиях жизненного цикла (ЖЦ) объекта, с учетом его поведения как элемента более крупной системы на всех последующих этапах: разработки, производства, эксплуатации, модернизации и утилизации. Таким образом, фокус смещается с проектирования изделия на проектирование его полного жизненного цикла и создаваемой им ценности.

Фундаментальной основой для этого сдвига послужила конвергенция нескольких ключевых технологических направлений: повсеместная цифровизация, развитие вычислительных мощностей и алгоритмов искусственного интеллекта (ИИ), распространение интернета вещей (IoT) и облачных вычислений [2]. Их синтез позволил перейти от статических цифровых моделей к динамическим, связанным с физическим миром кибер-физическим системам, что и составило технологическую суть стратегического проектирования.

1. Теоретико-методологические основы стратегического проектирования

Стратегическое проектирование базируется на синтезе нескольких устоявшихся научных и инженерных подходов:

1. Системный инжиниринг (Systems Engineering). Данный подход рассматривает проектируемый объект как сложную систему, взаимодействующую с внешней средой и состоящую из взаимосвязанных подсистем. Методы системного инжиниринга, такие как функциональное моделирование (например, с использованием IDEF0, N2-матриц), анализ требований и верификация, обеспечивают структурированное понимание проблемы, определение границ системы и установление корректных связей между ее элементами. В контексте стратегического проектирования системный инжиниринг задает архитектурный каркас, в который затем встраиваются конкретные технические решения [3].

2. Управление жизненным циклом (Life Cycle Management, LCM). Концепция LCM предполагает рассмотрение всех этапов существования объекта – от извлечения сырья для его производства до конечной утилизации – как единого, управляемого процесса. Стратегическое проектирование операционализирует LCM через количественные методы оценки жизненного цикла (Life Cycle Assessment, LCA) и расчета совокупной стоимости владения (Total Cost of Ownership, TCO).

Критически важным становится не минимизация первоначальной стоимости проекта, а оптимизация интегральных затрат и показателей эффективности на протяжении всего горизонта планирования [3].

3. Принцип «сквозной цифровой нити» (Digital Thread). Это концепция обеспечения непрерывного, согласованного потока данных, связывающего все этапы ЖЦ и всех участников процесса. Digital Thread преодолевает разрывы между изолированными «информационными островами» (CAD, CAE, CAM, ERP, MES, CMMS), обеспечивая прослеживаемость изменений, консистентность информации и основу для сквозной аналитики. Именно Digital Thread делает возможным переход от разрозненных файлов к единой управляемой цифровой модели объекта [2].

2. Технологический стек стратегического проектирования: компоненты и взаимодействие

Реализация методологических принципов возможна только при наличии развитого технологического стека, ядро которого образуют несколько взаимодополняющих платформ:

1. Информационное моделирование (BIM/IM). Технология Building Information Modeling (BIM), изначально развивавшаяся в строительной отрасли, эволюционировала в универсальную парадигму Information Modeling (IM). Ее суть заключается в создании не просто трехмерной геометрической модели, а структурированной базы данных, где каждый виртуальный объект (элемент) обладает набором атрибутов (данных). Стандартная расширяемая структура данных (например, в формате IFC – Industry Foundation Classes) позволяет хранить информацию о:

- Геометрии и пространственных отношениях;
- Временных параметрах и последовательности работ;
- Стоимостных характеристиках;
- Эксплуатационных и энергетических свойствах;
- Данных для управления объектом на всем жизненном цикле.

Таким образом, BIM-модель становится единым источником истины (Single Source of Truth), используемым для многодисциплинарной координации, выявления коллизий, планирования ресурсов и симуляции эксплуатационных сценариев еще на виртуальной стадии [5].

2. Системы управления жизненным циклом изделия (PLM). В то время как BIM/IM фокусируется на пространственно-физической модели и связанных с ней данных, PLM-системы (Product Lifecycle Management) обеспечивают управление бизнес-процессами, документацией, изменениями и конфигурациями на протяжении всего ЖЦ. Они являются платформой для реализации Digital Thread. Современные PLM-

платформы (например, Teamcenter, Windchill, 3DEXPERIENCE) обеспечивают:

- Управление структурой изделия (Product Structure Management);
- Управление изменениями и конфигурациями (Change & Configuration Management);
- Интеграцию с CAD/CAE/CAM-системами и корпоративными ERP-системами;
- Управление проектами и портфелями проектов.

Взаимодействие BIM и PLM создает комплексную среду: BIM предоставляет детализированную инженерную модель, а PLM управляет процессами работы с этой моделью, ее версиями, одобрениями и связью с производством и снабжением [2].

3. Интернет вещи (IoT) и кибер-физические системы (CPS). Данные технологии обеспечивают материализацию цифровых моделей и создание петли обратной связи. IoT-сенсоры, встроенные в физический объект (оборудование, строительную конструкцию), непрерывно собирают телеметрическую информацию: вибрацию, температуру, напряжение, нагрузки, расходы и т.д. Эта информация в реальном времени передается в цифровую модель. Таким образом, статичная модель трансформируется в цифровой двойник (Digital Twin) – динамическую виртуальную копию, отражающую текущее состояние и историю своего физического аналога [6, 7].

Цифровой двойник представляет собой не единый файл, а сложную программно-аппаратную систему, включающую:

- Модель данных (физические и функциональные характеристики);
- Модели симуляции и анализа (конечно-элементные, многодисциплинарные, потоковые);
- Алгоритмы машинного обучения для анализа данных и прогнозирования;
- Интерфейсы для визуализации и взаимодействия.

Через IoT и цифрового двойника реализуется принцип кибер-физической системы, где физические и цифровые процессы тесно интегрированы и координируются компьютерными алгоритмами [2, 6].

4. Искусственный интеллект и генеративное проектирование. ИИ кардинально расширяет возможности проектировщика на двух основных уровнях [4].

– Генеративное проектирование (Generative Design) – это итеративный процесс, в котором проектировщик задает целевые параметры (функциональные требования, нагрузки, ограничения по габаритам, материалы, методы производства) и критерии оптимизации

(минимизация массы, максимизация жесткости, снижение стоимости). Алгоритмы (часто на основе эволюционных вычислений или обучения с подкреплением) генерируют множество альтернативных вариантов геометрии, удовлетворяющих этим условиям. Результатом являются часто нетривиальные, бионические формы, оптимальные с точки зрения распределения материала, которые было бы невозможно разработать интуитивно или с помощью традиционных параметрических методов;

– Предиктивная и прескриптивная аналитика – на основе данных от цифрового двойника и исторических данных ИИ-модели способны предсказывать отказы, остаточный ресурс, пиковые нагрузки. Более того, системы прескриптивной аналитики предлагают конкретные рекомендации по действиям: когда и какой компонент обслужить, как скорректировать режим работы для экономии энергии, какая модернизация даст максимальный экономический эффект.

3. Реализация принципов стратегического проектирования на различных этапах жизненного цикла

Рассмотрим, как описанный технологический стек применяется на ключевых фазах.

1. Концептуальное и техническое проектирование. На этой стадии стратегическое проектирование проявляется в проведении многовариантного анализа (Multi-Disciplinary Optimization, MDO). Используя параметрические модели и облачные вычисления, инженеры могут проанализировать тысячи сценариев, оценивая влияние различных архитектурных и конструктивных решений не только на прочность и вес, но и на технологичность изготовления, стоимость эксплуатации, энергоэффективность и углеродный след. Генеративное проектирование активно используется для создания оптимальных несущих структур, теплообменников, кронштейнов [2].

2. Производство и строительство. Цифровая модель (BIM) становится основой для планирования производственных операций (4D), расчета потребностей в материалах (5D) и подготовки управляющих программ для станков с ЧПУ или роботов. Технология «цифрового заказа» (Digital Fabrication) позволяет напрямую передавать данные из модели на производственное оборудование, минимизируя ошибки. На стройплощадке использование планшетов с актуальными BIM-моделями, дополненной реальности (AR) для наложения модели на реальный объект и сканирования для контроля соответствия (Scan-to-BIM) становятся стандартными практиками [5].

3. Эксплуатация и обслуживание. Это область наиболее активного использования цифрового двойника [6, 7]. Модель, связанная с IoT-данными, позволяет:

– Проводить виртуальные тренировки персонала;
– Осуществлять мониторинг состояния в реальном времени;
– Переходить от планово-предупредительного к предиктивному обслуживанию, выполняя ремонт именно тогда, когда это необходимо, что снижает простой и затраты;
– Симулировать последствия модернизации или изменения режимов работы перед их реализацией в физическом мире;
– Управлять энергопотреблением интеллектуальных зданий и инфраструктур.

4. Утилизация и рециклинг. Стратегическое проектирование учитывает и конечную стадию жизненного цикла. В модель могут быть заложены данные о материалах компонентов, способах их соединения (разъемные, сварные, клееные), что облегчает планирование процессов демонтажа и переработки в конце срока службы, способствуя циркулярной экономике [3].

4. Сравнительный анализ парадигм проектирования

Для наглядного представления трансформации целесообразно провести сравнительный анализ (Таблица 1).

Сравнительный анализ ключевых параметров традиционного и стратегического подходов к проектированию выявляет фундаментальные различия по ряду критериев.

Объект управления в рамках традиционного подхода сосредоточен на документации и геометрической форме объекта. В противоположность этому, стратегическое проектирование управляет данными и процессами, охватывающими полный жизненный цикл.

Основной инструментарий также различен: традиционное проектирование опирается на CAD-системы (2D/3D), в то время как стратегическое использует интегрированные платформы, объединяющие технологии BIM, PLM, IoT и искусственный интеллект.

Основа для принятия решений в первом случае формируется на базе инженерного опыта, нормативных требований и линейных расчетов. Стратегический подход, напротив, базируется на многовариантном моделировании, симуляции и предиктивной аналитике, использующей большие данные (Big Data).

Критерий успеха проекта при традиционном проектировании заключается в соответствии техническому заданию в рамках установленных бюджета и сроков проектирования. Стратегическое проектирование нацелено на оптимизацию ключевых показателей, таких как совокупная стоимость владения (TCO), надежность, безопасность и устойчивость, на всем горизонте жизненного цикла объекта.

Взаимодействие дисциплин в традиционной парадигме чаще является последовательным или слабо согласованным, что приводит к частым коллизиям. Стратегический подход предполагает параллельное, координируемое взаимодействие через единую информационную модель, что позволяет выявлять и разрешать коллизии на ранних стадиях.

Взаимосвязь с производством и строительством в первом случае сводится к передаче пакета документации, что несет риски интерпретационных ошибок. Второй подход обеспечивает непрерывный обмен данными по принципу «цифровой нити» (Digital Thread) и цифровое сопровождение производственных процессов.

Итоговый результат традиционного проектирования — это комплект чертежей и спецификаций. Результатом стратегического проектирования становится цифровой двойник и управляемая им производственно-эксплуатационная экосистема.

Заключение

Стратегическое проектирование, подкрепленное комплексом цифровых технологий, представляет собой не просто новый набор программных инструментов, а формирующуюся научно-техническую дисциплину и производственную парадигму. Оно знаменует переход от фрагментированного, документо-ориентированного процесса к целостному, data-driven управлению созданием, существованием и развитием сложных инженерных систем. Преодоление существующих технологических, организационных и нормативных барьеров является актуальной задачей для промышленности, научного сообщества и регуляторов. Успешное внедрение принципов стратегического проектирования становится критическим фактором обеспечения технологического суверенитета, конкурентоспособности на мировом рынке и достижения целей устойчивого развития в XXI веке.

Список литературы

1. Шваб, К. Четвертая промышленная революция / К. Шваб ; пер. с англ. – М. : Эксмо, 2016. – 208 с.

2. Боровков, А. И. Современные методы и технологии проектирования в машиностроении: цифровые двойники, генеративный дизайн, аддитивные технологии : учебное пособие / А. И. Боровков [и др.] ; под ред. А. И. Боровкова. – СПб. : Изд-во Политехн. ун-та, 2021. – 252 с.

3. Тозик, А. В. Управление жизненным циклом продукции (CALS-технологии) : учебник / А. В. Тозик. – М. : Инфра-М, 2022. – 336 с.

4. Кабанов, А. Н. Предиктивная аналитика и машинное обучение для технического обслуживания промышленного оборудования :

коллективная монография / А. Н. Кабанов [и др.] ; отв. ред. А. Н. Кабанов. – М. : Инфра-Инженерия, 2023. – 198 с.

5. Калинин, М. А. BIM-технологии: информационное моделирование зданий и сооружений : учебник и практикум / М. А. Калинин, Д. А. Яковлев. – М. : Юрайт, 2023. – 383 с.

6. Рыжов, П. А. Разработка и применение цифровых двойников в высокотехнологичных отраслях промышленности / П. А. Рыжов [и др.] // Информационные технологии и вычислительные системы. – 2022. – № 4. – С. 64-76.

7. Тарасов, С. П. Цифровые двойники технологических процессов и производств : от концепции к практике / С. П. Тарасов, В. И. Скоробогатов // Автоматизация и современные технологии. – 2021. – № 11. – С. 470-475.

Анализ возможностей Siemens NX 10 и Компас-3D для автоматизации проектирования оснастки

В.С. Бакланова, e-mail: dexp51477@gmail.com
И.А. Караваева, e-mail: kirina205@gmail.com
В.А. Рыжков

Воронежский государственный университет

Аннотация. *Аннотация. Статья посвящена сравнительному анализу двух систем автоматизированного проектирования – Siemens NnpoекX 10 и Компас v23. В рамках учебного проекта были созданы одинаковые детали с использованием обеих систем. Проведен анализ интерфейса, удобства использования, гибкости, доступности API и общей направленности продуктов. Сделан вывод о сильных и слабых сторонах каждой системы.*

Ключевые слова: *Siemens NX, Компас-3D, САПР, проектирование, интерфейс, API, гибкость.*

Введение

Современные CAD-системы требуют глубокой интеграции с производственными процессами. В исследовании сравниваются Siemens NX 10 (NX Open на C++) – промышленный стандарт для сложных изделий и Компас-3D v23 (API на C#) – отечественное решение с поддержкой ГОСТ.

Siemens NX 10 зарекомендовал себя как мощная платформа с широкими возможностями для параметрического моделирования и интеграции с другими инженерными решениями. Интерфейс программы ориентирован на промышленное использование, включает развитую систему горячих клавиш и кастомизацию рабочих процессов. Серьезным преимуществом является открытость API и наличие многочисленных модулей, облегчающих разработку сложных изделий.

Компас v23, в свою очередь, представляет собой российскую разработку, активно развивающуюся в последние годы. Программа предлагает интуитивно понятный интерфейс, адаптированный под отечественные стандарты проектирования. Для образовательных целей и средних предприятий Компас демонстрирует приемлемый уровень функциональности и гибкости. Тем не менее, ему пока не хватает глубокой интеграции с внешними средами, а также зрелости и надежности в сравнении с NX.

© Бакланова В. С., Караваева И. А., Рыжков В. А. 2026

Введение в метод трансфера функций

Впервые метод трансфера функций был описан в 2012 году. Метод используется для нахождения плотных соответствий между трехмерными объектами в виде пар точек. Вместо работы над попарными соответствиями прямо, метод трансфера функций рассматривает перенос функций с вещественным значением с одного объекта на другой с помощью компактной матрицы трансфера, нахождение которой и является основной задачей метода. После нахождения матрицы трансфера с ее помощью возможен переход к попарным соответствиям. Далее рассматриваются некоторые понятия, требуемые для краткого описания метода.

Siemens NX 10 предоставляет мощные инструменты для разработки пользовательских интерфейсов (UI) прямо внутри системы. Это позволяет создавать удобные диалоговые окна, панели управления и даже полноценные приложения, интегрированные в NX.

Визуальный редактор NX для создания диалоговых окон поддерживает – различные элементов управления (кнопки, списки, поля ввода, таблицы); гибкую компоновку (вкладки, группы, динамическое изменение интерфейса)

Пример результата работы с данным инструментом можно увидеть на рис. 1.

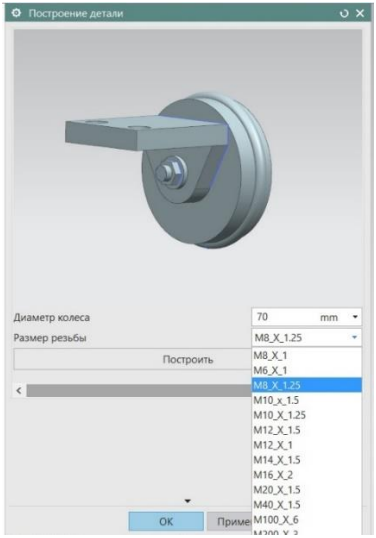


Рис. 1. Результат работы в визуальном редакторе Siemens NX

В отличие от NX, Компас-3D не имеет встроенных средств для разработки графического интерфейса.

Ограничения – нет визуального редактора диалоговых окон – интерфейс приходится писать вручную через WinForms или WPF; сложность интеграции – требуется отдельная разработка UI в Visual Studio.

Так, NX 10 дает полный контроль над интерфейсом, позволяя создавать профессиональные CAD-инструменты с минимальными затратами, а Компас-3D требует дополнительных усилий для разработки UI, что замедляет процесс автоматизации.

Создание модели на языках C++ (в NX) и C# (в Компас-3D) имеют как сходства, так и отличия. И в самих версиях библиотечных методов сигнатура может меняться; тем не менее, простейшая логика создания эскиза, простейших геометрических фигур, булевых операций, и пр. остаётся схожей. Ниже приведены фрагменты кода (листинг 1 и листинг 2) на данных языках для сравнения.

Листинг 2

Код для построения параллелепипеда в Siemens NX

```
int NXOPEN_dialogue::createBlock(
    const NXOpen::Point3d& basePoint,
    double length, double width, double height,
    UF_FEATURE_SIGN sign,
    tag_t* result_body
)
{   try
    {
        std::string length_str = std::to_string(length);
        std::string width_str = std::to_string(width);
        std::string height_str = std::to_string(height);
        tag_t block;
        double corner[3] = { basePoint.X, basePoint.Y,
basePoint.Z };
        char* block_dims[3] = {
            const_cast<char*>(length_str.c_str()),
            const_cast<char*>(width_str.c_str()),
            const_cast<char*>(height_str.c_str())
        };
        int status = UF_MODL_create_block1(sign, corner,
block_dims, &block);
        if (status != 0) {
            char fail_message[133];
            UF_get_fail_message(status, fail_message);
            uc1601(fail_message, 1);
            return status;
        }
```

```
        if (result_body != NULL) UF_MODL_ask_feat_body(block,
result_body);
        return 0;
    }
    catch (exception& ex)
    { NXOPEN_dialogue::theUI->NXMessageBox()->Show("Block
Creation", NXOpen::NXMessageBox::DialogTypeError, ex.what());
        return -1; }}
```

Листинг 3

Код для создания параллелепипеда на языке C# в Компас-3D

```
KompasObject kompas =
(KompasObject)Activator.CreateInstance(Type.GetTypeFromProgID(
"KOMPAS.Application.5"));
kompas.Visible = true;
kompas.ActivateControllerAPI();
ksDocument3D doc3D = (ksDocument3D)kompas.Document3D();
doc3D.Create(false, true);
ksPart part =
(ksPart)doc3D.GetPart((short)Part_Type.pTop_Part);
ksEntity mainSketch =
(ksEntity)part.NewEntity((short)Obj3dType.o3d_sketch);
ksSketchDefinition mainSketchDef =
(ksSketchDefinition)mainSketch.GetDefinition();
mainSketchDef.SetPlane(part.GetDefaultEntity((short)Obj3dType.
o3d_planeXOY));
mainSketch.Create();
ksDocument2D mainDrawing =
(ksDocument2D)mainSketchDef.BeginEdit();
mainDrawing.ksLineSeg(-26, -21, 26, -21, 1);
mainDrawing.ksLineSeg(26, -21, 26, 21, 1);
mainDrawing.ksLineSeg(26, 21, -26, 21, 1);
mainDrawing.ksLineSeg(-26, 21, -26, -21, 1);
mainSketchDef.EndEdit();
ksEntity mainExtrude =
(ksEntity)part.NewEntity((short)Obj3dType.o3d_baseExtrusion);
ksBaseExtrusionDefinition mainExtrudeDef =
(ksBaseExtrusionDefinition)mainExtrude.GetDefinition();
mainExtrudeDef.SetSketch(mainSketch);
mainExtrudeDef.directionType = (short)Direction_Type.dtNormal;
mainExtrudeDef.SetSideParam(true, (short)End_Type.etBlind,
14.0);
mainExtrude.Create();
```

Результаты работы представлены на рис. 2 и рис. 3.

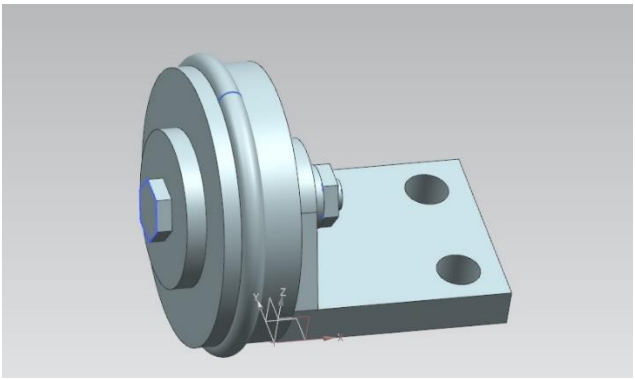


Рис. 2. Результат создания детали в Siemens NX

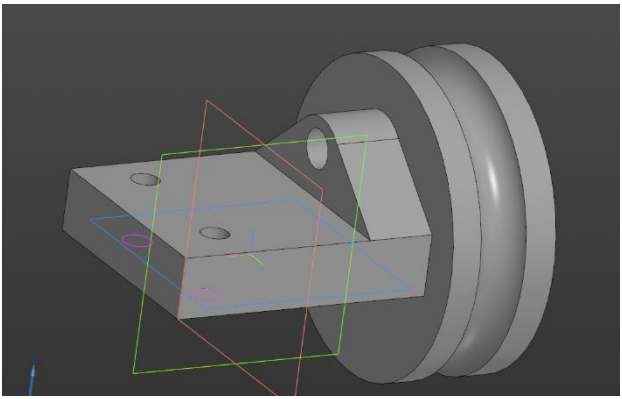


Рис. 3. Результат создания детали в Компас 3D

Заключение

Таким образом, NX и Компас занимают разные ниши: NX используется крупными промышленными предприятиями, в то время как Компас подходит для образовательных учреждений и малого бизнеса. В Siemens NX вложены значительные ресурсы, благодаря чему система достигла высокой степени зрелости и масштабируемости.

Список литературы

1. Siemens PLM Software. NX Open Programmer’s Guide [Электронный ресурс]. – Режим доступа: https://docs.plm.automation.siemens.com (ограниченный доступ).

2. АО «АСКОН». Руководство разработчика Kompas-3D. API v23 [Электронный ресурс]. – СПб.: АСКОН, 2025. – 1200 с. – Режим доступа: https://kompas.ru/api.

3. Махортов, С. Д. Основы автоматизации проектно-конструкторских работ в CAD/CAM/CAE системах / С. Д. Махортов, М. И. Чижов. – Воронеж: Изд-во Воронеж. гос. ун-та, 2023. – 215 с.

4. Чижов, М. И. Современные системы автоматизированного проектирования: сравнительный анализ и применение / М. И. Чижов // Информационные технологии в проектировании и производстве. – 2024. – № 3. – С. 45–52.

5. Грашин, А. А. Программирование в CAD-системах. От скриптов к полноценным приложениям / А. А. Грашин. – М.: ДМК Пресс, 2022. – 328 с.

6. Сазонов, В. В. Автоматизация проектирования технологической оснастки в среде Siemens NX / В. В. Сазонов // CAD/CAM/CAE Observer. – 2023. – Т. 15, № 4. – С. 30–38

Разработка пользовательского интерфейса для построения задачи расчета стрежневой системы

А. М. Гурьева, email: gureva_am@cs.vsu.ru[1]
А. С. Троценко, email: trotsenko93@mail.ru[1]

[1] Воронежский государственный университет

Аннотация. *В статье рассматривается подход к разработке пользовательского интерфейса и архитектуры специализированной системы автоматизированного проектирования (САПР), ориентированной на решение задачи расчета стержневых систем. Описываются архитектурные решения и особенности реализации интерфейса, основанные на паттерне MVC, которые могут быть использованы при разработке других узконаправленных инженерных САПР.*

Ключевые слова: *стержневая система, GUI, инженерное ПО, MVC, Qt, САПР.*

Введение

В настоящее время наблюдается рост интереса к разработке специализированных, узконаправленных систем автоматизированного проектирования (САПР), ориентированных на решение конкретных расчетных задач. В отличие от универсальных САПР, такие системы позволяют сократить избыточный функционал, упростить архитектуру и повысить эффективность работы за счёт адаптации под особенности конкретной предметной области.

Актуальность данной работы обусловлена необходимостью разработки обобщённого интерфейса, позволяющего ускорить процесс создания специализированных САПР. Ключевую роль в такой работе играет формирование универсального шаблона программного решения для выполнения поставленной задачи.

В работе рассматривается задача расчета стержневых систем. Данный класс задач является хорошо изученным и наглядным, что позволяет сосредоточиться не на вычислительных аспектах, а на принципах организации пользовательского интерфейса и архитектуры программного обеспечения.

Целью работы является разработка подхода к построению пользовательского интерфейса и архитектуры специализированного инженерного программного обеспечения, ориентированного на

© Гурьева А. М., Троценко А. С., 2026

интерактивное формирование расчетной модели и визуализацию результатов. Предлагаемые архитектурные и интерфейсные решения рассматриваются как элементы обобщенного шаблона разработки специализированных САПР и могут быть использованы при создании программных средств для решения различных задач инженерного анализа.

1. Архитектура и структура приложения

Разрабатываемое приложение построено на основе архитектурного паттерна MVC (Model-View-Controller), обеспечивающего разделение ответственности между расчетной логикой, пользовательским интерфейсом и управляющими компонентами. Использование данного паттерна позволило изолировать бизнес-логику расчетов от графического представления, что упростило расширение функционала и повторное использование компонентов приложения.

Для наглядного представления архитектуры и взаимодействия основных частей системы использована UML-диаграмма компонентов, отражающая структуру приложения.

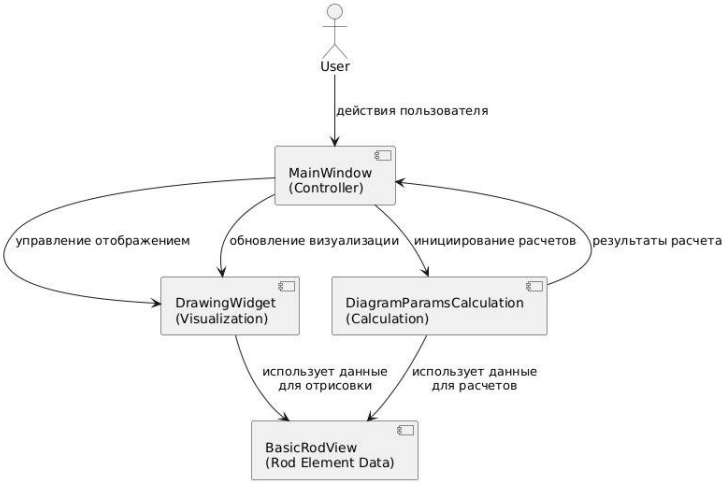


Рис. 1. UML-диаграмма компонентов приложения для расчета стержневой системы

Для детализации внутренней структуры приложения и описания основных сущностей расчетной и визуальной моделей использована UML-диаграмма классов. Диаграмма отражает ключевые классы системы и их взаимосвязи, обеспечивающие хранение параметров

элементов стержневой системы, выполнение расчетов и графическую визуализацию.

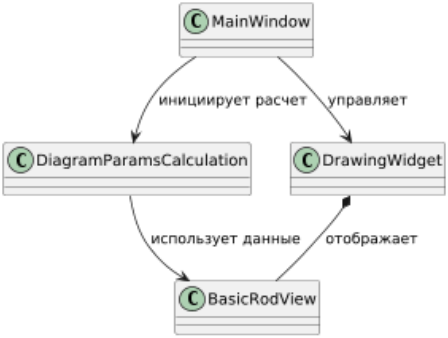


Рис. 2. UML-диаграмма классов приложения для расчета стержневой системы

В качестве архитектурного решения выбрана монолитная модель. Такой подход обеспечивает высокую производительность, поскольку все операции выполняются в рамках одного процесса без дополнительных затрат ресурсов на взаимодействие между модулями [1]. Данное свойство является важным для инженерного программного обеспечения, ориентированного на выполнение большого количества вычислительных операций в ограниченные временные интервалы.

В соответствии с паттерном MVC логическая структура приложения разделена на три слоя.

Слой Model реализует описание элементов стержневой системы и вычислительную логику, распределённую между классами хранения параметров и расчётным модулем.

Слой View отвечает за визуализацию расчетной модели и пользовательский интерфейс, включая графическое отображение стержней и элементов управления. Данный слой не содержит расчетной логики и выполняет функции отображения текущего состояния модели.

Слой Controller осуществляет обработку пользовательских действий, передачу данных в модель и обновление представления в соответствии с результатами расчетов.

Для организации взаимодействия между слоями используется механизм сигналов и слотов фреймворка Qt [2]. Такой подход обеспечивает слабую связанность компонентов и позволяет автоматически инициировать перерасчет и обновление визуализации при изменении параметров расчетной модели.

2. Реализация пользовательского интерфейса

Пользовательский интерфейс разработан с использованием виджетов Qt [3]. Реализован механизм интерактивного построения расчетной модели, при котором пользователь может изменять параметры элементов стержневой системы непосредственно на графическом представлении. Любое изменение параметров автоматически инициирует перерасчет и обновление визуализации, что обеспечивает оперативную обратную связь и повышает наглядность работы с моделью.

Нижняя часть главного окна приложения предназначена для управления параметрами. В данной области размещены элементы ввода параметров выбранного участка стержневой системы (длина, радиус, приложенная сила, продольная сила), а также общих параметров системы, таких как модуль упругости и сила, приложенная к концу системы. Активность элементов управления зависит от текущего состояния расчетной модели, что снижает вероятность ошибочных действий пользователя.

В правой части окна размещена область вывода результатов расчета. После выполнения вычислений результаты автоматически отображаются в табличной форме.

Для повышения удобства и безопасности работы реализован механизм подтверждения удаления элементов расчетной модели. При попытке удаления сегмента стержневой системы отображается модальное окно подтверждения, предотвращающее случайное удаление элементов.

Визуализация расчетной модели обновляется динамически с учетом действий пользователя. Например, при удалении сегмента оставшиеся элементы автоматически выравниваются, заполняя освободившееся пространство, что обеспечивает корректное и наглядное представление структуры стержневой системы.

Ключевые особенности разработанного пользовательского интерфейса:

- Интерактивность – возможность динамического редактирования расчетной модели с помощью мыши и автоматическое обновление визуализации;
- Наглядность – сочетание графического представления стержневой системы и табличного вывода результатов;
- Интуитивность – управление состояниями элементов интерфейса, подсветка выбранных элементов и использование модальных окон.

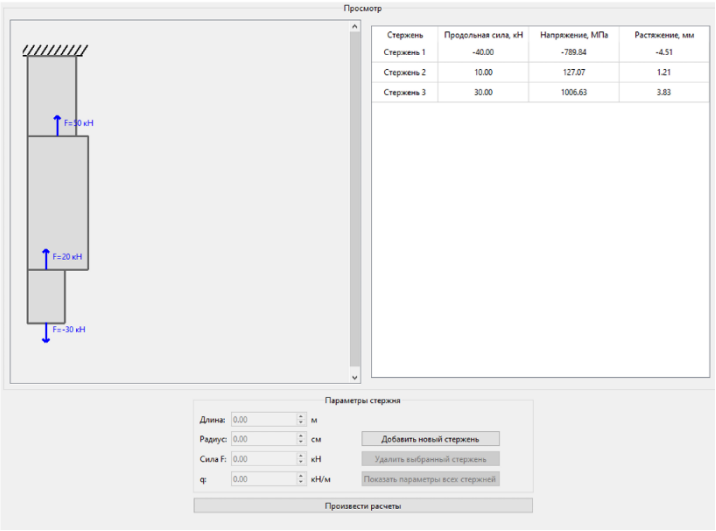


Рис. 3. Главное окно приложения и визуализация стержневой системы

3. Представление результатов расчета

Результаты расчетов отображаются в табличной форме, где для каждого участка стержневой системы приводятся числовые значения вычисленных параметров. Табличное представление обеспечивает однозначную интерпретацию результатов и упрощает анализ параметров отдельных элементов расчетной модели.

Использование таблиц позволяет применять полученные данные как для визуального анализа, так и для последующей верификации корректности расчетной схемы. Табличный вывод может рассматриваться как промежуточный или дополнительный этап перед построением эпюр.

Стержень	Продольная сила, кН	Напряжение, МПа	Растяжение, мм
Стержень 1	-40.00	-789.84	-4.51
Стержень 2	10.00	127.07	1.21
Стержень 3	30.00	1006.63	3.83

Рис. 4. Пример представления результатов расчета

Заключение

В рамках работы был сформирован и реализован подход к построению пользовательского интерфейса и архитектуры специализированного инженерного программного обеспечения для решения задачи расчета стержневой системы. Рассмотренное решение ориентировано на интерактивное формирование отображения стержня и наглядное представление результатов.

Использование изученной задачи расчета стержневых систем позволило сосредоточиться на обобщенных принципах разработки САПР, включая разделение ответственности между компонентами, организацию пользовательского интерфейса и обеспечение визуального отображения. Предложенные архитектурные и интерфейсные решения не зависят от конкретного метода расчета и могут быть адаптированы для широкого круга задач инженерного анализа.

Полученные результаты могут быть использованы при разработке узконаправленных САПР, ориентированных на автоматизацию инженерных расчетов. Дальнейшее развитие работы может быть связано с расширением предлагаемого шаблона разработки, а также с его применением к другим классам инженерных задач, таких как задачи кручения и изгиба стержней.

Список литературы

1. Монолитная архитектура: преимущества [Электронный ресурс]. – Режим доступа: https://habr.com/ru/companies/haulmont/articles/758780/ (дата обращения: 08.01.2026).

2. Qt Documentation. Signals & Slots [Электронный ресурс]. – Режим доступа: https://doc.qt.io/qt-6/signalsandslots.html (дата обращения: 14.01.2026).

3. Qt и C++. Введение в виджеты Qt [Электронный ресурс]. – Режим доступа: https://metanit.com/cpp/qt/1.4.php (дата обращения: 17.01.2026)

Автоматизированная система принятия решений в процессах в аграрной промышленности

В. С. Деханова, email: dehanovavika24@gmail.com [1]
В.В. Сокольников, email: Svp_kitp@mail.ru [2]
М.Г. Дибров, email: Mihail.dibrow@yandex.ru [2]

[1] Воронежский государственный технический университет

Аннотация. *В статье рассматривается проектирование мобильного приложения для поддержки принятия решения в аграрном секторе на основе мониторинга погодных условий и семантического сопоставления фактических данных с эталонными нормами культур. Приложение позволяет фермерам анализировать состояние посадок и принимать оперативное решение для снижения рисков и повышения урожайности.*

Ключевые слова: *цифровизация АПК, поддержка принятия решений, агроклиматический мониторинг, сельское хозяйство.*

Введение

Агропромышленный комплекс (АПК) всегда являлся более уязвимой и сложно автоматизированной отраслью за счет значительной зависимости от погодных условий. Качество и количество урожайности напрямую определяются такими показателями как влажность почвы и воздуха, температура и атмосферное давление, количество световых часов и осадков. Традиционные меры минимизации производственного ущерба, основанные на метеорологическом прогнозировании, теряют свою эффективность из-за несоответствия скорости реакции на сменяющуюся обстановку, а также несвоевременных рекомендаций в обновленных условиях. Это приводит к постоянным рискам полных или частичных потерь, что закономерно обуславливает колебания на рынке.

В тоже время в мире формируется тенденция на цифровизацию всех сфер, в рамках которой данные приобретают значительный вес и становятся одним из ключевых ресурсов развития, что также подтверждает российская государственная стратегия [1]. Цифровизация АПК подразумевает реализацию автоматизированных систем ориентированных на мониторинг состояний важных показателей, формирование масштабного хранилища справочных материалов касательно эталонных условий для различных культур, на основе

© Деханова В.С., Сокольников В.В., Дибров М.Г., 2026

которых после семантического сопоставления с текущими данными, формируются целевые рекомендации.

В данной статье будет описана реализация цифрового продукта в виде информационной системы для сбора данных и поддержки принятия решений в аграрных процессах. Внимание уделяется построению структуры обработки данных с учетом различных источников: внешних - метаданные формата JSON от стороннего сервиса через систему API; сформированных справочников с эталонными агротехническими нормативами.

Таким образом цель разрабатываемого проекта – создание информационного инструмента для пользователя на основе возможностей цифровых технологий, который позволит повысить эффективность принятия решений, основанных на статистических данных в аграрном секторе, и снизит риски, связанные с высокой изменчивости погодных условий.

1. Цифровая модель и функциональные требования

1. Концепция цифрового поля и объектная модель предметной области

Разработка автоматизированной системы принятия решений для аграрной промышленности основана на парадигме создания «цифрового двойника». Данная модель подразумевает создание виртуальной копии объекта и синхронизацию данных за счет использования цифровых технологий для постоянного мониторинга показателей. Данную методику также используют в настоящее время на территории Российской Федерации. Создано порядка 14 пробных «цифровых полей» распределённых под разные нужны сельского хозяйства для испытаний наиболее лучшего использования технологий [2].

Разработанная модель представляет следующие ключевые сущности и связь между ними (рис.1):

– Location – центральный объект, включающий описание земельного участка с атрибутами: тип почвы, наименование, площадь (г/м3), территориальное месторасположение по регионам Российской Федерации;

– Crop – выступает в роли нормативно-справочного компонента системы. Представляет собой статические данные, характеризующие эталонны для определенной культуры (диапазон оптимальной температуры, влажности воздуха и почвы)

– Planting – представляет технологическую операцию, связывая локацию с культурой, задавая сроки посева и сбора урожая. Данную сущность также можно охарактеризовать как центральный объект оперативного управления;

– WetherData – обеспечивает адаптацию и хранение динамических погодных показателей через интеграцию со сторонним API сервисом.

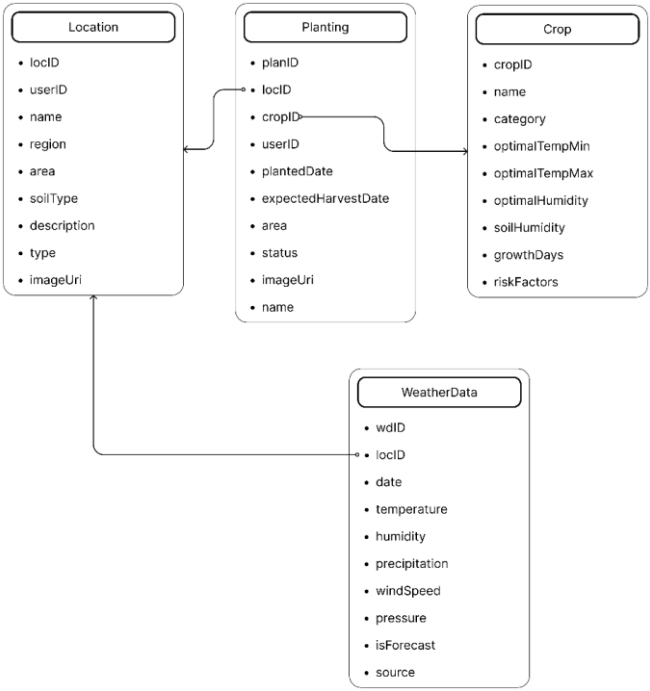


Рис. 1. Диаграмма объектной модели данных информационной системы

2. Функциональные требования

Основываясь на объектной модели, можно сформулировать следующие требования к АСППР:

– Требование постоянного мониторинга. Система должна обеспечивать непрерывное получение данных о состояниях основных показателей, что обеспечит своевременную реакцию на смену обстановки и предложит целевые рекомендации для поддержания эталонных показателей культур;

– Требование к автоматизированному анализу. На основе полученных данных необходимо реализовать автоматизированный анализ, выявляющий отклонения от эталонов, оценивающий риски и генерирующий рекомендации с типом действий и приоритетности;

– Требование к масштабируемости. Архитектура системы должна быть спроектирована с условием дальнейшего масштабирования без потери качества реализованного функционала. А также гарантировать возможность замены синтетических данных на модуль IoT интеграции без перепроектирования ядра системы.

2. Подпроцесс мониторинга

1. Архитектура подпроцесса и многоуровневая модель данных

Полнота и достоверность анализируемого контекста для принятия решений – необходимая составляющая системы, от которой напрямую завесит эффективность автоматизированной системы поддержки принятия решений (АСППР). В системе реализована трехуровневая модель интеграции данных:

– Статический уровень. Данные организованы в виде встроенных справочников регионов и культур. Последние является основой для семантического сравнения, поскольку задают эталон;

– Динамический внешний уровень. Обеспечивает поступление актуальных прогнозных данных со стороннего сервиса, такого как OpenWeatherMap API. Данные поступают в формате JSON, в системе же преобразуются в объектную модель WeatherData.

– Расчетный (синтетический) уровень. Формируется модулем имитационного моделирования для параметров, прямое измерение которых в прототипе на данном этапе разработки не реализовано. Несмотря на отсутствие данных от реальных IoT- датчиков, архитектура модуля подразумевает изолированное генерирование статистических релевантных данных, которые а) обеспечиваю полноценное тестирование; б) предусматривает простую замену на адаптер приема данных с устройств интернет вещей – датчиков.

2. Конвейер преобразования и агрегации данных

В реализуемой системе технологическим ядром подпроцесса выступает конвейер, который позволяет обеспечить сквозное трансформирование данных от внешних источников (полутруктурированных JSON-ответов) к структурным данным, организованным в реляционную модель.

Для получение актуальных погодных показателей системе необходимо отправить запрос к API. В прототипной реализации используется статические географические координаты, соответствующие городу Воронеж (широта и долгота), это сделано для упрощения тестирования. Архитектура предусматривает дальнейшую интеграцию с внешним API для получения GPS-координат, что позволит персонализировать систему. Листинг конечной точки для запроса текущей работы представлена ниже:

```
interface WeatherApi {
    @GET("data/2.5/weather")
    suspend fun getCurrentWeather(
        @Query("lat") lat: Double,
        @Query("lon") lon: Double,
        @Query("appid") apiKey: String,
        @Query("units") units: String = "metric",
        @Query("lang") lang: String = "ru"
    ): WeatherDataDTO
}.
```

Полученный от API ответ в формате JSON автоматически преобразуется в объект DTO (Data Transfer Object)- связующим звеном трансформации полуструктурированных данных JSON в типизированную внутреннюю модель. Часть функции-маппера для преобразования DTO представлена ниже:

```
fun WeatherDataDTO.toWeatherData(locID: Int): WeatherData {
 return WeatherData(
        locID = locID,
        date = dt * 1000, //конвертация времени
        temperature = main.temp,
        humidity = main.humidity,
        //…другие поля
)}.
```

После структуризации данные с уровня API, они объединяются с данными других уровней – справочными эталонами и синтетическими показателями. В результате система получает единый семантический контекст для анализа, который отображает фактическое состояние показателей на данный момент с учетом территориальное расположения и служит входом для модуля принятия решений.

3. Методы принятия решений на основе агрегированных данных

1. Детерминированный метод анализа отклонений

В системе реализован модель RecommendationGenerator, отвечающий за формирование правил. Его работа заключается в непосредственной инициализации ряба правил, которые представляют собой кортежи (условие, сообщение и приоритет), где условие – функция сопоставления с эталоном. Пример представлен ниже:

```
Triple(
            { soil: Double, _: Crop -> soil in 20.0..30.0 },
            "Почва начинает пересыхать...",
            "medium"
        ).
```

После срабатывания правила инициируется создание объекта Recomendation, с целевой командой, которая выбирается посредством анализа текста рекомендации через таблицу (actionTextMap):

```
val actionTextMap = listOf(
            "пересушена" to "Полить",
            "сух" to "Полить",
            "увлажнить" to "Полить",
            "переувлажн" to "Отключить полив"…)
```

2. Метод имитационного моделирования для сценария оценки рисков

На этапе прототипирование используется метод имитационного моделирование. Главная задача которого генерация синтетических, релевантных данных, которые обеспечивают полноту контекстам для тестирования алгоритмом принятия решений.

Для генерации при временном отсутствие прямых изменения используется модуль FakeWeatherDataGenerator. Алгоритм основан на двух принципах:

– Статистическая релевантность. Значения генерируются исходя из биологических допустимых норм для культуры, что обеспечивает моделирование реалистичных измерений климатических показателей.

– Контекст данных. Генерируемые данные привязаны к конкретному объекту Planting, что обеспечивает целостность цифровой модели.

Синтетический временной ряд, сформированный модулем FakeWeatherDataGenerator, используется для визуализации данных в виде графиков в фрагменте системы – AnalysisFragment. Метод generateFakeWeatherData() предоставляет данные за 30 дней, на основе которых отрисовывается сравнительные графики, где фактические (в данном случае синтетические) показатели сопоставляются с агрономическими нормами культур, что позволяет оценить состояние системы и отследить отклонение от оптимальных условий выбранной культуры.

Таким образом метод имитационного моделирование обеспечивает генерацию данных для тестирования работоспособности системы и поддерживает процесс принятия решений, за счет структурирование данных посредством их визуализации.

4. Реализация методов в прототипе системы

Прототип системы реализовывается для системы Android с использованием языка программирования Kotlin. Архитектура системы спроектирована на основе паттерна Model-View-ViewModel (MVVM) для обеспечения модульности, тестируемости и масштабируемости, и разделена на три логических слоя:

1. Слой данных(Data Layer). Включает модули взаимодействия с внешним API, локальной базой данных SQLite и генерацией

синтетических данных. Модуль синтеза данных (FakeWeatherDataGenerator) изолирован для последующей замены.

2. Слой бизнес-логии (Domain Layer). Содержит ядро системы принятия решений, реализовано независимо от интерфейса. RecommendationGenerator – ключевой компонент, который инкапсулирует систему алгоритмом анализа отклонений от агрономических эталонов.

3. Слой представления (Presentation Layer). Отвечает за визуализацию данных через мобильный интерфейс. Его задача упростить восприятие результатов работы бизнес-логики в удобном и понятном формате.

Данная организация явного разделения реализует принцип разделения ответственности (Separation of Concerns), что закладывает основу для вариантного развития системы, гарантирующую безболезненную для работоспособности замену компонентов.

Заключение

В работе был рассмотрена архитектура и принципы автоматизированной системы поддержки принятия решений для агропромышленного комплекса. Основное внимание уделено проектированию целостной функциональной модели, включающую модули мониторинга, формирующего контекст из разнородных источников (API, синтерз и эталоны), и детерминированного анализа, реализующего генерацию целевых рекомендаций.

Функциональная модель демонстрирует полный цикл формирование рекомендаций: начинаю от сбора информации, преобразование полуструктурированных данных (JSON-ответов), семантического сопоставление, заканчивая анализом, выявлением отклонений от эталонов культуры и генерацией рекомендаций.

Трехслойная структура, обеспечивает стратегическую гибкость системы: замена модуля синтезирования данных на адаптер IoT-датчиков или масштабирование за счет обогащения ядра алгоритмами машинного обучения без критического изменения архитектуры системы.

Внедрение АСППР в аграрный сектор важный этап цифровизации, который направлен на обеспечение минимизации рисов потери урожая. Система необходима для постоянного мониторинга и своевременного реагирование на смену показателей. Она снизит ответственность с людей, за счёт систематизации данных и корректных рекомендаций, основанных на анализе.

Список литературы

1. Указ Президента Российской Федерации от 09.05.2017 N 203 "О Стратегии развития информационного общества в Российской Федерации на 2017 - 2030 годы"/ КонсультантПлюс. [Электронный ресурс] – Режим доступа: https://www.consultant.ru/document/cons_doc_LAW_216363/ (дата обращения: 17.01.2026).

2. Щербакова А.С. Цифровая трансформация сельского хозяйства: от глобального до регионального уровня / А.С. Щербакова, В.А. Иванов, И.С. Мальцева [и др.] Journal of Agriculture and Environment. — 2024. — №11 (51). [Электронный ресурс] — Режим доступа: https://jae.cifra.science/archive/11-51-2024-november/10.60797/JAE.2024.51.10 (дата обращения: 17.01.2026)

3. Добровлянин, В. Д. Цифровизация сельского хозяйства: текущий уровень цифровизации в Российской Федерации и перспективы дальнейшего развития / В. Д. Добровлянин, Е. А. Антинескул // Цифровые модели и решения. – 2022. – № 2. [Электронный ресурс] - Режим доступа: https://cyberleninka.ru/article/n/tsifrovizatsiya-selskogo-hozyaystva-tekuschiy-uroven-tsifrovizatsii-v-rossiyskoy-federatsii-i-perspektivy-dalneyshego-razvitiya (дата обращения: 17.01.2026)

4. «Умные» фермы: как искусственный интеллект меняет сельское хозяйство [Электронный ресурс] - Режим доступа: https://www.rbc.ru/technology_and_media/14/06/2023/64802aae9a7947c6121756b7 (дата обращения: 17.01.2026)

5. Пугин, А. М. Система поддержки принятия решений по управлению процессами точного земледелия с использованием цифровых технологий / А. М. Пугин, Г. А. Саитова, Э. Р. Габдуллина // Современные наукоемкие технологии. – 2022. – № 10-2. – С. 226–230 [Электронный ресурс] – Режим доступа: https://top-technologies.ru/article/view?id=39374 (дата обращения: 17.01.2026)

ИИ как инструмент автоматизации типовых чертежей в CAD-системах

Д. Ю. Кириченко, email: dimavariator2006@gmail.com
И. А. Капустин, email: ikapustin16@gmail.com
В.В. Сокольников

Воронежский государственный технический университет

Аннотация. *Рассматриваются возможности применения инструментов искусственного интеллекта в системах автоматизированного проектирования для ускорения процесса создания типовых машиностроительных чертежей при работе в CAD-системах КОМПАС-3D, AutoCAD и Siemens NX.*

Ключевые слова: *искусственный интеллект, CAD-системы, автоматизация черчения, КОМПАС-3D, болт, гайка.*

Введение

Современное машиностроительное проектирование характеризуется высокой долей повторяющихся операций. При разработке конструкторской документации инженер-конструктор значительное время затрачивает на создание типовых элементов, таких как крепёжные изделия, стандартные валы, втулки и другие детали, параметры которых строго регламентированы нормативной документацией. Несмотря на развитие CAD-систем, процесс создания таких чертежей во многих случаях выполняется вручную, что приводит к увеличению сроков проектирования и росту вероятности ошибок [3, 4, 5].

Развитие технологий искусственного интеллекта позволило по-новому взглянуть на задачу автоматизации типового черчения. В настоящее время ИИ используется для генерации параметрических моделей, анализа стандартов и автоматического формирования чертежей на основе заданных параметров. В инженерной практике ИИ выступает не как самостоятельный проектировщик, а как «умный чертежник», ускоряющий рутинные операции.

1. Применение искусственного интеллекта в CAD-системах

Использование искусственного интеллекта в CAD-системах ориентировано прежде всего на задачи, которые хорошо формализуются и подчиняются строгим правилам. Типовые детали машиностроения имеют фиксированные геометрические параметры, определяемые

© Кириченко Д. Ю., Капустин И. А., Сокольников В. В., 2026

стандартами, что делает их удобным объектом для автоматической генерации. ИИ-ассистенты работают на основе шаблонов, параметрических зависимостей и баз нормативной документации, что позволяет формировать модели и чертежи по заданным исходным данным.

Взаимодействие инженера с ИИ осуществляется через задание параметров или текстовый запрос, в котором указывается тип детали, её размер и стандарт. На основе этого запроса ИИ анализирует параметры, выбирает соответствующий нормативный документ и формирует геометрию детали. Полученный результат используется в CAD-системе как основа для дальнейшей работы и может быть откорректирован вручную [1].

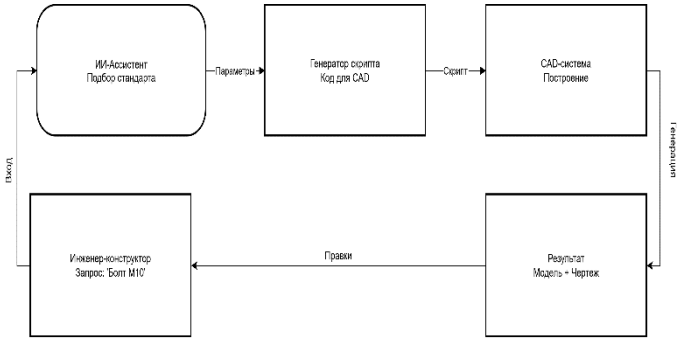


Рис. 1. Алгоритм работы ИИ-ассистента при автоматизированном создании типовой детали в CAD-системе

2. Автоматизация типовых деталей с использованием ИИ

Наиболее эффективным применение искусственного интеллекта является при работе с типовыми деталями, широко используемыми в машиностроении. К таким деталям относятся болты, гайки, шайбы, шпильки, а также простые валы и втулки. Геометрия этих изделий строго регламентирована, а различия между ними заключаются в ограниченном наборе параметров, таких как диаметр, длина и тип резьбы [3, 4].

При традиционном подходе инженер вручную создаёт модель детали, задаёт размеры и оформляет чертёж. Использование ИИ позволяет автоматизировать этот процесс. Достаточно указать обозначение детали, например болт М10 определённой длины, после чего ИИ формирует параметрическую модель, автоматически подбирая размеры в соответствии с выбранным стандартом. Далее создаётся рабочий чертёж с необходимыми видами и размерами.

Такой подход значительно сокращает время проектирования и снижает вероятность ошибок, связанных с неверным вводом размеров или несоответствием стандарту.

Таблица 1

Этапы работы ИИ-ассистента при создании типовой детали

Этап	Действие инженера	Действие ИИ	Результат
1	Задание параметров детали (тип, размер, стандарт)	Анализ запроса и определение типа изделия	Определён класс детали (болт, гайка, вал)
2	Указание стандарта (ГОСТ, ISO)	Выбор нормативной базы и типоразмера	Подбор стандартных геометрических параметров
3	Подтверждение параметров	Генерация параметрической 3D-модели	Готовая 3D-модель детали
4	Запрос на оформление чертежа	Автоматическая расстановка размеров и обозначений	Рабочий чертёж по ЕСКД
5	Проверка инженером	Корректировка по замечаниям	Готовая деталь для сборки

3. Применение ИИ в системе КОМПАС-3D

Система КОМПАС-3D широко используется в отечественной инженерной практике и содержит развитые средства параметрического моделирования и библиотеки стандартных изделий. Интеграция ИИ-подходов в рабочий процесс позволяет расширить эти возможности за счёт автоматического формирования моделей и чертежей по заданным параметрам [2].

При проектировании сборочных единиц инженер может использовать ИИ для быстрого создания крепёжных элементов или простых деталей. После задания параметров ИИ формирует модель, которая автоматически добавляется в сборку КОМПАС-3D. Аналогичным образом создаются валы и втулки, геометрия которых определяется ограниченным числом параметров. Это позволяет избежать ручного моделирования и ускорить процесс разработки конструкторской документации.

4. Преимущества и ограничения использования ИИ

Применение искусственного интеллекта в CAD-системах позволяет существенно сократить время создания типовых чертежей и снизить нагрузку на инженера-конструктора. Автоматизация рутинных операций способствует повышению унификации и стандартизации

конструкторской документации, а также снижает вероятность ошибок. Использование ИИ также является эффективным инструментом в учебном процессе, так как позволяет студентам быстрее осваивать принципы работы с CAD-системами.

В то же время искусственный интеллект имеет ограничения. Он эффективен при работе с типовыми и стандартизированными деталями, но не способен полностью заменить инженера при проектировании сложных или нестандартных изделий. В таких случаях ИИ может использоваться только как вспомогательный инструмент [2, 3].

Заключение

Искусственный интеллект в CAD-системах является перспективным направлением развития автоматизации машиностроительного проектирования. Использование ИИ позволяет ускорить процесс создания стандартных деталей, таких как болты, гайки и валы, особенно при работе в системах КОМПАС-3D, AutoCAD и Siemens NX. При этом ИИ не заменяет инженера-конструктора, а выполняет функции интеллектуального помощника, берущего на себя рутинные операции и повышающего эффективность проектирования.

Список литературы

1. Горошко, А. В. Системы автоматизированного проектирования в машиностроении : учеб. пособие / А. В. Горошко. – М. : Академия, 2012. – 336 с.

2. Анурьев, В. И. Справочник конструктора-машиностроителя. В 3 т. Т. 1 / В. И. Анурьев. – 9-е изд., перераб. и доп. – М. : Машиностроение, 2001. – 920 с.

3. Барсегян, А. А. Анализ данных и процессов : учеб. пособие / А. А. Барсегян, М. С. Куприянов, С. И. Елизаров. – 3-е изд., перераб. и доп. – СПб. : БХВ-Петербург, 2009. – 512 с.

4. ГОСТ 2.109–73. Основные требования к чертежам [Текст]. – Введ. 01.07.1974. – М. : Стандартинформ.

5. ГОСТ 2.307–2011. Нанесение размеров и предельных отклонений [Текст]. – Введ. 01.09.2012. – М. : Стандартинформ.

Архитектура подсистемы параметрического моделирования специализированной САПР на примере ступенчатого вала

В. И. Котолевский, email: kotolevsky_vi@cs.vsu.ru [1]
А. С. Троценко, email: trotsenko93@mail.ru [1]

[1] Воронежский государственный университет

Аннотация. *Работа посвящена разработке программной подсистемы специализированной системы автоматизированного проектирования, отвечающей за параметрическое формирование твердотельных моделей. Рассматривается архитектурный шаблон, основанный на разделении вычисления параметрических зависимостей и процедур геометрического построения, что позволяет упростить разработку и модификацию специализированных САПР. В качестве демонстрационного объекта используется задача построения параметрической модели ступенчатого вала. Описываются принципы формирования геометрии на основе набора управляющих параметров и пропорциональных зависимостей между ними. Рассматривается построение твердотельных моделей валов со сложными конструктивными элементами и их представление в формате обмена данными STEP*

Ключевые слова: *специализированные САПР, параметрическое моделирование, архитектура программной системы, твердотельное моделирование, CAD, STEP.*

Введение

В последние годы в инженерной практике всё чаще возникает потребность в разработке специализированных систем автоматизированного проектирования, ориентированных на решение узкого класса задач. В отличие от универсальных CAD-систем, такие решения нацелены на автоматизацию конкретных операций проектирования и позволяют сократить трудозатраты за счёт жёсткой привязки к предметной области. Однако разработка специализированных САПР, даже для относительно простых задач, требует значительных усилий, связанных с проектированием архитектуры программной системы и интеграцией геометрического ядра.

Одной из ключевых проблем при создании специализированных САПР является отсутствие универсальных архитектурных шаблонов, позволяющих быстро разрабатывать и модифицировать программные

© Котолевский В. И., Троценко А. С., 2026

подсистемы, отвечающие за параметрическое формирование геометрических моделей. Вопросы построения высокоуровневой структуры таких систем, обеспечивающей разделение ответственности между модулями и снижение связанности компонентов, активно рассматриваются в современных работах, посвящённых архитектуре специализированных САПР [1]. На практике же разработчики часто вынуждены каждый раз заново решать задачи организации вычисления параметрических зависимостей, управления геометрическим построением и взаимодействия между компонентами системы, что приводит к дублированию решений и усложняет сопровождение программных продуктов.

В рамках данной работы рассматривается подход к построению программной подсистемы специализированной САПР, ориентированной на параметрическое формирование твердотельных моделей. Основное внимание уделяется архитектурным аспектам системы, в частности разделению вычисления параметрических зависимостей и процедур геометрического построения, а также организации взаимодействия основных компонентов. Такой подход позволяет рассматривать подсистему как обобщённый шаблон, применимый при разработке специализированных САПР для различных инженерных задач.

В качестве демонстрационного примера используется задача параметрического построения ступенчатого вала. Выбор данного объекта обусловлен тем, что он является хорошо изученным и формализованным элементом машиностроения, для которого не требуется разработка новых методов моделирования. Это позволяет сосредоточиться не на особенностях геометрии конкретного изделия, а на анализе архитектурных решений и принципов построения программной системы, лежащих в основе разработки специализированных САПР.

1. Подход к параметрическому формированию геометрии (на примере ступенчатого вала)

В рамках данной работы параметрическое моделирование рассматривается как инструмент построения специализированных программных подсистем САПР. Используемая далее задача параметрического моделирования ступенчатого вала служит формализованным примером, позволяющим проиллюстрировать архитектурные решения и принципы организации программной системы [1]. В случае ступенчатых валов к таким параметрам относятся общая длина вала, диаметры отдельных цилиндрических участков, а также размеры и положение конструктивных элементов. Геометрия детали формируется автоматически на основе этих параметров без необходимости ручного редактирования отдельных элементов модели.

В рамках рассматриваемого подхода ступенчатый вал представляется в виде последовательности соосных сегментов, каждый из которых имеет заданную форму и размеры. Основными типами сегментов являются цилиндрические и конические участки, обеспечивающие формирование ступеней и плавных переходов между ними. Размеры сегментов определяются на основе управляющих параметров и пропорциональных зависимостей, что позволяет сохранять целостность геометрии при изменении исходных данных.

Дополнительные конструктивные элементы, такие как фаски и шпоночные пазы, формируются автоматически после построения основной геометрии вала. Параметры этих элементов задаются в привязке к соответствующим сегментам и масштабируются при изменении размеров вала. Такой подход позволяет получать геометрически согласованные модели, соответствующие требованиям машиностроительного проектирования.

Использование параметрического описания ступенчатого вала обеспечивает возможность быстрого получения различных вариантов конструкции на основе единой модели [2, 3]. Это особенно важно при проектировании типовых деталей, а также при необходимости адаптации конструкции под различные условия эксплуатации.

2. Архитектура и принцип построения трёхмерной модели

Формирование трёхмерной модели ступенчатого вала осуществляется на основе модульной архитектуры, разделяющей процессы параметризации и геометрического построения. Такой подход позволяет обеспечить независимость логики задания параметров от конкретных средств визуализации и пользовательского интерфейса. Общая структура взаимодействия основных компонентов системы представлена на рис. 1.

На первом этапе выполняется формирование набора геометрических параметров, определяющих размеры и пропорции вала. В качестве управляющих параметров используются общая длина вала и диаметры отдельных опорных цилиндрических участков. На их основе автоматически вычисляются размеры остальных сегментов и параметры конструктивных элементов.

На втором этапе осуществляется построение трёхмерной геометрии вала. Основная форма детали формируется путём последовательного формирования осесимметричных геометрических участков с последующим объединением их в единое твердотельное представление. После формирования базовой геометрии в модель добавляются конструктивные элементы, такие как фаски на торцевых поверхностях и шпоночные пазы на заданных участках.

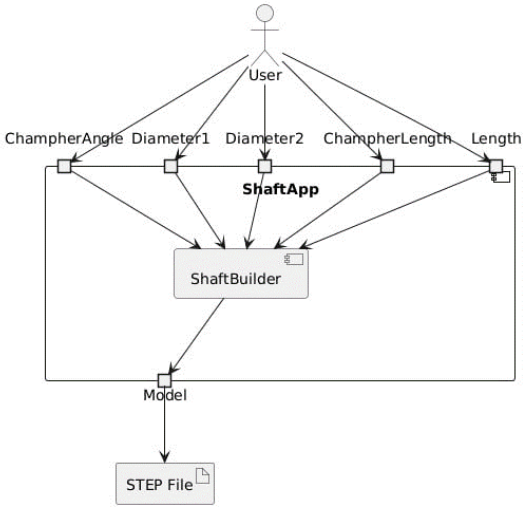


Рис. 1. Диаграмма компонентов системы параметрического построения ступенчатого вала

Построение геометрии выполняется с использованием геометрического ядра OpenCASCADE Technology, обеспечивающего операции создания примитивов, булевы операции и формирование твердотельных моделей в виде граничного представления [4]. Итоговая модель представляется в формате BRep, что обеспечивает корректное описание формы и топологии изделия.

Завершающим этапом является экспорт полученной трёхмерной модели в формат обмена данными STEP, предназначенный для передачи геометрической информации между различными CAD-системами и дальнейшего использования в процессах проектирования и производства.

3. Реализация параметрического построения ступенчатого вала

Реализация параметрического построения ступенчатого вала основана на использовании геометрического ядра OpenCASCADE Technology и включает этап вычисления геометрических параметров и этап формирования твердотельной модели. Логика параметризации и геометрического построения реализована в виде единого вычислительного конвейера, не зависящего от конкретного пользовательского интерфейса. Взаимосвязь основных классов, участвующих в процессе построения модели, показана на рис. 2.

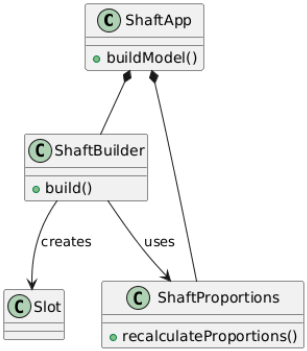


Рис. 2. Диаграмма классов основных компонентов системы

На первом этапе выполняется вычисление абсолютных размеров всех элементов вала. В качестве исходных данных используются общая длина вала и диаметры выбранных опорных цилиндрических участков. На их основе производится динамический пересчёт базового диаметра и связанных с ним геометрических параметров остальных сегментов. Такой подход позволяет автоматически изменять размеры всей конструкции при корректировке ключевых параметров без нарушения геометрической согласованности модели.

На втором этапе осуществляется построение твердотельной геометрии вала. Геометрические участки формируются в виде осесимметричных тел в абсолютных размерах, рассчитанных на предыдущем этапе, после чего последовательно объединяются в единое твердое тело с использованием булевых операций объединения [4]. Это обеспечивает корректное сопряжение сегментов и формирование целостной геометрии детали.

После построения основной формы вала выполняется добавление конструктивных элементов. Фаски формируются на торцевых поверхностях с использованием специализированных операций геометрического ядра, а шпоночные пазы создаются путём вычитания вспомогательных тел из основной геометрии вала. Размеры и пространственное положение пазов определяются параметрами соответствующих цилиндрических участков и автоматически пересчитываются при изменении геометрии вала. Общая последовательность этапов параметрического построения модели представлена на рис. 3.



Рис. 3. Диаграмма активностей процесса параметрического построения модели

Итоговая модель формируется в виде твердотельного объекта граничного представления и экспортируется в формат STEP [3. 5]. Это обеспечивает сохранение геометрической и топологической информации и возможность дальнейшего использования модели в различных CAD-системах для проектирования, анализа и подготовки производства. Пример полученной трёхмерной модели ступенчатого вала приведён на рис. 4.

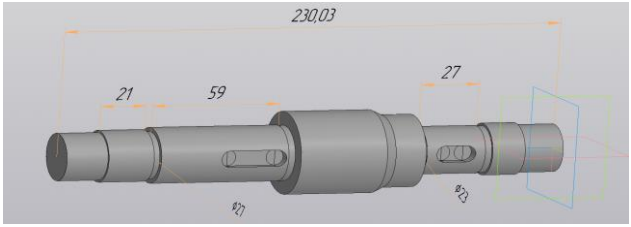


Рис. 4. Трёхмерная модель ступенчатого вала, сформированная параметрическим методом

Заключение

В работе рассмотрена архитектура программной подсистемы специализированной системы автоматизированного проектирования, предназначенной для параметрического формирования твердотельных моделей. Предложенный подход основан на разделении вычисления параметрических зависимостей и процедур геометрического построения, что позволяет упростить разработку и модификацию

специализированных САПР. Реализация подсистемы с использованием геометрического ядра OpenCASCADE Technology демонстрирует возможность изоляции прикладной логики от средств низкоуровневого моделирования и формирования корректных твердотельных моделей с последующим экспортом в формат STEP. Использование задачи построения ступенчатого вала в качестве примера позволяет сосредоточиться на архитектурных и методических аспектах разработки программной системы. Рассматриваемая подсистема может быть использована как задел для создания повторно применяемых архитектурных решений и фреймворков при разработке специализированных САПР для узкого класса инженерных задач.

Список литературы

1. Троценко А. С., Успехов А. А., Чижов М. И. Высокоуровневая структура модулей для построения специальных систем автоматизированного проектирования [Электронный ресурс]: науч. ст. // Моделирование, оптимизация и информационные технологии. – Электрон. журн. – Режим доступа: https://www.elibrary.ru/item.asp?id=54676401
2. Герасимов И. М., Кощеев В. Д. Основы автоматизированного проектирования [Электронный ресурс]: учеб. пособие. – СПб.: СПбГПУ, 2004. – Режим доступа: https://elib.spbstu.ru/dl/1158.pdf/download/1158.pdf
3. Солодилова Н. А., Маркова Т. В., Скрябнев А. Ю., Блинковский А. М. Параметрическое моделирование в САПР как ключевая компетенция современного инженера-машиностроителя [Электронный ресурс]. – СПб.: Санкт-Петербургский политехнический университет Петра Великого. – Режим доступа: https://elib.spbstu.ru/dl/2/id25-97.pdf/view
4. OpenCASCADE Technology. [Электронный ресурс]: Официальная документация. – Режим доступа: https://dev.opencascade.org/doc/overview/html/index.html
5. ISO 10303. Standard for the Exchange of Product Model Data (STEP) [Электронный ресурс]. – Режим доступа: https://www.iso.org/standard/63141.html

Кастомизация программы трехмерного моделирования NX

А. Е. Лесных, email: Alexey.Lesnyh@yandex.ru
И. В. Петренко, email: Lady00788@mail.ru
М. И. Чижов, email: mihailC@list.ru

Воронежский государственный университет

Аннотация. *В данной работе рассматривается кастомизация программы трехмерного моделирования NX с целью автоматизации создания трехмерных объектов, имеющих заданные геометрические характеристики.*

Ключевые слова: *трехмерное моделирование, кастомизация, программирование, CAD, NX Open API.*

Введение

Для тонкой настройки графического интерфейса программы NX и разработки собственных диалоговых окон, которые работают внутри основного окна NX, можно использовать инструмент «Разработчик пользовательского интерфейса», который появился начиная с NX версии 6.0. Собственные диалоговые окна позволяют значительно расширить основной функционал программы NX и автоматизировать выполнение повторяющихся операций. Кроме этого, можно запрограммировать создание сложных трехмерных моделей, которые обладают заданными геометрическими характеристиками. Таким образом, можно реализовать трехмерное параметрическое моделирование: пользователь задаёт некоторые общие геометрические характеристики детали, например, длина, ширина и т.д., а программа рассчитывает конфигурацию детали, которая будет удовлетворять заданным геометрическим характеристикам.

Разработка пользовательского интерфейса диалогового окна

Создание настраиваемого диалогового окна внутри основного окна программы NX осуществляется с помощью инструмента «Разработчик пользовательского интерфейса» [1]. Данный инструмент открывается при выборе команды «Разработчик пользовательского инструмента» в меню «Дополнительно» на закладке «Приложения» в ленте основного окна программы NX. После запуска данного инструмента в основном окне NX появляются графические элементы, позволяющие сформировать новое диалоговое окно, которое в будущем будет открываться внутри основного окна NX, и будет отвечать конкретным требованиям пользователя, т.е.

© Лесных А.Е., Петренко И.В., Чижов М.И., 2026

решать специализированную задачу, автоматизируя создание трехмерных деталей и сборок в NX.

Графический интерфейс инструмента «Разработчик пользовательского интерфейса» можно разделить на 3 части (см. рис. 1).

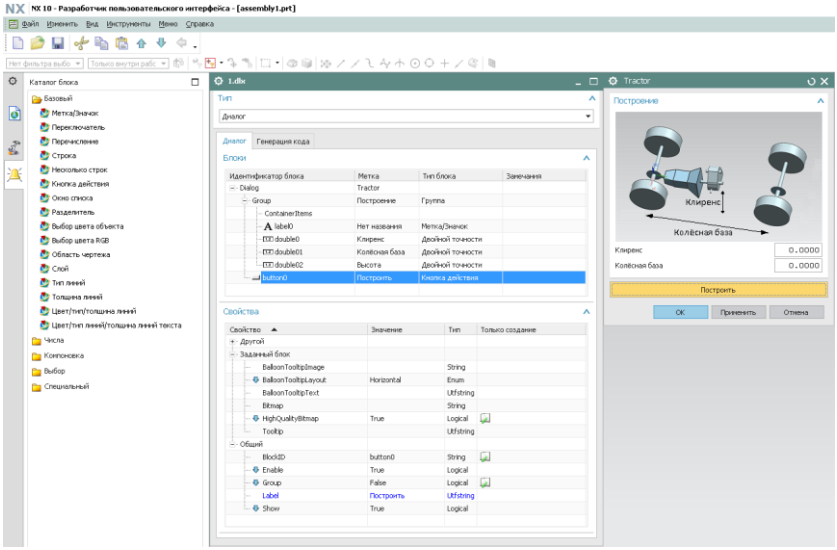


Рис. 1. Разработчик пользовательского интерфейса

В левой части экрана находится окно, которое называется «Каталог блока». Это окно содержит основные строительные элементы, которые используются для построения диалогового окна, такие как: «Метка/значок», «Переключатель», «Перечисление», «Строка» и т.д.

Второе основное окно инструмента «Разработчик пользовательского инструмента» расположено посередине экрана и называется: «Конструктор диалога». Это окно состоит из двух частей. В верхней части находится древовидная структура, которая называется: «Блоки» и представляет логическую структуру диалогового окна. Она отражает порядок элементов диалогового окна и позволяет сгруппировать несколько элементов. Для добавления элементов в эту структуру надо перетащить нужные элементы из левой части окна, из «Каталога блоков». Ниже структуры «Блоки» находится окно «Свойства», предназначенное для тонкой настройки каждого блока. Если выбрать нужный блок в дереве «Блоки», то в свойствах ниже можно настроить основные свойства блока, такие как: идентификатор блока, надпись и т.д.

Третье окно инструмента «Разработчик пользовательского инструмента» представляет окончательный вид диалогового окна, как оно будет выглядеть при запуске программы в окне NX.

Инструмент «Разработчик пользовательского интерфейса» генерирует файлы с расширениями *.cpp, *.hpp и *.dlx. В файле *.hpp приведено описание имен всех обработчиков событий на диалоге, в файле *.cpp – программный код для работы с диалогом, выполненным в стиле объектно-ориентированного программирования, а в файле *.dlx содержится описание, собственно диалога.

Для построения твердого тела из блоков (параллелепипедов) используется функция UF_MODL_create_block1. Она принимает следующие параметры:

UF_FEATURE_SIGN «sign»: определяет операцию булевого моделирования, выполняемую над блоком и целевым объектом;

double «corner_pt»: Трехмерные координаты точки, задающей начало (угловую точку) блока. Эта точка задается в абсолютной системе координат CSYS;

char* «edge_len»: Три значения длины ребер блока вдоль осей X, Y и Z соответственно;

tag_t «blk_obj_id»: Идентификатор объекта блока.

В следующем листинге 1 приводится фрагмент исходного текста функции, которая строит модель твердого тела (параллелепипеда).

Листинг 2

Создание параллелепипеда

```
inline void drawRazdatka(double klirens, double kolesnaBaza) {
    // Определение нейтральной булевой операции
    UF_FEATURE_SIGN sign = UF_NULLSIGN;
    // Задание размеров параллелепипеда
    double box_orig[3] = { -243,885,-570/2+klirens };
    char LenghtB[100] = "345";
    char WidthB[100] = "130";
    char HeightB[100] = "270";
    char* size[3] = { LenghtB, WidthB, HeightB };
    // Задание идентификатора
    tag_t box_obj;
    // Вызыв функции создания
    UF_MODL_create_block1(sign, box_orig, size, &box_obj);

}
```

Переменна типа UF_FEATURE_SIGN используется для обозначения различных типов операций над телами. Значение UF_NEGATIVE обозначает операцию вычитания, которая используется для удаления или

вырезания части тела из другого тела. Например, в листинге 2 приводится исходный текст функции, которая создает отверстие.

Листинг 3

Создание отверстия с помощью операции вычитания цилиндра

```
inline void drawStoikaHoleD16_5(double length, double width) {
    // Определение булевой операции как вычитание
    UF_FEATURE_SIGN sign = UF_NEGATIVE;
    // Задание размеров отверстия
    double hole_orig[3] = { 0, length-50+25, 22 };
    char height[125] = "25"; // Высота отверстия 25
    char diam[125];
    snprintf(diam, sizeof(diam), "%.2f", 16.5); // Диаметр
    double direction[3] = { 0, -1.0, 0 }; // Направление
против оси Y
    // Задание идентификатора
    tag_t hole_obj;
    // Создание отверстия
    UF_MODL_create_cyl1(sign, hole_orig, height, diam,
direction, &hole_obj);
}
```

В процессе работы бы создан программный продукт, реализующий создание компоновочной геометрии трансмиссии и тележки транспортного средства. В результате выполнения программы создается трехмерная компоновочная геометрия раздаточной коробки, которая входит в компоновочную схему трансмиссии, показанную на рис. 2.

Заключение

Создание компоновочной геометрии транспортной тележки реализовано в виде динамической библиотеки, создающей параметрические объемы основных агрегатов, при этом основные размеры задаются в поля диалогового окна. Это позволяет автоматизировать процесс проектирования при решении конструкторских задач с учетом вариантного моделирования.

Список литературы

1. Разработка прикладных модулей для системы автоматизированного проектирования «Siemens NX» : учебное пособие / А.И. Сергеев ; Оренбургский гос. ун-т. – Оренбург : ОГУ. – 2021. – 137 с.

2. Тихомиров, В.А. Разработка приложений для NX на языке C / В.А. Тихомиров. – Комсомольск-на-Амуре : Изд-во ФГБОУ ВПО «КнАГТУ», 2011. – 469 с.

3. Getting Started with NX Open. – Siemens Product Lifecycle Management Software, 2016. – 137 p.

4. Краснов, М. Unigraphics для профессионалов / М. Краснов, Ю. Чигишев. – М.: Изд-во «ЛОРИ», 2004. – 320 с.

5. Чижов М.И., Бредихин А.В., Ветохин В.В. : Подготовка производства: автоматизированное проектирования и средства поддержки жизненного цикла изделия. / - М.И. Чижов Воронеж: Диамат, 2021. – А.В. Бредихин, В.В. Ветохин, 167 с.

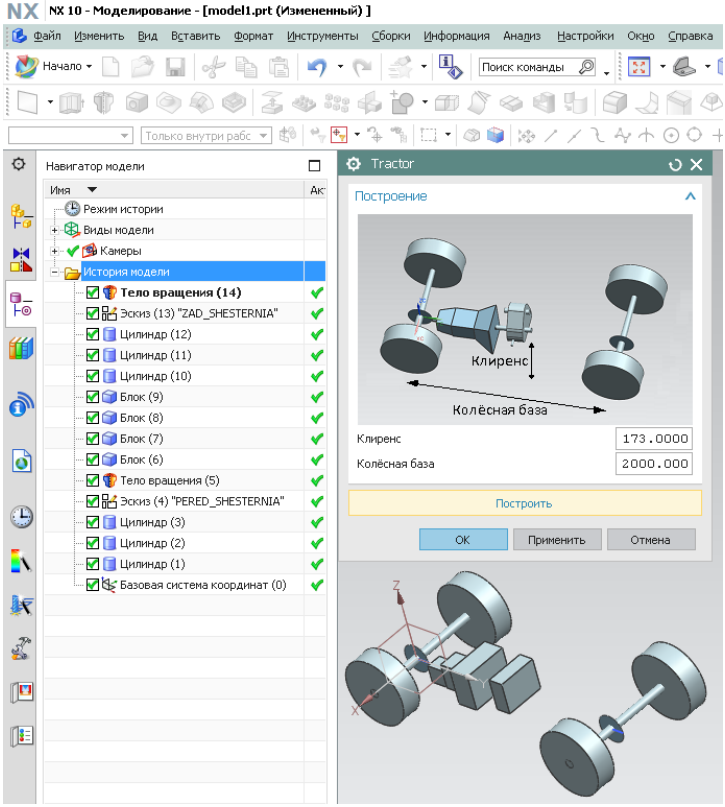


Рис. 2. Результат построение компоновочной схемы трансмиссии

Параллельное проектирование

К. Д. Поднебесная, e-mail: purplewomanx@gmail.com
В.В. Сокольников, e-mail: svp_kitp@mail.ru

Воронежский государственный технический университет

Аннотация. *В статье рассматриваются современные технологии параллельного проектирования как ключевого элемента цифровой трансформации машиностроительных отраслей. Анализируются методологические основы, архитектурные решения и программные платформы, обеспечивающие одновременную совместную работу распределённых команд над сложными инженерными объектами. Особое внимание уделяется интеграции параллельных процессов в системы САПР и механизмам управления конфигурацией. Приводятся примеры практической реализации в российской и мировой практике, а также оценка экономического эффекта от внедрения.*

Ключевые слова: *параллельное проектирование, САПР, распределённая разработка, управление конфигурацией, цифровая нить, PLM-системы.*

Введение

Современное машиностроение, особенно в таких высокотехнологичных отраслях как авиастроение, автомобилестроение и энергетическое машиностроение, сталкивается с беспрецедентными вызовами по сокращению сроков вывода продукции на рынок при одновременном повышении её сложности, качества и кастомизации. Традиционный последовательный (каскадный) подход к проектированию, когда работы над различными узлами и системами изделия ведутся поэтапно, становится критическим фактором, ограничивающим скорость и гибкость разработки [1].

Каждое выявленное на поздних стадиях несоответствие или требование изменения влечёт за собой дорогостоящие итерации по уже завершённым этапам. Технологии параллельного (совместного) проектирования предлагают принципиально иную парадигму, основанную на одновременной и скоординированной работе всех участников жизненного цикла изделия — от конструкторов и технологов до специалистов по производству и сервису на самых ранних этапах проекта. Эта методология, реализуемая на базе передовых САПР и PLM-систем, позволяет не только сократить общее время разработки на 30-

© Поднебесная К.Д., Сокольников В.В., 2026

50%, но и существенно повысить качество принимаемых решений за счёт комплексного учёта всех требований и ограничений [2].

1. Методологические основы параллельного проектирования

Сущность параллельного проектирования заключается не в простом распараллеливании задач между разными исполнителями, а в создании интегрированной среды, где все участники работают с актуальными, согласованными данными в реальном времени. Фундаментальным отличием от последовательного подхода является переход от модели «разработка-анализ-переработка» к модели непрерывной интеграции и валидации. Технологической основой для этого служит концепция «единого источника истины» — централизованного репозитория данных, где хранятся и управляются все цифровые модели, чертежи, спецификации и связанная с ними информация. Ключевыми методологическими принципами являются раннее вовлечение всех заинтересованных сторон (конструкторы, технологи, производственники, экономисты, специалисты по надёжности) в процесс проектирования, непрерывный обмен знаниями и ограничениями между смежными командами, использование цифровых прототипов и симуляций для верификации решений до этапа физического изготовления, а также итеративная разработка с постепенным уточнением деталей модели на основе поступающих данных и результатов анализа. Такой подход требует внедрения строгих процедур управления конфигурацией, которые обеспечивают контроль версий, отслеживание изменений и управление вариантами изделия, предотвращая конфликты и потерю данных при одновременной работе множества пользователей [2].

2. Архитектура и программные платформы для реализации

Техническая реализация технологий параллельного проектирования базируется на комплексной IT-инфраструктуре, ядром которой являются PLM-системы (Product Lifecycle Management). Такие платформы, как Siemens Teamcenter, Dassault Systèmes 3DEXPERIENCE, PTC Windchill или российский «Лоцман:PLM», выполняют роль интеграционного каркаса, связывающего разнородные инструменты в единый информационный контур. Типичная архитектура включает серверную часть (PLM-ядро) — централизованную базу данных для хранения всех объектов и их взаимосвязей, сервисы управления бизнес-процессами, правами доступа и жизненным циклом данных; интеграционные адаптеры для САПР, обеспечивающие прямую работу с CAD-моделями из среды PLM без необходимости ручного импорта/экспорта файлов; веб-клиенты и инструменты для совместной работы, доступные специалистам без установки тяжеловесных САПР; а также средства

моделирования и анализа (CAE), позволяющие запускать расчёты напрямую на актуальных данных из PLM. Особую роль играют облачные технологии и распределённые репозитории, которые позволяют организовать работу географически распределённых команд, в том числе в рамках глобальных цепочек поставок. Современные системы поддерживают гибридные модели развёртывания, обеспечивая необходимый баланс между производительностью, безопасностью данных и гибкостью доступа.

3. Преимущества и экономический эффект внедрения

Внедрение технологий параллельного проектирования приносит комплексный измеримый эффект на всех уровнях предприятия. Прежде всего, это сокращение сроков выхода на рынок за счёт ликвидации «времени ожидания» между последовательными этапами и возможности одновременной работы над разными аспектами изделия, что приводит к радикальному сокращению цикла «от концепции до продукта». Согласно отраслевым исследованиям, сокращение времени разработки может достигать 40-60% для сложных изделий. Во-вторых, достигается значительное снижение затрат на проектирование и переделки, поскольку раннее выявление коллизий в цифровой среде позволяет устранять их на этапе, когда стоимость изменений минимальна. Это уменьшает количество дорогостоящих физических макетов и испытаний, а также исключает затраты на переделку оснастки и инструмента на производстве. В-третьих, повышается качество и инновационность продукции за счёт совместной работы экспертов из разных областей на ранних стадиях, что позволяет находить более оптимальные, комплексные решения. Технологии вроде генеративного дизайна, работающие в среде параллельного проектирования, позволяют автоматически находить нетривиальные, оптимальные по массе и прочности формы конструкций. Наконец, улучшается управление знаниями и преемственность, так как все знания о проекте, принятые решения и их обоснования фиксируются в PLM-системе, становясь корпоративным активом, что упрощает адаптацию новых сотрудников, модификацию изделий и поддержку продукции.

4. Практические примеры и кейсы внедрения

Опыт ведущих промышленных компаний демонстрирует успешность применения технологий параллельного проектирования. В авиастроении консорциум Airbus для создания семейства самолётов A320neo использовал глобальную PLM-платформу, позволившую координировать работу тысяч инженеров и сотен поставщиков по всему миру. Параллельная разработка планера, силовой установки (двигатели

Pratt & Whitney PW1000G) и систем бортового оборудования позволила сократить сроки сертификации и вывода самолёта на рынок в условиях жёсткой конкуренции [2].

В российской практике значимым примером является Объединённая двигателестроительная корпорация (ОДК). При создании перспективного авиадвигателя ПД-8 была реализована модель параллельной работы конструкторских бюро (АО «ОДК-Авиадвигатель», АО «ОДК-Сатурн»), технологических институтов и серийных заводов на единой цифровой платформе. Это позволило синхронизировать разработку газогенератора, компрессора, турбины и систем управления, минимизировать количество натурных доводок и ускорить процесс выхода на лётные испытания.

Ещё одним показательным кейсом является автомобильная промышленность. Компания Tesla при проектировании новых моделей активно использует принципы параллельного проектирования, объединяя в единый процесс работу над шасси, силовой электроникой, аккумуляторными модулями и программным обеспечением. Это является одним из факторов, позволяющих компании осуществлять частые итеративные обновления модельного ряда.

5. Вызовы и перспективы развития

Несмотря на очевидные преимущества, внедрение технологий параллельного проектирования сопряжено с рядом вызовов. Основными из них являются высокие первоначальные инвестиции в приобретение лицензий на PLM-системы, мощное серверное оборудование, а также затраты на интеграцию и обучение персонала [2]. Не менее сложной является организационная трансформация, поскольку переход к параллельным процессам требует изменения сложившихся процедур, должностных инструкций и корпоративной культуры, что часто встречает сопротивление сотрудников. Кроме того, существуют проблемы с качеством исходных данных и стандартизацией, так как эффективность системы напрямую зависит от консистентности и структурированности загружаемых в неё данных, что требует внедрения строгих правил и стандартов предприятия.

Перспективы развития технологий связаны с несколькими ключевыми трендами. Во-первых, это конвергенция с цифровыми двойниками, где параллельное проектирование будет всё больше опираться на создание и использование комплексных цифровых двойников изделия, позволяющих не только проектировать, но и симулировать эксплуатацию, обслуживание и утилизацию [3]. Во-вторых, внедрение искусственного интеллекта и машинного обучения позволит алгоритмам автоматически выявлять потенциальные коллизии, предлагать альтернативные варианты решений, оптимизировать

маршруты согласований и предсказывать риски на основе анализа больших данных от предыдущих проектов. В-третьих, развитие облачных SaaS-решений и переход к облачным PLM-платформам по модели «программное обеспечение как услуга» снизит порог входа для средних и малых предприятий, обеспечив им доступ к передовым технологиям без крупных капитальных затрат. Наконец, будет углубляться интеграция в цепочки поставок, а технологии будут развиваться в сторону создания открытых экосистем, позволяющих безопасно и эффективно вовлекать в процесс параллельного проектирования внешних партнёров, поставщиков и субподрядчиков.

Заключение

Технологии параллельного проектирования перестали быть экзотической инновацией и становятся обязательным стандартом для промышленных компаний, стремящихся сохранить конкурентоспособность в условиях цифровой экономики. Их реализация на базе современных САПР и PLM-систем позволяет преодолеть ограничения традиционного последовательного подхода, обеспечивая беспрецедентное ускорение процессов разработки, снижение затрат и повышение качества сложной инженерной продукции. Успешное внедрение требует комплексного подхода, включающего не только выбор и настройку программного обеспечения, но и глубокую организационную трансформацию, инвестиции в обучение персонала и выработку новых стандартов работы. Для российской промышленности освоение этих технологий является стратегической задачей, от решения которой зависит технологический суверенитет и позиции на глобальных рынках высокотехнологичной продукции.

Литература

1. Спицнадель, В. Н. Основы системного анализа : учеб. пособие / В. Н. Спицнадель. — СПб. : Изд. дом «Бизнес-пресса», 2000. — 326 с.

2. Шляхтин, Г. С. Современные PLM-системы: управление жизненным циклом продукции / Г. С. Шляхтин, А. В. Костин. — М. : ДМК Пресс, 2018. — 288 с.

3. Семёнычев, В. К. Цифровые двойники и параллельное проектирование в машиностроении / В. К. Семёнычев // CAD/CAM/CAE Observer. — 2022. — № 3 (89). — С. 22–29.

Цифровая трансформация аналитики в управлении строительными проектами: современные тенденции и перспективы

Д. Д. Приладышева, email: darya.priladysheva@mail.ru.

ФГБОУ ВО «Санкт-Петербургский государственный экономический университет»

Аннотация. *В статье рассматриваются современные направления цифровой трансформации аналитики в управлении строительными проектами. Проанализированы технологические тенденции, определяющие развитие отрасли: внедрение интеллектуальных аналитических платформ, переход к событийно-ориентированным системам мониторинга, формирование адаптивных моделей анализа данных, а также интеграция архитектур межсистемного обмена. Представлены примеры актуальных подходов к применению цифровых технологий в российских строительных организациях и региональных программах цифровизации. Уделено внимание влиянию цифровых инструментов на повышение точности планирования, снижение рисков, оптимизацию ресурсного обеспечения и повышение прозрачности процессов. На основе анализа российских и международных практик выявлены перспективные направления дальнейшего развития цифровой аналитики, включая комплексное применение моделей прогнозирования, развитие цифровых двойников строительных объектов, формирование отраслевых платформ данных и стандартизацию механизмов информационного взаимодействия. Полученные результаты позволяют определить цифровую аналитику как ключевой фактор повышения эффективности строительных проектов и устойчивого развития строительной отрасли.*

Ключевые слова: *цифровая трансформация; строительные проекты; цифровая аналитика; информационное моделирование; цифровые платформы; прогнозные модели; строительная отрасль; управление проектами; цифровой двойник; межсистемная интеграция; предиктивная аналитика.*

Введение

Строительная отрасль России продолжает расти и усложняться: по данным Росстата, объём работ в строительстве в 2024 г. вырос на 2,1 % и достиг 16,8 трлн рублей [16]. Это сопровождается необходимостьюэффективного управления всё более масштабными и распределёнными проектами.

Современный строительный проект характеризуется высокой динамичностью параметров: изменчивостью стоимости строительных материалов, значительной вариативностью графиков поставок, прогнозными колебаниями спроса на рынке недвижимости, а также необходимостью постоянного контроля за качеством работ. Это формирует потребность в инструментальной среде, способной не только фиксировать текущие показатели, но и анализировать скрытые зависимости между ними. Интеллектуальная аналитика становится одним из механизмов, позволяющих обеспечить такой уровень управляемости. В российских организациях активное распространение получают системы, основанные на автоматизированной обработке данных с использованием алгоритмов кластеризации, регрессионных моделей, деревьев решений и нейронных сетей. Эти методы применяются для прогнозирования сроков выполнения работ, анализа влияния различных факторов на общий бюджет проекта, оптимизации загрузки ресурсов, моделирования движения персонала и техники на строительной площадке.

В условиях нарастающей сложности и масштабности строительных проектов цифровая трансформация становится условием повышения эффективности отрасли. Как отмечают эксперты, в цифровую эпоху выигрывают те компании, которые активно осваивают передовые технологии и умеют извлекать из них максимальную выгоду. Министерство строительства РФ подчеркивает, что в строительстве аккумулировано огромное количество данных из множества документов и ведомств, и управлять ими без цифровых инструментов практически невозможно. По словам заместителя министра К. Михайлика, цифровая трансформация открывает принципиально новый подход к инфраструктурным проектам: использование информационного моделирования ускоряет проектирование и экспертизу, повышает прозрачность процессов строительства и позволяет оптимизировать бюджетирование и снабжение. Это свидетельствует о высокой актуальности темы: перевод документации и бизнес-процессов в цифровой формат рассматривается в России как залог модернизации отрасли и улучшения качества строительства.

Анализ тенденций и перспектив цифровой трансформации аналитики в управлении строительными проектами

В аналитике управления строительными проектами используются разнообразные цифровые технологии. Цифровые платформы обеспечивают единое информационное пространство, в котором проектировщики, подрядчики, поставщики, управляющие компании и заказчики взаимодействуют в режиме реального времени. За счет них

удается минимизировать искажения данных, устранить повторяющееся ручное дублирование информации и сократить количество ошибок, возникающих при передаче документации между стадиями проекта. Основными функциональными элементами таких платформ становятся: централизованные базы данных, облачные хранилища, средства автоматизированного контроля версий документации и инструменты совместного моделирования.

