

УДК 681.3.06

Д.В. Коломиец

Одесский национальный университет им. И.И. Мечникова, Одесса

МНОГОМЕРНАЯ МОДЕЛЬ ПРОДАЖ ГИПЕРМАРКЕТА

В статье рассматривается процесс построения многомерной модели продаж крупного объекта оптово-розничной торговли. Определено назначение модели, указаны основные источники данных, детально рассмотрен процесс, методология и проблемы построения многомерного гиперкуба средствами Microsoft SQL Server Analysis Services. В завершении продемонстрированы возможности полученной модели, а также представлены варианты ее дальнейшего использования.

Ключевые слова: гиперкуб, OLAP, хранилище данных, модель, анализ, размещение, торговля.

Введение

Проблемы анализа огромного количества накопленных данных в различных системах поддержки принятия решений (СППР) являются особо острыми для большого числа крупных торговых организаций. Многие, так называемые гипермаркеты, обладают программным обеспечением (ПО), позволяющим вести оперативный кассовый и бухгалтерский учет, базирующийся в основном на разработках российской фирмы «1С» или аналогичных программных продуктах. Но, к сожалению, это программное обеспечение относится к классу OLTP систем – системам оперативного учета транзакций (Online Transactional Processing), которые не позволяют получать аналитические данные и характеристики, необходимые руководящему персоналу для принятия экономически обоснованных и взвешенных стратегических решений.

В работе предлагается способ построения многомерной модели данных, который сделает возможным проведение анализа объемов продаж в разрезе следующих измерений: категории товара, производители и поставщики товара, зоны размещения товаров в торговом центре и т.п. Особенностью данной модели является закладываемая изначально возможность анализа экономических результатов (например, объемов продаж) в привязке к пространственному размещению товара или группы товаров в виртуальной трехмерной модели торговывыставочного зала организации.

Основной материал

Куб хранилища данных (ХД), который позволит «объединять, просматривать и анализировать данные с точки зрения множественности измерений» [1], будем строить на основании двух первичных разнородных источников данных: базы данных системы оперативного учета продаж «1С» и базы данных приложения, служащего для моделирования трехмерного визуального представления торговых помещений с размещенными в них стеллажами с товарами [2]. База данных системы «1С» располагает информацией об остатках товаров на определенный момент времени, их ценах и поставщиках. Информационная система моделирования трехмерного торговых комплексов, наряду с учетом товаров и торгового оборудования, также позволяет:

- формировать карты торгового помещения и размещать оборудование на картах (Рис. 1);
- размещать товары на торговом оборудовании;
- выполнять просмотр виртуального трехмерного представления торгового зала (Рис. 2).

В качестве средства проектирования многомерной модели был выбран пакет Microsoft Analysis Services 2005.

Служба Analysis Services 2005 является частью продукта Microsoft SQL Server 2005.

Помимо Analysis Services SQL Server включает в себя также службы Integrations Services, позволяющую определять потоки данных, расписания и

соответствующие трансформации данных при помещении в хранилище, и Reporting Services, предназначенную для удобного построения аналитических отчетов. Вместе эти службы формируют ядро платформы бизнес аналитики, использующей SQL Server в качестве сервера данных.

