

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ВЫЧИСЛИТЕЛЬНЫЕ СЕТИ В ВУЗЕ «ВГУЭС»: ТЕХНОЛОГИЯ OLAP И ЕЕ ПРИМЕНЕНИЕ В EXCEL

■ Василенко К. А.

Информация всегда играла чрезвычайно важную роль в жизни человека. Общеизвестно высказывание о том, что тот, кто владеет информацией, тот владеет и миром. Не является исключением и деятельность предприятия или организации, в том числе и высшего учебного заведения, хранящие большое количество различной информации, требующее наличия специальных хранилищ, средств обработки и анализа.

В наше время высокий уровень развития аппаратных и программных средств позволяет фильтровать, хранить, обрабатывать информацию посредством баз данных. Большую роль для успешной деятельности организации играет оперативный анализ данных. Так, например, возьмем работу одного из структурных подразделений университета — деканат.

Большая часть выполняемой работы связана с анализом контингента обучающихся формированием отчетов в различных разрезах. Практически вся информация хранится в базе «1С: Управление образовательным учреждением». В связи с реорганизацией вуза пришлось отказаться от услуг франчайзера.

База перестала поддерживаться и обновляться, поэтому данные в ней ста-

ли не актуальными. Это очень усложнило работу сотрудников деканата, так как во много раз увеличилось время обработки данных в связи с необходимостью проверки и редактирования информации, получаемой из базы. Анализ данных так же требует соответствующих навыков сотрудников деканата, аппаратного и программного обеспечения. Для всего вышеперечисленного необходимы немалые материальные затраты и время, что проблематично для любого государственного учреждения. Решение проблем анализа было бы возможным, если бы удалось совместить знакомое всем сотрудникам деканата образовательного учреждения программное обеспечение (ПО) с функциями многомерного анализа.

Самым распространенным ПО является Microsoft Office, которое используется повсеместно и не требует специального обучения и крупных затрат. Такая возможность может быть реализована, если обратить внимание на OLAP — встроенный модуль в Microsoft Office Excel. OLAP (Online Analytical Processing) — это система поддержки многомерного оперативного анализа данных. Первым четкое определение OLAP дал в 1993 году Е. Ф. Кодд в статье, которая включала 12 правил.

Позже, в 1995 году, к ним были добавлены еще шесть менее известных правил. Все правила были разделены на четыре группы и названы «характеристиками».

Основные характеристики OLAP включают:

- многомерность модели данных. Частью этого требования считается возможность построения различных проекций и разрезов модели;
- интуитивные механизмы манипулирования данными;
- доступность;
- пакетное извлечение данных;
- архитектура «клиент-сервер»;
- прозрачность;
- многопользовательская работа.

Специальные характеристики:

- обработка ненормализованных данных;
- хранение OLAP результатов отдельно от исходных данных;
- выделение отсутствующих данных;
- обработка отсутствующих значений.

Характеристики построения отчетов:

- гибкое построение отчетов;
- стабильная производительность при построении отчетов;
- автоматическое регулирование физического уровня.

Управление размерностью:

- общая функциональность;
- неограниченное число измерений и уровней агрегирования;
- неограниченные операции между данными различных измерений.

Основными понятиями многомерной модели данных являются:

1. Показатель — это величина (обычно числового типа), которая собственно и является предметом анализа. Это, например, количество студентов, обучающихся на

определенном курсе, количество абитуриентов, сдавших вступительные экзамены, а также любая информация, связанная с контингентом учащихся. Один OLAP-куб может обладать одним или несколькими показателями.

2. Измерение — это множество объектов одного или нескольких типов, организованных в виде иерархической структуры и обеспечивающих информационный контекст числового показателя. Измерение принято визуализировать в виде ребра многомерного куба.
3. Объекты, совокупность которых и образует измерение, называются членами измерений. Члены измерений визуализируют как точки или участки, откладываемые на 145 осях гиперкуба. Например, временное измерение: Дни, Месяцы, Кварталы, Годы — наиболее часто используемые в анализе.

Как уже было сказано, объекты в измерениях могут быть различного типа, например «Студент» – «Национальность» или «Студент» – «Курс». Эти объекты должны быть организованы в иерархическую структуру так, чтобы объекты одного типа принадлежали только одному уровню иерархии.

4. Ячейка — атомарная структура куба, соответствующая конкретному значению некоторого показателя. Ячейки при визуализации располагаются внутри куба и здесь же принято отображать соответствующее значение показателя (рисунок 1).

Применительно к работе деканата, реализация указанного механизма позволяет проводить анализ контингента студентов по различным критериям.