На первом месте среди таких технологий – системы BI (Business Intelligence) и аналитического учёта, позволяющие интегрировать данные из ERP/CRM, систем управления проектом, финансов и других источников для формирования сводной управленческой отчётности. По данным Минстроя, российские платформы, например, Pragmacore, ЦУС, MStroy, EXON, «Адепт», уже включают модули построения цифровых моделей объектов, оценки рисков и прогнозирования динамики проектов. Активно внедряются на всех уровнях технологии Big Data и ИИ. Применение машинного обучения и предиктивной аналитики позволяет выявлять скрытые закономерности и заранее прогнозировать задержки или перерасход ресурсов. Так, современные ИИ-инструменты «анализируют большие объёмы данных и прогнозируют риски», что важно для планово-финансового контроля. Развиваются цифровые двойники строительных площадок – виртуальные модели реальных объектов, объединяющие данные с датчиков IoT, беспилотников и BIM, что даёт возможность в реальном времени отслеживать ход строительства, производительность и качество работ. Практика внедрения ИИ уже показала эффект в ряде направлений: например, корпорации создают внутрикорпоративных ИИ-агентов на основе LLM-моделей, которые отвечают на вопросы сотрудников по проектным регламентам и внутренним базам знаний. Также применяются технологии компьютерного зрения (CV) и оптического распознавания (OCR) для анализа проектной документации и чертежей – ИИ автоматически извлекает информацию из планов и спецификаций. Реже, но уже встречаются генеративные модели для более сложных инженерных задач, однако их широкое использование пока ограничено. [8]

Информационное моделирование (BIM/ТИМ) – отдельный класс технологий цифровой стройки. Согласно изданию РБК Trends, «технология информационного моделирования — это не просто упрощение проектирования, а построение четкой технологической цепочки: концепция и проектная документация – проект – строительство – эксплуатация». BIM-модель обеспечивает единую связанную 3D-модель объекта со встроенной информацией, которую могут

использовать все участники проекта. Например, один из экспертов отмечал, что BIM позволяет увидеть и скорректировать конструктивные конфликты (например, пересечения инженерных сетей), что неизбежно снижает количество ошибок и экономит средства. В России правительство обычно говорит не о BIM, а о широкой концепции «ТИМ» (технологий информационного моделирования) – комплексных подходах к цифровому ведению проектов. Заместитель министра К. Михайлик подчеркнул, что ТИМ-системы «сводят все элементы и процедуры строительства в единую цифровую вертикаль», повышая взаимосвязь этапов проекта. [10]

Концепция цифрового двойника расширяет BIM за счет реального времени: цифровой двойник – это «живая» виртуальная копия уже построенного объекта с постоянным потоком данных от сенсоров и систем здания, отражающая его актуальное состояние. Как пишет эксперт, создание цифрового двойника начинается с BIM-модели, которая на стройке постоянно актуализируется, например, с помощью лазерного сканирования. После ввода объекта в эксплуатацию такая «живущая» модель при подключении к датчикам становится точной цифровой копией здания, включающей данные от состава бетона до показателей энергопотребления. Практика показывает, что уже во время строительства актуализированная BIM-модель позволяет отслеживать ход работ и выявлять расхождения между проектом и реальностью, благодаря чему можно ускорить реализацию проекта примерно на четверть. Известные примеры в России включают «Лахта Центр» – здесь цифровой двойник в реальном времени контролировал деформации конструкции и обеспечивал безопасность строительства 462-метрового небоскрёба, а создаваемое в Москве электродепо «Бирюлёво-Западное» проектируется с цифровым двойником на отечественном программном обеспечении с учётом около 12 тысяч компонентов: от систем ОВиК до сетей канализации.

Одним из важнейших направлений цифровой трансформации становится переход от статичной и периодической аналитики к постоянному мониторингу проектных процессов. В отличие от практики, ориентированной на итоговые отчёты, современные системы стремятся к непрерывному формированию аналитических показателей, и такой подход предполагает внедрение инструментов автоматизированного сбора первичных данных: датчиков отслеживания прогресса работ, систем видеоаналитики, телеметрии строительной техники, интеграции с бухгалтерскими и складскими системами, анализа рабочих коммуникаций. В результате формируется цифровая среда, где данные

становятся не эпизодическим информационным ресурсом, а постоянным управленческим фактором.

Существенные изменения происходят и в методологии аналитической обработки данных. Традиционные методы, основанные на экспертных оценках и детерминированных моделях, дополняются и во многих случаях заменяются алгоритмами статистического моделирования, кластерного анализа, интеллектуальной оптимизации и машинного обучения. Они обеспечивают выявление скрытых зависимостей, прогнозирование отклонений и повышение точности планирования. Особенно важным становится использование решений, позволяющих выявлять сложные нелинейные связи между факторами. Так, динамические изменения цен на строительные материалы, ритмичность поставок, сезонные колебания погодных условий, поведение подрядчиков и степень загруженности техники формируют сложную схему факторов риска, которую невозможно описать традиционными моделями.

Важно также привести нормативные акты и государственные программы РФ, которые продемонстрируют ситуацию на уровне государственного управления. В России цифровизация строительной отрасли поддерживается на уровне национальных стратегий и нормативной базы. Уже в 2015 году были утверждены планы по поэтапному внедрению BIM в гражданском строительстве, а в 2018 году Указ Президента № 204 поставил задачи «цифрового строительства» и замены устаревших практик к 2024 году. По итогам последующих лет приняты важные решения: правительственное постановление № 1431 (2020) и приказ Минстроя № 614 (2024) определили правила формирования и содержания информационных моделей объектов. С 1 января 2022 года применение ТИМ стало обязательным на всех объектах госконтракта, а с июля 2024 года – и для крупных объектов долевого строительства и подрядчиков по городским заказам. Министерство строительства запустило проект «цифровой вертикали отрасли» – единую информационную систему управления стройками. Советник главы Минстроя Е. Звонарева сообщала, что к концу 2023 года в этот проект включены уже 74 региона. Планируется перевести весь документооборот в юридически значимый цифровой вид и создать единое информационное пространство для коммуникации между участниками. Законом № 653-ФЗ (2023) введён единый реестр требований - машиночитаемый свод норм, позволяющий автоматически сверять проектные решения с актуальными нормативами. В комплексе с «ГИС Стройкомплекс.РФ», единой цифровой платформой с классификаторами объектов, ресурсов и требований, это закладывает

основу для автоматизации проектирования: чиновники прогнозируют, что в перспективе специализированные ПО смогут выполнять порядка 90% рутинной проектной работы, автоматически подбирая оптимальные решения и сметные затраты.

Ведущие российские компании и региональные правительства уже демонстрируют практические результаты цифровизации. Например, корпорация «Дом.РФ» вместе с крупнейшими девелоперами создала отраслевое хранилище данных для обучения ИИ на основе Единой информационной системы жилищного строительства. В девелопменте появляются внутренние AI-решения: «Смарт Девелопмент Групп» создала ИИ-агента для ответов по внутренним регламентам и анализа данных. Компания «Самолёт» развила более десятка ИИ-моделей для строительства: системы контроля безопасности на площадках, мониторинга темпов работ, оценки объёмов, а также онлайн-сервис показа готовой отделки квартир в реальном времени. Такие решения помогают эффективно координировать крупные жилые комплексы. Практика внедрения BIM дает заметный эффект: так, по данным собеседника RBC Trends, при применении BIM-технологий и обнаружении коллизий один проект получил допвыручку порядка 160 млн рублей. Аналитики Dom.RF указывают, что около 17% цифровых решений в отрасли уже используют ИИ, причем наиболее активно – на этапе эксплуатации жилья, примерно треть решений связаны с «умным домом». [8]

С учётом российских реалий важно отметить статистику внедрения технологий. Исследование рынка показало, что в более чем 80 % строительных компаний бюджеты на цифровизацию не превышают 100 млн руб. в год – это меньше 1 % выручки. При этом 70 % девелоперов привлекают сторонние решения и подрядчиков. Основные приоритеты – цифровизация закупок, проектирования, документооборота, а также управления сроками и контроля качества. Большинство компаний (более 80 %) отмечают реальный измеримый эффект от цифровизации за последние 2–3 года. Например, перевод процессов в цифру в Тюменской области позволил сократить бумажную работу и резко ускорить обработку информации.

Зарубежные практики демонстрируют схожую логику развития. В Китае государственные корпорации по транспортному строительству повсеместно применяют BIM в инфраструктурных проектах. Москва даже заключила договор с крупнейшей китайской компанией CRCC об обмене технологиями BIM – к 2020 году в столице планируется полностью перейти на BIM для метростроя. В ОАЭ государство давно продвигает 3D-моделирование: новый терминал аэропорта Абу-Даби

реализован по требованию клиента полностью на BIM-платформе, что обеспечило координацию всех участников и снизило риски проекта. Европейские страны также жестко интегрируют цифровые стандарты: с 2014 года ЕС поощряет использование ИКТ при госзакупках, а правительства ряда стран, включая Великобританию, сделали применение BIM обязательным для госзаказа. Эти примеры демонстрируют, что мировые лидеры отрасли ставят цифровые технологии в основу управления проектами.

Опубликованные исследования свидетельствуют о постепенном росте цифризации, хотя до полного покрытия еще далеко. По данным совместного исследования «Дом.РФ» и Фонда «Сколково», сейчас лишь порядка 36% девелоперов применяют цифровые инструменты на этапе продаж и 30% – на этапе эксплуатации, тогда как на стадии инвестиционного планирования их используют единицы. При этом более 90% компаний предпочитают приобретать готовые IT-решения сторонних поставщиков, и лишь лидеры отрасли разрабатывают собственные системы. В рамках программы «умный город» упор делается на крупные города: по цифровым индексам лидерами считаются Москва, Санкт-Петербург, Екатеринбург и другие центры с развитой инфраструктурой и высоким уровнем инвестиций. При этом Минстрой планирует охватить «новые территории» и интегрировать в систему Росреестр и картографические данные, что пока требует дополнительных ресурсов.

Но внедрение цифровых аналитических решений сталкивается с рядом препятствий. Во-первых, сказывается консерватизм отрасли и сопротивление изменениям. Многие участники привыкли к бумажному документообороту и настороженно относятся к новым подходам. По оценке экспертов, основные барьеры – дефицит квалифицированных специалистов и финансовые ограничения, не позволяющие инвестировать в софт и IT-инфраструктуру. Во-вторых, все еще остаются значительной проблемой технические проблемы: отрасль страдает от «неоцифрованности» данных – разнородные чертежи, устаревшие планы, нерегулярные отчеты со стройки мешают обучению ИИ и аналитики. Интеграция новых систем с существующими процессами требует времени и больших доработок – нередко необходимо полностью перестраивать ИТ-инфраструктуру, что сопряжено с дополнительными затратами и рисками. Кроме того, запускаемые проекты «цифровой вертикали» зависят от синхронизации с другими госреестрами: на практике Минстрой отмечает, что требуется плотная работа с Росреестром, наполнением баз данными и вводом градостроительной документации в электронный вид, что несколько замедляет темпы

запуска системы в новых регионах. Наконец, вызывает затруднения импортозамещение, хотя отрасль ориентируется на отечественный софт, например, цифровой двойник депо «Бирюлёво» был создан на российской платформе, многие инструменты BIM традиционно зарубежные, и переход к локальным решениям требует дополнительных стандартов и обучения персонала.

Особое внимание в профессиональной среде уделяется вопросам интеграции аналитических систем. Цифровая трансформация аналитики невозможна без согласованного обмена данными между разнородными программными продуктами, используемыми проектировщиками, генподрядчиками, субподрядными организациями, поставщиками и контролирующими структурами. Проблема фрагментации информационных потоков остаётся одной из главных причин задержек, ошибок в оценке объемов работ и неточностей в прогнозировании сроков. В ответ на это разрабатываются стандарты обмена аналитическими данными, внедряются унифицированные форматы проектной информации и создаются открытые интерфейсы для интеграции различных систем. Рост числа таких решений расширяет возможности для построения комплексных аналитических моделей, охватывающих всю цепочку создания строительной продукции. [2]

Несмотря на сложности, цифровизация демонстрирует явные плюсы для экономики строительства. Реализация цифровых стратегий обещает существенное ускорение проектных работ, сокращение затрат на материалы и устранение «узких мест» за счет проактивного планирования. Как показывает практика, уже на уровне одного проекта внедрение BIM-моделей позволило выявить конструктивные ошибки, приведшие к дополнительной выручке в сотни миллионов рублей. Мировой опыт говорит о том, что использование цифровых двойников и BIM способно ускорять проект на 20–30%. Аналитики Минстроя сходятся во мнении, что цифровая «вертикаль» позволит оптимизировать процессы, тем самым снизить временные издержки и минимизировать ошибки при строительстве. В долгосрочной перспективе внедрение ИИ и машинной аналитики дает расчёт на перераспределение рутинных функций от человека к машинам – по прогнозам экспертов, «программные комплексы смогут выполнять 90% рутинного проектирования», включая расчёт стоимости элементов и подбор нормативов. Цифровые технологии именно так создают основу для более точного прогнозирования бюджета и сроков, а также для эффективного управления качеством. Можно предположить и значительные кадровые изменения, связанные с этим процессом.

Государство активно регулирует и поддерживает переход к цифровому строительству. Ведутся работы по введению новых технических регламентов и стандартов, синхронизированных с цифровыми системами, создаются единые классификаторы и термины. Особое внимание уделяется импортозамещению программных решений: в рамках «Стройкомплекса» и муниципальных проектов стимулируется использование отечественных BIM-платформ и аналитических сервисов. Например, московские власти развернули сотрудничество с отечественными разработчиками, а система «Стройкомплекс.Москва» переходит на российские программные продукты, о чем официально заявлялось в 2023 году. Подобная политика позволит снизить зависимость от зарубежного ПО и адаптировать инструменты под российские требования, что особенно важно при учёте санкционных рисков. Одновременно государство инвестирует в подготовку кадров и формирует фонды поддержки цифровых стартапов в области строительства. В итоге государственное регулирование и меры по локализации технологий призваны создать благоприятные условия, при которых эффективность цифровизации будет максимальной при минимальных рисках для отрасли.

Рынок цифровизации в строительстве стремительно растёт: по прогнозам, к 2028 г. его объём увеличится вчетверо. Импортозамещение и государственная поддержка дают толчок развитию отечественного ПО. Применение продвинутой аналитики уже приносит конкретные эффекты: как отмечает Минстрой, цифровые решения позволяют сократить затраты на строительство на 7–10 % и ускорить ввод объектов на 10–20 %. Достигается это за счёт уменьшения избыточных запасов, оптимизации логистики и качества планирования. Внедрение ИИ и машинного обучения делает управление проактивным: руководители вместо разрозненных отчётов получают единые дашборды с объективной картиной проектов и могут управлять ими на основе предиктивной аналитики. Таким образом, цифровая трансформация аналитики обеспечивает компаниям конкурентное преимущество и открывает доступ к крупнейшим строительным контрактам.

Заключение

Цифровая трансформация аналитики в строительстве уже перестала быть концептом будущего и становится насущной необходимостью. Применение BI-платформ, Big Data, ИИ, цифровых двойников и BIM/ТИМ позволяет российским компаниям значительно повысить эффективность проектов: сократить расходы, сдать объекты раньше срока и повысить качество строительства. Но для этого требуется не только внедрение технологий, но и развитие кадровых компетенций, улучшение

регуляторного механизма и адекватное финансирование инициатив. С принятием государственных стандартов и запуском отраслевых платформ создаются предпосылки для масштабного роста данных и аналитики. В перспективе цифровая модель управления – с единой средой данных и алгоритмическими инструментами – должна стать отраслевым стандартом. Те компании, которые осваивают эту модель в первую очередь, смогут существенно выиграть в конкурентной гонке и вывести российскую строительную отрасль на новый уровень прозрачности и технологической зрелости.

Литература

1. BN.ru. Запущена базовая версия платформы данных для строительства (новость) [Электронный ресурс]. — 25.11.2025. — Режим доступа: https://www.bn.ru/gazeta/news/284224/ (дата обращения: 05.12.2025).
2. CNews. «Самолёт» перешел на отечественную BI-платформу Visiology для анализа данных и принятия решений [Электронный ресурс]. — 22.07.2025. — Режим доступа: https://www.cnews.ru/news/line/2025-07-22_samolet_pereshel_na_otechestvennuyu (дата обращения: 05.12.2025).
3. CNews. ГК «Кортрос» внедрила систему продвинутой аналитики «Дельта BI» [Электронный ресурс]. — 17.09.2025. — Режим доступа: https://www.cnews.ru/news/line/2025-09-17_gk_kortros_vnedrila_sistemu (дата обращения: 05.12.2025).
4. MoJcement.ru. Объём работ в строительстве в РФ вырос на 2,5 % (январь–сентябрь 2024) [Электронный ресурс]. — 01.11.2024. — Режим доступа: https://jcement.ru/content/news/obem-rabot-v-stroitelstve-v-rf-vyros-na-2-5-/ (дата обращения: 02.12.2025).
5. NormaCS.info. «Строительная отрасль-2025: мы оттолкнулись от дна?» — отраслевой обзор [Электронный ресурс]. — 2025. — Режим доступа: https://www.normacs.info/articles/1149 (дата обращения: 02.12.2025).
6. Russia Smart City / SMLT. Цифровизация строительной отрасли — аналитический отчёт 2024 (PDF) [Электронный ресурс]. — 11.09.2024. — Режим доступа: https://russiasmartcity.ru/wp-content/uploads/2024/09/SMLT_Цифровизация_строительства_2024.pdf (дата обращения: 05.12.2025).
7. Sk.ru (Фонд «Сколково»). «Сколково и ДОМ.РФ оценят цифровую зрелость девелоперов» (новость) [Электронный ресурс]. — 25.10.2024. — Режим доступа: https://sk.ru/news/skolkovo-i-domrf-ocenyat-cifrovuyu-zrelost-developerov/ (дата обращения: 05.12.2025).

8. Stroygaz.ru. Цифровизация строительной отрасли в 2024 году (обзор) [Электронный ресурс]. — 01.11.2024. — Режим доступа: https://stroygaz.ru/publication/technologies/tsifrovizatsiya-stroitelnoy-otrasli-v-2024-godu/ (дата обращения: 05.12.2025).

9. ДОМ.РФ. Исследование «Цифровизация девелопмента» (ДОМ.РФ и Фонд «Сколково») [Электронный ресурс] / ДОМ.РФ. — 2025. — Режим доступа: https://наш.дом.рф/site/binaries/content/assets/domrf/тим/аналитические-материалы/исследование_цифровизации_девелопмента_2025.pdf (дата обращения: 05.12.2025).

10. РА (Национальное рейтинговое агентство). «СТОП, КРАН! Российская строительная отрасль: обзор сентябрь 2025» [Электронный ресурс]. — 2025. — Режим доступа: https://www.ra-national.ru/wp-content/uploads/2025/09/nra_stroitelnaja_otrasl_obzor_sentjabr_2025_verst.pdf (дата обращения: 02.12.2025).

11. Полипластик. «Строительство не сможет удержать темпы развития» — обзор 2025 года [Электронный ресурс]. — 14.04.2025. — Режим доступа: https://www.polyplastic.ru/press/news/2025/04/14/item22885 (дата обращения: 02.12.2025).

12. Росстат. Строительство. Выпущенные публикации. Объём работ по виду деятельности «Строительство» (публикации 2024–2025) [Электронный ресурс] / Федеральная служба государственной статистики (Росстат). — 2025. — Режим доступа: https://39.rosstat.gov.ru/publications_2/document/258338 (дата обращения: 05.12.2025).

13. Сайт «fomag.ru». Объём строительной отрасли РФ вырос на 2,1 % по итогам 2024 года — Росстат [Электронный ресурс]. — 07.02.2025. — Режим доступа: https://fomag.ru/news-streem/obem-stroitelnoy-otrasli-rf-vyros-na-2-1-po-itogam-2024-goda-rosstat/ (дата обращения: 02.12.2025).

14. ТАСС. В РФ объём строительных работ вырос в феврале на 11,9 % за год [Электронный ресурс]. — М., 02.04.2025. — Режим доступа: https://tass.ru/ekonomika/23580285 (дата обращения: 02.12.2025).

15. ТАСС. Объём работ в строительстве в РФ вырос на 7,4 % в январе 2025 года [Электронный ресурс]. — М., 05.03.2025. — Режим доступа: https://tass.ru/ekonomika/23317657 (дата обращения: 02.12.2025).

16. ТАСС. Объём строительной отрасли РФ вырос на 2,1 % по итогам 2024 года [Электронный ресурс]. — М., 07.02.2025. — Режим доступа: https://tass.ru/ekonomika/23084423 (дата обращения: 02.12.2025).

17. ТАСС. Хуснуллин заявил о росте объёма стройки в РФ более чем на треть с 2019 года [Электронный ресурс]. — М., 07.08.2024. — Режим доступа: https://tass.ru/ekonomika/21550477 (дата обращения: 12.12.2025).

Разработка системы визуализации позиционирования заготовок в роботизированной производственной ячейке

С. А. Ремнева, email: sofyaremneva@yandex.ru[1]
Н. А. Бондаренко, email: nik.bondarenko.farstayy@gmail.com [2]
А. А. Филимонова, email: filimonova_aa@inbox.ru [2]

[1] Воронежский государственный технический университет

Аннотация. *В статье представлена система веб-визуализации для контроля и корректировки положения заготовок в роботизированной сварочной ячейке. Разработанное решение обеспечивает оператору возможность интерактивного сопоставления CAD-модели с реальным положением детали и внесения пространственных корректировок при обнаружении расхождений. Система реализована на основе клиент-серверной архитектуры с использованием Python, FastAPI и геометрического ядра Open Cascade Technology на стороне сервера, а также React и Three.js на стороне клиента. Приведено обоснование выбранного технологического стека и описаны принципы работы системы, включая обработку STEP-моделей, организацию 3D-сцены и интерактивное управление. Разработанный инструмент служит основой для интеграции в контур адаптивной сварки и создания цифрового двойника производственной ячейки.*

Ключевые слова: *роботизированная сварка, веб-визуализация, позиционирование заготовок, CAD-модели, Open Cascade Technology, цифровой двойник.*

Введение

В современных роботизированных комплексах для финишной сварки ключевым этапом является точное позиционирование заготовки в рабочей зоне. Производственные системы используют 3D-сканеры для оцифровки физической детали и последующего сопоставления (матчинга) полученного облака точек с эталонной CAD-моделью. Успешность этого процесса непосредственно определяет корректность генерации траектории для робота-манипулятора. Однако на практике возможны ситуации значительного расхождения между реальным положением детали и её цифровой моделью, что требует ручного вмешательства оператора для корректировки позиции виртуального объекта. В связи с этим возникает научно-практическая задача создания специализированного инструмента веб-визуализации, который позволил

© Ремнева С. А., Бондаренко Н. А., Филимонова А. А., 2026

бы оператору в режиме реального времени контролировать процесс сканирования, оценивать результаты матчинга и при необходимости вносить точные корректировки в положение виртуальной модели относительно как сканированного облака точек, так и системы координат робота. Такой инструмент должен обеспечивать не только наглядное представление, но и гарантировать соответствие визуализации действительному положению детали в пространстве ячейки, выступая тем самым критически важным элементом человеко-машинного интерфейса в адаптивном сварочном контуре.

1. Архитектура и компоненты цифрового двойника производственной ячейки

Роботизированная сварочная ячейка выступающая объектом разработки цифрового двойника ориентирована на выполнение автоматизированных операций по финишной обварке крупногабаритных пресобранных металлоконструкций высокой массы. Основное назначение системы заключается в обеспечении стабильных качественных характеристик сварных соединений и существенном росте производительности труда за счет минимизации участия человека в опасных и трудоемких производственных циклах. Общее пространство роботизированного комплекса занимает площадь 12х9 метров что позволяет разместить все необходимое периферийное оборудование и обеспечить безопасные зоны обслуживания. При этом непосредственно зона обработки заготовок ограничена прямоугольным участком размером 9х3 метра расположенным в центральной части ячейки что диктуется спецификой габаритов целевых изделий и кинематическими возможностями оборудования.

Технологическую основу комплекса составляет сверхмощный шестиосевой промышленный робот манипулятор KUKA KR 1000 серии KR C4 Titan. Выбор данного типа оборудования обусловлен необходимостью работы с массивными конструкциями и использования специализированного тяжелого сварочного инструментария обеспечивающего высокую жесткость и точность позиционирования в условиях динамических нагрузок. Робот смонтирован на модуле линейного перемещения что значительно расширяет его рабочую область вдоль длинной стороны зоны обработки и позволяет беспрепятственно перемещаться между различными участками заготовки. Оснащение манипулятора включает в себя высокотехнологичный комплекс сварочного оборудования Fronius в состав которого входят источник питания блок подачи проволоки и специализированная горелка адаптированная для продолжительных циклов эксплуатации.

Для реализации механизмов адаптивного управления и построения точного цифрового двойника ячейка интегрирована с комплексом систем машинного зрения. В состав измерительного модуля входит 3D сканер PhoXi предназначенный для первичной оцифровки геометрии и определения пространственного положения заготовки а также лазерный сканер ScanSPA осуществляющий мониторинг параметров сварного шва непосредственно в процессе формирования соединения. Поддержание работоспособности инструмента обеспечивается автоматической станцией очистки горелки которая в регламентном режиме производит удаление брызг металла и обновление антипригарного покрытия. Технологический процесс функционирования ячейки организован как строгая последовательность операций начинающаяся с анализа облака точек и сопоставления реального положения детали с эталонной моделью. На основе полученных данных система управления корректирует траекторию движения манипулятора учитывая фактические отклонения сборки. Завершение каждого цикла сварки сопровождается процедурой автоматического обслуживания рабочего органа что гарантирует стабильность характеристик при многосменной работе комплекса. Таким образом архитектура ячейки на базе робота Титан позволяет решать задачи по сварке особо тяжелых узлов обеспечивая гибкость переналадки и высокую степень автоматизации пространственного контроля.

2. Архитектурное решение и технологический стек реализации

Для решения поставленной задачи был выбран подход на основе веб-технологий, обеспечивающий кроссплатформенность, лёгкость развёртывания и возможность интеграции с существующими промышленными системами. Архитектура решения строится по классической двухуровневой клиент-серверной модели, где серверная часть отвечает за обработку геометрических данных, а клиентская предоставляет интерактивный интерфейс визуализации в браузере.

Выбор технологического стека обусловлен долгосрочной целью создания специализированного программного обеспечения для робототехнических комплексов (ПО для РТК), способного обеспечивать полный цикл работы — от загрузки и обработки CAD-моделей до управления технологическими процессами. Подобное ПО требует применения промышленных инструментов для обеспечения надёжности, производительности и масштабируемости при работе со сложными производственными данными.

Серверный компонент разработан с использованием языка Python и фреймворка FastAPI, выбранных благодаря их асинхронной архитектуре, подходящей для обработки длительных операций ввода-вывода и

одновременного обслуживания множества клиентов. Для обработки геометрических данных использована библиотека pythonocc-core, представляющая собой Python-привязки к геометрическому ядру Open Cascade Technology (OCCT) — промышленной платформе, обеспечивающей полную поддержку стандарта STEP (ISO 10303-21). Выбор OCCT обусловлен требованием корректной обработки сложных многокомпонентных сборок с сохранением иерархии и атрибутов, что соответствует реальным потребностям при подготовке управляющих программ для сварки.

Обработка данных реализована в виде конвейера: приём и валидация STEP-файла через REST-эндпоинт, его парсинг с использованием STEPControl_Reader, декомпозиция сборки на твёрдые тела через TopExp_Explorer и тесселяция алгоритмом BRepMesh_IncrementalMesh с параметром отклонения 0,05 мм. Результатом тесселяции является файл в формате STL, выбранном благодаря его универсальности и совместимости с веб-технологиями.

Для управления данными технологической подготовки используется реляционная база данных PostgreSQL с асинхронным драйвером asyncpg. Такой выбор обеспечивает целостность данных, поддержку сложных запросов и закладывает основу для работы с полным объёмом производственной информации, включая библиотеки моделей, параметры сварки, историю корректировок и связи между технологическими операциями.

Клиентский компонент представляет собой веб-приложение на React, взаимодействующее с сервером через REST API. Для 3D-визуализации используется библиотека Three.js, интегрированная через @react-three/fiber и @react-three/drei. Данный стек позволяет создавать интерактивные интерфейсы для пространственного манипулирования сложными геометрическими объектами прямо в браузере, без необходимости установки дополнительного программного обеспечения.

Предложенная архитектура не только решает конкретную задачу визуализации и корректировки позиционирования, но и формирует фундамент для разработки полнофункционального программного обеспечения управления РТК, способного в будущем охватывать весь технологический цикл — от загрузки моделей и настройки параметров сварки до планирования заданий и оперативного контроля выполнения работ.

3. Реализация системы визуализации и принцип ее работы

Программная реализация модуля графического отображения и анализа конструктивной электронной структуры изделия выполнена с использованием библиотеки Three.js, предоставляющей

высокоуровневый интерфейс к графическому API WebGL. Данный подход обеспечивает возможность рендеринга сложной трехмерной графики непосредственно в среде веб-браузера, используя аппаратное ускорение видеокарты клиента, без необходимости установки дополнительных плагинов или специализированного программного обеспечения. Инициализация графического ядра происходит при монтировании основного компонента редактора, в ходе которого создается виртуальное сценическое пространство. Для обеспечения корректного восприятия пространственной глубины и масштаба объектов в сцену интегрированы вспомогательные визуальные инструменты: сетка плоскости основания и оси координат, соответствующие мировому пространству. Визуализация осуществляется через перспективную камеру, параметры которой, включая угол обзора и диапазоны отсечения, настроены для оптимального отображения объектов машиностроения средних габаритов. Освещение сцены реализовано по комбинированной схеме, включающей источник рассеянного света для заполнения теней и направленный источник, создающий объем и блики на гранях моделей, что облегчает визуальную идентификацию формы оператором.

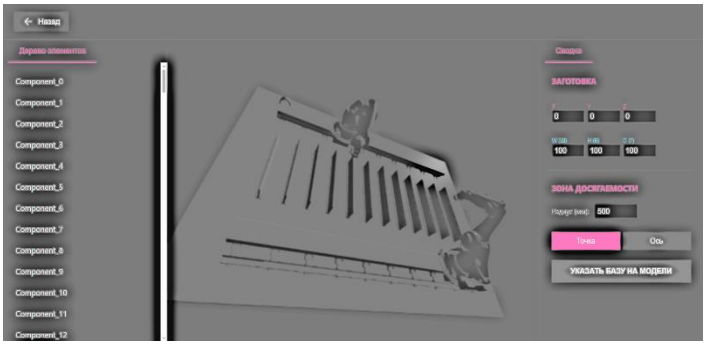


Рис. 1. Общий вид интерфейса приложения с загруженной сборкой и деревом элементов

Загрузка геометрических данных в приложение реализована как многоступенчатый асинхронный процесс. На первом этапе клиентская часть приложения обращается к серверному API для получения метаданных о составе модели, включая иерархию компонентов и ссылки на геометрические файлы. После получения списка компонентов инициируется параллельная загрузка геометрии, поддерживающая форматы STL и OBJ. Для каждого загружаемого файла создается полигональная сетка (Mesh), которой присваивается уникальный

идентификатор, соответствующий структуре исходного обменного файла. В процессе импорта производится нормализация геометрии: вычисляются габаритные контейнеры (Bounding Box) каждого элемента, определяется их геометрический центр, и производится соответствующая трансформация вершин. Это позволяет корректно позиционировать детали относительно друг друга в единой системе координат сборки, даже если исходные файлы имели смещения. Для оптимизации производительности рендеринга ссылки на все созданные трехмерные объекты сохраняются в специализированной структуре данных (Map), что обеспечивает доступ к любому элементу по его идентификатору за константное время.

Важнейшим элементом человеко-машинного взаимодействия в разработанной системе является механизм селекции объектов. Для реализации выбора деталей на трехмерной сцене используется метод отслеживания лучей (Raycasting). При клике пользователя мышью по области холста, координаты курсора преобразуются в нормализованное пространство устройства, после чего из позиции камеры через точку клика проецируется луч. Система анализирует пересечения данного луча со всеми полигональными сетками в сцене и определяет ближайший к наблюдателю объект. Идентификация выбранного элемента позволяет реализовать двустороннюю синхронизацию интерфейса: при клике на модель в 3D-окне автоматически подсвечивается и прокручивается соответствующий пункт в иерархическом дереве компонентов, и, наоборот, выбор элемента в дереве фокусирует камеру и выделяет объект на сцене. Визуальная индикация выбора реализована через динамическое изменение свойств материала объекта, в частности, наложение эмиссионного свечения, что позволяет оператору однозначно идентифицировать активный компонент даже внутри плотной компоновки сборки.

Помимо функций просмотра, разработанное приложение включает в себя функциональный блок моделирования рабочей зоны технологического оборудования, что критически важно для задач подготовки производства. В сцену введен управляемый параметрический объект — параллелепипед, имитирующий заготовку или зону обработки. Управление положением и габаритами данного объекта реализовано гибридным способом: оператор может вводить точные численные значения размеров (ширина, высота, глубина) и координат через интерфейсные панели, либо осуществлять интерактивное перемещение объекта в плоскости с помощью клавиатурных команд, что эмулирует ручное управление положением инструмента. Система в реальном

времени обновляет матрицу трансформации объекта, пересчитывая положение его вершин в мировом пространстве.

Ключевой алгоритмической особенностью системы является модуль расчета и визуализации зоны досягаемости робота-манипулятора. Данный функционал позволяет оценить возможность обработки конкретной детали или узла без физических испытаний. Зона досягаемости формируется динамически на основе выбранных пользователем базовых компонентов конструкции. Алгоритм вычисляет объединенный габаритный контейнер для группы выбранных деталей, определяет его геометрический центр и, в зависимости от режима работы, строит визуализацию зоны. Реализовано два режима построения зоны: точечный и осевой. В точечном режиме зона представляется в виде полупрозрачной сферы заданного радиуса с центром в геометрическом центре выбранной базы. В осевом режиме, предназначенном для длинномерных узлов или линейных осей робота, система автоматически определяет доминирующее измерение габаритного контейнера и генерирует зону в форме капсулы. Капсула представляет собой цилиндр с полусферическими окончаниями, длина и ориентация которого соответствуют габаритам базовой детали, а радиус задается параметрами робота.

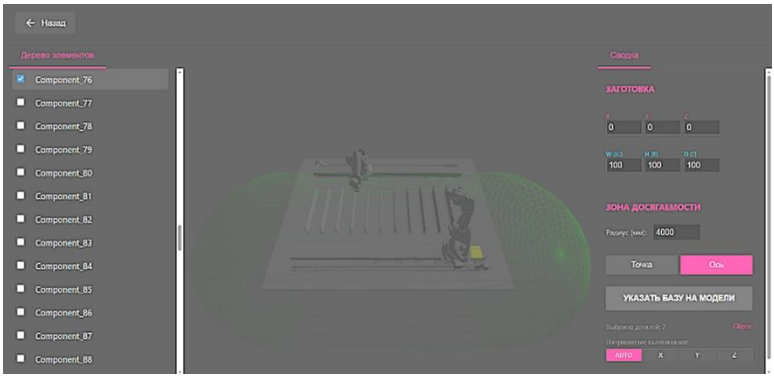


Рис. 2. Визуализация зоны досягаемости в режиме «Ось» вокруг выбранной детали и панель настроек зоны справа

Для автоматизации проверки технологичности реализован алгоритм детекции коллизий и вхождения в зону. Программа непрерывно анализирует взаимное расположение имитатора заготовки и зоны досягаемости. Алгоритм проверки базируется на вычислении дистанции от восьми вершин габаритного параллелепипеда заготовки до

геометрического примитива зоны (сферы или осевого сегмента капсулы). В случае капсулы используется математический аппарат нахождения кратчайшего расстояния от точки до отрезка прямой в трехмерном пространстве. На основе количества вершин, находящихся внутри допустимого радиуса, система формирует визуальную обратную связь: если объект полностью находится в зоне, он окрашивается в зеленый цвет; если частично — в желтый; если полностью выходит за пределы — в красный. Данная цветовая индикация позволяет технологу мгновенно оценивать корректность позиционирования и необходимость корректировки технологического процесса.

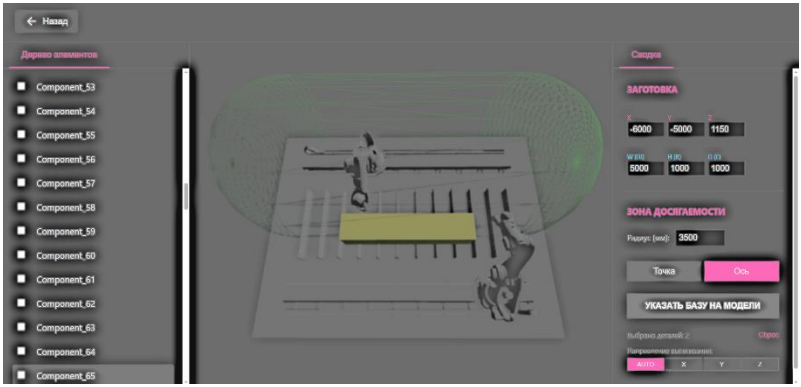


Рис. 3. Пример работы алгоритма проверки: заготовка, окрашенная в желтый при частичном выходе из полупрозрачной зоны досягаемости

Архитектура приложения построена на принципах реактивного программирования. Любые изменения параметров, будь то радиус зоны досягаемости, размеры заготовки или выбор базовых компонентов, вызывают мгновенный пересчет геометрических параметров и обновление сцены. Это достигается за счет использования хуков состояния и эффектов React, которые отслеживают зависимости переменных и триггерят перерисовку только необходимых элементов сцены, минимизируя нагрузку на процессор. Для обеспечения плавности анимации и отзывчивости интерфейса при навигации используется цикл рендеринга, привязанный к частоте обновления экрана. Таким образом, разработанный модуль не просто визуализирует статичную геометрию, а представляет собой интерактивный инструмент пространственного анализа, позволяющий на этапе подготовки производства выявлять

потенциальные проблемы геометрической несовместимости оборудования и изделия.

Заключение

Разработанная система веб-визуализации представляет собой эффективный инструмент для контроля и корректировки положения заготовок в роботизированной сварочной ячейке. Интеграция технологий React, Three.js и Open Cascade Technology обеспечила создание интуитивно понятного и производительного интерфейса, позволяющего оператору взаимодействовать с цифровым двойником ячейки в реальном времени. Система не только решает задачу визуализации и ручной коррекции, но и служит основой для замкнутого адаптивного контура управления, где цифровой двойник выступает связующим звеном между человеческим решением и автоматизированным исполнением. Перспективы развития системы связаны с её дальнейшей интеграцией в промышленные контуры управления, расширением функционала интерактивной трансформации моделей и реализацией полномасштабного цифрового двойника, способного к самооптимизации в условиях реального производства.

Список литературы

1. Коичи, М. WebGL : программирование трехмерной графики / М. Коичи, Л. Роджер ; пер. с англ. А. Н. Киселев. — Москва : ДМК Пресс, 2015. — 494 с. — ISBN 978-5-97060-146-4.

2. Лупачев, В. Г. Механизация и автоматизация сварочного производства : учеб. пособие / В. Г. Лупачев. — Минск : РИПО, 2021. — 346 с. — ISBN 978-985-7253-62-3.

3. Хортон, А. Разработка веб-приложений в ReactJS / А. Хортон, Р. Вайс ; пер. с англ. Р. Н. Рагимова. — Москва : ДМК Пресс, 2016. — 254 с. — ISBN 978-5-94074-819-9.

4. pythonocc-core: Python bindings for OpenCASCADE [Электронный ресурс]. — Режим доступа: https://dev.opencascade.org/project/pythonocc (дата обращения: 28.12.2025).

Секция 9
Инфокоммуникационные системы, системы телекоммуникаций, радиоэлектроника, обработка сигналов

Электродинамический анализ сверхширокополосной конической щелевой антенны для мобильного радиоэлектронного комплекса

А. В. Богословский [1], e-mail: razinkovsergey@rambler.ru,
Д. Н. Борисов [2], e-mail: email: borisov@sc.vsu.ru,
С. Н. Разиньков [1], e-mail: bogosandrej@yandex.ru

[1] Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н. Е. Жуковского и Ю. А. Гагарина» (г. Воронеж)
[2] Воронежский государственный университет

Аннотация. *На основе краевой задачи электродинамики для нерегулярной двухсторонней щелевой линии с присоединенной круглой щелевой полостью построена модель сверхширокополосной конической щелевой антенны для мобильного радиоэлектронного комплекса. С применением компьютерной программы электродинамического моделирования CST Microwave Studio проведен анализ коэффициента усиления и коэффициента стоячей волны (по напряжению) указанной антенны. Продемонстрированы возможности достижения высоких показателей пространственно-частотной избирательности передачи-приема сигналов при ограниченных массогабаритных характеристиках антенны.*

Ключевые слова: *коническая щелевая антенна, резонатор, раскрыв, область преобразования импеданса, коэффициент усиления антенны, коэффициент стоячей волны, электродинамическая модель*

Введение

Разработка сверхширокополосных (частотно-независимых) антенн для мобильных радиоэлектронных комплексов [1] включает совместное решение задач по достижению высоких показателей пространственно-частотной избирательности передачи (приема) сигналов и сокращению массогабаритных характеристик для размещения на носителях [2, 3].

В диапазоне сверхвысоких частот находят применение конические щелевые антенны. В отличие от волноводных структур они не обладают значительными размерами и не имеют жестких ограничений на полосы рабочих частот, обусловленных возбуждением высоких гармоник токов. В отличие от приемоизлучающих структур с прямоугольными щелями для согласования антенн указанного типа с распределительными

© Богословский А. В., Борисов Д. Н., Разиньков С. Н., 2026

(питающими) линиями не требуются сложные в исполнении устройства трансформации. В отличие от полосковых антенн конические щелевые антенны характеризуются повышенными показателями направленности при близких значениях коэффициентов перекрытия диапазонов рабочих частот [2, 4]. За счет монотонного изменения профиля щели импеданс антенны в области раскрыва преобразуется к значению, установленному для согласующего устройства с симметрирующим трансформатором [4]. При этом ввиду предельно малых потерь энергии сигналов в области преобразования импеданса обеспечивается ее высокая концентрация в секторе рабочих углов при излучении и достигается значительное усиление волновых процессов, возбуждающих антенную нагрузку, в режиме приема.

Первый вариант сверхширокополосной планарной (коллинеарной) антенны с шириной щели, изменяющейся по экспоненциальному закону, получившей название TSA – Tapered Slot Antenna, разработан в 1979 году.

Диапазонные и направленные свойства сверхширокополосной конической щелевой антенны [2, 4] находятся по результатам анализа функциональных взаимосвязей комплексных амплитуд токов в областях раскрыва, преобразования импеданса и в резонаторе с пространственно-частотным (пространственно-временным) распределением стороннего возбуждающего электромагнитного поля на основе решения краевой задачи электродинамики [3]. Эти взаимосвязи составляют методическую базу для восстановления параметров приемоизлучающей конструкции с требуемыми показателями пространственно-частотной избирательности передачи (приема) сигналов.

В предлагаемой работе с применением компьютерной программы электродинамического моделирования CST Microwave Studio построена электродинамическая модель и проведен анализ сверхширокополосной конической щелевой антенны для мобильного радиоэлектронного комплекса.

Цель работы – исследование показателей пространственно-частотной избирательности конической щелевой антенны при передаче-приеме сигналов в сверхширокополосном радиоэлектронном комплексе.

1. Основные положения электродинамической модели сверхширокополосной конической щелевой антенны

При построении электродинамической модели конической щелевой антенны будем полагать, что ее конструкция, выполненная из идеально проводящего материала, нанесена на диэлектрическую подложку в виде печатной платы.

Базовыми элементами антенны для излучения (приема) сигналов в диапазоне рабочих частот сверхширокополосного радиоэлектронного комплекса являются:

– нерегулярная двухсторонняя щелевая линия с круглой щелевой полостью;

– питающий фидер типа нерегулярной полосковой линии со шлейфом на основе радиальной линии.

Вследствие выбора конфигурации питающего фидера со шлейфом реализуется щелевой способ возбуждения антенны при малых уровнях паразитного излучения (приема) волновых процессов, что позволяет реализовать требуемые диапазонные и направленные свойства. Профиль нерегулярной щелевой линии с полостью концентрирует энергию сигналов внутри [2], способствуя подавлению негативных проявлений антенного эффекта и достижению высоких показателей направленности передачи-приема [5, 6].

Для оценки пространственно-частотной избирательности передачи-приема сигналов проведем расчет коэффициента усиления антенны, для характеристики качества согласования с нагрузкой будем использовать коэффициент стоячей волны (по напряжению).

При построении электродинамической модели конической щелевой антенны нерегулярная двухсторонняя щелевая линия разделяется на два участка – раскрыв и область преобразования импеданса, подключенная к ней круглая щелевая полость рассматривается как резонатор [4, 5]. Для расчета параметров конструкции симметрирующего трансформатора [5], согласующего приемоизлучающую структуру с распределительной линией, его входное волновое сопротивление полагается равным 50 Ом.

Выбранное модельное представление антенны позволяет находить амплитудно-фазовое распределение поверхностных токов на раскрыве и в области преобразования импеданса путем суммирования базисных функций краевых задач электродинамики с весовыми коэффициентами, равными значениям комплексных амплитуд токов на множестве точек дискретизации нерегулярной двухсторонней щелевой линии. В области резонатора токи устанавливаются по значениям, номинально заданным для распределительной линии, согласованной через симметрирующий трансформатор.

Электродинамическая модель сверхширокополосной конической щелевой антенны построена в системе компьютерного моделирования приемоизлучающих структур диапазона сверхвысоких частот CST Microwave Studio из состава вычислительного комплекса CST Studio Suite [7].

При моделировании конфигурация антенны формировалась путем компоновки деталей, содержащихся в библиотеке программы [7]. Для расчета поверхностных токов, по распределению [3] которых находятся поляризационные компоненты и показатели пространственно-частотной избирательности передачи-приема волновых процессов, задействовался вычислитель Frequency Domain Solver.

Решение краевой задачи электродинамики выполнялось методом конечных элементов [3, 7]. В соответствии с ним сетчатое разбиение [7] приемоизлучающей поверхности, установленное для метода моментов [4], заменялось дискретизацией объема пространства, где размещалась антенна [7]. В интересах достижения высокой точности восстановления комплексных амплитуд токов и компонентов электромагнитного поля, контролируемой по предварительно установленной остаточной невязке граничных условий [3, 5], в местах расположения электрически малых фрагментов и соединений отдельных областей антенны осуществлялось адаптивное изменение шага сетки разбиения с коррекцией позиций для размещения ее узлов [7].

2. Анализ электродинамических характеристик конической щелевой антенны

На основе построенной электродинамической модели была проведена оценка характеристик конической щелевой антенны с линейными размерами 240×200 мм, предназначенной для работы в ультраширокополосном диапазоне частот $f = 1,25\ldots20$ ГГц (рис. 1). Приемоизлучающая структура выполнена на печатной плате из стеклотекстолита Rogers RO4003 толщиной 1,5 мм, металлизированного медной фольгой.

Из результатов электродинамического моделирования следует, что объемная диаграмма направленности на частоте $f = 5$ ГГц (рис. 2) формирует выраженный главный лепесток, для которого ширина по уровню -3 дБ составляет 61,2° в плоскости $\varphi = 90°$ и 128,6° в плоскости $\varphi = 0°$, при максимальном значении коэффициента усиления КУ=5,7 dBi и уровнях боковых лепестков в этих сечениях -12 дБ и -12,4 дБ соответственно.

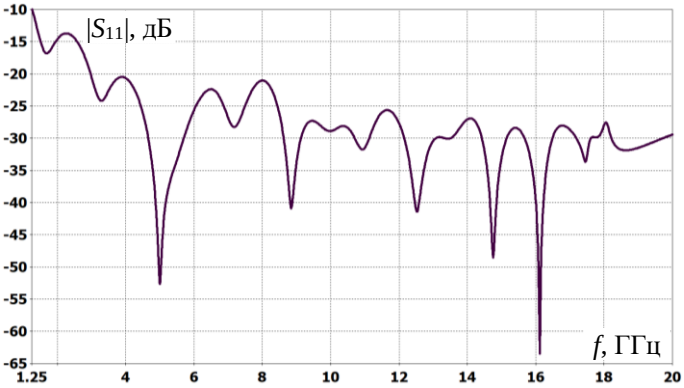


Рис. 1. Коэффициент отражения антенны

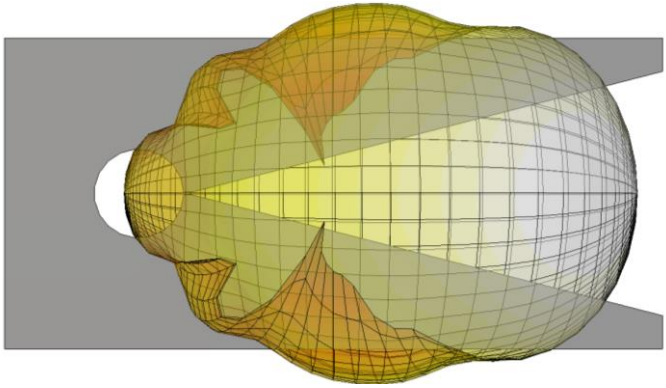


Рис. 2. Объемная диаграмма направленности антенны на частоте $f = 5$ ГГц

Частотная зависимость коэффициента усиления антенны в направлении максимального излучения (рис. 3) характеризуется средним наклоном приблизительно 0,6 дБ/ГГц, что отражает закономерное увеличение эффективной апертуры при увеличении частоты. В нижней части диапазона $f_1 = 1,7 - 7$ ГГц неравномерность КУ составляет 7,8 дБ, а соответствующая амплитуда отклонений КУ (определяемая как половина разности между максимальным и минимальным значениями усиления в поддиапазоне) достигает ±3,9 дБ,

что что связано с интенсивным увеличением усиления на низких частотах. В центральной части диапазона $f_2 = 7 - 12$ ГГц неравномерность равна 1,7 дБ, а амплитуда отклонений – около ±0,85 дБ, что указывает на формирование более стабильной частотной характеристики усиления. В верхней части диапазона $f_3 = 12 - 20$ ГГц неравномерность усиления уменьшается до 1,4 дБ, а амплитуда отклонений не превышает ±0,7 дБ, что свидетельствует о высокой равномерности усилительных свойств антенны и устойчивости её направленных характеристик на высоких частотах.

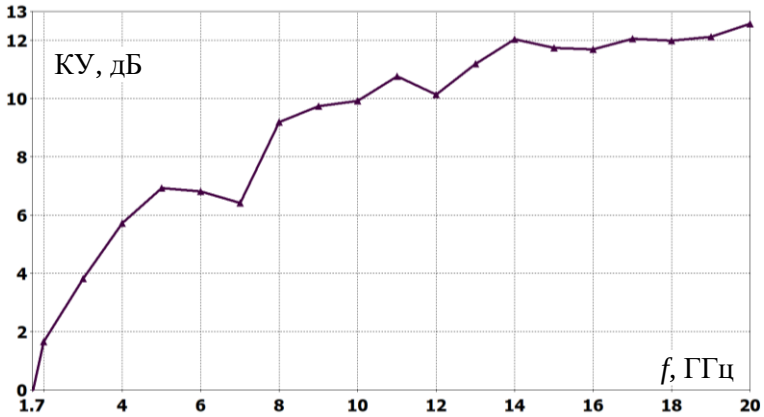


Рис. 3. Частотная зависимость коэффициента усиления

Заключение

С использованием пакета электродинамического моделирования приемоизлучающих структур CST Microwave Studio, входящего в состав вычислительного комплекса CST Studio Suite [7], проведен анализ конической щелевой антенны, которая вследствие высоких показателей пространственно-частотной избирательности [2, 3, 5] передачи (приема) сигналов и малых массогабаритных характеристик находит широкое применением в мобильных радиоэлектронных комплексах. Исследованы распределения коэффициента усиления и коэффициента отражения в диапазоне рабочих частот. Получены оценки уровней боковых лепестков и ширины главного луча диаграммы направленности антенны в плоскостях азимута и угла места.

Представленные результаты составляют методическую основу для решения задачи синтеза антенны, состоящей в рациональном выборе параметров конструкции для реализации требуемой пространственно-

частотной избирательности передачи (приема) сигналов, и обоснования предложений по разработке симметрирующих и согласующих устройств диапазона сверхвысоких частот [3, 5].

Литература

1. Автоматизированные системы управления, радиосвязи и радиоэлектронной борьбы / М. Л. Артемов [и др.] ; под общ. ред. М. Л. Артемова. – М. : Радиотехника. – 2021. – 556 с.
2. Антенны, СВЧ-устройства и их технологии / Д. Д. Габриэльян [и др.] ; под общ. ред. А. А. Косогора. – М. : ФИЗМАТЛИТ. – 2023. – 400 с.
3. Устройства СВЧ и антенны [Текст]. В 2 ч. Ч. 2. Теория и техника антенн / В. А. Неганов, Д. С. Клюев, Д. П. Табаков ; под ред. В. А. Неганова. – Москва : ЛЕНАНД. – 728 с.
4. Фролов, А. А. Моделирование характеристик антенн СВЧ-диапазона / А. А. Фролов, С. В. Гирич, В. П. Заярный // Известия Волгоградского государственного университета. – 2008. – № 4. – С. 112–114.
5. Неганов, В. А. Теория и применение устройств СВЧ / В. А. Неганов, Г. П. Яровой ; под ред. В. А. Неганова. – М. : Радио и связь. – 2006. – 720 с.
6. Рязанов, И. Г. Анализ и синтез широкополосной планарной щелевой антенны с экспоненциальным изменением ширины щели для широкополосного доступа / И. Г. Рязанов, А. А. Бякин, О. А. Белоусов // Вопросы современной науки и практики. – 2013. – № 2 (46). – С. 297–306.
7. Курушин, А. А. Проектирование СВЧ-устройств в среде CST Microwave Studio / А. А. Курушин, А. Н. Пластиков. – М. : МЭИ. – 2011. – 155 с.

Оценка точности пеленгационной локализации источника радиоизлучения при отсутствии прямой видимости

Д. Н. Борисов, email: borisov@sc.vsu.ru
Д. А. Семечев, email: kprvo2014@yandex.ru

Воронежский государственный университет

Аннотация. *В работе исследуется точность определения местоположения источника радиоизлучения на плоскости по серии азимутальных измерений, выполняемых двумя мобильными пеленгаторами на борту БПЛА, движущихся по заданным траекториям. Для оценки координат источника применяется метод взвешенных наименьших квадратов, использующий всю совокупность полученных пеленгов, а качество решения характеризуется как фактической ошибкой локализации в имитационном эксперименте, так и теоретической неопределённостью, определяемой по ковариационной матрице оценки. Для приближения моделирования к реальным условиям учитывается влияние препятствий и режима отсутствия прямой видимости (NLOS), который вызывает систематическое смещение пеленгов, увеличение дисперсии угловой ошибки и пропуски измерений, что в итоге усиливает зависимость результата от геометрии наблюдения и корректности учёта статистики ошибок.*

Ключевые слова: *пеленгация, метод взвешенных наименьших квадратов, NLOS, моделирование.*

Введение

Пеленгационный метод определения местоположения источника излучения находит применение в задачах радиоразведки, радиомониторинга и поисково-спасательных операций благодаря использованию исключительно угловой информации и отсутствию требований к точной временной синхронизации, характерной для дальномерных подходов, при сравнительно простой аппаратной реализации. Однако, измеряя лишь направление на цель, пеленгатор требует для восстановления координат анализа геометрии взаимного расположения наблюдателей и статистической обработки погрешностей, где итоговая точность локализации критически зависит как от выбранной конфигурации траекторий БПЛА, так и от характеристик ошибок угловых измерений. В рамках данной работы выполняется моделирование и анализ характеристик алгоритма пеленгационной локализации: сначала

© Борисов Д. Н., Семечев Д. А., 2026

для базового сценария в условиях прямой видимости, а затем для более сложной ситуации с учётом препятствий и режимов распространения сигнала вне прямой видимости.

На практике основными причинами деградации точности пеленгационной локализации являются участки с отсутствием прямой видимости (Non-Line-of-Sight, NLOS) и явление многолучевости. Отражения и дифракция радиоволн изменяют направление прихода сигнала, что приводит не только к увеличению дисперсии угловой ошибки, но и к появлению систематического смещения, которое невозможно устранить простым накоплением числа измерений. В условиях NLOS также возможны интервалы полного отсутствия корректных наблюдений, когда пеленгатор не способен сформировать оценку направления. Подобные пропуски данных сокращают объём полезной информации, в результате чего итоговое решение становится более чувствительным к геометрии оставшихся измерений и размеру доступной выборки. Поэтому при сравнительном анализе сценариев прямой видимости (Line-of-Sight, LOS) и условий NLOS целесообразно рассматривать не только фактическую ошибку локализации в моделировании, но и теоретическую неопределённость, получаемую из ковариационной матрицы оценки. Эта неопределённость служит важным индикатором качества решения в условиях, когда истинное положение источника неизвестно.

В данной работе рассматривается сценарий, в котором два БПЛА, двигаясь по известным траекториям, измеряют азимут на неподвижный источник радиоизлучения. Оценка координат источника формируется по накопленному набору угловых измерений, а для повышения устойчивости к неоднородной точности пеленгов применяется взвешенный метод наименьших квадратов (МНК), широко распространённый в задачах пеленгационной локации [1–3]. Данный подход относится к классу пассивных методов локации, где наблюдатели, не излучая собственных сигналов, оценивают направление прихода уже существующего излучения цели, что обеспечивает скрытность и энергоэффективность системы, но одновременно повышает требования как к качеству угловых измерений, так и к оптимальности взаимного расположения наблюдателей [1, 4].

1. Математическая модель и алгоритм

Оценка координат источника радиоизлучения по пеленгам относится к классическим задачам пассивной радиолокации и радиотехнической разведки. Практическая значимость таких задач связана с необходимостью поиска источников помех и несанкционированных передатчиков, локализации аварийных

радиомаяков, мониторинга радиоэлектронной обстановки, а также сопровождения и идентификации целей при ограниченном объеме априорной информации. Ключевое достоинство пассивного подхода заключается в том, что измерительная система не излучает собственных сигналов и использует уже существующее излучение цели, что обеспечивает скрытность и энергетическую эффективность работы, но одновременно предъявляет повышенные требования к точности угловых измерений и к оптимальной геометрии расположения наблюдателей.

В отличие от дальномерных методов, основанных на измерении времени прихода сигнала (ToA), разности времен прихода (TDoA) или уровня его мощности (RSS), пеленгационный подход опирается исключительно на угловую информацию. Это упрощает аппаратную реализацию системы и позволяет эффективно применять её на значительных расстояниях, хотя итоговая точность локализации при этом существенно зависит от взаимного расположения приёмников и статистических характеристик ошибок измерений. Как демонстрируют исследования [5, 6], даже незначительная погрешность в определении угла способна привести к существенному увеличению ошибки оценки координат при неоптимальной геометрии наблюдения, особенно в условиях, когда к случайной составляющей ошибки добавляется систематическое смещение.

Рассмотрим сценарий, в котором два БПЛА, следуя заданным траекториям, осуществляют дискретные измерения пеленга на неподвижный источник радиоизлучения. В рамках данного сценария задача сводится к оценке координат фиксированной точки по серии направлений, полученных из различных позиций наблюдателей. Ключевым фактором, определяющим точность и информативность решения, является выбор самих траекторий, поскольку геометрия пересечения пеленгов напрямую влияет на чувствительность к неизбежным ошибкам измерений. В реальных условиях задача осложняется наличием препятствий и нестационарным характером распространения сигнала, включая режимы вне прямой видимости, которые могут приводить как к значительному росту дисперсии ошибок, так и к возникновению зон временного отсутствия сигналов.

Рассмотрим систему координат XOY на плоскости, в которой положение источника полагается неизменным на протяжении всего времени движения БПЛА:

$$r_s = \begin{bmatrix} x_s \\ y_s \end{bmatrix}. \quad (1)$$

При заданных уравнениях движения (1) знание траекторий позволяет использовать координаты БПЛА на каждом дискретном временном шаге k:

$$r_i(k) = \begin{bmatrix} x_i(k) \\ y_i(k) \end{bmatrix}, \; i = 1, 2 . \tag{2}$$

На каждом временном шаге k каждый БПЛА осуществляет измерение пеленга на неподвижный источник, регистрируя текущий угол направления:

$$\theta_i^{\text{meas}}(k) = \operatorname{atan} 2(y_s - y_i(k), x_s - x_i(k)) + v_i(k), \tag{3}$$

где $v_i(k)$ – случайная погрешность измерения.

В реальных условиях дисперсия ошибок пеленгования непостоянна и изменяется в зависимости от расстояния до цели, текущих условий приема сигнала и состояния окружающей среды. В ситуации, когда отсутствует прямая видимость источника, к случайной ошибке дополнительно добавляется систематическое смещение измерений:

$$\theta_i^{\text{meas}}(k) = \theta_i(k) + b_i(k) + v_i(k), \tag{4}$$

где $b_i(k)$ –смещение, возникающее из-за условий различной природы.

Алгоритм основан на методе взвешенных наименьших квадратов (WLS), который определяет положение цели путем минимизации суммы квадратов поперечных отклонений от всех измеренных пеленгов [7–9]. Этот общепринятый подход естественным образом учитывает разную точность измерений через весовые коэффициенты и обеспечивает геометрически интерпретируемое и вычислительно эффективное решение.

Пусть по измеренному азимуту строится единичный направляющий вектор:

$$u_m = \begin{bmatrix} \cos\theta_m \\ \sin\theta_m \end{bmatrix}, \tag{5}$$

и соответствующий ему ортогональный вектор:

$$u_m^{\perp} = \begin{bmatrix} -\sin\theta_m \\ \cos\theta_m \end{bmatrix}. \tag{6}$$

Индекс m применяется для сквозной нумерации всех пеленгов, полученных от обоих БПЛА в различные моменты времени.

Поперечное отклонение (невязка) оцениваемой точки r относительно пеленга, заданного из позиции наблюдателя r_m, определяется выражением:

$$d_{\perp,m} = (u_m^{\perp})^T (r - r_m). \quad (7)$$

С учетом данного выражения задача оценки координат источника формулируется в терминах взвешенного метода наименьших квадратов:

$$\hat{r}_s = argmin \sum_{m=1}^{M} \omega_m d_{\perp,m}^2 \perp, \quad (8)$$

где ω_m – весовые коэффициенты, соответствующие каждому измерению.

В случае, когда все веса равны, задача сводится к обычному МНК. Если же веса различаются, что соответствует неодинаковой точности измерений, метод становится взвешенным (WLS), что является стандартным подходом в задачах локализации и параметрического оценивания [2–4, 7].

$$H = \sum_{m=1}^{M} w_m (I - u_m u_m^T), \quad b = \sum_{m=1}^{M} w_m (I - u_m u_m^T) r_m . \quad (9)$$

Численное решение данной оптимизационной задачи сводится к нахождению точки, удовлетворяющей:

$$\hat{r}_s = H^{-1} b . \quad (10)$$

Ошибка измерения угла, заданная в радианах, непосредственно влияет на точность определения координат через соответствующее ей поперечное отклонение на местности. Для малых угловых ошибок удобно использовать линейное приближение:

$$\sigma_{\perp,m} \approx \rho_m \sigma_{\theta,m} , \quad (11)$$

где ρ_m – расстояние от источника до позиции m-го измерения, а $\sigma_{\theta,m}$ – среднеквадратическое отклонение (СКО) угловой ошибки измерения. Данное соотношение показывает, что при одинаковой угловой точности σ_θ измерения, выполненные с большей дистанции, приводят к пропорционально большему разбросу поперечного отклонения, и их вклад в итоговое решение следует уменьшать.

В рамках рассматриваемой модели весовой коэффициент для каждого измерения логично выбрать обратно пропорциональным квадрату этой поперечной неопределённости.

$$w_m = 1 / \sigma_{\perp,m}^2 . \quad (12)$$

Качество итоговой оценки характеризуется двумя ключевыми показателями. В рамках моделирования первый показатель – это фактическая ошибка локализации, рассчитываемая как отклонение полученной оценки от известного истинного положения источника.

$$\varepsilon(k) = \left\| \hat{r}_s(k) - r_s \right\|. \quad (13)$$

Данный показатель является доступным исключительно в условиях моделирования, поскольку в реальных задачах истинные координаты цели остаются неизвестными. Тем не менее, его использование позволяет эффективно анализировать сходимость алгоритма и проводить сравнительную оценку его работы в различных сценариях, включая условия без помех и при их наличии.

Вторым ключевым показателем качества является прогнозируемая неопределённость оценки, определяемая по её ковариационной матрице. В рамках приближения метода взвешенных наименьших квадратов данная ковариация выражается через нормальные уравнения.

$$C \approx H^{-1}, \quad \sigma_{\text{pred}} = \sqrt{\text{trace}(C)} \quad (14)$$

где trace() обозначает след матрицы – сумму элементов ее главной диагонали.

Данный критерий важен тем, что, в отличие от фактической ошибки $\varepsilon(k)$, прогнозируемая неопределенность σ_{pred} может быть рассчитана непосредственно в реальной системе без знания истинного положения источника и, следовательно, служить оперативным индикатором достоверности полученной оценки [2–4,7]. Уменьшение величины σ_{pred} свидетельствует об улучшении геометрии наблюдения и накоплении информативных измерений.

2. Моделирование сценария с учетом прямой видимости

В рамках вычислительного эксперимента рассматривается работа алгоритма WLS при последовательном накоплении пеленгов, поступающих от двух БПЛА, перемещающихся по заданным траекториям в плоской области в условиях сохранения прямой видимости между источником и наблюдателями на всем интервале моделирования. На каждом шаге по времени для каждого БПЛА формируется очередное угловое измерение, после чего производится пересчет оценки координат источника $\hat{r}$ на основе всей аккумулированной совокупности пеленгов (рис. 1). Поскольку эксперимент проводится в среде имитационного моделирования, истинное положение источника r_s считается априори известным, что позволяет непосредственно вычислять ошибку локализации, анализировать характер сходимости алгоритма и сравнивать его поведение на различных участках траекторий движения (рис. 2).

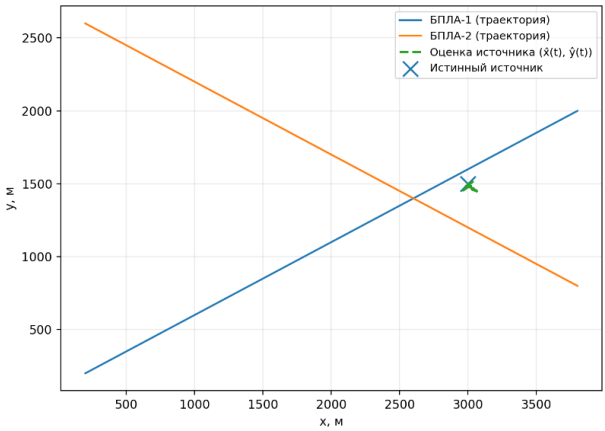


Рис. 1. Траектории БПЛА и оценка координат ИРИ для сценария без помех

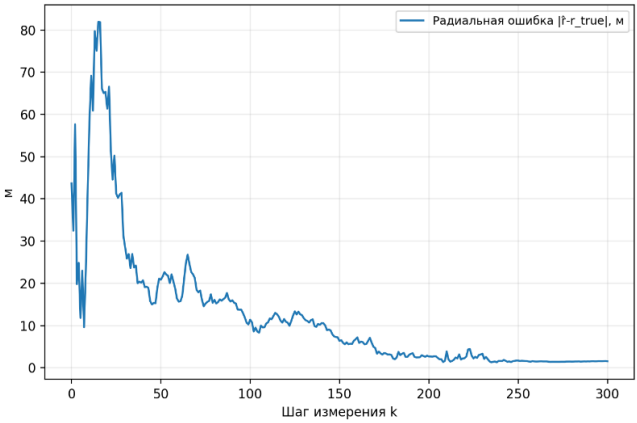


Рис. 2. Ошибка оценки координат ИРИ для сценария без помех

На графике движения БПЛА можно наблюдать, что на начальном этапе количество измерений является недостаточным для устойчивой пеленгационной локализации, что приводит к заметным смещениям оценки при поступлении новых пеленгов, особенно в ситуациях, когда геометрия наблюдений оказывается близкой к вырожденной и направления пеленгов пересекаются под малыми углами. По мере накопления данных в процессе движения БПЛА геометрия наблюдения

становится более информативной, а оценка, полученная методом WLS, стабилизируется в окрестности истинного положения источника, что проявляется в уменьшении амплитуды колебаний оценки и снижении ошибки $\left|\hat{r} - r_s\right|$ локализации на последующих шагах. Согласованная динамика прогнозируемой неопределенности, рассчитанной по ковариационной матрице оценки, демонстрирует устойчивую тенденцию к снижению, что соответствует росту информативности накопленного набора пеленгов и подтверждает, что выбранная схема взвешивания обеспечивает корректную обработку данных при неодинаковом вкладе отдельных измерений, определяемом их «геометрической ценностью».

Таким, образом, проведенное моделирование для сценария с прямой видимостью подтверждает устойчивую сходимость алгоритма WLS-оценивания при последовательном накоплении пеленгов и наглядно демонстрирует, что качество получаемого решения определяется не только уровнем углового шума измерений, но и геометрией расположения БПЛА в момент каждого измерения.

3. Моделирование сценария с учетом препятствий и отсутствия прямой видимости источника радиоизлучения

Для приближения модели к реальным условиям рассматривается среда с препятствиями, которые могут нарушать прямую видимость между БПЛА и источником. Если прямая видимость отсутствует, измерение классифицируется как режим NLOS, в рамках данной работы приводящий к исключению такого пеленга из обработки. Это уменьшает общее число измерений и ухудшает геометрическую обусловленность задачи на соответствующих участках траектории [10]. В модели влияние NLOS может быть описано систематическим смещением

$$\theta_i(k) = atan2(k) + b_i(k) + \nu_i(k), \qquad (15)$$

где $b_i(k)$ – систематическое смещение углового измерения, возникающее в режиме отсутствия прямой видимости.

Вычислительная схема WLS при этом остается неизменной, однако критическое значение приобретает адаптивная оценка текущего качества пеленгации. Для этого вычисляются угловые невязки между измеренным пеленгом и ожидаемым направлением на текущую оценку положения источника. Динамическая оценка угловой погрешности формируется по скользящему окну последних невязок с использованием робастной статистики – медианного абсолютного отклонения (MAD), которое менее чувствительно к выбросам по сравнению со стандартным отклонением [11]:

$$\sigma_\theta = 1.4826 \cdot MAD. \tag{16}$$

Полученная оценка σ_θ связывается с поперечной ошибкой через текущую дальность и используется для автоматического расчета весов в WLS:

$$\sigma_\perp \approx \rho\sigma_\theta, \tag{17}$$

$$w = \frac{1}{\sigma_\perp^{\,2}}. \tag{18}$$

Следовательно, в условиях ухудшения качества измерений весовые коэффициенты автоматически снижаются, что ослабляет влияние ненадёжных пеленгов на итоговую оценку координат источника и повышает робастность алгоритма.

На рис. 3 показано, как при обходе препятствий один из БПЛА проходит через зоны с ухудшенной геометрией наблюдения. Это приводит к временному снижению информативности набора пеленгов, что проявляется в увеличении дрейфа оценки координат. При возвращении в зоны с благоприятной конфигурацией и достаточным числом измерений оценка стабилизируется вблизи истинного положения источника, демонстрируя адаптивность предложенного подхода к изменяющимся условиям.

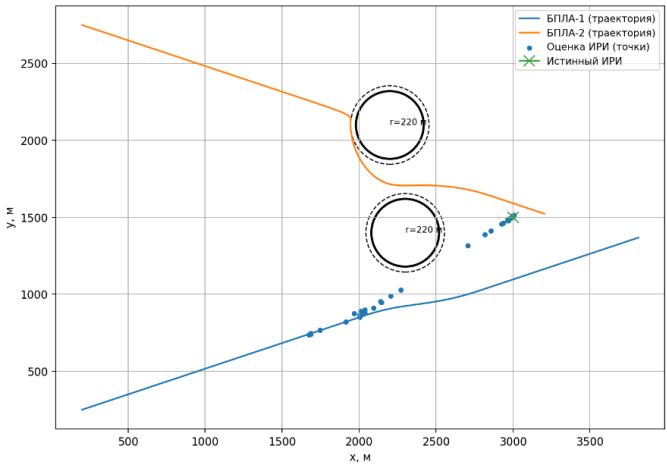


Рис. 3. Траектории БПЛА, препятствия и оценка координат ИРИ (NLOS-сценарий)

Заключение

В работе рассматривается задача пеленгационной локализации источника радиоизлучения двумя БПЛА, движущимися по известным траекториям, где для получения оценки координат используется метод взвешенных наименьших квадратов (WLS), применяемый к совокупности угловых измерений и демонстрирующий свою устойчивость. В рамках базового сценария с прямой видимостью (LOS) показана стабилизация оценки, сопровождающаяся снижением прогнозируемой неопределенности по мере накопления пеленгов и улучшения геометрии наблюдения.

Для приближения к реальным условиям был рассмотрен сценарий с наличием препятствий и режимом NLOS, где включение адаптивного механизма оценки угловой погрешности на основе робастной статистики (MAD), рассчитываемой по остаткам измерений, позволило автоматически снижать влияние зашумленных данных, сохраняя работоспособность алгоритма. Полученные результаты свидетельствуют, что на участках с NLOS наблюдается усиленный дрейф оценки координат, устраняемый при возвращении в зоны прямой видимости.

Перспективы дальнейших исследований связаны с развитием модели искажений в NLOS-режиме, а также с оптимизацией траекторий БПЛА, направленной на максимизацию информативности геометрии пеленгации при учёте операционных ограничений.

Литература

1. Statistical theory of D.F. fixing / R. G. Stansfield // Journal of the IEE, Part IIIA: Radiocommunication. – 1947. – Vol. 94(15). – P.762 –770.
2. Torrieri, D. J. Statistical Theory of Passive Location Systems / D. J. Torrieri // IEEE Transactions on Aerospace and Electronic Systems. – NY : Springer New York, 1990. – P. 151–166.
3. Doğançay, K. Bearings-only target localization using total least squares / K. Doğançay // Signal Processing. – 2005. – Vol. 85(9). – P. 1695–1710.
4. Foy, W. H. Position-Location Solutions by Taylor-Series Estimation / W. H. Foy // IEEE Transactions on Aerospace and Electronic Systems. – 1976. – Vol. AES-12, № 2. – P. 187–194
5. Wymeersch, H. Cooperative Localization in Wireless Networks / H. Wymeersch, J. Lien, M. Z. Win // Proceedings of the IEEE. – 2009. – Vol. 97, № 2. – P. 427–450.
6. Huber, P. J. Robust Statistics / P. J. Huber, E. M. Ronchetti. – John Wiley & Sons, Inc., 1981.

7. Locating the nodes: cooperative localization in wireless sensor networks / N. Patwari [et al.] // IEEE Signal Processing Magazine. – 2005. – Vol. 22, № 4. – P. 54–69.

8. Gustafsson, F. Mobile positioning using wireless networks: possibilities and fundamental limitations / F. Gustafsson, F. Gunnarsson // IEEE Signal Processing Magazine. – 2005. – Vol. 22(4). – P. 41–53.

9. NLOS identification and mitigation for localization based on UWB experimental data / S. Maranò S. [et al.] // IEEE Journal on Selected Areas in Communications. – 2010. – Vol. 28, № 7. – P. 1026–1035.

10. Guvenc, I. NLOS Identification and Mitigation for UWB Localization Systems / I. Guvenc, C.-C. Chong, F. Watanabe // Proc. IEEE Wireless Communications and Networking Conference. – Hong Kong : 2007. – P. 1571–1576.

11. MathWorks (Hampel) [Электронный ресурс]. – Режим доступа: https://www.mathworks.com/help/signal/ref/hampel.html

Численное моделирование переходных процессов в импульсных диодах методом конечных разностей

А. С. Величкина, email: nastyavelichkina65@gmail.com
А. А. Кононов, email: kononov@phys.vsu.ru
Г. К. Усков, email: uskov@phys.vsu.ru

Воронежский государственный университет

Аннотация. *В работе рассматривается применение метода конечных разностей для получения приближенного решения дифференциального уравнения в частных производных, описывающего переходные процессы в полупроводниковом диоде. На примере изученной ранее модели диода с накоплением заряда продемонстрированы этапы практического применения данного способа, выполнено сравнение с точным решением, приводятся типовые частные случаи результатов.*

Ключевые слова: *диод с накоплением заряда, распределение концентрации носителей, диффузионно-дрейфовое уравнение, численное решение, конечные разности, вычисление плотности тока.*

Введение

Импульсные диоды являются немаловажным компонентом элементной базы современных радиотехнических устройств [1–5]. Дополнительный интерес к подобным полупроводниковым приборам обусловлен, в частности, тем, что они могут быть использованы для формирования сверхширокополосных сигналов (включая сверхкороткие импульсы); в числе прочего, здесь находят применение диоды с накоплением заряда (ДНЗ), переходные процессы в которых характеризуются быстрым восстановлением высокого обратного сопротивления [2, 4].

Высокоточные модели подобных диодов включают в себя учет различных физических явлений, приводящих к нелинейным эффектам [6–10]. Это, в свою очередь, нередко основывается на введении в уравнения тех или иных эмпирических зависимостей, что как правило, делает невозможным получение аналитического решения; численные же методы зачастую чрезмерно адаптированы к конкретным частным случаям или слишком требовательны к производительности оборудования [6, 7, 9, 11–12]. Отсутствие единой и общепринятой технологии приближенного расчета интересующих характеристик говорит об актуальности проблемы разработки подобных алгоритмов.

© Величкина А. С., Кононов А. А., Усков Г. К., 2026

В настоящей работе подробно обсуждается один относительно несложный метод численного решения дифференциального уравнения, описывающего процессы в полупроводниковой среде. Явный вид всех рабочих соотношений при этом будет составлен на примере достаточно простой модели [2, 13–17], для которой предшествующие исследования позволили получить точное решение в аналитической форме (что будет ниже использовано для проверки приближенного метода).

1. Исходная математическая модель

Постановка задачи для диода с накоплением заряда такова. Считается, что полупространство $x > 0$ соответствует базе диода (для определенности — *n*-типа), в которой будет исследоваться поведение концентрации неосновных носителей заряда (то есть, в данном случае, дырок) $p(x,t)$. Эта функция на практике оказывается убывающей до нуля при достаточно больших значениях x, что соответствует конечной глубине проникновения дырок в *n*-область с физической точки зрения, а с математической — обосновывает правомерность использования модели с «полубесконечной базой». Сама полупроводниковая среда характеризуется рядом параметров, из которых здесь наиболее важны коэффициент диффузии D_p и среднее время жизни дырок τ_p, а также напряженность электрического поля E (предполагаемого постоянным и направленным вдоль оси x); для нее имеет место диффузионно-дрейфовое уравнение

$$\frac{\partial p(x,t)}{\partial t} = D_p \frac{\partial^2 p(x,t)}{\partial x^2} - \mu_p E \frac{\partial p(x,t)}{\partial x} - \frac{p(x,t)}{\tau_p}, \qquad (1)$$

где μ_p — коэффициент подвижности дырок, в простых случаях выражаемый посредством соотношения Эйнштейна — Смолуховского

$$\mu_p = \frac{D_p q}{kT} \qquad (2)$$

через элементарный заряд $q = 1{,}602\,176\,634 \cdot 10^{-19}\ Кл$, постоянную Больцмана $k = 1{,}380\,649 \cdot 10^{-23}\ Дж/К$ и абсолютную температуру T.

Все переходные процессы имеют место в моменты времени $t > t_0$ и изначально задаются законом изменения граничной плотности тока $j(t)$, описывающим включение прямого тока и последующую смену полярности. Распределение дырок связано с этим током условиями

$$p(x,t_0)=0,\quad -qD_p\frac{\partial p(x,t)}{\partial x}+q\mu_p Ep(x,t)\Big|_{x=0}=j(t), \tag{3}$$

смысл первого из которых в том, что при $t\le t_0$ дырок в базе нет (таким образом, мы пренебрегаем их равновесной концентрацией), второе же обеспечивает совпадение плотности тока на границе с $j(t)$.

Это продолжается до момента t_1, когда (уже при $j(t)<0$) вблизи *p-n* перехода концентрация неосновных носителей снижается до нуля:

$$p\left(0,t_1\right)=0; \tag{4}$$

как следствие, диод закрывается, и наступает 2-я фаза. Для нее в силе остается уравнение (1), а вместо (3) действуют условия

$$p\left(x,t_1\right)=p_1\left(x\right),\quad p\left(0,t\right)=0, \tag{5}$$

где $p_1(x)$ — распределение, конечное для 1-й фазы и начальное для 2-й. Модифицируя граничное соотношение из (3) с учетом (5), вычисляем плотность тока рассасывания:

$$j(t)=-qD_p\frac{\partial p(x,t)}{\partial x}\Big|_{x=0}. \tag{6}$$

Аналитическое решение поставленной задачи имеет довольно сложную форму. Явный вид точных расчётных соотношений был впервые получен в работах [13, 14], а затем исследован более подробно в [15, 16, 17]. В 1-й фазе

$$p\left(x,t\right)=u\left(x_n,t_n\right)\cdot e^{E_n x_n-\left(1+E_n^2\right)t_n},\quad t_n\in\left[t_{n0};t_{n1}\right], \tag{7}$$

где задействованы нормированные переменные

$$x_n=\frac{x}{L_p},\quad t_n=\frac{t}{\tau_p},\quad E_n=\frac{qEL_p}{2kT},\quad L_p=\sqrt{D_p\tau_p} \tag{8}$$

и вспомогательная функция

$$u\left(x_n,t_n\right)=\int_{t_{n0}}^{t_n}\frac{\nu(\tau)}{2\sqrt{\pi\left(t_n-\tau\right)}}e^{\frac{-x_n^2}{4(t_n-\tau)}}d\tau, \tag{9}$$

имеющая форму т. н. теплового потенциала [18–20] с плотностью

$$\nu(\tau)=-2f\left(\tau\right)+\frac{2E_n}{\sqrt{\pi}}\cdot\int_{t_{n0}}^{\tau}f\left(\tilde{\tau}\right)R\left(\tau,\tilde{\tau}\right)d\tilde{\tau}, \tag{10}$$

выражающейся с помощью резольвентного ядра [21, 22] вида

$$R(\tau,\tilde{\tau}) = \frac{1}{\sqrt{\tau-\tilde{\tau}}} + \sqrt{\pi}E_n e^{E_n^2(\tau-\tilde{\tau})}\left(\operatorname{erf}\left(E_n\sqrt{\tau-\tilde{\tau}}\right)-1\right), \tag{11}$$

и замены

$$f(t_n) \equiv -\frac{j(t)L_p}{qD_p}\cdot e^{\left(1+E_n^2\right)t_n} = -\frac{j\left(t_n\tau_p\right)}{q}\sqrt{\frac{\tau_p}{D_p}}\cdot e^{\left(1+E_n^2\right)t_n}. \tag{12}$$

Расчёт ведётся по формулам (7)–(12) до момента t_1, определяемого из условия (4), затем фиксируется распределение $p_1(x)$, после чего ток рассасывания во 2-й фазе может быть рассчитан, согласно [14–17], как

$$j(t) = -q\sqrt{D_p}\tau_p \cdot \frac{e^{-\left(1+E_n^2\right)\cdot\frac{t-t_1}{\tau_p}}}{2\sqrt{\pi}\left(t-t_1\right)^{3/2}}\int_0^\infty p_1\left(\xi L_p\right)\cdot\xi\cdot e^{\frac{-\xi^2\tau_p}{4(t-t_1)}-E_n\xi}\,d\xi. \tag{13}$$

Как можно видеть, точное решение представляет собой совокупность достаточно громоздких аналитических выражений, для своей практической реализации требующих численного интегрирования и нахождения корня нелинейной функции.

В более сложных моделях, предназначенных для учета нелинейных эффектов, подвижность дырок может зависеть от электрического поля в базе (которое, в свою очередь, уже не обязательно будет считаться постоянным); аналогичное верно и для коэффициента диффузии. Конкретные зависимости формулируются на основе тех или иных физических предположений, эмпирических закономерностей и так далее [6–10]; все подобные модификации усложняют, в первую очередь, уравнение (1), которое перестаёт быть линейным и уже практически никогда не поддаётся аналитическому решению.

В связи с этим, не возникает сомнений в актуальности вопроса о практических методах численного решения диффузионно-дрейфового уравнения, к рассмотрению одного из которых мы и приступаем далее.

2. Дискретизация модели. Разностные отношения

Для получения приближенного решения вышеописанная модель должна быть подвергнута процедуре пространственно-временной дискретизации. Вначале выберем некоторый фиксированный шаг по времени Δt и заменим частную производную в левой части уравнения (1) на конечное разностное отношение [23]:

$$\frac{\partial p(x,t)}{\partial t} \mapsto \frac{p(x,t+\Delta t)-p(x,t)}{\Delta t} = F\left[p(x,t)\right], \tag{14}$$

где $F[p(x,t)]$ — правая часть (1), взятая в момент t (в более сложных моделях в ней могут быть нелинейные члены и т. п.). Тогда на основе (14) получаем следующий алгоритм последовательного расчета $p(x,t)$:

$$p(x,t+\Delta t) = p(x,t) + \Delta t \cdot F\left[p(x,t)\right], \tag{15}$$

по аналогии с тем, как это делается в рамках одношагового метода Эйлера для обыкновенных дифференциальных уравнений [23, 24].

Затем дискретизируем пространственную координату с шагом Δx:

$$x_m = (m-1)\Delta x, \quad m = 1; 2; \ldots; M, \tag{16}$$

где количество точек M определяется максимальным учитываемым в моделировании значением координаты x_{max}, которое, в свою очередь, выбирается из условия достаточной малости $p(x_{max},t)$ (см. пример далее). В результате правая часть (1) принимает вид

$$F\left[p(x_m,t)\right] \mapsto F\left(p_m\right) = D_p \Delta_x^2 p_m - \mu_p E \Delta_x p_m - \frac{p_m}{\tau_p}, \tag{17}$$

где для краткости обозначено

$$p_m = p_m(t) = p(x_m,t) \tag{18}$$

и использованы центральные разностные отношения

$$\Delta_x p_m = \frac{p_{m+1} - p_{m-1}}{2\Delta x}, \quad \Delta_x^2 p_m = \frac{p_{m+1} - 2p_m + p_{m-1}}{(\Delta x)^2}. \tag{19}$$

Теперь, в соответствии с (15), (17) и (19), можно пошагово «обновлять» массив значений $\{p_2; p_3; \ldots; p_M\}$, вычисляя его для все более поздних моментов времени. Для p_1 же определения (19) утрачивают смысл, но в этой точке можно задействовать граничное условие из (3). Именно, дискретизировав его в виде

$$-qD_p \frac{p_2(t) - p_1(t)}{\Delta x} + q\mu_p E p_1(t) = j(t) \tag{20}$$

(с «упрощенным» разностным отношением), выразим отсюда

$$p_1(t) = \frac{D_p q p_2(t) + j(t)\Delta x}{q(D_p + \mu_p E \Delta x)}; \tag{21}$$

эта формула и будет использоваться для получения очередного значения p_1 после того, как «новое» значение p_2 уже найдено.

Далее этот процесс, как и в аналитическом решении, продолжается до момента t_1, когда p_1 обратится в ноль (в пределах заданной

допустимой погрешности). После этого вычисления продолжаются без изменений, но вместо (21) на каждом новом шаге полагается

$$p_1(t) = 0; \tag{22}$$

наконец, ток рассасывания рассчитывается с помощью соотношения

$$j(t) = -qD_p \frac{p_2(t)}{\Delta x}, \tag{23}$$

следующего из (21) — или дискретной версии (6) — с учетом (22).

Заметим, что при реализации на практике выбор параметров дискретизации должен осуществляться с учетом общих условий устойчивости явных разностных схем [23]. В данном случае (когда уравнение (1) — параболическое) справедлива рекомендация

$$\Delta t < \frac{1}{2D_p^2}(\Delta x)^2. \tag{24}$$

Проиллюстрируем изложенный метод на конкретном примере.

3. Пример применения численной методики

Зададим типовые значения параметров [4, 5]:

$$D_p = 18{,}0 \cdot 10^{-4}\ м^2/с; \quad \tau_p = 200\ нс; \quad T = 300\ K; \tag{25}$$

моделирование проведем на интервале времени от $-30\ нс$ до $60\ нс$, разделяя его на 20 000 частей, так что $\Delta t \approx 4{,}5\ пс$. По пространственной же координате ограничимся отрезком от нуля до $x_{max} = 100\ мкм$ (можно отметить, что эта величина с запасом превосходит диффузионную длину свободного пробега дырок, в данном случае равную $L_p \approx 19\ мкм$). Деление этого отрезка на 300 частей приводит к величине шага дискретизации $\Delta x \approx 0{,}334\ мкм$, удовлетворяющей требованию (24).

Для поля в базе диода рассмотрим 3 различных значения: $E_1 = 0$, $E_2 = -0{,}7 \cdot 10^4\ В/м$ (тормозящее) и $E_3 = 0{,}5 \cdot 10^4\ В/м$ (ускоряющее).

Ниже приведены графики результатов моделирования. На рис. 1 показаны временные и пространственные зависимости распределения концентрации неосновных носителей заряда (причем видно, что они, в основном, сосредоточены на глубинах $x \le 30\ мкм$), а на рис. 2 — поведение граничной плотности тока (диапазон ее значений — порядка $10^0\ А/м^2$ — выбран таким для наглядности [15]). В целом, наблюдается весьма хорошее совпадение с данными, полученными в работе [16] путем расчета по точным выражениям.

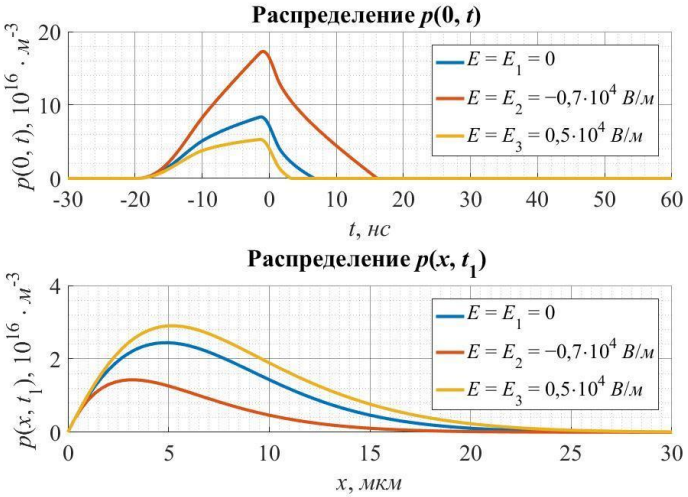


Рис. 1. Рассчитанное распределение неосновных носителей

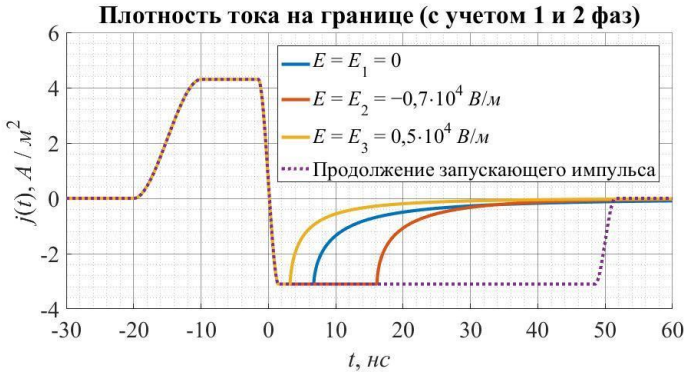


Рис. 2. Рассчитанная плотность тока на границе

Заключение

В работе изложен метод численного решения дифференциального уравнения в частных производных, используемого для моделирования импульсных процессов в базе полупроводникового диода с накоплением заряда. Даны все необходимые расчетные соотношения, допускающие непосредственное обобщение на более сложные физические модели; приводится условие устойчивости практической реализации.

Корректность изложенного приближенного метода подтверждена на конкретном численном примере, для которого расчет переходных процессов в ДНЗ при разных значениях электрического поля в базе привел в каждом случае к результатам, совпадающим с полученными ранее при тех же условиях с помощью строгого аналитического решения в явном виде.

Благодарности

Исследование выполнено за счет гранта Российского научного фонда № 24-19-00891, https://rscf.ru/project/24-19-00891/.

Для выполнения численных расчетов в работе использовалось оборудование учебно-научного дизайн-центра проектирования радиоэлектронных систем СВЧ, терагерцового и оптического диапазонов на отечественной электронной компонентной базе ФГБОУ ВО «ВГУ» в рамках реализации федерального проекта «Подготовка кадров и научного фундамента для электронной промышленности».

Список литературы

1. Рябинин, Ю. А. Стробоскопическое осциллографирование сигналов наносекундной длительности / Ю. А. Рябинин. – М. : «Советское радио», 1968. – 200 с.

2. Еремин, С. А. Полупроводниковые диоды с накоплением заряда и их применение / С. А. Еремин, О. К. Мокеев, Ю. Р. Носов. – 6-е изд., испр. и доп. – М. : «Советское радио», 1966. – 152 с.

3. Зи, С. Физика полупроводниковых приборов. В 2 книгах. Кн. 1. / С. Зи ; пер. с англ. В. А. Гергеля и В. В. Ракитина ; под ред. Р. А. Суриса. – 2-е изд., перераб. и доп. – М. : «Мир», 1984. – 456 с.

4. Kurata, M. Design Considerations of Step Recovery Diodes with the Aid of Numerical Large-Signal Analysis / M. Kurata // IEEE Transactions on Electron Devices. – 1972. – Vol. 19(11). – P. 1207–1215.

5. Демаков, Ю. П. Лекции по физическим основам микроэлектроники : учеб. пособие для вузов / Ю. П. Демаков. – Ижевск: Изд-во УИР ИжГТУ имени М. Т. Калашникова, 2023. – 368 с.

6. Modeling Sources of Nonlinearity in a Simple p-i-n Photodetector / Y. Hu [et al.] // Journal of Lightwave Technology. – 2014. – Vol. 32, no. 20. – P. 3710–3720.

7. Simulation of a partially depleted absorber (PDA) photodetector / Y. Hu [et al.] // Optics Express. – 2015. – Vol. 23, no. 16. – P. 20402–20417.

8. Pfirsch F. Note on Charge Conservation in the Transient Semiconductor Equations / F. Pfirsch, M. Ruff // IEEE Transactions on Electron Devices. – 1993. – Vol. 40, no. 11. – P. 2085–2087.

9. Dentan, M. Numerical Simulation of the Nonlinear Response of a p-i-n Photodiode Un-der High Illumination / M. Dentan, B. de Cremoux // Journal of Lightwave Technology. – 1990. – Vol. 8, no. 8. – P. 1137–1144.

10. Van Roosbroeck, W. Theory of the Flow of Electrons and Holes in Germanium and Other Semiconductors / W. van Roosbroeck // The Bell System Technical Journal. – 1950. – Vol. 29, no. 4. – P. 560–607.

11. Введение в физику полупроводниковых диодов и методы их проектирования с использованием высокопроизводительных вычислений : учебное пособие / Е. В. Волкова [и др.] ; отв. ред. Е. З. Грибова. – Нижний Новгород : Нижегородский госуниверситет, 2020. – 78 с.

12. Рындин, Е. А. Основы численных методов: теория и практика : электронное учебное пособие / Е. А. Рындин, И. В. Куликова, И. Е. Лысенко. – Таганрог : Издательство ЮФУ, 2015. – 217 с.

13. Расчёт переходных процессов в диоде с накоплением заряда на основе точного решения диффузионно-дрейфового уравнения / А. С. Величкина, [и др.] // VII Научный форум «Телекоммуникации: теория и технологии» : сб. тр. участников XXI Междунар. научно-практической конф. «Физика и технические приложения волновых процессов» (Самара, 6–8 ноября 2024 г.). – Самара, 2024. – С. 171–172.

14. Theoretical Calculation of the Charge Distribution and Resorption Current Density in a Step Recovery Diode / A. M. Bobreshov [et al.] // 2024 IEEE 9th All-Russian Microwave Conference (Moscow, 25–29 November 2024). – Moscow, 2024. – P. 463–466.