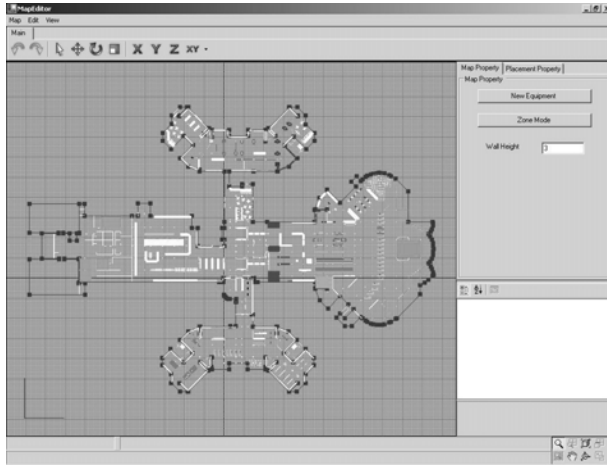


Рис. 1. Редактор карт торгового помещения

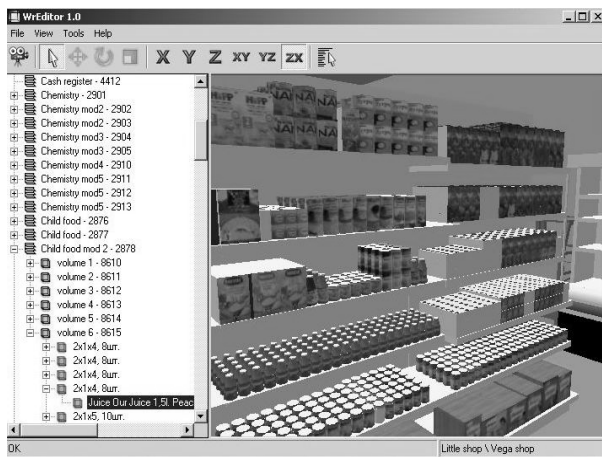


Рис. 2. Трехмерная сцена торгового помещения

Analysis Services в версии Microsoft SQL Server 2005 предлагает совершенно новый подход к моделированию, администрированию и выполнению запросов к данным, в основе которого:

- Унифицированная Многомерная Модель (Unified Dimensional Model, UDM), позволяющая различным клиентским приложениям получать доступ к данным многомерных баз данных без использования моделей каждого типа баз данных. При этом для определения моделей данных используется язык XML;
- расширенная поддержка языка запросов к многомерным данным (Multidimensional Expressions, MDX);
- использование тонкой клиентской архитектуры на базе протокола XML for Analysis (XML/A);

- увеличенная масштабируемость системы, снимающая ограничения на количество многомерных объектов, таких как атрибуты, иерархии, измерения, кубы и группы мер;

- богатый набор средств и инструментов для разработки приложений и управления серверами (SQL Server Business Intelligence Development Studio и Microsoft SQL Server Management Studio) [3].

Analysis Services 2005 не только обеспечивает возможность создавать измерения и кубы для анализа данных, но также поддерживает набор алгоритмов добычи данных (Time Series, Decision Trees, Clustering, Linear Regressions, Neural Network), которые могут обеспечить более глубокое понимание сути данных и выявить различные тенденции [4].

Основой гиперкуба будет являться таблица фактов – GoodsRests_Fact, включающей суррогатный первичный ключ, состоящий из комбинации значений ключевых полей таблиц измерений, а также анализируемые величины: остатки товаров, объемы продаж и стоимости товаров (Рис. 3).

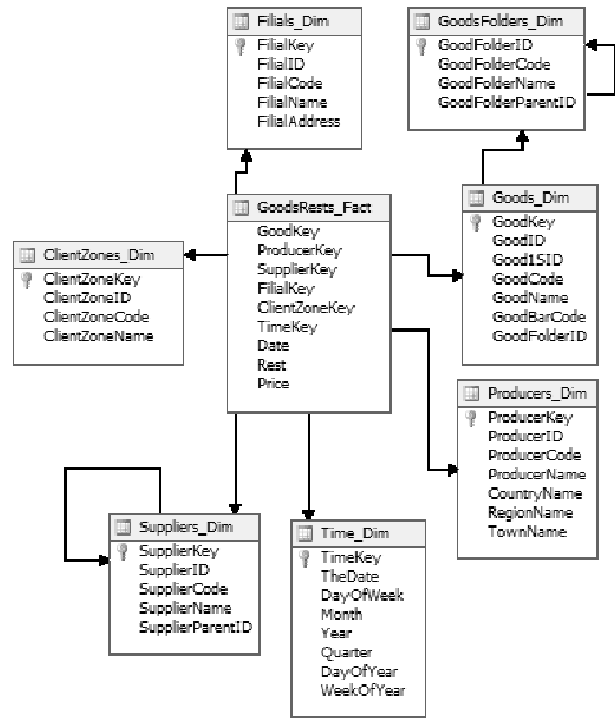


Рис. 3. Схема куба данных

Также будем использовать следующее множество таблиц-измерений:

- Goods_Dim – таблица измерения по товарам. Включает в себя следующие атрибуты: идентификационный номер товара, код, наименование, номер товара, используемый в системе «1С» и ссылка на категорию товара. Таблица также будет содержать в себе бинарную сериализованную информацию, характеризующую пространственное размещение товара в торговом комплексе. Наличие такого рода данных требует определения дополни-

тельных хранимых процедур и функций, которые будут выполнять математический расчет геометрических особенностей размещения (центр масс, расстояние товаров в пространстве), которые в дальнейшем будут использованы при определении классификационных функций.

- GoodsFolders_Dim – таблица категорий товаров. На самом деле, многомерный анализ экономических результатов удобнее проводить не в разрезе определенного товара, а в контексте группы логически связанных единиц товаров. Данное измерение обладает свойством иерархичности (например, безалкогольные напитки представлены следующими группами товаров: вода минеральная, вода фруктовая и соки.). Достигается иерархичность за счет наличия внешнего ключа GoodFolderParentID.

