

# Разработка новой функциональности для информационной системы поиска людей по интересам

Н. В. Оверчук, email: nkovrchk@gmail.com<sup>1</sup>

ФГБОУ ВО «Московский Авиационный Институт (Национальный Исследовательский Университет)»

***Аннотация.** В работе проведен анализ функциональных возможностей прототипа информационной системы для поиска людей по интересам, описаны существующие проблемы и недостатки решения и предложены идеи для его усовершенствования.*

***Ключевые слова:** веб-приложение, алгоритм построения взаимосвязей, распознавание изображений, анализ данных в социальных сетях, поиск по ключевым словам.*

## Введение

В современном мире внедрение информационных технологий позволило совершить огромный шаг вперед в развитии общества. С появлением сети Интернет люди получили возможность взаимодействовать на огромных расстояниях друг от друга с помощью персональных компьютеров и мобильных устройств. Для реализации этой возможности появилось множество сервисов и площадок, в частности социальные сети, имеющие множество инструментов для взаимодействия пользователей.

Пытаясь познакомиться с новым человеком в социальной сети, пользователь хотел бы иметь представление об этом человеке: чем он увлекается, какие у него хобби, что его интересует и т. д. Анализ подобной информации позволил бы людям заранее для себя решить, совпадают ли интересы этого человека с интересами пользователя или нет.

Данную проблему частично решают существующие социальные сети и приложения, предназначенные для знакомства с новыми людьми, но данные решения также обладают недостатками, а именно: наличие недостоверной и непроверяемой информации о пользователях, отсутствие или ограниченный набор функциональных возможностей для поиска людей и отсутствие в большинстве сервисов тематической информации (публикаций, аудио и видео материалов, сообществ), которая объединяет пользователей с одинаковыми интересами.

Для решения описанной проблемы был разработан прототип информационной системы (ИС), в которой пользователь может искать других людей по интересам, используя анализ открытых данных и в социальных сетях, который подтвердил возможность улучшения результатов решения задачи поиска людей по интересам.

В настоящий момент ИС обладает только функцией поиска по ключевым словам, но уже положительно зарекомендовала себя в выполнении данной задачи. На основании отзывов пользователей системы были предложены варианты функциональных возможностей, которые необходимо внедрить в систему, чтобы улучшить пользовательский опыт работы и расширить существующие возможности поиска людей по интересам.

### **1. Анализ функциональных возможностей текущего прототипа**

Информационная система «FriendFinder» представляет из себя веб-приложение с клиент-серверной архитектурой, доступ к которому осуществляется через веб-браузер.

На главной странице сервиса пользователю предлагается выполнить авторизацию в социальных сетях для дальнейшей работы. Данное решение избавляет человека от регистрации в сервисе и позволяет использовать существующие данные в социальных сетях, в которых пользователь зарегистрирован. Также это позволяет системе использовать программный интерфейс (далее «API»), предоставляемый социальной сетью, в которой выполнена авторизация.

После этой процедуры пользователь может ввести ключевые слова, по которым система будет выполнять поиск, а также параметры выстраиваемого графа — ширину и глубину. После выбора параметров начинается построение графа потенциальных друзей, представляющий из себя дерево, где корнем является сам пользователь, а узлами — профили социальных сетей, которые находятся среди друзей профилей других вершин графа. В вершинах графа расположены структуры данных, содержащие информацию о профиле анализируемого пользователя и ссылки на связанные вершины. Ширина графа определяет количество детей для каждого узла, а глубина — количество уровней графа. Граф строится до тех пор, пока его размер не будет соответствовать размеру, указанному пользователем на странице приложения.

В процессе построения учитывается вес найденной вершины, основанный на указанных приоритетах поиска. Перед началом построения графа программа предлагает ввести приоритет для ключевых слов, который должен учитываться во время анализа данных — от очень высокого до низкого. Например, если для пользователя

приоритетным интересом является робототехника, а наименее приоритетным футбол, то система при обработке данных будет учитывать в первую очередь информацию, связанную с робототехникой и в последнюю очередь с футболом.

Помимо этого, учитывается уровень вершины, расположенной от корня дерева, так как в реальной жизни о ближайших друзьях или знакомых друзей известно больше разнообразной информации, чем о незнакомых людях, хоть и у которых могут совпадать интересы с пользователем. По этой причине, приоритет при поиске будет отдаваться сначала вершинам графа, имеющих наименьшее количество связей с корнем и дальше по нисходящему приоритету остальным вершинам, расположенным на большем расстоянии.

После построения графа начинается процесс обработки полученных данных путем сопоставления информации в профиле с указанными ключевыми словами [5]. Так, система поддерживает возможность перевода полученного текста из профиля человека с помощью API Яндекса [1] для его дальнейшей обработки. Помимо этого, с помощью компьютерного зрения система определяет тематику опубликованного на странице пользователя изображения и сопоставляет её с введенными интересами.

Однако, объем анализируемых данных пользователя может быть увеличен, чтобы более точно определять интересы человека и сопоставлять их с ключевыми словами. Кроме того, набор функций системы мог бы быть расширен для улучшения упрощения работы пользователя или решения задачи поиска более оптимальным способом. Например, сервис не позволяет сохранять информацию о построенных графах для дальнейшей работы с ними. Из-за этого система при использовании одним и тем же человеком вынуждена выстраивать цепочку из людей каждый раз заново, расходуя ресурсы системы и тратя на это большое количество времени.