15. Величкина, А. С. Вычислительная реализация анализа переходных процессов в диоде с накоплением заряда / А. С. Величкина, А. А. Кононов, Г. К. Усков // Сб. материалов XXV Междунар. научно-практической конф. «Информатика: проблемы, методы, технологии» (Воронеж, 12–13 февраля 2025 г.). – Воронеж, 2025. – С. 1110–1119.

16. Величкина, А. С. Теоретический анализ зависимости распределения концентрации носителей и плотности тока рассасывания в диоде с накоплением заряда от величины электрического поля в базе / А. С. Величкина, А. А. Кононов, Г. К. Усков // Сб. ст. 12-й Всерос. научной школы-семинара «Взаимодействие сверхвысокочастотного, терагерцового и оптического излучения с полупроводниковыми микро- и наноструктурами, метаматериалами и биообъектами» (Саратов, 22–23 мая 2025 г.). – Саратов, 2025. – С. 102–106.

17. Бобрешов, А. М. Математическое моделирование переходных процессов в базе полупроводникового диода с накоплением заряда / А. М. Бобрешов, А. С. Величкина, Г. К. Усков // Инфокоммуникационные и радиоэлектронные технологии. – 2025. – Т. 8, № 5. – С. 704–717.

18. Тихонов, А. Н. Уравнения математической физики : учеб. пособие / А. Н. Тихонов, А. А. Самарский. – 6-е изд., испр. и доп. – М. : Изд-во МГУ, 1999. – 799 с.

19. Владимиров, В. С. Уравнения математической физики / В. С. Владимиров. – 4-е изд., испр. и доп. – М. : Наука, 1981. – 512 с.

20. Hassani, S. Mathematical Physics: A Modern Introduction to Its Foundations / S. Hassani. – Springer-Verlag, 1999. – 1043 p.

21. Зон, Б. А. Лекции по интегральным уравнениям / Б. А. Зон. – М. : «Высшая школа», 2004. – 92 с.

22. Васильева, А. Б. Интегральные уравнения / А. Б. Васильева, Б. А. Зон. – 2-е изд., стереот. – М. : ФИЗМАТЛИТ, 2002. – 160 с.

23. Корн, Г. Справочник по математике для научных работников и инженеров / Г. Корн, Т. Корн. – 2-е изд. – М. : Наука, 1970. – 720 с.

24. Бронштейн, И. Н. Справочник по математике для инженеров и учащихся втузов / И. Н. Бронштейн, К. А. Семендяев. – 13-е изд., испр. – М. : Наука, 1986. – 544 с.

Изучение модифицированного протокола канального уровня для организации связи в беспроводных сетях с направленными антеннами

В. А. Жигулин, email: obi4n@yandex.ru [1]
И. В. Пешков, email: ilvpeshko@gmail.com [1]

[1] Елецкий государственный университет им. И.А. Бунина

Аннотация. *Данная работа посвящена изучению способов организации беспроводных самоорганизующихся сетей с направленными антеннами. Описана проблема поиска соседних узлов при направленной связи. Представлен модифицированный протокол канального уровня для беспроводной сети с направленными антеннами, приводится описание принципов его работы, а также демонстрируются результаты моделирования данного протокола для условий городской застройки.*

Ключевые слова: *беспроводные самоорганизующиеся сети, ad-hoc сети, направленные антенны, MAC протоколы.*

Введение

Беспроводные самоорганизующиеся сети (БСС, беспроводные ad-hoc-сети) обладают рядом преимуществ над сетями традиционной архитектуры:

- Высокая отказоустойчивость.
- Гибкость и скорость развёртывания.
- Большое покрытие сети.
- Масштабируемость.

Одним из перспективных направлений развития БСС является внедрение в них направленных антенн. Плюсами такого подхода являются:

- Увеличение дальности связи.
- Снижение помех.
- Повышение пропускной способности.
- Повышение энергоэффективности.

Однако применение направленных антенн связано с рядом сложностей. Одна из ключевых проблем – это первоначальный поиск соседних узлов. Для установления связи узлу-передатчику и узлу-приёмнику нужно одновременно определить направления друг на друга и соответствующим образом направить антенны. Задача организации

© Жигулин В.А., Пешков И.В., 2026

направленной связи лежит на специализированных протоколах канального и сетевого уровней [1–5].

Цель данной работы – изучить модифицированный алгоритм организации связи канального уровня для беспроводных самоорганизующихся сетей с направленными антеннами. Для этого была разработана модель БСС, основанная на методе трассировки лучей, и проведено моделирование для условий городской застройки.

1. Описание алгоритма

В беспроводных сетях с направленными антеннами существует проблема обнаружения соседних узлов. Для установления связи узлу-передатчику и узлу-приёмнику нужно одновременно определить направления друг на друга и соответствующим образом перевести главный лепесток диаграммы направленности. Эта проблема особенно критична для сетей, где нет централизованного управления, а узлы могут иметь ограничения по размеру и вычислительной мощности.

В данной работе нами рассмотрен модифицированный алгоритм канального уровня, основанный на представленных в работах [3, 5] протоколах. Для поиска соседних узлов в предложенном алгоритме применяется подход с частичным формированием ДН, когда узел, производящий поиск соседей, передаёт служебное сообщение в направленном режиме, а остальные узлы осуществляют приём в ненаправленном. Данный подход позволяет достигнуть компромисса между дальностью связи и простотой конструкции узлов.

Рассмотрим этот алгоритм подробнее. Для его реализации каждый узел сети должен быть снабжён секторной антенной с переключаемой диаграммой направленности. Данная антенна может переключаться между несколькими фиксированными направлениями ДН, а также переходить в ненаправленный режим. Преимуществом подобной антенны является простота конструкции по сравнению с фазированными антенными решётками – для её работы не требуются фазовращатели или скоростные блоки цифровой обработки сигналов. Минусом подобной антенны является точность настройки – поскольку направления главного лепестка ДН заранее определены, невозможно точно направить луч в сторону приёмника.

Согласно рассматриваемому алгоритму, перед передачей полезных данных каждый узел должен обнаружить все соседние узлы и определить оптимальные сектора антенны для связи с ними. Для этого вводится специальный механизм обнаружения соседей (рис. 1), состоящий из двух этапов:

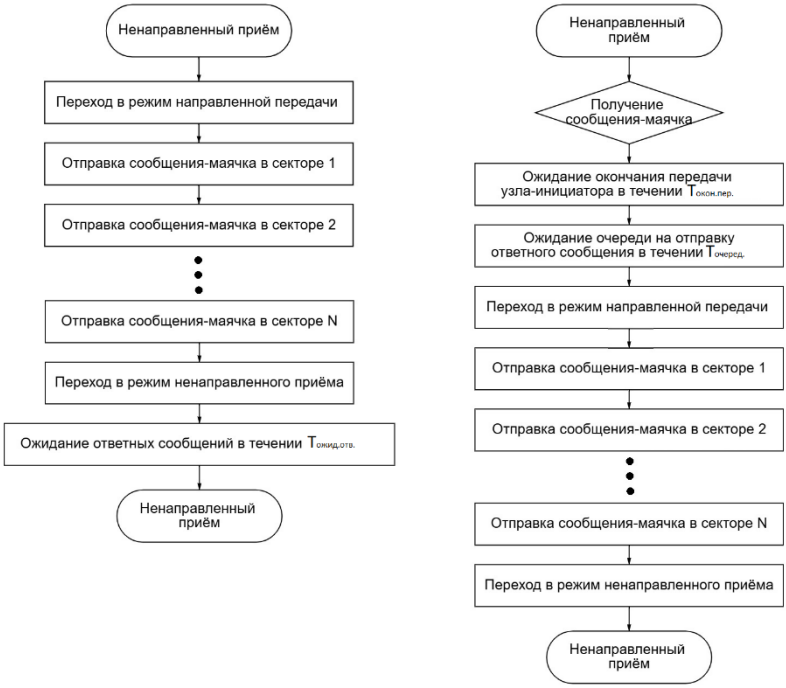


Рис. 1. Схема алгоритма обнаружения соседних узлов

Этап 1. Все узлы сети по умолчанию находятся в режиме ненаправленного приёма. В случае если один из узлов (узел-инициатор) решает обновить информацию о своих соседях, он последовательно переключает свою антенну в каждое из фиксированных направлений, и в каждом из них передаёт специальный сигнал-маячок, содержащий идентификатор узла-инициатора и номер сектора, который используется в данный момент для передачи этого сигнала. После завершения последней передачи узел переходит на приём в ненаправленном режиме.

Все соседи узла-инициатора, получившие один или несколько маячков, сохраняют в памяти мощность принятого сигнала и указанный в нём номер сектора, а также запускают таймер окончания передачи узла-инициатора. Значение таймера определяется на основе информации из первого принятого сообщения-маячка по следующей формуле:

$$T_{окон.пер.} = \left(N_{секторов.всего} - N_{сектора} \right) \cdot t_{маячка} \qquad (1)$$

где $t_{маячка}$ – время передачи одного сообщения-маячка узлом-инициатором; $N_{секторов.всего}$ – количество секторов в антеннах. За счёт данного таймера узлы-соседи определяют, когда узел-инициатор закончит передачу и перейдёт на приём.

Этап 2. По окончанию таймера узлы определяют, какой маячок был принят с наибольшей мощностью. Соответствующий этому маячку номер сектора указывает направление, которое узел-инициатор должен использовать для последующей передачи данных.

Узел-инициатор ждёт ответные сообщения в ненаправленном режиме. Поскольку узлы-соседи не знают, в каком направлении находится инициатор, они переключаются в направленный режим и начинают отправку ответного сообщения по очереди в каждом секторе (аналогично узлу-инициатору и сообщению-маячку). Ответное сообщение содержит собственный идентификатор узла-соседа, идентификатор узла-инициатора и номер оптимального сектора.

Для предотвращения коллизий от ответных сообщений на узле-инициаторе используется дополнительный таймер задержки отправления. Значение таймера определяется на каждом узле-соседе по следующей формуле:

$$T_{очеред.} = \left(N_{оптим.сектора} - 1\right) \cdot \left(t_{ответа} \cdot N_{секторов.всего}\right) \quad (2)$$

где $t_{ответа}$ – время передачи одного ответного сообщения. Только после окончания данного таймера узел-сосед начинает отправку. Таким образом, на узел-инициатор одновременно поступают сообщения от узлов, находящихся только в одном из секторов. За счёт этого снижается вероятность коллизии сигналов, однако увеличивается общее время на обнаружение.

Узел-инициатор ожидает ответные сообщения в течении:

$$T_{ожид.ответа} = N_{секторов.всего} \cdot \left(N_{секторов.всего} \cdot t_{ответа}\right) \quad (3)$$

На основе полученных ответных сообщений инициатор определяет оптимальные сектора для связи с соседними узлами. Далее данный механизм необходимо повторить для всех узлов сети.

Алгоритм отправки полезных данных представляет собой модификацию протокола CSMA/CA, адаптированную для направленной связи. Узлы сети всё так же по умолчанию находятся в режиме ненаправленного приёма. Узел, желающий отправить пакет данных (узел-инициатор), предварительно прослушивает эфир в ненаправленном режиме в течение интервалов DIFS и BackOff (как и в CSMA/CA), после чего переходит в направленный режим и отправляет сообщение-запрос

RTS сначала интересующему узлу (узлу-получателю), а затем по очереди в каждом секторе. После этого узел-инициатор снова переключается на приём, но уже в направленном режиме в направлении узла-получателя и начинает ожидать ответное сообщение-подтверждение CTS. Если в течение установленного периода CTS не получен, узел-инициатор делает вывод о занятости узла-получателя, увеличивает число попыток и начинает цикл отправки заново.

Получив сообщение RTS узел-получатель переключается в направленный режим, переключает сектор в направлении на узел-инициатор и ожидает, когда он закончит рассылку сообщений RTS и начнёт приём. Время ожидания определяется как:

$$T_{окончания.RTS} = \left(N_{секторов.всего} - 1\right) \cdot t_{RTS} \quad (4)$$

где t_{RTS} – время на отправку одного сообщения RTS. Соседние узлы, получившие RTS от узла-инициатора, устанавливают себе запрет на связь в секторе с узлом-получателем в течение указанного в сообщении времени (если среди их соседей есть этот узел-получатель). Однако при этом они всё ещё могут устанавливать связь с узлами в других секторах, реализуя тем самым пространственное разделение каналов связи.

После окончания ожидания узел-получатель отправляет CTS узлу-инициатору, после чего между ними происходит обмен пакетами данных как в оригинальном протоколе CSMA/CA.

2. Описание модели

Для изучения представленного алгоритма нами была подготовлена модель беспроводной сети в среде Matlab [6]. Основные параметры модели:

- Число узлов 5;
- Среднее расстояние между узлами 180 м;
- Высота крепления антенн 1,5 м;
- Частота 2,4 ГГц;
- Мощность передатчиков 10 дБм;
- Чувствительность приёмников -75 дБм.

Модель радиосвязи основана на методе трассировки лучей. Данный метод относится к классу детерминированных и основан на многолучевой модели распространения радиоволн [7]. Принцип его работы следующий: из точки передатчика (Tx на рис.2) формируется множество лучей на одинаковом угловом расстоянии друг от друга. Затем путь каждого луча моделируется по отдельности, с учётом взаимодействия с окружающими объектами.

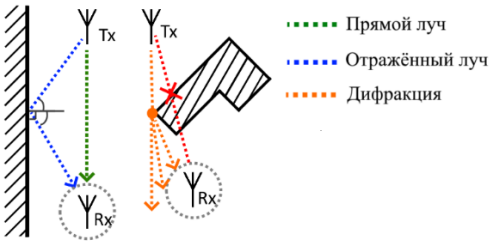


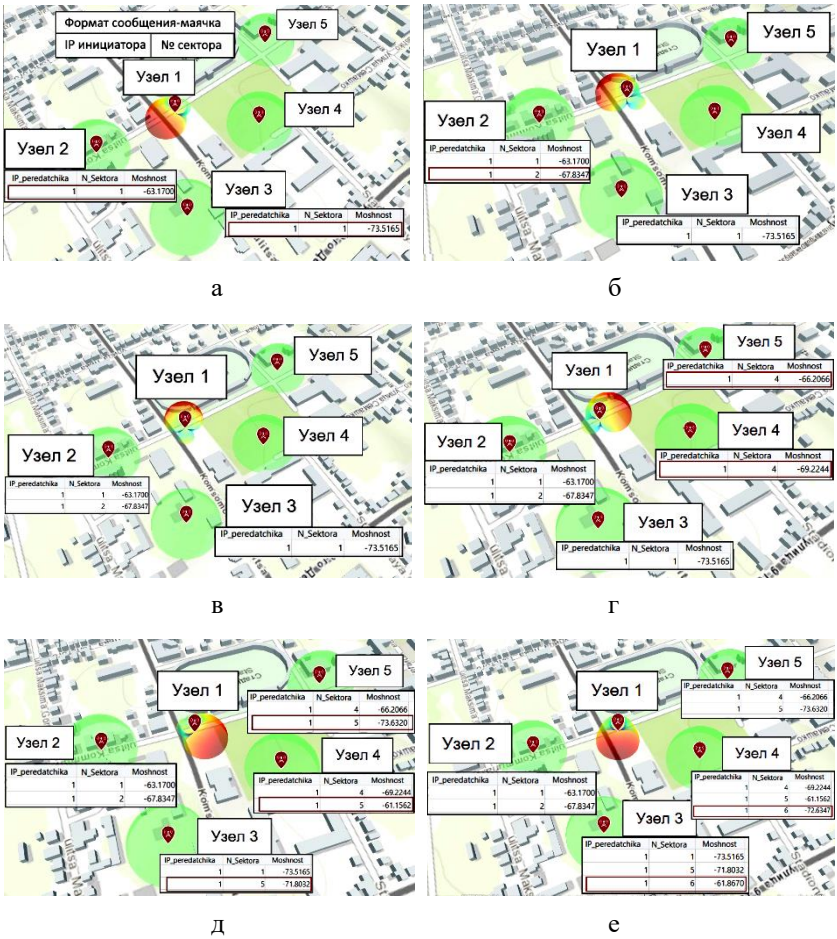
Рис. 2. Иллюстрация к методу трассировки лучей

Когда луч попадает на плоскую поверхность, он отражается под углом, равным углу падения в соответствии с законом отражения. Когда луч попадает на грань, он порождает множество дополнительных лучей, моделируя тем самым явление дифракции волн. Вокруг точки приёма (Rx на рис.2) формируется «сфера приёма». Если луч пересекает эту сферу, то модель считает его допустимым путём от Tx к Rx. Метод трассировки лучей обладает наибольшей точностью. Он хорошо подходит для моделирования распространения радиосигналов в сложных условиях городской застройки.

Кроме того, нами была подготовлена модель секторной антенной с переключаемой диаграммой направленности. Число секторов – 6. Ширина главного лепестка антенны составляет 62°. В качестве антенных элементов выступают микрополосковые антенны для диапазона 2.4 ГГц. В каждый момент времени подключён только один элемент антенны. Поворот главного луча ДН осуществляется путём переключения элементов одного за другим. Плюс такого подхода – простота конструкции по сравнению с фазированными и цифровыми антенными решётками. Однако, из-за того, что ДН разделена на фиксированные секторы, точно направить главный лепесток в нужном направлении может быть невозможно.

3. Результаты моделирования

Рассмотрим процесс работы сети с представленным алгоритмом. Вначале узлы должны обнаружить своих соседей. В памяти каждого узла хранится специальная таблица, в которой указываются идентификаторы соседних узлов и оптимальные сектора для связи с ними. Изначально узлы не имеют никакой информации о соседях, их таблицы пусты. Пусть узел 1 инициирует сбор информации о соседях. Он переключается в направленный режим в первом секторе (рис. 3.а) и начинает передачу сообщения-маячка.



а б

в г

д е

а – первый сектор, б – второй сектор, в – третий сектор, г – четвёртый сектор, д – пятый сектор, е – шестой сектор

Рис. 3. Первый этап поиска соседних узлов. Переключение секторов антенны узла-инициатора

В модели в этот момент происходит расчёт мощности сигнала на каждом узле. Если мощность выше чувствительности приёмника, то узел принимает сообщение и запоминает идентификатор узла-инициатора, номер сектора и мощность.

Узел-инициатор по очереди переключается между всеми секторами (рис. 3, а–е), после чего переходит в режим ненаправленного приёма (рис. 4). Остальные узлы, дождавшись окончания всех передач от узла-инициатора, определяют, при каком номере сектора мощность принятого сигнала была максимальной. Далее они ожидают окончания таймера, переключаются в направленный режим и начинают транслировать ответное сообщение, с указанием собственного идентификатора и номером оптимального сектора.

Передача этого сообщения осуществляется с переключением секторов, аналогично сообщению-маячку. На рис. 5 представлена итоговая таблица узла 1 после получения всех ответных сообщений.



а б

а – отправка ответного сообщения узлом 2,
б – отправка ответного сообщения узлом 3

Рис. 4. Второй этап поиска соседних узлов. Отправка ответных сообщений соседними узлами

IP_Priyomnika	N_Sektora
2	1
3	6
4	5
5	4

Рис. 5. Итоговая таблица секторов узла-инициатора

Далее кратко рассмотрим, как происходит отправка полезных данных. Пусть узел 1 решает отправить пакет данных узлу 3. Для этого он, после прослушивания эфира в ненаправленном режиме, отправляет сообщение RTS сначала узлу 3 (рис. 6.а), а затем и всем своим соседям (рис .6.б). Узел-получатель при этом уже перестроил свою антенну в направлении на узел-инициатор.



а б

а – отправка RTS узлу-получателю, б – отправка RTS остальным соседним узлам

Рис. 6. Отправка сообщения RTS узлом-инициатором

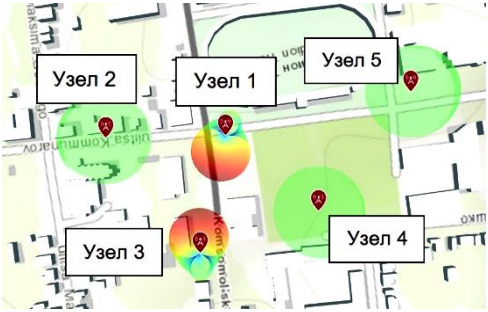


Рис. 7. Отправка сообщения CTS узлом-получателем

Далее узел 1 устанавливает сектор для связи с узлом 3, после чего между ними происходит обмен сообщениями согласно CSMA/CA.

Заключение

В данной работе был представлен модифицированный алгоритм канального уровня для беспроводной сети с направленными антеннами. Для разрешения проблемы поиска соседних узлов в предложенном алгоритме применён подход с частичным формированием ДН. Особенностью алгоритма является способ определения оптимального сектора антенны не на стороне передатчика, а на приёмниках.

Для проверки предложенного алгоритма была разработана модель беспроводной сети в среде Matlab. В работе описан метод построения модели, а также представлены результаты моделирования для условий городской застройки по методу трассировки лучей.

Список литературы

1. An adaptive MAC protocol for ad hoc networks with directional antenna / S. Bandyopadhyay [et al.] // IEICE Trans. Commun. – 2001. – Vol. 84, № 11.

2. Bao, L. Transmission scheduling in ad hoc networks with directional antennas / L. Bao, J. J. Garcia-Luna-Aceves // Proceedings of the 8th Annual International Conference on Mobile Computing and Networking (MobiCom 2002) (Atlanta, GA, USA, September 23-28, 2002). – 2002. – P. 48–58.

3. BeamNet: An Ad Hoc Network Testbed Using Beamforming Antennas / R. R. Choudhury [et al.] // 2005 IEEE 61st Vehicular Technology Conference (VTC 2005-Spring) (Stockholm, Sweden, May 30 – June 1, 2005). – 2005. – Vol. 2. – P. 673-677.

4. Inzillo, V. A deep learning clustering beamforming approach for future 6G mobile Ad Hoc networks / V. Inzillo, D. Garompolo // Computers and Electrical Engineering. – 2023. – Vol. 111.

5. MAC Protocol for Mobile Ad Hoc Networks using Directional Antennas / A. Nasipuri [et al.] // 2000 IEEE Wireless Communications and Networking Conference (WCNC 2000) (Chicago, IL, USA, September 23-28, 2000). – 2000. – Vol. 3. – P. 1213–1217.

6. Жигулин, В. А. Моделирование самоорганизующейся беспроводной сети в условиях городской застройки по методу трассировки лучей / В. А. Жигулин, И. В. Пешков // Радиолокация, навигация, связь : сб. тр. XXXI Международной научно-технической конференции : в 6 т. (Воронеж, 15–17 апр. 2025 г.). – Воронеж, 2025. – Т. 3. – С. 290–296.

7. MathWorks. Ray Tracing for Wireless Communications – MATLAB & Simulink [Электронный ресурс] : справочный центр. – Режим доступа: https://www.mathworks.com/help/comm/ug/ray-tracing-for-wireless-communications.html

Оптимизация пропускной способности MIMO-канала в условиях высокоскоростного движения на основе методов адаптивного разнесения и предварительного кодирования

М. А. Калюкин, email: mixkalyukin@mail.ru

ФГБОУ ВО «Елецкий государственный университет им. И. А. Бунина», г. Елец

Аннотация. *В работе рассматриваются проблемы эксплуатации MIMO-систем в условиях высокоскоростного движения (до 350 км/ч), в частности, рост пространственной корреляции, ведущий к деградации пропускной способности. Проводится сравнительный анализ двух методов компенсации: адаптивного изменения геометрии антенной решётки и пространственного предварительного кодирования на основе SVD-разложения. Моделирование выполнено на основе реалистичной модели канала WINNER II.*

Ключевые слова: *MIMO, высокоскоростное движение, пространственная корреляция, адаптивное разнесение, SVD-предварительное кодирование, пропускная способность, BER.*

Введение

Эксплуатация систем беспроводной связи в условиях высокоскоростного движения, характерных для современных железнодорожных магистралей, сопряжена с рядом фундаментальных проблем. При скоростях движения до 350 км/ч мобильная станция подвергается воздействию значительного доплеровского сдвига частоты, быстрому изменению условий распространения радиоволн и снижению точности оценки состояния канала. Особую сложность представляет обеспечение работы многоканальных MIMO-систем, эффективность которых критически зависит от пространственного разнесения сигналов [1].

Пространственная корреляция между элементами антенных решёток возрастает с увеличением скорости движения вследствие уменьшения углового разброса прибывающих сигналов. Это приводит к снижению ранга канальной матрицы, ухудшению её обусловленности и, как следствие, к деградации пропускной способности системы [2]. Для компенсации данного эффекта требуются специализированные методы адаптации, учитывающие динамику изменения условий распространения.

© Калюкин М. А., 2026

Целью настоящей работы является сравнительный анализ эффективности двух подходов к снижению пространственной корреляции в MIMO-системах: адаптивного изменения межэлементного расстояния антенн и пространственного предварительного кодирования на основе сингулярного разложения (SVD). На основе реалистичной модели канала WINNER II [3] проводится численное моделирование для оценки влияния указанных методов на пропускную способность и вероятность битовой ошибки в диапазоне скоростей от 0 до 350 км/ч.

1. Аналитическая модель пространственной корреляции

Основным фактором, ограничивающим производительность MIMO-систем при высоких скоростях, является пространственная корреляция, описываемая моделью Кронекера [2, 3], которая выражает канальную матрицу H через матрицы пространственной корреляции на приёме и передаче (1):

$$H = \sqrt{R_{Rx}} * H_{iid} * \sqrt{R_{Tx}^{T}} \quad (1)$$

где H_{iid} – матрица канала с независимыми комплексными гауссовскими коэффициентами $\sim CN(0,1)$, R_{Rx} – матрица пространственной корреляции на стороне приёмника, R_{Tx} – матрица пространственной корреляции на стороне передатчика.

Для унифицированной круговой антенной решётки (Uniform Circular Array, UCA) элементы матрицы пространственной корреляции определяются более сложным выражением, учитывающим геометрию решётки и статистические характеристики приходящих сигналов [3] справедлива формула (2):

$$[R]mn = J_0\left(\frac{2\pi d_{mn}}{\lambda}\right) + 2 \sum_{p \to \infty}^{1} J_{2p}\left(\frac{2\pi d_{mn}}{\lambda}\right) * \cos(2p\Delta\theta) \quad (2)$$

где $[R]mn$ – элемент матрицы корреляции, находящийся на пересечении m-й строки и n-го столбца; J_p – функция Бесселя первого рода порядка p, d_{mn} – физическое расстояние между m-м и n-м элементами антенной решётки. $\Delta\theta$ – угловой разброс (Angular Spread) прибывающих сигналов, характеризующий ширину распределения углов прихода.

При высоких скоростях движения $\Delta\theta$ уменьшается, что приводит к увеличению пространственной корреляции; Суммирование по p от 1 до ∞ учитывает высшие гармоники в разложении корреляционной функции.

2. Методы снижения пространственной корреляции

Метод оптимизации геометрии антенных решёток

Для снижения пространственной корреляции предлагается использовать адаптивную антенную конфигурацию с переменным межэлементным расстоянием. Эффективное расстояние между антеннами определяется как функция от текущей скорости движения (3):

$$d_{eff} = d_0 \cdot (1 + \alpha \cdot \frac{\upsilon}{\upsilon_{max}}) \quad (3)$$

где d_{eff} – эффективное (адаптивное) расстояние между антеннами, изменяющееся в зависимости от скорости движения; d_0 – базовое (номинальное) расстояние между антеннами при нулевой скорости или в стационарных условиях; α – коэффициент адаптации, безразмерный параметр в диапазоне $0 < \alpha < 1$, определяющий степень влияния скорости на изменение расстояния.

Модифицированная матрица корреляции с учётом адаптивного расстояния принимает вид (4), который учитывает не только изменённую геометрию, но и зависимость углового разброса от скорости [3]:

$$\left[R_{adapt}\right]_{mn} = J_0\left(\frac{2\pi d_{mn}^*}{\lambda} * \exp\left(\frac{-\sigma_\theta^2(v)}{2} * \left(\frac{2\pi d_{mn}^*}{\lambda}\right)^2\right)\right) \quad (4)$$

где $\left[R_{adapt}\right]_{mn}$ – элемент адаптированной матрицы корреляции; $mn*$ – адаптивное расстояние между m-й и n-й антеннами, вычисляемое по формуле выше с учётом их взаимного расположения; $\sigma_\theta^2(v)$ – дисперсия угла прихода сигналов, являющаяся функцией скорости v; Экспоненциальный множитель $\exp\left(\frac{-\sigma_\theta^2(v)}{2} * \left(\frac{2\pi d_{mn}^*}{\lambda}\right)^2\right)$ описывает уменьшение корреляции за счёт углового разброса. Чем больше угловой разброс, тем быстрее спадает корреляция с увеличением расстояния между антеннами.

Метод SVD-предварительного кодирования (SVD Precoding Method)

Для компенсации эффекта пространственной корреляции предлагается использовать алгоритм предварительного кодирования (precoding) на основе сингулярного разложения (Singular Value Decomposition, SVD) модифицированной канальной матрицы. Этот метод позволяет преобразовать коррелированный MIMO-канал в набор параллельных независимых подканалов.

Сингулярное разложение модифицированной канальной матрицы имеет вид (5) [4, 5]:

$$H_{mod} = U\Sigma V^{H} \quad (5)$$

где H_{mod} – модифицированная канальная матрица размерностью $Nr \times Nt$; U – унитарная матрица размерностью $Nr \times Nr$, столбцы которой являются левыми сингулярными векторами; Σ – диагональная матрица размерностью $Nr \times Nt$, на диагонали которой расположены сингулярные числа $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r \geq 0$, где $r = \min(Nr, Nt)$; – ранг матрицы; V – унитарная матрица размерностью $Nt \times Nt$, столбцы которой являются правыми сингулярными векторами;

Предварительное кодирование осуществляется с помощью матрицы V, при этом передаваемый сигнал модифицируется следующим образом (6):

$$X^* = V \cdot x \quad (6)$$

где x – исходный вектор передаваемого сигнала размерностью $Nr \times 1$; X^* – вектор предварительно закодированного сигнала размерностью $Nt \times 1$; Операция умножения $V * x$ представляет собой линейное преобразование, которое подготавливает сигнал для передачи через коррелированный канал.

Модифицированная канальная матрица с учётом декорреляции может быть выражена через исходную канальную матрицу и матрицу пространственной корреляции на передаче, согласно формуле (7):

$$H_{mod} = H \cdot R_{Tx}^{(-1/2)} \quad (7)$$

где H – исходная канальная матрица, полученная из модели Кронекера;

$R_{Tx}^{(-1/2)}$ – обратный квадратный корень матрицы корреляции на передаче. Эта операция компенсирует эффект пространственной корреляции на стороне передатчика;

Матрица $R_{Tx}^{(-1/2)}$ может быть вычислена как (8):

$$R_{Tx}^{(-1/2)} = U\Lambda^{(1/2)}U^{H} \quad (8)$$

где U – матрица собственных векторов R_{Tx}, Λ – диагональная матрица собственных значений.

3. Модифицированная модель канала

С учётом предложенных методов, модифицированная канальная матрица принимает вид, объединяющий адаптацию геометрии, предварительное кодирование (9):

$$H_{decorr} = \sqrt{R_{Rx,adapt}} \cdot H_{iid} \cdot \sqrt{R_{Tx,adapt}^{T}} \quad (9)$$

где H_{decorr} – окончательная декоррелированная канальная матрица; $R_{Rx,adapt}$ и $R_{Tx,adapt}$ – адаптированные матрицы пространственной корреляции на приёме и передаче соответственно, вычисленные с учётом переменного расстояния между антеннами;

Результирующая пропускная способность MIMO-системы с учётом декорреляции вычисляется по формуле Шеннона (10) для MIMO-каналов [5]:

$$C_{decorr} = \log_2 \det(I + \frac{SNR}{N_t} * H_{decorr} * H_{decorr}^{H}) \quad (10)$$

где C_{decorr} – пропускная способность декоррелированного канала в битах в секунду на герц (бит/с/Гц);

Предложенные методы позволяет существенно снизить пространственную корреляцию и сохранить преимущества MIMO-систем даже в экстремальных условиях высокоскоростного движения.

4. Методология и параметры моделирования

Для верификации аналитической модели проведено имитационное моделирование в MATLAB с использованием библиотеки WINNER II Channel Model. Конфигурация системы:

Антенные массивы:

BS: UCA (4 элемента, R=0.3 м), MS: UCA (4 элемента, R=0.05 м).

Сценарий: B4 (Outdoor-to-Indoor), условия NLOS.

Скорость движения: 0–350 км/ч.

Параметры сигнала: Несущая частота 2.4 ГГц, SNR = 15 дБ, модуляция BPSK.

5. Результаты моделирования

На рис. 1–4 представлены сравнительные характеристики трёх конфигураций системы в зависимости от скорости движения.

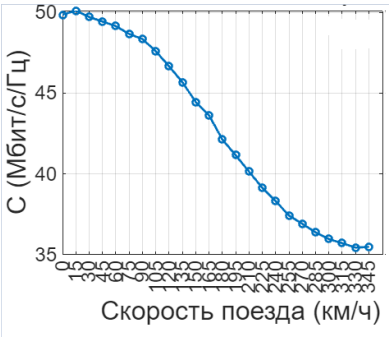
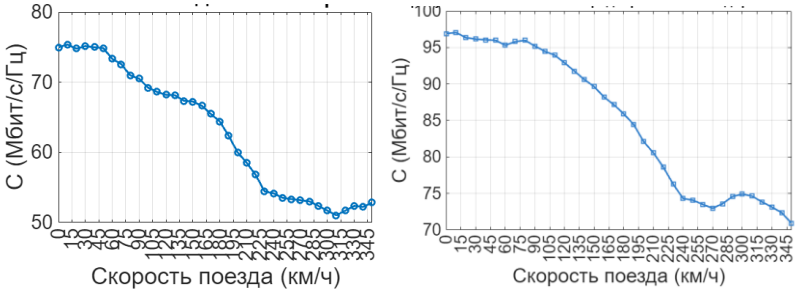


Рис. 1. Пропускная способность MIMO-системы



а б

а – с адаптивным разнесением, б – с SVD.

Рис. 2. Пропускная способность MIMO-системы в условиях адаптивного разнесения и предварительного кодирования.

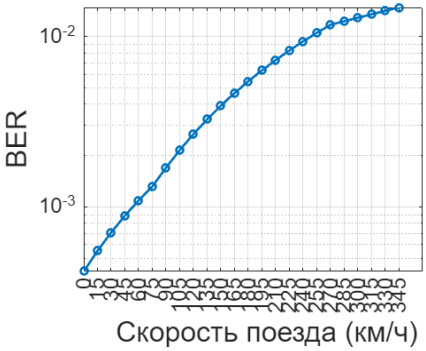


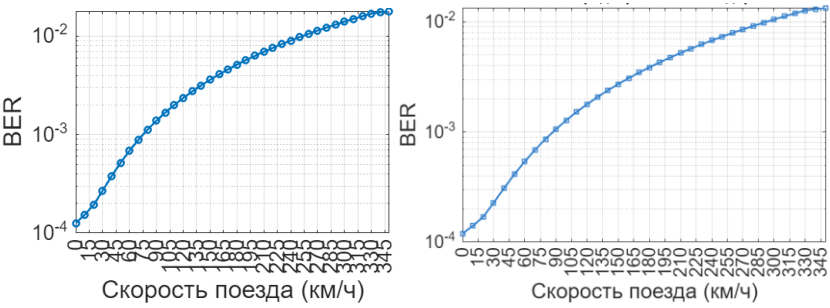
Рис. 3. Битовая ошибка MIMO-канала от скорости поезда

При расчете d_{eff} были получены следующие значения:

Скорость 0 км/ч: d_{eff_MS} : 0.050 м, d_{eff_BS} : 0.300 м

Скорость 170 км/ч: d_{eff_MS} : 0.062 м, d_{eff_BS} : 0.373 м

Скорость 350 км/ч: d_{eff_MS} : 0.075 м, d_{eff_BS} : 0.450 м



а б

а – с адаптивным разнесением, б – с SVD

Рис. 4. Битовая ошибка MIMO-канала от скорости поезда в условиях адаптивного разнесения и предварительного кодирования

Заключение

Проведённое исследование демонстрирует эффективность методов адаптивного разнесения антенн и SVD-предварительного кодирования для снижения пространственной корреляции в MIMO-системах при высокоскоростном движении. Полученные результаты свидетельствуют о следующем:

Метод адаптивного разнесения позволяет уменьшить пространственную корреляцию за счёт увеличения физического расстояния между антеннами пропорционально скорости движения, что обеспечивает улучшение пропускной способности по сравнению с базовой системой.

SVD-предварительное кодирование демонстрирует наибольшую эффективность, обеспечивая повышение пропускной способности и снижение BER по сравнению с базовой конфигурацией. Данный метод позволяет компенсировать пространственную корреляцию на уровне обработки сигналов без изменения физической конфигурации антенной системы.

Комбинированное применение обоих методов может обеспечить синергетический эффект, однако требует дополнительного исследования с учётом вычислительной сложности и требований к обратной связи в системе.

Предложенные подходы могут быть использованы при проектировании систем связи для высокоскоростного железнодорожного транспорта, где обеспечение стабильного качества связи при экстремальных скоростях движения является критически важным требованием.

Список литературы

1. Tse, D. Fundamentals of Wireless Communication / D. Tse, P. Viswanath. – Cambridge : Cambridge University Press, 2005. – 564 p.

2. Telatar, E. Capacity of multi-antenna Gaussian channels / E. Telatar // European Transactions on Telecommunications. – 1999. – Vol. 10, № 6. – P. 585–595.

3. WINNER II Channel Models [Электронный ресурс] : техн. отчёт IST-WINNER II, D1.1.2. – Режим доступа: https://www.cept.org/files/8339/winner2%20-%20final%20report.pdf – 2007

4. Salz, J. Effect of fading correlation on adaptive arrays in digital mobile radio / J. Salz, J. H. Winters // IEEE Transactions on Vehicular Technology. – 1994. – Vol. 43, № 4. – P. 1049–1057.

5. Keyholes, correlations, and capacities of multielement transmit and receive antennas / D. Chizhik [et al.] // IEEE Transactions on Wireless Communications. – 2002. – Vol. 1, № 2. – P. 361–368.

Разработка алгоритма акустического мониторинга воздушного пространства

Ю. А. Клименко, email: klm71165@mail.ru [1]
А. П. Преображенский, email: app@vivt.ru [1]
Я. Б. Кораблев, klm71165@mail.ru [2]
А. И. Родюков, email: klm71165@mail.ru [2]

[1] Воронежский институт высоких технологий
[2] ВУНЦ ВВС «Военно-воздушная академия имени профессора Н. Е. Жуковского и Ю. А. Гагарина», г. Воронеж

Аннотация. *В данной работе рассматривается проблема разработки и применения средств противодействия беспилотным летательным аппаратам. Для этой цели предложена система акустического мониторинга воздушного пространства для выявления БПЛА, их сопровождения, определения точных координат и уничтожения средствами ПВО. Предлагаемы алгоритм относится к программному обеспечению средств акустической локации. На основе данного алгоритма создана компьютерная программа.*

Ключевые слова: *акустика, локация, мониторинг, система, воздушное пространство, алгоритм, обнаружение, летательный аппарат.*

Введение

В современных условиях развития беспилотных технологий создаются новые средства поражения воздушного, морского и наземного базирования. Разрабатывается новая стратегия их применения. Особенно бурное развитие беспилотные технологии получили в военной сфере путем создания ударных средств огневого поражения людей, техники, объектов критической инфраструктуры в энергетике, промышленности, транспортной логистике [1]. Видим, что в современных условиях применение беспилотных авиационных средств наносит большой материальный урон различным типам инфраструктурных объектов, а также приводит к смертельному поражению людей, военнослужащих и гражданских лиц.

В данное время самым применяемым способом воздействия беспилотные летательные аппараты является огневое поражение с помощью вооружения различных видов [2]. Для этого применяются все

© Клименко Ю. А., Преображенский А. П., Кораблев Я. Б., Родюков А. И., 2026

их виды от применения средств противовоздушной обороны до обыкновенного стрелкового и использования охотничьих ружей.

Развитие технологий по созданию средств по противодействию беспилотным проектам заметно отстает от необходимости и не получило должного научного и технического прогресса в этой области, не считая, конечно, системы радиоэлектронной борьбы [3]. К сожалению, сфера применения средств радиоэлектронной борьбы тоже носит ограниченный характер воздействия. Создание высокоинтеллектуальных беспилотных средств позволяет решать вопросы управления беспилотными летательными аппаратами без использования радиосредств и оптоволокна [4].

Поэтому есть необходимость в развитии средств противодействия БПЛА, которые основаны на принципах простоты конструкции, надежности, экономичности и результативности применения [5].

Алгоритм акустического мониторинга воздушного пространства

Система акустического обнаружения предназначена для выявления движения беспилотных летательных аппаратов методом звуковой локации путем выявления звуковых сигнатур в окружающем воздушном пространстве, а также определения параметров беспилотных летательных аппаратов [6]:

– выявление факта пересечения БПЛА линии акустического контроля воздушного пространства;

– определения типа летательного аппарата;

– определение основных параметров полета (скорость, высота; координаты местоположения и другие);

– определение направления (маршрута) движения;

– определение массы перевозимого груза летательным аппаратом.

Структурная схема системы акустического обнаружения беспилотных летательных аппаратов представлена на рис. 1.

В состав системы акустического обнаружения беспилотных летательных аппаратов [7] входят следующие структурные элементы:

– акустические датчики регистрации звуковых сигналов (АД);

– акустические антенны приема звуковых сигналов (ААП);

– средства усиления звуковых сигналов (УЗС);

– средства аналого-цифрового преобразования звуковых сигналов (АЦП);

– средства цифровой обработки звуковых сигналов (СЦО);

– средства спектрального анализа звуковых сигналов (ССА);

– программное обеспечение обработки данных (ПО ОД);

– база данных (БД);

– средства связи (СС).



Рис. 1. Структурная схема системы акустического обнаружения беспилотных летательных аппаратов

На рис. 2 представлен алгоритм мониторинга воздушного пространства.

Шаг 1. Начало работы алгоритма.

Производится подача электропитания на все структурные технические элементы системы [8].

Для нормального функционирования системы обнаружения БПЛА необходимо вводить следующие данные:

– точные координаты размещения акустических антенных устройств (ААУ);

– данные о состоянии окружающей среды (температура воздуха, погодные условия);

– характеристики ААУ;

– картографические и топографические данные;

– эталонные акустические образцы различных типов БПЛА.

Шаг 3. Проверка работоспособности системы.

Проверка работоспособности заключается в проведении технологического контроля функционирования всех технических элементов структуры системы:

Шаг 4. Вывод о работоспособности системы.

После проведения проверка работоспособности системы принимается решение для дальнейшей работы алгоритма.

Если в системе сбои в работе элементов структуры отсутствуют, то принимается решение «Да» и производится переход к пункту алгоритма «Шаг 5».

Если в системе произошли сбои в работе элементов структуры, то принимается решение «Нет» и производится переход к пункту алгоритма «Шаг 3».

Шаг 3. Проверка работоспособности системы.

Проверка работоспособности заключается в проведении технологического контроля функционирования всех технических элементов структуры системы.

Шаг 4. Вывод о работоспособности системы.

После проведения проверка работоспособности системы принимается решение для дальнейшей работы алгоритма.

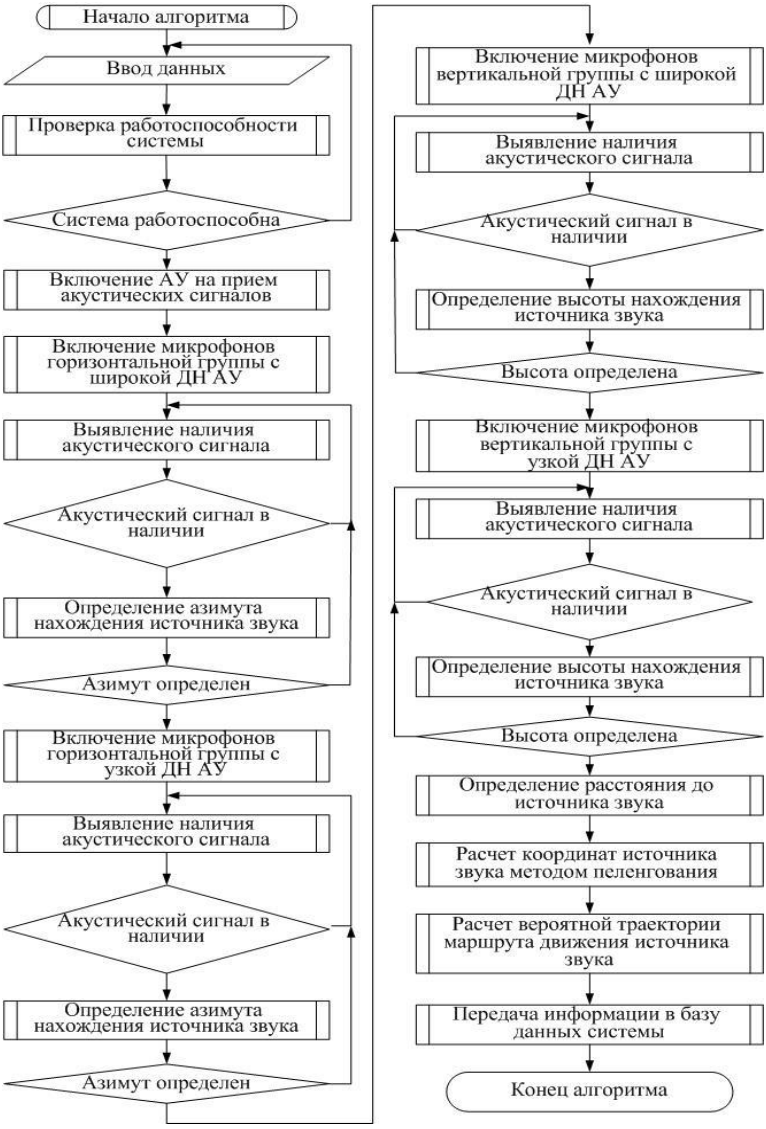


Рис. 2. Алгоритм акустического мониторинга воздушного пространства

Шаг 5. Включение АУ на прием акустических сигналов.

Если в системе сбои в работе элементов структуры отсутствуют, то принимается решение «Да» и производится переход к пункту алгоритма «Шаг 5».

Если в системе произошли сбои в работе элементов структуры, то принимается решение «Нет» и производится переход к пункту алгоритма «Шаг 3».

Производится включение ААУ на прием акустических сигналов.

Шаг 6. Включение микрофонов горизонтальной группы с широкой ДН.

Данная группа микрофонов позволяет «грубо» определять азимут расположения источника звука.

Шаг 7. Определение наличия звукового сигнала.

Определяется наличие на приеме акустического сигнала источника звука.

Шаг 8. Акустический сигнал в наличии?

Если осуществляется прием акустического сигнала микрофонами, то принимается решение «Да» и производится переход к пункту алгоритма «Шаг 9».

Если прием акустического сигнала микрофонами не осуществляется, то принимается решение «Нет» и производится переход к пункту алгоритма «Шаг 7».

Шаг 9. Определение азимута нахождения источника звука.

Отсчитывается азимут нахождения источника звука по углу поворота ААУ на основании приема звукового сигнала с максимальной амплитудой.

Шаг 10. Азимут определен?

Если азимут нахождения источника звука определен на основании приема звукового сигнала с максимальной амплитудой, то принимается решение «Да» и производится переход к пункту алгоритма «Шаг 11».

Если азимут нахождения источника звука не определен на основании приема звукового сигнала с максимальной амплитудой., то принимается решение «Нет» и производится переход к пункту алгоритма «Шаг 7».

Шаг 11. Включение микрофонов горизонтальной группы с узкой ДН.

Данная группа микрофонов позволяет «точно» определять азимут расположения источника звука.

Шаг 12. Определение наличия звукового сигнала.

Определяется наличие на приеме акустического сигнала источника звука.

Шаг 13. Акустический сигнал в наличии?

Если осуществляется прием акустического сигнала микрофонами с узкой ДН, то принимается решение «Да» и производится переход к пункту алгоритма «Шаг 14».

Если прием акустического сигнала микрофонами не осуществляется, то принимается решение «Нет» и производится переход к пункту алгоритма «Шаг 12».

Шаг 14. Определение азимута нахождения источника звука.

Отсчитывается азимут нахождения источника звука по углу поворота ААУ с микрофонами с узкой ДН на основании приема звукового сигнала с максимальной амплитудой.

Шаг 15. Азимут определен?

Если азимут нахождения источника звука определен на основании приема звукового сигнала с максимальной амплитудой, то принимается решение «Да» и производится переход к пункту алгоритма «Шаг 16».

Если азимут нахождения источника звука не определен на основании приема звукового сигнала с максимальной амплитудой, то принимается решение «Нет» и производится переход к пункту алгоритма «Шаг 12».

Шаг 16. Включение микрофонов вертикальной группы с широкой ДН.

Данная группа микрофонов позволяет «грубо» определять азимут расположения источника звука. Расчет производится на основе разности времени прихода акустического сигнала.

Шаг 17. Определение наличия звукового сигнала.

Определяется наличие на приеме акустического сигнала источника звука.

Шаг 18. Акустический сигнал в наличии?

Если осуществляется прием акустического сигнала микрофонами вертикальной группы с широкой ДН, то принимается решение «Да» и производится переход к пункту алгоритма «Шаг 19».

Если прием акустического сигнала микрофонами вертикальной группы не осуществляется, то принимается решение «Нет» и производится переход к пункту алгоритма «Шаг 17».

Шаг 19. Определение высоты (угла места) нахождения звукового сигнала.

Определяется высоты нахождения акустического сигнала источника звука.

Шаг 20. Высота (угол места) определена.

Определяется высота нахождения источника звука по углу поворота ААУ с микрофонами с широкой ДН на основании приема звукового сигнала с максимальной амплитудой.

Если высота нахождения источника звука определен на основании приема звукового сигнала с максимальной амплитудой, то принимается решение «Да» и производится переход к пункту алгоритма «Шаг 21».

Если азимут нахождения источника звука не определен на основании приема звукового сигнала с максимальной амплитудой., то принимается решение «Нет» и производится переход к пункту алгоритма «Шаг 17».

Шаг 21. Включение микрофонов вертикальной группы с узкой ДН.

Данная группа микрофонов позволяет «точно» определять азимут расположения источника звука.

Шаг 22. Определение наличия звукового сигнала.

Определяется наличие на приеме акустического сигнала источника звука.

Шаг 23. Акустический сигнал в наличии?

Если осуществляется прием акустического сигнала микрофонами вертикальной группы с узкой ДН, то принимается решение «Да» и производится переход к пункту алгоритма «Шаг 24».

Если прием акустического сигнала микрофонами вертикальной группы не осуществляется, то принимается решение «Нет» и производится переход к пункту алгоритма «Шаг 22».

Шаг 24. Определение высоты (угла места) нахождения звукового сигнала.

Определяется высота (угол места) нахождения акустического сигнала источника звука при использовании вертикальной группы ААП.

Шаг 25. Высота (угол места) определена.

Определяется высота нахождения источника звука по углу поворота ААУ с микрофонами с узкой ДН на основании приема звукового сигнала с максимальной амплитудой.

Если высота (угол места) нахождения источника звука определен на основании приема звукового сигнала с максимальной амплитудой, то принимается решение «Да» и производится переход к пункту алгоритма «Шаг 26».

Если высота (угол места) нахождения источника звука не определена на основании приема звукового сигнала с максимальной амплитудой, то принимается решение «Нет» и производится переход к пункту алгоритма «Шаг 22».

Шаг 26. Определение расстояния до источника звука.

Определяется расстояние до источника звука при применении следующих методов для оценки дальности нахождения объекта:

– метод пеленгования (триангуляция) [9] применяется при использовании нескольких разнесенных в пространстве приемных систем и наличии перемещаемого объекта;

– амплитудный метод предполагает проведение оценки дальности нахождения объекта по затуханию уровня сигнала;

– метод спектрального анализа по определению доплеровского смещения объекта или наличию в составе звукового сигнала низкочастотных компонентов [10, 11];

– горизонтальный разностно-дальномерный метод.

Шаг 27. Расчет координат источника звука методом пеленгования.

Координаты местоположения источника звука осуществляется методом интеллектуального акустического пеленгования на основе полученных данных отдельных ААП, ААП в составе группы и ААП в составе нескольких групп. Если известны точные географические координаты разнесенных в пространстве приемных антенн, то координаты источника звука в горизонтальной плоскости определяются достаточно просто использованием триангуляционного метода.

Шаг 28. Расчет вероятной траектории движения источника звука.

Шаг 29. Передача информации о результатах произведенных расчетных операций в базу данных системы.

Осуществляется передача следующих данных в БД и на интерфейс оператора:

– текущие координаты местоположения звукового объекта;

– основные характеристики звукового объекта: скорость (линейная и угловая), высота полета, курс и др.;

– тип летательного аппарата;

– прогноз вероятной траектории движения летательного аппарата;

– оценка вероятной зоны поражения;

– уровень угрозы от летательного аппарата.

Шаг 30. Конец алгоритма.

Заключение

В данной работе была проведена комплексная исследовательская работа, направленная на обнаружение беспилотных авиационных аппаратов с помощью системы акустической локации. Основной целью исследования являлось выявление эффективности и ограничений методов акустической локации, а также оценка их влияния на точность идентификации БПЛА. В ходе выполнения данной работы была достигнута поставленная цель, заключающаяся в оценке эффективности

акустической локации для обнаружения беспилотных авиационных аппаратов. В результате анализа существующих методов акустической локации было выявлено, что использование звуковых микрофонов и массивов акустических датчиков может значительно улучшить точность обнаружения БПЛА, однако на эффективность работы системы обнаружения оказывают влияние такие факторы, как шум, ветер и рельеф местности. Перспективные разработки в данном направлении внесут существенный вклад в систему противодействия беспилотным летательным аппаратам.

Список литературы

1. Иванов, Н. И. Теория и практика борьбы с шумом : учебник / Н. И. Иванов. – М. : Логос, 2010. – 424 с.
2. Smith, J. Introduction to Acoustic Localization Techniques [Электронный ресурс] / J. Smith, L. Johnson // Proceedings of the International Conference on Drone Detection. – Режим доступа: http://www.icdd2025.org/papers/acoustic_localization
3. Кузнецов, А. А. Современные подходы к акустической локации объектов / А. А. Кузнецов, В. В. Сидоров // Журнал радиофизики и электроники. – 2023. – № 4. – С. 45–52.
4. Петрова, Н. Н. Акустические методы обнаружения беспилотников / Н. Н. Петрова, А. А. Смирнов // Вестник научных исследований. – 2024. – № 2. – С. 30–37.
5. Иванова, Е. В. Применение акустической локации для мониторинга беспилотных летательных аппаратов / Е. В. Иванова, Д. Д. Коваленко // Научный вестник. – 2025. – № 3. – С. 50–56.
6. Brown, T. Advances in Acoustic Detection Technologies for UAVs / T. Brown, R. Green // Journal of Drone Technology. – 2023. – Vol. 12, № 1. – P. 15–22.
7. Сидоренко, А. А. Сравнительный анализ методов акустической локации для обнаружения беспилотников / А. А. Сидоренко, И. И. Михайлов // Труды конференции по радиолокации и акустике. – 2023. – С. 78–85.
8. Кузнецова, Е. В. Влияние рельефа местности на акустическую идентификацию беспилотников / Е. В. Кузнецова, А. А. Соловьев // Журнал акустических исследований. – 2024. – № 1. – С. 15–22.
9. Орлов, А. П. Акустическая локация в условиях urban: влияние шума и рельефа / А. П. Орлов, Д. Д. Коваленко // Труды конференции по акустике и радиотехнике. – 2023. – С. 90–97.
10. Михайлов, И. И. Влияние погодных условий на эффективность акустической локации беспилотников / И. И. Михайлов, В. В. Сидоров // Вестник научных исследований. – 2024. – № 3. – С. 20–27.

Разработка алгоритма идентификации БПЛА системой акустической локации

Ю. А. Клименко, email: klm71165@mail.ru [1]
А. П. Преображенский, email: app@vivt.ru [1]
Я. Б. Кораблев, klm71165@mail.ru [2]
А. И. Родюков, email: klm71165@mail.ru [2]

[1] Воронежский институт высоких технологий
[2] ВУНЦ ВВС "Военно-воздушная академия имени Н.Е. Жуковского и Ю. А. Гагарина

Аннотация. *В данной работе рассматривается проблема разработки и применения средств противодействия беспилотным летательным аппаратам. Для этой цели предложена система акустического мониторинга воздушного пространства для выявления БПЛА, их сопровождения, определения точных координат и уничтожения средствами ПВО. Предлагаемы алгоритм относится к программному обеспечению средств акустической локации. На основе данного алгоритма создана компьютерная программа.*

Ключевые слова: *акустика, локация, мониторинг, система, воздушное пространство, алгоритм, обнаружение, летательный аппарат.*

Введение

В современных условиях развития беспилотных технологий создаются новые средства поражения воздушного, морского и наземного базирования [1]. Разрабатывается новая стратегия их применения. Особенно бурное развитие беспилотные технологии получили в военной сфере путем создания ударных средств огневого поражения людей, техники, объектов критической инфраструктуры в энергетике, промышленности, транспортной логистике.

Развитие технологий по созданию средств по противодействию беспилотным проектам заметно отстает от необходимости и не получило должного научного и технического прогресса в этой области, не считая, конечно, системы радиоэлектронной борьбы. К сожалению, сфера применения средств радиоэлектронной борьбы тоже носит ограниченный характер воздействия [2]. Создание высоко-интеллектуальных беспилотных средств позволяет решать вопросы

© Клименко Ю. А., Преображенский А. П., Кораблев Я. Б., Родюков А. И., 2026

управления беспилотными летательными аппаратами без использования радиосредств и оптоволокна [3].

Поэтому есть необходимость в развитии средств противодействия БПЛА, которые основаны на принципах простоты конструкции, надежности, экономичности и результативности применения.

Алгоритм идентификации БПЛА

Процесс распознавания (идентификации) источника звука с присущими только ему акустическими характеристиками невозможен без исследования информативности параметров спектральной модели данного объекта.

Исследование информативности параметров источника звука для автоматизированного распознавания [4].

Во многих научных работах этому вопросу выделяется существенное место. Предложено большое количество разработанных алгоритмов по выявлению и анализу звуковых колебаний акустического источника. Задача настоящего исследования состоит в создании модели источника звука для распознавания (идентификация) акустических объектов на основе выявления некоторой информативности ее параметров. Конечно же, такие подходы в настоящее время для проведения сравнительного анализа алгоритмов невозможно решать без применения глубоких нейронных сетей.

Разберемся в вопросе построения спектральной модели источника звука. Проследим основные этапы ее построения.

Получается, что исходная модель определена во временной области и описывает определенную структуру источника звука [5].

Необходимо использовать модель, которая будет обладать вычислительной простотой и, в тоже время, эффективностью ее применения.

Для этой цели воспользуемся для решения практических задач звуковой идентификации:

– определение звукового объекта как беспилотного летательного аппарата в отличии от других звуковых объектов;

– идентификация звукового объекта;

– обратная фильтрация звукового сигнала.

Сформулируем исходную спектральную модель звукового источника во временной области и описывающую форму звукового объекта. Данная исходная модель представляет собой функцию времени t на одном периоде основного тона звукового сигнала [6]. Она определяется соотношением:

$$W = \begin{cases} W_1 = A_1 \cdot \sin\left(\dfrac{\pi \cdot t}{2 \cdot T_2}\right), 0 \le t \le T_2, \\ W_2 = (A_1 + A_2) \cdot \cos\left(\dfrac{\pi \cdot (t - T_1)}{2 \cdot (T_2 - T_1)}\right) - A_2, T_1 < t \le T_2, \\ W_3 = -A_2 \cdot \left(\dfrac{T_3 - t}{T_3 - T_2}\right)^2, T_2 < t \le T_3, \\ 0, T_3 < t \le T_0. \end{cases} \quad (1)$$

Модель определяет форму звукового объекта для T_0 при частоте f основного тона. Форма модели звукового источника представлена на рис. 1.

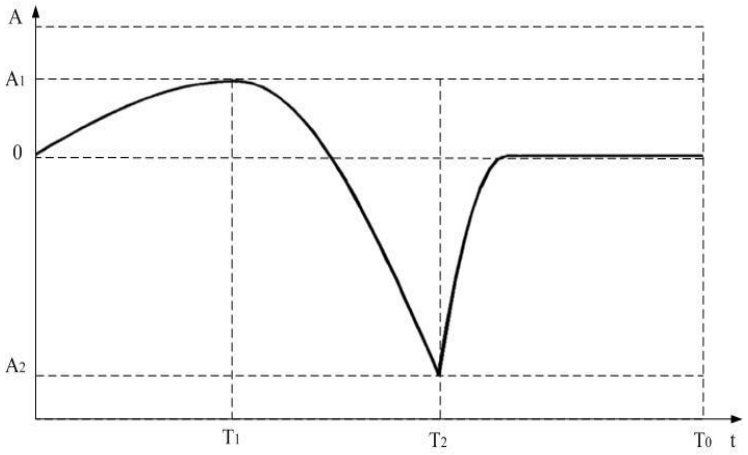


Рис. 1. Форма модели звукового источника

Определяем на графике следующие независимые параметры формы звукового источника согласно (1):

1) A_1 – максимальная амплитуда звукового импульса;

2) A_2 – минимальная амплитуда звукового импульса;

3) T_1 – момент времени при максимальной объемной скорости звукового импульса;

4) T_2 – момент времени при минимальной объемной скорости звукового импульса;

5) T_3 – момент времени при начальной фазе возникновения следующего звукового импульса;

6) T_0 – момент времени одного периода звукового импульса.

При рассмотрении (1) определяем, что A_1 является функцией соотношения независимых параметров A_1, T_1, T_2, T_3:

$$A_1 = \frac{A_2}{6T_2}\left[(6-3\pi)T_1 + (2\pi-6)T + \pi T_3\right] \quad (2)$$

Но при этом необходимо также предусмотреть автоматическую корректировку параметров T_1, T_2, T_3 для проведения аппроксимации звукового импульса, потому что велика вероятность искажения звукового сигнала помехами шума.

Для устранения данного явления необходимо модель (1)-(2) преобразовать в частотном диапазоне. Если функцию W представить в виде квазипериодической последовательности и провести дальнейшее преобразование путем разложения ее в ряд Фурье, то выводим значение n-ой комплексной амплитуды спектральной модели:

$$C_n = S_1 + S_2 + S_3 \quad (3)$$

где C_n – комплексная амплитуда спектральной модели (3).

Современные системы звукометрии, которые построены на основе глубоких нейронных сетей с высокой точностью могут определять звуковой объект по акустическим выборкам (сигнатурам), находящимся в БД, и зафиксированных в различных внешних условиях [7].

На рис. 2 приведен алгоритм проведения идентификации беспилотного летательного аппарата.

Поясним работу данного алгоритма:

Шаг 1. Начало алгоритма.

Запуск процесса работы алгоритма осуществляется после обнаружения источника звука и получения данных акустического объекта.

Шаг 2. Ввод данных.

Осуществляется поступление информации из БД и других внешних носителей для осуществления обработки получаемых акустических сигналов беспилотных летательных объектов

Шаг 3. Получение первичного акустического сигнала.

Первичный акустический сигнал поступает в обработку с акустических антенных устройств и систем преобразования звуковых колебаний.

Шаг 4. Предварительная обработка акустического сигнала.

Предварительная обработка предполагает выполнение следующих процедур:

– фильтрация низкочастотных шумов;

– подавление импульсных помех;

– нормализация амплитуды акустического сигнала.

Шаг 5. Проведение спектрального анализа акустического сигнала.

Процедура проведения спектрального анализа включает в себя:

– расчет спектрограммы акустического сигнала;

– выделение характерных частотных компонент;

– анализ модуляционных характеристик.

Шаг 6. Расчет спектрограммы акустического сигнала.

Расчет спектрограммы производится на определении следующих параметров акустического сигнала: время, частота, амплитуда.

Шаг 7. Выделение характерных частотных компонент.

При исследовании спектрограммы возможно выделить следующие частотные компоненты:

– основная частота вращения винтов летательного аппарата;

– основные частоты звукового излучения двигателя внутреннего сгорания (ДВС) летательного аппарата;

– гармоники основной частоты вращения винтов и работы ДВС;

– компоненты шумового сопровождения.

Шаг 8. Анализ модуляционных характеристик акустического сигнала:

При данном анализе определяются следующие модуляционные характеристики:

– частота модуляции (биения);

– глубина модуляции.

Шаг 9. Выделение звуковых сигнатур акустического сигнала.

Шаг 10. Формирование акустического паспорта.

При составлении акустического паспорта определяются:

– выделение уникальных звуковых сигнатур;

– оставление цифрового отпечатка.

Шаг 11. Выделение уникальных звуковых сигнатур.

– спектральные признаки (мел-кепстральные коэффициенты, центроиды, полосы);

– временными признаками (огибающая, автокорреляция);

– статистические признаки (дисперсии, эксцесс, асимметрия);

– матрица спектральных характеристик;

– временные метки изменений акустического сигнала.

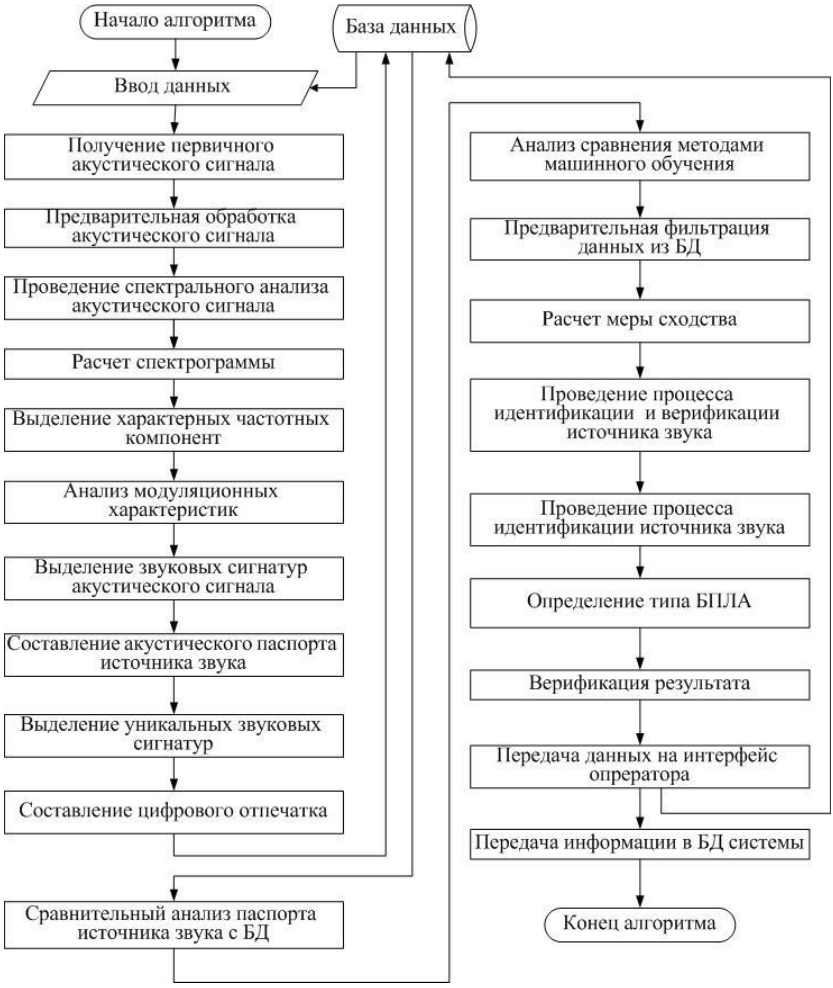


Рис. 2. Алгоритм идентификации беспилотного летательного аппарата

Шаг 12. Сравнительный анализ паспорта источника звука с БД.

При проведение спектрального анализа предполагается применение средств интеллектуального сравнения и классификации для выполнения следующих процедур:

– предварительная фильтрация данных из БД;

– анализ сравнения методами машинного обучения;

– расчет меры сходства.

Шаг 13. Анализ сравнения методами машинного обучения.

При проведении сравнительного анализа с помощью возможностей программ машинного обучения будут использоваться следующие методы:

– метод k-ближайших соседей (k-NN) для быстрого сопоставления данных;

– метод опорных векторов (SVM) для определения точной классификации;

– нейросетевые подходы (сверточные сети) для наиболее сложных случаев при проведении сравнительного анализа.

Шаг 14. Предварительная фильтрация данных из БД.

При проведении предварительной фильтрации данных производится:

– отбор подходящих вариантов моделей по частотному диапазону;

– исключение заведомо неподходящих моделей

Шаг 15. Расчет меры сходства (расстояния).

При проведении расчета меры сходства производится расчет следующих параметров:

– евклидово расстояние между векторами признаков;

– косинусная мера сходства;

– вероятность соответствия.

Шаг 16. Проведение процесса идентификации и верификации источника звука.

При проведении процесса идентификации и верификации осуществляются следующие процедуры:

– проведение процесса идентификации;

– определение типа БПЛА;

– верификация результата.

Шаг 17. Проведение процесса идентификации.

При проведении процесса идентификации производятся следующие действия:

– выбор наиболее вероятной модели;

– проверка достоверности и пороговых значений.

Шаг 18. Определение типа БПЛА.

При определении типа беспилотного средства предоставляются

следующие параметры звукового объекта с указанием степени достоверности:

– производитель летательного аппарата;

– модель летательного аппарата;

– уровень достоверности идентификации (в процентах, например «85%»);

– дополнительные характеристики, определяющие идентификацию данного летательного аппарата (например: количество винтов, тип двигателя и др.).

Шаг 19. Верификация результата.

При верификации осуществляются следующие действия:

– перекрестная проверка по нескольким временным сегментам;

– анализ устойчивости признаков во времени;

– учет воздействия внешних факторов состояния атмосферы.

Шаг 20. Передача данных на интерфейс оператора.

Помимо передачи данных в БД, необходимо оперативно осуществлять передачу полученной информации (координаты цели, метки цели, вероятная траектория движения) оператору на средства отображения (электронная карта, экран планшета и др.) сигналы оповещения (звуковые, световые, цветовые и др.).

Шаг 21. Передача информации в БД системы.

Полученная расчетная и корректирующая информация передаются в БД для хранения и дальнейшего использования.

Шаг 22. Конец работы алгоритма.

Заключение

На основании работы данного алгоритма при проведении исследования спектрограмм различных беспилотных летательных аппаратов можно сделать вывод о том, что производить обнаружение подлета беспилотного средства необходимо по низкочастотной части спектра, поскольку там сосредоточена основная энергия спектра звука двигателя и вращающихся винтов. Микрофоны должны, соответственно, быть рассчитаны на параметры работы в зоне приема звукового сигнала близкой к инфразвуку, т. к. звук на этих частотах испытывает минимальное затухание и может распространяться на большие расстояния.

В данной работе была проведена комплексная исследовательская работа, направленная на обнаружение и идентификацию беспилотных авиационных аппаратов с помощью системы акустической локации. Основной целью исследования являлось выявление эффективности и ограничений методов акустической локации, а также оценка их влияния на точность идентификации БПЛА. В ходе выполнения данной работы была достигнута поставленная цель, заключающаяся в оценке эффективности акустической локации для обнаружения беспилотных авиационных аппаратов. Перспективные разработки в данном

направлении внесут существенный вклад в систему противодействия беспилотным летательным аппаратам.

Список литературы

1. Клименко, Ю. А. Разработка пассивной системы определения позиционирования беспилотного летательного аппарата / Ю. А. Клименко, А. П. Преображенский // Современные проблемы радиоэлектроники и телекоммуникаций. Севастопольский государственный университет. – 2024. – № 7. – С. 32
2. Smith, J. Introduction to Acoustic Localization Techniques [Электронный ресурс] / J. Smith, L. Johnson // Proceedings of the International Conference on Drone Detection. – Режим доступа: http://www.icdd2025.org/papers/acoustic_localization
3. Кузнецов, А. А. Современные подходы к акустической локации объектов / А. А. Кузнецов, В. В. Сидоров // Журнал радиофизики и электроники. – 2023. – № 4. – С. 45–52.
4. Петрова, Н. Н. Акустические методы обнаружения беспилотников / Н. Н. Петрова, А. А. Смирнов // Вестник научных исследований. – 2024. – № 2. – С. 30–37.
5. Иванова, Е. В. Применение акустической локации для мониторинга беспилотных летательных аппаратов / Е. В. Иванова, Д. Д. Коваленко // Научный вестник. – 2025. – № 3. – С. 50–56.
6. Brown, T. Advances in Acoustic Detection Technologies for UAVs / T. Brown, R. Green // Journal of Drone Technology. – 2023. – Vol. 12, № 1. – P. 15–22.
7. Клименко, Ю. А. Разработка пассивной системы защиты от БПЛА / Ю. А. Клименко, А. П. Преображенский // Современные проблемы радиоэлектроники и телекоммуникаций. Севастопольский государственный университет. – 2024. – № 7. – С. 33.

Анализ уязвимостей протокола Wi Fi 6 (IEEE 802.11ax) к пассивному перехвату и методы защиты на канальном уровне

И.В. Климкин, email: nauchnajarota@yandex.ru
Г.А. Лахно, email: nauchnajarota@yandex.ru
Н.С. Уткин, email: nauchnajarota@yandex.ru

Межвидовой центр подготовки и боевого применения войск РЭБ
(учебный и испытательный)

Аннотация. *В статье проводится комплексный анализ уязвимостей протокола Wi Fi 6 (IEEE 802.11ax) к пассивному перехвату трафика. Рассматриваются особенности архитектуры протокола, включая технологию OFDMA и MU MIMO, создающие новые векторы атак. Предложены математические модели оценки вероятности перехвата и эффективности методов защиты на канальном уровне. Проведенное моделирование показало, что применение адаптивного шифрования с динамической сменой ключей и внедрение протоколов аутентификации на основе доверенных платформ снижают вероятность успешного пассивного перехвата на 91 %. Результаты исследования имеют практическую значимость для разработки защищенных беспроводных сетей следующего поколения.*

Ключевые слова: *Wi Fi 6, IEEE 802.11ax, пассивный перехват, уязвимости, защита на канальном уровне, OFDMA, криптографические методы, доверенная платформа.*

Введение

Современные беспроводные сети сталкиваются с возрастающими угрозами информационной безопасности, особенно в условиях массового внедрения технологий следующего поколения. Протокол Wi Fi 6 (IEEE 802.11ax), обеспечивающий высокую пропускную способность, вводит новые механизмы, которые могут создавать дополнительные уязвимости к пассивному перехвату трафика [1].

Актуальность исследования обусловлена несколькими факторами. Во-первых, согласно данным Cisco Annual Internet Report, к 2025 году количество устройств Wi Fi 6 превысит 4.7 миллиарда, что делает их привлекательной целью для злоумышленников. Во-вторых, особенности архитектуры Wi Fi 6, такие как ортогональное частотное разделение каналов (OFDMA) и многопользовательский MIMO (MU MIMO),

© Климкин И. В., Лахно Г. А., Уткин Н. С., 2026

создают новые векторы атак, недостаточно изученные в сфере информационной безопасности [2, 3].

Целью работы является анализ уязвимостей протокола Wi Fi 6 к пассивному перехвату и предложение более эффективного метода защиты на канальном уровне. Для достижения поставленной цели решаются следующие задачи: изучение архитектурных особенностей Wi Fi 6; идентификация уязвимостей к пассивному перехвату; разработка математических моделей оценки безопасности; предложение и оценка эффективности методов защиты.

1. Архитектурные особенности Wi Fi 6 и их уязвимости

Протокол IEEE 802.11ax вводит ряд инновационных технологий, повышающих эффективность работы беспроводных сетей. Ключевой технологией является OFDMA (Orthogonal Frequency Division Multiple Access), которая разделяет канал на множество поднесущих, распределяемых между пользователями. Математическая модель передачи с использованием OFDMA описывается следующим выражением:

$$s(t) = \sum X_k \cdot g(t - kT_s) \cdot e^{(j2\pi f_k t)}$$

где $s(t)$ –– общий сигнал во времени,X_k – данные (биты) передаваемые по k-тому каналу, $g(t - kT_s)$ – формирующий импульс для k-того подканала T_s – длительность символа), $e^{(j2\pi f_k t)}$ –– комплексная экспонента, создающая несущую частоту f_k для k -го подканала.

Несмотря на преимущества, технология OFDMA создает уязвимость к пассивному перехвату. Поскольку разные пользователи используют разные поднесущие одновременно, злоумышленник может перехватывать трафик конкретного пользователя, фокусируясь на соответствующих частотах. Вероятность успешного перехвата можно оценить как:

$$P_{\text{int ercept}} = 1 - e^{(-\lambda \cdot T_{capture})}$$

где $P_{intercept}$ –– вероятность перехвата (от 0 до 1, где 1 = 100%), λ – интенсивность атаки (сколько пакетов пытается перехватить злоумышленник в секунду), $T_{capture}$ – время, в течение которого злоумышленник пытается захватить сигнал.

Другой важной технологией является MU-MIMO (Multi-User Multiple Input Multiple Output), которая позволяет одновременную передачу данных нескольким пользователям. Однако это создает уязвимость к перехвату через пространственное разделение. Если злоумышленник имеет достаточное количество антенн, он может

разделить потоки данных и перехватить информацию конкретного пользователя.

На рис. 1 представлена схема пассивного перехвата в сети Wi-Fi 6 с использованием OFDMA.

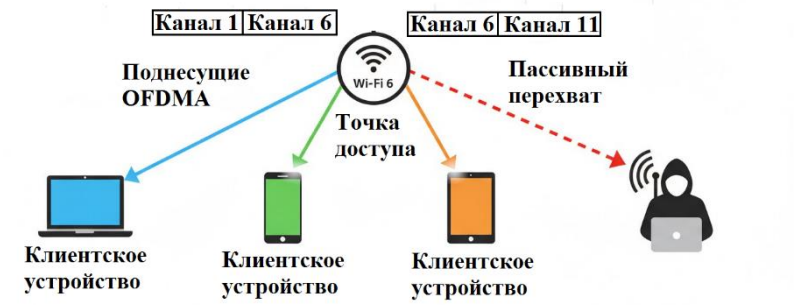


Рис. 1. Схема пассивного перехвата в сети Wi-Fi 6 с OFDMA

2. Методы защиты на канальном уровне

Для защиты от пассивного перехвата в сетях Wi-Fi 6 предлагается комплексный подход на канальном уровне. Первый метод – адаптивное шифрование с динамической сменой ключей. В отличие от статических ключей WPA3, предложенный метод использует динамическую генерацию ключей на основе временных меток и состояния канала:

$$K_{dynamic} = H(TS \oplus RSSI \oplus MAC_addr)$$

где $K_{dyna\ mic}$ – генерируемый динамический ключ шифрования, $H -$ – криптографическая хеш-функция (например, SHA-256), TS – текущая временная метка, $RSSI -$ – уровень сигнала в dBm (показывает качество связи), MAC_{addr} – уникальный MAC-адрес устройства, $\oplus$ – операция XOR (исключающее ИЛИ).

Эффективность метода оценивается коэффициентом снижения вероятности перехвата:

$$R_{reduction} = [\frac{P_{base} - P_{protected}}{P_{base}}] \cdot 100\%$$

где $R_{reduction}$ – снижение риска в процентах (%), P_{base} – вероятность перехвата без защиты, $P_{protected} -$ – вероятность перехвата с защитой.

Второй метод защиты – аутентификация на основе доверенных платформ (TPM) [4, 5]. Каждое устройство содержит аппаратный модуль TPM, который генерирует уникальные ключи и подтверждает

подлинность устройства перед передачей данных. Вероятность успешной аутентификации определяется как:

$$P_A = \left(\frac{N_{trusted}}{N_{total}} \right)^{e(-\alpha \cdot d)}$$

где P_A – вероятность успешной аутентификации, $N_{trusted} -$ –количество доверенных (зарегистрированных) устройств, $N_{total} -$ – общее количество устройств в сети, $e^{(-\alpha \cdot d)} -$ – экспоненциальное ослабление сигнала с расстоянием, α – коэффициент затухания сигнала (зависит от среды), $d -$ – расстояние от устройства до точки доступа.

На рис. 2 показана эффективность предложенных методов защиты в сравнении с базовыми решениями.

3. Экспериментальные результаты и обсуждение

Для оценки эффективности предложенных методов проведено моделирование в среде MATLAB с использованием пакета Communications Toolbox. Тестовая сеть включала 1 точку доступа Wi-Fi 6 и 7 клиентских устройств, симулирующих различную активность.

Результаты экспериментов представлены в таблице. Базовая конфигурация (WPA3) показала вероятность перехвата 0,45 при интенсивности атаки 10 пакетов/сек. Применение адаптивного шифрования снизило этот показатель до 0,12, а комбинация с TPM-аутентификацией – до 0,04.

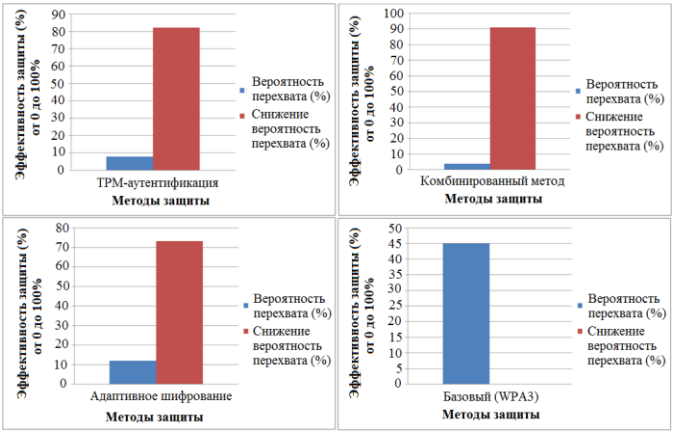


Рис. 2. Сравнение эффективности методов защиты от пассивного перехвата

Таблица

Результаты экспериментов по защите от пассивного перехвата

Метод защиты	Вероятность перехвата	Снижение, %	Накладные расходы, мс
Базовый (WPA3)	0,45	–	2,1
Адаптивное шифрование	0,12	73,3	3,8
TPM-аутентификация	0,08	82,2	5,2
Комбинированный метод	0,04	91,1	6,5

Как видно из таблицы, комбинированный метод обеспечивает наибольшую защиту, снижая вероятность перехвата на 91.1%, при этом накладные расходы остаются в допустимых пределах (6,5 мс).

График на рис. 3 демонстрирует зависимость вероятности перехвата от интенсивности атаки для разных методов защиты. Комбинированный метод сохраняет низкие показатели даже при высокой интенсивности атаки.

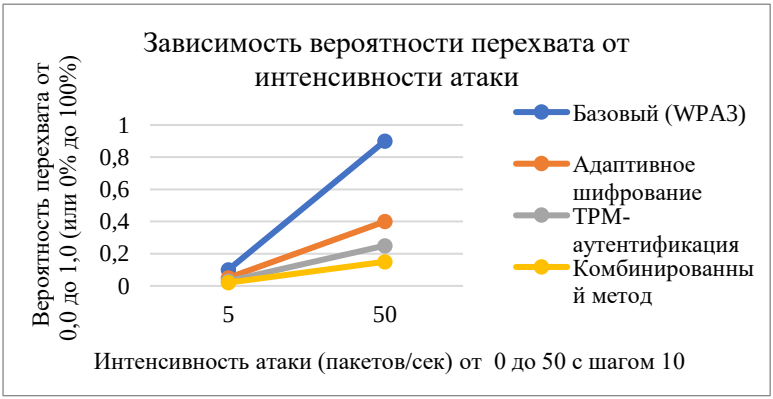


Рис. 3. Зависимость вероятности перехвата от интенсивности атаки

Заключение

Проведенное исследование показало, что протокол Wi-Fi 6 имеет специфические уязвимости к пассивному перехвату, связанные с особенностями OFDMA и MU-MIMO технологий. Предложенный комбинированный метод защиты, включающий адаптивное шифрование с динамической сменой ключей и TPM-аутентификацию, снижает

вероятность успешного перехвата на 91,1% при приемлемых накладных расходах.

Практическая значимость работы заключается в возможности внедрения разработанных методов в существующие системы беспроводной связи для повышения их информационной безопасности. Перспективы дальнейших исследований включают оптимизацию алгоритмов генерации ключей и интеграцию методов защиты с технологиями искусственного интеллекта для адаптивного реагирования на угрозы.

Список литературы

1. Киселев, В. И. Методы защиты беспроводных сетей от пассивного перехвата / В. И. Киселев, А. С. Петров // Информационные технологии и безопасность. – 2023. – Т. 18, № 2. – С. 89–102.

2. ГОСТ Р 57580.1-2017. Защита информации. Алгоритмы шифрования. – Введ. 2018-07-01. – М.: Стандартинформ, 2017. – 24 с.

3. Андрианов, В. И. Экспертные системы в области информационной безопасности / В. И. Андрианов, Г. Г. Романов, С. И. Штеренберг //Актуальные проблемы инфотелекоммуникаций в науке и образовании. – 2015. – С. 193–197.

4. Исследование способов удаленного перехвата трафика в корпоративных сетях / М. М. Ковцур [и др.] // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия. – 2021. – Т. 1. – С. 68–75.

5. Оценка рисков и угроз безопасности в среде «Умный дом» / А. М. Гельфанд [и др.] // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2020) : сб. научных статей IX Международной научно-технической и научно-методической конференция. – Санкт-Петербург : 2020. – С. 316–321.

Расчет тангенциальных компонент поля в апертуре антенны по результатам зондирования ближней зоны

А. А. Кононов, email: kononov@phys.vsu.ru
Г. К. Усков, email: uskov@phys.vsu.ru

Воронежский государственный университет

Аннотация. *В работе обсуждается метод восстановления тангенциальных компонент распределения поля на апертурной поверхности, расположенной вблизи излучающего устройства, путем обработки экспериментальных данных, получаемых с помощью малогабаритного приемного зонда, который последовательно перемещается по дискретной сетке точек в ближней зоне. На основе построенной ранее теоретической модели, связывающей сигнал на выходе подобного измерительного устройства с величиной поля в точке его расположения, получены аналитические выражения, позволяющие пересчитывать коэффициент передачи системы «антенна — зонд» в комплексную амплитуду определяемой тангенциальной компоненты вектора электрической напряженности. Корректность алгоритма подтверждена рассмотрением практического примера.*

Ключевые слова: *измерение электромагнитного поля, ближняя зона, апертурное распределение, малогабаритный зонд, диаграмма направленности.*

Введение

Экспериментальное определение электромагнитного поля в ближней зоне излучающего устройства является неотъемлемой частью широкого круга радиотехнических исследований, будь то проверка данных теоретического прогнозирования путем сопоставления с результатами непосредственного опыта или же проведение измерений с целью первичного сбора информации для последующей обработки и применения в построении моделей.

В то же время, среди практических способов измерения векторов напряженностей поля вблизи антенны — при всем их разнообразии — наблюдается ряд обстоятельств, приводящих, по всей видимости, к отсутствию какой-либо одной общепринятой методики [1–2]. Например, некоторые из методов базируются на пересчете «сразу всей» совокупности результатов измерений (производящихся несколько поодаль от непосредственной апертурной поверхности) в столь же

© Кононов А. А., Усков Г. К., 2026

«целостное» распределение токов на апертуре [3–9]; при этом нет возможности точечного исследования уже излученного поля (кроме того, ввиду неустойчивости задачи может возникать необходимость применения процедур регуляризации [6, 9]).

Альтернативный класс идей основан на использовании свойств взаимности приемных и передающих антенных устройств с целью связать принятый экспериментальным оборудованием сигнал и измеряемое поле в точке его расположения. Традиционный подход, по-видимому, заключается в представлении этого поля в виде суперпозиции плоских волн [2, 10–12], для каждой из которых затем может быть применена та или иная модель приема на антенну с известными свойствами [13–17]. Нужно заметить, однако, что подобное разложение по плоским волнам является математической абстракцией, которая не только снова охватывает «сразу всю» картину излученного поля, но и, как известно [2, 12], не может быть реализована в монохроматических пределах; в частности, для организации расчетов недостаточно было бы знать свойства измерительного зонда лишь на одной рабочей частоте.

В данной работе приводятся соображения по осуществлению пересчета сигнала, поступившего от принимающего устройства, в локальное значение измеряемого поля. Предшествующие исследования позволили построить аналитическую модель для вычисления этого сигнала при известном поле в предположении, что вышеупомянутый зонд имеет достаточно малые геометрические размеры — чтобы в его пределах поле могло считаться приблизительно постоянным [18].

Следует также добавить, что нас интересует векторная комплексная амплитуда поля, так как именно она включает в себя наиболее полные сведения о его структуре. Единственное упрощение, которое можно сделать, не умаляя общности — ограничиться исследованием лишь тангенциальных компонент, поскольку они содержат всю информацию, необходимую для расчета трехмерного вектора излученного поля в любой точке перед апертурой [19]. При этом указанное упрощение будет играть весьма принципиальную роль для осуществления расчетов в рамках предлагаемого метода, как это станет видно далее из основного текста работы.

1. Теоретические соотношения. Описание методики

Напомним для начала основные результаты, основанные на принципе взаимности и его следствиях [13–15]. Если приемный зонд по своим размерам достаточно мал для того, чтобы в непосредственной близости к нему внешнее поле могло считаться неизменным (как это было бы и при приеме плоской волны из дальней зоны), то для тока в подключенной к зонду нагрузке справедливо выражение [18]

$$\dot{I}_{rx} = \frac{i\lambda}{\sqrt{\pi}} \cdot \sqrt{\eta} \cdot \sqrt{\frac{\operatorname{Re} \dot{Z}_A}{Z_s}} \cdot \frac{\dot{\mathbf{E}}_{\mathbf{rx}}^{\mathbf{T}} \cdot \dot{\mathbf{D}}(\Omega)}{\dot{Z}_A + \dot{Z}_н}, \tag{1}$$

где λ — рабочая длина волны;

η — КПД зонда по излучению;

$\dot{Z}_A$ — входное комплексное сопротивление зонда в режиме передачи;

$Z_s \approx 120\pi$ *Ом* — волновое сопротивление свободного пространства;

$\dot{Z}_н$ — комплексное сопротивление нагрузки (по умолчанию 50 *Ом*);

$\dot{\mathbf{E}}_{\mathbf{rx}}$ — напряжённость принимаемого электрического поля;

$\dot{\mathbf{D}}$ — векторная комплексная диаграмма направленности (ДН) зонда;

$\Omega = (\theta; \varphi)$ — совокупность угловых переменных сферической системы координат, определяющая локальное направление распространения волны.

Сразу заметим: в (1) входят лишь самые базовые характеристики, которые предполагаются относительно легко определяемыми для любого антенного устройства. Напомним, что относительно ДН нами подразумеваются следующие соглашения [18, 20]: она должна быть нормирована по амплитуде в соответствии с тождеством

$$\oint_{\Omega} \left| \dot{\mathbf{D}}(\Omega) \right|^2 d\Omega = 4\pi, \tag{2}$$

а неоднозначность фазы ликвидируется выполнением условия, связывающего излучаемое антенной в режиме передачи поле $\dot{\mathbf{E}}_{\mathbf{tx}}$ с комплексной амплитудой полного входного тока $\dot{I}_{tx}$:

$$\dot{\mathbf{E}}_{\mathbf{tx}}(R, \Omega) = \dot{I}_{tx} \cdot \sqrt{Z_s \operatorname{Re} \dot{Z}_A} \cdot \sqrt{\frac{\eta}{4\pi}} \cdot \dot{\mathbf{D}}(\Omega) \cdot \frac{e^{-ikR}}{R}, \tag{3}$$

где R — расстояние до точки наблюдения;

$k = 2\pi/\lambda$ — волновое число).

Вводя комплексную действующую длину антенны

$$\dot{\mathbf{h}}(\Omega) = \frac{i\lambda}{\sqrt{\pi}} \cdot \sqrt{\eta} \cdot \sqrt{\frac{\operatorname{Re} \dot{Z}_A}{Z_s}} \cdot \dot{\mathbf{D}}(\Omega), \tag{4}$$

перепишем (1) в виде

$$\dot{I}_{rx} = \frac{\dot{\mathbf{E}}_{\mathbf{rx}}^{\mathbf{T}} \cdot \dot{\mathbf{h}}(\Omega)}{\dot{Z}_A + \dot{Z}_н}; \tag{5}$$

следовательно, по $\dot{I}_{rx}$ можно определить составляющую искомого поля $\dot{\mathbf{E}}_{\mathbf{rx}}$ вдоль вектора $\dot{\mathbf{h}}$.

Например, рассмотрим антенну в виде прямоугольного металлического рупора (рис. 1). Поле излучения такой антенны обладает преимущественно линейной поляризацией, значит, для его исследования можно использовать малогабаритный зонд в виде линейного электрического диполя (на рис. 1 раскрыв рупора параллелен плоскости xOy, поле поляризовано вертикально, то есть вдоль оси Oy; кроме того изображены векторы локальных направлений распространения волн).

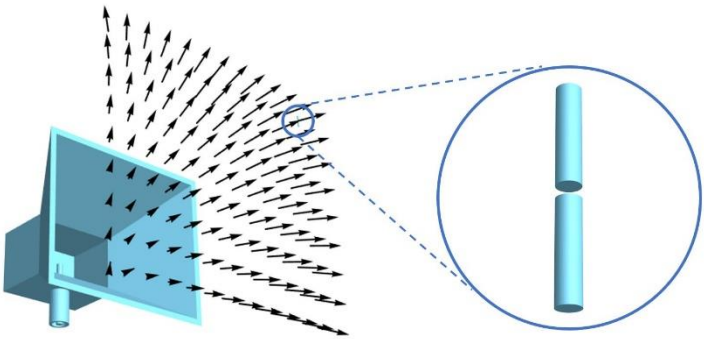


Рис. 1. Модель рупорной антенны и малогабаритного зонда

Для такой конфигурации из (5) следует рабочая формула:

$$\dot{I}_{rx} = \frac{\dot{E}_y \cdot \dot{h}_y(\Omega)}{\dot{Z}_A + \dot{Z}_н} \Rightarrow \dot{E}_y = \dot{I}_{rx} \cdot \frac{\dot{Z}_A + \dot{Z}_н}{\dot{h}_y(\Omega)}; \qquad (6)$$

остается только определиться с выбором выходящего в нее локального направления распространения волн, однако для плоских апертур его во многих случаях допустимо считать совпадающим с нормалью к апертурной поверхности (так что в данном случае компонента $\dot{h}_y$ должна вычисляться для направления прямо на раскрыв рупора).

На практике при получении опытных данных, как правило, непосредственно считываются замеры коэффициента передачи полученного «канала связи» (от антенны к зонду), т. е. комплексной величины отношения нормированных напряжений принятого сигнала в нагрузке зонда и падающего сигнала в линии подключения антенны. В

предположении, что оба устройства подключены к линиям передачи с вещественным волновым сопротивлением $\dot{Z}_н = 50 Ом$, получаем

$$\dot{S}_{11} = \frac{\dot{U}_{rx}}{\dot{U}_{tx}^{(i)}}, \tag{7}$$

где $\dot{U}_{tx}^{(i)}$ — напряжение падающей волны на входе антенны;

$\dot{U}_{rx}$ — напряжение на выходе приемного зонда:

$$\dot{U}_{rx} = -\dot{Z}_н \dot{I}_{rx}; \tag{8}$$

(знак минус возникает вследствие того, что принятый сигнал распространяется в соответствующей линии «от» зонда, т. е. в обратном — по отношению к определению $\dot{Z}_A$ — направлении [15, 21]).

