**Обработка и анализ информационных потоков: системы поддержки принятия решений**

Естественное развитие систем управления базами данных и управленческих систем привело к появлению совершенно новых систем поддержки принятия решений. Как правило, основной задачей таких систем является возможность работы с неструктурированными и слабоструктурированными массивами данных.

Системы поддержки принятия решений, несмотря на значительное распространение в крупном бизнесе и органах государственной власти, относятся к решениям, про которые принято говорить как о «будущем в ИТ-системах для среднего и малого бизнеса». Тем не менее здесь будут рассмотрены основные подходы, используемые в различных решениях такого рода.

Итак, системы поддержки принятия решений (СППР) могут быть необходимы в случае, если у бизнеса есть потребность в обработке больших объемов разнородной (постоянно поступающей) информации с последующим анализом и стратегическим планированием дальнейших действий. В СППР используется несколько так называемых аналитических технологий — методик, позволяющих по известным данным оценить значения неизвестных характеристик и параметров. Современные разработки в этой области предоставляют возможности учета большого потока разнородной информации об объекте исследования, и в то же время широкие возможности инфраструктурного анализа позволяют руководителю своевременно принять правильное решение. В зависимости от данных, с которыми эти системы работают, СППР условно можно разделить на оперативные и стратегические. Первые предназначены для немедленного реагирования на изменения текущей ситуации в управлении финансово-хозяйственными процессами компании. Вторые ориентированы на анализ значительных объемов разнородной информации, собираемой из различных источников. Важнейшей целью этих СППР является поиск наиболее рациональных вариантов развития бизнеса компании с учетом влияния различных факторов: конъюнктуры целевых для компании рынков, изменения финансовых рынков и рынков капиталов, изменения в законодательстве и т.д. На сегодняшний день аналитические системы практически не используются в среднем и малом бизнесе. Возможности прогнозирования и моделирования ситуаций (так называемые «ситуационные центры»), а также функции автоматизированного сбора данных и их обработки существуют немногим более чем у 5% предприятий SMB. Тем не менее с дальнейшим ростом объемов информации на предприятиях и удешевлением конечных решений такого класса, можно прогнозировать бурный рост рынка СППР. В качестве первичного источника данных для аналитических систем должны выступать СУБД организации, офисные документы, сеть Интернет. При этом должны учитываться как внутренние для организации данные, так и глобальные сведения (макроэкономические показатели, конкурентная среда и т. д.).

Хранилище данных — оптимальная база для построения аналитической системы (АС). Работа с таким хранилищем значительно увеличивает ее эффективность, поскольку одним из ключевых показателей АС является возможность быстро получить результат.

Следующий шаг на пути к принятию решения— выборка данных. Независимо оттого, в какой базе данных находятся необходимые сведения, лицо, принимающее решение не должно вникать в детали работы с СУБД. Поэтому необходим механизм, трансформирующий термины предметной области в запросы к конкретной БД. Дальнейшие шаги — это собственно анализ и представление конечных результатов. Существует два методологических подхода в таких системах: выработка рекомендаций (концепция data mining) и подготовка данных (OLAP).

**OLAP - средство составления отчетов на основе системы запросов**

OLAP (Online Analytical Processing) — технология, основанная на инструментах математической статистики, она применяется главным образом для анализа и отображения информации в виде многомерных структур, называемых также «кубы OLAP». Позволяет решать следующие задачи:

1. Подготовить базы данных (часто объемные и содержащие сложные взаимосвязи);

2. Организовать гибкий и удобный доступ к базам данных через мощные средства формирования запросов;

3. Получить результаты запросов в форме, максимально удобной для последующего анализа;

4. Использовать мощные генераторы отчетов.

Такой подход может быть очень полезен в том случае, если лицо, принимающее решение, использует компьютер только для извлечения необходимых данных, представления этих данных в структурированном, понятном виде, а выводы делает самостоятельно.

Представленное преобразование данных в трехмерную структуру — один из мощнейших инструментов технологии OLAP. Он отличается гибкостью: каждый пользователь может определять нужные многомерные проекции данных без каких-либо ограничений. Кроме того, в рамках этого метода существует возможность производить детализацию данных до нужного уровня. Таким образом, технологию OLAP стоит рассматривать как средство формирования и поиска запросов к базе данных (хранилищу данных). При этом функциональности OLAP явно недостаточно, если требуется более детальный анализ либо есть необходимость в автоматизированном поиске скрытых взаимодействий между объектами в представленном массиве информации.

**Data Mining - комплексный подход к интеллектуальному анализу данных**

В отличие от методов аналитической обработки информации и создания отчетов, концепция Data Mining предполагает обнаружение нетривиальных взаимосвязей между объектами данных, которые нужны для принятия решений. В частности, инструментарий выработки рекомендаций обладает следующими возможностями:

1. Формирование множества альтернативных вариантов решений;

2. Использование нескольких критериев оценки;

3. Учет важности критериев;

4. Выбор лучшего варианта, который выдается как рекомендация.

Выделяют пять типов закономерностей, которые позволяет выявлять Data Mining: классификация, кластеризация, регрессия, ассоциация, последовательность и прогнозирование. Кратко их можно охарактеризовать так:

1. Классификация — это отнесение объектов (наблюдений, событий) к одному из заранее известных классов;

2. Кластеризация —это группировка объектов (наблюдений, событий) на основе данных (свойств), описывающих сущность объектов. Объекты внутри кластера должны быть похожими друг на друга и отличаться от объектов, вошедших в другие кластеры. Чем больше похожи объекты внутри кластера и чем больше отличий между кластерами, тем точнее кластеризация;

3. Регрессия, в том числе задачи прогнозирования. Установление функциональнои зависимости между зависимыми и независимыми перемененными;

4. Ассоциация — выявление закономерностей между связанными событиями. Примером такой закономерности служит правило, указывающее, что из события X следует событие Y. Такие правила называются ассоциативными. Впервые это задача была предложена для нахождения типичных шаблонов покупок, совершаемых в супермаркетах, поэтому иногда ее еще называют анализом рыночной корзины (market basket analysis);

5. Последовательные шаблоны—установление закономерностей между связанными во времени событиями. Например, после события X через определенное время произойдет событие Y;

6. Анализ отклонений — выявление наиболее нехарактерных шаблонов.

Это все, что нужно сделать для автоматизации процесса извлечения данных. Все остальное делает лицо, принимающее решение. Различные методы просто дают разную информацию в разных видах: в простейшем случае это таблицы и диаграммы, в более сложном — модели и правила. Полностью исключить участие человека невозможно, поскольку выбранные данные не имеют никакого значения, пока не будут применены в конкретной предметной области. Таким образом, методы решения задачи по принятию решения не зависят от инструментария. Поэтому в рамках двух вышеописанных парадигм может существовать сколь угодно широкий набор инструментов. Говорить о действительно полнофункциональном решении можно только в том случае, если был охвачен весь список задач. По мнению руководителей крупнейших компаний, разрабатывающих специализированные информационно-аналитические системы и системы поддержки принятия решений, это направление должно стать приоритетным при информатизации бизнеса. Основная задача, решаемая при переходе на использование таких систем, — помочь организациям наладить контроль и управление, способствующие повышению эффективности, рациональности и качества оказываемых услуг.

**Список литературы**

IT спец № 07 ИЮЛЬ 2007