

МНОГОМЕРНАЯ МОДЕЛЬ ДАННЫХ ДЛЯ ПОСТРОЕНИЯ OLAP-КУБА

THE MULTIDIMENSIONAL DATA MODEL FOR BUILDING OLAP-CUBE

Макалада OLAP-кубту кураштыруу үчүн маалыматтардын көп өлчөмдүү моделдери каралган. Өлчөө, көп өлчөмдүү маалыматтардын үстүнөн негизги амалдар жана колдонуучулар үчүн функциялар талкууланган.

Ачык сөздөр: OLAP, көп өлчөмдүү анализ, өлчөө, кесилиш, кросс-таблица, консолидаация, детализация.

В статье рассматриваются многомерные модели данных для построения OLAP-куба. Обсуждены измерения, основные операции над многомерными данными и функции для пользователей.

Ключевые слова: OLAP, многомерный анализ, измерение, срез, кросс-таблица, консолидация, детализация.

The article deals with the multidimensional data model for building OLAP-cube. Discussed measurement, basic operations on multidimensional data and user function.

Keywords: OLAP, multidimensional analysis, measurement, section, cross table, consolidation, specification.

В процессе принятия решений пользователь генерирует некоторые гипотезы, которые необходимо проверять. Проверка гипотезы может быть осуществлена на основе информации об анализируемой предметной области. Причем не представляется возможным заранее предусмотреть, какие запросы будет формировать пользователь, т.к. генерация этих запросов производится «на лету».

К тому же требуется, чтобы анализ данных можно было осуществлять по множеству параметров, т.е. должна быть предусмотрена возможность выполнения многомерного анализа данных.

В связи с этим встает вопрос о том, как организовать данные, описывающие предметную область так, чтобы по ним было удобно проводить многомерный анализ.

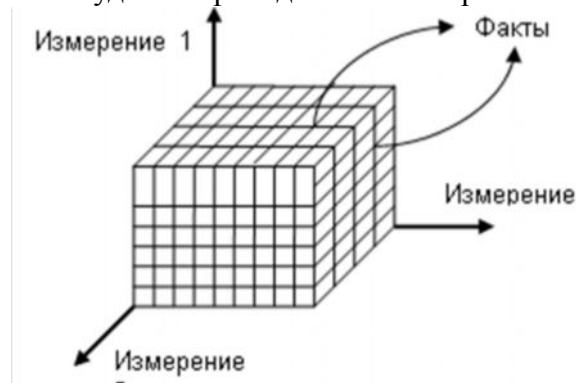


Рис.1. Графическое представление многомерной модели данных

Такой анализ данных целесообразно проводить на основе многомерной модели данных, которая лежит в основе построения большинства хранилищ данных.

Многомерная модель данных – это информационное содержимое предметной области в виде N-мерного куба (гиперкуба), где каждая ось соответствует измерению, представлено на рисунке 1.

Измерение – это последовательность значений одного из анализируемых параметров. На пересечении осей измерений находятся факты (меры), количественно характеризующие события, описываемые значениями измерений.

Многомерное представление данных позволяет легко выполнять ряд операций.

1. Срез – формирование подмножества данных, соответствующих фиксированному значению одного или нескольких измерений, т.е. построение проекции гиперкуба, представлено на рисунке 2.

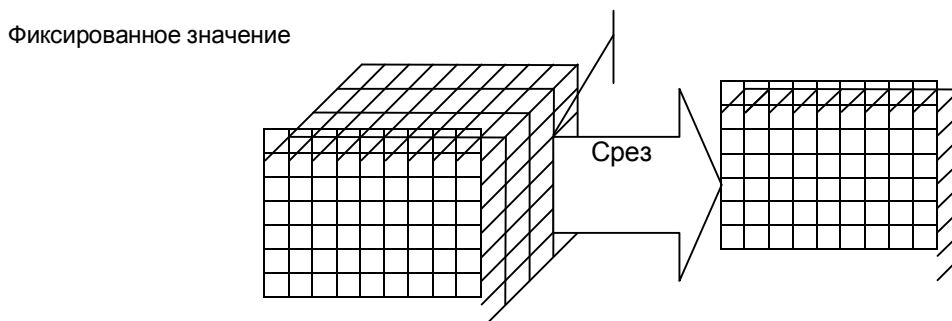


Рис.2. Операция среза многомерной модели данных

2. Вращение – изменение расположения измерений в представлении данных многомерной модели, представлено на рисунке 3.