Величина $\dot{U}_{tx}^{(i)}$, как обычно, выражается через полное напряжение $\dot{U}_{tx}$ и ток $\dot{I}_{tx}$ на входе исследуемой передающей антенны [15, 21]:

$$\dot{U}_{tx}^{(i)} = \frac{\dot{U}_{tx} + Z_н \dot{I}_{tx}}{2}; \tag{9}$$

соотношения (7)–(9) дополняют формулу (6), адаптируя ее для более удобного практического применения.

2. Пример определения тангенциальной компоненты

Для проверки качества функционирования вышеописанного метода вернемся к ситуации, изображенной на рис. 1. Численные значения геометрических и прочих параметров приведены в работе [18]. В качестве апертуры был взят участок плоскости, расположенной параллельно раскрыву рупора на расстоянии $45 мм$ от него и ограниченный до прямоугольного участка $90 \times 80 мм$ по осям x и y соответственно. Дипольный измерительный зонд последовательно перемещался с равномерным дискретным шагом по каждой оси (область подразделялась на 10 точек по горизонтали и 9 точек по вертикали), в итоге принимая каждое из 90 промежуточных положений; данные о зависимости величины S_{11} от координат точки его размещения были затем интерполированы для получения возможности отобразить их в виде более гладкой поверхности (рис. 2).

Далее на основе этого массива было восстановлено распределение вертикальной компоненты поля на апертуре (с использованием полученных в разделе 1 расчетных соотношений). Соответствующий график представлен на рис. 3; как можно видеть, он достаточно существенно отличается от предыдущего.

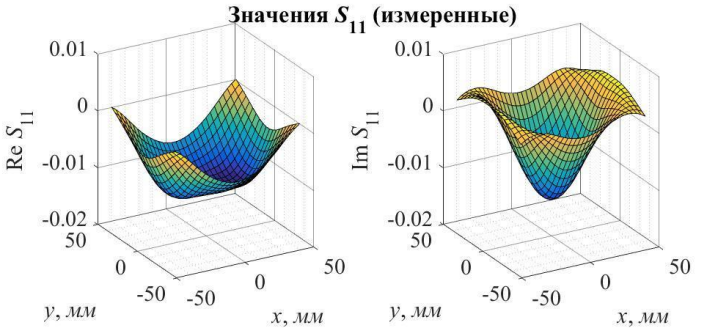


Рис. 2. Данные измерений, полученные моделированием

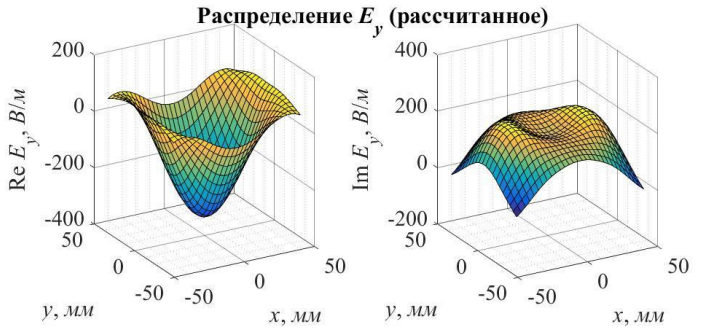


Рис. 3. Результат восстановления тангенциальной компоненты

Количественное сравнение рассчитанного здесь распределения $\dot{E}_y$ с данными, выгруженными непосредственно из модели в CST, дало следующие показатели. Среднеквадратичная ошибка восстановления вертикальной компоненты поля составила $\approx 7{,}935\ В/м$, максимальная же оказалась равной $\approx 16{,}022\ В/м$. С учетом того что минимальное и максимальное значения искомой составляющей на рассмотренном участке апертуры равны $\approx 51{,}9\ В/м$ и $\approx 354{,}4\ В/м$ соответственно, это можно считать весьма хорошим результатом. Наибольшие отклонения, как показало исследования, наблюдаются вблизи границ исследованного фрагмента апертурной поверхности.

Наконец, среднее значение относительной ошибки составило $\approx 4{,}2\%$; ввиду этого для «истинного» распределения поля отдельный график не построен, так как его отличия от поверхностей, показанных на рис. 3, были бы не вполне различимыми визуально.

Заключение

В работе представлен способ обработки экспериментальных данных, полученных с помощью перемещаемого в ближней зоне апертурной антенны малогабаритного измерительного зонда, с целью восстановления распределения величины излучаемого поля. Выведены все необходимые теоретические соотношения, для применения которых нужно иметь в распоряжении лишь наиболее фундаментальные характеристики приемного устройства — КПД, входное сопротивление и диаграмму направленности (определяемую, что особенно удобно, для дальней зоны и режима передачи). Еще одно преимущество изложенной методики — в ее локальном характере: для определения поля в одной точке не требуются данные об измерениях, проведенных в других частях апертуры.

Рассмотрен частный пример непосредственного восстановления распределения вертикальной составляющей поля излучения в ближней зоне рупорной антенны с линейной поляризацией по результатам моделирования значений комплексного коэффициента передачи от входа антенны к выходу малогабаритного зонда. Показано, что предлагаемый метод позволяет осуществлять расчет искомой компоненты с хорошей точностью (в то время как сам коэффициент передачи, до указанной обработки, может иметь качественно отличающееся распределение).

Благодарности

Исследование выполнено за счет гранта Российского научного фонда № 24-19-00891, https://rscf.ru/project/24-19-00891/.

Для выполнения численных расчетов в ходе верификации представленных в работе теоретических соотношений использовалось оборудование учебно-научного дизайн-центра проектирования радиоэлектронных систем СВЧ, терагерцового и оптического диапазонов на отечественной электронной компонентной базе ФГБОУ ВО «ВГУ» в рамках реализации федерального проекта «Подготовка кадров и научного фундамента для электронной промышленности».

Список литературы

1. Baird, R. C. A Brief History of Near-Field Measurements of Antennas at the National Bureau of Standards / R. C. Baird, A. C. Newell, C. F. Stubenrauch // IEEE Transactions on Antennas and Propagation. –1988. – Vol. 36, no. 6. – P. 727–733.

2. IEEE Recommended Practice for Near-Field Antenna Measurements. – Approved 20 August 2012. – New York : The Institute of Electrical and Electronics Engineers, Inc., 2012. – 102 p.

3. Методы и методики экспериментального исследования радиотехнических характеристик антенн наземных комплексов управления / А. В. Калинин [и др.] // Ракетно-космическое приборостроение и информационные системы. – 2016. – Т. 3, № 3. – С. 54–62.

4. Данилов, И. Ю. Диагностика апертурных распределений антенн путем измерений в зоне ближнего излученного поля / И. Ю. Данилов, Ю. Е. Седельников // Журнал радиоэлектроники. – 2016. – № 1. – С. 1–26.

5. Седельников, Ю. Е. Диагностика амплитудно-фазового распределения с использованием метода фокусировки при измерении в ближней зоне излучения / Ю. Е. Седельников, Н. А. Тутьяров, Т. Р. Шагвалиев // Вестник ПГТУ. –2020. – № 4(48). – С. 40–47.

6. Седельников, Ю. Е. Диагностика апертурных антенн по измеренным распределениям электромагнитных полей в зоне ближнего излученного поля / Ю. Е. Седельников, И. Ю. Данилов // Сб. материалов Международных научно-технических конференций «Проблемы техники и технологий телекоммуникаций (ПТиТТ-2014); Оптические технологии в телекоммуникациях (ОТТ-2014)» (Казань, 18–21 ноября 2014 г.). – 2014. – Т. 2. – С. 102–104.

7. Данилов, И. Ю. Измерение апертурных распределений антенн модифицированным методом фокусировки / И. Ю. Данилов // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. – 2016. – № 3. – С. 13–19.

8. Седельников, Ю. Е. Метод фокусировки при измерении амплитудно-фазового распределения в ближней зоне излучения / Ю. Е. Седельников, Т. Р. Шагвалиев // Вестник ПГТУ. Сер.: Радиотехнические и инфокоммуникационные системы. –2020. – № 1(45). – С. 30–36.

9. Седельников, Ю. Е. Измерение апертурных распределений антенн в зоне ближнего излученного поля / Ю. Е. Седельников, Т. Р. Шагвалиев // Сб. тр. XXXVII Междунар. научно-технической конф. «Радиолокация, навигация, связь» (Воронеж, 28–30 сентября 2021 г.). – Воронеж, 2021. – Т. 3. – С. 213–218.

10. Kerns, D. M. Theory of Diffraction in Microwave Interferometry / D. M. Kerns, E. S. Dayhoff // Journal of Research of the National Bureau of Standards. –1960. – Vol. 64B, no. 1. – P. 1–13.

11. Kerns, D. M. Correction of Near-Field Antenna Measurements Made with an Arbitrary but Known Measuring Antenna / D. M. Kerns // Electronics Letters. – 1970. – Vol. 6, no. 11. – P. 346–347.

12. Kerns, D. M. Plane-Wave Scattering Matrix Theory of Antennas and Antenna-Antenna Interactions : NBS Monograph 162 / D. M. Kerns. – Washington, DC : U. S. Government Printing Office, 1981. – 179 p.

13. Айзенберг Г. З. Антенны УКВ: в 2 т. / Г. З. Айзенберг, В. Г. Ямпольский, О. Н. Терёшин ; под ред. Г. З. Айзенберга. – М. : «Связь», 1977. – 384 с.

14. Наденeнко С. И. Антенны / С. И. Наденeнко. – М. : Связьиздат, 1959. – 552 с.

15. Сазонов, Д. М. Антенны и устройства СВЧ / Д. М. Сазонов. – М. : Высш. шк., 1988. – 432 с.

16. Stutzman, W. Polarization in Electromagnetic Systems / W. Stutzman. – 2nd ed. – Artech House, 2018. – 378 p.

17. Mott, H. Polarization in Antennas and Radar / H. Mott. – Wiley, 1986. – 297 p.

18. Кононов, А. А. Аналитическая модель малогабаритного зонда для измерений электромагнитного поля / А. А. Кононов, Г. К. Усков // Сб. материалов XXV Междунар. научно-практической конф. «Информатика: проблемы, методы, технологии» (Воронеж, 12–13 февраля 2025 г.). – Воронеж, 2025. – С. 1120–1127.

19. Кононов, А. А. Способы теоретического расчета векторной структуры электромагнитного поля в ближней зоне апертурного излучателя / А. А. Кононов, Г. К. Усков // Инфокоммуникационные и радиоэлектронные технологии. – 2025. – Т. 8, № 5. – С. 687–703.

20. Кононов, А. А. Вычисление поляризационных диаграмм направленности с учетом пространственной ориентации излучающих устройств / А. А. Кононов, Г. К. Усков // Сб. тезисов XI Международной заочной научно-технической конференции «Информационные технологии. Радиоэлектроника. Телекоммуникации» (Тольятти, 29 ноября 2024 г.). – Тольятти, 2025. – С. 224–234.

21. Калинин, В. И. Введение в радиофизику / В. И. Калинин, Г. М. Герштейн ; под ред. В. И. Калинина. – М. : Гос. изд-во технико-теоретической литературы, 1957. – 660 с.

Концептуальная модель совместной обработки видео и радиолокационных данных для обнаружения БПЛА в различных условиях

В. В. Плеве, email: zond1908@rambler.ru
А. В. Шешко, email: inokendiy.sigizmundovich@yandex.ru

ВУНЦ ВВС «ВВА», г. Воронеж

Аннотация. *Предложена концептуальная модель совместной обработки видео- и радиолокационных данных для создания устойчивой системы обнаружения малоразмерных беспилотных летательных аппаратов (БПЛА). Модель основана на взаимной компенсации слабых сторон отдельных каналов: оптического (различные погодные условия) и радиолокационного (помехи от фона и малая эффективная площадь рассеяния цели). Архитектура модели включает модули оценки контекста, адаптивного объединения данных и разрешения конфликтов. Рассмотрены пути реализации на основе методов глубокого обучения и вероятностных подходов. Обсуждаются преимущества, ограничения и практическая значимость модели для защиты критических объектов.*

Ключевые слова: *обнаружение БПЛА, совместная обработка данных, видеоаналитика, радиолокация, малозаметные цели, адаптивное объединение, микродоплеровская подпись.*

Введение

Рост числа беспилотных летательных аппаратов (БПЛА) сопровождается увеличением угроз их несанкционированного использования. Существующие системы обнаружения, основанные на одном типе датчиков, имеют принципиальные ограничения: оптические системы уязвимы к погодным условиям и освещению, а радиолокационные — к сильным помехам от фона и малой отражающей способности целей. Возникает потребность в создании более совершенных систем обнаружения БПЛА [1].

Цель работы — разработка концептуальной модели, системно описывающей совместную обработку видео и радиолокационных данных для создания адаптивной и устойчивой системы обнаружения.

Научная новизна заключается в формализации принципов взаимной компенсации каналов, предложении архитектуры с динамической оценкой контекста и интеграции механизмов адаптивного объединения данных.

1. Физико-информационные основы обнаружения БПЛА: слабые стороны каналов и их взаимная компенсация

Оптический (видео) канал предоставляет большое количество визуальных признаков [2-3]:

- Морфологические признаки: форма, контур, соотношение сторон, текстура корпуса и винтов, цвет.
- Пространственно-временные признаки: траектория движения, скорость, ускорение, характер манёвров.
- Признаками радиолокационного канала представлена следующими группами:
- Микродоппплеровские спектры: наиболее отличительный признак БПЛА. Вращение лопастей винтов модулирует отражённый сигнал, создавая в спектрограмме характерные гармонические структуры. Эти спектры несут информацию о количестве и длине лопастей, скорости вращения, что является уникальной «подписью» типа БПЛА.
- Кинематические параметры: Точные радиальные скорость, дальность и угловые координаты, не зависящие от освещённости.
- Параметры ЭПР: Флюктуации эффективной площади рассеяния, которые могут быть использованы для статистического распознавания.

Ключевая идея модели основана на принципе взаимной компенсации — формализация взаимодействия между каналами. Данные одного канала должны автоматически восполнять информационные провалы другого. Принцип взаимной компенсации реализуется в следующих сценариях:

- Сценарий 1 (туман, ночь): Радар уверенно обнаруживает и отслеживает цель, предоставляя её кинематический трек. Оптический канал, хотя и не может предоставить чёткое изображение, может использовать грубую позицию от радара для пристального анализа области интереса.
- Сценарий 2 (сильные помехи, низкая ЭПР): Радар теряет цель на фоне помех. Однако оптический канал, не подверженный этому типу помех, может продолжать сопровождение по визуальным признакам, предоставляя радару данные для уточнения траектории.
- Сценарий 3 (идентификация и снижение ложных тревог): Радар выдаёт гипотезу о классе объекта на основе микродоппплеровской подписи, что позволяет отличать БПЛА от птиц с похожей ЭПР. Оптический канал выполняет

окончательную визуальную проверку, снижая вероятность ошибочной классификации.

Таким образом, объединённая система превосходит сумму своих частей, обеспечивая обнаружение там, где каждая система в отдельности терпит неудачу.

2. Архитектура модели совместной обработки

Общая концепция модели - каскадная архитектура с обратной связью для иерархической обработки данных. Полная цепочка: датчики - предобработка и синхронизация - модуль ситуационной осведомленности - параллельное извлечение признаков - модуль адаптивного слияния (EADW) - ассоциация и сопровождение модуль слияния решений - принятие решения. Обратная связь позволяет уточнять область интереса и проводить калибровку с ключевыми модулями:

Модуль оценки контекста. Анализирует текущие условия (освещенность, погоду, уровень помех) для определения потенциальной надежности каждого датчика.

Модуль адаптивного объединения данных. Динамически распределяет «вес» каждого канала в итоговом решении на основе оценок контекста и достоверности текущих измерений. В сложных условиях влияние ненадежного канала автоматически снижается.

Модуль разрешения конфликтов. Сопоставляет данные от разных источников во времени и пространстве, поддерживает целостные траектории целей и разрешает противоречия с использованием вероятностных методов.

Модуль оценки контекста. Анализирует текущие условия (освещенность, погоду, уровень помех) для определения потенциальной надежности каждого датчика [4]. Контекст модуля включает:

– Время суток (день/ночь/сумерки), формализуемое, например, как (1).

$$c = 0.5 * \left[\sin\left(2\pi * t / 24\right) + 1 \right] \quad (1)$$

– Метеоусловия (видимость, влажность, осадки).

– Освещённость для видео-канала: (2).

$$c = \left(\log10\left(L+1\right) - \log10\left(L_{min}+1\right)\right) / \left(\log10\left(L_{max}+1\right) - \log10\left(L_{min}+1\right)\right) \quad (2)$$

Фоновая загруженность (помеховый уровень для радара, сложность текстуры для видео).

Извлечение пространственных признаков включает использование сверточных нейронных сетей (CNN). Например, модифицированную сеть ResNet-50, настроенную на сохранение деталей

мелких объектов, и модуль двухпутевого внимания (Dual-Path Attention Module, DPAM). Этот модуль параллельно анализирует канальное внимание (определяя важные цветовые или тепловые каналы) и пространственное внимание (выделяя значимые области кадра), усиливая информативные части данных. Происходит получение характеристик движения (траектория, скорость, ускорение, тип маневров) с использованием алгоритмов оптического потока (определяющих смещение пикселей между кадрами). На выходе модуль формирует одномерные векторы, образуя временные ряды для выявления закономерностей в движении.

При этом для **обработки оптических данных** предлагается использовать следующие методы [5]:

- для глубокого обучения: модели YOLOv4 (точность mAP 89.2%), YOLOv5 (mAP 91.5%), YOLOv8 (mAP 93.1%), Faster R-CNN (mAP 88.7%).
- для классического машинного обучения: метод опорных векторов (SVM, точность 85.3%) и метод k-ближайших соседей (K-NN, точность 82.1%).
- для обработки групп близко расположенных целей: стратегия с модифицированным алгоритмом подавления немаксимумов (Non-Maximum Suppression, NMS). Его порог срабатывания динамически меняется в зависимости от локальной плотности объектов.

Для **извлечения радиолокационных признаков** предлагается использовать следующие ключевые аспекты и методы их извлечения [7-8]:

- микродоплеровские спектры – вращение винтов БПЛА (50-500 Гц) модулирует отраженный сигнал, создавая на спектрограмме (визуализации спектра во времени) характерные гармонические структуры в виде вертикальных полос. Для их выделения предлагается использовать метод ортогонального согласованного преследования (OMP) для подавления фоновых помех при низком соотношении сигнал/шум (SNR до -10 дБ), а также кратковременное преобразование Фурье (STFT) с окном Ханна длиной 128 отсчетов и перекрытием 50% для построения спектрограммы.
- параметры эффективной площади рассеяния (ЭПР), которые для малых БПЛА колеблется от -30 до -10 дБсм². Её разброс анализируют через статистические модели (распределения Релея и Райса).

– кинематические параметры: дальность (точность до 0.5 м) и радиальная скорость (точность до 0.1 м/с).

Обработку радарных данных предлагается осуществлять следующими способами: исходные квадратурные выборки сигнала (IQ-данные) преобразовывать с помощью STFT в амплитудный спектр, затем применять логарифмическое преобразование в децибелы. Итоговая спектрограмма будет иметь размерность 1000×128. Для извлечения может использоваться двухветвевая сеть. Основная ветвь (ConvNeXt) отвечает за общие признаки, а облегченная U-Net — за детали микродопплеровских спектров. Признаки сливаются с помощью механизма кросс-внимания и адаптивной пирамиды признаков (AFP) для многомасштабного анализа.

Модуль адаптивного объединения данных динамически распределяет «вес» каждого канала в итоговом решении на основе оценок контекста и достоверности текущих измерений. В сложных условиях влияние ненадежного канала автоматически снижается. Алгоритм работы модуля: для каждой модальности m формируется входной вектор контекста, уверенность детектора $P_m(t)$ и информационная энтропия признаков $H_m(t)$. Предварительный вес $\alpha'_m(t)$ вычисляется через небольшую нейронную сеть. Веса нормализуются с помощью функции Softmax: $\alpha_{видео}(t)$ и $\alpha_{радар}(t) = 1 - \alpha_{видео}(t)$. Происходит слияние признаков (3).

$$F_{fused}(t) = \alpha_{видео}(t) \cdot F_{видео}(t) + \alpha_{радар}(t) \cdot F_{радар}(t) \qquad (3)$$

Модуль разрешения конфликтов сопоставляет данные от разных источников во времени и пространстве, поддерживает целостные траектории целей и разрешает противоречия с использованием вероятностных методов. Данный модуль содержит логический и вероятностный аппарат для корректного сопоставления разных обнаружений обоих каналов. Используются:

– алгоритмы глобальной ассоциации данных и трекинг на основе фильтра Калмана.

– теория доказательств Демпстера-Шафера (D-S) для количественной оценки неопределённости и разрешения конфликтов между решениями в разных каналах. При высоком коэффициенте конфликта K активируется механизм переоценки надёжности источников.

временное моделирование для работы с последовательностями кадров и улучшения построения траектории.

3. Методологические подходы к реализации модели

Путь на основе глубокого обучения предполагает использование двухканальных (или более) нейросетевых архитектур.

Раннее и промежуточное слияние: Видео и радиолокационные спектрограммы подаются на вход сети, которая учится совместному представлению.

Используются механизмы внимания, которые автоматически учатся выделять и объединять наиболее значимые признаки из разных источников данных.

Вероятностный подход основан на теории вероятностей и часто более интерпретируем.

Применение байесовских сетей**,** фильтров Калмана и теории Демпстера-Шафера для объединения данных, оценки неопределенности и принятия решений в условиях конфликта информации.

4. Потенциал использования модели и ограничения ее использования

Ожидаемые преимущества модели подтверждаются исследованиями и включают: высокую вероятность обнаружения (до 98%) в широком диапазоне условий, снижение числа ложных тревог, всепогодность и возможность уверенной классификации целей.

Фундаментальные ограничения связаны с физикой датчиков (дальность, чувствительность) и вычислительной сложностью обработки данных в реальном времени, особенно при высокой плотности целей.

Практическая значимость заключается в создании методологической основы для систем охраны периметра критически важных объектов (аэропорты, АЭС), где требуется надежное круглосуточное обнаружение малоразмерных БПЛА.

Заключение

Предложенная модель задает принципы построения эффективных гибридных систем обнаружения БПЛА, позволяет приступить к разработке надёжных, всепогодных и адаптивных средств обнаружения, способных работать в сложных урбанизированных и природных ландшафтах. Такие средства обнаружения позволят решать задачи обнаружения несанкционированных БПЛА для объектов критической инфраструктуры, которую не могут в полной мере решить автономные системы из-за их принципиальных ограничений. Они могут быть внедрены в распределенные сетевые системы и интегрированы в контур «обнаружение — классификация — нейтрализация», тем самым совершенствуя текущие средства защиты от таких угроз.

Список литературы

1. Василии, Н. Я. Беспилотные летательные аппараты / Н. Я. Василии. – Минск : «Попурри», 2003. – 272 с.

2. Васильев, В. В. Дальность действия оптических средств наблюдения беспилотных летательных аппаратов / В. В. Васильев, Р. В. Джуган, А. П. Манин // Полёт. – 2015. – № 10. – С. 3–8

3. Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление. / Б. А. Алпатов [и др.]. – М. : Радиотехника – 2008. – 176 с.

4. Васильев, В. В. Концептуальные вопросы построения оптических систем обеспечения безопасности полётов, контроля параметров движения и ориентации летательных аппаратов в пространстве / В. В. Васильев, А. П. Манин, В. В. Челахов // Полёт. – 2013. – № 9. – С.54–59.

5. Панорамные оптические средства наблюдения беспилотных летательных аппаратов / В. В. Васильев В. В. [и др.] // Полёт. – 2018. – № 4. – С. 3–8.

6. Годунов, А. И. Комплекс обнаружения и борьбы с малогабаритными беспилотными летательными аппаратами / А. И. Годунов, С. В. Шишков, Н. К. Юрков // Надёжность и качество сложных систем. – 2014. – № 2. – С. 62–69.

7. Справочник по радиолокации [Текст]. В 2 т. Т. 1. Основы радиолокации / ред. М. И. Сколник ; пер. с англ. А. Я. Брейтбарта [и др.] ; под ред. Я. С. Ицхоки. –М. : Сов. радио – 1976. – 454 с.

8. Теоретические основы радиолокации / под ред. Я.Д. Ширмана. – М. : Сов. радио – 1970. – 560 с.

Электродинамический анализ решеток электрических вибраторов для радиоэлектронных комплексов со сниженным уровнем вторичного электромагнитного излучения

С. Н. Разиньков, e-mail: razinkovsergey@rambler.ru [1],
О. Э. Разинькова, e-mail: olgarazinkovaic@icloud.com[1,2],
А. В. Савченко, e-mail: nomenate1@gmail.com [1]

[1] Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н. Е. Жуковского и Ю. А. Гагарина» (г. Воронеж)
[2] Воронежский государственный технический университет

Аннотация. *На основе метода интегральных уравнений при представлении передаваемых (принимаемых) волновых процессов в виде пространственно-временных функций построены электродинамические модели решеток электрических вибраторов с диапазонными устройствами коммутации и широкополосными распределительными линиями. Разработана методика анализа и исследованы показатели пространственно-частотной избирательности приемопередающих структур для радиоэлектронных комплексов со сниженным уровнем вторичного электромагнитного излучения.*

Ключевые слова: *Антенная решетка, электрический вибратор, диапазонное устройство коммутации, распределительная линия, метод интегральных уравнений, некорректная задача.*

Введение

Одной из важных задач построения радиоэлектронных комплексов является разработка антенных систем, характеристики которых, с одной стороны, устанавливают пространственно-частотную избирательность передачи (приема) радиосигналов, с другой стороны, создают уровень вторичного электромагнитного излучения, способного негативно влиять на условия совместимости [1] и показатели эффективности выполнения целевых функций различных устройств [1, 2]. Улучшение направленных и диапазонных свойств антенн сопровождается возрастанием мощности рассеяния волновых процессов [1].

Таким образом, поиск параметров конструкции антенных систем в интересах достижения технических характеристик, в наиболее полной мере удовлетворяющих требованиям к радиоэлектронным комплексам, необходимо проводить на основе электродинамического анализа [1, 3]

© Разиньков С.Н., Разинькова О.Э., Савченко А.В., 2026

с выявлением закономерностей распределения электромагнитного поля и токов приемоизлучающих структур.

В предлагаемой работе с использованием метода интегральных уравнений [1], примененного для постановки и решения некорректных задач теории поля [4], построены модели и методика анализа решеток электрических вибраторов для радиоэлектронных комплексов с малым уровнем вторичного электромагнитного излучения.

Перечень моделей включает в себя:

– модель решетки электрических вибраторов с диапазонными устройствами коммутации каналов для исследования передачи (приема) сигналов со структурной и энергетической скрытностью;

– модель решетки электрических вибраторов с широкополосными распределительными линиями для исследования пределов уменьшения мощности рассеиваемого поля при рассогласовании приемоизлучающей структуры с фидерными трактами.

Цель работы – исследование путей наращивания эффективности передачи (приема) сигналов при сохранении низких уровней вторичного электромагнитного излучения решеток электрических вибраторов.

1. Модель решетки электрических вибраторов с диапазонными устройствами коммутации каналов

Математической основой для разработки модели решетки является модель идеально проводящего круглого цилиндра с бесконечно тонкими стенками, облучаемого широкополосным волновым процессом [5]. Суть моделирования заключается в формализованном описании процедуры преобразования пространственно-временной структуры излучаемых (принимаемых) сигналов с их распределением в антенных нагрузках при взаимодействии между элементами антенной системы через вторичное электромагнитное излучение.

Построение электродинамической модели антенной решетки с диапазонными устройствами коммутации каналов включает четыре этапа.

1. Аналитическое представление электрических полей вибраторов на основе модели идеально проводящего цилиндра в приближении малого радиуса поперечного сечения. При малом радиусе поперечного сечения вибратора распределение поверхностных токов не содержит вариаций по окружности поперечного сечения [5, 6]; азимутальные компоненты токов антенных элементов равны нулю, а их продольные составляющие могут быть представлены нулевыми пространственными гармониками при разложении в ряды Фурье.

2. Аналитическое представление суперпозиции полей вибраторов с эквивалентными осевыми токами; получение интегральных уравнений

для установления взаимосвязей электрического поля с распределениями токов решетки при устранении особенностей ядер за счет выбора точек интегрирования на осях, а точек наблюдения на поверхностях антенных элементов.

3. Применение условия излучения Зоммерфельда [3] и граничных условий [1, 3] для суперпозиции электрических полей на поверхностях вибраторов; получение при задании электромагнитных связей в решетке [6] системы интегральных уравнений для пространственно-временного распределения токов.

4. Аппроксимация токов последовательностями пространственных и временных базисных функций и на основе дискретизации операторов полученной системы интегральных уравнений формирование системы линейных алгебраических уравнений (СЛАУ) относительно множества отсчетов искомых токов в пространственно-временной области.

Для обеспечения устойчивости частичного обращения матричного оператора СЛАУ и сходимости последовательностей для аппроксимации токов их весовые коэффициенты вычисляются при равенстве интервала дискретизации поверхности решетки и пространственной длительности шага по времени, выбираемого в соответствии с теоремой Котельникова.

За счет пространственно-временного представления передаваемых (принимаемых) сигналов становится возможным анализ переходных процессов возбуждения приемоизлучающих структур, энергетической доступности и пространственной избирательности передачи (приема) сигналов со структурной и энергетической скрытностью.

2. Модель решетки электрических вибраторов с широкополосными распределительными линиями

Модель решетки электрических вибраторов с широкополосными распределительными линиями содержит четыре этапа.

1. Аналитическое представление электрических полей вибраторов при плавном изменении профилей в областях разрывов поверхностей [3] для подключения распределительных линий.

2. Аналитическое представление эквивалентных осевых токов на поверхности приемоизлучающей структуры и фиктивных токов [3, 6] в зазорах вибраторов

3. Применение условия излучения Зоммерфельда и граничных условий для суперпозиции электрических полей приемной структуры со сшиванием эквивалентных и фиктивных токов вибраторов на границах областей определения; получение системы интегральных уравнений для пространственно-временного представления токов антенной решетки.

4. Аппроксимация токов произведениями последовательностей пространственных и временных базисных функций и получение СЛАУ

относительно массива коэффициентов, восстанавливающих решение краевой задачи для антенной решетки.

По результатам частичного обращения матричного оператора СЛАУ при последовательном продвижении по времени оцениваются:

– пространственно-частотная избирательность передачи (приема) сигналов, в том числе при реализации мер по обеспечению структурной и энергетической скрытности;

– степень снижения мощности вторичного электромагнитного излучения антенной системы при преднамеренном рассогласовании с антенными нагрузками.

3. Методика анализа решеток электрических вибраторов для радиоэлектронных комплексов с малым уровнем вторичного электромагнитного излучения

Суть методики анализа решеток электрических вибраторов для радиоэлектронных комплексов с малым уровнем вторичного электромагнитного излучения заключается в выполнении двух этапов.

1. Получение интегральных уравнений для пространственно-временных преобразований сигналов при передаче (приеме). Данная процедура реализуется с применением электродинамических моделей антенных решеток с диапазонными устройствами коммутации каналов и широкополосными распределительными линиями.

2. Решение системы интегральных уравнений для установления закономерностей пространственно-временных преобразований сигналов при выполнении следующих шагов:

– аппроксимация волновых процессов множеством отсчетов в пространственно-временной области при дискретизации длительности в соответствии с теоремой Котельникова и формирование СЛАУ для нахождения токов вибраторов в виде дискретных значений комплексных амплитуд;

– решение СЛАУ методом последовательного продвижения по времени, позволяющее восстановить переходные процессы возбуждения вибраторов [6] и исследовать показатели пространственно-частотной избирательности передачи (приема) сигналов [2], в том числе имеющих структурную и энергетическую скрытность [5, 6], при контроле уровней вторичного электромагнитного излучения.

Заключение

На основе метода интегральных уравнений для пространственно-временного распределения волновых процессов разработаны модели решеток электрических вибраторов с диапазонными устройствами коммутации каналов в режиме передачи (приема) и широкополосными

распределительными линиями. На базе представленных моделей построена методика электродинамического анализа антенных решеток вибраторного типа для радиоэлектронных комплексов в высокими показателями пространственно-частотной избирательности и малым уровнем вторичного электромагнитного излучения.

Разработанные модели и методика анализа решеток электрических вибраторов, в отличие от ранее известных [1–3], предназначенных для исследования характеристик антенных систем при передаче (приеме) узкополосных (монохроматических) волновых процессов [1, 3], за счет пространственно-временного представления токов антенных элементов позволяют исследовать пространственно-частотную избирательность структур, возбуждаемых широкополосными сигналами.

Таким образом, расширяются возможности модельного описания и теоретического исследования радиоэлектронных комплексов, в работе которых применяются излучения со структурной и энергетической скрытностью, а также оценки контрастности объектов в полосе частот вторичного электромагнитного поля [5, 6].

Работа выполнена при поддержке Министерства науки и высшего образования (проект № FZGM-2025-0002).

Литература

1. Бузова, М.А. Электродинамический анализ сложных антенных комплексов на основе интегральных уравнений / М. А. Бузова. – М. : Радиотехника, 2025. – 352 с.

2. Антенны, СВЧ-устройства и их технологии / Д. Д. Габриэльян [и др.] ; под общ. ред. А. А. Косогора. – М. : ФИЗМАТЛИТ, 2023. – 400 с.

3. Неганов, В.А. Современная теория и практические применения антенн / В. А. Неганов, Д. П. Табаков, Г. П. Яровой. – М.: Радиотехника, 2009. – 720 с.

4. Тихонов, А. Н. Методы решения некорректных задач / А. Н. Тихонов, В. Я. Арсенин. – М. : Наука, 1986. – 288 с.

5. Попова, О. Э. Рассеяние радиоимпульсов с прямоугольной и гауссовской огибающей на идеально проводящем круглом цилиндре конечной длины / О. Э. Попова, С. Н. Разиньков // Физика волновых процессов и радиотехнические системы. – 2008. – Т. 11. – № 4. – С. 18–25.

6. Попова, О. Э. Нестационарный электродинамический анализ излучения и приема негармонических сигналов решетками вибраторов / О. Э. Попова, С. Н. Разиньков // Физика волновых процессов и радиотехнические системы. – 2008. – Т. 11. – № 3. – С. 83–93.

Анализ передачи-приема сигналов в радиолиниях с зеркальными параболическими антеннами

О.Э. Разинькова, e-mail: olgarazinkovaic@icloud.com

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина» (г. Воронеж)
Воронежский государственный технический университет

Аннотация. *На основе спектрального представления волновых процессов при задании парциальных показателей пространственно-частотной избирательности передающих и приемных устройств получено энергетическое уравнение для передачи-приема сигналов в радиолиниях с зеркальными параболическими антеннами. Исследована энергетическая эффективность информационного обмена по прямому радиоканалу при использовании резонансных, диапазонных и частотно-независимых облучателей зеркальных антенн. При предельных уровнях средней мощности излучения получены оценки скорости доведения сообщений при передаче битов информации импульсами амплитудно-модулированной последовательности с пассивной паузой для приема двоичных сигналов известной формы и их последующей оптимальной обработки.*

Ключевые слова: *Прямой радиоканал, зеркальная параболическая антенна, частотно-независимый, диапазонный и резонансный облучатель, энергетическое уравнение передачи-приема, парциальный коэффициент направленного действия антенны*

Введение

Зеркальные параболические антенны с узкими главными лучами и низкими уровнями боковых лепестков диаграмм направленности [1, 2], обеспечивающие высокие энергетические потенциалы радиоканалов, эффективно используются в радиорелейных станциях и терминалах спутниковой связи. Процессы передачи-приема сигналов с применением таких антенн характеризуются энергетической доступностью абонентов, скрытностью [3, 4], защищенностью от несанкционированного доступа к информации и деструктивных воздействий [2]. Вследствие оснащения рефлекторами, фокусирующими энергию излучаемого (принимаемого) электромагнитного поля [1], зеркальные антенны обладают наилучшими функциональными свойствами в классе приемопередающих структур,

© Разинькова О. Э., 2026

обеспечивающими интенсивный обмен сообщениями по радиоканалам при ограничениях на полосы частот и мощность излучаемых сигналов. Кроме того, они отличаются меньшей стоимостью и более простой технологией производства по сравнению с активными фазированными антенными решетками [2].

Оптимизация их пространственно-частотной избирательности [2] может выполняться за счет рационального выбора фокусных расстояний и размеров раскрывов рефлекторов, а также формирования амплитудно-фазового распределения токов облучателей [5]. Спектральная плотность энергии электромагнитного поля антенны зависит от вида передаточной функции облучателя, которая, как отмечено в [6], в пределах главного луча диаграммы подлежит достаточно точной аппроксимации степенной функцией. Показатель степени для частотно-независимых, диапазонных и резонансных облучателей соответственно принимается равным 0, 1, 2. Нормировка спектрального представления парциальных коэффициентов направленного действия (КНД) антенн находится в виде отношения их величин на центральной частоте диапазона рабочих частот к значению этой частоты.

В предлагаемой работе при спектральном представлении волновых процессов, распространяющихся в радиоканале прямой видимости [2], и показателей пространственно-частотной избирательности [6] устройств получены энергетические уравнения передачи-приема [6, 7] и проведен анализ энергетической эффективности информационного обмена [6, 8] в радиолиниях с зеркальными параболическими антеннами.

Цель работы – исследование скорости передачи сигналов при использовании зеркальных антенн, содержащих частотно-независимые, диапазонные и резонансные облучатели.

1. Модель радиоканала прямой видимости с зеркальными параболическими антеннами

Энергетическая эффективность передачи-приема сигналов в канале прямой видимости определяется по уровню энергии, поступающей на вход приемной антенны через среду распространения непосредственно от передающей системы без проявления эффектов дифракции на трассе распространения [3, 4].

В этой связи при построении модели радиолинии будем полагать, что антенны размещаются на высотах, исключающих закрытие трассы и замирания (федингы) амплитуд волновых процессов при отражении от локальных препятствий. Расстояние между позициями развертывания средств передачи и приема сигналов значительно превосходит границы дальней зоны антенн. Вследствие чего энергия излучения в радиоканале

затухает пропорционально дальности до источника во второй степени, а волновой процесс на раскрыве приемной антенны имеет плоский фронт.

Рефлекторы передающей и приемной антенн выполнены в виде идеально проводящих параболоидов вращения с равными значениями радиусов раскрыва и фокусных расстояний. Точки фокусов рефлекторов совмещены с геометрическими центрами облучателей. Размеры зеркал и фокусные расстояния выбраны из условия нахождения каждой точки на внутренних поверхностях рефлекторов в дальних зонах облучателей. В результате при малых электрических размерах облучателей выполняется синхронное возбуждение раскрывов [1] с минимизацией потерь энергии сигналов вследствие малой результативности перехвата рефлекторами передающей и приемной антенн.

Облучатели антенн идеально согласованы с фидерными трактами по электрическим цепям в диапазоне рабочих циклических частот.

Для нахождения предельной скорости информационного обмена по радиоканалу в качестве носителя информации выберем регулярную последовательность амплитудно-модулированных импульсов, в каждом из которых содержится один бит информации, с пассивной паузой [8].

При передаче-приеме по главным лучам диаграмм направленности антенн сигналы претерпевают дополнительную частотную модуляцию, вид и параметры которой определяются формой передаточных функций облучателей и рефлекторов. За счет отражения от рефлектора импульс, возбуждающий антенну, приобретает вид своей первой производной [9]. Поэтому структура последовательности импульсов на входе приемника считается известной, что позволяет выполнять посимвольный прием двоичных сигналов с прогнозируемым изменением частотно-временных параметров без анализа распределений парциальных КНД в секторах углов работы антенн. Спектральная плотность энергии электрического поля антенны определяется амплитудным спектром [6, 8, 9] излучаемого (принимаемого) сигнала.

Таким образом, в соответствии с приведенными выше условиями для модельного представления передачи-приема волновых процессов получим энергетическое уравнение радиолинии, связывающее энергию сигнала на входе приемника со спектральным распределением КНД передающей и приемной антенн, спектральной плотностью излучаемой энергии при установленной дистанции информационного обмена.

Согласно [6, 9], энергетическое уравнение находится в результате интегрирования спектрального энергетического уравнения радиоканала в полосе рабочих циклических частот при аппроксимации предельного уровня спектральной плотности энергии излучения кусочно-постоянной функцией. Максимально допустимый уровень излучения в диапазоне

частот работы радиолинии [7, 8] определяется отношением наибольшего значения средней мощности передатчика [8] к произведению скорости передачи и ширины полосы сигнала.

Полученное энергетическое уравнение позволяет проводить анализ передачи-приема сигналов в радиолиниях и выполнять рациональный выбор параметров конструкций антенн для достижения требуемой интенсивности информационного обмена при ограничениях на уровни средней мощности и полосы рабочих частот передатчиков.

2. Исследование информационного обмена в радиолиниях с зеркальными параболическими антеннами

С использованием модели радиолинии проведен анализ передачи-приема сигналов в системе радиосвязи с зеркальными параболическими антеннами.

При базовых уровнях скорости дальность передачи сообщений по каналам прямой видимости с зеркальными антеннами, содержащими диапазонные облучатели, превосходит длину радиолинии с антеннами, оснащенными резонансными облучателями, в 3,2…3,35 раза. За счет применения антенн с диапазонными и частотно-независимыми облучателями дистанция связи возрастает в 1,4 и 1,55 раза. В системе радиорелейной связи уровня STM-16 достигается требуемая скорость информационного обмена 2,5 Гбит/с.

При использовании зеркальных антенн с частотно-независимыми и диапазонными облучателями наиболее высокие уровни излучения характерны для высокочастотных составляющих сигналов за счет роста парциальных КНД по мере увеличения циклической частоты. Прием волновых процессов антеннами с частотно-независимыми облучателями осуществляется с минимальными потерями энергии, обусловленными фильтрацией низкочастотных составляющих спектров [7]. Вместе с тем, такие антенны имеют низкую эффективность возбуждения рефлектора последовательностями видеоимпульсов, спектральные функции которых содержат постоянные составляющие на нулевой циклической частоте.

В результате замены зеркальных антенн с частотно-независимыми облучателями на частотно-независимые гипотетические антенны энергия принимаемого сигнала с относительной полушириной полосы частот 0,5 возрастает в 1,33 раза вследствие увеличения эффективных площадей в нижней части диапазона циклических частот, парирующей малую эффективность излучения низкочастотных процессов. При парциальном КНД частотно-независимой антенны на центральной циклической частоте энергия излучаемого (принимаемого) сигнала возрастает относительно уровня, характерного для зеркальной антенны с частотно-

независимым облучателем, пропорционально площади рефлектора и полуширине полосы частот во второй степени.

Прирост энергии сигнала, принимаемого антенной, оснащенной частотно-независимым облучателем, по сравнению с резонансными и диапазонными антеннами обусловлен уменьшением эквивалентной ширины его спектра по сравнению с полосой импульса, возбуждающего передающую антенну, за счет дифференцирования при излучении и приеме.

Заключение

На основе спектрального энергетического уравнения радиоканала прямой видимости с парциальными показателями пространственно-частотной избирательности передачи-приема сигналов [8, 9] построена модель и получены результаты анализа энергетической эффективности радиолиний с зеркальными параболическими антеннами. Исследованы закономерности изменения показателей энергетической эффективности каналов информационного обмена при применении зеркальных антенн с частотно-независимыми, диапазонными и резонансными облучателями.

При предельных уровнях мощности излучения найдены значения скорости передачи амплитудно-модулированных импульсов регулярной последовательности. Продемонстрированы возможности реализации требуемых информационных потоков в системе радиорелейной связи уровня STM-16 при базовой скорости.

Полученные методические результаты способствуют развитию методов анализа и синтеза [2], базовых технологий построения антенн с высокими показателями пространственно-частотной избирательности передачи-приема сигналов [3, 4] в условиях наращивания требований к интенсивности информационного обмена и ограничений на предельно допустимые уровни средней мощности передатчиков.

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (проект № FZGM-2025-0002).

Литература

1. Антенны / Ю. Т. Зырянов [и др.]. – СПб. : Лань. – 2016. – 416 с.
2. Антенны, СВЧ-устройства и их технологии / Д. Д. Габриэльян [и др.] ; под общ. ред. А. А. Косогора. – М. : ФИЗМАТЛИТ. – 2023. – 400 с.
3. Фролов, О. А. Антенны и фидерные тракты для радиорелейных линий связи / О. А. Фролов. – М. : Радио и связь. – 2001. – 416 с.
4. Фролов, О. А. Антенны для земных станций спутниковой связи / О. А. Фролов. – М. : Радио и связь. – 2000. – 376 с.

5. Бахрах, Л. Д. Зеркальные сканирующие антенны: теория и методы расчета / Л. Д. Бахрах, Л. И. Галимов. – М. : Наука. – 1981. – 389 с.

6. Богословский, А. В. Дальность обнаружения воздушных объектов широкополосными системами активно-пассивной радиолокации / А. В. Богословский, С. Н. Разиньков, О. Э. Разинькова // Антенны. – 2023. – № 2 (282). – С. 32–43.

7. Разиньков, С. Н. Эффективность подавления сверхширокополосных радиолокационных станций маскирующими помехами / С. Н. Разиньков, О. Э. Разинькова // Физика волновых процессов и радиотехнические системы. – 2012. – Т 15. – № 3. – С. 67–74.

8. Авдеев, В. Б., Авдеева Д. В. Уравнения радиосвязи при использовании сверхширокополосных сигналов / В. Б. Авдеев, Д. В. Авдеева // Телекоммуникации. – 2002. – № 9. – С. 2–10.

9. Авдеев, В. Б. Энергетические характеристики антенн и антенных систем при излучении и при приеме сверхширокополосных сигналов и сверхкоротких импульсов / В. Б. Авдеев // Антенны. – 2000. – № 7 (62). – С. 5–26.

Анализ среднего квадратического отклонения погрешностей определения дальности пуска средства поражения при двухпозиционном пеленговании

Д. В. Смирнов, email: dm.sm1rnov@mail.ru
А. А. Донцов, email: addoncov1@mail.com

ВУНЦ ВВС «Военно-Воздушная академия им. Н. Е. Жуковского и Ю. А. Гагарина»

Аннотация. *Используя данные об угловых координатах средства поражения и пространственные координат группы летательных аппаратов на основе двухпозиционного пеленгования летательными аппаратами, оснащенными комплексами индивидуальной защиты, получены значения среднего квадратического отклонения погрешностей определения дальности пуска средства поражения при двухпозиционном пеленговании. Выполнен анализ вычисленных средних квадратических отклонений погрешностей для заданных значений параметров информационного обеспечения при различных дальностях пуска и высотах полета летательных аппаратов. Методом имитационного моделирования проверена сходимость теоретических и практических значений.*

Ключевые слова: *двухпозиционное пеленгование, комплекс индивидуальной защиты, летательный аппарат, погрешность, средство поражения, среднеквадратическое отклонение.*

Введение

В условиях современной геополитической обстановки в стране и в мире наблюдается рост террористических угроз со стороны различных незаконных вооруженных формирований. Основная цель деятельности таких формирований – пугать общество, посеять страх и неуверенность, дестабилизировать обстановку и подорвать доверие людей к государству и общественным институтам. Так, в частности, одним из объектов для проведения такой деятельности являются гражданские и военные летательные аппараты (ЛА), выполняющие транспортные задачи (перевозку пассажиров, доставку гуманитарных грузов) в зонах вооруженных конфликтов и граничащих с ними территориях. Средством поражения (СП) чаще всего выступают переносные зенитно-ракетные комплексы (ПЗРК). Они получили широкое распространение в странах Ближнего востока, Средней Азии, Украине. В эксплуатации они

© Смирнов Д. В., Донцов А. А., 2026

мобильны и просты, что позволяет в короткие сроки подготовить стрелка и спланировать террористический акт из скрытной позиции. Как правильно, доставка гуманитарной помощи населению, перевозка людей осуществляется группой ЛА, полёт и посадка которых проходит в зонах возможных пусков СП. На этих этапах, необходимо обеспечить их защиту.

Для защиты ЛА от СП с оптическими головками самонаведения (ОГС) реализованы комплексы индивидуальной защиты (КИЗ) [1] и [2]. В системе [1] для защиты ЛА от управляемых ракет с инфракрасными (ИК) головками самонаведения определяют факт пуска одной или нескольких ракет, генерируют лазерное излучение посылают в точку нахождения СП. Благодаря этому СП получает ложную информацию о местонахождении цели. Вышеуказанные процедуры повторяются для каждого последующего СП. Важнейшую роль при реализации такого способа играет подсистема информационного обеспечения, включающая датчики обнаружения в секторах обзора и высокоточную гиростабилизированную систему сопровождения СП по факелу двигателя на основе матричного фотоприемника ИК диапазона. Подсистема информационного обеспечения позволяет определять угловые координаты атакующего СП с точностью, достаточной для наведения источника лазерных помех с малой расходимостью. Такая система предназначена для индивидуальной защиты ЛА от СП с ИК головкой самонаведения и наиболее эффективна при одиночных и повторных пусках. Возникает трудность при одновременных пусках с разных ракурсов. Информация, получаемая от КИЗ, не включает сведений о дальности до точки пуска, необходимой для распределения временного ресурса на постановку помех. Полагаем, что в составе группы ЛА есть минимум два ЛА, оснащенных КИЗ. Тогда, одним из возможных вариантов получения такой информации является вычисление дальности пуска на этапе обнаружения факта пуска методом двухпозиционного пеленгования и дальнейшей передаче этих данных всем ЛА в группе. Для оценки возможности применения такого способа необходимо оценить среднее квадратическое отклонение (СКО) погрешностей определения дальности пуска.

1. Постановка задачи

Рассмотрим один из вариантов полета группы ЛА в зоне возможных пусков СП. В состав группы включены два ЛА, оснащенные КИЗ, решающие задачу информационного обеспечения и защиты [3]. Они обнаруживают факты пуска СП, оповещают экипажи всех ЛА в группе о факте пуска СП. С помощью навигационных данных о параметрах полёта

ЛА, оснащенных КИЗ $\left\{X_{\text{КИЗ}_{1,2}}, Y_{\text{КИЗ}_{1,2}}, Z_{\text{КИЗ}_{1,2}}\right\}$, и ЛА в составе основной группы $\left\{X_{\text{ЛА}_{1,2,3,4}}, Y_{\text{ЛА}_{1,2,3,4}}, Z_{\text{ЛА}_{1,2,3,4}}\right\}$, начальной базы пеленгования (расстояния между ЛА, оснащенными КИЗ) B_0, углов пеленгования СП в горизонтальной и вертикальной плоскости $\theta_{1,2}$, ракурсе атаки СП α осуществляется определение дальности до СП $D_{1,2}$ и расчет дальности относительно атакуемых ЛА по результатам двухпозиционного пеленгования $D_{\text{п}}$ путем усреднения результатов измерения. Геометрическая постановка задачи представлена на рис. 1.

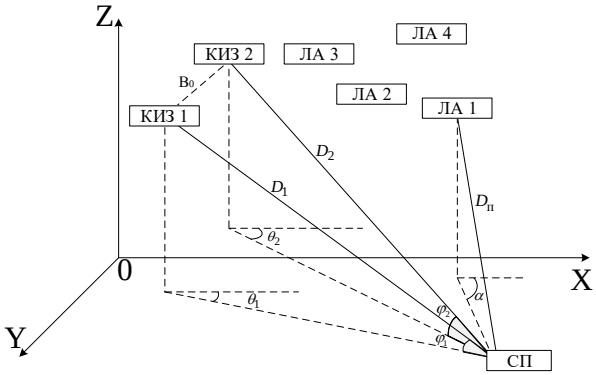


Рис. 1. Геометрическая постановка задачи

СКО погрешностей определения дальности пуска $D_{1,2}$ ЛА, оснащенными КИЗ, примем равными σ_{D_1} и σ_{D_2}. На основе имеющихся данных о параметрах полета и пусках СП значения этих СКО погрешностей каждым ЛА, без учета индексов, вычислим с помощью выражения (1) [4]:

$$\sigma_{\text{D}} = \frac{1}{D}\sqrt{P + T} \qquad (1)$$

где

$$P = 2\sigma_0^2\left(\begin{pmatrix} X_{\text{КИЗ}} - X_{\text{ЛА}} - \\ - Z_{\text{КИЗ}}\cot(\varphi)\cos(\theta) \end{pmatrix}^2 + \begin{pmatrix} Y_{\text{КИЗ}} - Y_{\text{ЛА}} - \\ - Z_{\text{КИЗ}}\cot(\varphi)\cos(\theta) \end{pmatrix}^2\right) +$$
$$+ Z_{\text{КИЗ}}^2\sigma_0^2 + \begin{pmatrix} \cot(\varphi)\cos(\theta)(X_{\text{КИЗ}} - X_{\text{ЛА}} - Z_{\text{КИЗ}}\cot(\varphi)\cos(\theta)) + \\ + \cot(\varphi)\sin(\theta)(Y_{\text{КИЗ}} - Y_{\text{ЛА}} - Z_{\text{КИЗ}}\cot(\varphi)\cos(\theta)) \end{pmatrix}^2 \sigma_0^2 \quad (2)$$

$$T = \begin{pmatrix} Z_{\text{КИЗ}}\cot(\varphi)\sin(\theta)(X_{\text{КИЗ}} - X_{\text{ЛА}} - Z_{\text{КИЗ}}\cot(\varphi)\cos(\theta)) + \\ + \cot(\varphi)\cos(\theta)(Y_{\text{КИЗ}} - Y_{\text{ЛА}} - Z_{\text{КИЗ}}\cot(\varphi)\cos(\theta)) \end{pmatrix}^2 \sigma_{\text{уг}}^2 +$$
$$+ \begin{pmatrix} (Z_{\text{КИЗ}}\cos(\theta) + Z_{\text{КИЗ}}\cot^2(\varphi)\cos(\theta)) \\ (X_{\text{КИЗ}} - X_{\text{ЛА}} - Z_{\text{КИЗ}}\cot(\varphi)\cos(\theta)) + \\ + (Z_{\text{КИЗ}}\sin(\theta) + Z_{\text{КИЗ}}\cot^2(\varphi)\cos(\theta)) \\ (Y_{\text{КИЗ}} - Y_{\text{ЛА}} - Z_{\text{КИЗ}}\cot(\varphi)\cos(\theta) \end{pmatrix}^2 \sigma_{\text{уг}}^2 \quad , \quad (3)$$

где σ_0 - СКО погрешностей определения собственных координат, $\sigma_{\text{уг}}$ - СКО погрешностей определения угловых координат. Полагаем, что системы, установленные на ЛА, оснащенные КИЗ, идентичны.

На основе заданного пространственного положения и выражений для определения дальности пуска при однопозиционном пеленговании получим выражение для случая двухпозиционного пеленгования путем усреднения результатов измерений и проведем анализ СКО погрешностей для данного способа.

2. Основные результаты

Принимая по внимание тот факт, что с помощью ЛА, оснащенных КИЗ, выполняется определение одного параметра – дальности, погрешности измерений которых независимы друг от друга, можно уменьшить СКО погрешностей определения дальности за счёт усреднения результатов измерений [5].

Выражение для СКО погрешностей определения дальности при двухпозиционном пеленговании можно записать в виде:

$$\sigma_{\text{D}_\text{п}} = \sqrt{\left(\frac{1}{\sigma_{\text{D}_\text{п}1}^2} + \frac{1}{\sigma_{\text{D}_\text{п}2}^2}\right)^{-1}} \quad (4)$$

где σ_{D_1} - СКО погрешностей определения дальности до точки пуска с первого ЛА, оснащенного КИЗ, а σ_{D_2} - со второго.

Рассмотрим как влияют на СКО погрешностей определения дальности $\sigma_{D_п}$ погрешности определения собственных координат σ_0 и угловых координат СП $\sigma_{уг}$ при следующих исходных данных: $X_{КИЗ_1} = 0$ м, $Y_{КИЗ_1} = 0$ м, $Z_{КИЗ_1} = 300$ м, $X_{КИЗ_2} = 0$ м, $Y_{КИЗ_2} = 1000$ м, $Z_{КИЗ_2} = 300$ м , $X_{ЛА} = 1000$ м, $Y_{ЛА} = 0$ м, $Z_{ЛА} = 300$ м, $\alpha = 30$ град при типовых дальностях пуска СП 3, 4 и 5 км.

Графически, относительно одного атакуемого ЛА, заданные начальные условия представлены на рис. 2

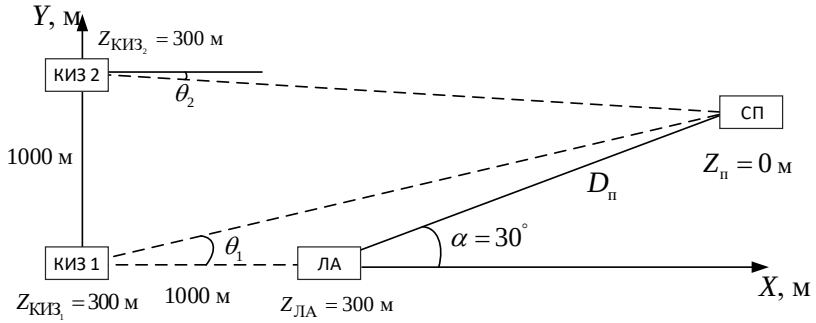


Рис. 2. Исходные данные для расчета

Ниже на рис. 3-5 показаны графики зависимостей СКО погрешностей определения дальности от параметров информационного обеспечения при различных дальностях пуска СП.

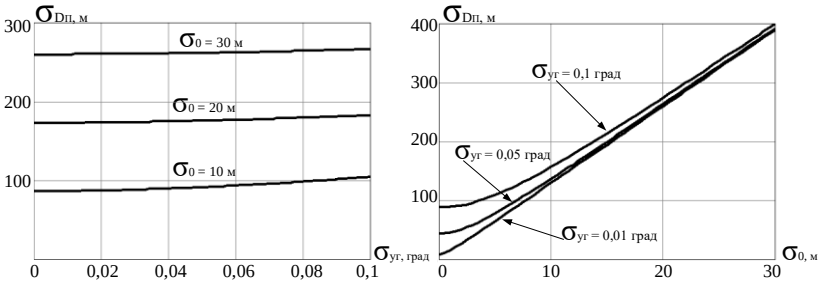


Рис. 3. Графики зависимости погрешностей определения дальности $\sigma_{D_п}$ от СКО погрешностей пеленгования $\sigma_{уг}$ и определения координат σ_0 , при дальности пуска $D_п = 3000$ м

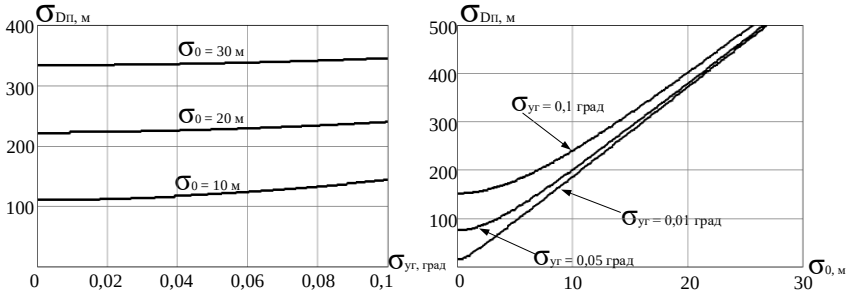


Рис. 4. Графики зависимости погрешностей определения дальности $\sigma_{D_п}$ от СКО погрешностей пеленгования $\sigma_{уг}$ и определения координат σ_0 , при дальности пуска $D_п$ = 4000 м

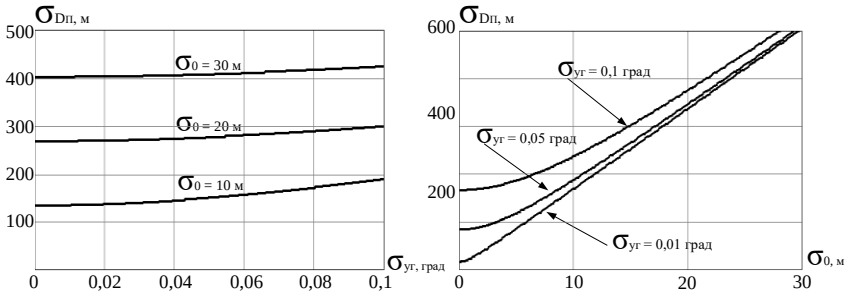


Рис. 5. Графики зависимости погрешностей определения дальности $\sigma_{D_п}$ от СКО погрешностей пеленгования $\sigma_{уг}$ и определения координат σ_0 , при дальности пуска $D_п$ = 5000 м

По графикам на рис. 3 ,4 и 5 можно сделать вывод о характере влияния СКО погрешностей определения пространственных координат вертолетов и угловых координат СП на СКО погрешностей определения дальности при различных дальностях пуска. Так, для $D_п$ = 3000 м (рис. 3) СКО погрешностей определения дальности $\sigma_{D_п}$ при $\sigma_{уг}$ =0,05 град и изменении ошибки определения пространственных координат σ_0 от 0 до 30 метров увеличивается с 28 до 267 метров. При σ_0 = 20 м и $\sigma_{уг}$ меняющейся в диапазоне от 0 до 0,1 град, $\sigma_{D_п}$ увеличивается с 177 до 186 м метров. При пуске ЗУР с дальности 5000 м (рис. 5) для вышеуказанных

значений параметров информационного обеспечения $\sigma_{уг}$ и σ_0 для каждого из графиков, значение СКО погрешностей определения дальности $\sigma_{D_п}$ увеличивается от 66 до 408 метров и от 268 до 300 метров соответственно. Анализ показал, что каждый из параметров информационного обеспечения влияет на значение погрешности определения дальности пуска. В большей степени влияние оказывает погрешность определения координат собственного местоположения, а погрешность определения угловых координат СП в рассматриваемом диапазоне значений вносит вклад в ошибку не более, чем на 11 %.

Рассмотрим влияние высоты вертолетов прикрытия на величину погрешности определения дальности пуска при двухпозиционном пеленговании. Пространственное положение вертолетов: $X_{КИЗ_1}$ = 0 м, $Y_{КИЗ_1}$ = 0 м, $Z_{КИЗ_1}$ = 0 - 1000 м, $X_{КИЗ_2}$ = 0 м, $Y_{КИЗ_2}$ = 1000 м, $Z_{КИЗ_2}$ = 0 - 1000 м, $X_{ЛА}$ = 1000 м, $Y_{ЛА}$ = 0 м, $Z_{ЛА}$ = 300 м, α = 30 град.

На рис. 6 представлены графики, связывающие СКО погрешностей определения дальности и высоту полета при различных дальностях пуска $D_п$ для погрешностей определения координат местоположения и угловых координат типовых оптических и навигационных систем, устанавливаемых на ЛА - σ_0 = 20 м, $\sigma_{уг}$ = 0,02 град.

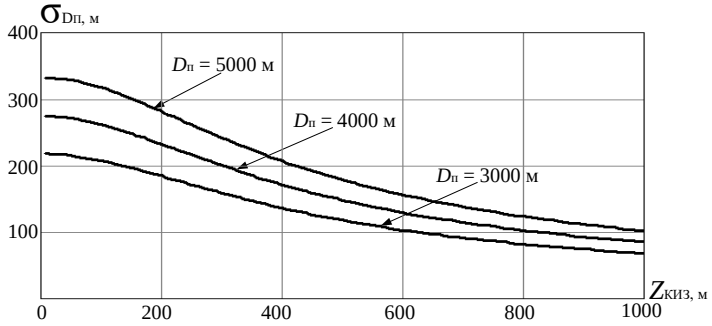


Рис. 6. Влияние высоты полета ЛА, оснащённых КИЗ, на СКО погрешностей определения дальности при типовых дальностях пуска СП

Из графиков на рис. 6 видно, что при увеличении высоты полета вертолетов прикрытия уменьшается СКО погрешностей определения дальности и наоборот. Так, при дальности пуска 3000 метров и

уменьшении высоты полета с 600 до 100 метров, СКО погрешностей определения дальности увеличивается от 100 до 210 метров. При дальности пуска 5000 метров – от 160 до 320 метров.

В целях проверки теоретических результатов разработана имитационная модель в среде MATLAB. Входными данными для модели являются: пространственные координаты ЛА, истинная дальность до СП, направление на СП от атакуемого ЛА, погрешности измерений координат ЛА и угловых координат СП, база пеленгования. Для получения статистических данных об СКО ошибок определения дальности проводилось 50000 испытаний, для каждого из которых использовались пространственные координаты с ошибками в пределах СКО, и также вычислялись угловые координаты СП. С использованием полученной совокупности статистических данных проведен анализ теоретических и полученных в результате моделирования значений СКО погрешностей определения дальности пуска при двухпозиционном пеленговании.

На рис. 7 представлен график зависимостей СКО погрешностей определения дальности от дальности пуска ЗУР.

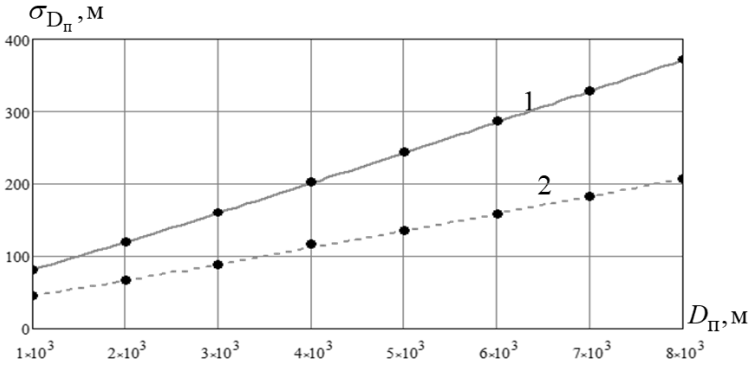


Рис. 7. График зависимостей СКО погрешностей определения дальности от дальности пуска СП

Кривая **1** получена при расчете для следующих исходных данных: $X_{КИЗ_1} = 0$ м, $Y_{КИЗ_1} = 0$ м, $Z_{КИЗ_1} = 300$ м, $X_{КИЗ_2} = 0$ м, $Y_{КИЗ_2} = 1000$ м, $Z_{КИЗ_2} = 300$ м, $X_{ЛА} = 1000$ м, $Y_{ЛА} = 0$ м, $Z_{ЛА} = 300$ м, $\alpha = 30$ град. При расчете погрешности определения дальности $\sigma_{D_п}$ использовались известные параметры оптических и навигационных систем, устанавливаемых на

ЛА: $\sigma_{уг}$ = 0,02 град, σ_0 = 20 м. Кривая **2** характеризует зависимость $\sigma_{D_п}$ от дальности пуска СП при том же пространственном положении ЛА, но для КИЗ с модернизированными параметрами информационного обеспечения: $\sigma_{уг}$ =0,01 град, σ_0 =10 м. Точками показаны значения погрешностей, полученных в результате имитационного моделирования.

Анализируя график на рис. 7, можно сделать вывод, что для типовых параметров информационного обеспечения ЛА погрешность определения дальности до точки пуска СП в диапазоне дальностей с 1 до 8 км и высоте полета 300 метров увеличивается от 80 до 370 м. В сравнении с этим, для модернизированного комплекса СКО погрешностей будет составлять от 45 до 200 м, что более чем в 1,8 раз точнее. Так отметим, что расхождение между СКО погрешностей, полученных в результате моделирования и теоретических расчетах, не превышает 1,1%, что подтверждает сходимость полученных результатов

Таким образом, использование данных о дальности на этапе обнаружения факта пуска СП позволит ранжировать угрозы по дальности и в дальнейшем на этапе сопровождения и постановке помех распределить циклограммы работы КИЗ, с целью эффективной защиты группы ЛА.

Заключение

На основе полученного с использованием основных положений теории ошибок измерений и цифровой обработки радиолокационной информации выражения, выполнен анализ СКО погрешностей определения дальности пуска СП при двухпозиционном пеленговании. Процедура расчета отличается тем, что расчет СКО погрешностей определения дальности пуска СП по защищаемым ЛА проводится с помощью системы пеленгования ЛА, оснащенных КИЗ. Анализ СКО погрешностей показал, что каждый параметр информационного обеспечения влияет на результирующие значение погрешностей. В большей степени влияние оказывает точность определения пространственных координат в рассматриваемом диапазоне значений каждого из параметров. Так же для достижения более точного определения дальности необходимо, чтобы ЛА, оснащенные КИЗ, осуществляли полет выше. Так, для высоты полета 800 метров в среднем в диапазоне возможных пусков СП от 1 до 8 км СКО погрешностей будет равно 100-120 метров, что учитывая высокую скорость СП и ограниченный интервал времени на сопровождение СП, позволит более точно распределить циклограммы работы КИЗ на постановку помех и выполнение защитных действий атакуемыми ЛА.

Список литературы

1. Патент 2511513C2 Российская Федерация, МПК 7 F41H, 11/02. Способ и система защиты воздушных судов от ракет переносных зенитных ракетных комплексов / В. В. Бутузов, С. Д. Великанов, С. Г. Гаранин, В. П. Иванов, А. В. Кислецов, В. С. Яцык ; заявитель и патентообладатель ОАО «НИИ «Экран». – № 2012135335/11 ; заявл. 16.08.12 ; опубл. 10.04.14, Бюл. № 10.

2. Патент 96553U1 Российская Федерация, МПК B64D 45/00. Бортовой комплекс индивидуальной защиты летательного аппарата от управляемых ракет с инфракрасными головками самонаведения / В. В. Бутузов, С. Н. Вагонов, А. И. Кобзарь, В. С. Казаков ; заявитель и патентообладатель ЗАО «СКБ «Зенит», ФГУП «НИИ «Экран». – № 2010113325/11 ; заявл. 07.04.10 ; опубл. 10.08.2010, Бюл. № 22.

3. Патент 2851131C1 Российская Федерация, МПК F41H 11/02. Способ защиты группы летательных аппаратов от управляемых боеприпасов с оптическими головками самонаведения/ А. А. Донцов, Д. В. Смирнов, С. С. Кущев, П. Е. Кулешов; заявитель и патентообладатель ВУНЦ ВВС «ВВА», №2025106125 ; заявл. 14.03.25 ; опубл. 19.11.25, Бюл. №32.

4. Свид. №2025685000 Российская Федерация. Свидетельство об официальной регистрации программы для ЭВМ. Программа расчета погрешностей определения дальности до объекта при однопозиционном пеленговании объекта матричными фотоприемниками / заявитель и правообладатель Смирнов Д.В. (RU). – №2025684406 ; заявл.16.09.25 ; опубл. 19.09.2025, Реестр программ для ЭВМ. – 1 с.

5. Фарина, А. Цифровая обработка радиолокационной информации. Сопровождение целей / А. Фарина, Ф. Студер ; пер. с англ. А. М. Бочкарева ; под ред. А. Н. Юрьева. М. : Радио и связь, 1993. – 319 с.

Алгоритм определения параметров отраженного зондирующего лазерного излучения от винтов беспилотного летательного аппарата малого класса

А. В. Чёнгин, e-mail: alexsvist168@mail.ru
А. Ю. Козирацкий, e-mail: aleksandrkozirackiy@mail.ru

ВУНЦ ВВС «ВВА» им. профессора Н. Е. Жуковского и Ю. А. Гагарина (г. Воронеж)

Аннотация. *В работе представлен обобщённый статистически оптимальный подход к совместной оценке амплитуды сигнала и частоты биений, возникающих при когерентном гетеродинном зондировании вибрирующих лопастей квадрокоптеров. Алгоритм основан на синтезе метода максимального правдоподобия для обработки сигналов и энергетической модели лазерного локационного сигнала. Выведены последовательные соотношения для оценивания параметров сигнала и аналитические выражения для их дисперсий, в условиях высокого отношения сигнал/шум. Особое внимание уделено решению обратной задачи: определению минимально необходимой энергии зондирующего сигнала, обеспечивающей требуемые вероятности правильного обнаружения и ложной тревоги, а также заданную точность измерения параметров вибрации.*

Ключевые слова: *вибрация лопастей, квадрокоптер, метод максимального правдоподобия, оценка параметров сигнала.*

Введение

Дистанционное обнаружение, распознавание и сопровождение малоразмерных беспилотных летательных аппаратов (БЛА) противника является сложной задачей современной противовоздушной обороны. Одним из информативных признаков для их выявления служит вибрационное поле, создаваемое вращающимися несущими винтами. Лазерное гетеродинное зондирование представляет собой эффективный метод бесконтактного и скрытного контроля таких вибраций. Измеряемые параметры — амплитуда и частота лазерных биений — несут информацию о типе роторной системы и её режиме работы, что позволяет идентифицировать цель и оценить её состояние, формируя основу для принятия решений по противодействию [1].

© Чёнгин А. В., Козирацкий А. Ю., 2026

1. Вероятностная модель наблюдаемого сигнала и вывод функционала правдоподобия

В рамках одного такта измерения, длительность которого $T_{\text{и}}$ определяется временем наблюдения за одной лопастью (предпочтительно охватывающим несколько периодов вращения), выходной сигнал фотоприёмного устройства моделируется как:

$$x(t) = S(t, A_{\text{с}}, f_{\text{г}}, f_{D_{\max}}) + n(t), \quad 0 \le t \le T_{\text{и}}, \tag{1}$$

– $x(t)$ – наблюдаемая реализация;

– $S(t, A_{\text{с}}, f_{\text{г}}, f_{D_{\max}})$ – полезный сигнал;

– $n(t)$ – аддитивный белый гауссовский шум;

Полезный сигнал, возникающий при смешении отражённой от вращающейся лопасти и опорной волн, имеет вид колебания с синусоидальной частотной модуляцией (ЧМ) и амплитудной модуляцией, описываемой огибающей $S_0(t)$:

$$x(t) = S(t, A_{\text{с}}, f_{\text{г}}, f_{D_{\max}}) = A_{\text{с}} S_0(t) \cos[2\pi f_{\text{с}} t + \beta \sin(2\pi f_{\text{г}} t + \varphi_0) + \phi], \tag{2}$$

– $A_{\text{с}}$ – амплитуда биений сигнала;

– $S_0(t)$ – нормированная огибающая сигнала;

– $f_{\text{с}}$ – промежуточная (несущая) частота биений;

– β – индекс частотной модуляции;

– $f_{\text{г}}$ – частота вращения лопасти;

– $f_{D_{\max}}$ – максимальный доплеровский сдвиг;

– φ_0 – начальная фаза вращения;

ϕ – случайная начальная фаза сигнала;

Для применения метода максимального правдоподобия (ММП) необходимо получить функционал правдоподобия $L(\theta; x(t))$, где $\theta = (A_{\text{с}}, f_{\text{г}}, f_{D_{\max}})^{\text{T}}$ - вектор оцениваемых параметров. Усредняя по равномерно распределенной случайной фазе ϕ, получаем известное выражение для сигнала со случайной начальной фазой [2]:

$$L(A_{\text{с}}, f_{\text{г}}, f_{D_{\max}}) = C \exp\left(-\frac{A_{\text{с}}^2 E_{\text{ог}}}{2N_0}\right) I_0\left(\frac{2A_{\text{с}} Z(f_{\text{г}}, f_{D_{\max}})}{N_0}\right), \tag{3}$$

– C – константа, не зависящая от оцениваемых параметров;

– $I_0(\bullet)$ – функция Бесселя первого рода нулевого порядка;

– $E_{\text{ог}} = \int_0^{T_{\text{и}}} S_0^2(t)dt$ – энергия нормированной огибающей сигнала;

– $Z(f_{\text{г}}, f_{D_{\max}})$ – огибающая нелинейного корреляционного интеграла;

В отличии от случая с постоянной частотой, опорный сигнал теперь должен повторять структуру ЧМ:

$$Z(f_{\text{г}}, f_{D_{\max}}) = \left| \int_0^{T_{\text{и}}} x(t) S_0(t) \exp\left[-j\beta \sin(2\pi f_{\text{г}} t + \hat{\varphi}_0) e^{-j2\pi f_{\text{с}} t} dt \right] \right|. \qquad (4)$$

Здесь $\hat{\varphi}_0$ – оценка начальной фазы вращения, которую также необходимо найти, что делает задачу четырёхпараметрической. На практике для упрощения часто ищут максимум Z по $f_{\text{г}}, f_{D_{\max}}$ и f_0, что эквивалентно согласованной фильтрации с набором опорных сигналов, имеющих различную частоту и глубину ЧМ.

2. Алгоритм совместной оценки параметров методом максимального правдоподобия

Оценки максимального правдоподобия $(\hat{A}_{\text{с}}, \dot{f}_{\text{г}}, \dot{f}_{D_{\max}})$ формально определяются как аргументы глобального максимума функционала правдоподобия. Удобнее работать с его логарифмом:

$$\Lambda(A_{\text{с}}, f_{\text{г}}, f_{D_{\max}}) = \ln I_0\left(\frac{2A_{\text{с}} Z(f_{\text{г}}, f_{D_{\max}})}{N_0} \right) - \frac{A_{\text{с}}^2 E_{\text{ог}}}{2N_0}, \qquad (5)$$

Непосредственное совместное решение этой задачи аналитически невозможно. Однако структура функционала позволяет предложить эффективный квазиоптимальный двухэтапный алгоритм, основанный на свойстве монотонного возрастания $I_0(x)$.

Оценки параметров ЧМ находятся независимо от амплитуды:

$$(f_{\text{г}}, f_{D_{\max}}) = \arg \max_{f_{\text{г}}, f_{D_{\max}}} Z(f_{\text{г}}, f_{D_{\max}}). \qquad (6)$$

Обозначим максимальное достигнутое значение:

$$Z_{\max} = Z(f_{\text{г}}, f_{\text{D}_{\max}}). \qquad (7)$$

После определения оценок $f_{\text{г}}, f_{D_{\max}}$ задача становится одномерной [3]. Оценка амплитуды находится из условия $\partial\Lambda / \partial A_{\text{с}} = 0$. Используя свойство $\partial I_0(x) / \partial x = I_1(x)$, получаем неявное уравнение:

$$\frac{2Z_{\max}}{N_0}\cdot\frac{I_1\left(\frac{2\hat{A}_c Z_{\max}}{N_0}\right)}{I_0\left(\frac{2\hat{A}_c Z_{\max}}{N_0}\right)}-\frac{A_c^2 E_{or}}{2N_0}=0. \qquad (8)$$

В практически важном случае высокого отношения сигнал/шум, когда аргумент функций Бесселя велик, справедливо приближение $I_1(x)/I_1(x)\approx 1$. Тогда уравнение принимает явный линейный вид:

Эта формула имеет ту же интерпретацию, что и в случае вибрации: оценка амплитуды пропорциональна максимальной корреляции $Z_{\max}$, достигнутой при оптимальном согласовании с моделью ЧМ-сигнала.

3. Анализ потенциальной точности оценок

Для формирования требований к энергетике системы необходимо определить теоретически достижимую точность оценок параметров вращения. Нижние границы для дисперсий любых несмещённых оценок устанавливаются неравенством Крамера-Рао через обратную матрицу Фишера J^{-1}, элементы которой: $J_{ij}=-E[\partial^2\Lambda/\partial\theta_i\partial\theta_j]$, где $\theta=(A_c, f_r, f_{D_{\max}})^T$ [4].

В условиях высокого ОСШ для сигнала с синусоидальной ЧМ дисперсии оценок стремятся к следующим границам. Дисперсия оценки амплитуды $\sigma(\hat{A}_c)$:

$$\sigma(\hat{A}_c)\geq\frac{N_0}{E_{or}}. \qquad (9)$$

Дисперсия оценки амплитуды $\sigma(\hat{A}_c)$:

$$\sigma(\dot{f}_r)\geq\frac{3}{2\pi^2\frac{E_c}{N_0}T_и^3}\cdot\frac{1}{g(\beta)}, \qquad (10)$$

- E_c/N_0 – отношение энергии сигнала к спектральной плотности шума;
- $E_c\approx A_c^2E_{or}/2$;
- $g(\beta)$ – функция, зависящая от индекса модуляции;
- $Z(f_r, f_{D_{\max}})$ – огибающая нелинейного корреляционного интеграла;

Дисперсия оценки максимального доплеровского сдвига $\dot{f}_{D_{\max}}$:

$$\sigma(\dot{f}_{D_{\max}}) \geq \frac{3}{2\pi^2 \frac{E_c}{N_0} T_и^3 \tau_{эфф}} \cdot h(\beta), \tag{11}$$

– $\tau_{эфф}$ – эффективная длительность зондирующего импульса;

Точность оценки $f_{D_{\max}}$ (а значит, и тангенциальной скорости) улучшается с ростом времени измерения $T_и$ и эффективной длительности импульса $\tau_{эфф}$, но не так быстро, как точность оценки $f_г$.

Для обеспечения заданной точности измерений $\sigma_{f_г}^{треб}$, $\sigma_{f_{D_{\max}}}^{треб}$ необходимо, чтобы выполнялись неравенства:

$$\frac{3}{2\pi^2 \frac{E_c}{N_0} T_и^3 g(\beta)} \leq (\sigma_{f_г}^{треб})^2, \quad \frac{3}{2\pi^2 \frac{E_c}{N_0} T_и h(\beta)} \leq (\sigma_{f_{D_{\max}}}^{треб})^2. \tag{12}$$

4. Статистическая модель обнаружения сигнала в лазерном локационном канале

Задача обнаружения факта наличия сигнала от вращающейся лопасти на фоне шумов решается в рамках теории статистического обнаружения. Оптимальный обнаружитель при случайной начальной фазе основан на сравнении статистики $Z_{\max}^2$ (или её монотонной функции) с порогом h [5].

Для случая накопления $n_н$ независимых измерений (импульсов) вероятности правильного обнаружения P_D и ложной тревоги P_{FA} связываются приближённым соотношением через интегральную функцию Лапласа $\Phi(z)$:

$$P_D = \Phi\left\{\sqrt{\frac{S_ф n_н}{1+S_ф}} - \frac{1}{\sqrt{1+S_ф}} \Phi^{-1}(1-P_{FA})\right\}, \tag{13}$$

– $S_ф = s/b$ – отношение сигнал/шум в фотонных единицах;

– s – среднее число сигнальных фотоэлектронов за импульс;

– b – среднее число фоновых фотоэлектронов за импульс;

– $n_н$ – количество накапливаемых импульсов;

Разрешая это уравнение относительно s, получаем неявное выражение для требуемого числа сигнальных фотоэлектронов:

$$s = \frac{1+S_{\text{ф}}}{S_{\text{ф}} n_{\text{н}}}\left[\Phi^{-1}(P_{\text{D}}) + \frac{\Phi^{-1}(1-P_{\text{FA}})^2}{\sqrt{1+S_{\text{ф}}}}\right], \tag{14}$$

Это уравнение решается итерационно относительно s, поскольку $S_{\text{ф}} = s/b$. Обозначим решение как $s_{\text{треб}}(P_{\text{D}}, P_{\text{FA}}, n_{\text{н}}, b)$.

5. Энергетическая модель формирования фотонного сигнала в лидарном канале

Число сигнальных фотоэлектронов s связано с энергией зондирующего импульса $E_{\text{т}} = P_{\text{пик}}\tau$ уравнением лидарного канала для точечной цели:

$$s = \frac{\eta_{\text{q}} E_{\text{т}} \lambda A_{\text{a}} T_{\text{атм}}^2 t_{\text{опт}} \langle \sigma_{\text{ЭПР}} \rangle}{4\pi h c \Omega_p L^4}, \tag{15}$$

- η_{q} – квантовая эффективность фотоприёмника;
- h – постоянная планка;
- c – скорость света;
- A_{a} – эффективная площадь приёмной апертуры;
- Ω_p – телесный угол расходимости лазерного луча;
- L – дальность до цели;
- $T_{\text{атм}}$ – коэффициент одностороннего пропускания атмосферы;
- $t_{\text{опт}}$ – коэффициент пропускания приёмной оптики;
- $\langle \sigma_{\text{ЭПР}} \rangle$ – средняя эффективная площадь рассеяния лопасти;

6. Определение минимально необходимой энергии импульса

Теперь мы можем синтезировать все компоненты для решения задачи: определения минимальной энергии импульса $E_{\text{т}}^{\text{треб}}$, обеспечивающей требуемые характеристики обнаружения и точности.

Из требований к точности (раздел 3) определяем минимально необходимое отношение сигнал/шум $(E_c/N_0)^{\text{треб}}$, необходимое для выполнения неравенств Рао-Крамера $\sigma_{f_{\text{r}}}^{\text{треб}}$ и $\sigma_{f_{D\max}}^{\text{треб}}$. Поскольку $E_{\text{с}} \infty s$, это задаёт нижнюю границу для числа фотоэлектронов $s_{\text{точн}}$.

Из требований к обнаружению (раздел 5) итерационно находим требуемое число фотоэлектронов $s_{\text{обн}} = s_{\text{треб}}(P_{\text{D}}, P_{\text{FA}}, n_{\text{н}}, b)$.

Итоговое требуемое число фотоэлектронов выбирается как максимальное из двух $s_{\text{итог}} = \max(s_{\text{обн}}, s_{\text{точн}})$. На практике для

вращающихся целей с большим временем наблюдения часто ограничивающим фактором является именно обнаружение, а не точность оценки частоты.

Подставляем $s_{\text{итог}}$ в уравнение лидарного канала и разрешаем его относительно $E_{\text{т}}$:

$$E_{\text{т}}^{\text{треб}} = P_{\text{пик}}\tau = \frac{4\pi\Omega_p L^4 hc}{\eta_{\text{q}}\lambda A_{\text{a}} T_{\text{атм}}^2 t_{\text{опт}} \left\langle \sigma_{\text{ЭПР}} \right\rangle} s_{\text{итог}}. \qquad (16)$$

Заключение

В проведенном исследовании разработана модель сигнала с синусоидальной ЧМ, описывающая отражение от вращающейся лопасти. Предложен квазиоптимальный двухэтапный алгоритм оценки амплитуды, частоты вращения и максимального доплеровского сдвига. Выполнен синтез алгоритма оценки с моделями обнаружения и энергетики лазерного канала. Выведено ключевое расчётное соотношение для определения минимальной энергии зондирующего импульса по заданным требованиям к надёжности и точности.

Литература

1. Чёнгин, А. В. Методический подход к энергетическому расчету пространственно-распределенной системы сложения мощностей когерентных источников оптического диапазона длин волн / А. В. Чёнгин, А. Ю. Козирацкий, О. В. Смынтына // Радиолокация, навигация, связь : сб. тр. XXXI Международной научно-технической конференции (г. Воронеж, 15–17 апреля 2025 г.) : в 6 т. / Воронежский государственный университет ; АО «Концерн “Созвездие”». – Воронеж : Издательский дом ВГУ, 2025. – Т. 3. – С. 8–15.
2. Борейшо, А. С. Лазеры: устройство и действие : учебное пособие / А. С. Борейшо, С. В. Ивакин. – СПб. : Изд. «Лань», 2016. – 304 с.
3. Либенсон, М. Н. Взаимодействие лазерного излучения с веществом (силовая оптика). В 2 ч. Ч. II. Лазерный нагрев и разрушение материалов : учебное пособие / М. Н. Либенсон, Е. Б. Яковлев, Г. Д. Шандыбина ; под общ. ред. В. П. Вейко. – СПб. : НИУ ИТМО, 2014. – 181 с.
4. Беспризованных, В. Г. Нелинейная оптика : учеб. пособие / В. Г. Беспризованных, В. П. Первадчук. – Пермь : Изд-во Пер. гос. Техн. Ун-та, 2011. – 200 с.
5. Оптико-электронные системы авиационного вооружения: учебник для слушателей и курсантов ВВУЗов ВВС / А. М. Краснов [и др.] ; под ред. А. М. Краснова. – М. : изд. ВВИА им. проф. Н. Е. Жуковского, 2007. – 1272 с.

Секция 10
Информационные технологии, вычислительный анализ и методы математической физики

Компьютерный эксперимент по фильтрации в волноводе на основе акустической голографии

В. А. Башкарев, email: bashkarev.edu@yandex.ru [1]
С. А. Пересёлков, email: pereselkov@yandex.ru [1]
В. М. Кузькин, email: kumiov@yandex.ru [2]
А. С. Пересёлков, email: pereselkov@yandex.ru [1]

[1] Воронежский государственный университет
[2] Институт общей физики им. А.М. Прохорова РАН

Аннотация. *Рассмотрен алгоритм построения высокоэффективного фильтра гидроакустических сигналов на основе методов интерферометрической обработки. Показана эффективность такого фильтра в условиях высокого уровня аддитивных помех. Методом численного моделирования оценена эффективность фильтра для широкого диапазона параметров приемника (частот и времени наблюдения) и источника (дистанций и значений скорости) при различных значениях отношения сигнал/помеха. На основе полученных результатов показана эффективность подхода фильтрации полезного сигнала при существенно неблагоприятной помеховой обстановке.*

Ключевые слова: *фильтрация сигналов, гидроакустический сигнал, интерферометрический подход, интерферограмма, двумерное преобразование Фурье, голограмма, функция углового распределения спектральной плотности.*

Введение

Интерферометрическая обработка сигналов (ИОС) [1–6] – перспективный метод подводной акустики, альтернативный согласованному полю [7–10]. Он основан на формировании устойчивой интерференционной картины широкополосного излучения в океаническом волноводе [1,2]. Интерференционные структуры, возникающие при взаимодействии нормальных мод, проявляются в частотно-пространственном и временно-частотном представлениях. Ключевые теоретические основы заложены С.Д. Чупровым, введшим понятие волноводного инварианта [1]; дальнейшие исследования [3–6] расширили потенциал ИОС в задачах анализа звукового поля и параметров среды. Одним из наиболее развитых направлений является голографическая обработка сигналов (ГОС) [7–35]. Теория формирования голограмм представлена в [7–11]; в [9] метод

© Башкарев В. А., Пересёлков С. А., Кузькин В. М., Пересёлков А. С. 2026

распространён на шумовые источники, в [10,11] — на низкие частоты в регулярных волноводах. Работы [12,13] посвящены повышению разрешения, [14–21] — устойчивости при внутренних волнах, [22–24] — адаптивным вариантам, [25–27] — применению в антеннах. Модовая селекция в мелководье рассматривалась в [28–30]. Большинство исследований [7–30] ограничивались скалярным давлением в низкочастотной области. Однако современные задачи требуют регистрации также компонент скорости звука, что ведёт к использованию векторно-скалярных приёмников (ВСП) в высокочастотном диапазоне [31–35], существенно расширяющих информативность и точность восстановления параметров источников и среды.

Подходы к построению эффективных фильтров для выделения информационных составляющих или подавления помех разнообразны. Априорная информация об особенностях сигнала и помех позволяет проектировать наиболее эффективные фильтры. Применение интерферометрического подхода позволяет усилить полезную компоненту входного сигнала, она коррелирована и накапливается квазикогерентно, и подавить помеховую из-за ее некоррелированности.

В настоящей работе рассматривается подход к построению полосового либо режекторного фильтра, дается оценка эффективности работы для случая регулярного океанического волновода по критерию улучшения отношения сигнал/помеха на выходе фильтра по отношению ко входу. В основу работы фильтра положен интерферометрический подход с использованием априорной информации о характере положения фокальных пятен, мест локализации энергии полезного сигнала на голограмме – двумерной спектральной плотности интерферограммы сигнала.

Формирование интерферограммы (интерференционной картины гидроакустического сигнала) в координатах время-частота:

$$I(f,t) = P(f,t)P^{*}(f,t), \qquad (1)$$

где $P(f,t)$ – давление звукового поля в точке приёма.

Получение голограммы (двумерной спектральной плотности $I(f,t)$) $F(\tau,\nu)$ применением к интерферограмме двумерного преобразования Фурье:

$$F(\tau,\nu) = \int_0^{\Delta t}\int_{f_1}^{f_2} \tilde{I}(f,t)\exp[i2\pi(\nu t - f\tau)]dfdt \qquad (2)$$

Здесь ν и τ – частота и время голограммы, соответствующие времени t и частоте f на интерферограмме. На голограмме особенности

характера накопления полезного сигнала и помеховой составляющей проявляются в формировании групп максимумов спектральной плотности в виде фокальных пятен, расположенных в узкой полосе, проходящей через начало координат.

Расчёт функции углового распределения спектральной плотности голограммы (функция обнаружения) $G(\chi)$, которая представляет собой зависимость интегрального значения спектральной плотности, рассчитанного вдоль луча, исходящего из начала координат, в зависимости от тангенса угла этого луча:

$$G(\chi) = \int_0^{\Delta\tau} |F(\tau, \chi\tau)| d\tau \qquad (3)$$

По результатам анализа нормированной в диапазон (0, +1) функции обнаружения рассчитывается фильтр – это матрица из 0 и 1 для вырезания узкой полосы на голограмме. В полосу попадают фокальные пятна, которые являются продуктом квазикогерентного накопления энергии полезного сигнала, равномерно распределенной по всей голограмме, энергия помеховой составляющей отфильтровывается. Угол наклона полосы определяется по максимуму функции обнаружения, ширина полосы рассчитывается по ближайшим к максимуму значениям функции обнаружения, где ее значение снижаются до 0.5.

2. Численный эксперимент

Моделирование проводилось для регулярного океанического волновода с постоянным профилем скорости звука по глубине равным 1500 м/с. Использовались все возможные комбинации из следующих значений сигнальной и акустической обстановки.

Отношение сигнал/помеха (по мощности) на входе: 0.0001, 0.001, 0.01, 0.1, 1.0, 10.0.

Глубина волновода 50 м, приемника 48 м, источника 10 м.

Дистанция между приемником и источником: 3 км, 5 км, 10 км, 25 км, 50 км, 100 км.

Источник радиально удаляется от приемника с постоянной скоростью: 0 м/с, 2 м/с, 5 м/с, 10 м/с, 15 м/с, 20 м/с.

Время наблюдения (накопления): 30 с, 60 с, 120 с, 240 с, 600 с.

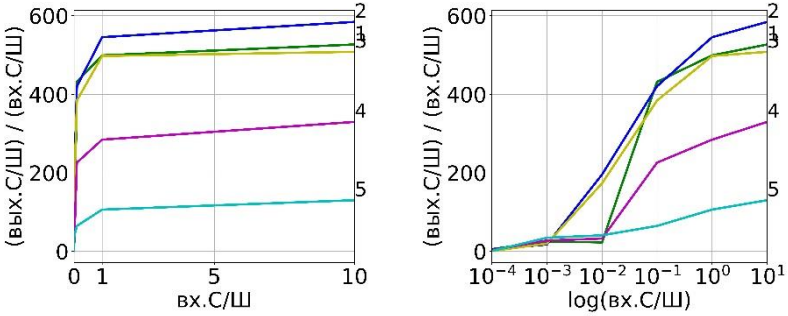
Полосы частот: 100-150 Гц, 300-350 Гц, 500-550 Гц, 700-750 Гц, 900-950 Гц.

На основе модового подхода [4] рассчитывалась интерферограмма полезного сигнала в точки приема. Исходя из множества значений отношения сигнал/помеха входного сигнала формировалась интерферограмма помеховой составляющей (как среднее по тридцати

реализациям случайного процесса). Суммарная интерферограмма входного сигнала использовалась для формирования голограммы и функции обнаружения, из которой рассчитывалась матрица фильтра. Матрица фильтра раздельно применялась к голограмме только полезного сигнала и голограмме только помеховой составляющей, которые обратным двумерным преобразованием Фурье возвращали нас к интерферограммам полезного сигнала и помеховой составляющей после фильтрации. Полученные таким образом интерферограммы использовались для расчета отношения сигнал/помеха после процедуры фильтрации и затем производилось сравнение с соответствующим значением отношения сигнал/помеха на входе фильтра.

