

СОЗДАНИЕ ИНФОРМАЦИОННОЙ СИСТЕМЫ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ ДАННЫХ ПО ГОСУДАРСТВЕННЫМ ЗАКУПКАМ

В.В. Костин, Ф.В. Яременко, А.Н. Бездушный

Введение

Создание информационной системы поддержки принятия решений на основе данных по закупкам является прямым ответом на соответствующую потребность. В первую очередь это связано с увеличивающимся количеством данных в открытом доступе по различным типам торгов – начиная от продажи имущества должников (банкротов) и заканчивая закупками негосударственных компаний.

В условия постоянно изменяющегося информационного мира компании, использующие специальные системы получения информации, ее обработки, поддержки принятия верного решения по времени подачи предложения, поиска оптимального контрагента получают значительное преимущество. Так, производитель товара крайне заинтересован в мониторинге различных рынков сбыта, объемов спроса на производимую продукцию, количества конкурентов и уровня цен, как на сам товар, так и на его доставку. Кроме того, имея возможность проанализировать действия конкурентов, их поведение при возрастании/снижении конкуренции, поставщик получает возможность спрогнозировать поведение участников рынка в изменяющихся условиях.

С другой стороны, заказчики, анализируя извлеченную информацию о рынке, получают возможность сопоставить свое предложение с реальными и потенциальными предложениями возможных конкурентов и грамотно построить свою ценовую политику.

Отдельного рассмотрения заслуживает их взаимная выгода. Кроме оптимизации поведенческой линии ценовой политике на рынке поставщики и заказчики имеют возможность рассмотреть большее количество партнеров и заключить соглашения с организациями, которые бы остались незамеченными при традиционном подходе.

На сегодняшний день активно развиваются проекты открытости информации. Один из них – это «открытое государство». В основе этого понятия лежит открытость и доступность данных, охватывающих все ветви власти, открытый диалог между государством и гражданами, приведение к общей содели всех проектов. Что образует почву для потенциального анализа этих данных.

Вторым схожим проектом стал проект Ивана Бегтина «открытые данные», направленный на открытие всех данных по происходящим в государственных структурах процессам. Кроме того, автор сайта является идейным вдохновителем создания различных проектов анализа этих данных.

Описанные выше примеры указывают на то, что в открытый доступ поступает все больше и больше информации, позволяющей проводить детальный анализ интересующей пользователя области. При этом не каждому пользователю доступна возможность приобрести дорогую и качественную СППР. Одним из решений данной проблемы является Software as a Service (SaaS) – «программное обеспечение как услуга». Такое программное обеспечение обслуживается и хранится на стороне поставщика, а заказчик получает у нему удаленный доступ. При этом затраты заказчика ограничиваются абонентской платой. Основное преимущество такого подхода – это возможность не платить за лицензию дорогостоящих продуктов, ограничив траты абонентской платой поставщику, которая на один порядок меньше полного пакета программы.

Задачи и Требования

Современные СППР направлены на решение различных задач. В зависимости от поставленной задачи определяются и требования. Перед рассматриваемой системой стояли следующие задачи:

1. Своевременный и полный сбор информации.
2. Наличие фильтров для возможности сужать область поиска.
3. Возможность писать запросы для полнотекстового поиска.
4. Возможность хранить архивные данные по уже прошедшим торгам.
5. Возможность анализировать имеющуюся информацию, используя как актуальную, так и архивную составляющую.
6. Возможность предоставлять пользователю информацию по той семантической составляющей, которая содержится в исходном запросе.

Для решения обозначенных выше задач были сформированы следующие требования:

1. Создание системы автоматического сбора информации, размещенной в на заданных торговых площадках.
2. Создание хранилища данных о прошедших торгах.
3. Создание классификаторов параметров, позволяющих проводить эффективную фильтрацию торгов.
4. Создание системы интеллектуального и семантического анализа данных.
5. Создание словарей ключевых терминов и системы семантических связей между ними.

6. Создание поисковой системы

В соответствии с описанными выше требованиями был сформирован план создания системы.

Реализация

В процессе создания обозреваемого программного средства программа разделена на независимые блоки, каждый из которых отвечает одному описанному выше требованию. Такой подход позволяет безболезненно проводить доработку и смену блоков и переход на другие системы, использующие более современные методы и технологии.

