

УДК 658

О.В. Луняков, доцент, канд. эконом. наук,

К.Н. Селищева, магистрант

Севастопольский национальный технический университет

Студгородок, г. Севастополь, Украина, 99053

E-mail: nicolaevna@mail.ru

ПРИМЕНЕНИЕ OLAP – ТЕХНОЛОГИЙ В АНАЛИЗЕ АССОРТИМЕНТА СЕТИ СУПЕРМАРКЕТОВ

В статье рассмотрены вопросы применения OLAP-технологий для решения различных оптимизационных задач, в частности для анализа и формирования оптимального ассортимента сети супермаркетов. Предложен вариант использования конкретного программного продукта Mondrian 2.0.1 для решения этой задачи.

Постановка проблемы. Основная цель торгового предприятия – получить желаемую прибыль. Пути ее достижения – удовлетворение спроса покупателей, повышение конкурентоспособности, активное развитие. Но в попытках добиться решения поставленных задач используются далеко не все методы и средства. Значительная часть продовольственных магазинов довольно неэффективно работает над ассортиментом, потому складываются ситуации, когда предложение не всегда соответствует спросу.

Как показывают опросы покупателей, именно ассортимент в наибольшей степени влияет на выбор места покупки. Поэтому анализ и подбор оптимального ассортимента является наиболее важным аспектом деятельности супермаркета.

В силу специфики ассортимента супермаркета (десятки тысяч наименований), сложность анализа ассортиментной политики очевидна. Поэтому возникает необходимость внедрения специализированного программного обеспечения на базе OLAP – технологий, которое позволило бы работать с большими объемами данных быстро и эффективно.

OLAP (On-Line Analytical Processing) – это технология комплексного многомерного анализа данных, ключевой компонент организации хранилищ данных [1].

OLAP-анализ (анализ в реальном масштабе времени) работает с данными, представленными в виде многомерных кубов, и позволяет построить отчеты в нужных руководителю срезах и ракурсах без обращения к программистам и ожидания нужной отчетной формы. OLAP также является удобным инструментом быстрого анализа больших объемов данных и наглядного отображения результатов в виде рисунков, графиков, диаграмм и таблиц. Этот класс программ адресован в основном высшему руководящему звену. Выявленные с помощью OLAP-технологии закономерности и тенденции – прекрасный материал для принятия правильных управленческих решений, а также для создания всевозможных отчетов и презентаций.

Анализ последних исследований и публикаций. Концепция OLAP была описана в 1993 году Эдгаром Коддом, известным исследователем баз данных и автором реляционной модели данных (см. E.F. Codd, S.B. Codd, and C.T. Salley, Providing OLAP (on-line analytical processing) to user-analysts: An IT mandate. Technical report, 1993). Основы применения аналитической обработки данных (OLAP) и хранилищ данных, а также возможности, предоставляемые современными OLAP-средствами, описаны такими авторами, как Алексей Федоров и Наталия Елманова [1]. В своей книге "Введение в OLAP-технологии Microsoft" они касаются вопросов создания многомерных баз данных, манипуляции OLAP -данными в microsoft excel, применения различных приложений для работы в OLAP. Барсегян А. и Куприянов М в своей книге «Методы и модели анализа данных: OLAP и Data Mining» [2] освещают основные направления в области анализа данных, в том числе оперативный (OLAP). В трудах многих авторов (например, Архипенков С.Я., Голубев Д. [3]) рассматриваются проблемы организации работ и выполнения проекта построения корпоративной информационно-аналитической системы с помощью хранилищ данных и OLAP. Большое внимание уделяется актуальному направлению расширения аналитических систем до масштабов World Wide Web в сетях интранет/Интернет. Этот аспект также рассматривается Джорджем Пеком [4]. Его книга посвящена построению масштабируемых Web-приложений и отчетов профессионального качества для любых баз данных с помощью мощного аналитического инструмента. Он также рассматривает использование многомерных функций OLAP в запросах и отчетах.

