

Поршнев С.В., Зеленская Е.В.

ОПЫТ ВНЕДРЕНИЯ РАЗДЕЛА «ОСНОВЫ OLAP-ТЕХНОЛОГИИ» В ИЗУЧЕНИИ ДИСЦИПЛИНЫ «БАЗЫ ДАННЫХ»

zel_kat@mail.ru

ГОУ ВПО УГТУ-УПИ

г. Екатеринбург

В статье приводится пример изложения основ OLAP-технологии студентам четвертого курса очного отделения специальности 230101 «Вычислительные машины, комплексы, системы и сети» при изучении дисциплины «Базы данных»

The article exemplifies the presentment OLAP-technology's basics to the 4th course full time students with a specialization in "Computing machinery, complexes, systems and networks" 230101, in studies of the "Data bases' discipline"

В настоящее время все большую популярность получают технологии многомерного анализа данных (OLAP-технологии) и методы нахождения в накопленных данных ранее неизвестных, нетривиальных и практически полезных знаний, необходимых для принятия решений в различных областях человеческой деятельности (Data Mining). Сегодня получили широкое распространение данные технологии в учетных системах торговых предприятий и в банковских системах [1]. Современные технологии Data Mining и OLAP позволяют превращать каждодневно пополняющиеся базы данных из склада мертвых данных на сервере в инструмент анализа и прогнозирования деятельности предприятия.

В нашей стране тенденцию нарастания интереса к этим технологиям легко прослеживается при анализе динамики количества статей по данной тематике публикуемых на сайте olap.ru. Так период с 1997 по 2005 было опубликовано всего 8 статей, в то время как за последнее 2 года более 170, т.е. произошел взрыв интереса к данным технологиям.

Интенсивное внедрение программных продуктов, поддерживающие технологии Data Mining и OLAP, требует подготовки специалистов, владеющих данными технологиями. Мы предлагаем ввести раздел «Основы OLAP-технологии», как завершающий этап в изучении дисциплины «Базы данных».

Для освоения данных технологий можно использовать программный продукт Microsoft SQL Server Analysis Services 2000, который представляет интегрированную среду для создания и работы с OLAP-кубами и моделями Data Mining.

Материал при изучении раздела «Основы OLAP-технологии» излагался в следующей последовательности:

- Введение в хранилища данных и OLAP
 - Концепции хранилищ данных и OLAP, требования к хранилищам данных и OLAP-средствам, логическая организация OLAP-данных, а также основные термины и понятия, применяемые при обсуждении многомерного анализа.

- Модели OLAP, технические аспекты многомерного хранения данных.
- Применение кубов OLAP.
- Понимание архитектуры Microsoft Analysis Services 2000
 - Обзор типичной архитектуры OLAP-служб, рассматриваемой на примере Microsoft Analysis Services – OLAP-сервера фирмы Microsoft, входящего в комплект поставки Microsoft SQL Server 2000.
 - Описание компонентов Analysis Services, репозитория метаданных, опции хранилища куба.
- Построение кубов, размерностей, задание мер
 - Работа с кубами
 - Создание измерений с помощью мастера и редактора.
 - Общие и частные измерения
 - Измерения класса "Время"
 - Стандартные измерения
 - Измерения типа "родитель-потомок"
 - Измерения с несбалансированной иерархией
 - Основные свойства уровня.
 - Работа с мерами.
- Обработка размерностей и кубов
 - Процесс обработки размерностей и кубов
 - Оптимизация процесса обработки кубов
 - Устранение неполадок в процессе обработки.
- Использование Excel в качестве клиента OLAP
 - Компоненты Microsoft Excel OLAP
 - Использование сводных таблиц Excel
 - Использование сводных схем
 - Работа с локальными кубами в виде файлов на основе реляционных данных, доступных с помощью OLE DB
- Работа с шаблонами Data Mining
 - Обзор алгоритмов Data Mining.
 - Алгоритм Microsoft Decision Trees
 - Построение модели на основе OLAP-куба
 - Построение модели на основе реляционной таблицы

Для проведения лабораторных работ была подготовлена база данных Korus, которая описывает деятельность торговой компании «Korus». Компания «Korus» занимается продажей торгового оборудования, которая имеет головной офис и около 30 филиалов по всей России. В базе содержатся сведения о всех продажах за период с 2003 по 2006 годы. Вся продукция классифицирована по категориям, моделям и товарам. На основе этой базы данных надо было построить:

- OLAP-куб продаж в Microsoft Analysis Services 2000 (рис. 1).

