

СПОСОБЫ И МЕТОДЫ АНАЛИЗА МНОГОМЕРНОГО БАЗЫ ДАННЫХ

*Эргашев Аслон Акрамович**ст. преподаватель,
Бухарский государственный университет,
Республика Узбекистан, г. Бухара
E-mail: aslonbek2008@gmail.com**Садикова Фируза Сафаровна**ассистент,
Бухарский государственный университет,
Республика Узбекистан, г. Бухара*

METHODS AND METHODS FOR ANALYSIS OF MULTIDIMENSIONAL DATABASE

*Ergashev Aslon Akramovich**Senior Lecturer,
Bukhara State University
Republic of Uzbekistan, Bukhara**Sadikova Firuza Safarovna**Assistant,
Bukhara State University
Republic of Uzbekistan, Bukhara*

АННОТАЦИЯ

В статье рассматриваются возможности многомерных баз данных, в частности многомерных кубов OLAP (On-Line Analytical Processing) при анализе больших объемов данных. Предоставлен обзор и особенности многомерной базы данных, обсуждаются действия, которые необходимо выполнить с многомерными базами данных для понимания структуры и возможностей куба OLAP. Для создания базы знаний описаны шаги, которые можно предпринять для создания и выполнения многомерной базы данных, которую можно собрать из различных источников, сохраняя в базу данных, а затем подготовить отчет с помощью анализа OLAP.

ABSTRACT

The article discusses the capabilities of multidimensional databases, in particular, multidimensional OLAP (On-Line Analytical Processing) cubes when analyzing large amounts of data. Provides an overview and features of a multidimensional database and discusses the steps you need to take with a multidimensional database to understand the structure and capabilities of an OLAP cube. To create a knowledge base, it describes the steps you can take to create and execute a multidimensional database that you can collect from various sources, save to a database and then prepare a report using OLAP analysis.

Ключевые слова: многомерные базы данных, OLAP-куб, сечение куба, операции выполняемые с кубом OLAP.

Keywords: multidimensional databases, OLAP cube, section of a cube, operations performed with an OLAP cube.

Введение. Многомерный анализ данных используется в наборе программных инструментов для принятия управленческих решений. Основан на базе знаний, возникающей при управлении процессами в конкретной системе и в результате обработки информации [1]. Принятие верного решения требует упрощения структуры и использования большого количества информации. Использование OLAP-кубов для анализа и визуализации больших данных является эффективным.

Цель методики и исследования. Определить принцип, свойства многомерного анализа данных, рассмотреть концепцию OLAP-куба, определить

измерения и атрибуты для анализа данных, а также определение зависимости размера (атрибута) и числовых значений (измерений) в многомерной базе данных по принципу OLAP-кубов, а также получение необходимой информации в виде простой двухмерной таблицы.

Благодаря высокой эффективности и простоте использования многомерные базы данных широко используются на практике на современных предприятиях и в организациях [5, стр. 2,3]. Многомерная база данных - это модель данных, как взаимосвязанная многомерная структура, подобная реляционной структуре [9].

Многомерная база данных. Многомерная база данных способна обрабатывать очень большие объемы данных с бесконечным числом полей [9]. Данные представляется в виде упорядоченного массива, что очень эффективно, поскольку поиск данных выполняется с помощью небольших блоков [8].

В многомерной базе данных данные представлены в виде числовых параметров или текстовых свойств, которые помогают принять правильное решение в результате анализа данных [8].

Есть три основных особенности многомерной базы данных:

1. На основе данных, собранных из разных источников, база данных просматривается и объединяется для анализа.
2. Новая информация находится путем разбивки выборок и поиска взаимосвязей.
3. Система OLAP отделяет знания от больших объемов данных [8,3].

Многомерный куб данных. OLAP-кубы. Многомерная база данных - представляют собой многомерный куб данных [7].

Структура простых электронных таблиц не меняется, в отличие от таблиц любого размера и иерархии, представленных в виде многомерных кубов. Несколько кубов образуют многомерную базу данных, а ячейки куба представляют собой измеренные значения [7, 2].