Помимо этого, в системе отсутствует визуализация построенного графа, чтобы пользователю демонстрировались результаты работы приложения, и он мог понимать, через какую цепочку людей он связан с найденным человеком.

Также на данный момент в системе анализируются только текстовые и фото публикации пользователя, что не дает возможность полноценно сформировать представление об его интересах.

Для устранения вышеописанных проблем, необходимо решить следующие задачи:

1. Реализовать хранение результатов работы из предыдущих сессий пользователя и визуализацию этих результатов для

улучшения пользовательского опыта взаимодействия с системой.

2. Учитывать при построении графа связанных людей предыдущие результаты работы программы с одним и тем же пользователем для ускорения работы системы и улучшения её эффективности.
3. Реализовать сбор и обработку информации из сообществ, в которых состоит пользователь и люди, связанные с пользователем с учетом тематики этих сообществ и публикуемых материалов. Данное решение увеличит набор анализируемых данных и позволит с большей точностью определять интересы человека для дальнейшего сопоставления их по ключевым словам.

## **2. Описание предлагаемых функциональных возможностей системы поиска**

Проанализировав существующую функциональность в системе, предлагается внести изменения в существующие функциональные возможности с целью их улучшения и повышения эффективности работы программы.

Для улучшения поиска необходимо добавить в анализируемую структуру данных информацию о сообществах, на которые подписан человек [4]. Социальные сети как правило позволяют явно указать тематику публикуемого материала и, используя API сервиса [1] можно получить информацию об этой тематике и использовать её для анализа профиля. В данном случае необходимо ограничиться только проверкой тематики сообщества [2], а не всех людей, состоящих в нём, так как обработка такого огромного количества связей для одного из множества людей является задачей, требующей огромных вычислительных мощностей, что негативно сказывается на эффективности работы системы.

Чтобы оптимизировать процесс построения графа взаимосвязанных людей, предлагается сохранять информацию о предыдущем результате работы программы с конкретным пользователем и учитывать её при последующем построении. Так, например, если человек после работы приложения добавил или удалил одно ключевое слово, изменил ширину или глубину графа, то необходимо изначально использовать предыдущий построенный граф и дополнять или уменьшать его в зависимости от указанных параметров. Благодаря этому решению можно оптимизировать алгоритм построения графа, что уменьшит время на выполнение этой задачи.

Помимо совершенствования существующих функций системы, предлагается реализовать новые функциональные возможности, которые позволят решать задачу поиска людей по интересам другими способами или улучшат взаимодействие с системой.

Необходимо внедрить в приложение возможность регистрации и авторизации пользователя для закрепления за ним результатов. В профиле сервиса должна быть доступна возможность просмотра предыдущих результатов поиска в социальных сетях, чтобы система не запрашивала его каждый раз у пользователя.

Также необходимо добавить возможность поиска людей по имени и фамилии среди друзей, используя ключевые слова. Таким образом можно найти конкретного человека, совпадающего по интересам, а не случайных людей в социальной сети. Для решения этой задачи необходимо анализировать цепочку людей в графе на предмет наличия друга с указанными фамилией и именем и, в случае совпадения, проверять интересы этого человека, сопоставляя с указанными ключевыми словами.

В качестве результатов работы программы будет отображаться цепочка людей, связывающих пользователя с этим человеком в окне браузера, что повышает информативность и наглядность результатов.

### **Заключение**

Предлагаемые функциональные решения для совершенствования существующего прототипа улучшают взаимодействие пользователя с системой, расширяя возможности информационной системы и уменьшая время на выполнение работы программы. Отображаемый в окне браузера выстроенный граф взаимосвязей демонстрирует наглядность результатов работы и дает возможность проследить цепочку людей, связывающих пользователя с найденным человеком. Анализ сообществ, в которых состоит человек в социальной сети, позволяет дополнить структуру анализируемых данных и повысить вероятность того, что человек совпадает по указанным интересам.

### **Список литературы**

1. API Переводчика [Электронный ресурс]. — документация. — режим доступа: <https://yandex.ru/dev/translate/> (дата обращения 07.12.2021)
2. API ВКонтакте. Получение списка сообществ пользователя [Электронный ресурс]. — документация. — режим доступа: <https://vk.com/dev/groups.get> (дата обращения: 08.12.2021)

3. API ВКонтакте. Получение ключа доступа [Электронный ресурс]. — документация. — режим доступа: [https://vk.com/dev/access\\_token](https://vk.com/dev/access_token) (дата обращения 08.12.2021)

4. Фомина Ю. В. Извлечение информации из социальных сетей [Текст] // Стратегические коммуникации в современном мире. - Красноярск, 2014.

5. Коршунов, А. Анализ социальных сетей: методы и приложения [Электронный ресурс] / А. Коршунов, И. Белобородов, Н. Бузун и др. // Труды ИСП РАН. 2014. №-1. URL: <http://cyberleninka.ru/article/n/analiz-sotsialnyh-setey-metody-i-prilozheniya> (дата обращения 13.12.2021).