Getting Started		Meta Data		Data	
Контрагенты	Все контрагенты				
Номенклатура	Вся номенклатура				
Обобщенные наименования	Все обобщенные наименования				
Признаки документов	Все				
Производители	Все производители				
Склады	Все склады				
ЦФУ	Все ЦФУ				
ЦФУ Товара	Все ЦФУ Товара				

			MeasuresLevel		
- Год	- Квартал	+ Месяц	Количество	Сумма продажи уе	Сумма НДС продажи уе
Все даты	Все даты Total		224 547 812,66	150 908 230,51	3 451 384,07
+ 2003	2003 Total				
- 2004	2004 Total		71 359 482,49	39 144 078,26	922 460,19
	- 1 кв. 2004	1 кв. 2004 Total	14 479 288,87	8 337 438,66	238 432,31
		+ 01.2004	4 646 368,49	2 225 253,15	68 899,44
		+ 02.2004	3 717 062,25	3 035 541,79	68 937,21
		+ 03.2004	6 115 858,13	3 076 643,72	100 595,66
	+ 2 кв. 2004	2 кв. 2004 Total	15 936 139,38	9 286 473,59	198 720,53
	+ 3 кв. 2004	3 кв. 2004 Total	17 316 794,83	9 473 718,04	236 256,68
	+ 4 кв. 2004	4 кв. 2004 Total	23 627 259,42	12 046 447,97	249 050,67

Рис. 1. Куб продаж в Microsoft Analysis Services 2000

- Показать данные из куба в сводной таблице Microsoft Excel (рис. 2) и построить диаграмму (рис.3)

	A	B	C	D	E	F	G	H
1	Признаки документов	Все						
2								
3	Количество	Год	Квартал					
4		2004		2004 Итог	2005		2005 Итог	Общий и
5	Контрагенты 02	1 кв. 2004	2 кв. 2004		1 кв. 2005	2 кв. 2005		
6	Алтайский край	65862	7059	72921	237877	106603	344480	417
7	Бурятия	258531	198879	457410	496972	332463	829435	1286
8	Волгоградская	1370246,23	796441	2166687,23	684074	289478	973552	314023
9	Вологодская	321309,44	196195,25	517504,69	1156341	1122180,25	2278521,25	279602
10	Казахстан	44723	316269	360992	23642	48534	72176	433
11	Кемеровская	563798	361012	924810	178070	216274	394344	1319
12	Краснодарский край	2	89300	89302	12	105735	105747	196
13	Красноярский край	202973	1531937	1734910	315392	294540	609932	2344
14	Курганская	696399,746	676256,68	1372656,426	88158	734372	822530	2195186
15	Московская	1383039,98	951056,76	2334096,74	3584343,5	438101,45	4022444,95	635654
16	Нижегородская	5857	-8917,96	-3060,96	65962	58225	124187	12112

Рис. 2. Куб продаж в Microsoft Excel

Секция 4. Внедрение информационно-коммуникационных технологий в преподавание учебных дисциплин

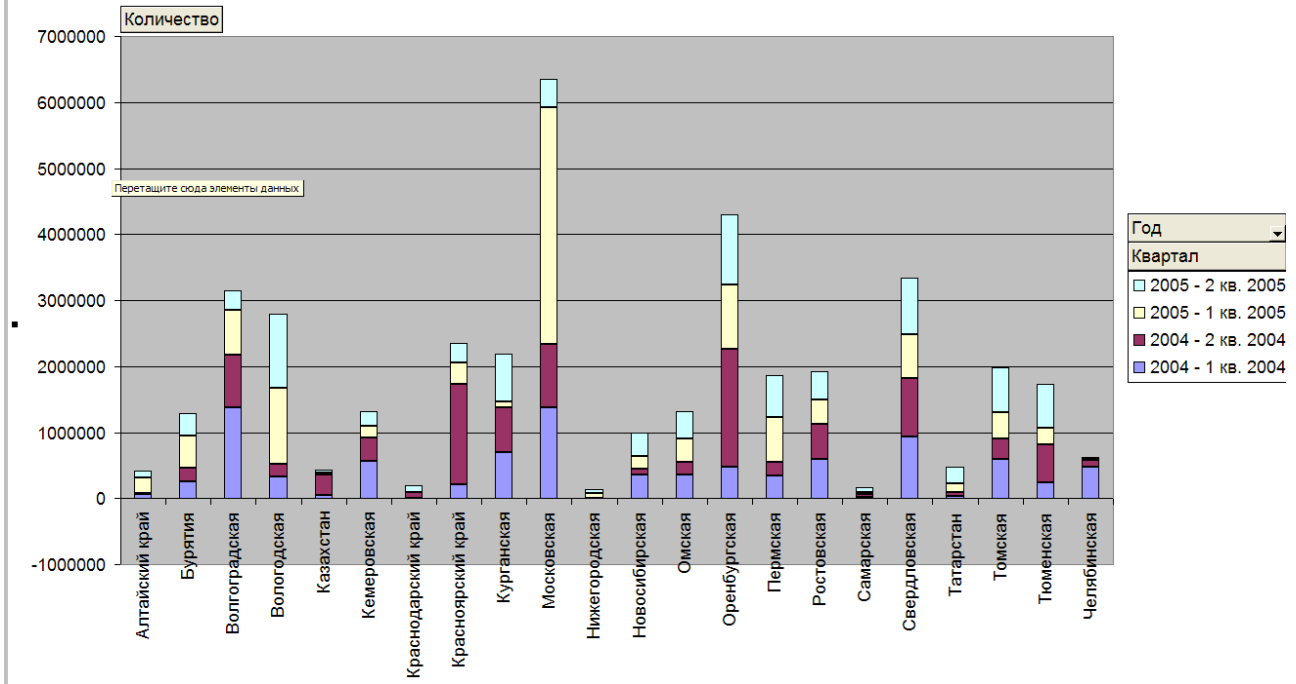


Рис. 3. Диаграмма по данным из куба продаж в Microsoft Excel

- Создать модель Data Mining по клиентам, используя алгоритм Microsoft Decision Trees (рис. 4).