3. Результаты численного эксперимента

Семейство кривых на рис. 1 демонстрирует влияние времени наблюдения на эффективность работы фильтра. На всех рисунках ниже (а) – линейный, (б) – логарифмический масштаб входного отношения сигнал/шум.



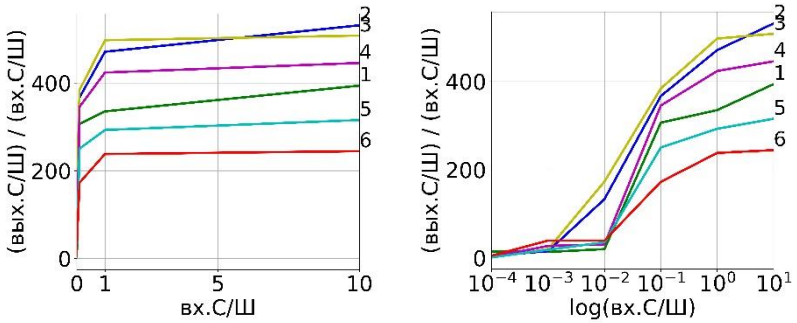
а б

Дистанция $r_0 = 10$ *км, скорость* $v = 5$ *м/с, частота* $\Delta f = 100 - 150$ *Гц.*
Время Δt *: кривая 1 – 30с, 2 – 60с, 3 – 120с, 4 – 240с, 5 – 600с.*

Рис. 1. Эффективность работы фильтра от отношения с/ш, зависимость от времени наблюдения

Универсальным значением времени наблюдения является интервал в 60 с. Значения 30 с и 120 с также обеспечивают хороший результат. Интервалы наблюдения в 240 с и 600 с дают худший результат, что объясняется эффектом «смазывания» интерференционной картины из-за движения источника. Зависимость эффективности работы фильтра от скорости источника представлены на рис. 2. Наблюдается ухудшение

качества работы фильтра с ростом скорости источника, при этом средние показатели эффективности достигаются при его неподвижности.



а б

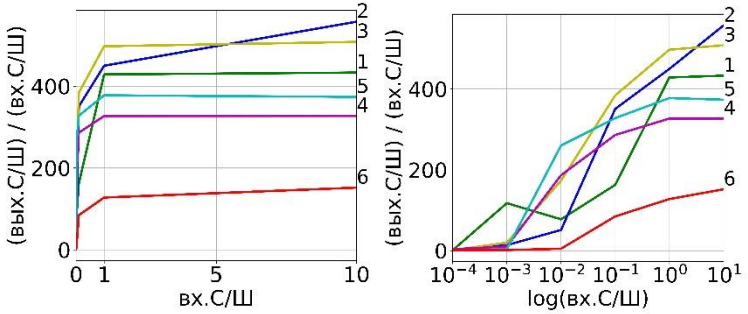
Дистанция $r_0 = 10$ *км, время наблюдения* $\Delta t = 120$ *с, частота* $\Delta f = 100 - 150$ *Гц. Скорость* v *: кривая 1 – 0м/с, 2 – 2м/с, 3 – 5м/с, 4 – 10м/с, 5 – 15м/с*

Рис. 2. Эффективность работы фильтра от отношения с/ш, зависимость от скорости источника

По рис. 3 можно выяснить качество работы предлагаемого фильтра в зависимости от дистанции между приемником и источником. Согласно него наблюдается кратное ухудшение работы фильтра с ростом дистанции.

Частотная закономерность качества работы фильтра представлена на рис. 4. Согласно него сделать однозначный вывод об эффективности использования определенного частотного диапазона не представляется возможным. Кривые, соответствующие различным полосам частот расположены достаточно плотно и многократно пересекаются. Для практики можно принять, что предлагаемый фильтр будет сопоставимо эффективно работать в широком диапазоне частот.

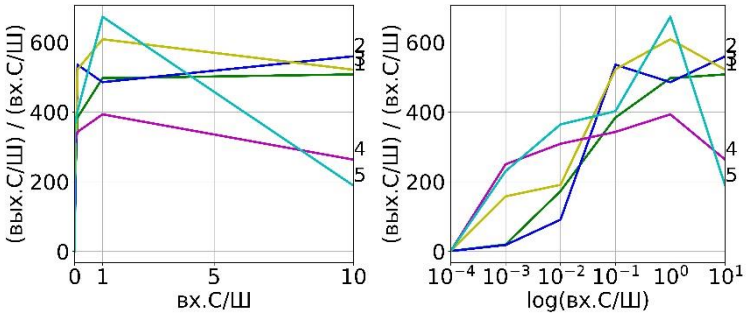
Каждый из представленных рисунков демонстрирует зависимость улучшения отношения сигнал/помеха на выходе фильтра по отношению к входному отношению в сотни раз, устойчивое улучшение наблюдается со значения отношения сигнал/помеха 0.1, а при некоторых сочетаниях параметров с 0.01.



а б

Скорость $v = 5$ *м/с, время наблюдения* $\Delta t = 120$ *с, частота* $\Delta f = 100 - 150$ *Гц. Дистанция* r_0 *: кривая 1 – 3км, 2 – 5км, 3 – 10км, 4 – 25км, 5 – 50км, 6 – 100 км*

Рис. 3. Эффективность работы фильтра от отношения с/ш, зависимость от начальной дистанции



а б

Дистанция $r_0 = 10$ *км, скорость* $v = 5$ *м/с, время наблюдения* $\Delta t = 120$ *с. Частота* Δf *: кривая 1 – 100-150Гц, 2 – 300-350Гц, 3 – 500-550Гц, 4 – 700-750Гц, 5 – 900-950Гц*

Рис. 4. Эффективность работы фильтра от отношения с/ш, зависимость от полосы частот

Полученные результаты подтверждают высокую эффективность работы фильтра. При меньших чем 0.001 входных отношениях сигнал/помеха улучшения не превышают десятков раз. Однако и такие

скромные результаты могут оказаться полезным при решении практических задач в крайне неблагоприятных помеховых условиях при отсутствии более эффективных способов обработки.

Заключение

Полученные в проведенной работе результаты моделирования демонстрируют высокую эффективность фильтрации входного сигнала с неблагоприятным уровнем сигнал/помеха на основе интерферометрического подхода. Для случая аддитивной гауссовской помехи показано, что при входном значении отношения сигнал/помеха по мощности равном или превышающем 0.1 наблюдается устойчивое улучшение отношения сигнал/помеха в сотни раз (от 100 до 400) для самого разнообразного сочетания параметров акустической обстановки.

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

Список литературы

1. Чупров СД. Интерференционная структура звукового поля в слоистом океане. В книге: Акустика океана. Современное состояние. Москва, Наука, 1982, стр. 71-91.
2. Weston D, Stevens K. Interference of wide-band sound in shallow water. J. Sound Vibr., 1972, 21(2):57-64.
3. Грачев Г. Теория инвариантов акустического поля в слоистых волноводах. Акуст. физ., 1993, 39(1):67-71.
4. Орлов ЕФ, Шаронов ГА. Интерференция звуковых волн в океане. Дальнаука, 1998.
5. Ianniello J. Recent developments in sonar signal processing. IEEE Signal Proc. Mag., 1998, 15(1):27-40.
6. Kuperman WA, D'Spain GL. Ocean acoustic interference phenomena and signal processing. In: Ocean Acoustic Interference Phenomena and Signal Processing; American Institute of Physics, San-Francisco, 2002.
7. Кузькин ВМ, Переселков СА, Кузнецов ГН. Спектрограмма и локализация источника звука в мелководном море. Акустический журнал, 2017, 63(4):406-418.
8. Кузькин ВМ, Переселков СА, Матвиенко ЮВ, Ткаченко СА. Обнаружение источника в диспергирующих средах. РЭНСИТ: Радиоэлектроника. Наносистемы. Информационные технологии, 2019, 11(3):337-344.

9. Kuz’kin VM, Pereselkov SA, Matvienko YV, Lyakhov GA, Tkachenko SA. Noise source detection in an oceanic waveguide using interferometric processing. Phys. Wave Phenom., 2020, 28(1):68-74.

10. Pereselkov SA, Kuz’kin VM. Interferometric processing of hydroacoustic signals for the purpose of source localization. J. Acoust. Soc. Am., 2022, 151:666-676.

11. Ehrhardt M, Pereselkov SA, Kuz’kin VM, Kaznacheev I, Rybyanets P. Experimental observation and theoretical analysis of the low-frequency source interferogram and hologram in shallow water. J. Sound Vibr., 2023, 544:117388.

12. Kuz’kin VM, Pereselkov SA, Kuznetsov GN, Grigor’ev VA. Resolving power of the interferometric method of source localization. Phys. Wave Phenom., 2018, 26(2):150-159.

13. Кузькин ВМ, Переселков СА, Грачев ВИ, Рыбянец ПВ, Ткаченко СА. Разрешение шумовых сигналов при наличии случайных неоднородностей океанической среды. РЭНСИТ: Радиоэлектроника. Наносистемы. Информационные технологии, 2024, 16(3):395-406.

14. Kuz'kin VM, Pereselkov SA, Zvyagin VG, Malykhin YuA, Prosovetskiy DYu. Intense internal waves and their manifestation in interference patterns of received signals on oceanic shelf. Physics of Wave Phenomena, 2018, 26(2):160-167.

15. Badiey M, Kuz’kin VM, Lyakhov GA, Pereselkov SA, Prosovetskiy DYu, Tkachenko SA. Intense internal waves and their manifestation in the interference patterns of received signals on oceanic shelf. Part II. Phys. Wave Phenom., 2019, 27(4):313-319.

16. Kaznacheeva ES, Kuz'kin VM, Pereselkov SA. Interferometric processing of hydroacoustic information in the presence of intense internal waves. Physics of Wave Phenomena, 2021, 29(3):278-284.

17. Kuz'kin VM, Pereselkov SA, Badiey M, Ladykin NV, Malykhin AYu, Tkachenko SA. Robustness of holographic processing of hydroacoustic signals in the presence of intense internal waves. Physics of Wave Phenomena, 2023, 31(5):346-354.

18. Badiey M, Kuz’kin VM, Pereselkov SA. Interferometry of hydrodynamics of oceanic shelf caused by intensive internal waves. Fundam. Appl. Hydrophys., 2020, 1:45-55.

19. Кузькин ВМ, Пересёлков СА, Казначеева ЕС, Грачев ВИ, Ткаченко СА, Рыбянец ПВ. Голографическая обработка движущихся источников в мелком море при наличии интенсивных внутренних волн. РЭНСИТ: Радиоэлектроника. Наносистемы. Информационные технологии, 2022, 14(2):197-204.

20. Кузькин ВМ, Пересёлков СА, Грачев ВИ, Ткаченко СА, Ладыкин НВ, Куцов МВ. Голографический метод локализации движущегося источника звука в присутствии интенсивных внутренних волн. РЭНСИТ: Радиоэлектроника. Наносистемы. Информационные технологии, 2023, 15(3):317-326.

21. Pereselkov S, Kuz'kin V, Ehrhardt M, Tkachenko S, Rybyanets P, Ladykin N. Three-dimensional modeling of sound field holograms of a moving source in the presence of internal waves causing horizontal refraction. J. Mar. Sci. Eng., 2023, 11:1922.

22. Кузькин ВМ, Матвиенко ЮВ, Пересёлков СА, Казначеева ЕС. Интерферометрическая обработка с использованием вертикальной линейной антенны. Вестник ВГУ. Серия: Физика. Математика, 2020, 2:14-23.

23. Кузькин ВМ, Пересёлков СА, Ткаченко СА, Казначеев ИВ. Реализация адаптивного интерферометрического метода локализации источника звука. II часть. Вестник ВГУ. Серия: Физика. Математика, 2020, 1:14-23.

24. Pereselkov S, Kuz'kin V, Lyakhov G, Tkachenko S, Kaznacheeva E. Adaptive Algorithms for Interferometric Processing. Phys. Wave Phenom., 2020, 28, 267-273.

25. Pereselkov S, Kuz'kin V, Lyakhov G, Tkachenko S, Kaznacheeva E. Adaptive algorithms for interferometric processing. Phys. Wave Phenom., 2020, 28(3):267-273.

26. Кузькин В.М., Пересёлков С.А., Ткаченко С.А., Казначеев И.В. Интерферометрическая обработка акустической информации с использованием протяженных антенн в диспергирующих средах. Радиоэлектроника. Наносистемы. Информационные технологии, 2020, 12(4):483-494.

27. Кузькин В.М., Пересёлков С.А., Матвиенко Ю.В., Ткаченко С.А., Просовецкий Д.Ю. Обнаружение шумового сигнала в океаническом волноводе горизонтальной антенной. РЭНСИТ: Радиоэлектроника. Наносистемы. Информационные технологии, 2022, 14(1):65-72.

28. Kuz'kin VM, Matvienko YuV, Pereselkov SA, Prosovetskii DYu, Kaznacheeva ES. Mode selection in oceanic waveguides. Physics of Wave Phenomena, 2022, 30(2):111-118.

29. Kuz'kin VM, Matvienko YuV, Pereselkov SA, Kaznacheeva ES, Tkachenko SA. Holographic method for mode selection in a shallow sea in the presence of intense internal waves. Physics of Wave Phenomena, 2022, 30(5):314-320.

30. Кузькин ВМ, Пересёлков СА, Казначеева ЕС, Грачев ВИ, Ткаченко СА, Рыбянец ПВ. Выделение мод шумового источника в мелком

море методом голографической интерферометрии в присутствии интенсивных внутренних волн. РЭНСИТ: Радиоэлектроника. Наносистемы. Информационные технологии, 2022, 14(3):279-286.

31. Казначеев ИВ, Кузнецов ГН, Кузькин ВМ, Пересёлков СА. Интерферометрический метод обнаружения движущегося источника звука векторно-скалярным приемником. Акуст. журн., 2018, 64(1):33-45.

32. Kuznetsov GN, Kuz'kin VM, Pereselkov SA, Kaznacheev IV, Grigor'ev VA. Interferometric method for estimating the velocity of a noise sound source and the distance to it in shallow water using a vector-scalar receiver. Phys. Wave Phenom., 2017, 25(4):299-306.

33. Kuz'kin VM, Pereselkov SA, Kuznetsov GN, Kaznacheev IV. Interferometric direction finding by a vector-scalar receiver. Physics of Wave Phenomena, 2018, 26(1):63-73.

34. Переселков СА, Кузькин ВМ, Кузнецов ГН, Просовецкий ДЮ, Ткаченко СА. Интерференционный метод оценки координат движущегося шумового источника в мелком море с использованием высокочастотных сигналов. Акустический журнал, 2020, 66(4):437-445.

35. Кузнецов ГН, Кузькин ВМ., Куцов МВ., Переселков СА. Интерференционная структура шумового поля движущегося источника в высокочастотном диапазоне. Известия РАН. Серия физическая, 2021, 85(6):907-912.

Расчёт углового распределения голограммы источника в гидроакустическом волноводе

В. А. Башкарев, email: bashkarev.edu@yandex.ru [1]
С. А. Пересёлков, email: pereselkov@yandex.ru [1]
В. М. Кузькин, email: kumiov@yandex.ru [2]
А. С. Пересёлков, email: pereselkov@yandex.ru [1]

[1] Воронежский государственный университет
[2] Институт общей физики им. А.М. Прохорова РАН

Аннотация. *В работе методами численного моделирования выполнен расчёт углового распределения голограммы широкополосного гидроакустического источника в мелководном акустическом волноводе. Проанализирована зависимость параметров углового распределения спектральной плотности сигнала от характеристик движущегося источника. Рассмотрены такие параметры, как положение максимума и ширина полосы на уровне полумаксимума. Проведено сопоставление результатов моделирования с полученными аналитическими зависимостями.*

Ключевые слова: *численное моделирование, гидроакустические сигналы, интерферограмма, голограмма, угловое распределение спектральной плотности, максимум функции обнаружения, ширина функции обнаружения, мелководный акустический волновод, интерференционный инвариант.*

Введение

Повышение требований по точности (и скорости) оценки параметров в задаче гидролокации инициируют разработку новых подходов в обработке акустических сигналов. Одним из таких подходов является метод на основе интерферометрической обработки – голографического метода. Голографическая обработка сигналов в своей сути соответствует интерферометрическим методам, которые делают возможным практически важное оценивание параметров в условиях неблагоприятной помеховой обстановки. Другим не менее важным обстоятельством в выборе голографических методов обработки гидроакустических сигналов является отсутствие потребности в знании передаточной функции акватории, которую на практике сложно получить и эволюцией которой во времени нельзя пренебречь.

© Башкарев В. А., Пересёлков С. А., Кузькин В. М., Пересёлков А. С., 2026

Алгоритм голографической обработки акустического сигнала, следующий: в течение некоторого времени, исчисляемого десятками секунд, происходит прием сигнала, который фиксируется в виде интерферограммы – записи интенсивности сигнала в координатах время-частота. Потребность времени накопления – это плата за отсутствие потребности в знании передаточной функции акватории. Применяя к интерферограмме двумерное преобразование Фурье получают голограмму, которая является спектральной плотностью интерферограммы принимаемого сигнала. На голограмме происходит концентрация полезного сигнала в виде фокальных пятен – максимумов спектральной плотности полезного сигнала в силу его квазикогерентности, с энергия помеховой составляющей такого не происходит, помеха случайна. На этих шагах одномерный сигнал преобразуется к двумерному и предварительно обрабатывается. Заключительным шагом перед извлечением требуемой информации о параметрах шумящего объекта информация вновь приводится к одномерному виду. На основе голограммы строится угловое распределение спектральной плотности (функция обнаружения). Проведение описанного преобразование входного сигнала обеспечивает высокую эффективность проводимых оценок параметров при отсутствии информации о среде распространения и одновременно низком отношении сигнал/шум (с/ш). Извлекаемая информация из функции обнаружения делает возможным решать задачи по разрешению нескольких источников (обнаружению реального количества источников, находящихся на одном направлении), оценке скорости, дистанции, пеленг. Функция обнаружения в некотором смысле является аналогом диаграммы направленности в радиолокации. Понимание поведения функции обнаружения от параметров акватории, шумящего объекта и условий приема является важной для практики задачей. В работе проводится такое исследование, результаты, полученные расчетным методом, сравниваются с аналитически полученными результатами. Объектом исследования были выбраны координата максимума и ширина полосы функции обнаружения на уровне ее полумаксимума. Проведено сравнение результатов аналитических методов и численного эксперимента. Численный расчеты реализованы на высокоуровневом язык программирования общего назначения Python с использованием специализированных библиотек.

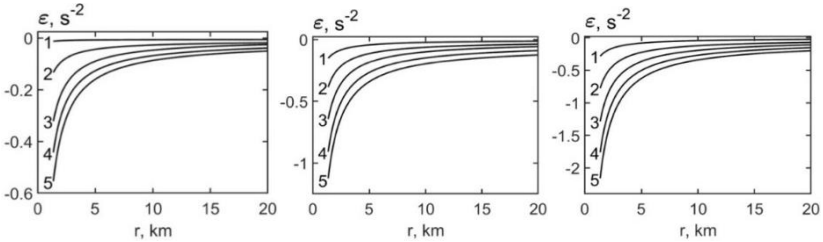
1. Численный эксперимент

Расчет и исследование зависимостей положения координаты максимума и ширины полосы функции обнаружения на уровне ее полумаксимума проведены для детерминированного случая, помеховая

составляющая отсутствует во входном сигнале. При проведении расчетов использовалась модовая модель [15]., как многократно экспериментально верифицированная для случая мелководного океанического волновода. К параметрам обработки отнесены время наблюдения Δt, частотный диапазон Δf, к параметрам движущегося источника начальная дистанция r_0 и радиальная скорость w.

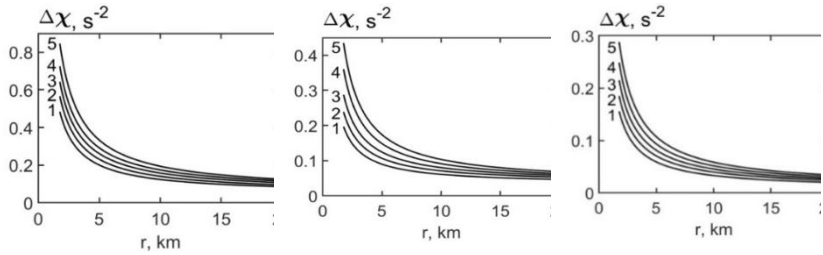
На рис. 1 и рис. 2 представлены зависимости координаты максимума ε и ширины полосы функции $\Delta\chi$ обнаружения от расстояния r для различных частотных диапазонов Δf_i. Величины $|\varepsilon|$ и $\Delta\chi$ уменьшаются с увеличением дистанции r. Расположение семейства кривых частотного диапазона Δf на рисунках демонстрирую схожую зависимость , с ростом частотного диапазона значение величин $|\varepsilon|$ и $\Delta\chi$ уменьшается.

Зависимости $|\varepsilon|$ и $\Delta\chi$ от радиальной скорости w при разных расстояниях для различных частотных диапазонов Δf_i продемонстрированны на рис. 3 и рис. 4. $|\varepsilon|$ увеличивается с ростом радиальной скорости w и расстояния r. При этом $\Delta\chi$ демострирует уменьшение. Рост радиальной скорости источника, частотного диапазона и уменьшение расстояния приводит к возрастанию $|\varepsilon|$ и $\Delta\chi$



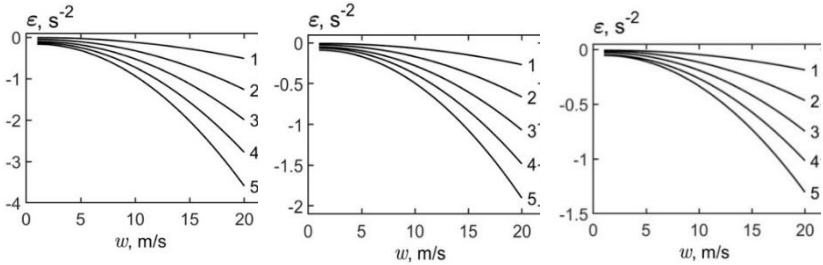
а б в

Рис. 1. Зависимость ε от r при различных w: (а) $w = 1$ м/с, (б) $w = 2$ м/с, (в) $w = 4$ м/с. Кривая 1– Δf_1, 2 – Δf_2, 3 – Δf_3, 4 – Δf_4, 5 – Δf_5. Время наблюдения $\Delta t = 60$ с.



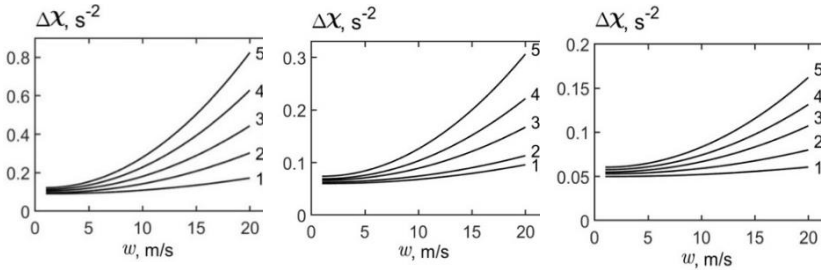
а б в

Рис. 2. Зависимость $\Delta\chi$ от r при различных Δt: (а) $\Delta t = 30\ c$, (б) $\Delta t = 60\ c$, (в) $\Delta t = 90\ c$. Кривая 1– Δf_1, 2 – Δf_2, 3 – Δf_3, 4 – Δf_4, 5 – Δf_5. $w = 2\ м/с$.



а б в

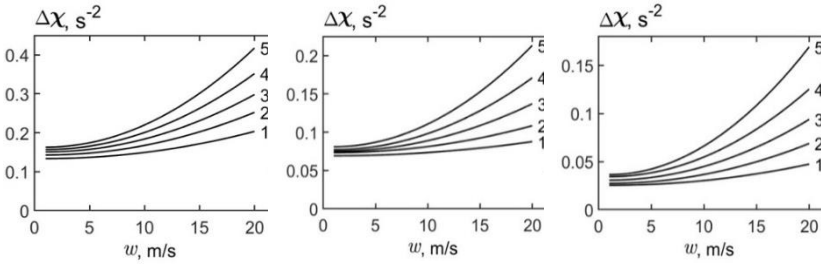
Рис. 3. Зависимость ε от w при различных r: (а) $r = 5\ км$, (б) $r = 10\ км$, (в) $r = 15\ км$. Кривая 1 – Δf_1, 2 – Δf_2, 3 – Δf_3, 4 – Δf_4, 5 – Δf_5. Время наблюдения $\Delta t = 60\ c$.



а б в

Рис. 4. Зависимость $\Delta\chi$ от w при различных r: (а) $r = 5$ *км*, (б) $r = 10$ *км*, (в) $r = 15$ *км*. Кривая 1 – Δf_1, 2 – Δf_2, 3 – Δf_3, 4 – Δf_4, 5 – Δf_5. Время наблюдения $\Delta t = 60$ *с*.

Поведение ширины полосы $\Delta\chi$ функции обнаружения при изменении радиальной скорости w для различных времен наблюдения Δt демонстрирует рис. 5. Наблюдается закономерность роста ширина полосы с ростом радиальной скорости и ее уменьшение при увеличении времени наблюдения.



а б в

Рис. 5. Зависимость $\Delta\chi$ от w при различных Δt: (а) $\Delta t = 30$ *с*, (б) $\Delta t = 60$ *с*, (в) $\Delta t = 90$ *с*. Кривая 1 – Δf_1, 2 – Δf_2, 3 – Δf_3, 4 – Δf_4, 5 – Δf_5. Расстояние $r = 10$ *км*.

При численных оценках ширины полосы (12) полагалось число мод $M = N$, что, естественно, должно приводить к заниженным оценкам. Величины ε и $\Delta\chi$, определяемые численным моделированием,

обозначаются нижним индексом «1», а определяемые по выражениям (11), (12) – нижним индексом «2». Первый пример рассмотрим на основе рис. 1б для расстояния $r = 5$ км. По результатам моделирования $\varepsilon_1 = -0.047, -0.32 \quad c^{-2}$ для частотных диапазонов Δf_1, Δf_5 соответственно. Согласно соотношению (11) $\varepsilon_2(\Delta f_1) = -0.053 \quad c^{-2}$ и $\varepsilon_2(\Delta f_5) = -0.30 \quad c^{-2}$. Значения ε_1 и ε_2 близки между собой для существенно разных частотных диапазонов. Второй пример рассмотрим на основе рис. 2б для расстояния $r = 5$ км. Согласно результатам моделирования $\Delta\chi_1(\Delta f_1) = 0.096 \quad c^{-2}$ и $\Delta\chi_1(\Delta f_5) = 0.18 \quad c^{-2}$. На основании соотношения (12) получаем $\Delta\chi_2(\Delta f_1) = 0.16 \quad c^{-2}$ и $\Delta\chi_2(\Delta f_5) = 0.042 \quad c^{-2}$. Для низкочастотного диапазона Δf_1, когда номера мод $M \approx N$, значения ширины полосы сопоставимы между собой. В высокочастотном диапазоне Δf_5 численная оценка ширины полосы существенно меньше модельной, $\Delta\chi_2(\Delta f_5) \ll \Delta\chi_1(\Delta f_5)$. Это объясняется заметным различием между номерами мод M и N. Третий и четвертый примеры рассмотрим на основе рис. 3б и рис. 4б для радиальной скорости $w = 5$ м/с. $\varepsilon_1(\Delta f_1) = -0.048 \quad c^{-2}$, $\varepsilon_1(\Delta f_5) = -0.28 \quad c^{-2}$, $\Delta\chi_1(\Delta f_1) = 0.068 \quad c^{-2}$, $\Delta\chi_1(\Delta f_5) = 0.083 \quad c^{-2}$. Согласно отношениям (11), (12) получаем $\varepsilon_1(\Delta f_1) = -0.066 \quad c^{-2}$, $\varepsilon_1(\Delta f_5) = -0.37 \quad c^{-2}$, $\Delta\chi_2(\Delta f_1) = 0.081 \quad c^{-2}$ и $\Delta\chi_2(\Delta f_5) = 0.021 \quad c^{-2}$. Результаты сопоставления характеристик функции обнаружения подтверждают закономерности, выявленные в первых двух примерах. В качестве пятого примера сопоставим результаты моделирования, представленные на рис. 5а для времени наблюдения $\Delta t = 30 \quad c$. На основании результатов моделирования $\Delta\chi_1(\Delta f_1) = 0.13 \quad c^{-2}$, $\Delta\chi_1(\Delta f_5) = 0.087 \quad c^{-2}$. Из (12) имеем $\Delta\chi_1(\Delta f_1) = 0.16 \quad c^{-2}$, $\Delta\chi_1(\Delta f_5) = 0.042 \quad c^{-2}$. Проведенная верификация расчетных закономерностей с аналитическими доказывают, что для случая малых радиальных скоростях, и расстояний порядка десяти километров, выражения (11), (12) правильно оценивают координаты максимума и ширины полосы функции обнаружения.

Заключение

В результате проведенных расчетов функции обнаружения подводного широкополосного источника акустического сигнала в мелководном океаническом волноводе получены закономерности

поведения таких параметром как положение максимума и ширины полосы на уровне полумаксимума с параметрами движения источника и характеристиками приема. Результаты модельных закономерностей сравнены с аналитическими зависимостями. Продемонстрирована практически важная совпадаемость расчетных и аналитических результатов.

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

Список литературы

1. Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Спектрограмма и локализация источника звука в мелком море // Акуст. журн. 2017. Т. 63. № 4. С. 406–418.

2. Казначеев И.В., Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Интерферометрический метод обнаружения движущегося источника звука векторно-скалярным приемником // Акуст. журн. 2018. Т. 64. № 1. С. 33–45.

3. Pereselkov S.A., Kuz'kin V.M. Interferometric processing of hydroacoustic signals for the purpose of source localization // J. Acoust. Soc. Am. 2022. Vol. 151. No. 2 P. 666−676.

4. Ehrhardt M., Pereselkov S., Kuz'kin V., Kaznacheev I., Rybyaners P. Experimental observation and theoretical analysis of the low-frequency source interferogram and hologram in shallow water // Sound and Vibration. 2023. Vol. 544. P. 117388.

5. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Rybyanets P., Ladykin N. Three-dimensional modeling of sound field holograms of a moving source in the presence of internal waves causing horizontal refraction // J. Mar. Sci. Eng. 2023. Vol. 11. No. 10, 1922.

6. Pereselkov S., Kuz'kin V., Ehrhardt M., Matvienko Yu., Tkachenko S., Rybyanets P. The formation of 2D holograms of a noise source and bearing estimation by a vector scalar receiver in the high-frequency band // J. Mar. Sci. Eng. 2024. Vol. 12. P. 704.

7. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Pereselkov A., Ladykin N. Influence of intense internal waves traveling along an acoustic path on source holographic reconstruction in shallow water // J. Mar. Sci. Eng. 2025. Vol. 13. P. 1409.

8. Кузькин В.М., Пересёлков С.А. Голографические методы обработки гидроакустических сигналов (Обзор) // Акуст. журн. 2025. Т. 71. № 1. С. 96–117.

9. Weston D., Stevens K. Interference of wide-band sound in shallow water // J. Sound Vibr. 1972. Vol. 21. No. 1. P. 57–64.

10. Бреховских Л.М., Лысанов Ю.П. Теоретические основы акустики океана. Л.: Гидрометеоиздат, 1982. 264 с.

11. Глущенко М.Ю., Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Хворостов Ю.А., Ткаченко С.А. Голографический метод идентификации шумового подводного источника в мелком море // Акуст. журн. 2024. Т. 70. № 3. С. 67−77.

12. Чупров С.Д. Интерференционная структура звукового поля в слоистом океане // Акустика океана. Современное состояние. М.: Наука, 1982. С. 71–82.

13. Кулаков В.Н., Мальцев Н.Е., Чупров С.Д. О возбуждении групп мод в слоистом океане // Акуст. журн. 1983. Т. 29. № 1. С. 74–79.

14. Кузькин В.М., Переселков С.А. Интерферометрическая диагностика гидродинамических возмущений мелкого моря. М: Ленанд, 2019. 200 с.

15. Kohler W., Papanicolaou G.C. Wave propagation in a randomly inhomogeneous ocean // Wave Propagation and Underwater Acoustics, eds. by J. B. Keller and J. S. Papadakis. Springer: Berlin Heidelberg New York, 1977. P. 153–223.

16. Kuz'kin V.M., Pereselkov S.A., Zvyagin V.G., Malykhin Yu., Prosovetskiy D.Yu. Intense internal waves and their manifestation in interference patterns of received signals on oceanic shelf // Phys. Wave Phenom. 2018. Vol. 26. № 2. P. 160–167.

17. Badiey M., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Prosovetskiy D.Yu., Tkachenko S.A. Intense internal waves and their manifestation in the interference patterns of received signals on oceanic shelf. Part II // Phys. Wave Phenom. 2019. Vol. 27. № 4. P. 313–319.

18. Kaznacheeva E.S., Kuz'kin V.M., Pereselkov S.A. Interferometric processing of hydroacoustic information in the presence of intense internal waves // Phys. Wave Phenom. 2021. Vol. 29. № 3. P. 278–284.

19. Кузькин В.М., Пересёлков С.А., Грачев В.И., Ткаченко С.А., Ладыкин Н.В., Куцов М.В. Голографический метод локализации движущегося подводного источника звука в присутствии интенсивных внутренних волн // РЭНСИТ. 2023. Т. 15. № 3. С. 317–326.

20. Kuz'kin V.M., Pereselkov S.A., Badiey M., Ladykin N.V., Malykhin A.Yu., Tkachenko S.A. Robustness of holographic processing of hydroacoustic signals in the presence of intense internal waves // Phys. Wave Phenom. 2023. Vol. 31. № 5. P. 346–354.

21. Кузькин В.М., Переселков С.А., Грачев В.И., Косенко И.М., Ладыкин Н.В., Переселков А.С. Локализация движущегося шумового

источника в мелководной акватории с нерегулярной батиметрией // РЭНСИТ. 2024. Т. 16. № 8. С. 1081–1088.

22. Kaznacheeva E.S., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Tkachenko S.A. Adaptive algorithms for interferometric processing // Phys. Wave Phenom. 2020. Vol. 28. № 3. P. 267–273.

23. Kuz'kin V.M., Pereselkov S.A., Rybyanets P.V., Tkachenko S.F. Adaptive method for holographic processing of broadband hydroacoustic signals // Phys. Wave Phenom. 2024. Vol. 32. № 5. P. 382–391.

24. Кузькин В.М., Пересёлков С.А., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Адаптивный алгоритм локализации шумовых подводных источников в мелководных акваториях // РЭНСИТ. 2024. Т. 16. № 4. С. 533–544.

25. Kuz'kin V.M., Kuznetsov G.N., Pereselkov S.A., Grigor'ev V.A. Resolving power of the interferometric method of source localization // Phys. Wave Phenom. 2018. Vol. 26. № 2. P. 150–159.

26. Кузькин В.М., Переселков С.А., Грачев В.И., Рыбянец П.В., Ткаченко С.А. Разрешение шумовых сигналов при наличии случайных неоднородностей океанической среды // РЭНСИТ. 2024. Т. 16. № 3. С. 395–406.

27. Kuz'kin V.M., Pereselkov S.A., Rybyanets P.V., Tkachenko S.A. Resolution of hydroacoustic noise signals under conditions of spatial-temporal variability of ocean environment // Phys. Wave Phenom. 2024. Vol. 32. № 4. P. 290–299.

28. Kuznetsov G.N., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Prosovetskiy D.Yu. Direction finding of a noise sound source // Phys. Wave Phenom. 2019. Vol. 27. № 3. P. 237–241.

29. Kuz'kin V.M., Pereselkov S.A., Kuznetsov G.N., Kaznacheev I.V. Interferometric direction finding by a vector-scalar receiver // Phys. Wave Phenom. 2018. Vol. 26. № 1. P. 63–73.

30. Кузькин В.М., Пересёлков С.А., Ткаченко С.А., Казначеев И.В. Интерферометрическая обработка акустической информации с использованием протяжённых антенн в диспергирующих средах // РЭНСИТ. 2020. Т. 12. № 4. С. 483–494.

31. Kuz'kin V.M., Pereselkov S.A., Matvienko Yu.V., Tkachenko S.A. Detection of a noise signal in an oceanic waveguide using a vertical array // Phys. Wave Phenom. 2021. Vol. 30. № 4. P. 323–329.

32. Кузькин В.М., Пересёлков С.А., Матвиенко Ю.В., Ткаченко С.А., Просовецкий Д.Ю. Обнаружение шумового сигнала в океаническом волноводе горизонтальной антенной // РЭНСИТ. 2022. Т. 14. № 1. С. 65–72.

33. Kaznacheev I.V., Kuz'kin V.M., Kutsov M.V., Lyakhov G.A., Pereselkov S.A. Interferometry in acoustic-data processing using extended antennas. Space-time analogy // Phys. Wave Phenom. 2020. Vol. 28. № 4. P. 326–332.

34. Kuz'kin V.M., Matvienko Yu.V., Pereselkov S.A., Prosovetskii D.Yu., Kaznacheeva E.S. Mode selection in oceanic waveguides // Phys. Wave Phenom. 2022. Vol. 30. № 2. P. 111–118.

35. Kuz'kin V.M., Matvienko Yu.V., Pereselkov S.A., Kaznacheeva E.S., Tkachenko S.A. Holographic method for mode selection in a shallow sea in the presence of intense internal waves // Phys. Wave Phenom. 2022. Vol. 30. № 5. P. 314–320.

36. Кузькин В.М., Пересёлков С.А., Казначеева Е.С., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Выделение мод шумового источника в мелком море методом голографической интерферометрии в присутствии интенсивных внутренних волн // РЭНСИТ. 2022. Т. 14. № 3. С. 301–308.

37. Кузькин В.М., Пересёлков С.А., Грачев В.И., Логачев В.В., Косенко И.М. Голографический метод обработки гидроакустических сигналов высокочастотного источника в мелком море // РЭНСИТ. 2024. Т. 16. № 4. С. 545–554.

38. Матвиенко Ю.В., Кузькин В.М., Переселков С.А., Ладыкин Н.В., Грачев В.И. Обнаружение и локализация шумового подводного источника в мелководной акватории на фоне интенсивного судоходства // РЭНСИТ. 2025. Т. 17. № 2. С. 211–220.

39. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Rybyanets P., Ladykin N. Use of interference patterns to control sound field focusing in shallow water // J. Mar. Sci. Eng. 2023. Vol. 11. P. 559.

Исследование ультразвуковых фокусированных преобразователей для неразрушающего контроля

Д. А. Кравчук, e-mail: kravchukda@sfedu.ru
А. Д. Шаповалов, e-mail: arshapo@sfedu.ru

ФГАОУ ВО Южный федеральный университете Институт нанотехнологий, электроники и приборостроения

Аннотация. *Многообразие задач, возникающих при необходимости проведения неразрушающего контроля различных изделий, привело к разработке и определению характеристик преобразователей различных типов и необходимо выбрать те, которые обеспечивают наибольшую чувствительность, точность и стабильность регистрации сигналов при неразрушающем контроле.*

Ключевые слова: *Дефект, акустика, фокусируемый преобразователь*

Введение

Неразрушающий контроль (НК) представляет собой совокупность методов, позволяющих выявлять внутренние и поверхностные дефекты в материалах и изделиях без их разрушения. Это направление играет ключевую роль в обеспечении надёжности и безопасности конструкций в таких областях, как энергетика, нефтегазовая промышленность, транспорт, металлургия. Раннее выявление дефектов уменьшает риск выхода оборудования из строя; оптимальный выбор преобразователя снижает производственные потери, связанные с браком [1, 2].

Математическое моделирование

1. Фокусирующий преобразователь излучает ультразвуковой луч, сходящийся в заданной точке внутри материала. Это повышает разрешающую способность и чувствительность к малым дефектам по сравнению с плоскими (неподвижными) преобразователями. Фокусировка может быть акустической (с помощью линз) или электронной (фазированная решётка).

2. Описание процесса распространения поперечной акустической волны в биологической среде на основе уравнения движения, описывающее процесс распространения акустической волны в среде с учетом поперечной деформации [3, 4]:

© Кравчук Д. А., Шаповалов А. Д., 2026

$$F = \frac{R \cdot c_m}{c_l - c_m}, \tag{1}$$

где R – радиус кривизны акустической линзы, c_l – скорость звука в материале линзы (например, полимер, $c_l \approx 2700$ м/с), c_m - скорость звука в контролируемом материале (для стали $c_m \approx 5900$ м/с), ρ – плотность среды.

Глубина резкости (длина фокусной зоны) L оценивается как [3]:

$$L \approx \frac{D^2 f}{4 c_m}, \tag{2}$$

D – диаметр активной поверхности преобразователя, f – частота.

Поперечное разрешение в фокальной точке определяется дифракционным пределом (θ — угол раскрытия луча):

$$\Delta x \approx \frac{c_m}{2 f \sin\theta} \tag{3}$$

Для проведения дефектоскопии используются прямые и наклонные многоканальные кассеты пьезоэлектрические преобразователи (ПЭП), ориентированные таким образом, что позволяют выявлять продольные и поперечные дефекты.

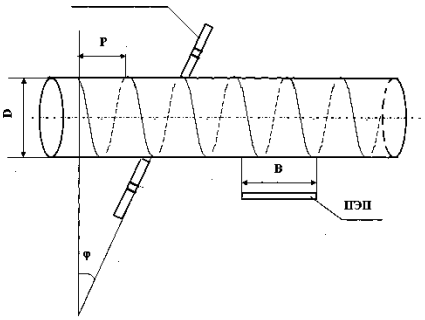


Рис. 1. Система сканирования труб

На рис. 1 приведены метод контроля труб при контроле продольных и поперечных дефектов от 73 до 273 мм. Толщина стенки контролируемых труб от 5 до 30 мм. D - диаметр контролируемой трубы; P - шаг сканирования; B - ширина применяемой кассеты ПЭП. ϕ -угол разворота колес рольганга [5].

Экспериментальные исследования характеристик разрабатываемого преобразователя проведем на сравнении с характеристиками, фокусирующих и не фокусирующих иммерсионных преобразователей, работающих на частотах 2,5 и 5 МГц [6].

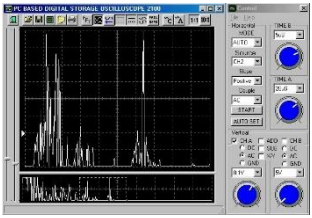


Рис. 2. Осциллограммы эхоимпульсов, полученных от дефекта, расположенного на плоскости: не фокусированный 2,5 МГц, 47 дБ (дефект и поверхность)

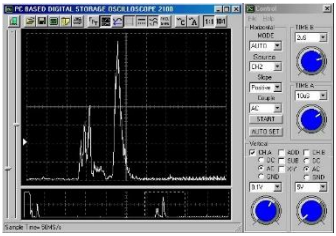


Рис. 3. Осциллограммы эхоимпульсов, полученных от дефекта, расположенного на плоскости: фокусированным преобразователем 5 МГц, 47 дБ (дефект и поверхность)

Заключение

Использование фокусирующих ультразвуковых преобразователей в системах контроля качества труб позволяет повысить точность обнаружения дефектов, снизить вероятность пропуска критических несоответствий и обеспечить надёжную эксплуатацию трубопроводов в экстремальных условиях нефтяной промышленности.

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

Список литературы

1. Барахатов В.А., Семухин А.С. Ультразвуковой комплекс. Приемник акустических сигналов. – Дефектоскопия, 2001, № 6, с 38 – 41.

2. Дедешко В.Н., Салюков В.В. Развитие системы диагностического обслуживания магистральных газопроводов. //Газовая промышленность, 2005, №8, с.15-18.

3. Бэтчелор Г.К. Волны сжатия в суспензии газовых пузырьков в жидкости // Механика: сб. переводов иностр. ст. 1968. Т. 109, № 3. С. 65-84.

4. Наука и Образование. МГТУ им. Н.Э. Баумана. Электрон. журн. 2015. № 02.

5. Кравчук Д.А., Кожаров В.Ю. Система неразрушающего контроля труб для нефтяной промышленности. В сборнике: Проблемы автоматизации. Региональное управление. Связь и автоматика (Паруса - 2016) Сборник трудов V Всероссийской научной конференции молодых ученых, аспирантов и студентов. Редколлегия: Фоменко О.А., Кирильчик С.В., Номерчук А.Я.. 2016. С. 119-124.

6. Кравчук, Д. А. Исследования преобразователей для неразрушающего контроля труб при производстве / Д. А. Кравчук, А. В. Ряжкин, С. А. Лыков // Радиолокация, навигация, связь: Сборник трудов XXX Международной научно-технической конференции. В 5-ти томах, Воронеж, 16–18 апреля 2024 года. – Воронеж: Издательский дом ВГУ, 2024. – С. 184-188. – EDN LFQGAU.

Анализ процессов нелинейного взаимодействия акустических волн с биотканями для оценки их структурных изменений

М. В. Лагута, e-mail: laguta@sfedu.ru
Д. А. Кравчук, e-mail: kravchukda@sfedu.ru

ФГАОУ ВО Южный федеральный университете Институт нанотехнологий, электроники и приборостроения

Аннотация. *В работе представлены результаты анализа нелинейных эффектов взаимодействия акустических волн с биологическими тканями. Их использование в качестве способа получения информационных характеристик позволяет повысить качество диагностики заболеваний на ранних стадиях.*

Ключевые слова: *Нелинейная акустика, эластография, поперечные волны.*

Введение

Актуальность исследования процессов нелинейного взаимодействия акустических волн с биологическими тканями обусловлена растущей потребностью в неинвазивных методах диагностики патологических изменений на ранних стадиях. Традиционные ультразвуковые методы, основанные на линейной акустике, зачастую недостаточно чувствительны к микроструктурным изменениям, происходящим в тканях при развитии заболеваний, таких как фиброз, воспаление или онкологические процессы. В то же время нелинейные акустические эффекты — например, генерация высших гармоник, модуляционные взаимодействия и зависимость скорости распространения от амплитуды — оказываются значительно более чувствительными к микроскопическим неоднородностям, вязкоупругим свойствам и степени организации коллагенового каркаса.

Современные достижения в области нелинейной акустики и цифровой обработки сигналов открывают возможности для создания новых диагностических подходов, основанных на количественной оценке нелинейных параметров среды. Такие параметры, как коэффициент нелинейности β или отношение амплитуд высших гармоник к основной, могут служить биомаркерами структурного состояния тканей. Особенно перспективно применение этих методов в онкологии, где изменения в микроархитектуре стромы часто предшествуют макроскопическим проявлениям опухолевого процесса. Это позволяет повысить точность

© Лагута М. В., Кравчук Д. А., 2026

дифференциальной диагностики и снизить количество ложноположительных результатов.

Кроме того, развитие вычислительных моделей, учитывающих как нелинейную упругость, так и диссипативные процессы в биотканях, способствует более глубокому пониманию механизмов формирования нелинейных акустических сигналов. Моделирование позволяет не только интерпретировать экспериментальные данные, но и оптимизировать параметры зондирующих сигналов для максимизации диагностической информативности. В совокупности это создаёт основу для разработки новых аппаратных и программных решений в медицинской ультразвуковой диагностике, ориентированных на раннее выявление структурных аномалий без использования ионизирующего излучения или контрастных агентов.

Структурные изменения тканей и органов сопровождает многие заболевания, протекающие в организме человека. В частности, таким изменениям подвергается печень. Важной задачей современных методов диагностики заболеваний печени является возможность исследование динамики протекания заболеваний, а также их дифференциация на ранних стадиях. Кроме того, нельзя забывать о таком важном факторе, как безопасность. Именно по этой причине биопсия, которая долгое время считалась «золотым стандартом» на сегодняшний день отошла на второй план. Также к ее недостатку можно отнести малую информативность в случае, когда площадь исследуемого органа значительно превышает размер изымаемого образца тканей. Методы акустической эластографии на сегодняшний день являются одними из наиболее перспективных с учетом перечисленного выше. Они также являются более безопасными по сравнению с рентгеном и имеют меньше ограничений в сравнении с МРТ. Однако большинство существующих устройств, использующих акустические поля основаны на приближениях линейной акустики, что существенно ограничивает диагностические возможности методов. Методы акустической эластографии основаны на анализе процесса распространения поперечной компоненты акустического сигнала, используемом для оценки вязкоупругих свойств органов и тканей человека. Это связано с тем, что скорость поперечной волны значительно меньше, чем скорость продольной, а значит ее относительные изменения будут удобнее для дифференциации. На основе относительного изменения скорости поперечных колебаний делают выводы о упругих характеристиках биотканей, таких как модуль Юнга [1-3].

Математическое моделирование

Описание процесса распространения поперечной акустической волны в биологической среде на основе уравнения движения,

описывающее процесс распространения акустической волны в среде с учетом поперечной деформации [4,5]:

$$\frac{\partial u}{\partial z} - \frac{\beta}{2\mu V_0^2} u \frac{\partial u}{\partial \tau} = \frac{\eta}{2\mu V_0^2} \frac{\partial^2 u}{\partial \tau^2}, \quad (1)$$

где $\tau \equiv t - \frac{z}{V_0}$ – время в движущейся системе координат, $V_0 = \sqrt{\frac{\mu}{\rho}}$ – скорость невозмущенной волны, ρ – плотность среды, z – координата в направлении распространении волны.

Для оценки нелинейных свойств среды введем параметр N, который выражается следующим образом [2,3]:

$$N = \frac{A_3}{A_1}, \quad (2)$$

где A_1 и A_3 – амплитуда первой и третьей гармоники.

Амплитуда n-й гармоники поперечного смещения определяется выражением [6]:

$$A_n(Z) = \frac{U_n(Z) A_0}{2\pi f_0 n} = \frac{J_n(nZ) A_0}{\pi f_0 n^2 Z}, \quad (3)$$

где A_0 – начальная амплитуда смещения в волне.

На рис. 1 приведены результаты расчета изменения параметра N с расстоянием для печени при частичном замещении функциональных клеток жировой тканью. На рис. 1: 1 – здоровая печень, 2 – печень с включением жировой ткани

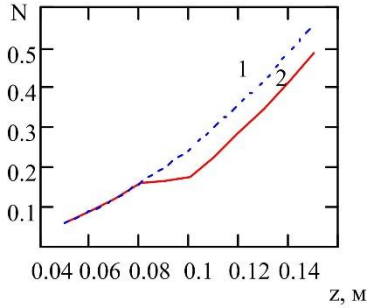


Рис. 1. Изменение N с расстоянием при наличии включений жировой ткани

При наличии в печени включения из жировых тканей размером 2 см, нелинейный параметр которого отличается на 14 %, параметр N на расстоянии 15 см меняется на 42%.

Также на рис. 2 приведены результаты расчетов изменения параметра N с расстоянием для различных стадий фиброза печени.

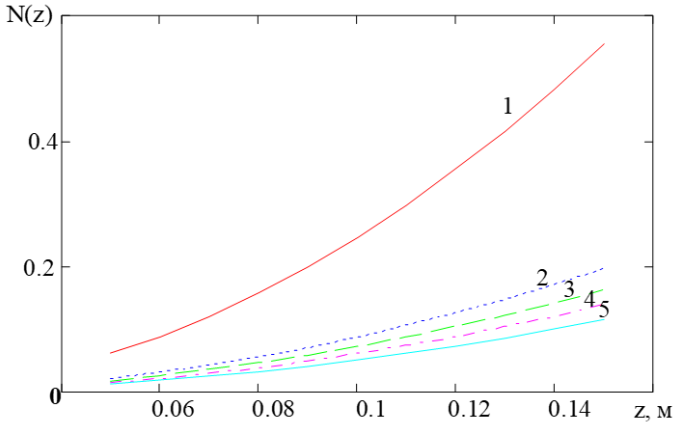


Рис. 2. Изменение параметра N с расстоянием при фиброзе различных стадий

На рис. 2: : 1 – F_0, 2 – F_1, 3 – F_2, 4 – F_3, 5 – F_4. Результат, представленный на рис. 2 показал, что изменение параметра N для стадий F_1 и F_2 составляет 26 %, для стадий F_2 и F_3 – 23 % при этом диапазоны значений модуля Юнга для этих стадий в некоторых случаях могут пересекаться.

Заключение

Проведенное моделирование позволяет сделать вывод, что вклад в процесс изменения параметра N с расстоянием вносят как непосредственно нелинейный параметр среды, так и жесткость ткани. В совокупности учет этих параметров позволит проводить более точную дифференциацию степени изменения структуры тканей и органов, в частность при развитии фиброза или гепатоза печени на ранних стадиях, что является сложной задачей для стандартных методов эластографии.

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

Список литературы

1. Демин, И. Ю. Современные акустические методы исследований в биологии и медицине: Учебно-методический материал по программе повышения квалификации «Хранение и обработка информации в биологических системах» / И. Ю. Демин, Н. В. Прончатов-Рубцов. – Нижний Новгород, 2007. 121 с.

2. Изранов, В. А. Физические основы эластографии печени / В. А. Изранов, Н. В. Казанцева // Вестник Балтийского федерального университета им. И. Канта. Серия: Естественные и медицинские науки. 2019. – №2. URL: https://cyberleninka.ru/article/n/fizicheskie-osnovy-elastografii-pecheni (дата обращения: 29.06.2024).

3. Лагута, М. В. Оценка упругих свойств биотканей на основе их нелинейного взаимодействия с акустическим полем / М. В. Лагута, Н. Н. Чернов // Сборник трудов XII Всероссийской научной конференции и молодёжной школы-семинара. ЭКОЛОГИЯ 2023 – МОРЕ И ЧЕЛОВЕК. (Таганрог. 21–23 сентября 2023). – Таганрог, 2023. – С. 31-36.

4. Лагута, М. В. Исследование нелинейного взаимодействия акустической волны с биотканями для целей акустической эластографии / М. В. Лагута, Д. А. Кравчук, Н. Н. Чернов // Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение. – 2025. – Т. 15, № 1. – С. 131–143.

5. Astapov, Y. Concretization of nonlinear constitutive relations by results of uniaxial compression and indentation experiments / Y. Astapov, A. Markin, M. Sokolova, D. Khristich // Journal of Physics: Conference Series. – 2021. – 1902(1). 012002.

6. Нефедов, Н. Н. О движении, усилении и разрушении фронтов в уравнениях типа бюргерса с квадратичной и модульной нелинейностью / Н. Н. Нефедов, О. В. Руденко // Доклады российской академии наук. математика, информатика, процессы управления. – Москва, 2020. – Т.493. – С. 26–31.

Численный расчет мультинаправленных векторно-скалярных пеленгаторов

Ю. В. Матвиенко, email: ymat33@yandex.ru[1]
С. А. Переселков, email: pereselkov@yandex.ru [2]
И. М. Косенко, email: kosenko.edu@yandex.ru [2]
Н. В. Ладыкин, email: kosenko.edu@yandex.ru [2]

[1] Институт проблем морских технологий ДО РАН
[2] Воронежский государственный университет

Аннотация. *В работе представлен численный расчёт мультинаправленных векторно-скалярных пеленгаторов, использующих несколько пар ортогональных каналов с общим фазовым центром для регистрации горизонтальных компонент колебательной скорости. Проанализирована их эффективность и показано, что такие системы обеспечивают равномерное снижение погрешности определения пеленга.*

Ключевые слова: *мультинаправленный приемник, векторно-скалярный пеленгатор, зависимость погрешности пеленгации от угла, отношение сигнал/помеха.*

Введение

Пассивные методы пеленгования низкочастотного шума требуют крупногабаритных антенн, непригодных для мобильных носителей [1–2]. Компактные векторные и векторно-скалярные приёмники уменьшают габариты, но классические схемы дают резкие угловые провалы точности из-за разбалансов каналов и неравномерного воздействия помех [3–6]. Предлагается архитектура мультинаправленного комбинированного приёмника, использующего несколько повернутых пар ортогональных каналов, что выравнивает точность за счёт независимых измерений и разбиения горизонтального сектора.

Голографическая обработка дополняет такую схему: её устойчивость опирается на интерференционные структуры, задаваемые источником [7–12], она эффективна в широком диапазоне частот [13–16], не требует априорных знаний о среде [17–23], позволяет также проводить оценку глубины источника [24–25], обладает высокой надежностью [26–27], проводит разделение близких объектов [28–30] и показывает высокую эффективность при пеленгации [31–33], поддерживает адаптивные методы [34–36], обеспечивает модовую селекцию в разных

© Матвиенко Ю. В., Переселков С. А., Косенко И. М., Ладыкин Н. В., 2026

типах волноводов [37–39] и превосходит классические схемы фокусировки [40].

1. Пеленгатор с одной парой ортогональных каналов

Наиболее ранним алгоритмом пеленгования подразумевается использование векторно-скалярного приемника. Такой приемник в силу ортогональности пары каналов с общим центром используется для измерения звукового давления P и векторов колебательных скоростей V_X и V_Y среды [4]. Компоненты интенсивности формируются на горизонтальных каналах приемника так:

$$I_X = RE\left(PV_X^*\right), I_Y = RE\left(PV_Y^*\right) \quad (1)$$

При использовании голографического метода пеленг для такого приемника оценивается следующим образом:

$$\theta = arctg\frac{I_Y}{I_X} \quad (2)$$

2. Пеленгатор с двумя парами ортогональных каналов

Две пары ортоканалов с общим центром предлагается разнести на 45 градусов. Поскольку датчиков теперь четыре, выходит извлечь больше компонент вектора интенсивности:

$$I_{X1} = RE\left(PV_{X1}^*\right), I_{Y1} = RE\left(PV_{Y1}^*\right), I_{X2} = RE\left(PV_{X2}^*\right), I_{Y2} = RE\left(PV_{Y2}^*\right) \quad (3)$$

Оценка пеленга в таком случае производится по следующему выражению:

$$\theta = \frac{\sum_{i=1}^{N}\mu_i^2\theta_i}{\sum_{i=1}^{N}\mu_i^2} \quad (4)$$

В нём $N = 6$ слагаемых, определяемых как:

$$\begin{aligned}&\theta_1 = arctg\frac{I_{Y1}}{I_{X1}}, \theta_2 = \frac{\pi}{4} - arctg\frac{I_{Y2}}{I_{X2}}, \theta_3 = arctg\frac{\sqrt{2}I_{Y2} - I_{X1}}{I_{X1}},\\ &\theta_4 = arctg\frac{I_{Y1}}{\sqrt{2}I_{X2} + I_{Y1}}, \theta_5 = arctg\frac{I_{X1} - \sqrt{2}I_{X2}}{I_{X1}}, \theta_6 = arctg\frac{I_{Y1}}{\sqrt{2}I_{Y2} - I_{Y1}}\end{aligned} \quad (5)$$

3. Пеленгатор с тремя парами ортогональных каналов

В случае с тремя парами ортоканалов угол между ними задаётся в 30 градусов. Извлекаются шесть компонент вектора интенсивности.

$$I_{X1} = RE\left(PV_{X1}^{*}\right), I_{Y1} = RE\left(PV_{Y1}^{*}\right),$$
$$I_{X2} = RE\left(PV_{X2}^{*}\right), I_{Y2} = RE\left(PV_{Y2}^{*}\right), \quad (6)$$
$$I_{X3} = RE\left(PV_{X3}^{*}\right), I_{Y3} = RE\left(PV_{Y3}^{*}\right)$$

Пеленг также определяется соотношением (4), но $N = 15$ слагаемых определяются как

$$\theta_1 = arctg\frac{I_{Y1}}{I_{X1}}, \theta_2 = \frac{\pi}{6} - arctg\frac{I_{Y2}}{I_{X2}}, \theta_3 = \frac{\pi}{3} - arctg\frac{I_{Y3}}{I_{X4}}, \ldots \quad (7)$$

4. Численное моделирование

Результаты численного эксперимента для векторно-скалярного приемника для случая $q = 16$ дБ представлены на рис. 1. Рисунок содержит информацию о погрешности пеленгования σ_θ. Наблюдается зависимость цикличность ошибки пеленгования через каждые 45 градусов. Пики максимальной ошибки на графике порядка 5 градусов соответствуют направлениям 0, 90, 180, 270 градусов. Минимальная ошибка, менее 1 градуса, фиксируется для пеленга 45, 125, 225, 315.

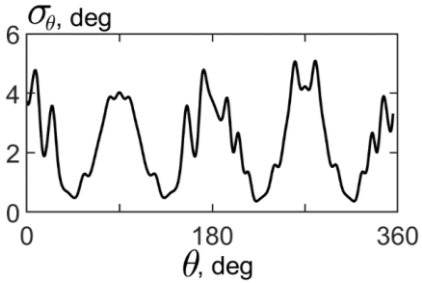


Рис. 1. Погрешность оценки пеленга σ_θ для приемника с одной парой каналов ($q = 16$ дБ)

На рис. 2 представлен результат для отношения сигнал/шум $q = 6$ дБ для случая мультинаправленного приемника с двумя парами каналов, повернутыми на 45 градусов. Сравнивая графики можно сделать вывод для такой конструкции приемника произошло уменьшение ошибки для всего диапазона углов. В дополнение к уже существующим локальным минимумам сформировались ещё четыре сектора с осями в 0, 90, 180, 270 градусов, в которых точность оценки гораздо выше.

Наконец, как показано на рис. 3, добавление ещё одной пары ортоканалов позволяет ещё сильнее равномерно снизить погрешность пеленгации. Всего формируются 16 секторов шириной 22.5 градуса, с максимальной точностью оценки. Максимальная погрешность также заметно уменьшается.

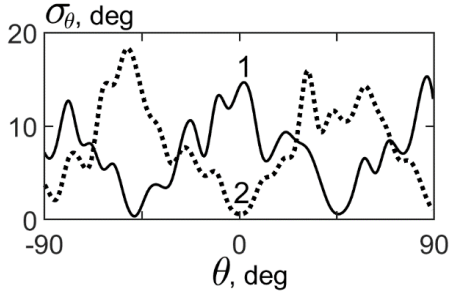
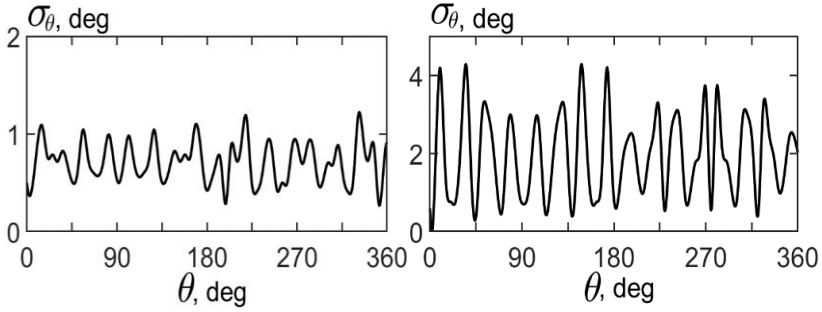


Рис. 2. Погрешность оценки пеленга σ_θ для приемника с двумя парами каналов ($q = 6$ дБ)



(а) (б)

Рис. 3. Зависимость погрешности пеленгования σ_θ для приемника с тремя парами каналов: а) $q = 16$ дБ, б) $q = 6$ дБ

Заключение

Рассмотрены модели мультинаправленного комбинированного векторно-скалярного пеленгатора, использующего несколько пар ортоканалов с общим центром. В результате теоретического анализа и численного моделирования показано, что применение подобных систем позволяет равномерно снижать погрешность пеленгации по всему угловому диапазоную. Для каждой системы выписаны соответствующие

выражения для оценки пеленга. Применение данного метода вместе с голографическим методом позволяет достичь высокой помехоустойчивости. Данный эффект достигается за счет более мелкого разбиения диапазона.

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

Список литературы

1. Бурдик В.С. Анализ гидроакустических систем. — Ленинград: Судостроение, 1988. 392 с.
2. Селезнев И.А., Ясников А.И. Перспективы применения подводных глайдеров для океанографии и освещения подводной обстановки. Обзор по материалам зарубежной печати // Подводные исследования и робототехника. 2023. № 1 (43). С. 4–13.
3. Скребнев Г.К. Комбинированные гидроакустические приемники. — Санкт-Петербург: Элмор, 1997. 200 с.
4. Щуров В.А. и др. Мобильные акустические комбинированные приемные системы на основе автономных необитаемых подводных аппаратов // Подводные исследования и робототехника. 2012. № 2(14). С. 4–12.
5. Захаров К.Л. Частотно-угловые характеристики гидроакустического сигнала при использовании векторно-фазового метода // Устойчивое инновационное развитие: проектирование и управление (электронное научное издание). 2014. Т. 10, № 3 (24). С. 7.
6. Матвиенко Ю.В., Хворостов Ю.А., Кулешов В.П. Особенности применения скалярно-векторных приемников звука в системах контроля подводной обстановки локальных районов // Подводные исследования и робототехника. 2022. № 4 (42). С. 4–15.
7. Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Спектрограмма и локализация источника звука в мелком море // Акуст. журн. 2017. Т. 63. № 4. С. 406−418.
8. Кузькин В.М., Пересёлков С.А. Голографические методы обработки гидроакустических сигналов (Обзор) // Акуст. журн. 2025. Т. 71. № 1. С. 96−117.
9. Pereselkov S.A., Kuz'kin V.M. Interferometric processing of hydroacoustic signals for the purpose of source localization // J. Acoust. Soc. Am. 2022. Vol. 151. No. 2 P. 666−676.
10. Ehrhardt M., Pereselkov S., Kuz'kin V., Kaznacheev I., Rybyanets P. Experimental observation and theoretical analysis of the low-frequency

source interferogram and hologram in shallow water // Sound and Vibration. 2023. Vol. 544. P. 117388.

11. Глущенко М.Ю., Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Хворостов Ю.А., Ткаченко С.А. Голографический метод идентификации шумового подводного источника в мелком море // Акуст. журн. 2024. Т. 70. № 3. С. 67−77.

12. Pereselkov S., Kuz'kin V., Ehrhardt M., Matvienko Y., Tkachenko S., Rybyanets P. The Formation of 2D Holograms of a Noise Source and Bearing Estimation by a Vector Scalar Receiver in the High-Frequency Band. // J. Mar. Sci. Eng. 2024, 12, 704.

13. Казначеев И.В., Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Интерферометрический метод обнаружения движущегося источника звука векторно-скалярным приемником // Акуст. журн. 2018. Т. 64. № 1. С. 33-45.

14. Kuznetsov G.N., Kuz'kin V.M., Pereselkov S.A., Kaznacheev I.V., Grigor'ev V.A. Interferometric method for estimating the velocity of a noise sound source and the distance to it in shallow water using a vector-scalar receiver // Phys. Wave Phenom. 2017. Vol. 25. No. 4. P. 299-306.

15. Переселков С.А., Кузькин В.М., Кузнецов Г.Н., Просовецкий Д.Ю., Ткаченко С.А. Интерференционный метод оценки координат движущегося шумового источника в мелком море с использованием высокочастотных сигналов // Акуст. журн. 2020. Т. 66. № 4. С. 437-445.

16. Кузнецов Г.Н., Кузькин В.М., Куцов М.В., Переселков С.А. Интерференционная структура шумового поля движущегося источника в высокочастотном диапазоне // Изв. РАН. Сер. физ. 2021. Т. 85. № 6. С. 907-912.

17. Kuz'kin V.M., Pereselkov S.A., Zvyagin V.G., Malykhin Yu.A., Prosovetskiy D.Yu. Intense internal waves and their manifestation in interference patterns of received signals on oceanic shelf // Phys. Wave Phenom. 2018. Vol. 26. No. 2. P. 160-167.

18. Badiey M., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Prosovetskiy D.Yu., Tkachenko S.A. Intense internal waves and their manifestation in the interference patterns of received signals on oceanic shelf. Part II. Phys. Wave Phenom., 2019, 27(4):313-319.

19. Кузькин В.М., Пересёлков С.А., Казначеева Е.С., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Голографическая обработка движущихся источников в мелком море при наличии интенсивных внутренних волн // РЭНСИТ: Радиоэлектроника. Наносистемы. Информационные технологии. 2022. Т. 14. № 2. С. 197-204.

20. Кузькин В.М., Пересёлков С.А., Грачев В.И., Ткаченко С.А., Ладыкин Н.В., Куцов М.В. Голографический метод локализации

движущегося источника звука в присутствии интенсивных внутренних волн // РЭНСИТ: Радиоэлектроника. Наносистемы. Информационные технологии. 2023. Т. 15. № 3. С. 317-326.

21. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Rybyanets P., Ladykin N. Three-dimensional modeling of sound field holograms of a moving source in the presence of internal waves causing horizontal refraction // J. Mar. Sci. Eng. 2023. Vol. 11. P. 1922.

22. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Pereselkov A., Ladykin N. Influence of Intense Internal Waves Traveling Along an Acoustic Path on Source Holographic Reconstruction in Shallow Water. // J. Mar. Sci. Eng. 2025, 13, P. 1409.

23. Pereselkov S., Kuz'kin V., Ehrhardt, M., Tkachenko S., Pereselkov A., Ladykin N. Effects of Irregular Bathymetry on Hologram Formation of Moving Source in ShallowWater. // J. Mar. Sci. Eng. 2025, 13, P. 1775.

24. Беседина Т.Н., Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Определение глубины источника звука в мелком море на фоне интенсивного шума // Акуст. журн. 2015. Т. 61. № 6. С. 718-728.

25. Kuznetsov G.N., Kuz'kin V.M., Pereselkov S.A., Prosovetskiy D.Yu. Wave method for estimating the sound source depth in a oceanic waveguide // Phys. Wave Phenom. 2016. Vol. 24. No. 4. P. 310-316.

26. Кузькин В.М., Пересёлков С.А., Матвиенко Ю.В., Ткаченко С.А. Обнаружение источника в диспергирующих средах // РЭНСИТ. 2019. Т. 11. № 3. С. 337–344.

27. Матвиенко Ю.В., Кузькин В.М., Переселков С.А., Ладыкин Н.В., Грачев В.И. Обнаружение и локализация шумового подводного источника в мелководной акватории на фоне интенсивного судоходства // РЭНСИТ. 2025. Т. 17. № 2. С. 211–220.

28. Kuz'kin V.M., Kuznetsov G.N., Pereselkov S.A., Grigor'ev V.A. Resolving power of the interferometric method of source localization // Phys. Wave Phenom. 2018. Vol. 26. № 2. P. 150–159.

29. Кузькин В.М., Переселков С.А., Грачев В.И., Рыбянец П.В., Ткаченко С.А. Разрешение шумовых сигналов при наличии случайных неоднородностей океанической среды // РЭНСИТ. 2024. Т. 16. № 3. С. 395–406.

30. Kuz'kin V.M., Pereselkov S.A., Rybyanets P.V., Tkachenko S.A. Resolution of hydroacoustic noise signals under conditions of spatial-temporal variability of ocean environment // Phys. Wave Phenom. 2024. Vol. 32. № 4. P. 290–299.

31. Kuz'kin V.M., Pereselkov S.A., Kuznetsov G.N., Kaznacheev I.V. Interferometric direction finding by a vector-scalar receiver // Phys. Wave Phenom. 2018. Vol. 26. № 1. P. 63–73.

32. Kuznetsov G.N., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Prosovetskiy D.Yu. Direction finding of a noise sound source // Phys. Wave Phenom. 2019. Vol. 27. № 3. P. 237–241.

33. Кузькин В.М., Пересёлков С.А., Грачев В.И., Логачев В.В., Косенко И.М. Голографический метод обработки гидроакустических сигналов высокочастотного источника в мелком море // РЭНСИТ. 2024. Т. 16. № 4. С. 545–554.

34. Kaznacheeva E.S., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Tkachenko S.A. Adaptive algorithms for interferometric processing // Phys. Wave Phenom. 2020. Vol. 28. № 3. P. 267–273.

35. Kuz'kin V.M., Pereselkov S.A., Rybyanets P.V., Tkachenko S.F. Adaptive method for holographic processing of broadband hydroacoustic signals // Phys. Wave Phenom. 2024. Vol. 32. № 5. P. 382–391.

36. Кузькин В.М., Пересёлков С.А., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Адаптивный алгоритм локализации шумовых подводных источников в мелководных акваториях // РЭНСИТ. 2024. Т. 16. № 4. С. 533–544.

37. Kuz'kin V.M., Matvienko Yu.V., Pereselkov S.A., Prosovetskii D.Yu., Kaznacheeva E.S. Mode selection in oceanic waveguides // Phys. Wave Phenom. 2022. Vol. 30. No. 2. P. 111-118.

38. Kuz'kin V.M., Matvienko Yu.V., Pereselkov S.A., Kaznacheeva E.S., Tkachenko S.A. Holographic method for mode selection in a shallow sea in the presence of intense internal waves // Phys. Wave Phenom. 2022. Vol. 30. No. 5. P. 314-320.

39. Кузькин В.М., Пересёлков С.А., Казначеева Е.С., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Выделение мод шумового источника в мелком море методом голографической интерферометрии в присутствии интенсивных внутренних волн // РЭНСИТ: Радиоэлектроника. Наносистемы. Информационные технологии. 2022. Т. 14. № 3. С. 279-286.

40. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Rybyanets P., Ladykin N. Use of interference patterns to control sound field focusing in shallow water // J. Mar. Sci. Eng. 2023. Vol. 11. P. 559.

Алгоритм пеленгования на основе разнесенных векторно-скалярных приемников

С. А. Переселков, email: pereselkov@yandex.ru [1]
В. М. Кузькин, email: kumiov@yandex.ru [2]
И. М. Косенко, email: kosenko.edu@yandex.ru [1]
Н. В. Ладыкин, email: ladykin.edu@yandex.ru [1]

[1] Воронежский государственный университет
[2] Институт общей физики им. А.М. Прохорова РАН

Аннотация. *В работе рассматривается алгоритм пеленгования шумового подводного аппарата на основе разнесённых векторно-скалярных приёмников в рамках развития адаптивных голографических методов. Приведены результаты численного моделирования в условиях низкого входного отношения сигнал/помеха.*

Ключевые слова: *голографическая обработка, векторно-скалярный приемник, пеленгование, помехоустойчивость, численное моделирование.*

Введение

В настоящее время активно развивается голографическая обработка звуковых сигналов, излучаемых подводными малогабартиными источниками и принимаемых на основе систем из векторно-скалярных приемников (ВСП) [1–23]. Она базируется на работах предшественников по интерферограмме звуковых полей, формируемых источниками звука из-за особенностей волновода [24, 25]. Благодаря экспериментальной верификации [19–22] показано, что голографический подход, в отличие от известных методов обработки [26–30], эффективно решает задачу пассивной гидроакустики, позволяет обнаруживать источник и оценивать его параметры: радиальную скорость, удаление, пеленг. Отдельным преимуществом голографической обработки является независимость от априорных знаний о среде. Это достигается благодаря адаптивным алгоритмам с использованием нескольких ВСП. В работе представлено развитие одногоиакрнр метода, показаны результаты численного эксперимента на фоне низкого входного отношения сигнал/шум.

© Переселков С. А., Кузькин В. М., Косенко И. М., Ладыкин Н. В., 2026

1. Пеленгация с использованием двух ВСП

Рассмотрим однородные регулярный волновод. На рис. 1 показаны: S – источник сигнала, $Q_{1,2}$ – ВСП. Источник движется с постоянной скоростью.

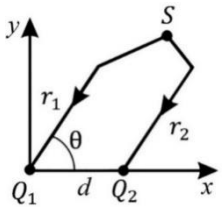


Рис. 1. Геометрия задачи

В ходе наблюдения на векторных каналах приёмника накапливаются независимые временные реализации компонент колебательных скоростей частиц жидкости $V_{X,Y}(t)$. Когерентные, отличающиеся помехоустойчивостью, интерферограммы формируются как $R(f,t)=V_X(f,t)V_X^*(f,t)$ и $L(f,t)=V_Y(f,t)V_Y^*(f,t)$. Для повышения контрастности из них вырезаются средние значения. Голограммы формируются путем двумерного преобразования Фурье, обладающего высокой фокусирующей способностью [33]:

$$F_R(\tau,\nu)=\int_0^{\Delta t}\int_{f_1}^{f_2} R(f,t)\exp[i2\pi(\nu t-f\tau)]dfdt \qquad (1)$$

$$F_L(\tau,\nu)=\int_0^{\Delta t}\int_{f_1}^{f_2} L(f,t)\exp[i2\pi(\nu t-f\tau)]dfdt \qquad (2)$$

Здесь ν и t – частота и время голограммы. После перехода в новую область исчезают трудности, связанные с низким входным отношением сигнал/шум, неоднородностями среды. Для перехода к функции одной переменной строится угловое распределение голограммы как интеграл вдоль лучей, выходящих из начала координат:

$$G_{R,L}(\chi)=\int_0^{\Delta\tau}\left|F_{R,L}(\tau,\chi\tau)\right|d\tau \qquad (3)$$

Пеленг в таком случае может быть вычислен как:

$$\left|tg\theta\right|=\frac{G_L(\varepsilon)}{G_R(\varepsilon)} \qquad (4)$$

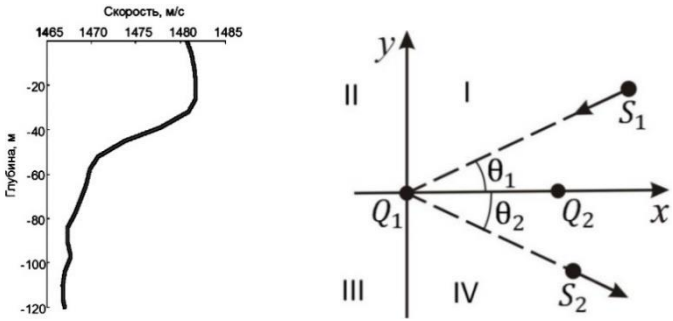
Но такая оценка пеленга неоднозначна с точностью до четверти, т.е. значения $\pm\theta$ и $\pi \pm \theta$ неотличимы. Неоднозначность устраняется согласно алгоритму. Рассчитываем корреляцию между интерферограммами в фиксированный момент времени $\gamma = L(f,t)/R(f,t)$. Если $\gamma > 0$, то пеленг расположен в первом и третьем квадрантах, иначе – во втором и четвертом. Следующий шаг подразумевает вариантность. Первый вариант: применяем одномерное преобразование Фурье по частоте. Для второго варианта применяем преобразование Фурье по времени. В первом случае исследуем знак временной задержки $\tilde{t}$ между огибающими сигнала $S_R(Q_1,t)$ и $S_R(Q_1,t+\tilde{t})$. Во втором случае знак частотного сдвига $\tilde{f}$ между спектрами $u_R(Q_1,f)$ и $u_R(Q_2,f+\tilde{f})$. Если знак положительный, источник в первой или четвертой четверти. Если отрицательный – во второй или третьей.

2. Численный эксперимент

Глубина волновода $H = 120\ м$. На рис. 2(а) показана зависимость скорости звука от глубины. Отношение плотности грунта и воды $\rho = 1.8$, комплексный показатель преломления $n = 0.84(1+i0.03)$. Геометрическое взаимное расположение источников показана на рис. 2(б). Глубина расположения приемника $z_q = 100\ м$, они удалены на $d = 100\ м$. S_1 и S_2 расположены в первой и четвёртой четвертях, на глубине $z_s = 80\ м$, направления на них $\theta_{1,2} = \pm 30^\circ$. Первый источник приближается к приемникам со скоростью $u = 3\ м/с$, а второй удаляется с той же скоростью. В начальный момент времени источники удалены от Q_1 на равное расстояние $r = 7\ км$.

Объекты являются источниками белый шум, помеха характеризуется также равномерно распределенным спектром мощности. Ширина полосы $\Delta f = 80 - 100\ Гц$, время эксперимента $\Delta t = 90\ с$. Поле регистрируется в течение $\Delta T = 5\ с$, интервал между регистрациями $\delta T = 0.1\ с$. Всего таких интервалов 17. Входное отношение сигнал/помеха равно -22.22 дБ.

На рис. 3 демонстрируются в качестве результата моделирования графики интерферограммы, голограммы, углового распределения спектральной плотности голограммы



а б

Рис. 2. Профиль скорости звука (а) и схема расположения источников и приёмников (б)

Наблюдается отсутствие интерференционных полос при высоком отношении сигнал/шум. Фокальные пятна отчетливо просматриваются. Функция углового распределения голограммы демонстрирует данное явление более отчетливо. Отношение абсолютных значений максимумов (4) показывает, что для источника S_1 возможные значения пеленгов: 30.7°, 149.3°, 210.7°, 329.3°, а для второго источника – 29.1°, 150.9°, 209.1°, 330.9°.

На рис. 4 показан результат очистки интерферограммы и голограммы от шума. Для источника S_1 величины $R(f,t)$ и $L(f,t)$ идентичны. А вот для S_2 интерферограммы различаются знаком.

На рис. 5 видно, что для источника S_1 отношение корреляции очищенных интерферограмм положительно, а для источника S_2 – отрицательно. Исключения составляют значения частот, при которых $R(f_*) = L(f_*)$. Благодаря этому и можно сопоставить источники с их четвертями.

Наконец, на рис. 6 приведен модуль взаимокорреляционной функции, определяемой как:

$$B(\tilde{f}) = \int_{-\infty}^{\infty} S_R(Q_1, f) S_R(Q_2, f + \tilde{f}) df \qquad (5)$$

Для обоих источников максимум приходится на частотный сдвиг $\tilde{f} > 0$, следовательно источник S_1 расположен в первой четверти, а S_2 – в четвертой. Оценки пеленгов 30.7° и -29.1°.

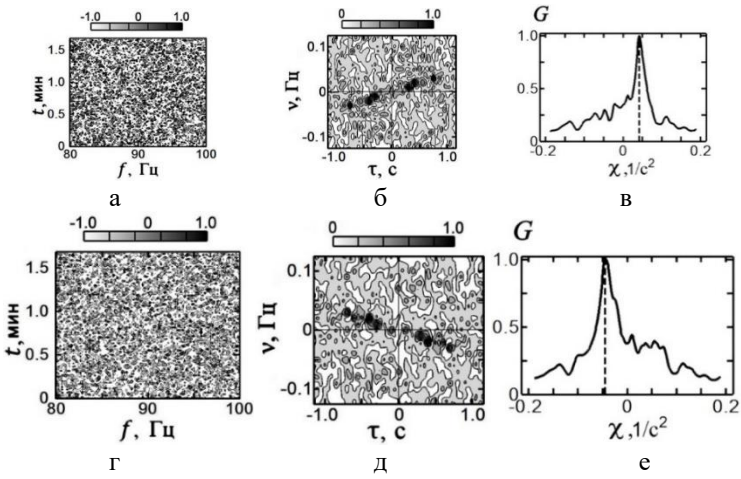


Рис. 3. Интерферограммы (а, г), голограммы(б, д) и угловые распределения голограмм (в, е) для источников S_1 (а – в), S_2 (г – е) и приемника Q_1

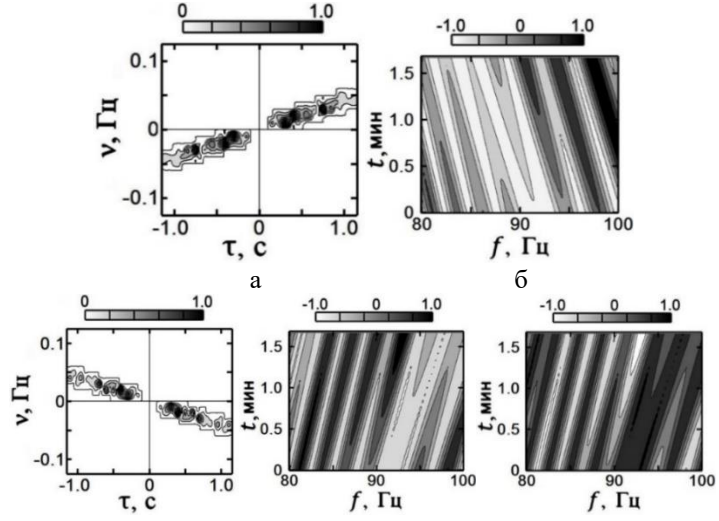


Рис. 4. Нормированные очищенные голограммы и интерферограммы для источников S_1 (а, б), S_2 (в – д) и приемника Q_1

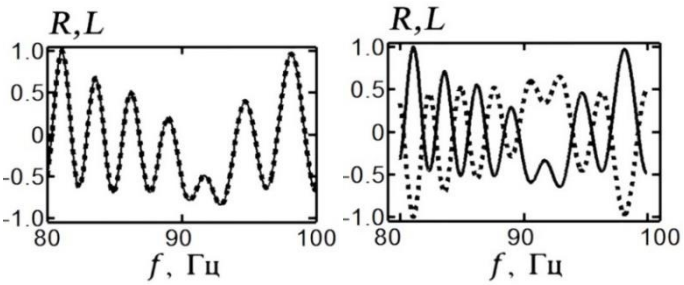


Рис. 5. Интерферограммы при $t = 0$ для источников S_1 и S_2

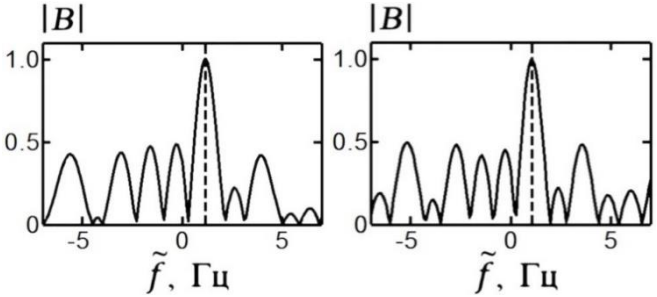


Рис. 6. Нормированная заимокорреляционная функция для источников S_1 (слева) и S_2 (справа).

Заключение

В работе представлен голографический метод пеленгации с использованием двух векторно-скалярных приёмников, использующихся в рамках адаптивных подходов. Важно, что такой метод не требует априорныз знаний о среде распространения. В рамках численного эксперимента показана работоспособность предложенного способа пеленгации даже для высокого отношения сигнал/шум.

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

Список литературы

1. Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Спектрограмма и локализация источника звука в мелком море // Акуст. журн. 2017. Т. 63. № 4. С. 406–418.

2. Казначеев И.В., Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Интерферометрический метод обнаружения движущегося источника звука векторно-скалярным приемником // Акуст. журн. 2018. Т. 64. № 1. С. 33–45.

3. Pereselkov S.A., Kuz'kin V.M. Interferometric processing of hydroacoustic signals for the purpose of source localization // J. Acoust. Soc. Am. 2022. Vol. 151. No. 2 P. 666−676.

4. Ehrhardt M., Pereselkov S., Kuz'kin V., Kaznacheev I., Rybyaners P. Experimental observation and theoretical analysis of the low-frequency source interferogram and hologram in shallow water // J. Sound and Vibration. 2023. Vol. 544. P. 117388.

5. Кузькин В.М., Пересёлков С.А., Грачев В.И., Ткаченко С.А., Ладыкин Н.В., Куцов М.В. Голографический метод локализации движущегося подводного источника звука в присутствии интенсивных внутренних волн // РЭНСИТ. 2023. Т. 15. № 3. С. 317–326.

6. Kuz'kin V.M., Pereselkov S.A., Badiey M., Ladykin N.V., Malykhin A.Yu., Tkachenko S.A. Robustness of holographic processing of hydroacoustic signals in the presence of intense internal waves // Phys. Wave Phenom. 2023. Vol. 31. № 5. P. 346–354.

7. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Rybyanets P., Ladykin N. Three-dimensional modeling of sound field holograms of a moving source in the presence of internal waves causing horizontal refraction // J. Mar. Sci. Eng. 2023. Vol. 11. No. 10, 1922.

8. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Pereselkov A., Ladykin N. Influence of intense internal waves traveling along an acoustic path on source holographic reconstruction in shallow water // J. Mar. Sci. Eng. 2025. Vol. 13. P. 1409.

9. Кузькин В.М., Переселков С.А., Грачев В.И., Косенко И.М., Ладыкин Н.В., Переселков А.С. Локализация движущегося шумового источника в мелководной акватории с нерегулярной батиметрией // РЭНСИТ. 2024. Т. 16. № 8. С. 1081–1088.

10. Kuz'kin V.M., Pereselkov S.A., Rybyanets P.V., Tkachenko S.F. Adaptive method for holographic processing of broadband hydroacoustic signals // Phys. Wave Phenom. 2024. Vol. 32. № 5. P. 382–391.

11. Кузькин В.М., Пересёлков С.А., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Адаптивный алгоритм локализации шумовых подводных источников в мелководных акваториях // РЭНСИТ. 2024. Т. 16. № 4. С. 533–544.

12. Кузькин В.М., Переселков С.А., Грачев В.И., Рыбянец П.В., Ткаченко С.А. Разрешение шумовых сигналов при наличии случайных

неоднородностей океанической среды // РЭНСИТ. 2024. Т. 16. № 3. С. 395–406.

13. Kuz'kin V.M., Pereselkov S.A., Rybyanets P.V., Tkachenko S.A. Resolution of hydroacoustic noise signals under conditions of spatial-temporal variability of ocean environment // Phys. Wave Phenom. 2024. Vol. 32. № 4. P. 290–299.

14. Kuz'kin V.M., Pereselkov S.A., Matvienko Yu.V., Tkachenko S.A. Detection of a noise signal in an oceanic waveguide using a vertical array // Phys. Wave Phenom. 2021. Vol. 30. № 4. P. 323–329.

15. Кузькин В.М., Пересёлков С.А., Матвиенко Ю.В., Ткаченко С.А., Просовецкий Д.Ю. Обнаружение шумового сигнала в океаническом волноводе горизонтальной антенной // РЭНСИТ. 2022. Т. 14. № 1. С. 65–72. (Kuz'kin V.M., Pereselkov S.A., Matvienko Yu.V., Tkachenko S.A., Prosovetskii D.Yu. Detection of a noisy signal in an oceanic waveguide using a horizontal array // RENSIT. 2022. Vol. 14. No. 1. P. 65–72.)

16. Kaznacheev I.V., Kuz'kin V.M., Kutsov M.V., Lyakhov G.A., Pereselkov S.A. Interferometry in acoustic-data processing using extended antennas. Space-time analogy // Phys. Wave Phenom. 2020. Vol. 28. № 4. P. 326–332.

17. Kuz'kin V.M., Matvienko Yu.V., Pereselkov S.A., Prosovetskii D.Yu., Kaznacheeva E.S. Mode selection in oceanic waveguides // Phys. Wave Phenom. 2022. Vol. 30. № 2. P. 111–118.

18. Kuz'kin V.M., Matvienko Yu.V., Pereselkov S.A., Kaznacheeva E.S., Tkachenko S.A. Holographic method for mode selection in a shallow sea in the presence of intense internal waves // Phys. Wave Phenom. 2022. Vol. 30. № 5. P. 314–320.

19. Кузькин В.М., Пересёлков С.А., Грачев В.И., Логачев В.В., Косенко И.М. Голографический метод обработки гидроакустических сигналов высокочастотного источника в мелком море // РЭНСИТ. 2024. Т. 16. № 4. С. 545–554.

20. Матвиенко Ю.В., Кузькин В.М., Переселков С.А., Ладыкин Н.В., Грачев В.И. Обнаружение и локализация шумового подводного источника в мелководной акватории на фоне интенсивного судоходства // Радиоэлектроника. Наносистемы. Информационные технологии. 2025. Т. 17. № 2. С. 211–220.

21. Глущенко М.Ю., Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Хворостов Ю.А., Ткаченко С.А. Голографический метод локализации шумового подводного источника в мелком море // Акуст. ж. 2024. Т. 70. № 3. С. 67–77.

22. Pereselkov S., Kuz'kin V., Ehrhardt M., Matvienko Yu., Tkachenko S., Rybyanets P. The formation of 2D holograms of a noise source and bearing

estimation by a vector scalar receiver in the high-frequency band // J. Mar. Sci. Eng. 2024. Vol. 12. P. 704.

23. Кузькин В.М., Пересёлков С.А. Голографические методы обработки гидроакустических сигналов (Обзор) // Акуст. журн. 2025. Т. 71. № 1. С. 96–117.

24. Weston D., Stevens K. Interference of wide-band sound in shallow water // J. Sound Vibr. 1972. Vol. 21. No. 1. P. 57–64.

25. Чупров С.Д. Интерференционная структура звукового поля в слоистом океане // Акустика океана. Современное состояние. М.: Наука, 1982. С. 71–82.

26. Jackson D.R. Ewart T.E. The effect of internal waves on matched-field processing // J. Acoust. Soc. Am. 1994. Vol. 96. № 5. P. 2945–2955.

27. Baggeroer A.B., Kuperman W.A., Schmidt H. Matched field processing: Source localization in correlated noise as an optimum parameter estimation problem // J. Acoust. Soc. Am. 1998. Vol. 83. № 2. P. 571–587.

28. Dosso S.E., Nielsen P.L., Wilmut M.J. Data error covariance in matched-field geoacoustic inversion // J. Acoust. Soc. Am. 2006. Vol. 119. № 1. P. 208–219.

29. Baggeroer A.B. Why did applications of MFP fail, or did we not understand how to apply MFP? // Proc. 1st Int. Conf. and Exhibition. Underwater Acoustics / Eds. Papadakis J.S., Bjorno L. Corfu Island, Greece, 2013. P. 41–49.

30. Сазонтов А.Г., Малеханов А.И. Согласованная пространственная обработка сигналов в подводных звуковых каналах (Обзор) // Акуст. журн. 2015. Т. 61. № 2. С. 233–253.

31. Kuznetsov G.N., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Prosovetskiy D.Yu. Direction finding of a noise sound source // Phys. Wave Phenom. 2019. V. 27. No. 3. P. 237−241.

32. Weston D., Stevens K. Interference of wide-band sound in shallow water // J. Sound Vibr. 1972. Vol. 21. No. 1. P. 57–64.

33. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Rybyanets P., Ladykin N. Use of interference patterns to control sound field focusing in shallow water // J. Mar. Sci. Eng. 2023. Vol. 11. P. 559.