- Producers_Dim – таблица производителей товаров. Помимо скалярных характеристик (номер производителя, его код и имя) данное измерение включает в себя информацию о географическом размещении производителя товара, что в свою очередь дает возможность получать агрегированную информацию о продажах товаров производителей из определенной страны, города, т.е. также обладает иерархичной формой представления.

- Suppliers_Dim – таблица поставщиков товаров, которые характеризуются уникальным номером, кодом, именем, а также атрибутом (SupplierParentID), определяющим иерархию подразделений поставщиков.

- Filials_Dim – таблица измерений по филиалам торговых организаций. Позволяет анализировать объемы продаж товаров в разрезе филиалов. Таким образом, у аналитиков и управляющих появляется возможность оценки работы подразделения, как в общем, так и по конкретным группам товаров.

- ClientZones_Dim – таблица измерений зон визуальной представленности товара. Зона представленности товара определяет его геометрическое размещение в торговом зале и характеризует уровень услуг, предоставляемых предприятием поставщику. Данное измерение даст возможность проведения анализа объемов продаж с учетом условий сотрудничества организации с поставщиком товара.

- Time_Dim – таблица временных измерений. Позволит получать и просматривать данные о продажах во временном разрезе, группируя по дням недели (выходные или рабочие дни), месяцам, кварталам, годам [5].

Построенная в виде гиперкуба модель, позволяет получать экономические характеристики, представленные в виде сводных таблиц, в которых аналитик сам может выбирать, по каким измерениям или группам измерений проводить анализ. Над таким гиперкубом могут выполняться следующие операции:

- срез (Slice) – формирование подмножества многомерного массива данных, соответствующего одного или нескольких элементов измерений. Например, выбор подмножества значений объемов продаж за весенний период 2006 года. На Рис. 4 продемонстрирована возможность выполнения среза данных по конкретному филиалу торгового комплекса («Gastello»).

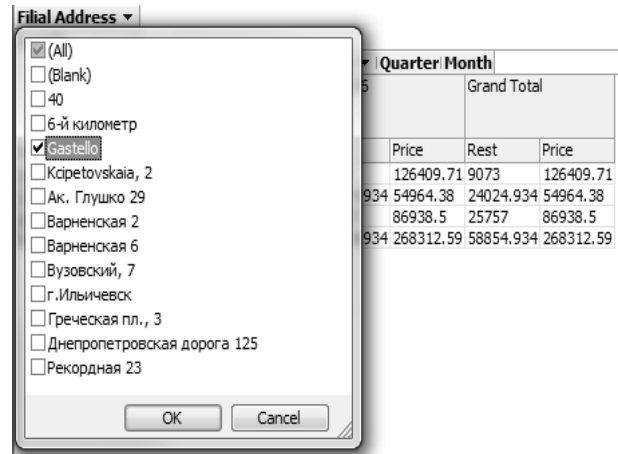


Рис. 4. Срез по торговому филиалу

- вращение (Rotation) – изменение расположения измерений, представленных на отображаемой странице или отчете. В среде Microsoft SQL Server Business Intelligence Development Studio операция вращения реализуется простым перетаскиванием измерений поверх просматриваемого куба данных.

- консолидация (Drill Up) и детализация (Drill Down) – операции, которые определяют переход вверх по направлению от детального представления к агрегированному и наоборот. Например, при анализе объемов продаж производителей из Украины можно выполнить операцию Drill Down для измерения «Регион», тогда будут отображены такие его элементы, как «Одесская область», «Черкасская область», «Киевская область» и т.п. В результате дальнейшей детализации элемента «Черкасская область» будут отображены элементы «Черкасс», «Умань», «Жашков» и т.д. На Рис. 5 представлен пример детализации по категории товара (детализация от категории «Безалкогольных напитков» к таким группам как «Вода минеральная», «Вода фруктовая» и «Соки»).

Выводы

Построенная модель может служить основой для проведения так называемого процесса «добычи данных» (Data Mining). При этом спектр решаемых методами Data Mining задач может быть достаточно широк. Например, мы можем поставить задачу классификации производителей или поставщиков в зависимости от объемов реализации произведенной/поставленной ими продукции.