Выделение нерешенных частей проблемы. Все вышеперечисленные авторы рассматривали в качестве примера платные OLAP-продукты. Однако, в современных условиях украинские предприятия малого и среднего бизнеса не могут позволить себе покупку дорогого программного обеспечения. Поэтому в работе предлагается, как вариант, применение бесплатной и открытой OLAP-системы Mondrian 2.0.1.

Цель и задачи статьи. Цель данной публикации состоит в изучении возможностей применения OLAP – технологий для анализа ассортимента сети супермаркетов на базе Mondrian 2.0.1. Задачи статьи:

- определение сущности OLAP - технологий;

- проектирование хранилища данных для анализа ассортимента сети супермаркетов ООО «Ален»;
- построение OLAP-куба в программе Mondrian 2.0.1 для анализа ассортимента сети супермаркетов ООО «Ален».

Результаты исследования. OLAP-технология дает возможность работать с многомерными данными, обеспечивает доступ ко всей, связанной с анализируемыми данными информации независимо от места и способа её хранения.

OLAP использует стандартные методики анализа бизнеса: "Динамика итоговых значений", "ABC анализ", "Структурный анализ", "Корреляция", а также встроенный математический аппарат для исследования массивов данных (определение минимумов и максимумов, интерполяция и экстраполяция, основные функции статистического анализа).

OLAP-технологии позволяют получать обобщения вместо цифр, тенденции вместо статистики и сценарии развития предприятия, зависящие от различных внутренних и внешних факторов (в том числе и от принимаемых решений).

Рассмотрим использование методов OLAP - обработки данных для анализа ассортимента сети супермаркетов (на примере программного продукта Mondrian 2.0.1).

Mondrian 2.0.1 является бесплатным программным продуктом, что дает ему существенное преимущество.

Важным достоинством Mondrian 2.0.1 является открытость его архитектуры, т.е. пользователи сами могут дополнять и изменять функциональный состав системы по желанию.

Оптимизация бизнес процесса по формированию ассортимента будет осуществляться с помощью составления OLAP-кубов данных (агрегированных таблиц).

Mondrian работает с данными непосредственно в реляционной базе данных. Как только Mondrian считывает часть данных, он размещает их в своем кэше. Это значительно упрощает установку программы.

Например, для того чтобы получить отчет, содержащий одно число: общая выручка по всем торговым комплексам сети супермаркетов за 2006 год, Mondrian генерирует запрос и отправляет его в базу данных. БД выполняет его за несколько минут. Все записи за год считываются, и затем агрегируются в одну сумму. Все, что нужно для выполнения этой операции – это наличие агрегированных таблиц.

Агрегированные таблицы существуют одновременно с таблицами фактов*. Рассмотрим агрегированные таблицы на примере схемы хранилища данных, разработанной для сети супермаркетов ООО «Ален» в г. Севастополе (рисунок 1).

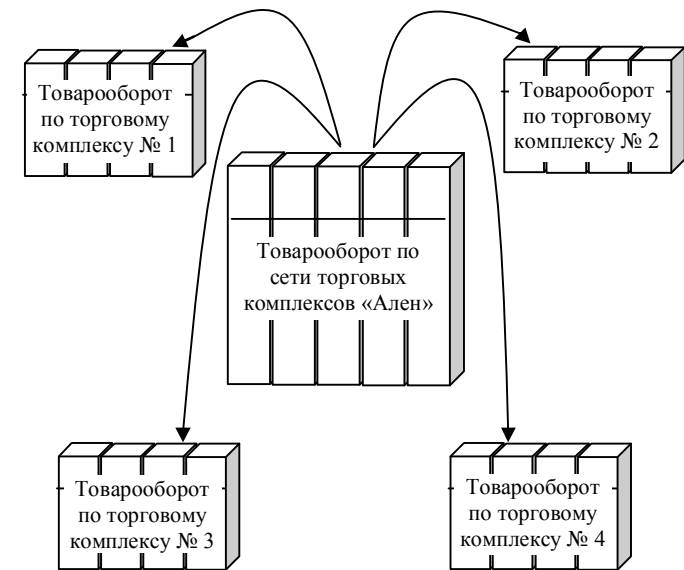


Рисунок 1 – Пример схемы хранилища данных для сети ООО «Ален»

Схема, приведенная на рисунке 1, содержит одну таблицу фактов «Товарооборот по сети супермаркетов» и четыре таблицы измерений** «Товарооборот по торговому комплексу».