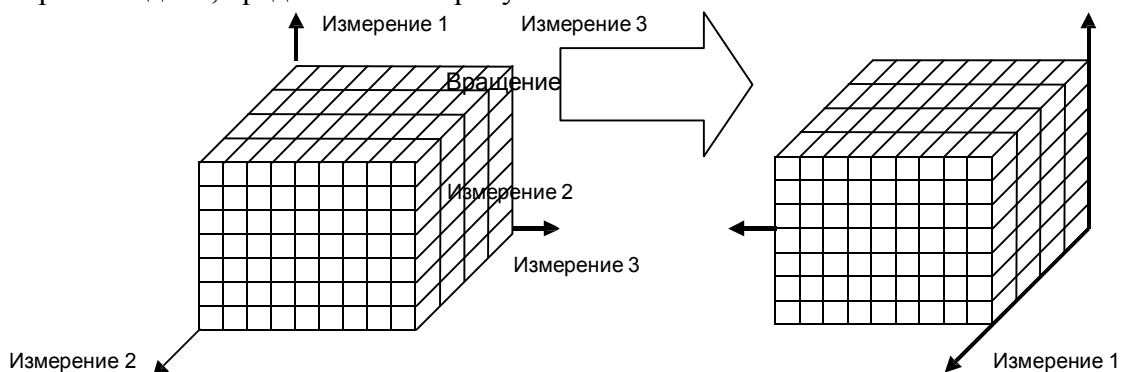


Рис. 3. Операция вращения многомерной модели данных

Операция вращения обычно применяется при отображении многомерных данных в отчете в виде кросс-таблиц или кросс-диаграмм.

Кросс-таблица представляет собой размещение многомерных данных на плоскости в виде сводной таблицы.

Измерения могут быть в строках и столбцах кросс-таблицы, а факты с применением какой-либо агрегирующей функции отображаются на пересечении измерений, т.е. на пересечении строк и столбцов.

В качестве агрегирующих функций чаще всего используются: минимум (min), максимум (max), сумма(sum), среднее (average), количество(count).

Кросс-таблица состоит из следующих элементов:

- заголовки строк;
- заголовки столбцов;
- матрица с фактами.

Изобразить геометрическую интерпретацию гиперкуба с размерностью более 3-х не представляется возможным. Но в принципе число используемых измерений может быть любым.

Однако следует помнить, что использование большого числа измерений не рекомендуется, в первую очередь потому, что осмысление и интерпретация результатов аналитиком в этом случае могут быть затруднены.

3. Консолидация и детализация – переход от детального представления данных к агрегированному (обобщенному) и наоборот, отражено на рисунке4 и рисунке5.

При выполнении операции детализации отображаются все записи, которые вносят свой вклад в формирование значения выделенной ячейки, имеющей агрегированное значение в многомерной модели.

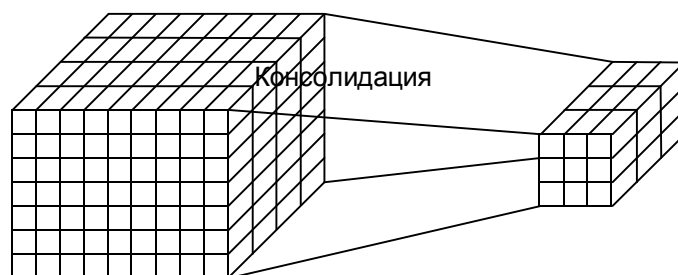


Рис. 4. Операция консолидации многомерной модели данных

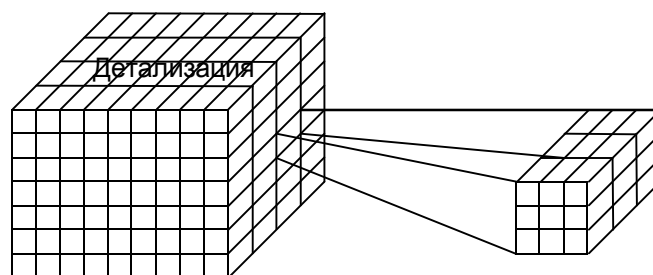


Рис. 5. Операция детализации многомерной модели данных

По кросс-таблицам, полученным в результате консолидации, строятся кросс-диаграммы, позволяющие анализировать тренд.

Тренд позволяет увидеть тенденции, которые обычно скрыты из-за большого разброса значений, наличия отклонений, не типичных для отображаемого процесса.

Таким образом, представленные операции многомерной модели данных предоставляют конечному пользователю ряд аналитических и навигационных функций:

- расчеты и вычисления по нескольким измерениям;
- анализ трендов;
- выборка подмножеств данных для просмотра;
- получение обобщенных (агрегированных) значений;
- переход к детальным данным, лежащим в основе анализа;
- вращение кросс-таблиц отображаемых данных.

Список литературы

1. Мак-Дональд М., Шпушта М. MicrosoftASP.NET 2.0 с примерами на C# [Текст] / М.Мак-Дональд. – М.: Вильямс, 2007. – 1408 с.
2. Малик С. MicrosoftADO.NET 2.0 [Текст] С.Малик. – М.: Вильямс, 2007. – 560 с.
3. Виейра Р. Программирование баз данных Microsoft SQL Server 2005 [Текст] / Р.Виейра. – М.: Диалектика, 2007. – 832 с.
4. Для подготовки данной работы были использованы материалы с сайтов:
[-http://www.sourceforge.net/projects/zedgraph/](http://www.sourceforge.net/projects/zedgraph/)
[-http://www.microsoft.com/downloads/details.aspx?FamilyId=7287252C-402E-4F72-97A5-E0FD290D4B76&displaylang=en](http://www.microsoft.com/downloads/details.aspx?FamilyId=7287252C-402E-4F72-97A5-E0FD290D4B76&displaylang=en)