Система автоматического сбора данных реализована следующим образом – к различным источникам открытых данных отправляются xml запросы. При этом запрашивается вся информация о торгах – не только объявления или победители, но и поведение его участников (порядок повышения цен, активность, что позволит вычислить не только готовность контрагента к повышению предложения, но и его реакцию на изменение ситуации по торгам). Полученные данные по торгам собираются в базу оперативных данных. После закрытия торги (когда уже никакие изменения в них вноситься более не будут) переносятся в хранилище данных для оптимизации работы системы.

Хранилище данных создается на основе СУБД открытого доступа PostgreSQL. Одним из основных свойств хранилища является невозможность вносить изменения в хранящиеся там данные, поэтому туда попадает только информация по уже завершенным торгам, которая не подлежит дальнейшему изменению. На основе хранилища данных реализуются средства OLAP-анализа, его визуального и графического отображения, средства интеллектуального анализа данных. Система интеллектуального анализа данных включает в себя технологии кластеризации, карт Кохонена. Разрабатываются другие подходы анализа. Максимально эффективно используя свойства системы интеллектуального анализа данных можно выделять некие поведенческие императивы потенциальных конкурентов. Так, на основе собранной информации можно оценить готовность конкурента активно участвовать в торгах, максимальный процент потенциального превышенияповышения первоначальной цены, склонность подавать предложение в первые или последние минуты торгов.

В системе создаются полные словари классификаторов параметров торгов. Хранение и взаимосвязи классификаторов реализуются в виде owl-онтологии. Данное средство выбрано, потому что оно предоставляет возможность задавать неограниченное число связей между сущностями и отсутствие ограничений на эти связи. Путем интеллектуального анализа обучающего текста в отдельные сущности выделяются омонимы. Задаются синонимичные связи между терминами. На основе owl-онтологий создается система семантического анализа, которая позволяет проверять потенциальную достоверность гипотез и оценивать вероятность того или иного события.

Модуль поисковой системы создается на основе аналогичных систем для работы с английским языком. При этом, при адаптации для русского языка создается полностью новый перечень словоформ. Индексирование хранилища данных позволяет существенно уменьшить время выполнения запроса. Описанная выше owl онтология применяется для выполнения поисковых запросов не только по исходно заданному набору слов, но и по синонимам этих слов, а созданный на основе интеллектуального анализа данных перечень омонимов позволяет пользователю отфильтровать побочные значения введенного термина. Таким образом, поиск проводится не по формально написанному слову, а по семантике, которую в него вкладывают автор запроса.

Степень готовности и выводы

На данный момент реализован вариант СППР в виде системы «Росгосзаказ» (портал <http://rosgoszakaz.ru>). Эта система проводит сбор информации по торговым площадкам, собирает данные воедино и позволяет проводить выборку путем задания фильтров и полнотекстовых запросов. Написанные ранее запросы пользователя сохраняются для будущего использования. Система интеллектуального анализа данных находится в стадии тестирования, система семантического анализа находится в стадии разработки.

Данная система уже показала свою эффективность в качестве удобного информационного портала по поиску государственных заказов. Так, мэрия г.Орел за пол года сэкономила 23 млн. руб. при ее использовании.

ЛИТЕРАТУРА:

1. Атаева О.А., Бездушный А.Н. Приведение данных. Онтологический подход. // Научный сервис в сети интернет, 2009.
2. Мерцалов А.Н. Автоматизация процессов планирования зерновых ресурсов мукомольного производства. М, 2009.
3. Аbruков В.С., Николаева Я.Г., Макаров Д.Н., Сергеев А.А. Разработка моделей социальных явлений с помощью средств "Data Mining"// Чувашский государственный университет, г. Чебоксары, 2008
4. Яременко Ф.В., Бездушный А.Н. К вопросу выражения требований федеральных законов в виде формальных спецификаций бизнес-процессов// "Современные проблемы математики, информатики и биоинформатики", 11 - 14 октября 2011 г., Академгородок, Новосибирск, Россия.

5. Костин В.В., Бездушный А.Н. К вопросу создания информационно-аналитической системы месторождений редких металлов. // "Современные проблемы математики, информатики и биоинформатики", 11 - 14 октября 2011 г., Академгородок, Новосибирск, Россия.
6. Филатов В.А., Щербак С.С., Хайрова А.А. Разработка высокоэффективных средств создания и обработки онтологических баз знаний. <http://shcherbak.net/razrabotka-vysokoeffektivnyx-sredstv-sozdaniya-i-obrabotki-ontologicheskix-baz-znanij/>