

Разработка системы поддержки принятия решений для прогнозирования оттока клиентов на базе АП Loginot

Н. Ю. Прокопенко, email: prokopenko_nu@mail.ru

Нижегородский государственный архитектурно-строительный

***Аннотация.** Показана необходимость применения специализированного программного обеспечения и современных методов анализа данных для прогнозирования оттока клиентов в телекоммуникационных компаниях.*

***Ключевые слова:** системы поддержки принятия решений, аналитическая платформа Loginot, хранилище данных, Data Mining, клиентская аналитика, прогнозирование.*

Введение

Развитие реальной конкуренции на телекоммуникационном рынке приводит к необходимости внедрять наиболее прогрессивные методы управления, изыскивать внутренние резервы, выявлять конкурентные преимущества. Специфика телекоммуникационной отрасли заключается в обслуживании многочисленной абонентской базы, нужно анализировать трафик клиентов, структурировать входящий поток обращений, формировать различную отчетность. При современной ситуации на рынке телекоммуникационных услуг, использование CRM-стратегии для прогнозирования оттока клиентов (их лояльности) является необходимым условием успешной деятельности компании, предоставляющей услуги связи.

Для модернизации или расширения инфраструктуры, а также для создания и продвижения новых услуг от оператора требуется внедрение эффективных схем удержания существующих и привлечения новых абонентов, и один из путей достижения этих целей заключается в извлечении всей полезной информации из имеющихся у телекоммуникационных фирм данных об абонентах и применении систем бизнес-аналитики и современных интеллектуальных методов анализа данных Data Mining, Big Data, Machine Learning и других технологий анализа данных [1].

Тенденции последних лет в развитии аналитических систем заключаются в интеграции средств аналитической обработки, алгоритмов извлечения знаний, управления метаданными и

визуализации результатов на одной программной аналитической платформе. Эти аналитические системы созданы, чтобы помочь в планировании и управлении различными бизнес-процессами и в решении некоторых бизнес-задач. К подобным информационно-аналитическим инструментам можно отнести аналитическую платформу Loginom (разработка компании «ООО Аналитические технологии Loginom Company»), реализующая методы продвинутой аналитики с использованием подхода «low-code», визуального проектирования и объектно-ориентированного подхода.

В рамках реализации данного проекта была определена следующая структура СППР на базе АП Loginom:

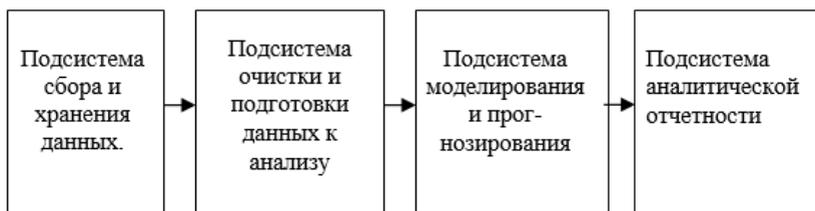


Рис. 1. Структура системы поддержки принятия решений

1. Разработка хранилища данных

Вся информация о клиентах компании консолидируется в специализированном хранилище данных. Хранилище данных ориентировано на решение именно задач анализа со специфичными для этих задач механизмами хранения данных. Использование единого хранилища позволяет гарантировать непротиворечивость данных и централизованное хранение, а также автоматически обеспечивает всю необходимую поддержку процесса анализа данных.

Чтобы ХД выполняло функции, соответствующие его основной задаче – поддержке процесса анализа данных, оно должно удовлетворять требованиям, сформулированным Р. Кимбаллом, одним из авторов концепции ХД [2]:

- высокая скорость получения данных из хранилища;
- автоматическая поддержка непротиворечивости данных;
- возможность получения и сравнения срезов данных;
- наличие удобных средств для просмотра данных в хранилище;
- обеспечение целостности и достоверности хранящихся данных.

Спроектированное ХД представлено ниже на рисунке 2. Хранилище данных содержит специальный семантический слой, обеспечивающий возможность работы с ним пользователю без

необходимости вникать в особенности хранения данных – пользователь оперирует привычными терминами: «дата», «абонент», «обращение».