Далее составим схему хранилища данных для одного торгового комплекса ООО «Ален» (рисунок 2).

* Таблица фактов – основная таблица хранилища данных, содержит сведения об объектах или событиях, совокупность которых будет в дальнейшем анализироваться.

** Таблица измерений – структурный элемент хранилища данных, содержащий неизменяемые либо редко изменяемые данные.

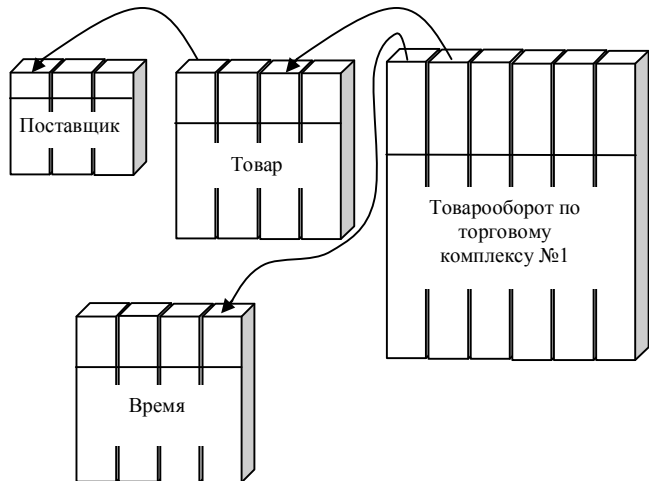


Рисунок 2 – Пример схемы хранилища данных для одного торгового комплекса ООО «Ален»

Схема, приведенная на рисунке 2, содержит одну таблицу фактов «Товарооборот по торговому комплексу» и три таблицы измерений (Время, Товар, Поставщик). Теперь можно создать агрегированную таблицу (рисунок 3).

Год	Квартал	Поставщик	Наименование товара	Товарный штрих-код	Суммарный объем продаж в нат. выр.	Мин-ый объем продаж в нат. выр.	Макс-ый объем продаж в нат. выр.	Суммарный объем продаж в ден. выр.	Счетчик строки
Агрегированная таблица									

Рисунок 3 – Пример агрегированной таблицы для хранилища данных по торговому комплексу № 1

Таблица фактов может содержать 0 и более агрегированных таблиц. Агрегированная таблица группирует измерения таблицы фактов в одну или несколько размерностей. Например, если конкретная колонка в таблице фактов содержит количество продаж какого-то продукта в заданный день в заданном магазине, то агрегированная таблица суммирует информацию помесечно. Такая агрегированная таблица представляет собой 1/30 размера таблицы фактов. Таким образом, выполнение запросов значительно ускорится при работе с агрегированными таблицами.

Можно создать таблицу, агрегирующую данные не только по месяцам, но и по отдельным торговым комплексам. ООО «Ален» владеет девятью продовольственными магазинами, следовательно, агрегированная таблица будет составлять 1/270 от размера первоначальной таблицы фактов.

Агрегированную таблицу, представленную на рисунке 3, можно построить с помощью Mondrian-2.0.1. Программа обращается к хранилищу данных (рисунок 2) и затем проектирует OLAP-кубы (агрегированные таблицы) по запросу пользователя.

Рабочий лист Mondrian представлен на рисунке 4.