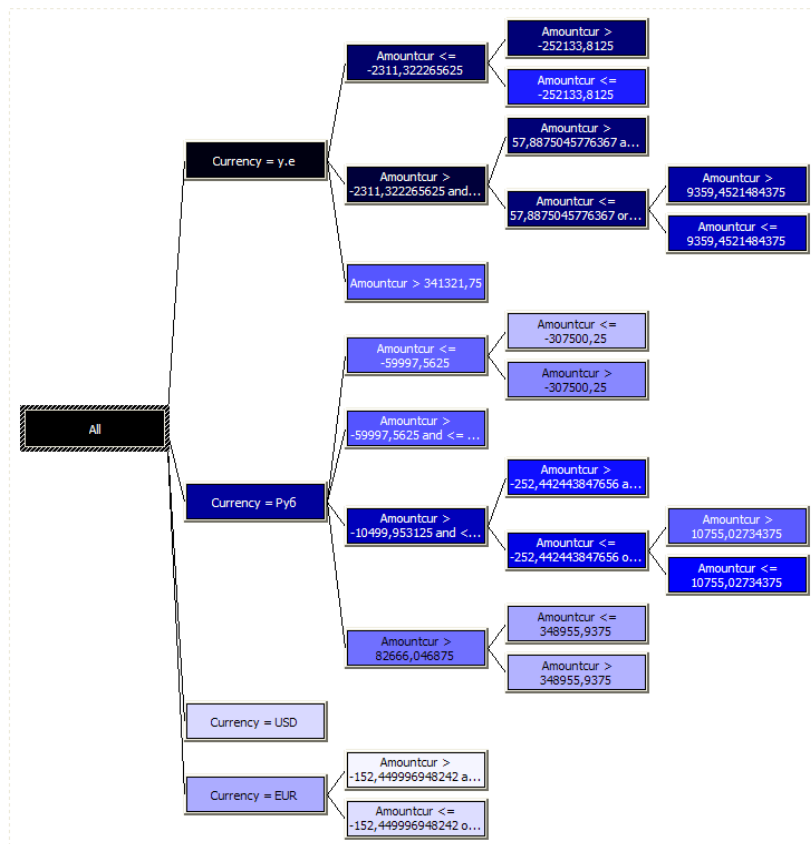


Рис. 4. Модель Data Mining (алгоритм Microsoft Decision Trees)

Данный подход был апробирован на студентах четвертого курса очного отделения специальности 230101 «Вычислительные машины, комплексы, системы и сети» при изучении дисциплины «Базы данных».

Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining. – СПб.: БХВ – Петербург, 2004. – 336 с.: ил.

Поршнев С.В., Параничев А.В.

СТРУКТУРА КУРСА "СТАНДАРТЫ КАЧЕСТВА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ"

dreamworld13@yandex.ru

ГОУ ВПО УГТУ-УПИ

г. Екатеринбург

В статье предлагается и обсуждается структура курса «Стандарты качества программного обеспечения», предложенная авторами.

This paper is proposed and discussed framework of discipline «Standards of Software Quality». The framework was supposed by authors.

Оценивание качества программной продукции в соответствии с положениями международных стандартов, регламентирующих качество программного обеспечения, становится все более актуальной задачей в практике IT-компаний. Это обусловлено не столько причинами этического характера, сколько желанием повысить конкурентоспособность своей продукции.

Отметим, что в России проблема оценивания программного обеспечения стоит особенно остро, поскольку официально принятые международные стандарты в рассматриваемой области официально переводятся на русский язык с опозданием от года до 10-15 лет. С неизбежностью, данная ситуация отражается на качестве подготовки студентов, обучающихся по направлению «Информатика и вычислительная техника».

Для устранения отмеченных недостатков в учебный план подготовки по магистерской программе 230118 «Компьютерный анализ и интерпретация данных» включен курс «Стандарты качества программного обеспечения». При разработке структуры курса принималась во внимание последовательность изложения вопросов, связанных со стандартами качества программного обеспечения, предложенные в [1-3]. Рассмотрим их подробнее.

Авторами [1] предлагается в первую очередь рассматривать программную продукцию с позиции системы менеджмента качества (СМК), затем перейти к изучению стандартов, регламентирующих жизненный цикл (ЖЦ) программных средств (ПС), подробно останавливаясь на вопросах сопровождения и конфигурационного управления программными проектами; здесь также рассматриваются вопросы безопасности программного обеспечения. Далее предлагается рассмотреть модель зрелости процессов в стандартах [12] (отметим, что [12] рег-