Алгоритм формирования голограммы источника ЛЧМ-сигнала на основе дробного преобразования Фурье

С. А. Переселков, email: pereselkov@yandex.ru [1]
В. М. Кузькин, email: kumiov@yandex.ru [2]
И. М. Косенко, email: kosenko.edu@yandex.ru [1]
Н. В. Ладыкин, email: kosenko.edu@yandex.ru [1]

[1] Воронежский государственный университет
[2] Институт общей физики им. А.М. Прохорова РАН

Аннотация. *В работе представлен алгоритм формирования голограммы источника линейно частотно модулированного сигнала на основе дробного преобразования Фурье в мелководном волноводе. Показано, что использование FrFT позволяет формировать локализованные фокусные области даже в условиях интенсивного шума, положение которых связано с параметрами источника. Эффективность предложенного алгоритма подтверждена результатами численного моделирования в диапазоне частот 100–150 Гц.*

Ключевые слова: *дробное преобразование Фурье, голографический метод, интерферограмма, линейная частотная модуляция, мелководье.*

Введение

Дробное преобразование Фурье (FrFT), предложенное в 1980 году [1], расширяет классическое преобразование Фурье (FT), обеспечивая более обобщённый и гибкий анализ сигналов. Оно интерпретируется как вращение в плоскости «время–частота». Благодаря лучшей локализации спектра в согласованном порядке, FrFT активно применяется в квантовой механике, оптике, акустике и обработке сигналов. Наиболее эффективная реализация в этой области связана с STFrFT [2]. В работе представлен метод голографической обработки сигналов (ГОС) [3–6] на основе STFrFT. Современные подходы к ГОС основаны на том, что интерференционная картина, возникающая при излучении звука в океане, обладает высокой устойчивостью и несёт информативные пространственно-частотные признаки [3–8]. Структура и положение её локальных максимумов определяются характеристиками частотного спектра, параметрами волновода и кинематикой движущегося источника. ГОС методы для обнаружения и локализации источников, применимые в

© Переселков С. А., Кузькин В. М., Косенко И. М., Ладыкин Н. В. 2026

широком частотном диапазоне — от низких до высоких частот [9–12], демонстрируют устойчивость к шумам, минимальную чувствительность к изменчивости гидрологических условий и сохраняют эффективность в условиях сложной нестационарности среды [13–19]. Интеграция ГОС с алгоритмами, оценивающими глубину источника [20,21], обеспечивает высокую помехоустойчивость при решении задач контроля подводной обстановки: от выявления [22,23] и разделения объектов [24–26] до определения направления, радиальной скорости, расстояния [27–29] и глубины [20,21] малошумных подводных аппаратов. Помимо этого, ГОС позволяет применять адаптивные процедуры обработки [30–32], которые не требуют предварительных сведений о свойствах среды, а также выполнять модовую фильтрацию в различных типах волноводов — регулярных и нерегулярных [33–35]. По способности к пространственно-временной фокусировке он превосходит классические методы, включая обращение волнового фронта и временной реверс к положению исходного источника [36]. Однако применение ГОС к LFM-сигналам в условиях сильных помех ранее почти не изучалось. Настоящая работа посвящена построению интерферограммы LFM-источника и её преобразованию в голограмму с использованием FrFT. Полученные результаты численных экспериментов в диапазоне 100–150 Гц показывают, что предложенный метод обеспечивает устойчивую локализацию источника даже при уровне шума, существенно превышающем уровень сигнала.

1. Дробное преобразование Фурье

Дробное преобразование (FrFT) $S_p(u)$ для сигнала $s(t)$ описывается следующим образом:

$$S_p(u) = \int_{-\infty}^{+\infty} s(t) K_p(t,u)\,dt \tag{1}$$

где $p = 2\alpha/\pi$ (α – угол поворота преобразования), а $K_p(t,u)$ – ядро, определяемое как ($\delta(u)$ –функция Дирака)

$$K_p(t,u) = \begin{cases} \sqrt{(1 - i\cot\alpha)/(2\pi)}\exp\left(0.5i\left(t^2 + u^2\right)\cot\alpha - itu\csc\alpha\right) \\ \delta(u-t), \alpha = 2k\pi \\ \delta(u+t), \alpha = (2k+1)\pi \end{cases} \tag{2}$$

На практике обычно применяется оконное преобразование STFrFT $S_p(t,u)$, задаваемое соотношением. $w(t)$ – оконная функция, в нашем примере – функция Гаусса:

$$S_p(t,u) = \int_{-\infty}^{+\infty} s(\tau) w(\tau - t) K_p(\tau,u) d\tau \ , \qquad w(t) = (\pi\sigma^2)^{-1/4} e^{-t^2/(2\sigma^2)} \tag{3}$$

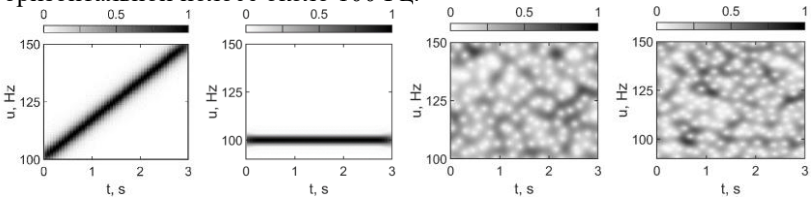
Использование FrFT эквивалентно повороту на произвольный угол во временно-частотной плоскости. Для LFM сигнала, например, это позволяет превратить наклонную спектральную линию в вертикальную, тем самым уплотняя её энергию. Подобный механизм существенным образом увеличивает эффективность голографического подхода.

2. Пеленгатор с одной парой ортоканалов

Расчёты выполнялись для полосы частот $\Delta f = 100 - 150$ Гц. Глубина волновода составляет $H = 50$ *м*, а скорость звука в воде - $c = 1500$ *м/с*. Дно моделируется комплексным показателем преломления $n_b = 0.84(1 + i0.03)$ и плотностью $\rho_b = 1.8$ *г/см*^{3}. Приёмник расположен на расстоянии $r_q = 0$ и глубине $z_q = 45$ *м*, тогда как источник в начальный момент времени находится на $r_s = 10$ *км* и глубине $z_s = 15$ *м* и движется к приемнику со скоростью 3 *м/с*. LFM-сигнал имеет единичную амплитуду и конечную длительность $\tau = 3$ *с*; вне этого интервала он равен нулю. Для уменьшения спектрального «размазывания» на начала и конец сигнала накладываются плавные фронты. Мгновенная частота линейно возрастает от $f_1 = 100$ Гц при $t = 0$ до $f_2 = 150$ Гц при $t = \tau = 3$ *с*, то есть реализуется up-chirp с наклоном $k = (f_2 - f_1)/\tau = 16.667$ Гц/с. Энергия сигнала сосредоточена в полосе 100–150 Гц (ширина полосы $\Delta f = 50$ Гц; центральная частота $f_0 \approx 125$ Гц), причём внутри полосы спектр имеет распределение близкое к равномерному.

На рис. 1 продемонстрирован STFrFT излученного и принятого LFM-сигнала $s(t)$ для двух порядков в случае отсутствия шума. Излученный сигнал при $p = 1$ (то есть стандартный STFT), показывает характерный up-chirp в виде яркого хребта, поднимающегося от 100 до 150 Гц за 0–3 с. В отличие от этого, при $p = 0.994$ ($\alpha \approx 1.561$ рад) преобразование фактически «поворачивает» принятый сигнал во временно–частотной области: спектр LFM-сигнала практически

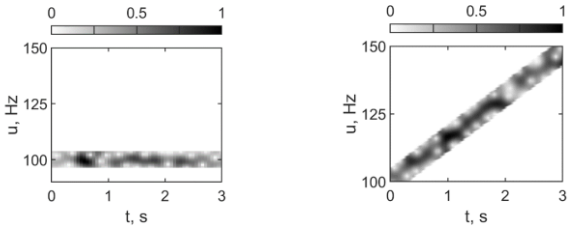
выпрямляется, и его энергия концентрируется в узкой, почти горизонтальной полосе около 100 Гц.



а б в г

Рис. 1. STFrFT излученного и принятого зашумлённого LFM-сигнала: (а),(в) $S(t,u)$ - STFrFT $s(t)$ при $p=1$; (б),(г) $S(t,u)$ - STFrFT $s(t)$ при $p=0.994$. Шум равен -30 дБ

На рис. 1(в, г) показан результат STFrFT принятого LFM-сигнала $s(t)$ при $SNR=-30$ дБ. На рис. 1(в), при $p=1$ (стандартный STFT), сильный диффузный шум, и LFM-сигнал неразличим. Но на рис. 1(г), при $p=0.994$, STFrFT эффективно де-чирпирует сигнал и концентрирует энергию в узкой, почти горизонтальной полосе около 100 Гц благодаря выбору дробного порядка, близкого к скорости изменения частоты LFM-сигнала. Это позволяет эффективно отфильтровать принятый LFM-сигнал от шума.



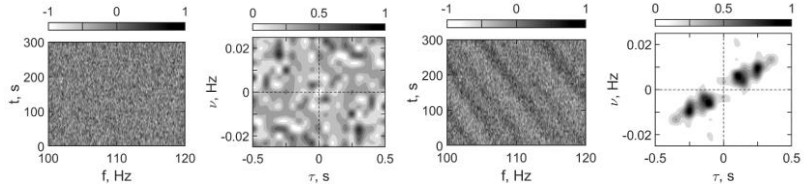
а б

Рис. 2. STFrFT принятого зашумлённого LFM-сигнала после фильтрации шума: (a) $S(t,u)$ - STFrFT $s(t)$, $p=0.994$; (b) $S(t,u)$ - STFrFT $s(t)$, $p=1$. Шум равен -30 дБ

На рис. 2 показан результат фильтрации шума в области STFrFT для зашумленного принятого LFM-сигнала при входном $SNR=-30$ дБ. Принятый LFM-сигнал фильтруется в области STFrFT, как показано на рис. 2(а), при порядке, согласованном с параметром LFM-сигнала (

$p = 0.994$). В этом случае представление сигнала эффективно дечирпируется, то есть энергия сжимается в узкую, почти горизонтальную полосу около 100 Гц. Такой подход обеспечивает высокую эффективность фильтрации. После фильтрации LFM-сигнала применяется обратное STFrFT ($p = 0.994$). Затем к очищенному LFM-сигналу применяется прямое STFrFT ($p = 1$). Рис. 2(б) показывает, что при использовании стандартного STFT энергия восстановленного принятого LFM-сигнала совпадает с ожидаемой структурой LFM-сигнала. Таким образом, мы получаем спектральное распределение очищенного от шума LFM-сигнала, которое может быть использовано для построения интерферограммы и голограммы источника.

На рис. 3(а, б) показаны результаты голографической обработки с помощью традиционного FT для LFM-сигнала в присутствии значительной шумовой помехи ($\Delta f = 100 - 120$ Гц, $v = 3$ *м/с*, $SNR = -30$ дБ). Отсутствие фильтрации в области STFrFT заметно ухудшает результат голографической обработки. В области интерферограммы на рис. 3(а) структура полностью маскируется широкополосным шумом, и наклонные квазипериодические интерференционные полосы, ожидаемые для движущегося источника, не просматриваются.



а б в г

Рис. 3. Голографическая обработка LFM-сигнала на основе традиционного FT: (а) интерферограмма; (б) голограмма. Голографическая обработка LFM-сигнала на основе STFrFT: (в) интерферограмма (г) голограмма. Шум равен -30 дБ

В области голограммы на рис. 3(б) энергия распределена диффузно. Фокальные пятна не удаётся локализовать: они сильно замаскированы шумом. Отсутствие хорошо выраженной структуры не позволяет определить координаты фокальных пятен и, следовательно, делает невозможной оценку параметров источника. Эти результаты подчёркивают необходимость шумоподавления в области STFrFT перед голографической обработкой сигнала в условиях низкого *SNR* (

$SNR \le -30$ дБ). На рис. 3(в, г) показаны результаты голографической обработки на основе STFrFT LFM-сигнала в присутствии значительной шумовой помехи ($\Delta f = 100 - 120$ Гц, $v = 3$ *м/с*, $SNR = -30$ дБ). Применение фильтра существенно улучшает результат голографической обработки. В области интерферограммы на рис. 3(в) структура во времени и частоте становится чёткой: на всей полосе видны наклонные квазипериодические интерференционные полосы с ненулевым наклоном $\delta f / \delta t$, что согласуется с частотным сдвигом из-за движения источника. В области голограммы на рис. 3(г) энергия фокусируется в компактных фокальных пятнах, смещённых относительно оси $v = 0$.

Голографическая обработка выполнена на основе STFrFT фильтрации позволяет получить следующие результаты (рис. 3). Интерферограмма демонстрирует ненулевой частотный сдвиг, $\delta f / \delta t \approx -0.036$ c^{-2} обусловленный движением источника. В области голограммы доминирующее фокальное пятно расположено при $\tau_1 = 1.15 \cdot 10^{-1}$ *c* и $\nu_1 = 3.81 \cdot 10^{-3}$ Гц, что согласуется с частотным сдвигом интерферограммы. Эти голографические параметры приводят к оценке параметров сигнала: $\dot{v} = -3.82$ *м/с* и $\dot{r}_0 = 9.6$ *км*. Эти значения хорошо согласуются с модельными параметрами движущегося источника. По сравнению с модельными параметрами ($v = 3$ *м/с*, $r_0 = 10$ *км*), ошибки составляют около 27.3% по скорости и 4% по дальности. Представленные результаты исследования демонстрируют, что использование подавления шума на основе STFrFT в качестве предварительного этапа в голографической обработке сигналов является высокоэффективным для LFM-сигналов от подвижного источника в мелководном волноводе при низком отношении сигнал/шум. При $SNR = -30$ дБ разработанный метод существенно понижает спектральный уровень шума, восстанавливает интерферометрическую структуру и формирует хорошо локализованные фокальные пятна в голограмме. Это повышает обнаруживаемость, а также точность и устойчивость последующих оценок параметров сигнала по сравнению с традиционной обработкой на основе FT.

Заключение

Предложен метод голографической обработки сигналов на основе FrFT для LFM-сигнала движущегося источника в мелководном волноводе. Показано, что такой метод обеспечивает надежное определение дальности и радиальной скорости источника даже при очень высоких уровнях шума ($SNR \le -30$ дБ). Голограмма слабо искажается из-за помех, фокальные точки выстраиваются вдоль прямой и напрямую

соответствуют параметрам источника. Для движущегося источника расстояние восстанавливается с погрешностью до 5%, скорость — с погрешностью до 30%: в отсутствие шума ошибки скорости и дальности составляют 15% и 1%, а при с/п = −30 дБ – 27% и 4%.

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

Список литературы

1. Namias V. The fractional order Fourier transform and its application to quantum mechanics // IMA J. Appl. Math. 1980. Vol. 25. P. 241–265.
2. Tao, R.; Li, Y.-L.; Wang, Y. Short-time fractional Fourier transform and its applications. IEEE Trans. Signal Process. 2010, 58, 2568–2580.
3. Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Спектрограмма и локализация источника звука в мелком море // Акуст. журн. 2017. Т. 63. № 4. С. 406−418.
4. Кузькин В.М., Пересёлков С.А. Голографические методы обработки гидроакустических сигналов (Обзор) // Акуст. журн. 2025. Т. 71. № 1. С. 96−117.
5. Pereselkov S.A., Kuz'kin V.M. Interferometric processing of hydroacoustic signals for the purpose of source localization // J. Acoust. Soc. Am. 2022. Vol. 151. No. 2 P. 666−676.
6. Ehrhardt M., Pereselkov S., Kuz'kin V., Kaznacheev I., Rybyanets P. Experimental observation and theoretical analysis of the low-frequency source interferogram and hologram in shallow water // Sound and Vibration. 2023. Vol. 544. P. 117388.
7. Глущенко М.Ю., Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Хворостов Ю.А., Ткаченко С.А. Голографический метод идентификации шумового подводного источника в мелком море // Акуст. журн. 2024. Т. 70. № 3. С. 67−77.
8. Pereselkov S., Kuz'kin V., Ehrhardt M., Matvienko Y., Tkachenko S., Rybyanets P. The Formation of 2D Holograms of a Noise Source and Bearing Estimation by a Vector Scalar Receiver in the High-Frequency Band. // J. Mar. Sci. Eng. 2024, 12, 704.
9. Казначеев И.В., Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Интерферометрический метод обнаружения движущегося источника звука векторно-скалярным приемником // Акуст. журн. 2018. Т. 64. № 1. С. 33-45.
10. Kuznetsov G.N., Kuz'kin V.M., Pereselkov S.A., Kaznacheev I.V., Grigor'ev V.A. Interferometric method for estimating the velocity of a noise

sound source and the distance to it in shallow water using a vector-scalar receiver // Phys. W. Ph. 2017. Vol. 25. No. 4. P. 299-306.

11. Переселков С.А., Кузькин В.М., Кузнецов Г.Н., Просовецкий Д.Ю., Ткаченко С.А. Интерференционный метод оценки координат движущегося шумового источника в мелком море с использованием высокочастотных сигналов // Акуст. журн. 2020. Т. 66. № 4. С. 437-445.

12. Кузнецов Г.Н., Кузькин В.М., Куцов М.В., Переселков С.А. Интерференционная структура шумового поля движущегося источника в высокочастотном диапазоне // Изв. РАН. Сер. физ. 2021. Т. 85. № 6. С. 907-912.

13. Kuz'kin V.M., Pereselkov S.A., Zvyagin V.G., Malykhin Yu.A., Prosovetskiy D.Yu. Intense internal waves and their manifestation in interference patterns of received signals on oceanic shelf // Phys. W. Ph. 2018. Vol. 26. No. 2. P. 160-167.

14. Badiey M., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Prosovetskiy D.Yu., Tkachenko S.A. Intense internal waves and their manifestation in the interference patterns of received signals on oceanic shelf. Part II. Phys. W. Ph., 2019, 27(4):313-319.

15. Кузькин В.М., Пересёлков С.А., Казначеева Е.С., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Голографическая обработка движущихся источников в мелком море при наличии интенсивных внутренних волн // РЭНСИТ. 2022. Т. 14. № 2. С. 197-204.

16. Кузькин В.М., Пересёлков С.А., Грачев В.И., Ткаченко С.А., Ладыкин Н.В., Куцов М.В. Голографический метод локализации движущегося источника звука в присутствии интенсивных внутренних волн // РЭНСИТ. 2023. Т. 15. № 3. С. 317-326.

17. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Rybyanets P., Ladykin N. Three-dimensional modeling of sound field holograms of a moving source in the presence of internal waves causing horizontal refraction // J. Mar. Sci. Eng. 2023. Vol. 11. P. 1922.

18. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Pereselkov A., Ladykin N. Influence of Intense Internal Waves Traveling Along an Acoustic Path on Source Holographic Reconstruction in Shallow Water. // J. Mar. Sci. Eng. 2025, 13, P. 1409.

19. Pereselkov S., Kuz'kin V., Ehrhardt, M., Tkachenko S., Pereselkov A., Ladykin N. Effects of Irregular Bathymetry on Hologram Formation of Moving Source in ShallowWater. // J. Mar. Sci. Eng. 2025, 13, P. 1775.

20. Беседина Т.Н., Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Определение глубины источника звука в мелком море на фоне интенсивного шума // Акуст. журн. 2015. Т. 61. № 6. С. 718-728.

21. Kuznetsov G.N., Kuz'kin V.M., Pereselkov S.A., Prosovetskiy D.Yu. Wave method for estimating the sound source depth in oceanic waveguide // Phys. W. Ph. 2016. Vol. 24. No. 4. P. 310-316.

22. Кузькин В.М., Пересёлков С.А., Матвиенко Ю.В., Ткаченко С.А. Обнаружение источника в диспергирующих средах // РЭНСИТ. 2019. Т. 11. № 3. С. 337–344.

23. Матвиенко Ю.В., Кузькин В.М., Переселков С.А., Ладыкин Н.В., Грачев В.И. Обнаружение и локализация шумового подводного источника в мелководной акватории на фоне интенсивного судоходства // РЭНСИТ. 2025. Т. 17. № 2. С. 211–220.

24. Kuz'kin V.M., Kuznetsov G.N., Pereselkov S.A., Grigor'ev V.A. Resolving power of the interferometric method of source localization // Phys. W. Ph. 2018. Vol. 26. № 2. P. 150–159.

25. Кузькин В.М., Переселков С.А., Грачев В.И., Рыбянец П.В., Ткаченко С.А. Разрешение шумовых сигналов при наличии случайных неоднородностей океанической среды // РЭНСИТ. 2024. Т. 16. № 3. С. 395–406.

26. Kuz'kin V.M., Pereselkov S.A., Rybyanets P.V., Tkachenko S.A. Resolution of hydroacoustic noise signals under conditions of spatial-temporal variability of ocean environment // Phys. W. Ph. 2024. Vol. 32. № 4. P. 290–299.

27. Kuz'kin V.M., Pereselkov S.A., Kuznetsov G.N., Kaznacheev I.V. Interferometric direction finding by a vector-scalar receiver // Phys. W. Ph. 2018. Vol. 26. № 1. P. 63–73.

28. Kuznetsov G.N., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Prosovetskiy D.Yu. Direction finding of a noise sound source // Phys. W. Ph. 2019. Vol. 27. № 3. P. 237–241.

29. Кузькин В.М., Пересёлков С.А., Грачев В.И., Логачев В.В., Косенко И.М. Голографический метод обработки гидроакустических сигналов высокочастотного источника в мелком море // РЭНСИТ. 2024. Т. 16. № 4. С. 545–554.

30. Kaznacheeva E.S., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Tkachenko S.A. Adaptive algorithms for interferometric processing // Phys. W. Ph. 2020. Vol. 28. № 3. P. 267–273.

31. Kuz'kin V.M., Pereselkov S.A., Rybyanets P.V., Tkachenko S.F. Adaptive method for holographic processing of broadband hydroacoustic signals // Phys. W. Ph. 2024. Vol. 32. № 5. P. 382–391.

32. Кузькин В.М., Пересёлков С.А., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Адаптивный алгоритм локализации шумовых подводных источников в мелководных акваториях // РЭНСИТ. 2024. Т. 16. № 4. С. 533–544.

33. Kuz'kin V.M., Matvienko Yu.V., Pereselkov S.A., Prosovetskii D.Yu., Kaznacheeva E.S. Mode selection in oceanic waveguides // Phys. W. P. 2022. Vol. 30. No. 2. P. 111-118.

34. Kuz'kin V.M., Matvienko Yu.V., Pereselkov S.A., Kaznacheeva E.S., Tkachenko S.A. Holographic method for mode selection in a shallow sea in the presence of intense internal waves // Phys. W. Ph. 2022. Vol. 30. No. 5. P. 314-320.

35. Кузькин В.М., Пересёлков С.А., Казначеева Е.С., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Выделение мод шумового источника в мелком море методом голографической интерферометрии в присутствии интенсивных внутренних волн // РЭНСИТ. 2022. Т. 14. № 3. С. 279-286.

36. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Rybyanets P., Ladykin N. Use of interference patterns to control sound field focusing in shallow water // J. Mar. Sci. Eng. 2023. Vol. 11. P. 559.

Компьютерный эксперимент по оценке дальности обнаружения источника звука в нерегулярном волноводе.

С. А. Переселков, email: pereselkov@yandex.ru [1]
В. М. Кузькин, email: kumiov@yandex.ru [2]
И. М. Косенко, email: kosenko.edu@yandex.ru [1]
Н. В. Ладыкин, email: ladykin.edu@yandex.ru [1]

[1] Воронежский государственный университет
[2] Институт общей физики им. А.М. Прохорова РАН

Аннотация. *Представлены результаты численного моделирования по оценке дальности обнаружения подводного источника. Выполнено сравнение в условиях регулярного и нерегулярного волновода. В качестве нерегулярности рассматривается батиметрия в форме берегового клина.*

Ключевые слова: *голографический метод, нерегулярный волновод, береговой клин, горизонтальная рефракция волн, дальность обнаружения, отношение сигнал/шум.*

Введение

Голографический метод, ведущая своё развитие с работы С.Д. Чупрова [1] на данные момент является крайне перспективным направлением пассивной гидроакустики. Изображение источника в нем строится посредством двумерного преобразования Фурье, обеспечивая квазикогерентное накопление сигнала вдоль полос интерферограммы и некогерентное накопление помех. Ранее показана связь спектральной плотности голограммы с параметрами источника [2–4], дана схема разделения нескольких источников [5, 6], подтверждена устойчивость к неоднородностям среды [7–18], а также разработаны адаптивные методы при отсутствии данных о среде [19–22].

Ранее [23–28] методом численного моделирования была исследована задача обнаружения подводного движущегося источника широкополосного шума. Рассматривалась задача на случай мелководного волновода с нерегулярной батиметрией в виде берегового клина. В результате были получены выводы о дальности обнаружения источника, шумящего в диапазоне 300–350 Гц в условиях пространственно-временной изменчивости среды. Нерегулярность такого вида приводит к явлению горизонтальной рефракции мод. В работе использованы экспериментальные спектры аппарата и фоновой помехи [29]. Для

© Переселков С. А., Кузькин В. М., Косенко И. М., Ладыкин Н. В. 2026

рассмотренного одиночного приемника выполнено сравнение регулярной и нерегулярной акватории. Вероятности правильного обнаружения при заданных вероятностях ложной тревоги определены по критерию Неймана–Пирсона.

Численное моделирование

На рис. 1 показана геометрия задачи. Приемник представлен одиночным приемником и трехэлементной горизонтальной антенной ($B = 3$). Глубина волновода изменяется по закону $z(x, y) = H_0 + ky$. Здесь $H_0 = 50$ *м*, $k = tg\alpha = 0.01$. Параметры однородного жидкого дна: комплексный показатель преломления $n = 0.84(1 + i0.013)$, плотность $\rho = 1.8$ *г/см*^{3}. Скорость звука в воде $c = 1480$ *м/с*. Глубина волновода $H = 60$ *м*.

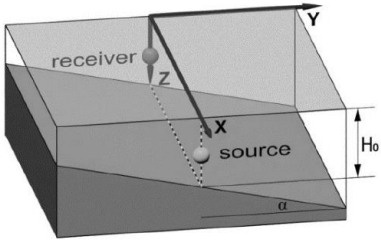


Рис. 1. 3D схема экспериментальной задачи

Источник удаляется от приемника со скоростью $w = 2$ *м/с*. В момент времени $t_0 = 0$ горизонтальное расстояние $r_0 = 200$ *м*. Источник акустических колебаний расположен на глубине $z_s = 30$ *м*, приемник $z_q = 50$ *м*. Частотный диапазон $\Delta f = 300 - 350$ *Гц*. Средняя частота спектра $f_0 = 325$ *Гц*. Время наблюдения $\Delta t = 30$ *с*, длительность шумовой реализации $t_1 = 1$ *с*, интервал между ними $t_2 = 0.5$ *с*, количество таких временных интервалов $J = \Delta t/(t_1 + t_2) = 20$. Число мод $M = 10$.

В полосе $\Delta f = 50$ *Гц* шаг дискретизации по частоте $\delta f = 0.2$ *Гц*, число гармонических составляющих $K = 251$. В полосе Δf : средняя мощность помехи $N_{\Delta f} = -43.01$ дБ, средняя мощность шумоизлучения подводного аппарата на расстоянии $P_{\Delta f}(r_0) = -30.51$ дБ [29].

Результат в виде зависимости средней мощности шумоизлучения $P_{\Delta f}(r)$ для одиночного приемника в зависимости от расстояния r представлен на рис.2. Причиной осцилляций мощности излучения являются межмодовые интерференции. Для расстояния 12 км уровень сигнала составляет $P_{\Delta f} = -54.32$ дБ и начиная с этого расстояния осцилляции поля практически полностью сходят на нет. Начиная с расстояния 12 км поле звукового давления формируется одной модой, интерференционная картина в таких условиях не формируется. Горизонтальные пунктиры соответствуют значениям средних мощностей -42.97, -52.97 дБ для входных отношений с/п $q_0 = 0.04, -9.96$ дБ. Вертикальными пунктирами отмечены расстояния $r = 1.57, 9.46$ *км* для этих отношений.

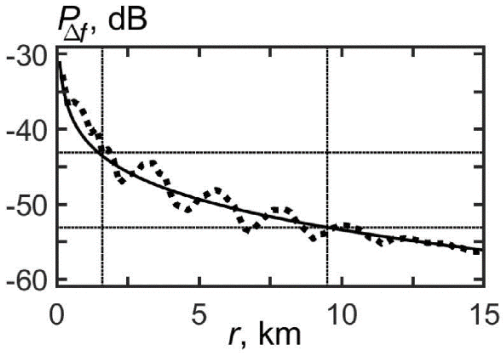


Рис. 2. Зависимость средней мощности шумоизлучения $P_{\Delta f}(r)$ от расстояния r

На рис. 3 показаны искажения поля в результате влияния горизонтальной рефракции. Заметно, что такой влияния по отношению к моде структуре носит избирательный характер. Влияние горизонтальной рефракции больше для большего номера моды. Объяснение данной закономерности связано с тем, что при увеличении номера моды уменьшается горизонтальное волновое число. В нашем эксперименте угол рефракции α горизонтальных лучей в точке приема на расстоянии $r = 10$ *км* для десятой моды оценивается как $\alpha = 7^{\circ}$.

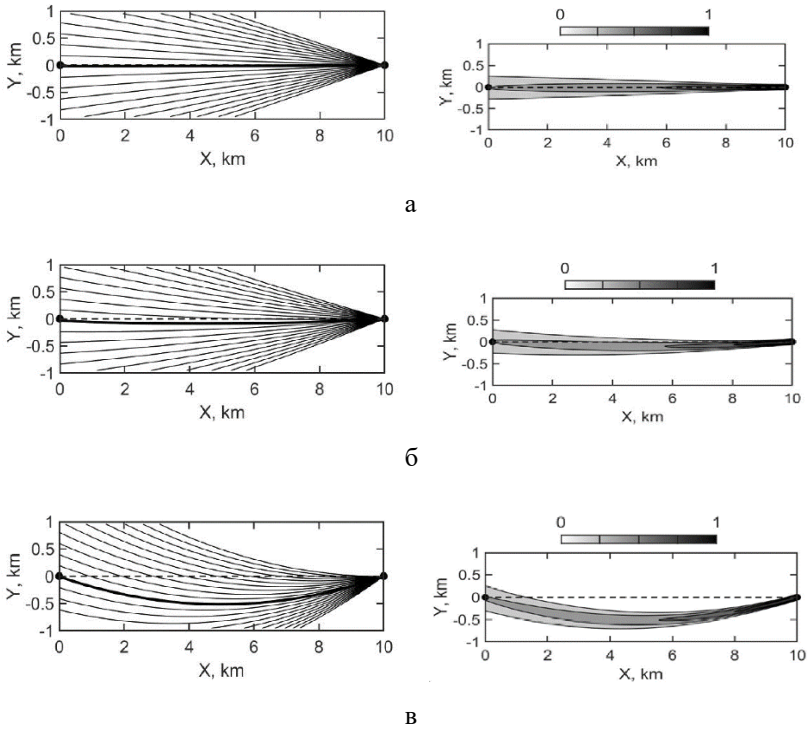
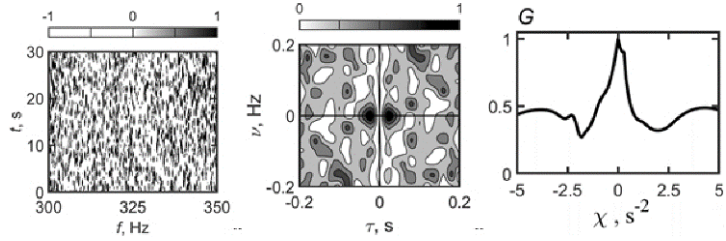


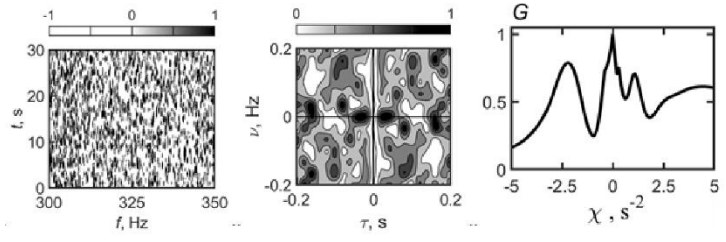
Рис. 3. Нормированная горизонтальная структура поля на частоте $f_0 = 325$ *Гц* . Слева траектории горизонтальных лучей, справа распределения мод. (а) 1-я мода, (б) 4-я мода, (в) 10-я мода.

На рис. 4 показаны результаты голографической обработки в переходной области. Из рисунков делаем вывод о расположении максимального расстояния обнаружения между шумящего объекта в диапазоне от $r = 9.2$ *км* до $r = 9.8$ *км* .

При этом в отсутствие нерегулярности в виде берегового клина максимальная дальность обнаружения для тех же самых параметров среды возрастает до 11 км. Можно сделать вывод, что нерегулярность батиметрии в данном случе уменьшает дальности обнаружения на 14 %.



а



в

Рис. 4. Нормированные интерферограммы (слева), голограммы (по центру) и функции обнаружения (справа). (а) $r = 9.2$ *км*, (б) $r = 9.8$ *км*.

Заключение

В работе задействован алгоритм оценки максимальной дальности обнаружения звукового источника. Рассмотрена ситуация регулярного волновода и нерегулярной батиметрии в виде берегового клина, что приводит к горизонтальной рефракции мод. Проведён численнй эксперимент для малогабартиного подводного аппарата в полосе 300-350 Гц с радиальной скоростью 2 $м/с$. Сделаны расчёты вероятности верного обнаружения для заданного входного отношения сигнал/шум на максимальной дальности обнаружения.

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

Список литературы

1. Чупров С.Д. Интерференционная структура звукового поля в слоистом океане // Акустика океана. Современное состояние. М.: Наука, 1982. С. 71–82.

2. Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Спектрограмма и локализация источника звука в мелком море // Акуст. журн. 2017. Т. 63. № 4. С. 406–418.

3. Казначеев И.В., Кузнецов Г.Н., Кузькин В.М., Пересёлков С.А. Интерферометрический метод обнаружения движущегося источника звука векторно-скалярным приемником // Акуст. журн. 2018. Т. 64. № 1. С. 33–45.

4. Pereselkov S.A., Kuz'kin V.M. Interferometric processing of hydroacoustic signals for the purpose of source localization // J. Acoust. Soc. Am. 2022. Vol. 151. No. 2 P. 666−676.

5. Kuz'kin V.M., Kuznetsov G.N., Pereselkov S.A., Grigor'ev V.A., Resolving power of interferometric method of source localization. // Phys. Wave Phenom., 2018 Vol. 26 No. 2 P.150-159.

6. Глущенко М.Ю., Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Хворостов Ю.А., Ткаченко С.А. Голографический метод локализации шумового подводного источника в мелком море // Акуст. ж. 2024. Т. 70. № 3. С. 67–77.

7. Kuz'kin V.M., Pereselkov S.A., Zvyagin V.G., Malykhin A.Yu., Prosovetskiy D.Yu., Intense internal waves and their manifestation in interference patterns of received signals on oceanic shelf // Phys Wave Phenom., 2018 Vol. 26 No. 2 P.160-167.

8. Badiey M., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Prosovetskiy D.Yu., Tkachenko S.A., Intense internal waves and their manifestation in the interference patterns of received signals on oceanic shelf. // Phys Wave Phenom., 2019 Vol. 27 No. 4 P.313–319.

9. Kaznacheeva E.S., Kuz'kin V.M., Pereselkov S.A., Interferometric processing of hydroacoustic information in the presence of intense internal waves. // Phys Wave Phenom., 2021 Vol. 29 No. 3 P.278–284.

10. Кузькин В.М., Пересёлков С.А., Казначеева Е.С., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Голографическая обработка движущихся источников в мелком море при наличии интенсивных внутренних волн. // РЭНСИТ, 2022, Т. 14. № 2. С. 197–204.

11. Кузькин В.М., Пересёлков С.А., Грачев В.И., Ткаченко С.А., Ладыкин Н.В., Куцов М.В. Голографический метод локализации движущегося источника звука в присутствии интенсивных внутренних волн. // РЭНСИТ, 2023, Т. 15. № 3. С.317–326.

12. Kuz'kin V.M., Pereselkov C.A., Badiey M., Ladykin N.V., Malykhin A.yu., Tkachenko S.A. Robustness of holographic processing of hydroacoustic signals in the presence of intense internal waves. // Phys. Wave Phenom., 2023, Vol. 31 No. 5 P. 346–354.

13. Кузькин В.М., Переселков С.А., Грачев В.И., Рыбянец П.В., Ткаченко С.А. Разрешение шумовых сигналов при наличии случайных неоднородностей океанической среды. // РЭНСИТ, 2024, Т. 16 № 3. С. 395–406.

14. Kuz'kin V.M., Pereselkov S.A., Rybyanets P.V., Tkachenko S.A. Resolution of hydroacoustic noise signals under conditions of spatial–temporal variability of ocean environment. // Phys. Wave Phenom., 2024, Vol. 32 No. 4. P. 290–299.

15. Pereselkov S.A., Kuz'kin V.M. Interferometric processing of hydroacoustic signals for the purpose of source localization. // J. Acoust Soc Am., 2022, Vol. 151 No. 2. P. 666–676.

16. Ehrhardt M., Pereselkov S., Kuz'kin V., Kaznacheev I., Rybyaners P. Experimental observation and theoretical analysis of the low-frequency source interferogram and hologram in shallow water // J. Sound and Vibration. 2023. Vol. 544. P. 117388.

17. Pereselkov S., Kuz'kin V., Ehrhardt M., Tkachenko S., Rybyanets P., Ladykin N. Three-dimensional modeling of sound field holograms of a moving source in the presence of internal waves causing horizontal refraction // J. Mar. Sci. Eng. 2023. Vol. 11. No. 10, 1922.

18. Pereselkov S., Kuz'kin V., Ehrhardt M., Matvienko Yu., Tkachenko S., Rybyanets P. The formation of 2D holograms of a noise source and bearing estimation by a vector scalar receiver in the high-frequency band // J. Mar. Sci. Eng. 2024. Vol. 12. P. 704.

19. Kaznacheeva E.S., Kuznetsov G.N., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A. Measurement capability of the interferometric method of sound source localization in the absence of data on the waveguide transfer function. // Phys. Wave Phenom., 2019, Vol. 27 No. 1. P. 73–78.

20. Kaznacheeva E.S., Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Tkachenko S.A. Adaptive algorithms for interferometric processing. // Phys. Wave Phenom., 2020, Vol. 28 No. 3. P.267–273.

21. Кузькин В.М., Пересёлков С.А., Грачев В.И., Ткаченко С.А., Рыбянец П.В. Адаптивный алгоритм локализации шумовых подводных источников в мелководных акваториях // РЭНСИТ. 2024. Т. 16. № 4. С. 533–544.

22. Kuz'kin V.M., Pereselkov S.A., Rybyanets P.V., Tkachenko S.F. Adaptive method for holographic processing of broadband hydroacoustic signals // Phys. Wave Phenom. 2024. Vol. 32. № 5. P. 382–391.

23. Казначеева Е.С., Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Хворостов Ю.А. Оценка дальности обнаружения малогабаритного подводного аппарата по его шумовому полю. // Подводные исследования и робототехника, 2021, Т. 4 № 38 С. 88–93.

24. Kuz'kin V.M., Pereselkov S.A., Tkachenko S.A., Matvienko Yu.V., Khvorostov Yu.A. Range of detection of underwater sound sourse. // Phys. Wave Phenom., 2023, Vol. 31 No. 5 P. 339–345.

25. Кузькин В.М., Пересёлков С.А., Грачев В.И., Рыбянец П.В., Ткаченко С.А. Предельная дальность обнаружения подводного шумового источника с применением голографической обработки. // РЭНСИТ, 2024, Т.16 № 1. С. 129–136.

26. Кузькин В.М., Матвиенко Ю.В., Пересёлков С.А., Рыбянец П.В. О возможности оценки предельной дальности обнаружения подводных глайдеров. // Подводные исследования и робототехника, 2024 Т. 2 № 48 С. 68-75.

27. Кузькин В.М., Пересёлков С.А., Матвиенко Ю.В., Ткаченко С.А. Обнаружение источника в диспергирующих средах. // РЭНСИТ, 2019, Т.11 № 3 С. 337–344.

28. Kuz'kin V.M., Lyakhov G.A., Pereselkov S.A., Matvienko Yu.V., Tkachenko S.A. Noise-source detection in an oceanic waveguide using interferometric processing. // Phys. Wave Phenom., 2020, Vol. 28 № 1 P. 68-74.

29. Хворостов Ю.А., Матвиенко Ю.В., Кузькин В.М., Пересёлков СА, Ткаченко СА. Спектральные характеристики шумового поля малогабаритного автономного необитаемого подводного аппарата в дальней зоне. // Подводные исследования и робототехника, 2022, Т. 4 № 42 С. 84–88.

Алгоритм для разделения волн акустического каротажа

С. А. Переселков, email: pereselkov@yandex.ru [1]
В. М. Кузькин, email: kumiov@yandex.ru [2]
А. С. Переселков, email: pereselkov.edu@yandex.ru [1]

[1] Воронежский государственный университет
[2] Институт общей физики им. А.М. Прохорова РАН

Аннотация. *Предложен голографический метод фильтрации корпусной волны при акустическом каротаже во время бурения. Метод основан на применении двумерного преобразования Фурье к фазово-корреляционной диаграмме. Эффективность метода продемонстрирована при обработке акустических сигналов трехэлементного прибора «Горизонт-90 ВАК», предназначенного для измерения параметров упругих волн, генерируемых в породах электроакустическим излучателем. Представлены фазово-корреляционные диаграммы, голограммы и спектрограммы принятых сигналов до и после фильтрации корпусной волны. Метод позволил идентифицировать продольные и поперечные волны и оценить скорости их распространения в породе.*

Ключевые слова: *акустический каротаж в процессе бурения, диагностические волны, корпусная волна, фазово-корреляционная диаграмма, двумерное преобразование Фурье*

Введение

Акустический каротаж — это метод геофизических исследований скважин, предназначенный для изучения акустических свойств горных пород. Он основан на измерении скорости распространения и затухания звуковых волн, что позволяет оценивать пористость, состав и характер насыщения пород [1-3]. В приборах акустического каротажа используются излучатели и приёмники, установленные в цилиндрическом корпусе. В скважине формируется гидродинамическая волна, которая взаимодействует со стенками скважины и породой. В результате возникают продольные (P), поперечные (S), поверхностные (Stoneley) и корпусная (C) волны. Диагностическими являются P-, S- и Stoneley-волны, тогда как C-волна создаёт помехи. Перспективным направлением является акустический каротаж в процессе бурения (ALWD), позволяющий получать данные в реальном времени. Однако основная проблема метода — сильное влияние корпусной волны,

© Переселков С. А., Кузькин В. М., Переселков А.С., 2026

затрудняющее выделение полезного сигнала. Для подавления корпусной волны применяются конструктивные и алгоритмические методы [4,5]. Конструктивные методы изменяют конструкцию бурильной колонны, но ограничены требованиями к её прочности. Алгоритмические методы основаны на обработке сигналов и позволяют отделять полезные волны от помех [6-11].

В данной работе предложен метод обработки сигналов для малоканальных систем, основанный на двумерном преобразовании Фурье фазово-корреляционной диаграммы. Это позволяет разделять диагностические и корпусные волны. Метод был протестирован на трёхэлементном приборе «Горизонт-90 ВАК» и показал возможность определения скоростей продольных и поперечных волн.

1. Обработка фазокорреляционной диаграммы

ФКД — это интерференционная «картина» давления в $s(z,t)$ координатах «глубина–время» (z,t), формируемая при последовательной регистрации откликов на разных глубинах. Отдельные волны приходят на выбранные горизонты с разными задержками, зависящими от скоростей и расстояний между приёмниками, поэтому ФКД можно рассматривать как линейную суперпозицию вкладов разных волн.

Идея метода заключается в применении к ФКД двумерного преобразования Фурье (далее «голограмма» по аналогии с подходами в подводной акустике [12-23]):

$$F(k,\nu)=\int_{z_1}^{z_2}\int_{t_1}^{t_2} s(z,t)\exp(-i(zk+\nu t))dtdz \qquad (1)$$

где k и ν — «волновое число» и частота голограммы; (z_1,z_2) и (t_1,t_2) задают окно формирования.

Ключевое наблюдение: из-за различий дисперсионных свойств компоненты волнового пакета локализуются в разных областях плоскости (k,ν). В рассматриваемой задаче энергию C- и поверхностных волн (при близких условиях распространения по глубине) ожидаемо видеть около оси частот ν, тогда как P/S-волны дают характерные «вертикальные» структуры (полосы) при изменении условий с глубиной. Это позволяет в голограмме «вырезать» области, соответствующие C и St, и восстановить очищенную ФКД обратным преобразованием Фурье [24].

Дополнительное преимущество — помехоустойчивость за счёт квазикогерентного накопления по глубине. Оценка выигрыша по отношению «сигнал/помеха» может быть записана в виде [14]:

$$q \approx q_0 J\gamma\ , \qquad J = \frac{\Delta z}{\delta z}\ , \qquad \gamma = \frac{D_n}{D_p}\ , \tag{2}$$

где q_0 — входное отношение «сигнал/помеха», δz — шаг дискретизации по глубине, D_p и D_n — площади областей голограммы, занятых соответственно спектральной плотностью полезных компонент и интерференционного фона.

2. Экспериментальные результаты обработки данных ALWD

Демонстрация проведена на данных испытательной скважины, полученных серийным трёхэлементным прибором Горизонт-90 ВАК (один излучатель S и два приёмника давления Q_1, Q_2). Расстояние от излучателя до Q_1: $l_1 = 3.2$ м; межприёмный интервал $\Delta l = 0.5$ м. Спектр возбуждения сосредоточен в диапазоне 1–60 кГц, шаг по глубине $\delta z = 0.2$ м.

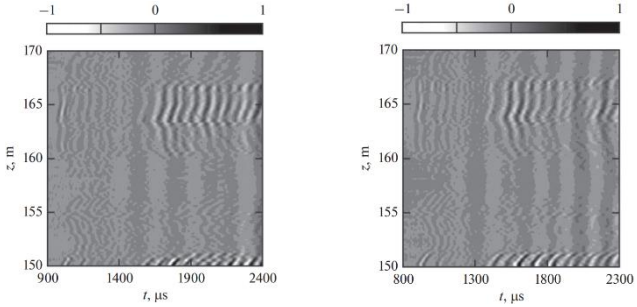


Рис. 1. Фрагменты нормированной фазо-корреляционной диаграммы: (a) Q1 и (b) Q2

Для иллюстрации выбраны фрагменты ФКД на глубинах 151–160 м и в окнах времени 800–1380 мкс (для Q_1) и 900–1560 мкс (для Q_2), где диагностические сигналы плохо различимы на фоне корпусной волны (Fig. 1). После построения голограмм (Fig. 2) выполнено подавление вкладов St и C в спектральной области, что наглядно проявляется в голограммах после фильтрации (Fig. 3) и приводит к восстановлению упорядоченного волнового пакета P / S.

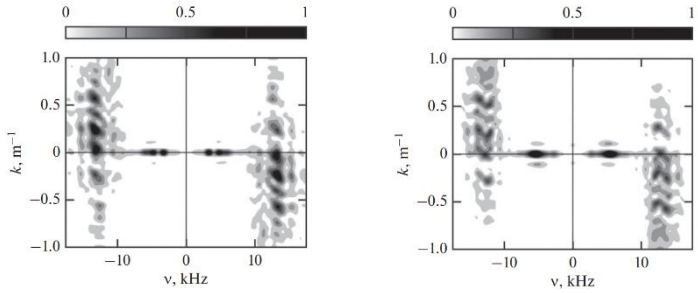


Рис. 2. Нормированные голограммы фрагментов фазо-корреляционной диаграммы: (a) Q1 и (b) Q2

Скорость выбранной волны оценивалась по межприёмной задержке:

$$u = \frac{\Delta l}{\Delta t}, \qquad \Delta t = t_2 - t_1, \tag{3}$$

где t_1, t_2 — моменты прихода волны на Q_1 и Q_2. Для горизонта $z \approx 157$ – 157.5 м получены времена прихода $t_{P1} = 897.96$ мкс, $t_{P2} = 987.50$ мкс и $t_{S1} = 1518.37$ мкс, $t_{S2} = 1716.67$ мкс, что даёт оценки скоростей $u_P \approx 5.58$ км/с и $u_S \approx 2.52$ км/с.

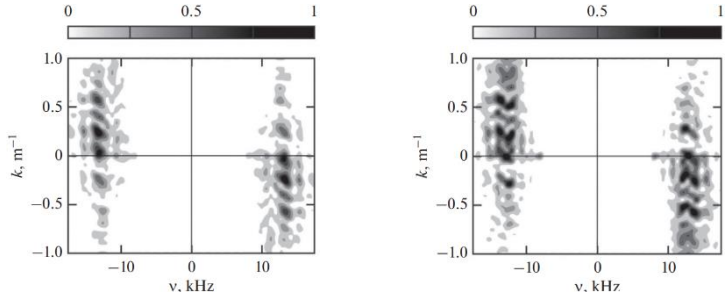


Рис. 3. Нормированные голограммы фрагментов фазо-корреляционной диаграммы с отфильтрованными волнами: (a) Q1 и (b) Q2

После фильтрации основная энергия P / S -волн концентрируется в диапазоне примерно 1–15 кГц; спектрограммы до/после фильтрации в исходной статье получены оконным преобразованием Фурье [24].

Заключение

Предложена и экспериментально показана голографическая схема подавления корпусной волны C в двухприёмной системе ALWD: (i) формирование ФКД, (ii) переход в область (k, ν) с помощью двумерного преобразования Фурье, (iii) спектральная фильтрация областей, соответствующих C и поверхностным волнам, (iv) восстановление очищенной ФКД обратным преобразованием. Метод обеспечивает выделение P- и S-волн и позволяет оценивать их скорости распространения в породе по межприёмной задержке, что важно для дальнейшей интерпретации упругих свойств.

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

-ратуры

1. Close D., Cho D., et al. The sound of sonic: A historical perspective and introduction to acoustic logging // CSEG Rec. 2009. Vol. 34. № 5. P. 34-43.
2. Peterson E. W. Acoustic wave propagation along fluid-filled cylinder // J. Appl. Phys. 1974. Vol. 45. № 8. P. 3340-3350.
3. Keller J. D. Acoustic wave propagation in composite fluid-saturated media // Geophysics. 1989. Vol. 54. № 12. P. 1554-1563.
4. Wang H., Tao G., Fehler M. C. Investigation of the high-frequency wavefield of an off-center monopole acoustic logging-while-drilling tool // Geophysics. 2015. Vol. 80. № 4. P. D329-D341.
5. Ji Y. Monopole collar wave characteristics for acoustic logging while drilling // Wave Motion. 2019. Vol. 90. P. 66-81.
6. Ахметсафин Р. Д., Ахметсафина Р. З. Две вычислительные реализации скоростной фильтрации волновых форм акустического каротажа // Geophysics. 2016. № 3. С. 78-84.
7. Lu J. Q., Ju X. D., et al. Azimuthally acoustic logging tool to evaluate cementing quality // J. Geophys. Eng. 2014. Vol. 11. № 4.
8. Xu S., Zhuang C., et al. Acoustic source characterization for a logging while drilling tool // J. Acoust. Soc. Am. 2018. Vol. 144. № 3.
9. Lin N., Wang J., et al. Technical properties and application examples of the newest acoustic logging while drilling tools // Pet. Drill. Tech. 2006. Vol. 31. P. 73-76.
10. Zhang K., Tan B., et al. Design of a new acoustic logging while drilling tool // Sensors. 2021. Vol. 21. № 13.

11. Ахметсафин Р. Д. О возможности фильтрации корпусной волны акустического каротажа, выполняемого во время бурения // Каротажник. 2024. № 327 (1). С. 137-142.

12. Pereselkov S. A., Kuz'kin V. M. Interferometric processing of hydroacoustic signals for the purpose of source localization // J. Acoust. Soc. Am. 2022. Vol. 151. № 2. C. 666-676.

13. Ehrhardt M., Pereselkov S., et al. Experimental observation and theoretical analysis of the low-frequency source interferogram and hologram in shallow water // J. Sound Vib. 2023. Vol. 544.

14. Kuz'kin V. M., Pereselkov S. A. Holographic methods of underwater acoustic signals processing: A review // Acoust. Phys. 2025. Vol. 71. № 1. C. 80-100. Kuznetsov G.N. et al. Spectrogram and localization of a sound source in a shallow sea. // Acoust. Phys. – 2017. – 63(4). – P.449-461.

15. Kuz'kin V.M. et al. Detection of a source in dispersive media. // RENSIT. – 2019. – 11(3). – P.337-344.

16. Kuz'kin V.M. et al. Noise source detection in an oceanic waveguide using interferometric processing. // Phys. Wave Phen. – 2020. – 28(1). – P.68-74.

17. Kuz'kin V.M. et al. Resolving power of the interferometric method of source localization. // Ph. W. Ph. – 2018. – 26(2). – P.150-159.

18. Kuz'kin V.M. et al. Intense internal waves and their manifestation in interference patterns of received signals on oceanic shelf. // Ph. W. Ph. – 2018. – 26(2). – P.160-167.

19. Badiey M. et al. Intense internal waves and their manifestation in the interference patterns of received signals on oceanic shelf. Part II. // Ph. W. Ph. – 2019. – 27(4) – P.313-319.

20. Kuz'kin V.M. et al. Interferometric processing of hydroacoustic information in the presence of intense internal waves// Ph. W. Ph. – 2021. – 29(3). – P.278-284.

21. Badiey M. et al. Interferometry of hydrodynamics of oceanic shelf caused by intensive internal waves // Fund. Ap. Hydr. – 2020. – 1. – P.45-55.

22. Ehrhardt M. et al. Three-dimensional modeling of sound field holograms of a moving source in the presence of internal waves causing horizontal refraction // J. Mar. Sci. Eng. – 2023. –11. – P.1922.

23. Ehrhardt, M. et al. The Formation of 2D Holograms of a Noise Source and Bearing Estimation by a Vector Scalar Receiver in the High-Frequency Band // J. Mar. Sci. Eng. – 2024. – 12. – P.704.

24. Harris F. J. On the use of windows for harmonic analysis with the discrete Fourier transform // Proc. IEEE. 1978. Vol. 66. № 1. C. 51-83.

Компьютерный эксперимент по оценке глубины источника волн в волноводе

С. А. Переселков, email: pereselkov@yandex.ru [1]
В. М. Кузькин, email: kumiov@yandex.ru [2]
А. С. Переселков, email: pereselkov.edu@yandex.ru [1]
Л. Н. Попова, email: popova.edu@yandex.ru [1]

[1] Воронежский государственный университет
[2] Институт общей физики им. А.М. Прохорова РАН

Аннотация. *Представлен пассивный акустический метод дистанционного зондирования в мелководных условиях, позволяющий оценивать глубину источника звука с использованием одного приемника. Оценка глубины осуществляется путем сопоставления измеренных соотношений амплитуд мод с теоретическими значениями*

Ключевые слова: *пассивный метод, источник звука, оценка глубины, одиночный приемник, моды звукового поля.*

Введение

В работе представлен пассивный акустический метод дистанционного зондирования в мелководье ([1, 2]), позволяющий оценивать глубину источника звука с использованием одного приемника. Предполагается, что сигнал источника является импульсным, что позволяет реализовать модальный фильтр в частотно-временной области на основе модовой дисперсии. Оценка глубины осуществляется путем сопоставления измеренных соотношений амплитуд мод с теоретическими значениями.

1. Методика оценки глубины источника

Мелководный волновод глубиной $H(r,z)$ в координатах (r,z). Сигнал, излучаемый точечным источником S на частоте $\omega = 2\pi f$ с глубины z_s, принимается точечным приемником Q на глубине z_q. Источник S находится на горизонтальном расстоянии r от приемника Q. Звуковое поле источника S в точке приемника Q представлено как сумма распространяющихся мод:

© Переселков С. А., Кузькин В. М., Переселков А. С., Попова Л. Н. 2026

$$p(r, z_s, z_q, \omega) = \sum_{m=1}^{M} A_m(r, z_s, z_q, \omega) \exp\left(i h_m(\omega) r\right) \quad (1)$$

где A_m и h_m — амплитуда и волновое число моды $\phi_m(z)$ с номером m. Здесь M — число распространяющихся мод.

Предполагается, что сигнал источника является импульсным или имеет короткую длительность, что позволяет реализовать модальный фильтр в частотно-временной области на основе модовой дисперсии. Временная зависимость $\zeta(\tau)$ принятого сигнала в частотной полосе $|\omega - \omega_0| \le \Delta\omega / 2$. Для определения $\zeta(\tau)$ применяем преобразование Фурье (1):

$$\zeta(\tau) = \sum_{m=1}^{M} \zeta_m(\tau), \quad (2)$$

где $\zeta(\tau)$ — временная зависимость сигнала моды $\phi_m(z)$ с номером m:

$$\zeta_m(\tau) = \int_{\omega_0 - \Delta\omega/2}^{\omega_0 + \Delta\omega/2} A_m(r, z_s, z_q, \omega) \exp\left(i k_m(\omega) r - i\omega\tau\right) d\omega. \quad (3)$$

Считаем амплитуду A_m медленно изменяющейся функцией по сравнению с фазой, что позволяет вынести ее из интеграла при $\omega = \omega_0$. В результате получаем следующее выражение для временной зависимости моды $\zeta_m(\tau)$:

$$\zeta_m(\tau) = A_m(r, z_s, z_q, \omega_0) \Delta\omega \exp\left\{ i \left[h_m(\omega_0) r - \omega_0 \tau \right] \right\} \times \quad (4)$$

$$\times \frac{\sin\left[\Delta\omega(\tau - \tau_m)/2\right]}{\left[\Delta\omega(\tau - \tau_m)/2\right]} \quad (5)$$

где положение главного максимума сигнала моды $\zeta_m(\tau)$ имеет значение $\tau = \tau_m$. Здесь $\tau_m = \dfrac{r}{u_m(f_0)}$ — время распространения m-й моды, а $u_m = \dfrac{d(\omega)}{dh_m}$ — групповая скорость. Рассмотрим отношение амплитуд двух соседних мод $A_m(r, z_s, z_q, f)$ и $A_n(r, z_s, z_q, f)$:

$$\chi_{mn}^{(\mathrm{ex})}(z_s, z_q, f) = \frac{A_m(r, z_s, z_q, f)}{A_n(r, z_s, z_q, f)} \tag{6}$$

Здесь «(ex)» обозначает экспериментальный результат. Введем функцию $\Theta_{mn}(z)$ как разность между теоретическим $\chi_{mn}^{(th)}$ и экспериментальным $\chi_{mn}^{(\mathrm{ex})}$ значениями:

$$\Theta_{mn}(z) = \left|\chi_{mn}^{(\mathrm{th})}(z, z_q, f) - \chi_{mn}^{(\mathrm{ex})}(z_s, z_q, f)\right| \tag{7}$$

где

$$\chi_{mn}^{(\mathrm{th})}(z, z_q, f) = \frac{\phi_m^{(\mathrm{th})}(z, f)\phi_n^{(\mathrm{th})}(z_q, f)}{\phi_n^{(\mathrm{th})}(z, f)\phi_m^{(\mathrm{th})}(z_q, f)} \tag{8}$$

Предложенный метод основан на уравнениях (5)–(9). Оценка глубины источника — это значение $z = z'$, для которого выражение (6) обращается в нуль:

$$\Theta_{mn}(z') = 0 \tag{9}$$

2. Параметры численного моделирования

Параметры численного моделирования. Глубина водного слоя $H = 53$ м. Скорость звука в придонном слое $c = 1700$ м/с. Плотность придонного слоя $\rho = 1.8$ г/см^3. Параметр поглощения дна $\alpha = 0.01$. Глубина приемника $z_q = 52$ м. Глубина источника $z_s = 15$ м. Принятый сигнал рассматривался в частотной полосе $\Delta f = 120 - 160$ Гц.

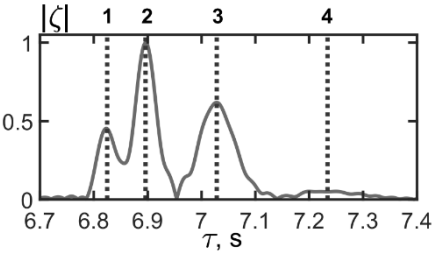


Рис. 1. Временная зависимость $|\zeta(\tau)|$

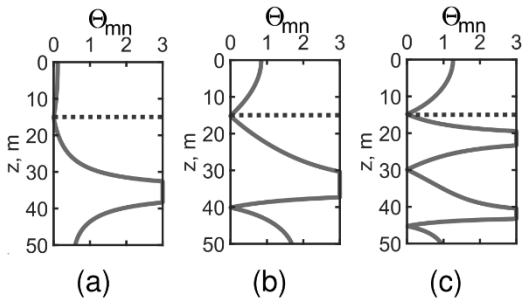


Рис. 2. Зависимость $\Theta_{mn}(z)$

Результаты моделирования представлены на Рис. 1 и Рис. 2. Нормированная временная зависимость $|\zeta(\tau)|$ принятого сигнала показана на Рис. 1. Принятый сигнал содержит четыре моды: $A_1 = 0,54$, $A_2 = 1,0$, $A_3 = 0,6$, $A_4 = 0,05$. Функция $\Theta_{mn}(z)$, рассчитанная для этих значений A_m, показана на Рис. 2 для различных пар мод: (a) $m = 1$, $n = 2$; (b) $m = 3$, $n = 2$; (c) $m = 4$, $n = 3$;. Как видно, из представленных результатов получаемая оценка глубины источника соответствует модельному значению: $z_s = 15$ м.

Заключение

Таким образом, представленные результаты численного моделирования демонстрируют высокую эффективность рассмотренного в работе метода оценки глубины источника звука с использованием одного приемника. Разработанный оценщик глубины имеет четкий и хорошо определенный экстремум. Результаты численного моделирования к истинному значению (ошибка менее 1 метра). В отличие от этого, ошибка оценки глубины источника в альтернативных подходах составляет десятки метров ([1,2]). Предложенный подход адаптирован для компактной аппаратной реализации в системах подводного мониторинга с использованием АНПА с приемниками. В будущих исследованиях планируется комбинировать предложенный метод оценки глубины с голографической обработкой сигналов, разработанной в работах [3-25]. Физические и математические принципы формирования голограмм в океаническом волноводе описаны в [4-7]. В статье [5] метод обобщен для шумового источника. В [6,7] был проанализирован для низкочастотный диапазон в регулярных волноводах. В работах [8,9] исследована разрешающая способность нескольких источников. В

работах [10-17] проанализирована эффективность метода в присутствии внутренних волн. Адаптивные алгоритмы разработаны в статьях [18]. В работе [19,20] методы обобщены на случай антенных сигналов. В статьях [21,22] разработаны методы модовой селекции в мелководном волноводе. Применение метода в случае векторно-скалярного приёмника в высокочастотном диапазоне рассмотрено в работах [23-25].

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

Список литературы

1. Yang T. C. A method of range and depth estimation by modal decomposition // J. Acoust. Soc. Am. – 1987. – 82. – P.1736-1745.
2. Shang E. C. Source depth estimation in waveguides // J. Acoust. Soc. Am. – 1985. – 77. – P.1413-1418.
3. Kuznetsov G.N. et al. Spectrogram and localization of a sound source in a shallow sea. // Acoust. Phys. – 2017. – 63(4). – P.449-461.
4. Kuz'kin V.M. et al. Detection of a source in dispersive media. // RENSIT. – 2019. – 11(3). – P.337-344.
5. Kuz'kin V.M. et al. Noise source detection in an oceanic waveguide using interferometric processing. // Phys. Wave Phen. – 2020. – 28(1). – P.68-74.
6. Pereselkov S.A. et al. Interferometric processing of hydroacoustic signals for the purpose of source localization. // J. Acoust. Soc. Am. – 2022. – 151. – P.666-676.
7. Ehrhardt M. et al. Experimental observation and theoretical analysis of the low-frequency source interferogram and hologram in shallow water. // J. Sound Vibr. – 2023. – 544. – P. 117388.
8. Kuz'kin V.M. et al. Resolving power of the interferometric method of source localization. // Ph. W. Ph. – 2018. – 26(2). – P.150-159.
9. Grachev V.I. et al. Resolution of noise signals in the presence of random inhomogeneities of the oceanic medium. // RENSIT. – 2024. – 16(3). – P.395-406.
10. Kuz'kin V.M. et al. Intense internal waves and their manifestation in interference patterns of received signals on oceanic shelf. // Ph. W. Ph. – 2018. – 26(2). – P.160-167.
11. Badiey M. et al. Intense internal waves and their manifestation in the interference patterns of received signals on oceanic shelf. Part II. // Ph. W. Ph. – 2019. – 27(4) – P.313-319.

12. Kuz'kin V.M. et al. Interferometric processing of hydroacoustic information in the presence of intense internal waves// Ph. W. Ph. – 2021. – 29(3). – P.278-284.

13. Badiey M. et al. Robustness of holographic processing of hydroacoustic signals in the presence of intense internal waves//Ph. W. Ph. – 2023. – 31(5). – P.346-354.

14. Badiey M. et al. Interferometry of hydrodynamics of oceanic shelf caused by intensive internal waves // Fund. Ap. Hydr. – 2020. – 1. – P.45-55.

15. Kuz'kin V.M. et al. Holographic processing of moving sources in shallow sea in the presence of intensive internal waves // RENSIT. – 2022. – 14(2). – P.197-204.

16. Kuz'kin V.M. et al. Holographic method for localizing a moving sound source in the presence of intensive internal waves. // RENSIT. – 2023. – 15(3). – P.317-326.

17. Ehrhardt M. et al. Three-dimensional modeling of sound field holograms of a moving source in the presence of internal waves causing horizontal refraction // J. Mar. Sci. Eng. – 2023. –11. – P.1922.

18. Pereselkov S. et al. Adaptive Algorithms for Interferometric Processing // Ph. W. Ph. – 2020. –28. – P.267-273.

19. Kuz'kin V.M. et al. Interferometric processing of acoustic information using extended antennas in dispersive media // RENSIT. – 2020. – 12(4) . – P.483-494.

20. Matvienko Yu.V. et al. Detection of a noise signal in an oceanic waveguide using a horizontal antenna // RENSIT. – 2022. – 14(1). – P.65-72.

21. Matvienko Yu.V. et al. Mode selection in oceanic waveguides // Ph. W. Ph. – 2022. – 30(2). – P.111-118.

22. Matvienko Yu.V. et al. Holographic method for mode selection in a shallow sea in the presence of intense internal waves // Ph. W. Ph. – 2022. – 30(5). – P.314-320.

23. Kuznetsov G.N. et al. Interferometric method for detecting a moving sound source using a vector-scalar receiver // Acoust. Phys. – 2018. – 64(1) . – P.33-45.

24. Kuznetsov G.N. et al. Interferometric method for estimating the coordinates of a moving noise source in shallow sea using high-frequency signals // Acoust. Phys. – 2020. – 66(4). – P.437-445.

25. Ehrhardt, M. et al. The Formation of 2D Holograms of a Noise Source and Bearing Estimation by a Vector Scalar Receiver in the High-Frequency Band // J. Mar. Sci. Eng. – 2024. – 12. – P.704.

Предобработка 2D-снимков для повышения качества анализа нейросетями

И. Б. Старченко, email: istarchenko@donstu.ru [1]
В. С. Белоусова, email: viktoria.belousova.work@yandex.ru [1]
Д. А. Кравчук, email: kravchukda@sfedu.ru [2]

[1] Донской государственный технический университет
[2] Южный федеральный университет

Аннотация. *В работе проведён анализ методов предварительной обработки 2D-снимков, имеющих ключевое значение для задач компьютерного зрения и анализа изображений. Рассмотрены существующие библиотеки, предназначенные для предобработки визуальных данных. В качестве практического применения представлен пример фильтрации изображений с использованием операции свертки с помощью библиотеки OpenCV на языке Python. Полученные результаты позволяют говорить о значительном улучшении качества входных данных.*

Ключевые слова: *свёртка, компьютерное зрение, 2D снимки, методы предварительной обработки визуальных данных, фильтрация изображений, OpenCV, Python.*

Введение

В современном мире технологии искусственного интеллекта (ИИ) развиваются стремительными темпами. Одним из ключевых направлений в этой области являются свёрточные нейронные сети (СНС), которые используются для анализа изображений [1]. Для обучения СНС необходим набор данных, а успешность обучения и дальнейшая классификация с помощью нейросети во многом определяется этапом предварительной обработки данных [2].

Предварительная обработка данных — это ряд действий, которые выполняются над данными до обучения модели [3]. Среди таких действий:

1. выравнивание;
2. нормализация;
3. смещение;
4. повороты;
5. вырезки и другие операции.

© Старченко И. Б., Белоусова В. С., Кравчук Д. А. 2026

Основной целью предварительной обработки 2D-изображений является снижение уровня шумов, это достигается благодаря применению различных фильтров и помогает уменьшить случайные помехи, сохраняя при этом важные детали изображения, что позволяет сократить число ошибочных выводов и улучшить выявление характеристик.

Рассмотрим методы для предварительной обработки данных [4, 5]:

1. Нормализация – предполагает собой преобразование значений пикселей в заданный диапазон для улучшения скорости и стабильности процесса обучения.
2. Устранение шумов – подразумевает сглаживание или фильтрацию изображений для снижения случайных колебаний в значениях пикселей.
3. Изменение размеров – это масштабирование изображения до необходимых размеров или разбиение исходного изображения на фрагменты заданного размера для последующего обучения.
4. Аугментация – это искусственное увеличение объёма данных с помощью различных преобразований (поворотов, отражений, изменения масштаба и добавления шума).

Проанализируем существующие библиотеки для реализации предобработки, представленные в таблице 1.

Таблица 1

Библиотеки для предобработки [6]

Библиотека	**Описание**	**Применение**
NumPy	Предназначена для работы с математическими вычислениями, методами линейной алгебры и статистики. Работает с многомерными массивами и матрицами	Используется для представления изображений в виде массивов, выполнения арифметических, логических и других операций над массивами
Pandas	Используется для анализа и предварительной обработки данных	Помогает считывать данные из разных источников (например, CSV-файлов), фильтровать их, приводить к единому виду

Продолжение таблицы 1

OpenCV	имеет доступ к различным алгоритмам и моделям для компьютерного зрения (SIFT, SURF, ORB, HOG, Haar, LBP, DNN, CNN, RNN).	Применяется для чтения и записи изображений, применения различных фильтров, обнаружения объектов, обработки изображений
Scikit-Image	Надстройка над NumPy служит для обработки изображений: фильтрации, сегментации, геометрических преобразований и др.	Используется для предобработки изображений: коррекции яркости и контраста, удаления шума, выделения границ и других преобразований
Pillow	Предназначена для манипулирования изображениями: открытия, сохранения в разных форматах, преобразования, применения фильтров и эффектов.	Помогает в обработке изображений: изменении размера, поворота, обрезки и т. д.
Mahotas	Используется для компьютерного зрения и обработки изображений, основанная на NumPy.	Применяется для быстрой обработки изображений, поддерживает различные форматы
Detectron2	Предназначена для обнаружения объектов и семантической сегментации на изображениях с использованием глубокого обучения.	Используется для более сложных задач предобработки, таких как сегментация изображений и обнаружение объектов
Librosa	Служит для работы с аудио, но также может быть полезна для извлечения признаков из изображений, представленных в виде сигналов.	Применяется для извлечения низкоуровневых признаков из изображений, преобразованных в сигнальный формат

Окончание таблицы 1

TensorFlow	Используется для построения и обучения нейронных сетей, включая свёрточные нейронные сети.	Включает инструменты для аугментации данных, нормализации, изменения размера изображений и других шагов предобработки
PyTorch	Фреймворк необходим для глубокого обучения, поддерживающий тензорные вычисления и построение нейронных сетей.	Содержит модули для преобразования и аугментации изображений, нормализации данных и других этапов предобработки

Таким образом выбор конкретной библиотеки зависит от целей и задач проекта, уровня подготовки данных, используемого фреймворка для обучения моделей и специфики работы с данными.

Целью данной статьи является рассмотрение на примере основных методов предварительной обработки 2D-изображений для обучения СНС, анализ существующих библиотек для предобработки данных.

1. Методы

Для реализации предобработки изображений при работе с нейросетями в данной работе используются:

– OpenCV — библиотека для компьютерного зрения, которая предоставляет функции для обработки изображений, включая фильтрацию и преобразования.

– scikit-image — библиотека, предоставляющая алгоритмы обработки изображений и компьютерного зрения, включая фильтрацию и преобразования.

– Keras — высокоуровневая библиотека для глубокого обучения, предоставляет инструменты для предобработки данных.

Далее, чтобы оценить качество предобработки можно использовать следующие метрики:

– Для классификации — точность, полнота, F1-Score и ROC-AUC, позволяющие анализировать способность модели различать классы.

– Для задач сегментации — метрики IoU (Intersection over Union) и Dice Coefficient, которые оценивают степень совпадения предсказанных и истинных масок.

2. Результаты

Рассмотрим фильтрацию изображений с использованием свертки OpenCV в Python. Произведем фильтрацию изображения с помощью 2D-свертки с функцией cv2.filter2D(). Свертка происходит между исходным изображением и ядром. Ядро – это еще один массив, который обычно меньше исходного изображения и определяет действие фильтрации. Ядро может быть высокочастотным, низкочастотным или настраиваемым, которое может обнаруживать определенные функции в изображении.

Фильтр низких частот представляет собой процесс усреднения, но с весами и размахом усреднения в зависимости от формы и содержимого ядра. Фильтр высоких частот похож на детектор границ. Он дает высокий уровень при значительном изменении значений соседних пикселей. Кроме того, можно использовать настраиваемый фильтр для обнаружения кругов, квадратов или некоторых нестандартных форм, которые вы хотите обнаружить на изображении.

Приведем пример с двумерной сеткой, где фильтр нижних частот представляет собой массив 5 × 5 со всеми единицами и усреднением. Затем применим свертку между исходным изображением и ядром с помощью функции cv2.filter2D().

```
import numpy as np
import cv2 #read image
img_src = cv2.imread('./images_2/barkhany2.jpg') #prepare the
5x5 shaped filter
kernel = np.array([[1, 1, 1, 1, 1],
                   [1, 1, 1, 1, 1],
                   [1, 1, 1, 1, 1],
                   [1, 1, 1, 1, 1],
                   [1, 1, 1, 1, 1]])

kernel = kernel/sum(kernel) #filter the source image

img_rst = cv2.filter2D(img_src,-1,kernel) #save result image

cv2.imwrite('./images_2/result2.jpg',img_rst)

#img_src = cv2.imread('./images_2/barkhany.jpg')

# Display the images using Matplotlib
plt.figure(figsize=(22, 12))

plt.subplot(1, 4, 1)
img_src = cv2.cvtColor(img_src, cv2.COLOR_BGR2RGB)

plt.imshow(img_src)
plt.title('RGB Image')
plt.axis('off')
```

```
plt.subplot(1, 4, 2)
img_rst = cv2.cvtColor(img_rst, cv2.COLOR_BGR2RGB)

plt.imshow(img_rst)
plt.title('Result Image')
plt.axis('off')
```

Результат показан на рис. 1.

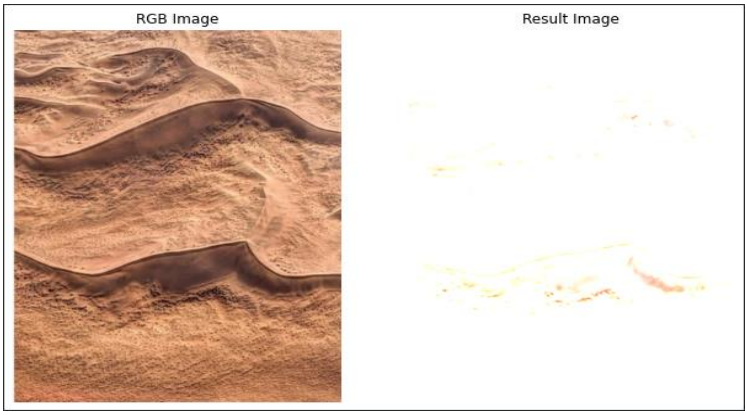


Рис. 1. Применение фильтра нижних частот (ядро 5 × 5 с усреднением)

В следующем примере фильтр верхних частот представляет собой массив 3 × 3, который является переменной ядра в приведенной ниже программе. Для реализации необходимо применить свертку между исходным изображением и ядром с помощью функции cv2.filter2D()

```
import numpy as np
import cv2 #read image
img_src = cv2.imread('./images_2/barkhany2.jpg') #edge
detection filter
kernel = np.array([[0.0, -1.0, 0.0],
                   [-1.0, 4.0, -1.0],
                   [0.0, -1.0, 0.0]])
kernel = kernel/(np.sum(kernel) if np.sum(kernel)!=0 else 1)

#filter the source image
img_rst = cv2.filter2D(img_src,-1,kernel) #save result image

cv2.imwrite('result3.jpg',img_rst)
plt.figure(figsize=(22, 12))

plt.subplot(1, 4, 1)
img_src = cv2.cvtColor(img_src, cv2.COLOR_BGR2RGB)
```

```
plt.imshow(img_src)
plt.title('RGB Image')
plt.axis('off')

plt.subplot(1, 4, 2)
img_rst = cv2.cvtColor(img_rst, cv2.COLOR_BGR2RGB)
plt.imshow(img_rst)
plt.title('Result Image')
plt.axis('off')
```

Результат показан на рис. 2.

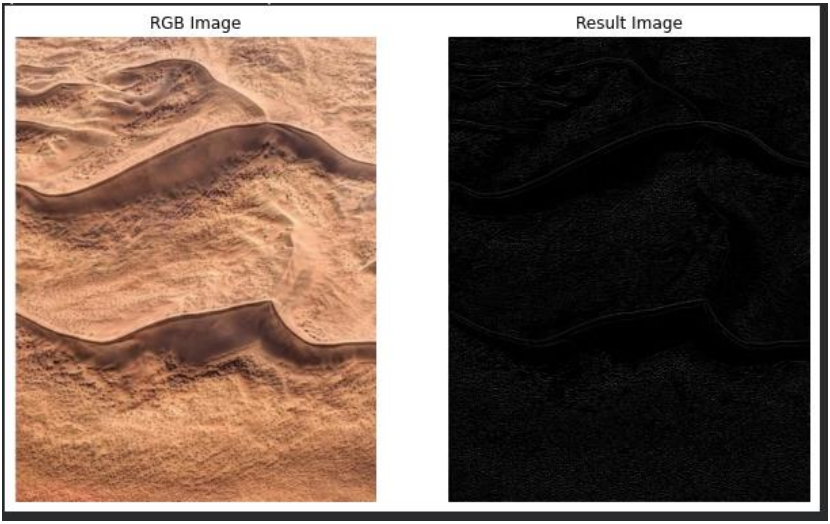


Рис. 2. Применение фильтра нижних частот (ядро 3 × 3 с усреднением)

Выходное изображение выглядит так, как будто вся зернистая информация сохранена, а остальная часть исчезла. Далее если произвести изменение массива ядра на следующий, то информация о цвете будет сохранена с выделенными высокочастотными пиксельными областями.

```
img_src = cv2.imread('./images_2/barkhany2.jpg') #edge
detection filter

kernel = np.array([[0.0, -1.0, 0.0],
                   [-1.0, 5.0, -1.0],
                   [0.0, -1.0, 0.0]])
kernel = kernel/(np.sum(kernel) if np.sum(kernel)!=0 else 1)

#filter the source image
```

```
img_rst = cv2.filter2D(img_src,-1,kernel) #save result image

cv2.imwrite('result4.jpg',img_rst)
plt.figure(figsize=(22, 12))

plt.subplot(1, 4, 1)
img_src = cv2.cvtColor(img_src, cv2.COLOR_BGR2RGB)
plt.imshow(img_src)
plt.title('RGB Image')
plt.axis('off')

plt.subplot(1, 4, 2)
img_rst = cv2.cvtColor(img_rst, cv2.COLOR_BGR2RGB)
plt.imshow(img_rst)
plt.title('Result Image')
plt.axis('off')
```

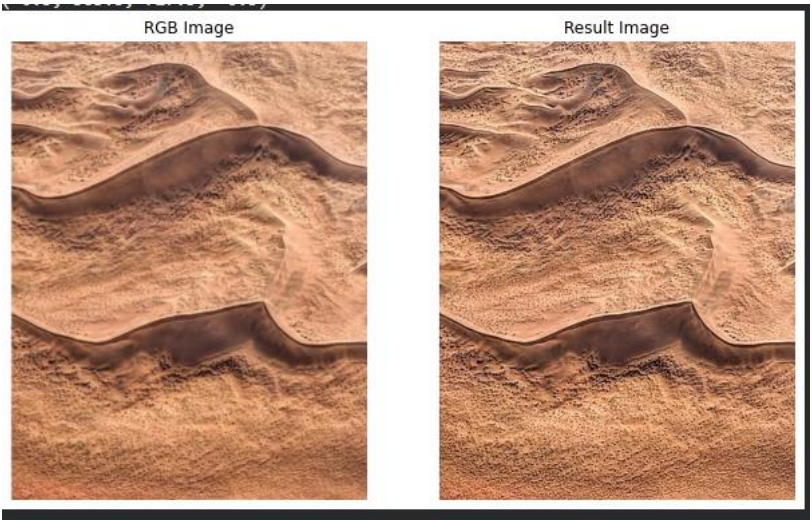


Рис. 3. Влияние конфигурации ядра свёртки высокочастотных пиксельных областей с сохранением цветовой информации

В следующем примере определим собственное ядро, в котором оно будет определять только горизонтальные линии, если они есть.

```
Import numpy as np
import cv2 #read image
img_src = cv2.imread('./images_2/barkhany2.jpg') #kernal
sensitive to horizontal lines
kernel = np.array([[-1.0, -1.0],
                   [2.0, 2.0],
                   [-1.0, -1.0]])
```

```
kernel = kernel/(np.sum(kernel) if np.sum(kernel)!=0 else 1)
#filter the source image
img_rst = cv2.filter2D(img_src,-1,kernel) #save result image
cv2.imwrite('result5.jpg',img_rst)

plt.figure(figsize=(22, 12))

plt.subplot(1, 4, 1)
img_src = cv2.cvtColor(img_src, cv2.COLOR_BGR2RGB)
plt.imshow(img_src)
plt.title('RGB Image')
plt.axis('off')

plt.subplot(1, 4, 2)
img_rst = cv2.cvtColor(img_rst, cv2.COLOR_BGR2RGB)
plt.imshow(img_rst)
plt.title('Result Image')
plt.axis('off')
```

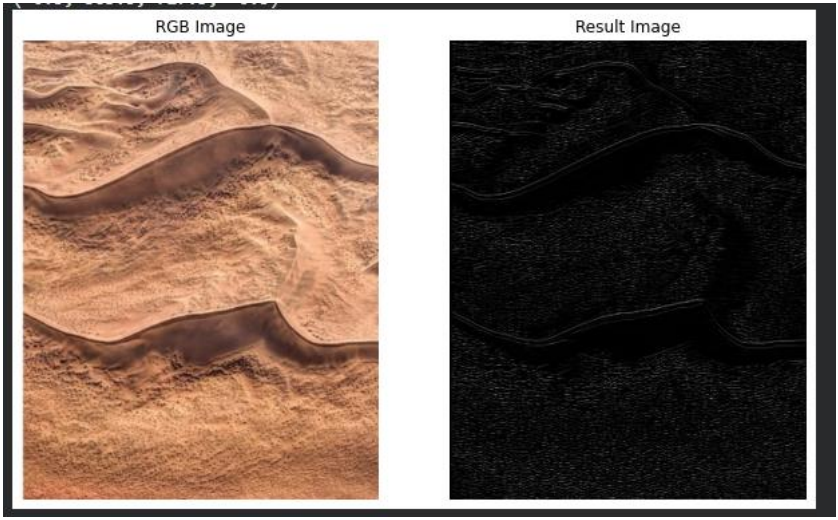


Рис. 4. Детектирование горизонтальных линий на изображении с применением фильтра свёртки

Заключение

Предварительная обработка данных играет ключевую роль в успешном обучении свёрточных нейронных сетей. Использование различных методов предобработки и библиотек позволяет значительно улучшить качество входных данных и, как следствие, точность и устойчивость разрабатываемых моделей ИИ.

OpenCV и Scikit-Image предлагают широкий набор алгоритмов для обработки изображений, Mahotas подойдёт для быстрой обработки и манипуляций с изображениями, а TensorFlow и PyTorch предоставляют полный пайплайн для обучения моделей, включая инструменты для предобработки данных.

Таким образом, предварительная обработка данных и правильный выбор библиотек для её реализации – неотъемлемые этапы в разработке эффективных моделей машинного обучения для анализа изображений.

Благодарности

Исследование поддержано грантом Российского научного фонда № 23-61-10024, https://rscf.ru/project/23-61-10024/.

Список литературы

1. Ерохин Д. Ю., Ершов М. Д. Современные сверточные нейронные сети для обнаружения и распознавания объектов //Цифровая обработка сигналов. – 2018. – Т. 3. – С. 64-69.
2. Starchenko I. B. et al. Image Reconstruction Algorithm of Optoacoustic Signal Transformation Based on Neural Networks //International Conference on Physics and Mechanics of New Materials and Their Applications. – Cham : Springer Nature Switzerland, 2023. – С. 571-576.
3. Dehbozorgi P., Ryabchykov O., Bocklitz T. A systematic investigation of image pre-processing on image classification //IEEE Access. – 2024. – Т. 12. – С. 64913-64926.
4. Абрамов Н. С. и др. Современные методы интеллектуальной обработки данных ДЗЗ //Программные системы: теория и приложения. – 2018. – Т. 9. – №. 4 (39). – С. 417-442.
5. Шапошникова Н. В., Ганжа Я. С. Исследование влияния предварительной обработки данных и параметров настройки нейронной сети на эффективность ее работы //Решетневские чтения. – 2018. – Т. 2. – С. 177-179.
6. Python 3.11 URL: https://www.python.org/ (дата обращения: 18.01.2026).

Школа-конференция «Информатика в образовании»

Секция 1
Методика преподавания информатики

Разработка алгоритма решения олимпиадной задачи по информатике «Достройте дороги» на основе понятия эйлерова графа

Ю. В. Кулаков, email: kulak@list.ru

Тамбовский государственный технический университет

Аннотация. *Разработан алгоритм решения олимпиадной задачи по информатике «Достройте дороги», основанный на понятии эйлеровых и полуэйлеровых графов.*

Ключевые слова: *олимпиадная задача «Достройте дороги», контроль состояния дорог, минимальное количество дорог, достраиваемые дороги, маршрут министра транспорта, граф сети дорог, связный граф, компоненты связности графа, эйлеров цикл, эйлерова цепь.*

Введение

Известная олимпиадная задача по информатике и программированию «Достройте дороги» была предложена организаторами олимпиад в качестве одной из задач III Всероссийской заочной олимпиады по информатике [1].

1. Постановка олимпиадной задачи

В Тридесятом государстве есть N городов, все города пронумерованы числами от 1 до N. Между городами построены дороги – каждая дорога соединяет между собой два города.

Царь Тридесятого государства уволил министра транспорта за то, что дороги были в очень плохом состоянии. Новый министр транспорта, чтобы не повторить судьбу своего предшественника, решил, что он будет лично контролировать состояние дорог. А именно он решил, что раз в год он лично будет объезжать все дороги.

При этом министр транспорта очень ценит свое время, и считает, что если в процессе объезда ему придется дважды проехать по какой-то дороге, то когда он будет ехать по этой дороге второй раз, ему уже не придется проверять ее состояние, и это будет недопустимой тратой времени.

Министр транспорта живет в городе номер 1, и поэтому хочет, чтобы его путешествие начиналось в этом городе. Заканчиваться путешествие должно в городе номер K, где каждый год будет проходить

© Кулаков Ю. В., 2026

всегосударственное совещание по вопросам планирования ремонта дорог на следующий год.

Определите, какое минимальное количество дорог нужно построить в Тридесятом царстве в дополнение к уже существующим дорогам, чтобы можно было выполнить все требования министра транспорта, а именно, чтобы он мог, выехав из города номер 1, проехать по каждой дороге ровно один раз и в итоге приехать в город номер K.

Формат входных данных. Вводится число N – количество городов в Тридесятом царстве ($1 \le N \le 100$). Далее вводится число M – количество дорог в Тридесятом царстве ($1 \le M \le 10000$). Далее идет M пар чисел – каждая пара задает номера городов, соединяемых соответствующей дорогой. Все дороги двухсторонние, т. е. по дороге можно ездить в любую сторону. Между некоторыми городами может быть несколько дорог. Возможны дороги из города в него же. В последней строке входных данных находится число K – номер города, где заканчивается маршрут министра ($1 \le K \le N$).

Формат выходных данных. Выведите минимальное количество дорог, которое необходимо построить в Тридесятом царстве. Затем выведите, какие именно дороги надо построить.

Примеры входных и выходных данных, а также комментарии к ним приведены в таблице.

Таблица

Входные данные, выходные данные и комментарии

Входные данные	**Выходные данные**	**Комментарии**
4 2 2 3 3 4 1	2 1 2 1 4	4 города и 2 дороги: из 2 в 3 и из 3 в 4. Т. к. министр транспорта живет в городе 1, и совещание также будет в городе 1, необходимо достроить 2 дороги — из 1 в 2 и из 1 в 4, тогда министр сможет из 1-го города поехать во 2-й, оттуда — в 3-й, оттуда — в 4-й и вернуться в 1-й
6 5 1 2 2 3 3 4 4 2 6 6 1	2 2 6 6 1	Маршрут министра может быть, например, таким: 1, 2, 3, 4, 2, 6, 6. В 5-м городе вообще нет дорог, поэтому туда министру заезжать не нужно. А вот дорогу, ведущую из 6-го города в него же, он проверить обязательно должен

Окончание таблицы

2 4 1 2 1 2 1 1 1 2 2	0	Маршрут министра может быть таким: 1, 1, 2, 1, 2. Новых дорог достраивать не нужно

2. Первый пример решения олимпиадной задачи из условия олимпиадной задачи

Пусть N = 4 – в Тридесятом царстве четыре города. Пусть число M = 2 – в Тридесятом царстве две дороги. Пусть две пары чисел, задающих дороги, будут {2, 3} и {3, 4} – первая (двусторонняя) дорога соединяет города 2 и 3, а вторая (двусторонняя) дорога соединяет города 3 и 4. Пусть число K = 1 – в городе номер 1 (выделен красным цветом) заканчивается маршрут движения министра транспорта.

Граф сети дорог Тридесятого царства для первого примера представлен на рис. 1.

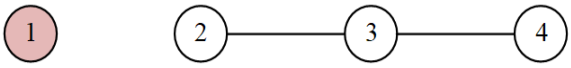


Рис. 1. Начальная сеть дорог Тридесятого царства для первого примера

В данном случае необходимо (и достаточно) достроить две дороги – одну из города номер 1 в город номер 2 и другую из города номер 1 в город номер 4.

Граф сети дорог Тридесятого царства с достроенными дорогами (представленными штриховыми линиями) для первого примера представлен на рис. 2.

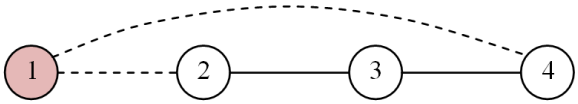


Рис. 2. Конечная сеть дорог Тридесятого царства для первого примера

Тогда министр транспорта для проверки состояния дорог царства сможет из родного города номер 1 поехать по достроенной дороге {1, 2} в город номер 2, из него поехать по контролируемой дороге {2, 3} в город номер 3 и оттуда по контролируемой дороге {3, 4} в город номер 4. Затем

и из него вернуться по достроенной дороге {4, 1} в город номер 1, где должно состояться всегосударственное совещание по вопросам планирования ремонта дорог на следующий год.

Заметим, что в первом примере граф начальной сети дорог (см. рис. 1) был несвязным и состоял из двух компонент связности. При этом одна (первая) из компонент связности представляла собой изолированную вершину (номер 1), а другая (вторая) – простую цепь (с концевыми вершинами 2 и 4). Причем изолированная вершина (номер 1) является и родным городом министра транспорта, из которого он начал свое путешествие, и конечным городом этого путешествия.

В данном случае для выполнения всех требований министра транспорта нам потребовалось в дополнение к уже существующим дорогам построить (минимально) две дороги. При этом первая построенная дорога связала изолированную вершину графа (первую компоненту связности) с одной из концевых вершин второй компоненты связности, а вторая построенная дорога – ту же самую изначально изолированную вершину графа с другой концевой вершиной второй компоненты связности.

Отметим, что в результате построения двух дополнительных дорог граф конечной сети дорог (см. рис. 2) представляет собой эйлеров граф, который содержит, например, эйлеров цикл ({1, 2}, {2, 3}, {3, 4}, {4, 1}). Этот эйлеров цикл определяет возможный маршрут ((1, 2), (2, 3), (3, 4), (4, 1)) движения министра транспорта для осуществления личного контроля состояния всех дорог Тридесятого государства. Этот маршрут начинается в родном для министра транспорта городе номер 1 и заканчивается в городе номер K = 1, где будет проходить всегосударственное совещание по вопросам планирования ремонта дорог на следующий год.

3. Второй пример решения олимпиадной задачи из условия олимпиадной задачи

Пусть N = 6 – в Тридесятом царстве шесть городов. Пусть число M = 5 – в Тридесятом царстве пять дорог. Пусть пять пар чисел, задающих дороги, будут {1, 2}, {2, 3}, {3, 4}, {4, 2} и {6, 6} – первая (двусторонняя) дорога соединяет города 1 и 2, а вторая (двусторонняя) дорога соединяет города 2 и 3, третья – города 3 и 4, четвертая – города 4 и 2, наконец, пятая – ведет из города номер 6 в него же. Пусть число K = 1 – в городе номер 1 (выделен красным цветом) заканчивается маршрут движения министра транспорта.

Граф сети дорог Тридесятого царства для второго примера представлен на рис. 3.

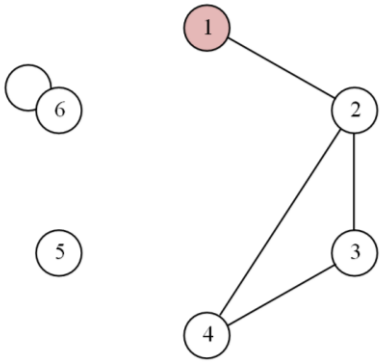


Рис. 3. Начальная сеть дорог Тридесятого царства для второго примера

В данном случае, как и в первом примере, необходимо (и достаточно) достроить две дороги, но одну из города номер 1 в город номер 6, а другую из города номер 6 в город номер 2.

Граф сети дорог Тридесятого царства с достроенными дорогами для второго примера представлен на рис. 4.

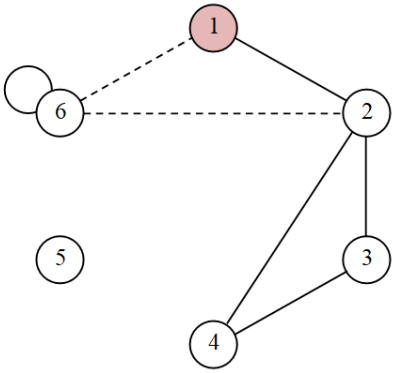


Рис. 4. Конечная сеть дорог Тридесятого царства для второго примера

Тогда министр транспорта для контроля состояния дорог царства сможет из родного города номер 1 (по контролируемым дорогам) поехать в город номер 2, из него поехать в город номер 3, оттуда в город номер 4 и из него вернуться в город номер 2. Затем по достроенной дороге {2, 6} поехать в город номер 6, проехать по контролируемой дороге из города

номер 6 в него же и, наконец, по достроенной дороге {6, 1} поехать в город номер 1, где должно состояться всегосударственное совещание по вопросам планирования ремонта дорог на следующий год. При этом маршрут министра транспорта не проходит через город номер 5, а по условию олимпиадной задачи и не должен проходить, поскольку вокруг города номер 5 вообще нет дорог.

Заметим, что во втором примере граф начальной сети дорог (см. рис. 3), также как и в первом примере, был несвязным, но состоял из трех компонент связности. При этом одна (первая) из компонент связности представляла собой подграф, носителем которого являлось множество вершин {1, 2, 3, 4}. Данное множество включает вершину (номер 1), которая обозначает и родной город министра транспорта, из которого он начал свое путешествие, и конечный город этого путешествия.

Другая (вторая) компонента связности графа начальной сети дорог представляла собой изолированную вершину без петель (номер 5). Причем данная изолированная вершина не является ни родным городом министра транспорта, из которого он не начал свое путешествие, ни конечным городом этого путешествия.

Третья компонента связности графа начальной сети дорог представляла собой изолированную вершину с петлей (номер 6). Данная изолированная вершина также как и вершина номер 5 не является ни родным городом министра транспорта, из которого он не начал свое путешествие, ни конечным городом этого путешествия. Но министр транспорта в соответствии с условием задачи обязан посетить ее, поскольку должен проверить состояние дороги, ведущей из города номер 6 в него же.