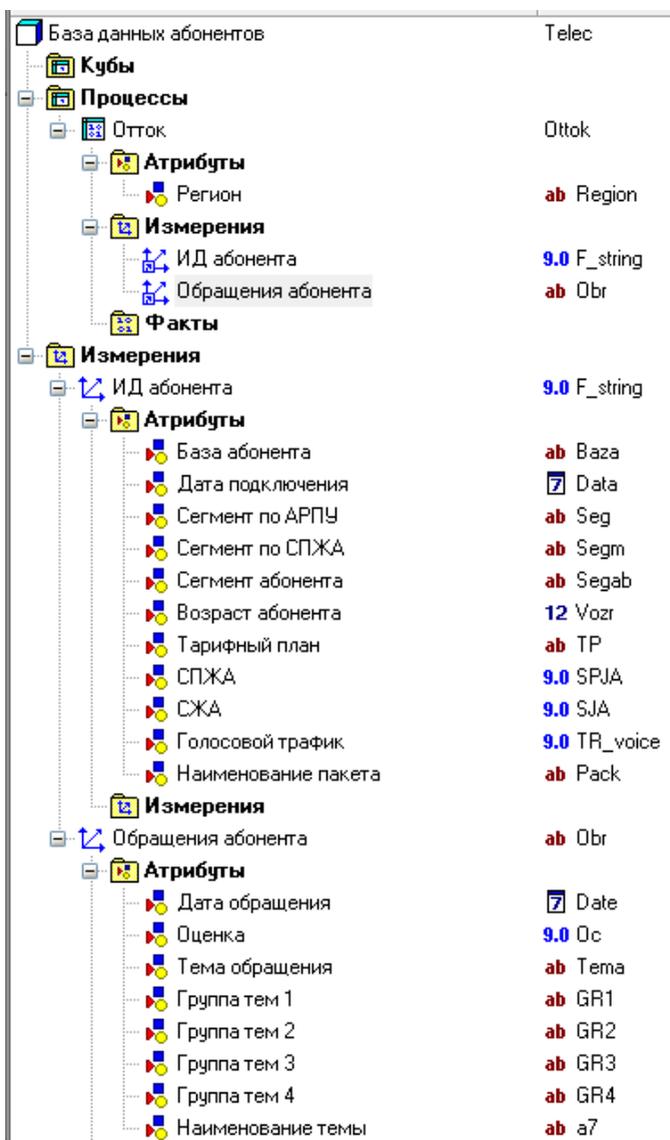


Рис. 2. Структура хранилища данных

После создания структуры хранилища данных оно представляет собой «пустое» ХД с настроенным семантическим слоем. В таком виде оно готово к загрузке в него данных из внешних структурированных источников. Для этого написан соответствующий сценарий, который должен выполнять следующие функции: импорт данных; опциональную предобработку данных, например, очистку или преобразование формата; загрузку данных в измерения и процессы хранилища данных.

В АП Loginom предусмотрены средства, позволяющие выполнять все действия автоматически в пакетном режиме. Благодаря этому большую часть аналитической обработки можно проводить по регламенту, например, когда нагрузка на вычислительные ресурсы сервера и ресурсы сети интернет минимальна. При помощи любого планировщика заданий можно настроить выполнение данного сценария по заданному регламенту. Таким образом, написанный сценарий автоматически может работать на новых данных с экспортом результатов обработки в хранилище данных.

Созданное ХД позволяет обеспечить целостность и непротиворечивость данных, их централизованное хранение, обеспечит возможность корректного построения прогнозных моделей на основе технологий Data Mining.

2. Реализация системы поддержки принятия решений

В процессе работы над проектом были собраны статистические данные по активности клиентов и их оттоку в компании «Мегафон», реализованы модули очистки данных, кластеризации, прогнозирования и аналитической отчетности в АП Loginom Community, используя встроенные low-code инструменты и визуальное проектирование [3].

В функциональность low-code-инструментов закладывается возможность доработки с помощью графического интерфейса и несложных скриптов, не требующих профессиональных навыков программирования.

Концепция low-code оказывается особенно полезной в случаях, когда требуется:

- увеличить скорость реализации и внедрения нового функционала информационных систем;
- выполнять быстрое прототипирование приложений и их компонентов;
- подключить к реализации логики обработки бизнес-экспертов;
- создать среду, позволяющую специалистам обмениваться знаниями в понятной визуальной форме.

Преимущества low-code подхода к аналитике дает возможность больше использовать готовые компоненты, а визуальное проектирование позволяет решать задачи анализа с меньшими затратами сил разработчиков. Так, например сценарий анализа данных в АП Loginom настраивается визуально (рис. 3). Пользователь видит, как обрабатываются данные на любом этапе, что упрощает поиск ошибок в расчётах. Изменения в разработанные сценарии вносятся просто: достаточно заменить один компонент обработки другим или переопределить логические связи. Гибкость и вариативность в применении компонентов закладывается через параметризацию, архитектура компонентов должна исключать дублирование логики обработки, в сценариях применяется наследование или ссылки на фрагменты обработки.