		Year ▼ Quarter Month		Grand Total	
		2006			
Level 02	Level 03	Rest	Price	Rest	Price
☑ Алкогольные напитки		54438	758458.26	54438	758458.26
☑ Безалкогольные напитки	Вода Минеральная	118523	287365.85	118523	287365.85
	Вода Фруктовая	11292	23245.2	11292	23245.2
	Соки	13463.604	18401.22	13463.604	18401.22
	Total	143278.604	329012.27	143278.604	329012.27
☑ Консервированная продукция		152342	513931	152342	513931
Grand Total		350058.604	1601401.53	350058.604	1601401.53

Рис. 5. Детализация по категории товара

Вследствие такой классификации поставщики могут быть отнесены к группе крупных поставщиков, средних и мелких. Более того, полученная информация может быть использована, например, для определения факта оправданности выделения конкретному заказчику определенной зоны представленности.

Обладая информацией о пространственных координатах размещения товаров в торговых залах, указанная модель позволит проводить анализ объемов продаж групп товаров, расположенных рядом. Т.е., имея информацию о реализации некоторых товаров за значительный промежуток времени, можно судить о наличии или отсутствии закономерности продаж от степени близости размещения товаров в зале [6].

Таким образом, благодаря удачному симбиозу источников данных (данных об остатках товаров за продолжительный период времени и информационной системы поддержки трехмерного размещения товаров), предлагаемая модель позволяет ставить и решать принципиально новые задачи в сфере анализа продаж с учетом «правильности» размещения в торговом зале. При этом полученные результаты могут быть полезны маркетологам, исследующим проблему более качественного способа предоставления услуг, мэрчандайзерам, изучающим способы оптимальной выкладки товаров в торговом зале, а также управленческому персоналу для принятия экономически обоснованных стратегических решений.

Список литературы

1. Барсегян А.А. Методы и модели анализа данных: OLAP и Data Mining / А.А. Барсегян. – СПб.: БХБ-Петербург, 200.
2. Информационно-моделирующая система поддержки принятия решений для сверхбольших обслуживающих комплексов / П.В. Назаренко, Э.А. Винницкий, Е.В. Голушиков, Т.И. Петрушина // Интеллектуальные системы принятия решений и прикладные аспекты информационных технологий: труды Международной научной конференции, Евпатория (Крым), 18-21 мая 2005 года. – Т. 1. – С. 96-99.
3. Бергер А. Microsoft SQL Server 2005 Analysis Services. OLAP и многомерный анализ данных / А. Бергер, И. Горбач. – СПб.: БХБ-Петербург, 2007
4. Харинатх С. SQL Server Analysis Services 2005 и MDX / Сивакумар Харинатх, Стивен Куинн. // Диалектика. – 2008.
5. Коломієць Д.В. Побудова многомерної моделі продаж гіпермаркета / Д.В. Коломієць, Т.І. Петрушина // Інноваційні проблеми автоматизації управління бізнесом: матеріали Першої міжнародної науково-практичної конференції. – Житомир: ИПСТ, 2007. – С. 13-16.
6. Коломієць Д.В. Аналіз об'єму продаж гіпермаркету / Д.В. Коломієць // Сучасні проблеми прикладної математики та інформатики: матеріали XIV Всеукраїнської наукової конференції. – Львів: Львівський національний університет ім. Франка, 2007. – С. 78-79.

Поступила в редколлегию 16.04.2011

Рецензент: д-р физ.-мат. наук, проф. В.Г. Шевчук, Одесский национальный университет им. И.И.Мечникова, Одесса.

БАГАТОВИМІРНА МОДЕЛЬ ПРОДАЖІВ ГІПЕРМАРКЕТУ

Д.В. Коломієць

В статті розглядається процес побудови багатовимірної моделі продажів великого підприємства оптово-роздрібною торгівлю. Визначено призначення моделі, зазначені головні джерела даних, деталізовано розглянутий процес, методологія та проблеми побудови багатовимірної моделі засобами Microsoft SQL Server Analysis Services. Наприкінці продемонстровані можливості отриманої моделі, а також наведені варіанти її подальшого використання.

Ключові слова: гіперкуб, OLAP, сховище даних, модель, аналіз, розміщення, торгівля.

MULTIDIMENSIONAL HYPERMARKET SALES MODEL

D.V. Kolomyiets

Multidimensional hypermarket sales model creation process is shown in the article. Purpose of the model is defined in the beginning of the work. Then the main data sources of the model are specified and methodology and common problems of usage Microsoft SQL Server Analysis Services are described in details. Abilities of the created model are shown at the end of the article as well as the way this model can be used in future analysis.

Keywords: hypercube, OLAP, data warehouse, model, analysis, placement, sales.