Измеритель	
Товарный комплекс	Продукт
± Все	± Все
936108	12 332 012,00

Рисунок 4 – Рабочий лист Mondrian-2.0.1

Охарактеризуем функциональность кнопок программного меню Mondrian-2.0.1, представленного на рисунке 4:

Значок позволяет открыть OLAP-навигатор. С помощью OLAP-навигатора можно самостоятельно спроектировать OLAP-куб (многомерную агрегированную таблицу данных).

Значок позволяет спроектировать OLAP-куб с помощью запроса на языке MDX (Multidimensional Expressions).

Спроектированный OLAP-куб также представлен на рисунке 4. Его можно отсортировать по убыванию (значок). Кнопки позволяют менять внешний вид OLAP-куба. Проверить свойства OLAP-куба можно с помощью кнопки . Значок скрывает ячейки с нулевым значением, транспонирует OLAP таблицу, позволяют спроектировать и отобразить диаграмму, интегрирует данные в MS Excel.

Это только часть функций, которые реализует Mondrian-2.0.1. Программа позволяет быстро и просто формировать структуру требуемых запросов и отчетов, в течение 5-10 секунд получать отклик на сформированные запросы и пересчитывать отчеты, применять различные методики анализа, в том числе и свои собственные, моделировать и прогнозировать всевозможные бизнес-ситуации.

На отечественных предприятиях подобный анализ часто осуществляют через MS Excel. Однако, программа Mondrian-2.0.1 обладает рядом преимуществ перед MS Excel:

- Mondrian в отличие от MS Excel является бесплатным программным продуктом;
- открытость архитектуры Mondrian-2.0.1;
- Mondrian реализует интерактивный анализ больших и маленьких объемов информации;
- Mondrian позволяет исследовать информацию многомерно;
- высокоскоростное выполнение запросов за счет использования агрегированных таблиц в Mondrian-2.0.1;
- в Mondrian математические расчеты реализованы через MDX (язык многомерного описания данных);
- Mondrian реализует мгновенное обновление данных;
- в Mondrian в отличие от MS Excel отсутствуют ограничения по объему таблиц.

Mondrian-2.0.1 – наиболее удобный инструмент анализа больших объемов данных, например, таких как ассортимент крупной сети супермаркетов. Грамотное использование возможностей Mondrian и их совершенствование (программный продукт является открытым, т.е. позволяет самостоятельно дорабатывать и расширять его функции) может повысить качество ассортиментной политики любой сети супермаркетов и скорость реакции на изменения факторов внешней среды.

Выводы по результатам исследования. В результате проведенного исследования спроектировано хранилище данных, построены агрегированные таблицы на основе OLAP-продукта Mondrian-2.0.1 и даны конкретные рекомендации по использованию методов OLAP-обработки данных для анализа ассортимента любой сети супермаркетов (на примере ООО «Ален»).

Научная новизна и перспективы дальнейшего развития исследования. В работе впервые предложено использование программного продукта Mondrian-2.0.1 для анализа ассортимента сети супермаркетов. Спроектировано хранилище данных, построена агрегированная таблица и OLAP-куб в Mondrian-2.0.1. Поднятая проблема требует дальнейшего решения путем дополнения, изменения, доработки и расширения функционального состава системы. Также остро стоит проблема русификации программы. Работа в этом направлении авторами уже ведется.

Библиографический список

1. Федоров А., Введение в OLAP-технологии Microsoft / А. Федоров, Н. Елманова. — М.: Диалог — МИФИ, 2002. — 325 с.
2. Методы и модели анализа данных: OLAP и Data Mining / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод. — СПб.: БХВ-Петербург, 2002. — 401 с.
3. Архипенков С.Я. Хранилища данных. От концепции до внедрения / С.Я. Архипенков, Д. Голубев. — М.: Диалог — МИФИ, 2002. — 256 с.
4. Пек Джордж Seagate Crystal Reports. Полное справочное руководство / Джордж Пек. — М.: Лори, 2001. — 345 с.

Поступила в редакцию 27.03.2007 г.