В данном случае, как и в первом примере, для выполнения всех требований министра транспорта нам потребовалось в дополнение к уже существующим дорогам построить (минимально) две дороги. При этом первая построенная дорога ({2, 6}) связала вершину номер 2 графа (из первой компоненты связности) с изолированной вершиной с петлей номер 6 (третьей компонентой связности), а вторая построенная дорога ({6, 1}) – изолированную вершину графа с петлей номер 6 с концевой вершиной номер 1 (из первой компоненты связности).

Отметим, что в результате построения двух дополнительных дорог граф конечной сети дорог (см. рис. 4) представляет собой эйлеров граф, который содержит эйлеров цикл, например, ({1, 2}, {2, 3}, {3, 4}, {4, 2}, {2, 6}, {6, 6}, {6, 1}). Этот эйлеров цикл определяет, например, возможный маршрут ((1, 2), (2, 3), (3, 4), (4, 2), (2, 6), (6, 6), (6, 1)) движения министра транспорта. Он начинается в родном для министра транспорта городе номер 1 и заканчивается в городе номер K = 1, где

будет проходить всегосударственное совещание по вопросам планирования ремонта дорог на следующий год.

4. Третий пример решения олимпиадной задачи из условия олимпиадной задачи

Пусть N = 2 – в Тридесятом царстве два города. Пусть число M = 4 – в царстве четыре дороги. Пусть четыре пары чисел, задающих дороги, будут {1, 2}, {1, 2}, {1, 1} и {1, 2} – первая дорога соединяет города 1 и 2, вторая дорога также соединяет города 1 и 2, третья – город 1 с самим собой, и, наконец, четвертая – города 1 и 2. Пусть число K = 2 – в городе 2 (выделен красным цветом) заканчивается маршрут движения министра транспорта. Город 1, в котором живет министр транспорта и в котором начинается его путешествие, выделен зеленым цветом.

Граф сети дорог Тридесятого царства для третьего примера представлен на рис. 5.

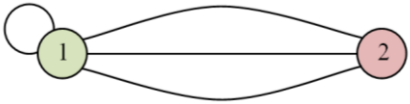


Рис. 5. Начальная сеть дорог Тридесятого царства для третьего примера

В данном случае, в отличие от первого и второго примеров, новых дорог достраивать не нужно, то есть необходимо достроить ноль дорог.

Действительно, министр транспорта для контроля состояния дорог царства сможет, например, из родного города номер 1 (по контролируемым дорогам) поехать в него же, затем поехать в город номер 2, оттуда вернуться в город номер 1 и из него снова поехать в город номер 2, где должно состояться всегосударственное совещание по вопросам планирования ремонта дорог на следующий год.

Заметим, что в третьем примере граф начальной сети дорог (см. рис. 5), в отличие от первого и второго примеров, был связным, то есть состоял из одной компоненты связности. При этом носителем графа является множество вершин {1, 2}. Данному множеству принадлежит и вершина (номер 1), которая обозначает родной город министра транспорта, из которого он начал свое путешествие, и конечный город (номер 2) этого путешествия.

Отметим, что и без построения дополнительных дорог граф сети дорог представляет собой полуэйлеров граф, который содержит эйлерову цепь, например, ({1, 1}, {1, 2}, {1, 2}, {1, 2}). Эта эйлерова цепь определяет, например, возможный маршрут ((1, 1), (1, 2), (2, 1), (1, 2))

движения министра транспорта для осуществления личного контроля состояния всех дорог Тридесятого государства. Он начинается в родном для министра транспорта городе номер 1 и заканчивается в городе номер K = 2, где будет проходить всегосударственное совещание по вопросам планирования ремонта дорог на следующий год.

5. Разработанный алгоритм решения олимпиадной задачи «Достройте дороги»

В результате проведенного анализа трех примеров входных данных, выходных данных и комментариев к ним разработан алгоритм решения олимпиадной задачи «Достройте дороги».

Представим его словесно по шагам.

Шаг 1. По заданным входным данным построить граф G начальной сети дорог Тридесятого царства с заданным числом вершин (городов) N и заданными ребрами (двусторонними дорогами).

Шаг 2. Проверить степень каждой из N вершин графа G на четность.

Шаг 3. Отметить на графе G начальную вершину маршрута министра транспорта (город номер 1) зеленым цветом и конечную вершину его маршрута (город номер K) красным цветом.

Шаг 4. Определить компоненты связности G1, G2, G3, … , Gn графа G.

Шаг 5. Если n = 1 (граф G связен), то перейти к следующему шагу; в противном случае перейти к шагу 10.

Шаг 6. Если номер конечной вершины маршрута министра транспорта K равен 1 и все вершины графа G имеют четную степень, то вывести значение ноль (граф G эйлеров, в нем существует эйлеров цикл и, следовательно, новых дорог достраивать не нужно) и КОНЕЦ алгоритма; в противном случае перейти к следующему шагу.

Шаг 7. Если номер конечной вершины маршрута министра транспорта K не равен 1 и только вершины 1 и K графа G имеют нечетную степень, то вывести значение ноль (граф G полуэйлеров, в нем существует эйлерова цепь с концевыми вершинами 1 и K, следовательно, новых дорог достраивать не нужно) и КОНЕЦ алгоритма. В противном случае перейти к следующему шагу.

Шаг 8. При K равном 1 в граф G добавлять ребра между парами вершин с нечетными степенями, выводить и подсчитывать их до тех пор, пока все вершины не станут иметь четную степень.

При K не равном 1 в граф G добавлять ребра между вершинами так, чтобы только вершина 1 и вершина K стали иметь нечетную степень. Выводить и подсчитывать их. Например, соединить ребром вершины 1 и K, если только они в графе G имеет четную степень.

Шаг 9. КОНЕЦ алгоритма.

Шаг 10. Если среди n компонент связности G1, G2, G3, … , Gn графа G существуют изолированные вершины без петель, отличные от начальной и конечной вершины, то удалить их из графа.

Шаг 11. Связать оставшиеся компоненты связности графа в одну компоненту, наводя мосты между ними, при этом выводить эти мосты и подсчитывать их. Наводить мосты следует преимущественно между вершинами компонент связности с нечетными степенями. Полученный связный граф назвать графом G.

Шаг 12. Вернуться к шагу 5.

6. Применение разработанного алгоритма для первого примера входных данных из условия олимпиадной задачи

Рассмотрим и проиллюстрируем протокол исполнения разработанного алгоритма для решения олимпиадной задачи «Достройте дороги» для первого примера входных данных.

Шаг 1. По заданным входным данным построим граф G начальной сети дорог Тридесятого царства с заданным числом вершин (городов) N = 4 и двумя заданными ребрами (двусторонними дорогами) {2, 3} и {3, 4} (см. рис. 6).

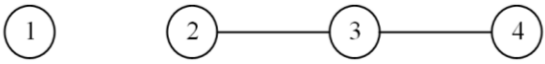


Рис. 6. Граф начальной сети дорог Тридесятого царства для первого примера

Шаг 2. В графе G вершины 1 и 3 имеют четную степень, а вершины 2 и 4 – нечетную степень.

Шаг 3. Отметим на графе G начальную вершину маршрута министра транспорта (город номер 1) зеленым цветом и конечную вершину его маршрута (город номер K = 1) красным цветом (см. рис. 1). Заметим, что при этом красный цвет закрыл зеленый цвет (начальная и конечная вершины маршрута совпадают).

Шаг 4. Определим компоненты связности G1 = {1} и G2 = {2, 3, 4} графа G. Число компонент связности графа n равно 2.

Шаг 5. Поскольку $n \neq 1$ (граф G несвязен), перейдем к шагу 10.

Шаг 10. Поскольку среди двух компонент связности G1 и G2 графа G отсутствуют изолированные вершины без петель, отличные от начальной и конечной вершины, то никакие вершины из графа не удаляем.

Шаг 11. Свяжем оставшиеся компоненты связности G1 и G2 графа в одну компоненту, наведя мост, например {1, 2}, между ними, поскольку в

первой компоненте связности отсутствуют вершины нечетной степени, а степень вершины 2 второй компоненты связности нечетная. Отправим этот мост на вывод, и он будет первой достроенной дорогой между городами Тридесятого царства (см. рис. 7).

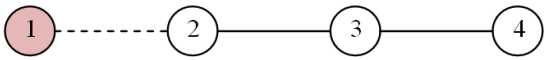


Рис. 7. Сеть дорог для первого примера с первой достроенной дорогой

Шаг 12. Вернемся к шагу 5.

Шаг 5. Поскольку граф G связен, то перейдем к следующему шагу.

Шаг 6. Номер конечной вершины маршрута министра транспорта K равен 1, но не все вершины графа G имеют четную степень, поэтому перейдем к следующему шагу.

Шаг 7. Номер конечной вершины маршрута министра транспорта K равен 1, поэтому перейдем к следующему шагу.

Шаг 8. Поскольку K равно 1 в граф G добавим ребро между единственной парой вершин с нечетными степенями 1 и 4. Отправим добавленное ребро {1, 4} на вывод. Это ребро будет обозначать вторую достроенную дорогу. Заметим, что теперь все вершины графа стали иметь четную степень (см. рис. 2).

Шаг 9. КОНЕЦ алгоритма.

Заключение

Разработан алгоритм решения олимпиадной задачи «Достройте дороги», который использует понятие эйлерова графа.

Решение данной олимпиадной задачи может быть использовано учащимися при подготовке к участию в олимпиадах по информатике и программированию, а также при изучении темы «Эйлеровы графы» в курсе информатики и при подготовке к сдаче Единого государственного экзамена по информатике.

Список литературы

1. Олимпиады по программированию. III Всероссийская заочная олимпиада школьников по информатике, 2008/09 учебный год [Электронный ресурс]: Задача F. Достройте дороги. – Режим доступа: https://olympiads.ru/zaoch/2008/problems/f.shtml

О логарифмической линейке замолвим мы слово, или зачем изучать историю вычислительной техники на уроках информатики

Л. Ю. Черницына, email: chernicynalyu@mgpu.ru

Институт развития профильного обучения, Государственное автономное образовательное учреждение высшего образования города Москвы «Московский городской педагогический университет»

Аннотация. *В статье рассматривается актуальность изучения истории вычислительной техники на примере логарифмической линейки.*

Ключевые слова: *информатика, история вычислительной техники, патриотическое воспитание, логарифмическая линейка, техника вычислений с помощью логарифмической линейки.*

Введение

Согласно федеральной рабочей программе по информатике для 7–9 классов, в 7-м классе в рамках раздела «Цифровая грамотность» предусмотрено изучение истории развития компьютеров, включая знакомство с поколениями компьютеров, и программного обеспечения [1, с. 6; 2, с. 6]. Однако в существующих учебниках этот элемент содержания либо практически отсутствует, либо раскрыт крайне поверхностно. Например, тема, с которой ранее традиционно начиналось изучение информатики – история вычислительной техники (от истоков счета до современных компьютеров), – сейчас часто остается без должного внимания.

1. Почему важно знакомить учащихся с историей вычислительных устройств

Без знакомства с историей развития вычислительной техники в СССР и уникальными достижениями отечественной школы создания ЭВМ у обучающихся может сложиться ошибочное представление об отсутствии в нашей стране значимых разработок в этой области. Для формирования объективной картины и ценностного отношения к отечественному культурному, историческому и научному наследию, владения достоверной информацией о передовых мировых и отечественных достижениях в области информатики и информационных технологий [1, с. 13; 2, с. 13], необходимы, в частности, знания о серии мощных вычислительных машин БЭСМ (Большая Электронная Счетная

© Черницына Л. Ю., 2026

Машина, 1960-1970 гг.), чьи характеристики в период их создания превосходили многие зарубежные аналоги, а также о первом в мире серийном компьютере на основе троичной логики – «Сетунь» (1958 – 1959 гг.).

Знакомство учащихся с историей отечественных технологий и инструментов, которые обеспечили прорывные достижения, является важным элементом патриотического воспитания, позволяя на конкретных примерах продемонстрировать мощь инженерной мысли, интеллект и упорство ученых в условиях технологических вызовов эпохи.

Принято считать, что эра электронных вычислительных машин началась в 1946 году с запуска в работу универсальной цифровой вычислительной машины (ЭВМ) ENIAC (Electronic Numerical Integrator and Computer) в США [3]. В Советском Союзе первая цифровая ЭВМ М-1 была создана под руководством И.С. Брука и введена в эксплуатацию в 1950-1951 годах [4]. Это закономерно вызывает вопрос: какими средствами обеспечивались сложнейшие расчеты в науке и технике до появления первых ЭВМ?

До широкого внедрения электронно-вычислительной техники (конец 1970-х – начало 1980-х годов) основными инструментами для вычислений служили логарифмические линейки и механические арифмометры. Эти устройства были незаменимыми помощниками инженеров, конструкторов и ученных. Именно с помощью таких инструментов, как логарифмическая линейка и арифмометр, велись расчеты, легшие в основу грандиозных проектов XX века.

Так, например, выдающийся советский конструктор Сергей Павлович Королёв и его команда использовали их при создании первого искусственного спутника Земли (1957 г.), пилотируемых космических кораблей и другой ракетно-космической техники в период с 1946 по 1966 годы. Это наглядное свидетельство того, что даже до эры суперкомпьютеров отечественные специалисты, опираясь на глубокие знания, изобретательность и «ручные», инструменты, могли решать задачи космического масштаба. Лишь с 1980-х годов их окончательно вытеснили портативные электронные калькуляторы.

2. Логарифмическая линейка

Логарифмическая линейка представляет собой аналоговое вычислительное устройство, принцип работы которого основан на представлении чисел в логарифмическом масштабе. Это позволило заменить операции умножения и деления на сложение и вычитание отрезков, осуществлять возведение в степень и извлечение корней.

Важной особенностью вычислений на логарифмической линейке была их приближенность. Точность вычислений с помощью логарифмической линейки зависела от длины шкалы: для стандартных линеек длиной 25 см она составляла 3–4 значащих цифры. Это требовало от инженеров, конструкторов и ученых понимания физического смысла задачи, оценки порядка величин и критического анализа результата.

На протяжении более 300 лет логарифмическая линейка оставалась практически единственным инструментом, с помощью которого выполнялись сложнейшие расчеты. Кто и когда изобрел это уникальное вычислительное устройство?

Первое вычислительное устройство, использующее логарифмическую шкалу, было создано английским математиком Эдмундом Гантером в 1622 году, вскоре после публикации первых логарифмических таблиц. Устройство Гантера представляло собой одну шкалу длиной 60 см, а операции сложения и вычитания отрезков выполнялись с помощью циркуля [6].

Значительный шаг в эволюции этого инструмента сделал Уильям Отред (Англия). В 1632 году он описал в трактате «Круги пропорций» конструкцию, состоящую из двух концентрических кругов со шкалами, которые можно было вращать друг относительно друга. Это ключевое усовершенствование позволило избавиться от циркуля и производить вычисления путем совмещения шкал [6].

В 1654 году Роберт Биссакар (Англия) предложил принципиально новую, прямоугольную конструкцию. Его линейка состояла из основного корпуса с вырезом, в котором скользил движок с нанесенными шкалами. Эта форма стала классической [6].

Современный вид логарифмической линейке придал французский офицер Амеде Мангейм в 1850 году, добавив прозрачный бегунок (курсор) с визирной линией. Этот элемент значительно повысил точность и удобство расчетов. Конструкция Мангейма считается канонической и оставалась практически неизменной до конца эпохи широкого использования линеек [6].

3. Устройство логарифмической линейки

Логарифмическая линейка состоит из трех основных компонентов (рисунок 1):

1. Корпус (G) – неподвижная часть, на которой нанесены основные шкалы.

2. Движок (Q) – центральная подвижная часть, вставляемая в корпус. Содержит дублирующие и дополнительные шкалы. Движок можно перемещать влево и вправо относительно корпуса.

3. Бегунок (В) – прозрачная рамка со стеклом, перемещающаяся вдоль линейки. На стекле нанесена визирная линия, которая служит для точного совмещения и считывания значений со шкал.

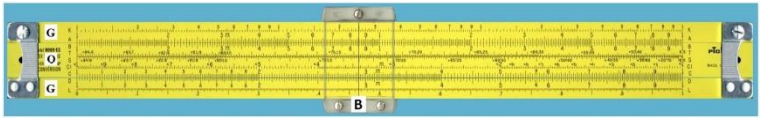


Рис. 1. Внешний вид логарифмической линейки

4. Основные вычислительные возможности

При наличии соответствующих шкал, с помощью логарифмической линейки можно выполнять следующие основные вычислительные операции:

– арифметические действия: умножение, деление;

– работа со степенями и корнями: возведение в степень (целую, дробную) и извлечение корня (квадратного, кубического и др.);

– работа с логарифмами: нахождение десятичного (lg) и натурального (ln) логарифмов числа, вычисление антилогарифмов (например, 10 в степени x);

– тригонометрические вычисления: определение синусов, косинусов, тангенсов и котангенсов углов (обычно в диапазоне от 0,5° до 90°);

– прочие операции: решение пропорций, вычисление процентов, манипуляция порядками величин (умножение/деление на степени десяти), вычисление экспоненциальных функций, гиперболических функций (на продвинутых инженерных моделях).

Это делало ее незаменимым инструментом для инженерных и физических расчетов, например, мощности, силы тока, сопротивления и др.

5. Образовательные ресурсы

В сети Интернет можно найти различные образовательные онлайн-ресурсы, с помощью которых можно изучить принцип работы логарифмической линейки. Например:

1. На сайте «Математические этюды» представлена наглядная анимация умножения на логарифмической линейке [7].

2. Интерактивный симулятор логарифмической линейки позволяет самостоятельно выполнять расчеты [8].

6. Практический пример

Разберем пример умножения чисел 260 и 0,3.

Алгоритм выполнения действий:

1. Представление чисел в форме мантиссы (с дробной частью). Так как максимальное значение на основной шкале обычно равно 10 (или 1,0), числа приводят к стандартному виду:

260 = 2,6 * 10^2 и 0,3 = 3,0 * 10^(-1).

На шкалах будем работать с мантиссами 2,6 и 3,0. Так как учащиеся 7 класса на момент знакомства с логарифмической линейкой обычно еще не знакомы с понятием «мантисса», то вместо этого понятия можно использовать фразу «десятичная дробь».

2. Установка первого множителя. На основной шкале корпуса (шкала D) находят значение 2,6 и совмещают с ним визирную линию бегунка (рисунок 2).

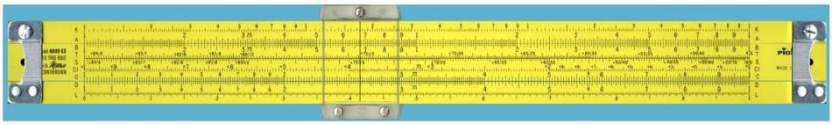


Рис. 2. Установка первого множителя

3. Позиционирование движка. Движок сдвигают так, чтобы начало его основной шкалы (шкала C, значение 1,0 или «индекс») совпало с визирной линией (рисунок 3).



Рис. 3. Позиционирование движка

4. Нахождение результата. На шкале C находят второй множитель 3. Напротив него на шкале D корпуса считывают результат произведения мантисс: ~7,8 (рисунок 4).

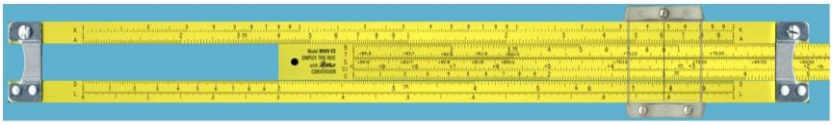


Рис. 4. Нахождение результата

Укрупненный вид нахождения результата показан на рисунке 5.

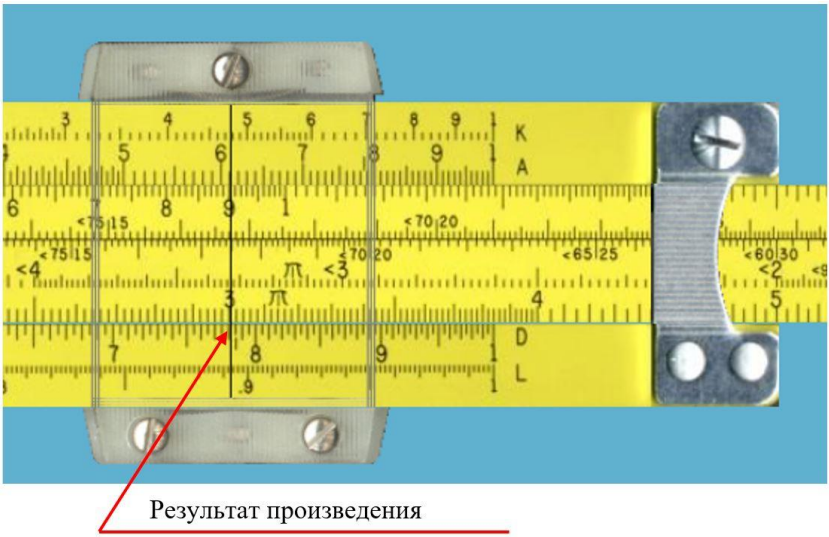


Рис. 5. Укрупненный вид результата

5. Определение порядка. Суммируем порядки:

$$10^2 * 10^{(-1)} = 10^1.$$

Таким образом, окончательный результат:

$$7{,}8 * 10^1 = 78.$$

Заключение

Изучение истории таких инструментов, как логарифмическая линейка и арифмометр, выходит за рамки простого исторического экскурса. Оно позволяет оценить эволюцию вычислительной техники, понять условия, в которых работали пионеры отечественной космонавтики и других высокотехнологичных отраслей, и осознать, что технологический прогресс определяется не только мощностью аппаратуры, но и интеллектом, образованностью и целеустремленностью людей. Интеграция этих знаний в образовательный процесс способствует формированию у подрастающего поколения чувства уважения к научно-техническому наследию страны и гордости за достижения прошлого.

Список литературы

1. Федеральная рабочая программа основного общего образования. Информатика (базовый уровень) (для 7-9 классов образовательных организаций). М.: Институт стратегии развития образования. 2023. 45 с.

2. Федеральная рабочая программа основного общего образования. Информатика (углубленный уровень) (для 7–9 классов образовательных организаций). М.: Институт стратегии развития образования. 2023. 51 с.

3. Истовый инженер. Калькулятор, который смог: краткая история первого программируемого компьютера [Электронный ресурс]. – Режим доступа : https://engineer.yadro.com/note/eniac/

4. М-1: Как зарождалась ЭВМ [Электронный ресурс]. – Режим доступа : https://www.cnews.ru/articles/2017-04-13_istoriya_otechestvennyh_it_avtomaticheskaya_tsifrovaya_vychislitelnaya

5. Златопольский Д. М. Логарифмическая линейка: предок калькулятора // Компьютерные инструменты в школе. 2019. № 6. С. 59-68.

6. Большая Российская Энциклопедия [Электронный ресурс]. – Режим доступа : https://bigenc.ru/c/logarifmicheskaia-lineika-8c82a1

7. Математические этюды. Этюды. Логарифмическая линейка [Электронный ресурс]. – Режим доступа : https://clck.ru/3RRzyN

8. Симулятор логарифмической линейки [Электронный ресурс]. – Режим доступа : https://www.antiquark.com/sliderule/sim/n909es/virtual-n909-es.html

Секция 2
ЕГЭ по информатике

Компетенции в области динамического программирования: аналитический и программный методы решения заданий 23 ЕГЭ по информатике

В. С. Попов, email: popov_vlad@mail.ru [1,2,3]

[1] Тверской государственный университет
[2] МГТУ им. Н.Э. Баумана
[3] ГАОУ ВО МГПУ

Аннотация. *В статье рассматриваются два подхода к решению задания 23 ЕГЭ по информатике на динамическое программирование: аналитический и с помощью рекурсивной функции на языке программирования Python. Представлен обзор характерных ошибок, возникающих у учащихся при написании программы для решения данного задания.*

Ключевые слова: *компетенции, компетенции в педагогике, информатика, динамическое программирование, ЕГЭ по информатике, задание 23 ЕГЭ по информатике.*

Введение

Задание 23 ЕГЭ по информатике не претерпевает существенных изменений уже 10 лет – с 2016 года [1, с. 86-87]. Тем не менее, обучающиеся испытывают существенные затруднения с выполнением данного задания. Так, процент выполнения задания 23 ЕГЭ по информатике долгие годы находится на уровне 50 % [1, с. 87-88; 2, с. 327-328], по результатам ЕГЭ по информатике 2024 года с данным заданием справились 51,1 % выпускников [3, с. 26-28], в 2025 году – 51,8 % [4, с. 24-26].

Цель статьи – рассмотреть аналитическое решение задания 23 ЕГЭ по информатике и решение на языке программирования, а также частые ошибки, допускаемые учениками при выполнении данного задания программным кодом. Рассмотрим следующее задание: «Исполнитель ИИ умеет выполнять следующие три операции: вычесть 1, вычесть 4, целочисленно разделить на 3. Сколько существует программ, преобразующих число 19 в число 2, если траектория вычислений включает число 14 и не включает число 8?».

© Попов В. С., 2026

1. Решение задания 23 ЕГЭ по информатике на языке программирования с использованием рекурсивной функции

Верное решение вышеприведённого задания на языке Python представлено на рис 1. Для вычисления количества программ в заданиях такого типа, как правило, в настоящее время экзаменуемыми используется рекурсивная функция [5, 6]. Она имеет краткую запись на языке программирования (рис. 1), быстро выполняется и, как следствие, не требует от обучающегося значительных усилий для правильного решения идентичных заданий.

```
def f(x, y):
    if x < y or x == 8:
        return 0
    if x == y:
        return 1
    return f(x - 1, y) + f(x - 4, y) + f(x // 3, y)
print(f(19, 14) * f(14, 2))
```

Рис. 1. Вариант верного решения задания 23 ЕГЭ по информатике в Python

В листинге на рис. 1 реализована рекурсивная функция f(x, y), вычисляющая количество программ для получения из числа x числа y. Если начальное число x меньше конечного числа y, невозможно получить число y из числа x, последовательно выполняя представленные в задании операции (вычесть 1, вычесть 4, целочисленно разделить на 3) – в этом случае функция возвращает 0, поскольку не существует программ для получения большего числа y из меньшего числа x. Также функция f(x, y) возвращает 0 в случае, если текущим числом является число 8, запрещённое в траектории вычислений. Для записи этой части базы рекурсии используется условный оператор: if x < y or x == 8: return 0.

Если числа x и y равны, существует ровно одна программа для получения числа y из числа x – пустая программа: if x == y: return 1.

В остальных случаях количество программ получения числа y из числа x определяется как сумма количества программ получения числа y из чисел (x − 1), (x − 4), (x // 3), где // – оператор целочисленного деления: return f(x - 1, y) + f(x - 4, y) + f(x // 3, y).

Поскольку по заданию траектория вычислений должна включать число 14, количество программ получения из числа 19 числа 2 через число 14 определяется как произведение соответствующих функций: f(19, 14) * f(14, 2). Результатом вычислений (ответом на задание) является число 69.

2. Аналитическое решение задания 23 ЕГЭ по информатике

С 2012 по 2020 годы задания ЕГЭ по информатике, включая данное, выполнялись аналитически, без компьютеров. Аналитическое решение задания 23 ЕГЭ по информатике заключается в постепенном подсчёте количества программ для получения чисел от заданного до целевого.

Напомним, что для указанных действий (вычесть 1, вычесть 4, целочисленно разделить на 3) следует вычислить количество программ получения из числа 19 числа 2, причём траектория вычислений включает число 14 и не включает число 8. Приведём аналитическое решение этого задания в виде последовательного вычисления значений функции для различных аргументов:

F(19) = 1 (существует одна пустая программа для получения числа 19 из числа 19)

F(18) = F(19) = 1

F(17) = F(18) = 1

F(16) = F(17) = 1

F(15) = F(16) + F(19) = 1 + 1 = 2

F(14) = F(15) + F(18) = 2 + 1 = 3

F(13) = F(14) = 3 (поскольку число 14 входит в траекторию вычислений, невозможно получить число 13 из числа 17 вычитанием 4 и при этом включить в траекторию вычислений число 14)

F(12) = F(13) = 3

F(11) = F(12) = 3

F(10) = F(11) + F(14) = 3 + 3 = 6

F(9) = F(10) + F(13) = 6 + 3 = 9

F(8) = 0 (по условию задания число 8 не включается в траекторию вычислений)

F(7) = F(11) = 3

F(6) = F(7) + F(10) = 3 + 6 = 9

F(5) = F(6) + F(9) = 9 + 9 = 18

F(4) = F(5) + F(12) + F(13) + F(14) = 18 + 3 + 3 + 3 = 27

F(3) = F(4) + F(7) + F(9) + F(10) + F(11) = 27 + 3 + 9 + 6 + 3 = 48

F(2) = F(3) + F(6) + F(6) + F(7) = 48 + 9 + 9 + 3 = 69

Аналитическое решение заданий 23 ЕГЭ по информатике на динамическое программирование освещено в статьях [5, с. 34, 36-37, 40-42; 7, с. 108-109; 8].

3. Примеры ошибок при решении задания 23 ЕГЭ по информатике программным кодом

При выполнении данного задания с использованием написанной программы ученик допустил сразу несколько ошибок (рис. 2 – настоящий

программный код, написанный учеником при решении представленного в статье задания).

```
def f(x, y):
    if x < y and x == 8:
        return 0
    if x == y:
        return 1
    return f(x - 1, y) * f(x - 4, y) * f(x % 3, y)
print(f(19, 14) + f(14, 2))
```

Рис. 2. Ошибочное решение задания на динамическое программирование учеником

Типичные ошибки при решении заданий данного типа с использованием программного кода включают: ошибки в логических операциях (рис. 2, строка 2), использование операции получения остатка от деления вместо целочисленного деления (рис. 2, строка 6), умножение результатов выполнения функций при их рекурсивном вызове вместо сложения (рис. 2, строка 6), сложение результатов вызовов функции вместо умножения (рис. 2, строка 7), использование неверных операторов сравнения, неверная вложенность фрагментов программного кода, запись неверных операций (например, вычитание 3 вместо вычитания 4), неверная запись включённых и исключённых из траектории вычислений чисел.

Заключение

В статье рассмотрены решения задания 23 ЕГЭ по информатике на языке программирования с использованием рекурсивной функции, аналитическим методом, приведены частые ошибки, допускаемые обучающимися при выполнении заданий данного типа. Изучение различных методов решения задания способствует формированию компетенций в области динамического программирования.

Список литературы

1. Попов В. С., Алефиренко Е. А., Черницына Л. Ю. Рекурсивная функция для решения задания на динамическое программирование: от беззнаниевой формы компетенции к знанию и пониманию // Информатика в школе. 2024. Т. 23, № 5. С. 85-94. EDN: MDGTEQ. DOI: 10.32517/2221-1993-2024-23-5-85-94.

2. Попов В. С., Абросимова-Романова Л. А. Анализ рисков выпускников и выделение необходимых компетенций в связи с изменением среднего процента выполнения заданий ЕГЭ по

информатике за 2019-2023 годы и введением компьютерного формата экзамена // Новые информационные технологии в образовании: Сборник научных трудов XXIV Международной научно-практической конференции, Москва, 30–31 января 2024 года. Москва: ООО «1С-Паблишинг», 2024. С. 326-330. EDN: RJIRBK.

3. Крылов С. С. Методические рекомендации для учителей, подготовленные на основе анализа типичных ошибок участников ЕГЭ 2024 года по информатике. М.: ФГБНУ «Федеральный институт педагогических измерений», 2024. 28 с.

4. Крылов С. С. Методические рекомендации для учителей, подготовленные на основе анализа типичных ошибок участников ЕГЭ 2025 года по информатике. М.: ФГБНУ «Федеральный институт педагогических измерений», 2025. 26 с.

5. Попов В. С., Алефиренко Е. А. Решение задания № 23 ЕГЭ по информатике на динамическое программирование: программный код на языке Python, использование обратных операций и проверка ответа // Потенциал. Журнал для старшеклассников и учителей. 2024. № 6. С. 33-44.

6. Попов В. С. Решение усложнённых заданий № 23 ЕГЭ по информатике на динамическое программирование, содержащих требования к командам // Потенциал. Журнал для старшеклассников и учителей. 2025. № 5. С. 40-47.

7. Ерофеев А. Н. Решение задач динамического программирования при подготовке к ЕГЭ по информатике // Актуальные проблемы модернизации математического и естественно-научного образования. Сборник научных трудов по материалам Всероссийской научно-методической конференции (Балашов, 15 мая 2020 года). Балашов: Саратовский источник, 2020. С. 107–110. EDN: RUYALN.

8. Муцурова З. М. Динамическое программирование в задачах ЕГЭ // Известия Чеченского государственного педагогического университета. Серия 2: Естественные и технические науки. 2019. Т. 17. № 1 (20). С. 44–47. EDN: VJYMNW.

Секция 3
Искусственный интеллект в образовании

Искусственный интеллект в образовательном процессе высшей школы

А. Е. Беляев, email: bae11@yandex.ru [1, 2]
М. М. Беляева, email: mmb12@ yandex.ru [1]

[1] Федеральное государственное бюджетное образовательное учреждение высшего образования «Рыбинский государственный авиационный технический университет имени П. А. Соловьева»
[2] Муниципальное образовательное учреждение средняя общеобразовательная школа «Образовательный комплекс «Взлёт» Центр образования № 23 имени С.И. Грудинского

Аннотация. *В данной работе рассмотрены отдельные аспекты, возможности и риски применения систем искусственного интеллекта (СИИ) в деятельности преподавателей и студентов высшей школы.*

Ключевые слова: *искусственный интеллект, образовательный процесс, большая языковая модель.*

Введение

В настоящее время системы искусственного интеллекта получают широкое распространение в различных сферах жизни человека и общества. Идея внедрения технологий искусственного интеллекта (ИИ) привлекает многих исследователей, руководителей, чиновников. Широкое распространение СИИ получают и в образовательном процессе. Актуальным является вопрос корректности и рамок использования таких систем в образовательном процессе.

1. Понятие системы искусственного интеллекта

Искусственный интеллект в широком смысле – это общее название технологий, которые позволяют вычислительным системам имитировать «умное» поведение человека: учиться, анализировать, принимать решения, генерировать информацию [1] Это широкое понятие, охватывающие разные классы СИИ. Далее под СИИ понимаются системы, основанные на машинном обучении. Но и этот термин охватывает большой класс моделей искусственного интеллекта.

Для корректности описания использования ИИ в образовательном процессе необходимо различать несколько классов СИИ.

К первому классу можно отнести специализированные СИИ – это системы, направленные на решение определенных задач, как правило, это

© Беляев А. Е., Беляева М. М., 2026

задачи классификации или распознавания объектов. Такие системы применяются в промышленности, дефектоскопии, медицине, охране правопорядка и других сферах жизни. Их применение в образовательном процессе не вызывает широкого резонанса, так как, как правило, они являются изучаемыми инструментами, с которыми предстоит работать специалистам. Их существование и применение подлежит обсуждению в специализированных областях, для которых и разработаны такие системы.

Ко второму классу можно отнести большие языковые модели (БЯМ). Это самый мощный и современный тип языковых моделей (часть группы генеративных систем ИИ). Они обучены на огромных массивах текстовых данных и способны «понимать» и генерировать текст. Такие системы получили широкое распространение не только на профессиональном, но и на бытовом уровне. Для начала использования таких систем не нужно (на первый взгляд) специального обучения, так как они реализованы в виде чат-ботов – условных собеседников по переписке. Интерфейс работы с ними прост и понятен. Однако именно эта кажущаяся простота может нести в себе определенные риски использования таких систем без понимания принципов их работы.

Необходимо понимать, что БЯМ обучены на открытых текстовых данных, в обучающий набор могли попасть как достоверные, так и недостоверные данные. Качество работы БЯМ с данными зависит от стратегии обучения и настройки. Это влечет за собой возможность появления в ответах БЯМ недостоверной информации. Иногда системы начинают «фантазировать», что, как правило, ограничивается разработчиками, чтобы не вводить пользователей в заблуждение.

БЯМ получили широкое распространение в образовательном процессе, особенно в высшей школе. Далее в данной работе будет более подробно рассмотрено применение именно таких СИИ в образовательном процессе.

2. Образовательный процесс в высшей школе как место применения систем искусственного интеллекта

Образовательный процесс в высшей школе многогранное явление, в котором существует множество мест приложения СИИ.

Особенностью образовательного процесса в высшей школе можно считать то, что его можно рассматривать как диалектическое взаимодействие преподавателя и студента. Можно рассмотреть образовательный процесс как ряд стадий и взаимодействий (рисунок).

Можно выделить несколько мест приложения возможностей систем искусственного интеллекта в образовательном процессе:

– подготовка преподавателя к занятию;

– выполнение студентом заданий;
– проверка преподавателем заданий;
– анализ преподавателем результатов работы;
– самоподготовка студентов.

Отдельно можно выделить такую область как организация образовательного процесса, она тоже может быть местом использования СИИ. В данном случае в эту область не попадают вопросы организации взаимоотношений преподавателей и студентов с образовательной организацией (контроль посещаемости/работы и другие), так как они относятся не к сфере непосредственно образовательного процесса, а к сфере организации работы коллектива (где могут применяться специализированные СИИ), что выходит за рамки данной работы.

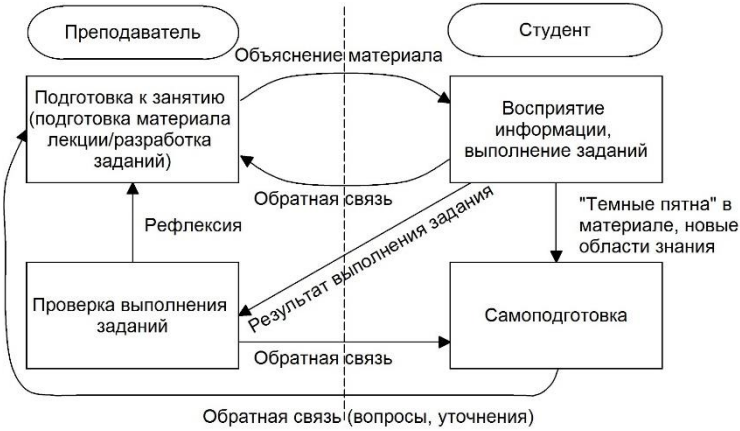


Рисунок. Примерная схема взаимодействия преподавателя со студентом

Таким образом, целесообразно рассмотреть применение СИИ преподавателем, студентом и административным сотрудником.

3. Искусственный интеллект в работе преподавателя

В работе педагога ИИ может быть применён на разных этапах.

3.1 Возможности и риски применения искусственного интеллекта при подготовке к занятиям.

Одним из основных видов внеаудиторной работы преподавателя является подготовка к занятиям. При этом подготовка к лекционным, практическим, лабораторным и семинарским занятиями отличается.

При подготовке к лекциям преподавателю необходимо найти информацию по теме занятия, систематизировать её, а также, что

особенно важно для качественного образования студентов и развития навыков мышления, вскрыть зависимости, тенденции, причины проблем по теме занятия. Именно на развитие профессионального, критического, нетривиального, системного мышления необходимо делать упор в современном образовании. Это связано с тем, что доступность информации высока, но её необходимо уметь искать, оценивать и использовать, поэтому, особенно важно учить студентов мыслить.

При подготовке к лекционным занятиям преподаватель может использовать БЯМ для поиска информации. Они позволяют быстро «просмотреть» источники информации, получить выжимку, а также найти интересные моменты. Во многом подспорьем преподавателю является поисковые системы с искусственным интеллектом (как подвид БЯМ, но анализирующие страницы в Интернете). Однако, как правило, сгенерированный моделями текст может содержать недостоверную информацию. Если запрос не популярен, система не имеет достаточного количества информации по теме, ответ всё равно будет сгенерирован, но из запроса будут взяты наиболее употребительные слова и составлен фактически новый запрос [2]. Информация, полученная от БЯМ, требует критического осмысления, проверки, верификации по альтернативным источникам, проверки ссылок, которые сформировала БЯМ, так как некоторые системы способны генерировать несуществующие ссылки.

При передаче её в «сыром» виде существует риск введения студентов в заблуждение, появление нестыковок, ошибок. Кроме того, на данный момент, БЯМ слабо способны выявлять тенденции, причины и следствия, неописанные в каком-либо источнике. БЯМ способны поддаваться убеждению со стороны пользователя и транслировать его.

Таким образом, наиболее корректным будет использование БЯМ для «умного» поиска информации с дальнейшей профессиональной переработкой полученных результатов преподавателем. Кроме того, СИИ могут использоваться для перевода иностранной литературы, однако, необходимо, чтобы данный перевод оценивался преподавателем со знанием предметной области и знанием иностранного языка. Это необходимо для того, чтобы избежать ошибок перевода, некорректной трактовки контекста и перевода терминов.

При подготовке к семинарским занятиям СИИ могут быть полезны для поиска и генерации интересных тем для обсуждения, дискуссии.

Аналогично можно использовать подобные системы для создания вариаций задания на лабораторную или практическую работу (с необходимостью дальнейшей проверки).

Необходимо с осторожностью использовать БЯМ при создании контрольных и/или лабораторных заданий «с нуля», так как системы не

знают лекционного материала конкретного преподавателя и могут сгенерировать задания, которые ему не соответствуют, что вызовет существенные затруднения у студентов.

Однако БЯМ достаточно неплохо показывают себя при необходимости составить задания на основе загруженного в неё материала. Система анализирует лекционный материал и способна генерировать задания по нему. Это может быть полезно для облегчения труда преподавателя, особенно при составлении тестовых заданий, где необходимо придумать несколько неправильных ответов, что может быть затруднительно при большом объёме информации.

Однако изменение условий проведения образовательного процесса требует от преподавателей отходить от тестовых заданий, так как они достаточно легко выполняются студентами с использованием БЯМ. Это требует при составлении контрольных заданий разрабатывать вопросы и задания, направленные на проверку не фактических знаний, а на умение анализировать информацию, оценивать её, выявлять закономерности.

3.2 Возможности и риски применения искусственного интеллекта при проверке выполнения заданий

Применение БЯМ при проверке студенческих работ может повысить эффективность труда преподавателя, но также несёт определенные риски [3].

БЯМ могут анализировать письменные работы студентов, проверяя грамматику, орфографию, стилистику и структуру изложения, они помогают быстро выявить типичные ошибки.

Использование СИИ при проверке работ может повысить объективность оценки за счёт исключения человеческого фактора и возможной предвзятости.

При использовании специализированной БЯМ, адаптированной под конкретные задачи, возможна настройка СИИ на генерацию индивидуальных рекомендаций обучающимся по улучшению качества знаний.

В то же время применение БЯМ при проверке работ вызывает некоторые риски. БЯМ при проверке работ опирается на формальные критерии, упуская важные моменты креативности и индивидуальный подход студента к решению задачи, что может негативно сказаться на мотивации к развитию творческого, исследовательского мышления студентов.

Помимо этого, для точной оценки необходимо наличие предварительно размеченного (оцененного преподавателем) большого набора исходных данных и тренировка системы на нём. Подготовка такого набора является дополнительной значительной нагрузкой на

преподавателя. Кроме того, осведомленные о способе проверки студенты могут преднамеренно затруднять автоматическую проверку работ, используя сложные обороты и фразы, трудные для распознавания БЯМ, с целью скомпрометировать средства оценивания конкретного преподавателя.

Следует обратить внимание на то, что такие системы могут хорошо обрабатывать ответы в электронном виде; в случае, если ответ был написан «от руки» возможны серьёзные проблемы с распознаванием почерка, особенно если проверяются «случайные» работы, а не используется специальная модель, тренированная распознавать почерки определенных людей.

Особым моментом является вопрос автоматизированной обработки персональных данных студентов даже в обезличенном виде.

Можно сделать вывод, что данные системы способны проверить тестовые задания, задания с кратким ответом, но плохо могут проверять задания с развернутым ответом, особенно, требующие оценки мыслительного процесса. Это связано со сравнением ответов, а не рассуждений. Кроме того, велик риск игнорирования индивидуальных психологических и социальных особенностей конкретных студентов.

Следовательно, корректным является применения преподавателем СИИ для проверки орфографии, пунктуации, а также других формальных признаков (в том числе кратких ответов) в работах. Можно применять БЯМ для формулирования комментариев студентам, с целью предоставления обратной связи.

3.3 Возможности и риски применения искусственного интеллекта при анализе результатов работ

Важным этапом в работе преподавателя является анализ результатов выполнения какой-либо работы. На основе наблюдения за студентами во время выполнения работы, качества выполнения работы в целом и конкретных заданий в частности преподаватель должен проводить рефлексию – при необходимости корректировать лекционный материал, подробнее освещать плохо усвоенные студентами темы или вопросы, а также при необходимости балансировать работы по сложности заданий. Кроме того, необходимо проводить со студентами краткие разборы работы с анализом распространенных ошибок.

СИИ могут помочь определить наиболее выполняемые и наименее выполняемые задания, если предоставить им результаты работы. Более того, в таком виде данные о выполнении работы могут быть использованы для выявления коррелирующих между собой заданий (заданий, выполняемость которых связана). Эти данные могу быть полезны преподавателю для подстройки дальнейшей работы.

Однако необходимо отметить, что для такого анализа результаты работы должны быть представлены в определенном виде, что может являться для преподавателя дополнительной работой.

4. Искусственный интеллект в деятельности студента

Как и в работе преподавателя, в деятельности студента искусственный интеллект может быть применен на разных этапах.

4.1 Возможности и риски применения искусственного интеллекта при выполнении заданий

Использование студентами БЯМ при выполнении заданий является самым распространенным на данный момент вариантом использования СИИ в образовательном процессе [4]. Актуальным является вопрос корректности использования СИИ при выполнении заданий.

Важно отметить, что существует несколько вариантов применения БЯМ при выполнении заданий.

Многие современные студенты выбирают наиболее простой путь выполнения задания – делегирование его выполнения СИИ. Впоследствии полученное решение зачастую просто копируется студентом и выдаётся за свое. Это наиболее губительный для образования сценарий использования БЯМ. При таком подходе студент получает некое решение, но он сам до конца не понимает не то что оснований и процесса его получения, но и правильности. В данном случае БЯМ выступает аналогом решебника в школе.

С учётом того, что целью обучения в высшем учебном заведении является не только освоение фактических знаний, но и получение навыков выработки решений, зачастую преподаватели требуют при сдаче задания устных пояснений или ответов на контрольные вопросы, как правило, в случае простого копирования студенты не способны аргументировать решение. Более того, в некоторых случаях, если решение, полученное от искусственного интеллекта, не является верным или оптимальным, студенты крайне негативно воспринимают критику данного решения преподавателем, утверждая, что формально задание выполнено. В этом случае упускается из виду, что задания на практические и лабораторные работы зачастую является моделью или фрагментом реальных задач и выполнять его нужно не для формальной сдачи работы, а для освоения реальных практических навыков.

Можно отметить такой общий недостаток сгенерированных ответов как некоторая стилистическая несвязность текста, наличие логических и, порой, фактических ошибок. Другим недостатком является то, что хоть информация берется системой из открытых источников, она не может указать, откуда конкретно взята информации, а при попытке заставить указать источники это может быть сделано формально (как, впрочем,

зачастую это и делается студентами). При этом возможно искажение информации, представленной в источнике, неверная её трактовка, что негативно сказывается на авторском праве.

С таким использованием студентами БЯМ необходимо бороться. В первую очередь необходимо проводить разъяснительную работу, касающуюся корректного использования современных технологий в деятельности студента. Помимо этого, необходимо пресекать попытки такой формальной сдачи работы без её понимания. На данный момент существуют специализированные СИИ, которые обучаются распознавать работы, выполненные с помощью БЯМ.

Другим, более благоприятным для образовательного процесса вариантом использования СИИ студентами при выполнении работы является выполнение задания с помощью БЯМ и дальнейший анализ решения, с возможной корректировкой запросов к системе. Кроме того, можно задавать уточняющие вопросы к системе, с целью понять, почему она выполнила задание именно таким образом. Здесь многое зависит от того, на что опиралась система при выполнении задания. Если источники были качественными, а задание выполнено верно, то система способна «объяснить» ход решения студенту. Проблемой является то, что нельзя заранее знать, как система выполнит задание.

В таких случаях необходимо, чтобы студент понимал хотя бы идею верного решения, примерный путь и «наставлял» БЯМ на него. В этом случае, решение может быть верным, и студент получит знания и навыки его осуществления. Как правило, в таких случаях студент способен рассказать ход решения и некоторые причины выбора определенных подходов. Однако при небольшом изменении задания у студента могут возникнуть затруднения по самостоятельному изменению решения.

Наиболее благоприятным вариантом использования БЯМ при выполнении задания является использования их в качестве дополнительного источника информации, справочника, а также «умной» поисковой системы. В этом случае студент сам выполняет задание, а по возникающим вопросам получает справочную информацию от СИИ. В этом случае достигается одна из основных педагогических задач – освоение методов системного мышления, самостоятельного мышления с использованием доступных источников информации. Студент критически осмысливает предложения БЯМ и может на их основе предлагать собственные варианты решения задачи. Если БЯМ использовались таким образом, то при сдаче работы у студента не возникает проблем с аргументацией решения, а также с адаптацией решения при изменении задания.

4.2 Возможности и риски применения искусственного интеллекта при самостоятельной работе студента

Важным элементом образовательного процесса является самоподготовка студентов. Здесь существует достаточно большой простор для применения СИИ. Студент может использовать их для поиска новой информации (с учётом особенностей работы с БЯМ) или задавать БЯМ вопросы, которые стесняется задать преподавателю. Например, если во время лекции студент не понял какой-либо фрагмент и не имеет возможности переспросить преподавателя, он может попытаться попросить БЯМ объяснить этот фрагмент.

Разумное применение БЯМ при самоподготовке студентов с последующим критическим осмыслением и верификацией полученной информации, а также с обсуждением непонятных или сложных моментов с преподавателем является одним из наиболее благоприятных сценариев использования СИИ студентами.

На сегодняшний день цель обучения в высшей школе заключается не просто в достижении студентами академических результатов, но и в развитии личности, социализации, формировании мировоззрения. Оптимизация процесса получения знаний студентами с помощью ИИ, с одной стороны, позволяет осваивать больший объем информации, но с другой стороны, может ограничить формирование мыслительных операций по пониманию, получению и созданию нового знания [4].

Кроме того, необходимо, чтобы сохранялось общение преподавателя и студента с целью социализации студента, получения преподавателем обратной связи и реализации диалектического взаимодействия, приводящего к получению качественного образования.

5. Искусственный интеллект при организации образовательного процесса

На данный момент существует ряд систем, направленных на комплексное внедрение СИИ в образовательный процесс, в том числе с целью его организации [5]. Такие системы позволяют на основе аудио- и видеозаписи урока, всего процесса взаимодействия преподавателя и обучающихся организовывать образовательный процесс. Данные системы анализируют поведение, действия обучающихся и преподавателя на уроке, формируют отчёт по каждому учащемуся, с учётом его вовлеченности, могут предлагать индивидуальные задания. Кроме того, анализируется поведение преподавателя, его речь, действия, интонации и формируется отчёт об уроке с указанием негативных моментов. Данные отчёты могут автоматически поступать администратору образовательной организации для принятия каких-либо решений (в том числе на основе рекомендаций СИИ).

Внедрение таких систем вызывает существенную перестройку педагогического коллектива, обучающихся и их законных представителей. В некоторых случаях, особенно на этапе внедрения и подстройки системы под новых обучающихся, система может выдавать некорректные рекомендации. Преподаватель вынужден [5] вручную размечать записи уроков для лучшего распознавания обучающихся. Такие системы поднимают вопрос защиты персональных данных учащихся и преподавателей. Все данные передаются для обработки оператору системы, что несёт в себе существенные риски и может быть негативно воспринято всеми участниками образовательного процесса.

Заключение

Можно сделать вывод, что системы искусственного интеллекта, включая большие языковые модели активно входят в образовательный процесс. При этом они являются инструментом, которым каждый участник образовательного процесса может распорядиться по своему усмотрению. При корректном использовании БЯМ могут стать существенным подспорьем, как преподавателю, так и студенту.

Однако необходимо учиться и учить использовать БЯМ с учётом особенностей их работы. Неаккуратное использование БЯМ и подмена образовательного процесса «игрой двух моделей», когда одна модель составляет задания, а другая их выполняет, может привести к деградации и существенному ухудшению качества образования.

Список литературы

1. Ландшафт основных терминов в области генеративного AI, их взаимосвязь и употребление [Электронный ресурс] : Habr URL : https://habr.com/ru/articles/950600/ (дата обращения 13.01.2026).

2. Искусственный интеллект в российском высшем образовании: текущее состояние и перспективы развития / С. Г. Давыдов, Н. Н. Матвеева, Н. В. Адемукова, А. А. Вичканова // Университетское управление: практика и анализ. – 2024. – Т. 28, № 3. – С. 32-44.

3. Букина Т.В. Искусственный интеллект в образовании: современное состояние и перспективы развития // Общество: социология, психология, педагогика. – 2025. – № 1. – С. 76-83.

4. Гасанова Р. Р., Романова Е. А. Искусственный интеллект в высшей школе: проблемы, возможности, риски // Вестник РУДН. Серия: Информатизация образования. – 2024. – № 4. – С. 501-505.

5. О. В. Трунова Искусственный интеллект – мощный инструмент для оптимизации и улучшения качества образовательного процесса // Материалы XXV Международной научно-практической конференции им. Э.К. Алгазинова. – Воронеж, 2025. – С. 1236-1241.

Мультиагентная система для автоматизированной генерации учебных заданий

В. В. Кириллов, email: kirillov@mirea.ru

РТУ МИРЭА

Аннотация. *Рассмотрена задача построения мультиагентной системы для мультимодального анализа образовательного контента. Предложена архитектура специализированных агентов на основе больших языковых и визуально-языковых моделей, обеспечивающая генерацию адаптивных заданий.*

Ключевые слова: *Мультиагентные системы (MAS), мультимодальные модели, большие языковые модели (LLM), компьютерное зрение (CV), Искусственный интеллект в образовании.*

Введение

Современная образовательная экосистема претерпевает фундаментальные изменения под влиянием процессов цифровизации и экспоненциального роста объёмов информации Электронные курсы, онлайн-платформы, системы дистанционного обучения и корпоративные LMS используют разнообразные форматы представления информации [1].

В этих условиях возрастает потребность в интеллектуальных системах, способных интегрировать методы анализа данных, генеративные модели и механизмы адаптации обучения в единую когнитивную инфраструктуру.

Одной из наиболее трудоемких и интеллектуально затратных задач для преподавателя является разработка адаптивных проверочных заданий, которые не только оценивали бы усвоение фактов, но и проверяли глубокое понимание, умение применять знания в новых ситуациях и интегрировать информацию из разных источников. Существующие инструменты автоматической генерации вопросов (Automatic Question Generation – AQG) в основном ориентированы на обработку текстовой информации, извлекая факты и преобразуя их в вопросы закрытого типа [2]. Однако они оказываются неэффективными перед лицом современного мультимодального контента, где ключевая смысловая нагрузка может быть распределена между текстом, визуальным рядом и структурными элементами. В результате генерируемые задания оказываются поверхностными, не затрагивают

© Кириллов В. В., 2026

межпредметные связи и не способствуют развитию критического мышления, оставаясь в парадигме проверки памяти, а не понимания материала.

В последние годы значительное развитие получили методы машинного обучения и глубоких нейронных сетей, в том числе большие языковые модели (LLM) и визуально-языковые модели (VLM), демонстрирующие высокую эффективность при анализе текстовых и визуальных данных. Эти модели открывают новые возможности для автоматического анализа образовательного контента и генерации учебных заданий. Традиционные же подходы к обработке образовательного контента строились на конвейере узкоспециализированных моделей: классификаторов именованных сущностей (NER) для текста, детекторов объектов для изображений, парсеров для таблиц. Такой подход, хотя и обеспечивал высокую точность в узких задачах, был трудоёмким в настройке и неспособным к обобщённому пониманию и рассуждению [3].

Одним из подходов к преодолению указанных ограничений является применение мультиагентных систем, в рамках которых задачи анализа и принятия решений распределяются между специализированными агентами. Каждый агент может быть ориентирован на обработку отдельной модальности или выполнение конкретной функции — анализа текста, интерпретации изображений, оценки уровня сложности материала или генерации адаптивных заданий. Взаимодействие агентов позволяет реализовать коллективное принятие решений и повысить устойчивость и гибкость системы в целом.

1. Мультиагентный подход

Мультиагентный подход особенно перспективен в сочетании с современными большими языковыми моделями, которые могут использоваться в качестве интеллектуального ядра отдельных агентов. Такая архитектура обеспечивает модульность системы, возможность независимого развития и обучения агентов, а также повышает прозрачность процессов принятия решений, что является важным требованием для образовательных приложений. Кроме того, мультиагентные системы позволяют реализовать адаптивную генерацию заданий с учетом как содержания учебного материала, так и индивидуальных особенностей обучаемого.

Особую актуальность рассматриваемая проблематика приобретает в медицинском образовательном контексте, где формирование клинического мышления требует анализа и интерпретации информации, представленной в различных модальностях. Обучение клинической диагностике предполагает работу с текстовыми описаниями симптомов,

медицинскими изображениями (рентгенография, МРТ, КТ, УЗИ), биомаркерами и временными рядами физиологических сигналов. Игнорирование одной из модальностей приводит к искажению представлений о клинической картине и снижает качество подготовки специалистов.

AI-агенты на базе LLM в медицинском образовании сейчас формируют несколько ключевых линии развития: симуляционные «виртуальные пациенты» и контекстно-зависимые ассистенты-тьюторы.

В частности, работа, опубликованная в JMIR (2025), представляет обзор LLM-базированных виртуальных пациентов — автономных AI-агентов, имитирующих реальных пациентов с высокой степенью интерактивности: они способны вести диалог на естественном языке, демонстрировать невербальные реакции (включая зрительный контакт и мимику), а также интегрироваться в среды смешанной и виртуальной реальности (MR/VR) [4]. Такие системы позволяют студентам отрабатывать клиническое мышление в безопасной, масштабируемой и доступной среде без временных и географических ограничений.

Другое ключевое направление — персонализированная поддержка обучения. В исследовании, опубликованном в Nature npj Digital Medicine (2025), описывается RAG-базированный AI-ассистент, который отвечает только на основе авторизованных учебных материалов, снижая риск галлюцинаций и повышая доверие со стороны студентов [5]. Такой подход обеспечивает контекстно-зависимую помощь, интегрированную непосредственно в учебный процесс.

Идея мультиагентных систем существует десятилетиями, но её практическая реализация для сложных когнитивных задач с использованием LLM стала доступной лишь с появлением специализированных высокоуровневых фреймворков. Эти фреймворки абстрагируют низкоуровневые детали коммуникации и оркестрации, позволяя разработчику сосредоточиться на определении ролей, компетенций и протоколов взаимодействия агентов.

Фреймворк AutoGen предлагает модель взаимодействия агентов в виде диалогов, где каждый агент использует LLM для генерации сообщений и принятия решений. Такой подход упрощает реализацию коллективного рассуждения, однако слабо формализует структуру процессов и может приводить к непредсказуемому поведению системы.

LangGraph, в свою очередь, использует графовую модель управления агентами, позволяя явно задавать состояния, переходы и точки синхронизации. Это повышает контролируемость и воспроизводимость работы системы, что является важным фактором для образовательных и медицинских приложений [6].

CrewAI ориентирован на организацию команд агентов с чётко определёнными ролями и целями. Данный фреймворк удобен для прототипирования, однако требует дополнительной адаптации для задач мультимодального анализа и интеграции визуальных моделей.

2. Предлагаемая архитектура системы

Система рассматривается как совокупность взаимодействующих интеллектуальных агентов, каждый из которых решает специализированную задачу обработки данных или принятия решений. Такой подход обеспечивает независимость агентов, возможность их поэтапного развития и гибкость при добавлении новых модальностей или образовательных сценариев. Общая архитектура системы представлена на рис. 1.

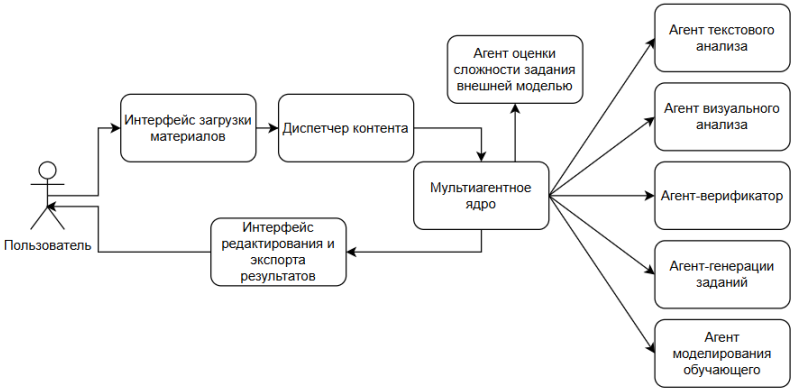


Рис. 1. Верхнеуровневая архитектура системы.

Входной интерфейс принимает мультимодальный образовательный ресурс, такой как PDF-файл лекции, веб-страницу с текстом и изображениями, видеоматериал занятия. Диспетчер контента - модуль, отвечающий за предварительную обработку, извлечение текста, выделение изображений, парсинг таблиц и метаданных. Результатом его работы является нормализованный пакет данных, распределённый по модальностям.

Мультиагентное ядро системы состоит из пула специализированных агентов и координатора. База знаний хранит доменные концепции, отношения и педагогические шаблоны для верификации и обогащения результатов анализа. Генератор заданий -модуль, который, используя объединённую семантическую модель от координатора и педагогические

правила, формулирует финальные задания различных типов. Выходной интерфейс представляет сгенерированные задания в структурированном формате (JSON, XML) или в виде готовых для импорта в стандартах образовательных платформ, таких как Moodle.

Агент текстового анализа предназначен для обработки текстовых компонентов образовательного контента, включая лекционные материалы и методические указания. Его основные функции — это извлечение ключевых понятий, терминов и сущностей, анализ логической структуры учебного материала, выявление предпосылок и следствий, заложенных в тексте, а также формирование структурированных представлений знаний, пригодных для последующего использования другими агентами.

Агент визуального анализа ориентирован на обработку изображений и визуальных элементов образовательного контента. В медицинском домене данный агент анализирует диагностические изображения, включая рентгенограммы, таблицы и схемы, с использованием визуально-языковых моделей. Результаты работы агента представляются в виде структурированных описаний, которые могут быть использованы как для генерации заданий, так и для формирования объяснений обучаемым.

Агент-критик выполняет функцию верификации и согласования результатов, полученных от других агентов. Его основная задача заключается в выявлении противоречий, логических несоответствий и потенциальных ошибок в выводах, сформированных на предыдущих этапах анализа.

Агент моделирования обучаемого формирует и обновляет модель обучаемого на основе его взаимодействия с системой. В модель включаются данные о результатах выполнения заданий, времени ответа, характере ошибок и предпочтительных формах представления информации.

Агент генерации адаптивных заданий отвечает за формирование учебных заданий, ориентированных на текущий уровень подготовки и образовательные цели обучаемого. Генерация заданий осуществляется с учетом педагогических инструкций, которые задают требования к форме, сложности и дидактическим целям заданий. Такие инструкции могут включать указания на допустимый уровень абстракции, необходимость использования определённых модальностей или акцент на развитие клинического мышления.

Для обеспечения необходимого уровня сложности генерируемых заданий в системе используется специализированный агент оценки сложности. Данный агент осуществляет оценку сложности заданий и

ответов с привлечением внешней языковой модели, не имеющей доступа к исходному образовательному контексту и материалам.

Взаимодействие агентов организовано в рамках кооперативной мультиагентной архитектуры, где каждый агент вносит вклад в формирование итогового образовательного продукта. Агент-критик и агент оценки сложности выполняют функции контроля качества и педагогической валидности, в то время как агенты анализа и генерации обеспечивают содержательную сторону процесса.

3. Практическая реализация прототипа системы

Практическая реализация предложенной мультиагентной системы была выполнена с использованием современных больших языковых и визуально-языковых моделей, а также специализированного фреймворка для организации агентного взаимодействия. При выборе моделей и инструментов учитывались требования к качеству обработки русскоязычного образовательного контента, мультимодальной интерпретации данных и возможности локального развертывания системы.

В качестве базовых моделей для реализации агентов текстового и визуального анализа были выбраны языковая модель Qwen3 и визуально-языковая модель Qwen3-VL размером около 30 миллиардов параметров. Модели семейства Qwen демонстрируют высокое качество обработки русскоязычных текстов, включая специализированную терминологию, что является критически важным для задач образовательного и медицинского домена.

Эксперименты проводились на графическом процессоре NVIDIA RTX 5090, обладающем достаточным объёмом видеопамяти для размещения моделей указанного размера. Развёртывание моделей производилось с помощью фреймворка vLLM, который обеспечивает высокую пропускную способность и низкую задержку при генерации текста за счёт использования передовых техник оптимизации. Использование локального вычислительного ресурса позволило обеспечить воспроизводимость экспериментов, снизить задержки при взаимодействии агентов и исключить зависимость от внешних облачных сервисов, что является важным аспектом для образовательных и медицинских приложений.

Для реализации взаимодействия между агентами был выбран фреймворк LangGraph, предоставляющий средства для построения агентных систем на основе графовой модели управления процессами. В отличие от диалогово-ориентированных фреймворков, LangGraph позволяет явно задавать структуру вычислительного графа, состояния агентов и условия переходов между этапами обработки. Выбор

LangGraph также обусловлен возможностью гибкой модификации архитектуры системы без существенной переработки кода, что важно при экспериментальном исследовании различных конфигураций агентов и паттернов взаимодействия.

Для оценки эффективности предложенной системы были рассмотрены несколько экспериментальных сценариев, ориентированных на анализ образовательного контента медицинской направленности. В качестве входных данных использовались учебные тексты, сопровождаемые диагностическими изображениями, характерными для обучения клиническому мышлению. На рис. 2 показан пример задания, сгенерированного агентной системой на основании учебного пособия по радиологии и рентгенологии [7].

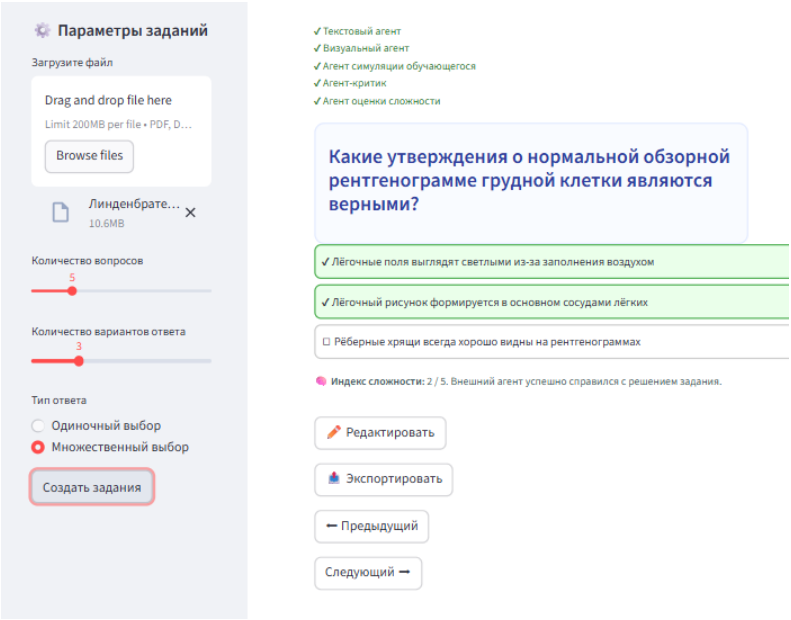


Рис. 2. Пример работы агентной системы.

В рамках данного примера был рассмотрен раздел учебного материала по лучевой анатомии легких. В интерфейсе разработанного streamlit приложения можно настроить параметры генерации ответов на вопрос, после чего агентная система предлагает варианты заданий, которые пользователь может изменить их и экспортировать в удобном

формате. Для оценки сложности вопроса использовалось обращение по API к внешнем языковым моделям, в рамках тестирования прототипа использовались модели Deepseek и GigaChat, в приведённом примере они успешно справились с подготовленным тестовым вопросом.

Результаты экспериментов с представленной системой показали, что использование мультиагентной архитектуры позволяет повысить согласованность мультимодального анализа образовательного контента по сравнению с монолитным использованием языковой или визуально-языковой модели. Привлечение агента-критика и независимой оценки сложности способствует снижению числа логических ошибок и повышению педагогической обоснованности генерируемых заданий.

Заключение

В данной работе была рассмотрена задача разработки мультиагентной системы для мультимодального анализа образовательного контента и генерации адаптивных заданий, ориентированной на применение в образовательных системах с высокой долей визуальной и текстовой информации. Особое внимание уделено медицинскому образовательному домену, где формирование клинического мышления требует комплексного анализа разнородных данных и повышенной интерпретируемости интеллектуальных решений.

Вместе с тем предлагаемая система обладает рядом ограничений, которые следует учитывать при её практическом применении. Во-первых, качество работы системы напрямую зависит от характеристик используемых языковых и визуально-языковых моделей, включая их способность корректно интерпретировать специализированную терминологию и визуальные признаки. Во-вторых, мультиагентная архитектура увеличивает сложность разработки и отладки системы, а также предъявляет дополнительные требования к вычислительным ресурсам при одновременной работе нескольких агентов. Несмотря на использование агента-критика и независимой оценки сложности, существует возможность генерации некорректных заданий система рассматривается исключительно как инструмент поддержки образовательного процесса, а не полной замены преподавателя.

Перспективы дальнейших исследований связаны с расширением набора анализируемых модальностей, включая временные ряды физиологических сигналов и аудиоданные, а также с разработкой формальных метрик оценки эффективности адаптивных заданий. Отдельный интерес представляет интеграция системы в реальные образовательные платформы и проведение пользовательских исследований с участием обучаемых и преподавателей.

Таким образом, результаты работы подтверждают перспективность применения мультиагентных и мультимодальных систем на основе современных языковых моделей для задач адаптивного обучения, особенно в медицинском образовательном домене, и создают основу для дальнейших исследований в области интеллектуальных образовательных технологий.

Список литературы

1. Overcoming Information Overload in Higher Education: The Power of Document Summarization [Электронный ресурс]: многопредмет. науч. журн. – Режим доступа: https://research.com/education/overcoming-information-overload-in-higher-education
2. Достижения в области автоматической генерации вопросов: обзор современных методов и технологий / Д. В. Лопаткин [и др.] // Вестник ВКТУ. – 2025. – № 2. – С. 172-181.
3. Парсинг сайтов с помощью нейронных сетей: лучшие подходы и практики [Электронный ресурс]: – Режим доступа: https://xmldatafeed.com/parsing-sajtov-s-pomoshhyu-nejronnyh-setej-luchshie-podhody-i-praktiki
4. Embracing the Future of Medical Education With Large Language Model–Based Virtual Patients: Scoping Review / Jianwen Zeng [et al.] // Journal of Medical Internet Research – 2025. – № 27.
5. Thesen, T. A generative AI teaching assistant for personalized learning in medical education / T. Thesen, S. H. Park //npj Digital Medicine. – 2025. – № 8, 627.
6. LangChain Documentation: LangGraph overview [Электронный ресурс]: документация фреймворка. – Режим доступа: https://docs.langchain.com/oss/python/langgraph/overview
7. Линденбратен, Л. Д. Медицинская радиология (Основы лучевой диагностики и лучевой терапии), Учеб. для студентов мед. вузов / Л. Д. Линденбратен, И. П. Королюк. – 2-е изд., перераб. и доп. – М. : Медицина, 2000. – 670 с.

Инструменты генеративного ИИ в академической деятельности вуза: взгляд студентов

А. В. Сычев, email: sav@cs.vsu.ru

Воронежский государственный университет

Аннотация. *В докладе приведены результаты анкетирования бакалавров, специалистов и магистрантов об использовании ими инструментов генеративного искусственного интеллекта для учебной и научно-исследовательской работы. Анкетирование проводилось в конце первого семестра и охватило около 200 студентов. В анкете содержалось более 30 вопросов, которые охватывали самые разнообразные аспекты использования инструментов генеративного искусственного интеллекта, оценку возможностей инструментов и этических аспектов их применения. Проведенное исследование результатов анкетирования показывает, что инструменты генеративного ИИ стали массовым и повседневным инструментом в учебно-исследовательской практике студентов. Студенты демонстрируют прагматичный и избирательный подход, используя эти инструменты преимущественно как интеллектуального помощника для объяснения тем, поиска информации и подготовки к экзаменам. При этом они способны разделять этичные и неэтичные практики применения, но отсутствие четких институциональных правил создает зону неопределенности. Ключевым выводом является запрос студенческого сообщества не на запреты, а на формирование ИИ-грамотности через образовательные инициативы и прозрачные регламенты.*

Ключевые слова: *генеративный искусственный интеллект, большие языковые модели, галлюцинации, критическое мышление, генерация контента, этика искусственного интеллекта, академическая честность, ИИ-грамотность, анкетирование.*

Введение

Широкое распространение инструментов генеративного искусственного интеллекта (ГенИИ) в корне меняет академическую реальность современного университета. На переднем крае этой трансформации находятся студенты, которые активно адаптируют технологии в свою учебную практику, формируя новый, сложный и зачастую противоречивый взгляд на их роль. В отличие от дискуссий

© Сычев А. В., 2026

среди администраторов и преподавателей, отношение студентов к инструментам ГенИИ представляет собой сочетание прагматизма, осторожного оптимизма и осознанных этических тревог.

Текущие данные показывают, что использование ГенИИ студентами стало массовым явлением: до 85% учащихся высшей школы применяли его для учебных задач за последний год [1]. Однако ключевой вопрос заключается не в факте использования, а в его характере. Студенческое сообщество демонстрирует удивительную зрелость, разделяя допустимые и недопустимые практики. Наиболее распространенными и социально одобряемыми являются вспомогательные функции: генерация идей (55%), использование в качестве тьютора (50%) и подготовка к экзаменам (46%). Эти паттерны говорят о стремлении использовать ИИ как интеллектуального партнера для углубления понимания, а не как простой инструмент для выполнения работы [2].

Одновременно студенты глубоко озабочены вопросами академической честности, и их взгляд на этот вопрос неоднозначен. С одной стороны, подавляющее большинство (69%) считают неэтичным использование ChatGPT для написания готовых эссе или экзаменационных ответов. С другой — лишь меньшинство видят этическую проблему в применении ИИ для предварительного исследования (30%) или мозгового штурма (20%). Этот этический водораздел формирует главный запрос студенческого сообщества к университетам: не запреты и полицейский контроль, а четкие правила, прозрачность и образовательные инициативы.

Студенты ожидают от вузов не карательных мер (например, в виде ПО для детектирования ИИ генерации), а обучения навыкам ИИ-грамотности. Они хотят понимать, как правильно формулировать эффективные промпты, критически оценивать выводы ИИ, распознавать «галлюцинации» и соблюдать академические нормы при использовании сгенерированного контента. Примечательно, что в то время, как студенты позитивно относятся к применению ИИ для поддержки обучения, они скептически настроены в отношении замены им ключевых функций преподавателя, таких как создание курсов, оценка работ и тем более полное проведение занятий по учебной дисциплине [3].

Тем не менее, внедрение ГенИИ сопряжено с серьезными рисками, которые также осознаются учащимися. К ним относятся: снижение навыков критического и творческого мышления из-за чрезмерной зависимости от алгоритмов, опасения по поводу конфиденциальности данных и алгоритмической предвзятости в системах оценивания, а также усиление образовательного неравенства, поскольку исследования

фиксируют разрыв в использовании и уверенности между разными группами студентов [4].

Таким образом, взгляд студентов на инструменты ГенИИ в вузе можно охарактеризовать как прагматично-сбалансированный. Они признают технологию как неизбежный и мощный элемент новой образовательной экосистемы, который может стать катализатором более персонализированного и эффективного обучения. Однако они настаивают на ответственной интеграции, где технология выступает «умным ассистентом», а не суррогатом интеллектуального труда, и требуют от университетов руководства в формировании этических и компетентных моделей ее использования [5].

Результаты анкетирования

1. Бакалавры и специалисты

На факультете компьютерных наук ВГУ по направлениям подготовки бакалавров 09.03.02 Информационные системы и технологии, 09.03.04 Программная инженерия, 10.03.01 Информационная безопасность и специалистов 10.05.01 Компьютерная безопасность первого года обучения в конце первого семестра 2025-2026 учебного года было проведено анонимное анкетирование. Общее количество анкет составило 160, распределение заполненных анкет по направления подготовки представлено на рис. 1. Анкета содержала 33 вопроса.



Рис. 1. Распределение анкет по направлениям подготовки студентов

По используемым студентами инструментам было выявлено заметное предпочтение зарубежных больших языковых моделей (БЯМ) (рис. 2).

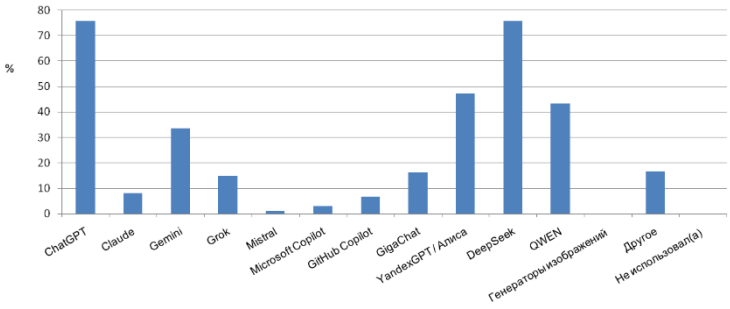


Рис. 2. Используемые студентами инструменты

Опыт использования инструментов Ген ИИ представлен на рис. 3. Видно, что всего лишь 10% студентов не имели опыта использования этих инструментов до поступления в вуз, а активно их использовали — 36%.

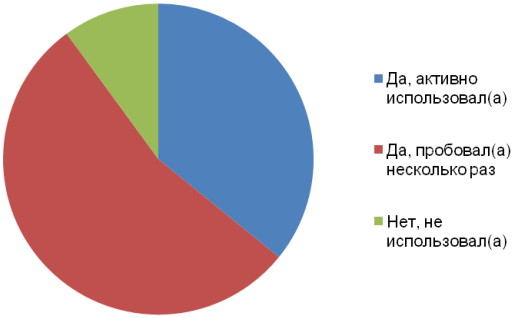


Рис. 3. Частота использования инструментов ГенИИ (до поступления в вуз)

Ответы на вопрос: «Для каких учебных задач вы используете генеративный ИИ?» распределились так, как это представлено на рис. 4.

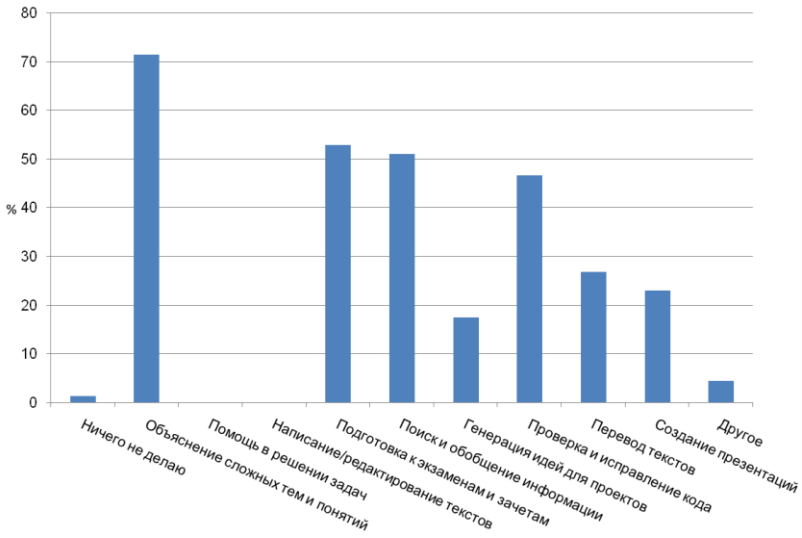


Рис. 4. Учебные задачи, для которых используется ГенИИ

В конце первого семестра характер использования ГенИИ студентами первого года обучения имел структуру, представленную на рис. 5. То есть половина всех студентов применяют эти инструменты несколько раз в неделю, а систематически – до трети от всех студентов.

Рассмотрим теперь оценку студентами используемых ими инструментов. Оценки давались по 5-балльной шкале. Результаты показаны на рис. 6.

Можно заметить, что наивысшие оценки получили такие характеристики, как «полезность для обучения» (средневзвешенный балл – 4.4), «понятность объяснений» (4.1 балла). Это коррелирует с тем, что наиболее часто студенты используют ГенИИ для задачи «Обяснение сложных тем и понятий» (более 70% ответов на рисунке 5). Далее идут задачи «Подготовка к экзаменам и зачетам» и «Поиск и обобщение информации». Еще одно важное для многих студентов применение этих инструментов – «Проверка и исправление кода».

Анкетирование подтверждает, что студенты имеют опыт получения ошибочных или неточных ответов от БЯМ (часто – 21%, иногда – 70%). При этом критичность студентов по отношению к получаемой от них информации отражена на рис. 7.

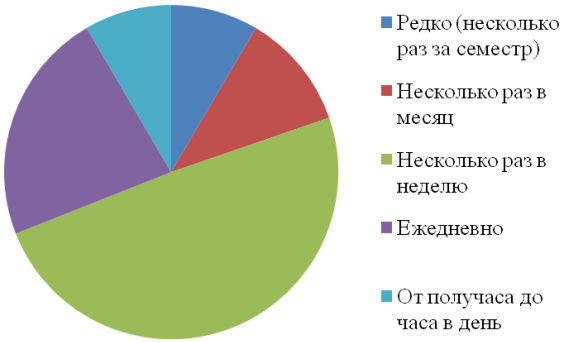


Рис. 5. Использование инструментов ГенИИ студентами в конце семестра

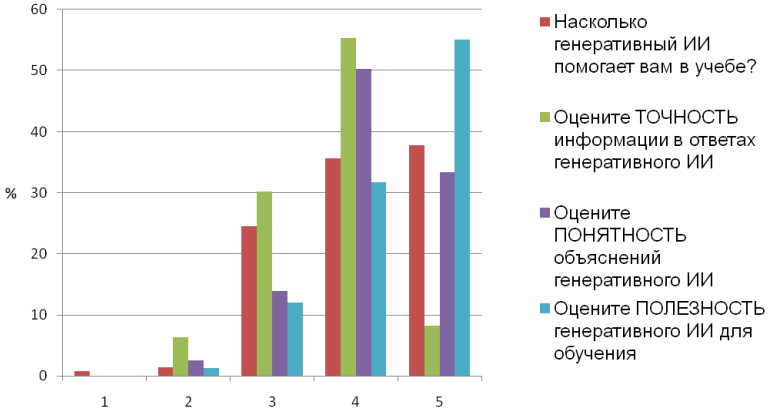


Рис. 6. Оценка используемых инструментов ГенИИ по качественным параметрам.

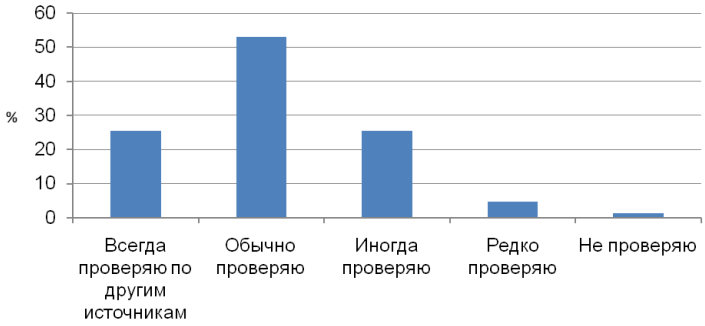


Рис. 7. Критическое отношение студентов к информации, получаемой от БЯМ.

Характер изменений в отношении использования ГенИИ студентами по итогам первого семестра отражается на рис. 8 и 9.

77% опрошенных заявили в целом об улучшении отношения к инструментам ГенИИ, 87% отметили увеличение частоты использования этих инструментов.

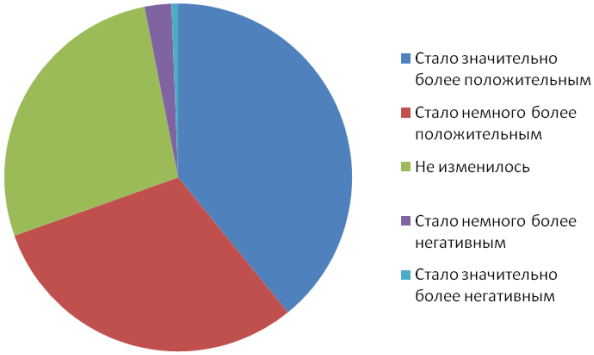


Рис. 8. Изменение отношения студентов к ГенИИ в течение семестра.

В качестве ключевых факторов, которые повлияли на изменение отношения студентов к ГенИИ были выбраны: личный опыт использования (72%), обсуждения с однокурсниками (46%), качество ответов ИИ (33%). К сожалению, рекомендации преподавателей и позиция университета/факультета повлияли лишь в малой степени: 18% и 6% соответственно.

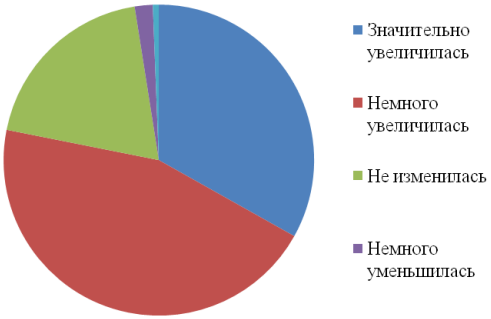


Рис. 9. Изменение частоты использования студентами инструментов ГенИИ в течение семестра.

В настоящее время консолидированная позиция вуза/факультета в отношении использования ГенИИ студентами отсутствует. Представление студентов об этой позиции отражено на рис. 10.

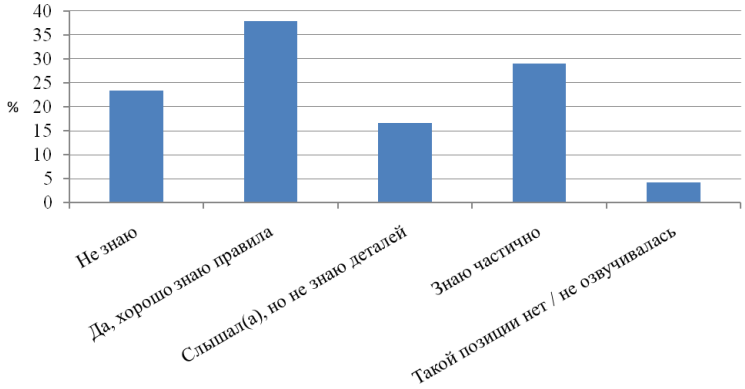


Рис. 10. Представление студентов о позицию университета/факультета относительно использования ГенИИ

Следующая группа вопросов в анкете касалась этических вопросов использования ГенИИ. На рис. 11 отражена готовность студентов информировать преподавателя об использованных ими инструментах ИИ при выполнении заданий. Около 70% предпочитают умолчать, если их прямо об этом не спрашивают.

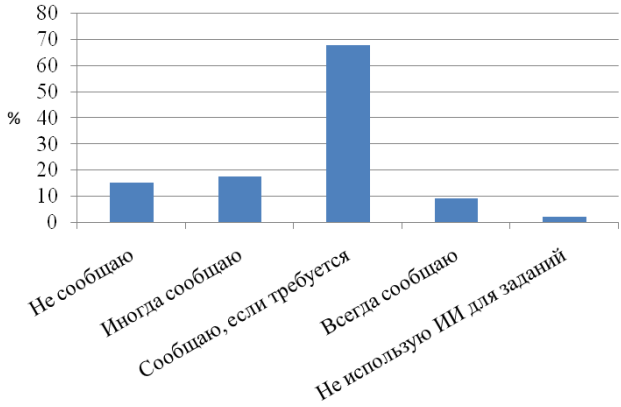


Рис. 11. Готовность студентов сообщать об использовании инструментов ГенИИ.

Оценка студентами этичности различных вариантов использования ГенИИ в своей академической деятельности по пятибалльной шкале (1 – полностью неэтично, 5 – полностью этично) отражена на рис. 12.

Приведенная на рисунке гистограмма в целом отражает понимание студентами базовых этических принципов академической честности при использовании ими инструментов ГенИИ.

Рефлексия студентов относительно влияния ИИ на их учебную деятельность представлена в виде оценок по 5-балльной шкале (1–сильно ухудшает, 5 – сильно улучшает) на рисунке 13. Самые высокие оценки получили «понимание материала» (взвешенный балл 4.1) и «успеваемость» (взвешенный балл 3.9).

Что касается опасений студентов относительно влияния генеративного ИИ на их обучение, 64% указали на отсутствие таковых, небольшие опасения выразили 26%, затруднились с ответом 19%.

Отдельный вопрос в анкете был относительно учебных дисциплин, для которых чаще всего использовался ГенИИ. Лидерами по числу упоминаний оказались программирование (42%), математический анализ (37%), основы российской государственности (27%), введение в программирование (25%).

Стоит также отметить, что имеется явно выраженный запрос студентов на большую степень информированности в университете о правильном использовании генеративного ИИ. Этого очень хотели бы 37% опрошенных, «скорее да» ответили 45%.

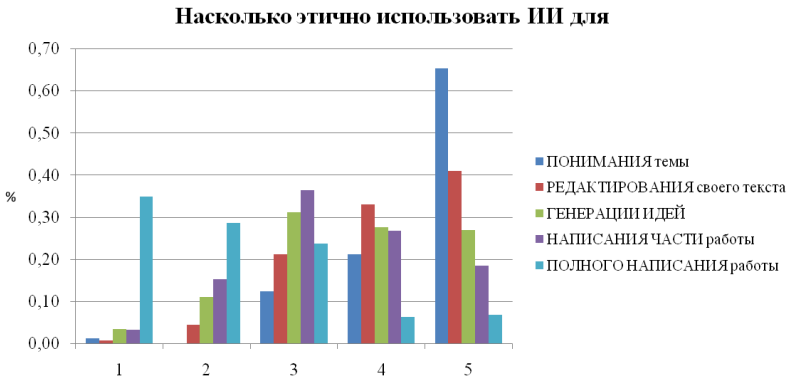


Рис. 12. Оценка студентами этичности различных вариантов использования ГенИИ.

На вопрос о том, что студенты хотели бы узнать или чему научиться в контексте ГенИИ, чаще встречались ответы: «как правильно писать промты», «получить более глубокое понимание работ».

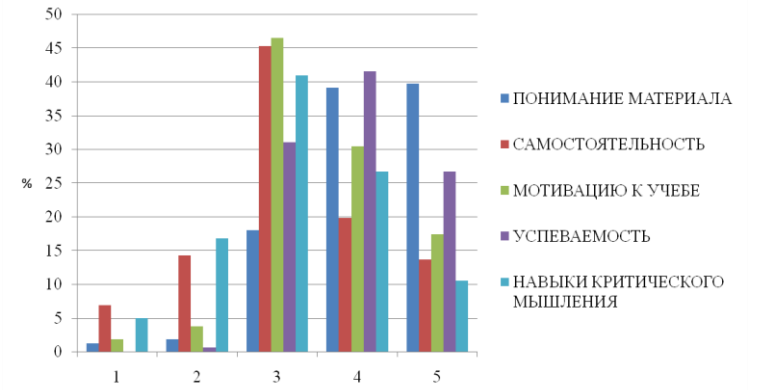


Рис. 13. Оценка студентами влияния ГенИИ на различные аспекты их учебной деятельности.

На вопрос о том, что студенты хотели бы узнать или чему научиться в контексте ГенИИ, чаще встречались ответы: «как правильно писать промты», «получить более глубокое понимание работы ИИ».

2. Магистранты

Для исследования аудитории магистрантов была разработана отдельная анкета, содержавшая 50 вопросов. Затем было проведено анонимное анкетирование, в котором приняли участие 10 магистров 2 года обучения по направлению 09.04.02 «Информационные системы» и 15 магистров первого года обучения по направлению 02.04.01 «Математика. Компьютерные науки».

Из опрошенных магистрантов, многие уже имеют опыт практической работы по специальности (рис. 14) и уже используют инструменты ГенИИ больше года (рис. 15).

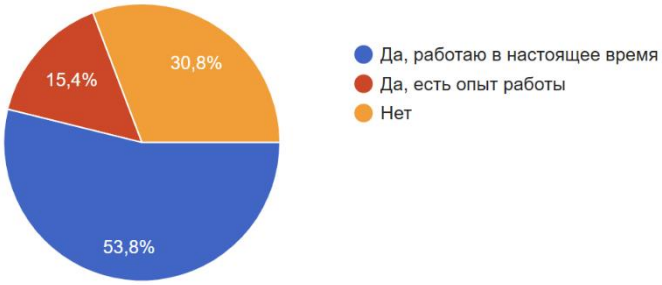


Рис. 14. Опыт работы по специальности



Рис. 15. Начало использования инструментов ГенИИ

По спектру используемых инструментов ГенИИ для учебы, исследований и работы распределение голосов представлено на рис. 16. 68% магистрантов пользуются бесплатными версиями инструментов.

Частота использования инструментов для различных сфер применения отражена на рисунке 17.

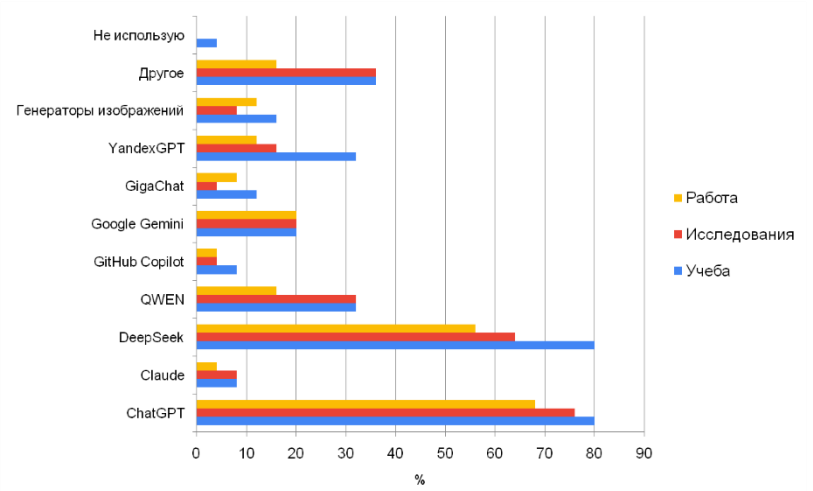


Рис. 16. Используемые инструменты ГенИИ

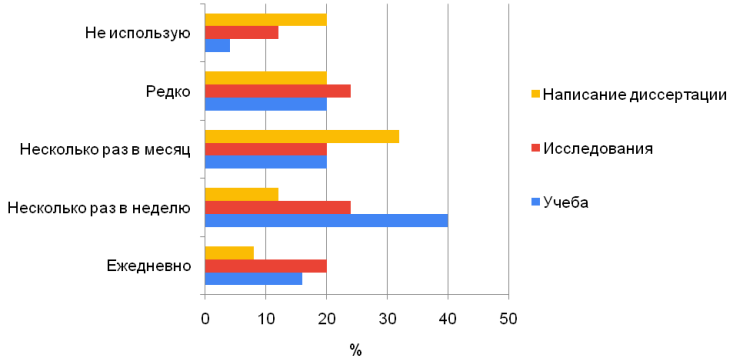


Рис. 17. Частота использования инструментов ГенИИ

Типичные задачи в рамках научно-исследовательской деятельности магистрантов, для которых ими применяются инструменты ГенИИ, отражены на рис. 18.