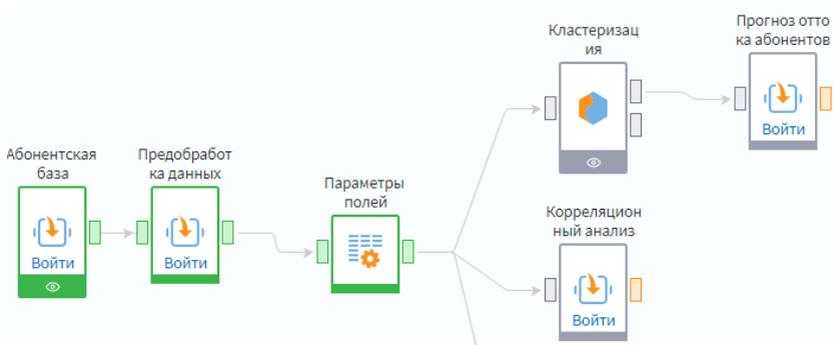


Рис. 3. Визуальное проектирование компонент сценария

Качество данных, которые собираются и консолидируются для анализа из различных источников, является одной из самых больших проблем аналитических технологий. С целью повышения качества данных используется комплекс методов и алгоритмов, получивших название «очистка данных». Для того чтобы правильно подготовить данные к анализу необходимо иметь стратегию их очистки, которая разрабатывается на основе знания структуры и особенностей источников, из которых поступают данные, характера самих данных, методики и цели их анализа. Эти задачи решает в сценарии узел «Предобработка данных». Этот компонент оценивает пригодность данных для анализа и прогнозирования, выявляет наличие ошибок в данных ряда (пропуски, дубликаты, противоречия), наличие аномальных значений, статистические характеристики.

АП Loginom имеет OLAP-ядро, способное на лету проводить многомерный анализ данных с возможностью визуализации результатов и экспорта в офисные приложения, базы данных и внешние сервисы.

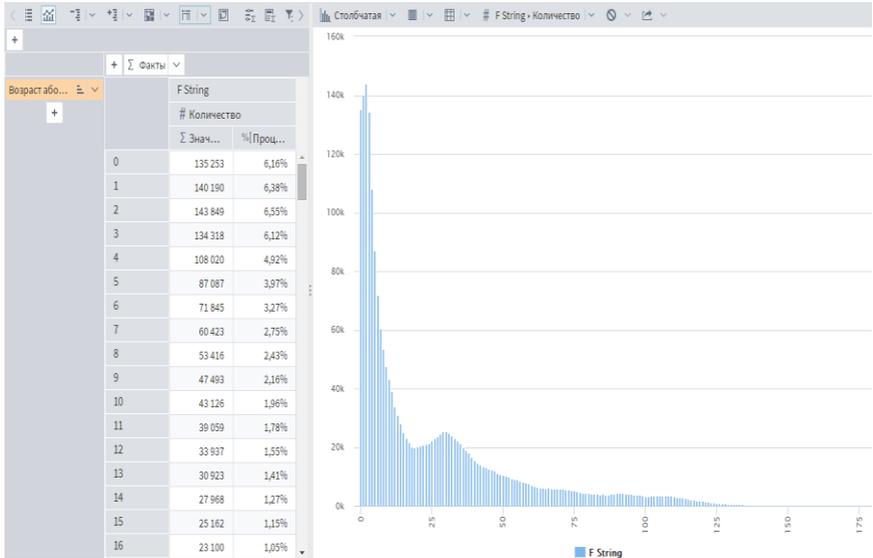


Рис. 4. OLAP-отчет «Отток в зависимости от возраста абонента»

Первый этап анализа выполняет сегментацию абонентской базы, которая заключается в разделении всех клиентов компании на группы, которые могут использовать разные предложения, и к которым необходимо прилагать разные маркетинговые усилия. Этот процесс является ключевым во всей цепочке работы, так как от того насколько успешно будет разделена база будет зависеть успех предлагаемых на последнем этапе предложений.

Методы интеллектуального анализа данных, в частности деревья решений и кластерный анализ, могут применяться для сегментирования клиентских баз. В этой работе с помощью обработчика Кластеризация, используя алгоритм g-means [2], было сформировано 9 кластеров. При сегментации для выделения наиболее прибыльных клиентов использовали показатели: средний доход от абонента, продолжительность отношений (сколько месяцев является клиентом), количество обращений в службу поддержки.

Следующий этап – прогноз оттока абонентов каждого кластера, используя обработчики нейронная сеть и логистическая регрессия. С

помощью реализованных моделей можно рассчитать вероятность ухода и вероятность отклика на новое предложение для конкретного клиента, а также оценить влияние различных факторов на отток. Оказалось, что самое высокое влияние на показатель оттока, по имеет показатель MOU – среднемесячное количество потраченных минут, входящего голосового трафика и ARPU – среднемесячные траты абонента, а показатели, связанные с возрастом абонента, имеют отрицательную корреляцию.

Заключение

Созданная библиотека компонентов сегментации и прогнозирования была протестирована на реальных данных компании Мегафон (база данных включала 2195000 записей об абонентах). Основными полями, характеризующие абонентов были их средние (за определенный период) показатели, связанные с входящим и исходящим трафиком, также были сформированы производные переменные, отражающие доли потребления типов услуг. При тестировании библиотеки сбоев и ошибок в работе компонентов обнаружено не было. Несмотря на работающий прототип, продукт требует доработок, прежде чем сможет быть подвергнут нагрузочному тестированию. Разработанная СППР в виде библиотеки компонентов на базе API Loginom даст возможность повысить эффективность маркетинговых кампаний за счет точного сегментирования клиентских баз, прогнозирования оттока абонентов и их отклика на рассылку о новых тарифах.

Список литературы

1. 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.
2. Паклин Н. Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям – СПб.: Питер, 2013. – 624 с.
3. Loginom: мало кодовая разработка (Low-code) [Электронный ресурс]. – Режим доступа: <https://wiki.loginom.ru/articles/low-code.html>