Например, требуется информация о студентах мужского пола, 3–5 курса, чеченской национальности, проживающих в г. Грозный.

Из числа всех студентов отбираются учащиеся 3–5 курса мужского пола. Затем выбираются парни чеченской национальности. Если провести параллель с работой технологии OLAP, то можно сказать, что получается куб данных, из которого выбираются студенты, проживающие в г. Грозный (рисунок 2).

Таким образом, проблема может быть решена при совмещении имеющегося в деканате ПО и функций многомерного анализа. Наиболее целесообразным в данном случае является использование программы Microsoft Office Excel, в которой, начиная с 2010 года, присутствует модуль анализа данных, использующий технологию OLAP. Успешное использование Microsoft Office Excel для многомерного анализа возможно при наличии не только самого приложения,

но и хранилища данных (в частности SQLсервера), а также самих данных, являющихся актуальными на данный момент.

Поскольку используемая ранее программа «1С: Управление образовательным учреждением» работала в клиент-серверном варианте, отказ от нее не приведет к потере данных на сервере и даст возможность для дальнейшего использования этих данных.

По завершении данной работы сотрудники деканата получают возможность осуществления многомерного анализа контингента студентов с наименьшими материальными затратами, максимальной экономией времени, с исключением необходимости специального обучения сотрудников или принятия на работу специалистов-аналитиков по работе с ПО для анализа данных.

Поскольку в настоящее время большая часть сведений находится в базе

	5 курс			
	4 курс			
	3 курс			
		Иванов	Петров	Сидоров
Русский				
Чеченец				
Турок				

Рисунок 1. Куб данных с тремя измерениями

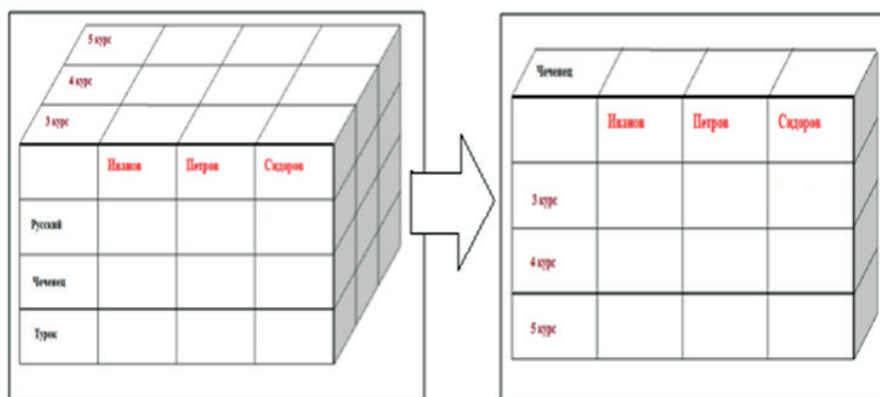


Рисунок 2 Детализация куба данных

ТС:Предприятие, а в дальнейшем предполагается переход на другое ПО (по указанию головного вуза), описание выше изложенного механизма может быть реализовано для упрощения работы деканата, так как и в дальнейшем будет использована технология «клиент-сервер».

Литература

1. Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. Методы и модели анализа данных. OLAP и DataMining: монография. Санкт-Петербург: «БХВ-Петербург», 2016. 335 с.
2. Михеев Р. В. VBA и программирование в Microsoft Office для пользователей: монография. Санкт-Петербург: «БХВ-Петербург», 2014. 364 с.
3. Елманова А. А., Федоров А. В. Введение в OLAP-технологии Microsoft: монография. М.: «Диалог — Диалог МИФИ», 2015. 727 с.
4. Баргер А. И. Microsoft SQL Server 2005 Analysis Services. OLAP и многомерный анализ данных: монография. Санкт-Петербург: «БХВ-Петербург», 2016. 928 с.
5. Громов Ю. Ю., Драчев В. О. Иванова О. Г. Информационная безопасность и защита информации: учебное пособие. Ст. Оскол: ТНТ, 2016. 384 с.
6. Вишневецкий В. М., Ляхов А. И., Портной С. Л., Шахнович И. Л. Широкополосные беспроводные сети передачи информации. М.: Техносфера, 2017. 126 с.
7. Кияев В., Граничин О. Информатизация предприятия. М.: Национальный Открытый Университет «ИНТУИТ», 2016. 235 с.
8. Мерит М., Полино Д. Безопасность беспроводных сетей /пер. с англ. А. В. Семенова. М.: Компания АйТи, ДМК Пресс, 2017. 288 с.