При оценке полезности инструментов голоса магистрантов распределились как представлено на рис. 19. Средневзвешенная оценка полезности получилась следующая: поиск литературы - 3.52 балла, понимание сложных концепций – 3.68 балла, генерация идей – 3.48 балла, написание кода – 3.44 балла, академическое письмо – 3.24 балла.

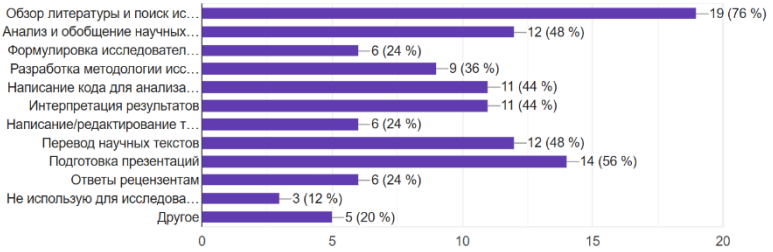


Рис. 18. Исследовательские задачи с применением инструментов ГенИИ

Оценка эффективности ГенИИ давалась с точки зрения влияния инструментов на скорость работы, качество текстов и глубину понимания темы (рис. 20). Средневзвешенные баллы по этим показателям составили: 4.0, 3.3, 3.7 баллов соответственно.

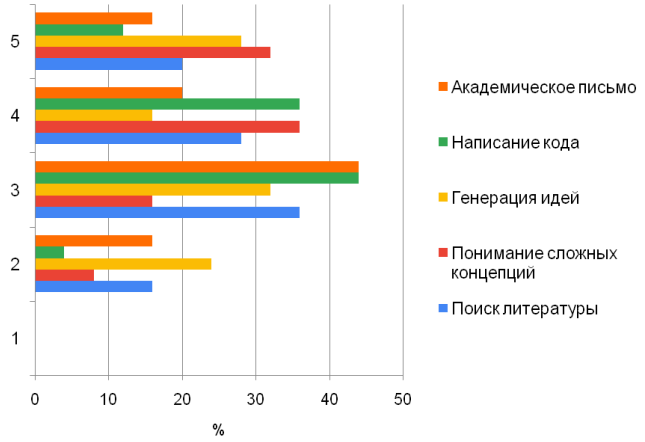


Рис. 19. Оценка полезности инструментов ГенИИ

Рис. 21 отражает саморефлексию магистрантов в отношении влияния использования ИИ на их навыки.

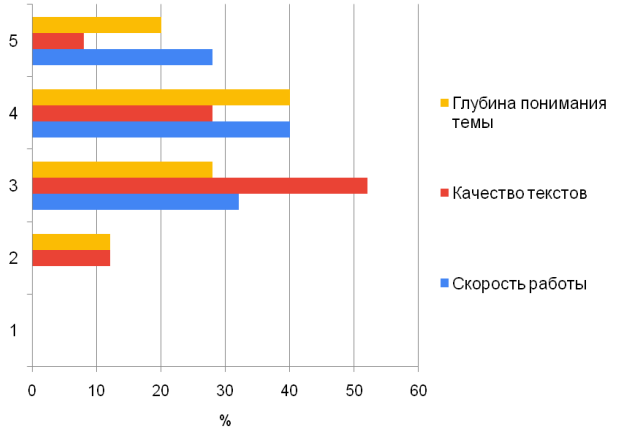


Рис. 20. Оценка эффективности инструментов ГенИИ

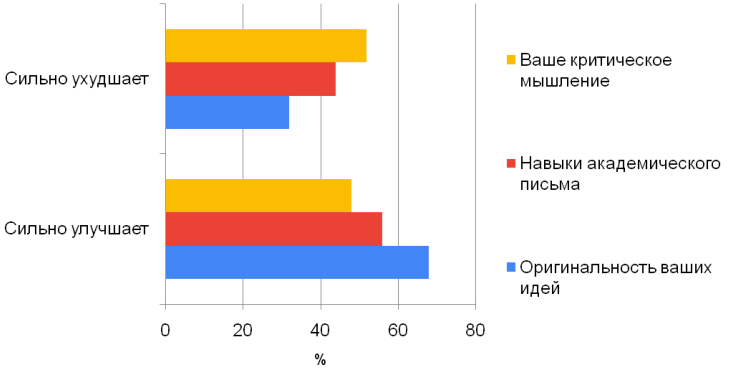


Рис. 21. Самооценка влияния инструментов ГенИИ на «мягкие» навыки магистрантов

По итогам осеннего семестра у некоторых магистрантов отношение к ГенИИ изменилось (рис. 22). При этом оценка возможностей ИИ не

изменилась у 56% из них, а у остальных примерно в одинаковых пропорциях оценка скорректировалась в лучшую или худшую сторону.

Этическая оценка применения ГенИИ представлена на рис. 23. При этом 44% магистрантов не указывают на использование ГенИИ в своих работах и только лишь 40% делают это только по требованию.

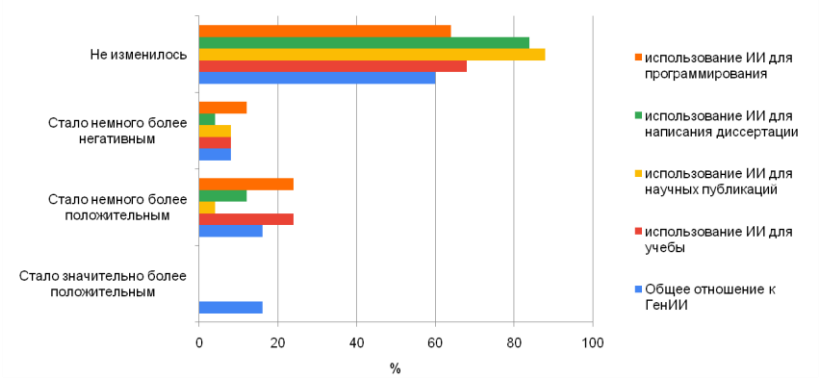


Рис. 22. Изменение отношения магистрантов к ГенИИ по итогам осеннего семестра

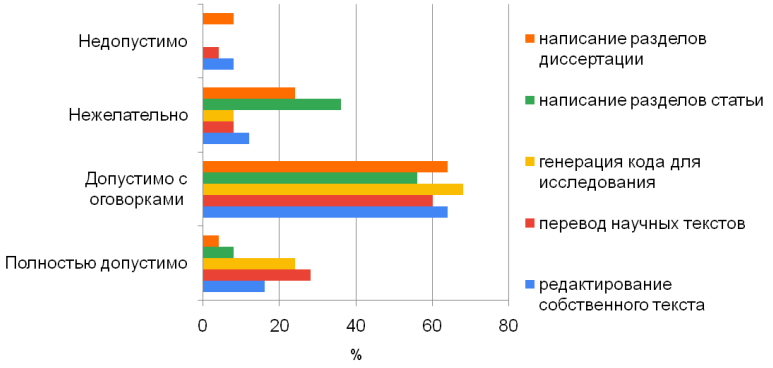


Рис. 23. Этическая оценка применения инструментов ГенИИ

При ответе на вопрос об опасениях, связанных с использованием ГенИИ, магистранты среди наиболее значимых указали зависимость от ИИ (52%) и обесценивание квалификации (40%).

На вопрос о необходимости специальных курсов/семинаров по использованию ИИ в исследованиях положительно ответили более половины опрошенных (обязательно – 20%, желательно – 44%).

Заключение

Проведенное эмпирическое исследование практик использования и восприятия инструментов ГенИИ студентами факультета компьютерных наук ВГУ позволяет сделать следующие основные выводы.

1. ГенИИ стал неотъемлемым элементом учебной и исследовательской практики студентов ФКН. Подавляющее большинство опрошенных (90% бакалавров) имели опыт использования ГенИИ до поступления в вуз, а в ходе первого семестра частота и интенсивность их применения значительно возросли. Важно отметить, что за один семестр наблюдалось значительное увеличение как частоты использования инструментов (87% респондентов) так и улучшение отношения к ним (77% респондентов).

2. Студенты демонстрируют прагматичный и вполне избирательный подход к использованию ГенИИ. Эта технология воспринимается не как средство для полного замещения интеллектуального труда, а скорее как «интеллектуальный помощник» или «партнер». Наиболее востребованными в студенческой среде являются такие функции инструментов ГенИИ как объяснение сложных понятий (>70% бакалавров), подготовка к экзаменам, поиск и обобщение информации, а также проверка кода. В исследовательской работе магистранты активно используют ГенИИ для поиска литературы, генерации идей и написания текстов.

3. Студенты демонстрируют критическое мышление в отношении технологий ИИ: они осознают риски получения неточной информации (91% сталкивались с ошибочными ответами БЯМ), опасаются развития зависимости от технологий (52% магистрантов) и возможного обесценивания их квалификации (40% магистрантов).

4. Этические границы использования ГенИИ студентами в целом осознаются, но требуют институционального закрепления. Студенты проводят четкую грань между этичным (мозговой штурм, объяснение тем) и неэтичным (написание готовых текстов работ) использованием. Однако при этом около 70% предпочитают не информировать преподавателя об использовании ГенИИ, если это явно не требуется. Что указывает на существующий «серый» сектор и неопределенность правил.

5. Ключевым запросом к университету является не запрет, а формирование ИИ-грамотности и четких регламентов. Большинство студентов либо не знают о позиции вуза по использованию ГенИИ, либо считают ее неозвученной. Существует и выраженная потребность в

образовательных инициативах: обучении эффективному промптингу, критической оценке ответов ИИ, распознаванию «галлюцинаций» и правилам академической честности при работе с генерируемым контентом. Более половины магистрантов поддержали введение специальных курсов по использованию ИИ в исследованиях.

Таким образом, для гармоничной интеграции технологий ГенИИ в образовательный процесс университету необходимо срочно перейти от выжидательной позиции к активному формированию политики ответственного использования ИИ. Эта политика должна включать в себя: 1) разработку прозрачных правил и стандартов цитирования; 2) внедрение учебных модулей по ИИ-грамотности в учебные планы; 3) поддержку преподавателей в адаптации методик преподавания и оценивания в новых условиях. Только такой комплексный подход позволит превратить генеративный ИИ из источника академических рисков в эффективный инструмент для усиления познавательной и исследовательской деятельности студентов.

Список литературы

1. Inside Higher Ed & The Generation Lab. (2025, August 29). Student Voice Survey: How AI Is Changing—Not 'Killing'—College. Inside Higher Ed. [Электронный ресурс] – Режим доступа: https://genain3.ie/how-ai-is-changing-not-killing-college/

2. University of Illinois Chicago (UIC). (2024, April 5). Report on Student Attitudes Towards AI in Academia. Office of the Vice Provost for Undergraduate Affairs and Academic Programs. [Электронный ресурс] – Режим доступа: https://learning.uic.edu/news-stories/report-on-student-attitudes-towards-ai-in-academia-at-uic/

3. University of Oxford. (n.d.). An introduction to the use of generative AI tools in teaching. Centre for Teaching and Learning. [Электронный ресурс] – Режим доступа: https://www.ctl.ox.ac.uk/ai-tools-in-teaching

4. Kalnitskaya, I. V. Риски использования технологий искусственного интеллекта в высшем образовании. Современные научные исследования и инновации. https://uronsk.ru/articles2/435/36/riski-ispolzovaniya-tehnologiy-iskusstvennogo-intellekta-v-vysshem-obrazovanii

5. Laverdiere, R., Henry, T., Parro, M., Allan, B., Alexander, S. Five Ways Higher Education Can Leverage Generative AI. [Электронный ресурс] – Режим доступа: https://web-assets.bcg.com/pdf-src/prod-live/five-ways-education-can-leverage-gen-ai.pdf

Секция 4
Подготовка ИТ-инженеров для решения задач цифровизации промышленности и бизнеса

Задача вычисления расстояния Левенштейна в интеллектуальном турнире Кибер-бой

Е. А. Алефиренко, email: alefirenkoea@mgpu.ru

Институт развития профильного обучения, Государственное автономное образовательное учреждение высшего образования города Москвы «Московский городской педагогический университет»

Аннотация. *В статье рассмотрена учебная задача вычисления расстояния Левенштейна и её аналитическое решение. Учебная задача была представлена в рамках интеллектуального турнира «Кибер-бой» для обучающихся инженерных классов московских школ.*

Ключевые слова: *Кибер-бой, интеллектуальный турнир, учебная задача, расстояние Левенштейна, динамическое программирование, информатика, инженерный класс.*

Введение

Интеллектуальный турнир «Кибер-бой» проводится в рамках реализации проектов предпрофессионального образования «Инженерный класс в московской школе», «ИТ-класс в московской школе». Организатор турнира – Институт развития профильного обучения Московского городского педагогического университета. Цель проведения турнира – стимулирование профессионального роста обучающихся через предоставление равных возможностей демонстрации результатов освоения содержания образовательных программ для участников вышеперечисленных проектов предпрофессионального образования [1, с. 1].

В данной статье рассмотрена учебная задача по информатике, предложенная для выполнения обучающимся десятых инженерных классов школ г. Москвы на первом квалификационном этапе интеллектуального турнира «Кибер-бой» в январе 2026 года.

В учебном задании обучающимся предложено ознакомиться с понятием расстояния Левенштейна и вычислить его для двух приведённых листингов программ.

Под учебным заданием Б.М. Бим-Бад понимает вид поручения учителя учащимся, в котором содержится требование выполнить какие-либо учебные (теоретические или практические) действия, при этом термины «задача» и «задание» рассматриваются Б.М. Бим-Бадом как синонимы. В.М. Давыдов определяет учебную задачу как задачу,

© Алефиренко Е. А., 2026

требующую от учащихся открытия и освоения в учебной деятельности всеобщего способа (принципа, закономерности) решения относительно широкого круга проблем и конкретно-практических задач. Рассмотренная в статье учебная задача соответствует приведённым выше определениям, поскольку требует от участников интеллектуального турнира усвоения нового понятия – «расстояния Левенштейна», вычисления этой метрики на примере и самостоятельного выполнения более сложного задания.

Для совершенствования методик и повышения качества ИТ- и инженерного образования, в частности на уровне среднего общего образования, необходимо систематически анализировать решения задач из олимпиад [2, 3], конкурсов и интеллектуальных турниров.

1. Учебное задание

Расстояние Левенштейна – это метрика, которая измеряет минимальное количество односимвольных операций (вставка, удаление, замена), необходимых для преобразования одной строки текста в другую. Чем меньше это число, тем больше строки похожи. Для идентичных строк оно равно нулю. Расстояние Левенштейна для строк «барокко» и «рококо» равно трём: для преобразования первого слова во второе требуется удаление двух первых символов и вставка символа «о».

На ЕГКР ученик решил выполнить следующее задание с использованием программного кода: «Исполнитель ИИ умеет выполнять следующие три операции: вычесть 1, вычесть 4, целочисленно разделить на 3. Сколько существует программ, преобразующих число 19 в число 2, если траектория вычислений включает число 14 и не включает число 8?»

Для выполнения вышеприведённого задания ученик написал программный код, содержащий сразу несколько ошибок (рисунок 1).

```
def f(x, y):
    if x < y and x == 8:
        return 0
    if x == y:
        return 1
    return f(x - 1, y) * f(x - 4, y) * f(x % 3, y)
print(f(19, 14) + f(14, 2))
```

Рис. 1. Программный код с ошибками, написанный учеником

Задание можно корректно выполнить с использованием следующего программного кода (рисунок 2).

```
def f(x, y):
    if x < y or x == 8:
        return 0
    if x == y:
        return 1
    return f(x - 1, y) + f(x - 4, y) + f(x // 3, y)
print(f(19, 14) * f(14, 2))
```

Рис. 2. Программный код для верного решения задания ЕГКР

Вычислите расстояние Левенштейна для двух приведённых листингов. В ответе укажите целое десятичное число – расстояние Левенштейна для двух листингов.

2. Решение учебного задания

Необходимо найти отличия в представленных листингах программ и подсчитать минимальное количество односимвольных операций для преобразования первого листинга во второй.

Действие № 1. Ключевое слово and заменено на ключевое слово or, для чего требуется выполнение трёх односимвольных операций: двух замен (символов «a» на «o» и «n» на «r») и одного удаления (удаляется буква «d»).

Действие № 2. Знаки операции «*» заменены на знаки «+» – требуется выполнение двух односимвольных замен.

Действие № 3. Оператор получения остатка от целочисленного деления «%» заменён на оператор «//», для чего символ «%» заменён на символ «/», после чего вставлен ещё один символ «/» – снова требуется выполнение двух односимвольных операций.

Действие № 4. Знак операции «+» в последней строке программного кода заменён на «*» – выполнена одна односимвольная замена.

Таким образом, общее количество односимвольных операций равняется 3 + 2 + 2 + 1 = 8.

Ответ: 8.

3. Динамическое программирование в учебной задаче

Классический алгоритм для вычисления расстояния Левенштейна использует динамическое программирование. Кроме того, в условии представленной учебной задачи содержится другое задание динамического программирования – и обучающимся при выполнении данной задачи предстоит внимательно прочитать и верно интерпретировать её постановку.

Программный код, представленный в листинге на рис. 2, верно решает включённое в задачу нахождения расстояния Левенштейна

задание динамического программирования – вычисление количества программ для заданных команд, чисел и траектории вычисления. С таким заданием обучающиеся, выбравшие в качестве дополнительного экзамена ЕГЭ по информатике, столкнутся в процессе подготовки и решения задания 23 ЕГЭ по информатике. Таким образом, в предложенной задаче не только осваивается понятие расстояния Левенштейна и проверяется умение его вычисления для двух заданных строк: обучающиеся также получают и актуализируют знания о способах решения других учебных заданий по теме динамического программирования.

Решение заданий на вычисление количества программ с использованием программного кода и аналитических методов описано в научных статьях [4, 5], объяснение принципов работы представленного рекурсивного программного кода содержится в статье «Рекурсивная функция для решения задания на динамическое программирование: от беззнаниевой формы компетенции к знанию и пониманию» [6]. Разнообразные способы вычисления количества различных результатов для заданных команд и длины программы представлены в статье [7].

Заключение

В статье впервые представлена учебная задача интеллектуального турнира «Кибер-бой» для обучающихся инженерных классов на вычисление расстояния Левенштейна и её аналитическое решение. В статье описаны внутрипредметные связи, заложенные в учебную задачу: её решение требует сравнения и анализа двух программных листингов для выполнения классического задания по динамическому программированию – вычисления количества программ.

Список литературы

1. Положение о проведении интеллектуального турнира «Кибер-бой». – М.: ИРПО ГАОУ ВО МГПУ, 2024. – 6 с.

2. Попов, В. С. Необходимые компетенции старшеклассников в области программирования для решения задания олимпиады школьников по криптографии / В. С. Попов // Математика, информатика, физика: проблемы и перспективы: Сборник научных статей международной научно-практической конференции, Оренбург, 24–25 апреля 2025 года. – Оренбург: Оренбургский государственный педагогический университет, 2025. – С. 425-432. – EDN: WIUGCC.

3. Попов, В. С. Развитие предпрофессиональных компетенций будущих специалистов в области информационной безопасности и различные методы решения комбинаторного задания отборочного этапа олимпиады школьников имени И.Я. Верченко по информатике 2024-2025

учебного года / В. С. Попов // Традиции и новации в профессиональной подготовке и деятельности: педагог, социальный работник, социолог, организатор работы с молодежью: Сборник научных трудов Международной научно-практической конференции, посвященной 80-летию Великой Победы, Году защитника Отечества, 155-летию Тверского государственного университета, Тверь, 28 марта 2025 года. – Тверь: Тверской государственный университет, 2025. – С. 562-567. – EDN: STMDPK.

4. Попов, В. С. Решение задания № 23 ЕГЭ по информатике на динамическое программирование: программный код на языке Python, использование обратных операций и проверка ответа / В. С. Попов, Е. А. Алефиренко // Потенциал. Журнал для старшеклассников и учителей. – 2024. – № 6. – С. 33-44.

5. Попов, В. С. Решение усложнённых заданий № 23 ЕГЭ по информатике на динамическое программирование, содержащих требования к командам / В. С. Попов // Потенциал. Журнал для старшеклассников и учителей. – 2025. – № 5. – С. 40-47.

6. Попов, В. С. Рекурсивная функция для решения задания на динамическое программирование: от беззнаниевой формы компетенции к знанию и пониманию / В. С. Попов, Е. А. Алефиренко, Л. Ю. Черницына // Информатика в школе. – 2024. – Т. 23, № 5. – С. 85-94. – EDN: MDGTEQ. – DOI: 10.32517/2221-1993-2024-23-5-85-94.

7. Попов, В. С. Компетенции для анализа результата исполнения алгоритма: задание вычисления количества различных результатов для заданных команд и длины программы в ЕГЭ по информатике / В. С. Попов, Е. А. Алефиренко, Л. Ю. Черницына // Современное педагогическое образование. – 2024. – № 10. – С. 62-68. – EDN: IQJPGM. – DOI: 10.24412/2587-8328-2024-10-62-68.

Модуль в Python для вычисления расстояния Левенштейна

Л. Ю. Черницына, email: chernicynalyu@mgpu.ru

Институт развития профильного обучения, Государственное автономное образовательное учреждение высшего образования города Москвы «Московский городской педагогический университет»

Аннотация. *В статье представлены возможности вычисления расстояния Левенштейна с помощью программирования в Python с использованием одноимённого модуля.*

Ключевые слова: *расстояние Левенштейна, Levenshtein distance, расстояние Дамерау-Левенштейна, редакционное расстояние, дистанция редактирования, инженерный класс, ИТ-класс, интеллектуальный турнир, Кибер-бой.*

Введение

Зачастую для исправления ошибок в словах и поиска дубликатов текстов, сравнения геномов и прочих полезных операций с символьными последовательностями используется расстояние Левенштейна, также известное как редакционное расстояние и дистанция редактирования. Оно является известной метрикой, характеризующей расстояние между двумя последовательностями символов. Расстояние Левенштейна равно минимальному количеству односимвольных операций (замена, вставка, удаление), выполняемых для преобразования исходной строки в другую. Данную метрику в 1965 году предложил советский математик Владимир Иосифович Левенштейн. Расстояние Левенштейна используется в таких научных областях как компьютерная лингвистика, теория информационных процессов и систем, биоинформатика.

Расширением понятия является расстояние Дамерау-Левенштейна – минимальное количество односимвольных операций, к которым добавлена операция транспозиции – перемены позиций двух соседних символов (аналогом может служить перемена мест двух соседних элементов массива при пузырьковой сортировке) – именно такая опечатка крайне часто встречается при наборе текстов.

Задание на вычисление расстояния Левенштейна содержалось в квалификационном этапе № 1 интеллектуального турнира «Кибер-бой» для обучающихся инженерных классов школ г. Москвы (см. статью «Задача вычисления расстояния Левенштейна в интеллектуальном

© Черницына Л. Ю., 2026

турнире Кибер-бой» [1]). Предполагалось преимущественно аналитическое выполнение данного задания обучающимися.

Федеральные рабочие программы по информатике подчёркивают необходимость реализации компетентностного подхода в школе [2]. Знания, умения, деятельностные установки для получения расстояния Левенштейна с использованием созданной программы могут быть рассмотрены как составляющие компетенций использования языков программирования для решения задач [3-5], в том числе при обработке текстовой информации [6], а также в качестве компонентов цифровых предпрофессиональных компетенций обучающихся [7-9].

1. Установка модуля Levenshtein

Для установки модуля Levenshtein в языке программирования Python следует выполнить команду в командной строке:

```
pip install levenshtein
```

2. Примеры использования модуля Levenshtein

Примеры использования функции distance модуля Levenshtein представлены на рисунке.

```
1  import Levenshtein
2
3  print(Levenshtein.distance('барокко', 'рококо')) # Вывод: 3
4
5  print(Levenshtein.distance('математика', 'информатика')) # Вывод: 5
6
7  first = '''def f(x, y):
8      if x < y and x == 8:
9          return 0
10     if x == y:
11         return 1
12     return f(x - 1, y) * f(x - 4, y) * f(x % 3, y)
13 print(f(19, 14) + f(14, 2))'''
14
15 second = '''def f(x, y):
16     if x < y or x == 8:
17         return 0
18     if x == y:
19         return 1
20     return f(x - 1, y) + f(x - 4, y) + f(x // 3, y)
21 print(f(19, 14) * f(14, 2))'''
22
23 print(Levenshtein.distance(first, second)) # Вывод: 8
```

Рисунок. Использование функции Levenshtein.distance для вычисления расстояния Левенштейна

В первом примере измеряется расстояние Левенштейна для слов «барокко» и «рококо». Данное расстояние равняется 3: для преобразования первого слова во второе требуется удаление двух первых

букв и вставка буквы «о», то есть три односимвольных операции. Во втором примере для преобразования слова «математика» в слово «информатика» потребовалось 5 односимвольных операций: добавление буквы «и» в начало слова и замена символов «мате» на «нфор». Наконец, для преобразования первой многострочной литеральной константы во вторую необходимо выполнение 8 односимвольных операций (см. статью Е. А. Алефиренко «Задача вычисления расстояния Левенштейна в интеллектуальном турнире Кибер-бой» [1]).

Заключение

В статье представлен пример использования модуля Levenshtein в языке программирования Python для вычисления расстояния Левенштейна с применением функции distance, приведён ряд вызовов данной функции и возвращаемых ею значений.

Список литературы

1. Алефиренко, Е. А. Задача вычисления расстояния Левенштейна в интеллектуальном турнире Кибер-бой // Информатика: проблемы, методы, технологии: сборник материалов XXVI международной научно-практической конференции. – Воронеж, Воронежский государственный университет, 11-12 февраля 2026 г. – Воронеж: ВГУ, 2026. [В печати]

2. Попов, В. С. Компетентностный подход в федеральных рабочих программах по информатике на уровне среднего общего образования / В. С. Попов, Е. А. Алефиренко, Л. Ю. Черницына // Текущие вызовы в подготовке кадров. Обучение специалистов по современным направлениям информационных технологий, кибербезопасности и ИКТ-электроники, актуальным для экономики данных: Сборник научных трудов, Тверь, 16–17 мая 2024 года. – Тверь: Тверской государственный университет, 2024. – С. 604-607. – EDN: BUWEMV.

3. Попов, В. С. Формирование компетенций старшеклассников в области программирования: решение задания 25 ЕГЭ по информатике с масками чисел / В. С. Попов // Информатика в школе. – 2025. – Т. 24, № 2. – С. 71-79. – DOI: 10.32517/2221-1993-2025-24-2-71-79. – EDN: ORKFAF.

4. Попов, В. С. Необходимые компетенции старшеклассников в области программирования для решения задания олимпиады школьников по криптографии / В. С. Попов // Математика, информатика, физика: проблемы и перспективы: Сборник научных статей международной научно-практической конференции, Оренбург, 24–25 апреля 2025 года. – Оренбург: Оренбургский государственный педагогический университет, 2025. – С. 425-432. – EDN: WIUGCC.

5. Попов, В. С. Компетенции в области программирования: семантические различия двойных неравенств на примере языков программирования Python и C++ / В. С. Попов, Е. А. Алефиренко, Л. Ю. Черницына // Проблемы гуманитарных наук и образования в современном мире: Сборник научных статей по материалам X Всероссийской (с международным участием) научно-практической конференции, Сибай, 22 ноября 2024 года. – Сибай: Уфимский университет науки и технологий, 2024. – С. 29-30. – EDN: XOOEMA.

6. Попов, В. С. Лексикографический порядок для сравнения строк: примеры сортировки на языках программирования и буква «Ё» / В. С. Попов, Е. А. Алефиренко, Л. Ю. Черницына // Мультилингвизм как социальный контекст языка: проблемы и перспективы: сборник научных статей по итогам работы IV Международной научной конференции, Армавир, 14–15 ноября 2024 года. – Киров: Межрегиональный центр инновационных технологий в образовании, 2024. – С. 291-298. – EDN: QOUVWF.

7. Попов, В. С. Европейская рамка цифровых компетенций педагога DigCompEdu: базовые компетенции, их оценка и неравномерность развития / В. С. Попов // Перспективные направления исследований в области наук об образовании: Сборник научных трудов I Международной научной конференции аспирантов, посвященной 155-летию Тверского государственного университета, Тверь, 04 апреля 2025 года. – Тверь: Тверской государственный университет, 2025. – С. 129-136. – EDN: PAEESK.

8. Попов, В. С. Формирование цифровых предпрофессиональных компетенций: примеры заданий для урока информатики по теме «Структуры данных» / В. С. Попов, Е. А. Алефиренко, Л. Ю. Черницына // Электронные средства и системы управления. Материалы докладов Международной научно-практической конференции. – 2024. – № 1-2. – С. 140-142. – EDN: OEROWO.

9. Попов, В. С. Пропедевтика структур данных в школьном курсе информатики как средство формирования цифровых предпрофессиональных компетенций: предметное содержание и требования / В. С. Попов, Е. А. Алефиренко, Л. Ю. Черницына // Электронные средства и системы управления. Материалы докладов Международной научно-практической конференции. – 2024. – № 1-2. – С. 143-146. – EDN: IBVFTH.

Секция 5
Электронное обучение и дистанционные образовательные технологии

Автоматизация учебного процесса. 1С:Автоматизированное составление расписания. «1С:Общеобразовательное учреждение»

О. В.Трунова, email: olga.trunowa2013@yandex.ru

Муниципальное общеобразовательное учреждение Калачеевская средняя общеобразовательная школа №6 имени Героя Советского Союза М. П. Краснолуцкого

Аннотация. *Данная статья посвящена подробному ознакомлению с функционалом и основными принципами работы программы «1С:Автоматизированное составление расписания. Школа». Она адресована администраторам образовательных учреждений, руководителям отделов, учителям-предметникам и другим специалистам, занимающимся организацией учебного процесса. Актуальна при организации дистанционного образования.*

Ключевые слова: *автоматизированное составление расписания, школьное расписание, управление образовательным учреждением, программное обеспечение для образовательных учреждений.*

Введение

Современная школа сталкивается с множеством вызовов, одним из которых является необходимость оптимизации административных процессов и повышение эффективности управления образовательным учреждением. Одной из таких задач является переход от традиционных бумажных журналов к электронным средствам фиксации успеваемости и другой важной информации [1-4].

В сегодняшнем мире информационных технологий электронные журналы становятся неотъемлемым элементом учебного процесса. Их внедрение направлено на устранение излишних бюрократических процедур, ускорение работы педагогов и администраторов, а также повышение прозрачности и объективности учета успеваемости учащихся.

Статья посвящена вопросам по автоматизированному составлению расписания в школе с использованием программы «1С:Общеобразовательное учреждение» [5].

© Трунова О. В., 2026

1. Автоматизированное составление расписания

Эффективное управление школой включает в себя множество аспектов, одним из важнейших является грамотное составление учебного расписания. Современные технологии позволяют существенно облегчить этот процесс благодаря специализированным программам автоматизации. Одной из наиболее популярных и удобных решений является программа «1С:Общеобразовательное учреждение». Она помогает оптимально распределить учебные часы, учесть особенности педагогов и учащихся, а также минимизировать риск возникновения конфликтных ситуаций.

Процесс формирования расписания делится на несколько ключевых этапов.

Этап 1. Подготовка исходных данных.

Первым делом необходимо собрать полную информацию обо всех педагогах, учениках, классах и объемах учебной нагрузки. Важно учитывать специфику предметов, продолжительность уроков и перерывов, режим работы школы и предпочтения педагогов относительно конкретных дней и временных интервалов.

Этап 2. Настройка программы.

Следующим этапом становится настройка самой программы «1С:Общеобразовательное учреждение»:

- определение режима работы школы,
- создание базового состава педагогического коллектива и учащихся,
- заполнение справочников предметов и аудиторий,
- ввод планового объема нагрузок для каждого педагога и класса.

Этап 3. Генерация расписания.

После ввода всей необходимой информации запускается модуль автоматической генерации расписания. Система учитывает введенные ранее ограничения и рекомендации, распределяя уроки оптимальным образом. При необходимости отдельные элементы расписания можно скорректировать вручную.

Этап 4. Проверка и коррекция.

Полученное расписание проверяется на возможные конфликты и несоответствия. Если обнаружены проблемы, система предложит варианты исправления либо допускает ручную правку отдельных элементов. После устранения возможных недостатков расписание считается готовым к публикации.

Этап 5. Распределение и публикация расписания.

Готовое расписание выводится в удобочитаемом виде и размещается на официальных ресурсах школы. Его можно распечатывать и распространять среди учащихся и педагогов в бумажном варианте.

2. Преимущества автоматизированного подхода

Использование программы «1С:Общеобразовательное учреждение» обеспечивает ряд преимуществ перед традиционными методами составления расписания:

- повышение точности планирования и исключение человеческих ошибок,
- значительное сокращение времени на формирование расписания,
- улучшение качества взаимодействия между учителями и учениками,
- снижение риска конфликтов и перегрузки педагогов.

Составить удобное и гибкое расписание занятий для учащихся – это сложный и трудоемкий процесс, при этом нужно учесть множество дополнительных условий, не допустив при этом накладок и ошибок. Данная программа составит расписания автоматически и позволит с легкостью его корректировать по ходу учебного процесса.

Программа позволяет:

– составлять расписание различной сложности в автоматическом, ручном и смешанном режимах; вести несколько сеток звонков;

– составлять расписание без привязки к сетке звонков, назначая для каждого занятия продолжительность и определяя произвольное время начала;

– учитывать время перемещения между учебными корпусами;

– выводить расписание на печать, по классам, преподавателям и помещениям, с возможностью быстрых настроек выводимой информации; копировать расписание прошлых периодов и корректировать его;

– учитывать требования СанПиН и ФГОС по последовательности проведения занятий, максимальной дневной нагрузке с учетом сложности занятий; вводить и учитывать сложность предметов/занятий/дисциплин в баллах;

– учитывать пожелания и возможностей учителей, классов учащихся, помещений;

– учитывать разбиение на подгруппы;

– составлять индивидуальные траектории для групп и отдельных учащихся;

– строить расписание для 1, 2-х и более смен;

– автоматически проверять расписание на ошибки, удобно их устранять;

– корректировать расписание с нужной периодичностью, сравнивать расписания;

– импортировать и экспортировать данные из «1С:Общеобразовательное учреждение»;

– формировать учебный план на основе готового шаблона базисного учебного плана, рекомендованного Министерством образования и науки РФ;

– составлять несколько расписаний и выбирать лучшее;

– подбирать и вести замены;

– вести учет аудиторного фонда образовательной организации;

– формировать отчеты об использовании помещений и проведенных занятиях.

Что входит в продукт:

– установочный диск с дистрибутивами:

– платформы «1С:Предприятие 8.3»;

– конфигурации «1С:Автоматизированное составление расписания. Школа»;

– комплект документации.

Заключение

Применение современных технологий и специализированных программ позволяет сделать работу образовательного учреждения более эффективной и комфортной для всех участников учебного процесса. Использование программы «1С:Общеобразовательное учреждение» способствует созданию оптимального учебного расписания, что положительно сказывается на качестве образования и удовлетворенности школьников и педагогов [6-9].

Актуальность автоматизированного составления расписания в школе с использованием программы «1С:Общеобразовательное учреждение» обусловлена рядом объективных факторов, характерных для современной образовательной среды.

Оптимизация учебного процесса: автоматика позволяет равномерно распределить нагрузку на преподавателей и учащихся, исключая возникновение перегрузок и конфликтов расписания.

Повышение эффективности администрирования: благодаря современным технологиям администрация школы получает инструмент для оперативного внесения изменений и своевременного реагирования на потребности образовательного процесса.

Экономия времени: процедура ручной компоновки расписания требует значительных усилий и времени, тогда как программное решение сокращает сроки составления графика и снижает риск ошибок.

Индивидуализация подхода: специальные алгоритмы помогают учитывать индивидуальные особенности и требования как педагогов, так и учащихся, создавая комфортную атмосферу для учебы и работы.

Соответствие требованиям законодательства: программа поддерживает учет нормативных документов и требований Министерства просвещения РФ, обеспечивая соответствие стандартам и нормативам.

Адаптация к современным реалиям: возможность гибкого управления расписанием важна в условиях дистанционных и смешанных форматов обучения. Таким образом, применение программы «1С:Общеобразовательное учреждение» для автоматизированного составления расписания позволяет повысить качество образовательного процесса, снизить административную нагрузку и обеспечить эффективное функционирование школы в современных условиях.

Список литературы

1. Гордеев, В. В. Управление образовательным учреждением: теория и практика / В. В. Гордеев. – Москва : Академия, 2019. – 320 с.
2. Кузнецов, А. А. Информационные технологии в образовании: практические рекомендации / А. А. Кузнецов, А. М. Шапиро. – СПб.: Изд-во СПГУ, 2020. – 280 с.
3. Семенов, А. Н. Электронные средства регистрации успеваемости в школе / А. Н. Семенов. – Новосибирск : Новосибирский университет, 2021. – 180 с.
4. Васильева, Л. П. Внедрение электронного журнала в образовательном учреждении: методическое пособие / Л. П. Васильева. – Екатеринбург : УралГУФК, 2022. – 160 с.
5. Официальный сайт компании «1С». – Режим доступа: http://www.1c.ru
6. Материалы портала «Российская школа XXI века». – Режим доступа: https://rschool21.ru
7. Харламов, А. Г. Современные технологии в управлении образованием / А. Г. Харламов. – Ростов-на-Дону : Феникс, 2022. – 240 с.
8. Левинсон, С. С. Электронный журнал как средство повышения эффективности образовательного процесса. – Волгоград : Волгоградский государственный университет, 2020. – 140 с.
9. Маслов, Н. Б. Информационные технологии в учебном процессе: реализация и сопровождение / Н. Б. Маслов. – Нижний Новгород : Нижегородский университет, 2021. – 220 с.

От цифрового архива к интеллектуальной среде: текущее состояние и перспективы развития виртуального музея SVM по истории информатики в Сибири

В. Н. Касьянов, email: kvn@iis.nsk.su
Т. А. Волянская, email: tanya@iis.nsk.su

Институт систем информатики им. А.П. Ершова СО РАН

Аннотация. *Представлены результаты комплексной модернизации созданного в ИСИ СО РАН виртуального музея SVM по истории информатики в Сибири. Ключевыми элементами модернизации стали создание современного адаптивного веб-интерфейса и разработка семантической онтологической модели данных. Данный подход обеспечивает переход от модели классического цифрового архива к интеллектуальной музейной среде, поддерживающей работу со связанными данными, гибкий контекстный поиск и персонализированный доступ. Рассматриваются текущее состояние проекта и стратегические перспективы его развития.*

Ключевые слова: *виртуальный музей, история информатики, адаптивный веб-интерфейс, онтологическая модель, семантическая интеграция данных.*

Введение

В настоящее время наряду с традиционными музеями широкое развитие получают цифровые, или виртуальные, музеи различной тематики, доступные в среде Интернет. Сайт «Музеи России» [1] содержит каталоги ссылок на многочисленные существующие виртуальные музеи в России и за рубежом.

Виртуальные музеи, являясь органической частью сети Интернет, своим присутствием в этой информационной среде открывают огромные возможности по развитию Интернета как среды по сохранению культурной истории, ее осмыслению и интерпретации, а также по формированию нового общественного сознания.

Музейные сайты реальных музеев могут, не нарушая безопасность хранящихся в них памятников культуры, дать возможность посетителям сайтов наслаждаться виртуальными представлениями своих культурных реликвий без ограничений на время и место. Ярким примером богатых возможностей, которые могут предоставляться посетителям

© Касьянов В. Н., Волянская Т. А., 2026

виртуального музея, является сайт Государственного Эрмитажа, одной из величайших сокровищниц мира [2].

Виртуальные музеи могут содержать не только виртуальные двойники реальных экспонатов, но и виртуальные экспонаты, которые являются мультимедийными цифровыми представлениями произвольных артефактов без каких-либо ограничений на их природу или текущее состояние. В частности, виртуальный экспонат может представлять уже исчезнувшею картину, самого художника, художественную школу или отдельное культурное событие.

Многие существующие виртуальные музеи не имеют в основе реальных музеев и содержат только виртуальные экспонаты. Среди них — широко известный Виртуальный компьютерный музей [3], посвященный истории вычислительной техники, прежде всего отечественной, а также создаваемый в Институте систем информатики СО РАН Сибирский виртуальный музей SVM [4-6], посвященный Сибирской школе информатики и программирования, основателем которой является академик Андрей Петрович Ершов [7-9].

Важное отличие музея SVM от традиционных (виртуальных и обычных) состоит в том, что авторы проекта с самого начало стремились сделать его адаптивным (настраиваемым на потребности конкретных пользователей) и открытым (предоставляющим широкому кругу пользователей удобные возможности по пополнению и развитию музея). Как открытый виртуальный музей SVM ориентируется не только на виртуальных посетителей, но и на виртуальных музейных работников [10-12]. Поэтому некоторые пользователи музея SVM могут не только пополнять его экспонатами и высказывать предложения и замечания, но и создавать свои авторские экскурсии и экспозиции.

Вместе с тем, используемая в имеющейся реализации SVM реляционная модель данных, основанная на PostgreSQL, ограничивала возможности представления контента, интеграции с внешними ресурсами и предоставления современных сервисов пользователям. Поэтому было решено осуществить модернизацию архитектуры SVM. В результате была проделана работа по двум взаимосвязанным направлениям: полному обновлению клиентской части (веб-интерфейса) и разработке новой семантической онтологической модели данных.

Эти изменения определяют текущее состояние SVM и формируют основу для его будущего развития в рамках концепции семантического веба и связанных открытых данных (Linked Open Data), что и является предметом рассмотрения в данной статье.

1. Современный адаптивный веб-интерфейс

Первым этапом модернизации стала полная переработка клиентской части – веб-интерфейса, целью которой было обеспечение безопасного, производительного и кроссплатформенного взаимодействия пользователей с контентом музея.

В рамках комплексной модернизации выполнена миграция веб-интерфейса SVM с устаревшей платформы PHP 4.0 на актуальную версию PHP 8.4 [13]. В результате достигаются следующие ключевые улучшения:

– повышенная безопасность, обеспечиваемая соответствием современным стандартам кибербезопасности и встроенными механизмами защиты;

– улучшенная стабильность и поддерживаемость за счет снижения количества ошибок, упрощения отладки и доступа к современным библиотекам;

– а также значительный прирост производительности благодаря оптимизации выполнения кода и сокращению времени обработки запросов.

Полностью переписаны и оптимизированы функциональные модули с учетом актуальных стандартов и синтаксических возможностей PHP 8.4. В частности, были переработаны модули авторизации и регистрации пользователей для внедрения современных протоколов аутентификации и безопасной обработки данных, а также заменена устаревшая обработка ошибок на систему исключений.

В результате проведенных работ архитектура и логика реализации основных модулей системы были модернизированы в соответствии с современными стандартами веб-разработки. Для повышения качества пользовательского взаимодействия был осуществлен переход на современные интерфейсные технологии, важнейшим из которых стал веб-стандарт HTML5 [14]. В результате были реализованы следующие возможности современного веба:

– семантическая разметка с использованием структурных блоков и смысловых тегов;

– расширенные возможности форм, включая специализированные типы полей, подсказки и встроенную валидацию;

– кросс-браузерная совместимость и адаптивный дизайн в полном соответствии со стандартами W3C.

Результатом данного перехода стала формальная подготовка интерфейсного слоя к интеграции с семантической онтологической моделью данных музея, что является критическим условием для

реализации концепции связанных открытых данных (Linked Open Data) и интеллектуального контекстного поиска.

Для обеспечения полноценной адаптивности интерфейса SVM решающее значение имели переход на CSS3 и разработка новой CSS-архитектуры. Архитектура построена на четырех взаимодополняющих принципах [15]:

– Медиазапросы (Media Queries) для условного применения стилей в зависимости от характеристик устройства, что обеспечивает корректное отображение на экранах любого разрешения.

– Современные системы раскладки: CSS Grid Layout – для создания сложных двумерных макетов, и Flexbox – для гибкого одномерного позиционирования компонентов. Их комбинация позволяет реализовать автоматическое перераспределение и масштабирование элементов.

– Адаптивная типографика, основанная на относительных единицах измерения, что гарантирует динамическое масштабирование текста и сохранение его читаемости на всех типах устройств.

– CSS-переменные (Custom Properties) для централизованного управления значениями стилей (цветовая палитра, размеры, отступы), что повышает поддерживаемость и целостность визуального дизайна.

Для оптимизации визуального контента все растровые графические элементы были переведены в современный формат WebP. Переход на данный формат позволил получить ряд преимуществ:

– Эффективное сжатие – уменьшение размера файлов на 25–35% по сравнению с JPEG при сохранении визуального качества, что напрямую ведет к ускорению загрузки страниц и экономии трафика.

– Расширенная функциональность – поддержка прозрачности (как в PNG) и анимации (как в GIF) в рамках одного формата.

– Широкая совместимость – поддержка более чем 98% современных браузеров, что делает решение практически универсальным.

Для полной адаптивности графики реализован комплекс CSS-методик: автоматическая подстройка размеров и разрешения изображений под характеристики экрана, использование srcset для выбора оптимальной версии файла и прогрессивных методов загрузки. В результате графический контент музея обеспечивает высокое качество отображения и скорость работы на любых устройствах.

Таким образом, результатом первого этапа модернизации стало создание современного, производительного и кроссплатформенного веб-интерфейса SVM, который решает ключевую задачу обеспечения доступности музейного контента для широкой аудитории на любых типах

устройств. Достигнутые техническая надежность и адаптивность создают стабильную и масштабируемую основу для интеграции интерфейса со вторым системообразующим компонентом модернизации: семантической онтологической моделью данных.

2. Семантическая онтологическая модель данных

Вторым и концептуально наиболее значимым компонентом модернизации является разработка семантической онтологической модели данных музея SVM. Данная модель призвана преодолеть принципиальные ограничения реляционной базы данных, обеспечив формальное представление знаний о предметной области истории информатики в Сибири для последующей семантической обработки и интеграции.

Существующая реляционная модель данных SVM, основанная на PostgreSQL, исчерпала свой потенциал для задач семантической интеграции. Ее жесткая схема затрудняла описание сложных полиморфных связей между сущностями (например, «ученый – участник – проект – публикация»), а отсутствие семантики делало невозможным контекстный поиск и интеграцию с внешними ресурсами. Это обусловило необходимость перехода к онтологической модели.

Разработка онтологии была выполнена в соответствии с общепринятой методологией семантического веба, что гарантирует интероперабельность и долгосрочную жизнеспособность модели. В качестве технологической основы были выбраны ключевые стандарты консорциума W3C [16-18]:

– RDF (Resource Description Framework) – фундаментальная графовая модель данных, обеспечивающая универсальное представление знаний в виде триплетов «субъект-предикат-объект».

– OWL 2 (Web Ontology Language) – выразительный язык описания онтологий, позволяющий определять сложные классы, свойства и выполнять логический вывод, что критически важно для адекватного моделирования связей в музейной среде.

– SPARQL (SPARQL Protocol and RDF Query Language) – мощный язык запросов, предназначенный для работы с графовыми данными RDF и закладывающий основу для реализации интеллектуального поиска в музее.

Для практической реализации был использован комплекс инструментов, включающий редактор Protégé, наиболее распространенный инструмент для инженерии онтологий, и Apache Jena Fuseki, предоставляющий надежный сервер для хранения данных и обработки SPARQL-запросов с открытой лицензией [19].

Для обеспечения гибкости, масштабируемости и четкого разделения концептуальных уровней была разработана трехуровневая онтологическая модель, включающая: ядро музейной онтологии (Museum Core Ontology), предметную онтологию (Domain Ontology) и онтологию управления доступом (Access Control Ontology).

Ядро музейной онтологии (Museum Core Ontology) формализует абстрактные, независимые от предметной области, универсальные концепции любого виртуального музея: Экспонат (Exhibit), Экспозиция (Exhibition), Экскурсия (Excursion), Зал (Hall). Экспонат – минимальная структурная единица музея, атомарный объект представления информации. Это центральный абстрактный класс, все конкретные типы экспонатов являются его подклассами. Экспозиция – тематически организованная совокупность экспонатов. Экскурсия – временная последовательность демонстрации экспонатов, объединенная общим нарративом. Зал – абстрактное понятие виртуального пространства для размещения и категоризации экспонатов.

Ядро определяет обобщенную иерархию типов залов, абстрагированную от предметной специфики и основанную на критерии доступности: Открытый зал (OpenHall) – пространство, доступное для всех посетителей; Зал-запасник (StorageHall) – пространство с ограниченным доступом, предназначенное для хранения и работы специалистов; Зал новых поступлений (NewArrivalsHall) – промежуточное пространство для временного размещения новых или непроверенных объектов. Такое абстрактное разделение обеспечивает концептуальную чистоту ядра и позволяет предметной онтологии на его основе создавать специализированные залы, связывая их через логические ограничения с конкретными типами экспонатов.

Предметная онтология (Domain Ontology) наполняет ядро конкретикой истории информатики. Наследуясь от абстрактного Экспоната, она определяет специализированные классы: Ученый (Scientist), Коллектив (Collective), Публикация (Publication), Документ архива (ArchivalDocument), Проект (Project), Событие (Event), Конференция (Conference), Вычислительная техника (ComputingEquipment). Каждый класс характеризуется уникальным набором атрибутов (Datatype Properties) и участвует в специфических объектных связях (Object Properties), что позволяет установить связи между различными типами объектов и отразить семантику предметной области.

На основе абстрактных классов залов создаются специализированные классы залов-запасников: Зал ученых (ScientistsHall), Зал коллективов (CollectivesHall), Библиотека

(LibraryHall), Архив (ArchiveHall), Зал проектов (ProjectsHall), Хроника событий (ChronicleHall), Зал конференций (ConferencesHall), Зал вычислительной техники (EquipmentHall), а также Зал подготовки экспозиций и экскурсий (PreparationHall). Строгие OWL-ограничения гарантируют, что в каждом зале могут находиться только экспонаты определенного типа (например, в Библиотеке – только Публикации), обеспечивая формальную целостность данных.

Онтология управления доступом (Access Control Ontology) определяет ролевую модель и системы прав доступа. Определена иерархия ролей от Посетителя (Visitor – незарегистрированный пользователь) до Директора (Director – полный доступ), включая специализированные должности: Волонтер (Volunteer), Экскурсовод (ExcursionGuide), Экспозитор (ExpositionCreator), Простой специалист (BasicSpecialist), Музейный работник (MuseumWorker), Биограф (Biographer), Коллективовед (CollectiveExpert), Библиотекарь (Librarian), Архивариус (Archivist), Хронолог (Chronicler), Секретарь (Secretary), Инженер (Engineer), а также Администраторы ресурсов (ResourceAdministrator). Модель реализует трехуровневую иерархию прав доступа: Посетители (Visitor) имеют доступ только к открытым залам; Специалисты (Specialist) – включают простых специалистов (BasicSpecialist: Volunteer, ExcursionGuide, ExpositionCreator), ответственных за создание контента, и музейных работников (MuseumWorker: Librarian, Archivist, и др.), обладающих правами редактирования; Администраторы (Administrator: ResourceAdministrator, Director) имеют полные права управления системой и ее компонентами. Для каждой роли прописаны права доступа к конкретным типам экспонатов и залов, реализуя детальную систему разграничения полномочий для сотрудников музея.

Для обеспечения интероперабельности и соответствия принципам Linked Open Data (LOD) модель будет связана с общепринятыми (стандартными) внешними онтологиями (Dublin Core, FOAF, CIDOC CRM) [20]. Это позволит однозначно идентифицировать сущности (например, связывать ученого с его профилем в VIAF или Wikidata) и открывает путь к интеграции SVM в глобальное семантическое пространство.

Разработанная онтологическая модель осуществляет принципиальный сдвиг парадигмы в работе с данными SVM: от изолированных таблиц — к связному графу знаний (Knowledge Graph). Такой подход обеспечивает следующие функциональные возможности:

– семантический (SPARQL) поиск, оперирующий смыслом контента, а не совпадением ключевых слов;

– навигация по ассоциативным связям между сущностями, позволяющая исследовать сложные нетривиальные отношения;

– автоматизированный логический вывод для получения новых, неочевидных знаний;

– семантическая интеграция с внешними ресурсами в рамках Linked Open Data.

Заключение

Реализация описанных этапов модернизации позволяет вывести виртуальный музей истории информатики в Сибири (SVM) на качественно новый уровень технологического развития.

Текущее состояние проекта характеризуется:

– наличием современного, безопасного и адаптивного пользовательского интерфейса;

– спроектированной и формализованной трехуровневой онтологической моделью данных, готовой к наполнению контентом и интеграции.

Ближайшие и стратегические перспективы развития включают:

– наполнение онтологической модели и миграцию существующих данных в RDF-представление;

– реализацию интеллектуального семантического поиска через интерфейс SPARQL-запросов, позволяющего исследователям находить сложные корреляции;

– интеграцию по стандартам Linked Open Data (LOD), установление связей с внешними авторитетными источниками (VIAF, Wikidata) для обогащения контекста и повышения видимости музея в академической среде;

– развитие персональных сервисов на основе онтологии: создание адаптивных экскурсионных маршрутов, тематических подборок и систем рекомендаций на основе интересов пользователя и логических связей между объектами;

– разработку семантически аннотированного API для взаимодействия со сторонними исследовательскими проектами и образовательными платформами.

Таким образом, комплексная модернизация заложила технологический фундамент для эволюции SVM от традиционного цифрового каталога к полноценной интеллектуальной информационно-исследовательской среде, способствующей сохранению, изучению и популяризации истории информатики в Сибири.

Список литературы

1. Сайт «Музеи России» [Электронный ресурс]. – Режим доступа: http://www.museum.ru/
2. Сайт «Государственный Эрмитаж» [Электронный ресурс]. – Режим доступа: http://www.hermitage.ru/
3. Виртуальный компьютерный музей [Электронный ресурс]. – Режим доступа: http://www.computer-museum.ru/
4. Касьянов, В. Н. Виртуальный музей истории информатики в Сибири / В. Н. Касьянов, Г. П. Несговорова, Т. А. Волянская // Проблемы программирования. – 2003. – № 4. – С. 82-91.
5. Kasyanov, V. N. SVM – Siberian Virtual Museum of Informatics History / V. N. Kasyanov // Innovation and the Knowledge Economy: Issues, Applications, Case Studies. – Amsterdam: IOS Press, 2005. – Part 2. – P.1014-1021.
6. Kasyanov, V. N. An open adaptive virtual museum of informatics history in Siberia / V. N. Kasyanov // IFIP International Federation for Information Processing. – New York: Springer, 2008. – Vol. 269. – P. 129-146.
7. Специальный выпуск памяти академика Андрея Петровича Ершова // Программирование. – 1990. – № 1. – 156 с.
8. Bjørner, D. Images of Programming. Dedicated to the Memory of A.P. Ershov / D. Bjørner, V. Kotov. – Amsterdam: North-Holland, 1991. – 190 p.
9. Касьянов, В. Н. Российское программирование в лицах: мои учителя / В. Н. Касьянов // Проблемы информатики. – 2014. – № 2. – С. 74-85.
10. Касьянов, В. Н. Музеи и Интернет: новые возможности / В. Н. Касьянов // Информационные технологии в гуманитарных исследованиях. – Новосибирск : Новосиб. гос. ун-т, 2006. – Выпуск 10. – С. 88- 96.
11. Касьянов, В. Н. Музеи и Интернет / В. Н. Касьянов // Современные информационные технологии и письменное наследие: от древних текстов к электронным библиотекам : материалы Международ. науч. конф., Казань, 26-30 авг. 2008 г. / Отв. ред.: В. Д. Соловьёв, В. А. Баранов. – Казань : Казан. гос. ун-т, 2008. – С. 132-135.
12. Kasyanov, V. N. An open adaptive virtual museum of information history in Siberia / V. N. Kasyanov // IFIP Advances in Information and Communication Technology. – Berlin: Springer, 2011. – Vol. 357. – P. 194-200.
13. PHP: Руководство по языку [Электронный ресурс]. – Режим доступа: https://www.php.net/manual/ru/

14. HTML Living Standard [Электронный ресурс]. – Режим доступа : https://html.spec.whatwg.org/multipage/

15. CSS Snapshot [Электронный ресурс]. – Режим доступа : https://www.w3.org/TR/CSS/

16. RDF 1.1 Concepts and Abstract Syntax [Электронный ресурс]. – Режим доступа : https://www.w3.org/TR/rdf11-concepts/

17. OWL 2 Web Ontology Language Document Overview [Электронный ресурс]. – Режим доступа : https://www.w3.org/TR/owl2-overview/

18. SPARQL 1.1 Overview [Электронный ресурс]. – Режим доступа : https://www.w3.org/TR/sparql11-overview/

19. Noy, N. F., McGuinness D. L. Ontology Development 101: A Guide to Creating Your First Ontology [Электронный ресурс]. – Режим доступа : https://protege.stanford.edu/publications/ontology_development/ ontology101.pdf

20. Концептуальная эталонная модель CIDOC CRM / пер. с англ. А. Б. Антопольского, Д. Ю. Гук. – Москва : ИНИОН РАН, 2024. – 400 с.

Цифровой помощник учителя для проведения уроков

Л. Н. Копалиани, email: ramaz-kopaliani@yandex.ru

Муниципальное бюджетное общеобразовательное учреждение
средняя общеобразовательная школа № 98 г. Воронежа

Аннотация. *Образовательный онлайн-сервис «Облако знаний» содержит цифровые уроки по 17 предметам. Разнообразие теоретического материала и практических заданий оказывают помощь учителю в объяснении нового материала, в проведении контрольных работ, а также помогают готовить обучающихся к ЕГЭ, ОГЭ и олимпиадам различного уровня.*

Ключевые слова: *онлайн-сервис «Облако знаний», универсальная библиотека ЦОК, цифровой образовательный контент, цифровой урок.*

Введение

Мир меняется, нас окружают скорость, технологии, огромные ресурсы. Время изучения новых возможностей обучения и коммуникации с обучающимися пришло неожиданно, но своевременно. Выпускники наших школ должны обладать умением ориентироваться быстро в разных неопределенных ситуациях, быть всегда готовым к самообучению, самостоятельно осуществлять поиск необходимой информации, проводить её анализ, формулировать цели и задачи, затем оценивать достигнутые результаты, осваивать инновационные технологии, быть лидером при достижении одних задач и исполнителем при выполнении других, уметь работать в команде.

Современное образование выдвигает высокие требования к уровню развития обучающихся, поэтому педагог стремится повысить активность и самостоятельность своих учеников, получить высокое качество знаний по изучаемому предмету, научить понимать абстрактные теоретические положения. Хорошим инструментом, который обеспечивает увлеченную и результативную работы обучающихся, является цифровой урок, который позволяет логически последовательно осваивать учебный материал и в том числе использовать специально разработанные средства оценки достижения образовательных результатов в соответствии с рабочей программой.

© Копалиани Л. Н., 2026

Цифровой помощник учителя для проведения уроков

Таким новым инструментом для меня и моих учеников стал образовательный онлайн-сервис «Облако знаний» [1, 2], который содержит цифровые уроки по 17 предметам. Теоретический материал и практические задания, которые представлены на платформе «Облако знаний», помогают мне в объяснении нового материала, в проведении контрольных, самостоятельных, проверочных работ, интегрировать межпредметные связи на примере реальных жизненных ситуаций, назначать задания и анализировать выполнение с помощью цифрового следа, а также этот онлайн-сервис оказывает помощь в подготовке обучающихся к ЕГЭ, ОГЭ и олимпиадам различного уровня.

Весь контент, представленный на «Облаке знаний», разработан методистами, соответствует федеральному перечню ЭОР, соответствует требованиям ФГОС и содержанию федеральных рабочих программ. Этот материал нацелен не только на приобретение знаний по предметам, но и на способность эти знания применять в реальных жизненных ситуациях, то есть эти задания позволяют формировать все типы учебных действий по ФГОС.

Использование цифрового контента данного онлайн-сервиса для фронтальной работы позволяет мне разнообразить по содержанию занятия по математике, даёт возможность безболезненно демонстрировать наглядный материал онлайн, помогает сэкономить время на назначение цифровых домашних работ, которые проверяются в автоматическом режиме. На данный момент нет возможности отправки отметок в электронный журнал, но в скором будущем такая функция на платформе появится.

На «Облаке знаний» доступна проверка индивидуальных заданий с развёрнутым ответом. Развернутые ответы в заданиях ЕГЭ и ОГЭ есть возможность проверить с помощью искусственного интеллекта, а также учитель может проверить ответы самостоятельно или согласиться с результатами проверки нейросетью.

«Облако знаний» помогает учителю планировать с обучающимися персональную образовательную траекторию, ставить нужные образовательные цели, задачи, изменять уровень сложности работы и назначать её учащемуся в зависимости от его подготовки, также есть возможность в анализе результатов учеников по конкретной работе, возможность провести статистику всего класса по предмету, а также статистику, позволяющую определить пробелы в знаниях обучающегося и универсальные действия, которые хуже всего поддаются школьнику.

На платформе есть функционал, с помощью которого учитель может узнать всё о сильных и слабых сторонах учеников, также доступна

матрица предметных дефицитов для отслеживания проблемных тем и умений обучающихся.

Цифровой тренажер онлайн-сервиса «Облако знаний» выходит за рамки привычных тренажеров по подготовке к ЕГЭ, ОГЭ, так как он направлен на развитие функциональных компетенций обучающихся. Цель данного тренажера научить учащихся эффективно применять полученные в школе математическую грамотность, читательские навыки и естественнонаучные представления в повседневной жизни.

Контент «Облака знаний» представлен на базовом и углублённом уровнях, что позволяет восполнить дефициты содержания рабочих программ по предметам.

Содержание учебного предмета «Математика» на данном онлайн-сервисе представлено так, что сначала располагаются все материалы по учебному курсу алгебра, затем по геометрии и вероятности и статистике. Это помогает быстро найти нужный материал по курсу и позволяет наглядно увидеть план изучения.

Платформа «Облако знаний» вошла в число ведущих поставщиков цифровых уроков в Универсальной библиотеке ЦОК в рамках ФГИС «Моя школа», которая является единой платформой для доступа к верифицированным ЭОР, прошедшим экспертизу.

Таким образом, платформа «Облако знаний» позволяет мне уйти от таких рутинных процессов как планирование и проверка выполненного материала, так как на платформе они автоматизируются, что помогает экономить учебное время. Благодаря автоматизированной проверке большей части заданий, существует возможность осуществлять оперативную обратную связь по результатам выполнения отдельных заданий, то есть цифровая платформа помогает ученику самостоятельно учиться, а учителю учить, выполняя различные роли: методиста, тьютора, психолога.

Платформа «Облако знаний» способствует формированию образовательной среды, которая ориентирована на развитие каждого обучающегося в соответствии с выбранной персональной траекторией, направленной на развитие навыков 21 века.

Заключение

«Облако знаний» обеспечивает учеников всеми образовательными ресурсами для обучения на одном едином онлайн-сервисе. Обучение с помощью цифровых уроков позволяет мне развивать интеллектуальные способности обучающихся, помогает видеть каждого ученика и строить обучение под его потребности, а также развивать ответственность, самостоятельность, предметно-пространственную среду образовательной организации, которая ориентирована на развитие

каждого обучающегося в соответствии с выбранной персональной траекторией.

Цифровые уроки «Облака знаний», на мой взгляд, являются фундаментом, который поможет обучающимся не просто изучить материал, а по-настоящему понять логику цифрового мира и чувствовать себя в нём уверенно и безопасно.

Если педагог стремится повысить активность и самостоятельность обучающихся, получить высокое качество знаний по изучаемому учебному курсу, научить обучающихся понимать абстрактные теоретические положения, то онлайн-сервис «Облако знаний» для него станет одним из часто используемых инструментов.

Список литературы

1. Козленко, А. Г. Облако знаний. Школа: Методическое пособие / А. Г. Козленко, Д. И. Мамонтов. – Долгопрудный: ФИЗИКОН, 2016. – 133 с.

2. Козленко, А.Г. Облако знаний. Школа: Методическое пособие / А. Г. Козленко, Д. И. Мамонтов. – Долгопрудный: ФИЗИКОН, 2018. – 111 с.

Секция 6
Переход на профессиональные стандарты: требования к IT-компетенциям

Клавиатурный почерк как вид поведенческой биометрии обучающихся для развития цифровых компетенций

С. А. Шульгиненко, email: superkran9@gmail.com

МГТУ им. Н. Э. Баумана

Аннотация. *В статье рассматривается клавиатурный почерк как вид поведенческой биометрии и анализируется возможность его интерпретации в контексте формирования цифровых компетенций обучающихся. Показано, что в условиях цифровизации образования возрастает значение осознанного цифрового поведения и понимания принципов идентификации пользователя в информационных системах. Клавиатурный почерк предлагается рассматривать не только как технический инструмент информационной безопасности, но и как компонент цифровой грамотности, связанный с развитием метапредметных навыков. Обоснована целесообразность включения данной тематики в содержание школьного курса информатики.*

Ключевые слова: *цифровые компетенции, клавиатурный почерк, поведенческая биометрия, цифровая грамотность, информационная безопасность, метапредметные навыки, информатика.*

Введение

Современный этап развития системы образования характеризуется активным внедрением цифровых и информационных технологий в учебный процесс. Использование электронных образовательных платформ, дистанционных форм обучения и цифровых сервисов требует формирования у обучающихся устойчивых цифровых компетенций, обеспечивающих эффективную и безопасную деятельность в цифровой среде [1, с. 474-476; 2, с. 1648-1652, 1654-1656]. Как отмечают Е. С. Полат и М. Ю. Бухаркина, цифровые технологии в образовании должны рассматриваться не только как инструмент обучения, но и как средство формирования универсальных и метапредметных навыков обучающихся [3].

Одним из важных аспектов цифровой грамотности и компетентности является понимание принципов функционирования информационных систем, включая способы идентификации и аутентификации пользователей. В этой связи особый интерес представляет использование биометрических методов, в том числе поведенческой биометрии, к которой относится клавиатурный почерк.

© Шульгиненко С. А., 2026

1. Понятие и классификация методов идентификации пользователя

Идентификация пользователя в информационных системах представляет собой процесс установления его личности на основе определённых признаков. В современных системах используются различные методы идентификации, которые условно можно разделить на методы, основанные на знании, владении и биометрических характеристиках пользователя.

Биометрические методы идентификации получили широкое распространение благодаря высокой степени надёжности и сложности подделки. В фундаментальном труде Handbook of Biometrics подчёркивается, что биометрия включает как физиологические, так и поведенческие характеристики человека, каждая из которых может использоваться для автоматизированного распознавания личности [4].

Поведенческая биометрия анализирует особенности взаимодействия пользователя с техническими средствами. В отличие от физиологических характеристик, поведенческие признаки формируются в процессе обучения и практики, что делает их особенно интересными с точки зрения образовательного контекста.



Рис. 1. Классификация методов идентификации пользователя

2. Клавиатурный почерк как вид поведенческой биометрии

Клавиатурный почерк представляет собой совокупность устойчивых временных характеристик ввода текста пользователем, таких как длительность удержания клавиш и интервалы между нажатиями. В научной литературе данный подход известен как keystroke dynamics и

рассматривается как один из перспективных методов биометрической аутентификации.

В классической работе Ф. Монроуза и А. Рубина показано, что ритм набора текста обладает индивидуальными и относительно устойчивыми характеристиками, позволяющими использовать его для идентификации пользователя [5]. Авторы подчёркивают, что клавиатурный почерк является ненавязчивым и практически бесплатным с точки зрения технической реализации методом биометрической идентификации, поскольку не требует дополнительного оборудования.

В отличие от статических биометрических характеристик, клавиатурный почерк изменяется в процессе обучения, что делает его показателем уровня владения клавиатурой и цифровыми навыками в целом.

3. Клавиатурный почерк в структуре цифровых компетенций обучающихся

Цифровые компетенции обучающихся включают цифровую грамотность, информационную безопасность и метапредметные навыки, а также все компетенции для эффективного использования современных цифровых технологий. По мнению А. В. Хуторского, метапредметный подход ориентирован на формирование универсальных способов деятельности, которые выходят за рамки отдельных учебных дисциплин [6].

Клавиатурный почерк может быть соотнесён с несколькими компонентами цифровых компетенций. В рамках цифровой грамотности он отражает уровень владения инструментами цифровой среды и формирование цифровой идентичности. С точки зрения информационной безопасности клавиатурный почерк иллюстрирует современные подходы к защите доступа к информационным системам. В метапредметном аспекте изучение данной тематики способствует развитию аналитического и критического мышления, а также осознанию взаимосвязи между технологией и поведением пользователя.



Рис. 2. Место клавиатурного почерка в структуре цифровых компетенций

Заключение

Клавиатурный почерк, являясь видом поведенческой биометрии, может быть полезен для формирования цифровых компетенций обучающихся. Его рассмотрение в образовательном процессе позволяет расширить представления обучающихся о принципах работы информационных систем и методах идентификации пользователя.

Использование данной тематики в школьном курсе информатики способствует формированию осознанного цифрового поведения, развитию цифровой грамотности и метапредметных навыков, что соответствует современным требованиям к результатам образования.

Список литературы

1. Попов, В. С. Трансформация ключевых компетенций модели Key competences for lifelong learning / В. С. Попов, Л. А. Абросимова-Романова // Текущие вызовы в подготовке кадров. Обучение специалистов по современным направлениям информационных технологий, кибербезопасности и ИКТ-электроники, актуальным для экономики данных: Сборник научных трудов, Тверь, 16–17 мая 2024 года. – Тверь: Тверской государственный университет, 2024. – С. 474-477. EDN: QUFFBF.

2. Попов, В. С. Ключевые компетенции для обучения в течение всей жизни в контексте компетенций и компетентностей в области информационных технологий / В. С. Попов, Л. А. Абросимова-Романова // Информатика: проблемы, методы, технологии: Материалы XXIV Международной научно-практической конференции им. Э.К. Алгазинова,

Воронеж, 14–15 февраля 2024 года. – Воронеж: Воронежский государственный университет, 2024. – С. 1648-1657. EDN: LMNJAT.

3. Полат, Е. С. Современные педагогические и информационные технологии в системе образования: учебное пособие для студ. высш. учеб. заведений / Е. С. Полат, М. Ю. Бухаркина. – 3-е изд., стер. М.: Издательский центр «Академия», 2010. – 368 с. – ISBN 978-5-7695-7057-5.

4. Jain, A. K. Handbook of Biometrics / A. K. Jain, P. Flynn, A. A. Ross (eds.). – New York: Springer Science+Business Media, 2008. – ISBN 978-0-387-71040-2.

5. Monrose, F. Keystroke dynamics as a biometric for authentication / F. Monrose, A. Rubin // Future Generation Compute Systems. – 2000. – Vol. 16, No. 4. – P. 351-359. – DOI: 10.1016/S0167-739X(99)00059-X.

6. Хуторской, А. В. Метапредметный подход в обучении: Научно-методическое пособие / А. В. Хуторской. – М.: Издательство «Эйдос»; Издательство Института образования человека, 2012. – 73 с. – ISBN 978-5-904329-18-1.

Содержание

Пленарные доклады

Зацаринный А. А. Об актуальности фундаментальных научных исследований в области искусственного интеллекта 4

Переселков С. А. Математические модели голографической обработки векторно-скалярных сигналов в гидроакустических волноводах 19

Секция 1. Методы и технологии разработки программных систем

Волченко П. В., Самойлов Н. К. Преимущества микрофронтенд архитектуры для разработки веб-приложений 41

Воронина И. Е., Полякова П. Г. Сравнительный анализ подходов к сбору и формализации требований к программному обеспечению 50

Дрокин М. Р., Михайлюк Е. А. Разработка инструмента для автоматической генерации тестов REST API на основе OpenAPI и семантических сценариев 56

Дылевский А. В., Хрипушин Д. А. Идентификация интегро-дифференцирующего звена 63

Занин Д. С., Михайлюк Е. А. Разработка бэкенд-сервиса для анализа и ранжирования смежных информационных технологий 68

Кислый К. А., Ефремов М. С. Анализ применимости различных форматов хранения аудиоданных при разработке чата 74

Молоков И. Ю., Михайлюк Е. А. Теоретические и методологические основы интеграции AI-ассистентов в процесс промышленной разработки ПО 83

Савинков А. Ю. Исследование временны́х характеристик выполнения задач реального времени в Linux 92

Ушаков В. А. Применение трансформерной модели BERT для автоматического парсинга библиографических ссылок 110

Секция 2. Компьютерное моделирование в фундаментальных и прикладных исследованиях.

Балакин В. С., Шипко Ю. В., Колычев О. В. Программный комплекс автоматизированного описания природно-климатической зоны в заданном районе.....117

Барсукова А. А., Козырь А. В. Адаптивная сегментация траектории и навигация на основе векторного поля.....126

Беспалов С. В. Исследование влияния атмосферы на пропускную способность инфракрасного излучения.....135

Борзунов С. В., Кургалин С. Д., Гончарова Н. Ю., Батищева Г. А., Асеева А. А., Петрищев К. О. Программа контроля антибиотикорезистентности патогенной микрофлоры в медицинских организациях.....140

Бутовецкая С. И., Храпов И. В., Каверина В. К. Автоматизация расчетов при изучении 7-мерных алгебр Ли.....150

Буховец А. Г., Бирючинская Т. Я., Горностаев А. К. Геометрический подход к интерпретации теоремы Островского.....158

Веремьёв Н. А., Маслобойщиков А. Н., Голунов А. С., Потапов Н. С. Алгоритм объединения разнородных метеорологических данных в единое «поле потенциальной опасности» для конвективных явлений.....164

Гедзенко Д. В., Бодрых Д. В. Модель метеорологического обеспечения авиационных задач на основе элементов линейной регрессии.....169

Кирносов С. Л., Бадыгин А. Э., Долгов Е. А. Модель поддержки принятия решений при выполнении авиационных задач с использованием элементов теории детерминированного хаоса.....175

Климкин И. В., Лахно Г. А., Уткин Н. С. Анализ уязвимостей протокола Wi Fi 6 (IEEE 802.11ax) к пассивному перехвату и методы защиты на канальном уровне.....180

Козирацкий Ю. Л., Паринов М. Л., Петренков С. В. Имитационная модель процесса управления диаграммами направленности антенн пеленгаторов при угломерно-корреляционном местоопределении источников электромагнитного излучения.....186

Козирацкий Ю. Л., Прохоров Д. В., Баркалов А. А. Компьютерная модель(экспресс-оценки точности процесса сопровождения целей телевизионной системой, основанного на определении энергетического центра формируемого составного изображения.....196

Костылева Л. Н., Гордеев М. С. Прогностическое моделирование уровня загрязнения приземного слоя атмосферного воздуха 204

Лакунов И. С., Алексеев В. В., Ионченко Е. С. Иммерсивное моделирование процедурного обучения на базе Unity: тенденции и практическое применение 209

Мартинез Э. Н., Ножкин В. С., Шербаков Г. Е. Модели тепловых потоков в тропических волнах.....217

Мельников А. В., Назарчук С. В., Кузнецов Д. Н. Анализ перспектив применения фторполимерных покрытий в противообледенительных системах малоразмерных беспилотных летательных аппаратов 222

Петров В. А. Модель фоноцелевой и помеховой обстановки 229

Попов В. В., Митюшин В. М. Метод прогноза слоистой облачности по территории Республики Беларусь с использованием гидродинамических показателей 235

Потапов А. Н., Кузнецов В. А. Системный анализ обоснования эксплуатационно-технических характеристик беспилотных летательных аппаратов средней и большой дальности при управлении техническим состоянием 242

Потапов А. Н., Кузнецов В. А. Системный анализ принятия решений при смешанном техническом обслуживании беспилотных летательных аппаратов средней и большой дальности 249

Разиньков С. Н., Борисов Д. Н., Богословский А. В. Синтез антенных решеток с широкими секторами снижения боковых лепестков при минимизации направлений контроля диаграмм направленности.....254

Соловьев А. В., Бакаев Г. Н., Нененко Б. В. Нейросетевое моделирование процесса образования радиационных туманов 261

Туркменова Р. Т. Обзор лагранжевых моделей переноса частиц (LPDM) и их вычислительных особенностей 268

Туровский Я. А., Крупенин С. С. Структура информационной системы сбора и обработки данных для систем обратной связи в миоэлектрических протезах верхней конечности 278

Ульшин И. И., Семенов П. Д. Прогнозирование количественных характеристик лесных пожаров на основе нейросетевых технологий 284

Феофилов Д. С., Феофилов С. В., Фомичева О. А. Синтез программного оптимального управления нелинейным объектом на основе DQN-агента 291

Холина А. А., Хапкин Д. Л. Анализ нейросетевых алгоритмов управления неустойчивыми динамическими объектами 300

Шабанов А. Р., Бочаров А. С. Компьютерное моделирование авиационного электроэнергетического комплекса 308

Шпинев А. В. Алгоритм повышения стойкости микросхем в условиях космоса 314

Шпинев А. В., Потапов А. Н. Анализ состояния вопроса современных компьютерных моделей для повышения радиационной стойкости микросхем в условиях космоса 322

Щегольков С. В., Рыжов А. И. Разработка учебно-тренировочных средств на платформе Unity для подготовки специалистов к работе со сложными техническими системами с использованием технологий виртуальной реальности 332

Секция 3. Технологии обработки и защиты информации

Будников С. А., Енютин Ю. А., Аксютин В. М. Исследование устойчивости функционирования объекта критической информационной инфраструктуры 340

Васильев С. В. Особенности выявления движения в приграничных областях кадра видеоряда 350

Веремеев В. Е., Ермаков М. В. Уязвимости машинного обучения: классификация и защита от атак отравления данных в моделях машинного обучения 355

Евстафиев А. Ф., Евстафиев Ф. А., Вышлов О. С. Алгоритм обнаружения импульсных радиосигналов в условиях частотно-селективных искажений и флюктуационных помех 363

Иванов К. А., Иванков А. Ю. Синтез и анализ алгоритмов неэталонной оценки качества изображений на основе глубокого машинного обучения 370

Иванова Д. Н., Иванков А. Ю. Семантический анализ текста с учетом диалогового контекста 379

Каладзе В. А., Шаповалов Н. А. Моделирование угловых характеристик вектора направления на объект по его двумерной проекции 384

Куликов С. С. Сравнительная оценка подходов к автоматической сегментации и радиомическому анализу КТ грудной клетки 390

Небогина Д. Р., Иванков А. Ю. Сравнительный анализ алгоритмов обучения с подкреплением для компьютерных игр 399

Паламарь И. Н., Гагарина А. И. Алгоритм синтеза бимодальной свертки с кросс-доменным согласованием для генерации профиля поверхности 406

Перова С. В., Акимов А. В., Гречишкин А. В. Классификация неисправностей пассивных радиочастотных устройств при помощи нейронных сетей с использованием корней комплексного коэффициента передачи 416

Редин Н. А., Гаршина В. В. Разработка системы анализа уязвимостей на основе баз знаний MITRE ATT&CK и MITRE D3FEND с использованием RAG технологий 425

Рогозин Р. Е. К вопросу анализа современных механизмов внимания для задач восстановления изображений 433

Родионов Р. И., Сирота А. А. Анализ алгоритмов обучения с подкреплением для управления авиационными и беспилотными системами 441

Руднев А. Е., Дрюченко М. А. Алгоритм мультиплексирования скрытых каналов при реализации стеганографического встраивания данных в JPEG изображения 448

Ступникова Д. А., Горностаев А. Д., Русанов А. В. Практические аспекты обеспечения информационной безопасности при разработке проектов для программируемых логических контроллеров в средах CODESYS 3.5 и Astra IDE 457

Храмов В. Ю., Стешковой А. С., Щеглов А. В. Модель машинного обучения обнаружения компьютерных атак в беспроводных сетях стандарта IEEE 802.11 466

Шатков П. В., Гаршина В. В. Автоматизированная система обезличивания персональных данных в медицинских текстовых документах на основе NLP с веб-интерфейсом 474

Шишкин С. А., Швырева А. В. Анализ методов повышения полноты извлечения релевантной информации в системах RAG 481

Ялвач С. М., Каладзе В. А. Микросервисная реализация остановочного узла субавтономной ИТС для прогнозирования времени прибытия общественного транспортного средства 493

Секция 4. Информационные системы и базы данных

Борисова А. А. Сравнительный анализ SQL и DL-Query при обработке данных юридической предметной области 500

Варданян Н. А. Внедрение BI-системы в малый бизнес: принципы, архитектура и практические рекомендации 512

Гатауллина А. А. Архитектура данных корпоративных систем кадрового электронного документооборота 517

Гордеев С. Ю. Корпоративные информационные системы нового поколения: как микросервисы и гибридные базы данных ломают монолит 525

Косыгин В. А. Принципы построения и архитектура единой информационной платформы для цифровой трансформации организаций в сфере недвижимости 533

Скибин Д. Д. Сравнительный анализ подходов к сбору и формализации требований к программному обеспечению 541

Тали Д. И. Обеспечение связности компонентов электронных документов на основе бинарного дерева Меркла 546

Секция 5. Прикладное моделирование и E-business

Бригадина А. И., Копытина Е. А. Разработка системы для автоматизации ключевых процессов дополнительного образования 555

Великодный В. А., Сирота Е. А. Реализация нейронечеткой модели для цифровой обработки изображений 561

Зацаринный А. А., Меденников В. И. Об одном подходе применения технологий общего назначения в рамках проекта «Экономика данных» 570

Иванов Д. А., Алексеев В. В. Результаты вычислительного эксперимента по оценке эффективности применения модели управлении зонами покрытия в спутниковой IoT-инфраструктуре 579

Илларионов И. В., Гришаев Н. О. Система хранения и организации значимой для управления изменениями в SAP S/4 HANA корпоративной информации 588

Илларионов И. В., Каплий В. В. Извлечение транспортных расходов из общего объема выручки в SAP S/4 HANA 599

Илларионов И. В., Крупчак Д. Р. Разработка алгоритма автоматизированного определения приоритета заявок технической поддержки на основе метода анализа иерархий (МАИ) в системе SAP Service Desk 613

Лавлинская С. С., Астахова И. Ф., Копытина Е. А. Разработка системы автоматизированного сопровождения психологического консультирования на основе IOS-приложения с функцией аналитики 622

Лоскутова А. Е., Алейникова Н. А. Математические модели анализа качества образовательного процесса на основе цепей Маркова 630

Матвеев М. Г., Свиридова Е. Н., Богданова Н. С. Сравнительный анализ решения задачи нечеткой регрессии с использованием классических подходов и W-алгебры 640

Меньших В. В., Гришаева В. Р. Подходы к восстановлению информации с использованием методов нечисловой статистики 650

Меньших В. В., Никитенко А. Н. Задачи распределения ресурсов элементов мультиагентных систем в условиях «мягкого» управления 658

Петренко М. С., Сирота Е. А. Модифицированный алгоритм кумулятивных сумм (Adaptive CUSUM) для обнаружения разладок в временных рядах метеоданных 666

Полумеев М. С., Копытина Е. А. Создание сервиса автоматизации поиска работы и стажировок для студентов 675

Попова П. Н., Копытина Е. А. Разработка веб-приложения для создания индивидуальных планов развития сотрудников IT-компаний.....680

Секция 6. Интеллектуальные информационные системы, компьютерная лингвистика, технологии информационного поиска

Абрамян А. В. Разработка и внедрение предметно-ориентированного GPT-ассистента для поддержки математического обучения.....687
Алейников С. А., Карлова Е. Н. Создание виртуальных ассистентов закрытого контура и их применение в образовательной, научной, и повседневной деятельности военного вуза 697
Бабкина Е. А., Гаршина В. В. Разработка системы выявления галлюцинаций в ответах больших языковых моделей 707
Барков А. А. Применение мультимодальной большой языковой модели для первичной аналитической интерпретации визуальных материалов, публикуемых хакерскими группировками.....714
Борисенко В. И., Пименов В. И. Верифицируемая система поддержки принятия решений на основе больших языковых моделей.....727
Гришкин М. А. Применение методов машинного обучения для прогнозирования негативных факторов в системах распределения энергии.....737
Калушев К. А. Структура системы управления коллаборативным роботом-манипулятором для свободного манипулирования объектами внешней среды 744
Крыжановская Ю. А. Применение машинного обучения для классификации кибератак 755
Кутепов Н. С., Ткаченко К. Н. Анализ современных моделей синтеза речи в задачах межъязыковой адаптации 761
Ландовский В. В., Гриф М. Г., Королькова О. О. Разработка системы хранения и подготовки данных для машинного перевода русского жестового языка 768
Михин Д. А., Толстошеин Н. С. Принципы построения онтологической модели организационно-технической системы для задачи мониторинга состояния её элементов на основе семантического анализа.....775
Пашигорев К. И., Головской В. А. Формирование информационного критерия классификации входных данных для большой языковой модели.....782
Полицына Е. В., Полицын С. А. Использование программных инструментов автоматического анализа текстов на русском языке на базе

алгоритмических методов и инструментов на основе нейросетевых технологий для решения производственных и научных задач 792
Скопин Е. С. Использование статических анализаторов в генерации программного кода на основе больших языковых моделей 803
Смотрова К. В., Гаршина В. В. Возможности применения технологии Retrieval Augmented Generation в системах поддержки принятия решений при разработке информационных систем 808
Солодовникова Я. В., Гаршина В. В. Использование алгоритмов обучения с подкреплением для мультиагентных систем в задачах управления складом 814
Солопанов Н. Д., Яковлева М. В., Санько В. А. Сравнительная характеристика методов машинного обучения в задачах управления БПЛА в рамках эргатической системы 821
Сычев А. В. Автоматическая оценка релевантности научных статей тематикам конференции с помощью большой языковой модели 828
Сычев А. В., Дымов А. И. Модель педагогических измерений на основе мультимодальных генеративных технологий ИИ 840
Толстошеин Н. С., Михин Д. А. Применение искусственного интеллекта в системе поддержки принятия решений нефтехимических производств 854
Тухтасинов М. Т. Моделирование расписания занятий на основе генетического алгоритма в системе высшего образования 859

Секция 7. Суперкомпьютерные вычисления и BigData. Квантовые информационные системы

Бозоров Р. М., Клинских А. Ф. Квантовые алгоритмы поиска в машинном обучении: проблемы, методы и технологии 870

Ивонин Д. А., Клинских А. Ф. Экспериментальное сравнение квантовых классификаторов для бинарной и мультиклассовой задач ..875

Коваль А. С. Разработка загрузочных образов для СнК.....882

Кургалин С. Д., Борзунов С. В. Квантовые и параллельные вычисления: креативный подход 888

Новиковский К. В., Преображенский А. П. Метод FROST для распределённой оптимизации с отслеживанием градиента на ориентированных графах 893

Потапов А. Г. Особенности использования сопроцессоров Intel Xeon Phi в обучении студентов параллельному программированию 902

Хаустов И. С., Клинских А. Ф. Библиотеки и инструменты квантового ИИ для Quantum PCA (qPCA) 907

Секция 8. САПР и цифровые технологии

Акимов Н. В., Сокольников В. В. Стратегическое проектирование.....919

Бакланова В. С., Караваева И. А., Рыжков В. А. Анализ возможностей Siemens NX 10 и Компас-3D для автоматизации проектирования оснастки.....927

Гурьева А. М., Троценко А. С. Разработка пользовательского интерфейса для построения задачи расчета стрежневой системы 933

Деханова В. С., Сокольников В. В., Дибров М. Г. Автоматизированная система принятия решений в процессах в аграрной промышленности.....939

Кириченко Д. Ю., Капустин И. А., Сокольников В. В. ИИ как инструмент автоматизации типовых чертежей в CAD-системах 947

Котолевский В. И., Троценко А. С. Архитектура подсистемы параметрического моделирования специализированной САПР на примере ступенчатого вала 951

Лесных А. Е., Петренко И. В., Чижов М. И. Кастомизация программы трехмерного моделирования NX.....958

Поднебесная К. Д., Сокольников В. В. Параллельное проектирование.....963

Приладышева Д. Д. Цифровая трансформация аналитики в управлении строительными проектами: современные тенденции и перспективы 968

Ремнева С. А., Бондаренко Н. А., Филимонова А. А. Разработка системы визуализации позиционирования заготовок в роботизированной производственной ячейке 979

Секция 9. Инфокоммуникационные системы, системы телекоммуникаций, радиоэлектроника, обработка сигналов

Богословский А. В., Борисов Д. Н., Разиньков С. Н. Электродинамический анализ сверхширокополосной конической щелевой антенны для мобильного радиоэлектронного комплекса 989

Борисов Д. Н., Семечев Д. А. Оценка точности пеленгационной локализации источника радиоизлучения при отсутствии прямой видимости 996

Величкина А. С., Кононов А. А., Усков Г. К. Численное моделирование переходных процессов в импульсных диодах методом конечных разностей 1007

Жигулин В. А., Пешков И. В. Изучение модифицированного протокола канального уровня для организации связи в беспроводных сетях с направленными антеннами 1017

Калюкин М. А. Оптимизация пропускной способности MIMO-канала в условиях высокоскоростного движения на основе методов адаптивного разнесения и предварительного кодирования 1027

Клименко Ю. А., Преображенский А. П., Кораблев Я. Б., Родюков А. И. Разработка алгоритма акустического мониторинга воздушного пространства 1035

Клименко Ю. А., Преображенский А. П., Кораблев Я. Б., Родюков А. И. Разработка алгоритма идентификации БПЛА системой акустической локации 1044

Климкин И. В., Лахно Г. А., Уткин Н. С. Анализ уязвимостей протокола Wi Fi 6 (IEEE 802.11ax) к пассивному перехвату и методы защиты на канальном уровне 180

Кононов А. А., Усков Г. К. Расчет тангенциальных компонент поля в апертуре антенны по результатам зондирования ближней зоны 1059

Плеве В. В., Шешко А. В. Концептуальная модель совместной обработки видео и радиолокационных данных для обнаружения БПЛА в различных условиях 1068

Разиньков С. Н., Разинькова О. Э., Савченко А. В. Электродинамический анализ решеток электрических вибраторов для радиоэлектронных комплексов со сниженным уровнем вторичного электромагнитного излучения 1075

Разинькова О. Э. Анализ передачи-приема сигналов в радиолиниях с зеркальными параболическими антеннами 1080

Смирнов Д. В., Донцов А. А. Анализ среднего квадратического отклонения погрешностей определения дальности пуска средства поражения при двухпозиционном пеленговании 1086

Чёнгин А. В., Козирацкий А. Ю. Алгоритм определения параметров отраженного зондирующего лазерного излучения от винтов беспилотного летательного аппарата малого класса 1096

Секция 10. Информационные технологии, вычислительный анализ и методы математической физики

Башкарев В. А., Пересёлков С. А., Кузькин В. М., Пересёлков А. С. Компьютерный эксперимент по фильтрации в волноводе на основе акустической голографии.....1104

Башкарев В. А., Пересёлков С. А., Кузькин В. М., Пересёлков А. С. Расчёт углового распределения голограммы источника в гидроакустическом волноводе.....1114

Кравчук Д. А., Шаповалов А. Д. Исследование ультразвуковых фокусированных преобразователей (для неразрушающего контроля) 1124

Лагута М. В., Кравчук Д. А. Анализ процессов нелинейного взаимодействия акустических волн с биотканями для оценки их структурных изменений 1128

Матвиенко Ю. В., Переселков С. А., Косенко И. М., Ладыкин Н. В. Численный расчет мультинаправленных векторно-скалярных пеленгаторов.....1133

Переселков С. А., Кузькин В. М., Косенко И. М., Ладыкин Н. В. Алгоритм пеленгования на основе разнесенных векторно-скалярных приемников 1141

Переселков С. А., Кузькин В. М., Косенко И. М., Ладыкин Н. В. Алгоритм формирования голограммы источника ЛЧМ-сигнала на основе дробного преобразования Фурье 1150

Переселков С. А., Кузькин В. М., Косенко И. М., Ладыкин Н. В. Компьютерный эксперимент по оценке дальности обнаружения источника звука в нерегулярном волноводе 1160

Переселков С. А., Кузькин В. М., Переселков А. С. Алгоритм для разделения волн акустического каротажа.....1168

Переселков С. А., Кузькин В. М., Переселков А. С., Попова Л. Н. Компьютерный эксперимент по оценке глубины источника волн в волноводе.....75

Старченко И. Б., Белоусова В. С., Кравчук Д. А. Предобработка 2D-снимков для повышения качества анализа нейросетями 1180

ШКОЛА-КОНФЕРЕНЦИЯ «ИНФОРМАТИКА В ОБРАЗОВАНИИ»

Секция 1. Методика преподавания информатики

Кулаков Ю. В. Разработка алгоритма решения олимпиадной задачи по информатике «Достройте дороги» на основе понятия эйлерова графа 1192

Черницына Л. Ю. О логарифмической линейке замолвим мы слово, или зачем изучать историю вычислительной техники на уроках информатики 1202

Секция 2. ЕГЭ по информатике

Попов В. С. Компетенции в области динамического программирования: аналитический и программный методы решения заданий 23 ЕГЭ по информатике 1210

Секция 3. Искусственный интеллект в образовании

Беляев А. Е., Беляева М. М. Искусственный интеллект в образовательном процессе высшей школы 1216

Кириллов В. В. Мультиагентная система для автоматизированной генерации учебных заданий 1226

Сычев А. В. Инструменты генеративного ИИ в академической деятельности вуза: взгляд студентов 1235

Секция 4. Подготовка ИТ-инженеров для решения задач цифровизации промышленности и бизнеса

Алефиренко Е. А. Задача вычисления расстояния Левенштейна в интеллектуальном турнире Кибер-бой 1253

Черницына Л. Ю. Модуль в Python для вычисления расстояния Левенштейна 1258

Секция 5. Электронное обучение и дистанционные образовательные технологии

Автоматизация учебного процесса. 1С:Автоматизированное составление расписания. «1С:Общеобразовательное учреждение» ... 1263

Касьянов В. Н., Волянская Т. А. От цифрового архива к интеллектуальной среде: текущее состояние и перспективы развития виртуального музея SVM по истории информатики в Сибири 1268

Копалиани Л. Н. Цифровой помощник учителя для проведения уроков..... 1278

Секция 6. Переход на профессиональные стандарты: требования к IT-компетенциям

Шульгиненко С. А. Клавиатурный почерк как вид поведенческой биометрии обучающихся для развития цифровых компетенций 1283

Научное издание

Информатика: проблемы, методы, технологии

Материалы
XXVI Международной научно-практической конференции

Воронеж,
12-13 февраля 2026 г.

Под общей редакцией Д. Н. Борисова

Издано в авторской редакции

Редакционная коллегия

Главный редактор:

Борисов Дмитрий Николаевич

Члены редакционной коллегии:

1 секция	Махортов Сергей Дмитриевич Нужных Алексей Васильевич
2 секция	Семенов Михаил Евгеньевич Толкачев Аким Владимирович
3 секция	Сирота Александр Анатольевич Швырева Анна Владимировна
4 секция	Ермаков Михаил Викторович
5 секция	Матвеев Михаил Григорьевич Копытин Алексей Вячеславович
6 секция	Гаршина Вероника Викторовна Сычев Александр Васильевич
7 секция	Кургалин Сергей Дмитриевич Сынкова Александра Евгеньевна
8 секция	Чижов Михаил Иванович
9 секция	Усков Григорий Константинович Корчагин Юрий Эдуардович Сынкова Александра Евгеньевна
10 секция	Пересёлков Сергей Алексеевич Косенко Иван Михайлович
Школа-конференция	
1-4,6 секции	Крыловецкий Александр Абрамович Атанов Артем Викторович
5 секция	Толстобров Александр Павлович