Многомерные OLAP-кубы помогают различать различные комбинации данных и анализировать данные, доступные в виде блока данных и извлекаемые по запросу [1]. Они также позволяет добавлять и редактировать новые данные, как реляционной базе данных.

Различные блоки образуют OLAP-куб, который упрощает направление и взаимосвязь данных [6]. Данные представлены в статистической форме, что позволяет выполнять различные действия.

Кубы могут быть бесконечными по размеру, но не трехмерными, что приводит к снижению быстродействия компьютера [11].

OLAP-куб можно рассматривать как логическую модель для представления многомерных данных, характеризуемых индикаторами и измерениями: $G = \langle D, F \rangle$ - гиперкуб (рисунок 1).

$F = \langle f_1, f_2, \dots, f_n \rangle$ - индикаторы гиперкуба (измерения): каждый индикатор имеет набор значений, которые количественно определяют анализируемый процесс.

$D = \langle d_1, d_2, \dots, d_m \rangle$ - измерения гиперкуба: каждое измерение представляет собой упорядоченный набор значений определенного типа. Измерения можно организовать в упорядоченную иерархическую структуру. Оси гиперкуба создаются из набора измерений:

$$d_1 = d_1^1, d_1^2, \dots, d_1^{k_1}$$

$$d_2 = d_2^1, d_2^2, \dots, d_2^{k_2}$$

.....

$$d_n = d_n^1, d_n^2, \dots, d_n^{k_n}$$

Одним из основных требований технологии OLAP является «прозрачность»: готовый многомерный куб должен быть представлен в удобной для пользователя форме, средства управления кубом должны быть интуитивно понятными, названия объектов анализа должны соответствовать терминологии.

Процесс OLAP-анализа характеризуется следующим набором действий с многомерными данными - консолидация, агрегированием (группировкой), вырезанием и вращением.

Формирование части куба заключается в определении значения (значений) определенного измерения, в котором размер куба уменьшается. Часть куба - это внутренний куб, содержащий все остальные измерения. Операция вращения заключается в изменении направления осей куба.

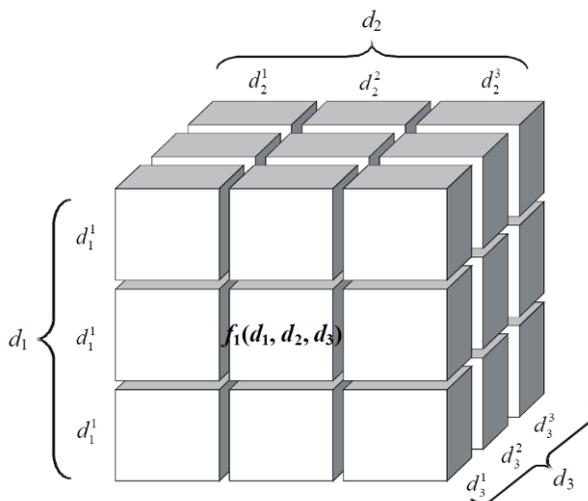


Рисунок 1. OLAP-куб

Для конечного пользователя обеспечивается высокая гибкость в принятии решений за счет возможности изменения результирующего внешнего вида OLAP-куба. Аналитику предоставляется не строго регламентированный отчет, а возможность использовать набор инструментов для творческого исследования проблемы. Свободное манипулирование данными позволяет легко получить требуемый набор данных.

OLAP-куб можно рассматривать как абстрактное представление выбранного подмножества реляционной базы данных [1].

Структура OLAP. Основными элементами куба OLAP являются:

1. Измерения или атрибуты - при измерении значения любого параметра откладываются на потом (это оболочка куба) [2].
2. Кубы разного размера (более трех) можно представить в виде иерархии структурных представлений, расположенных друг внутри друга [1].
3. Иерархия измерений представлена в виде структуры для создания агрегированных данных.
4. Измерение или числовые значения – это кубические данные (количественное описание измерения) для определенных значений.
5. Символы. Состоят из различных размеров внутри куба, имеющих символическое обозначение.

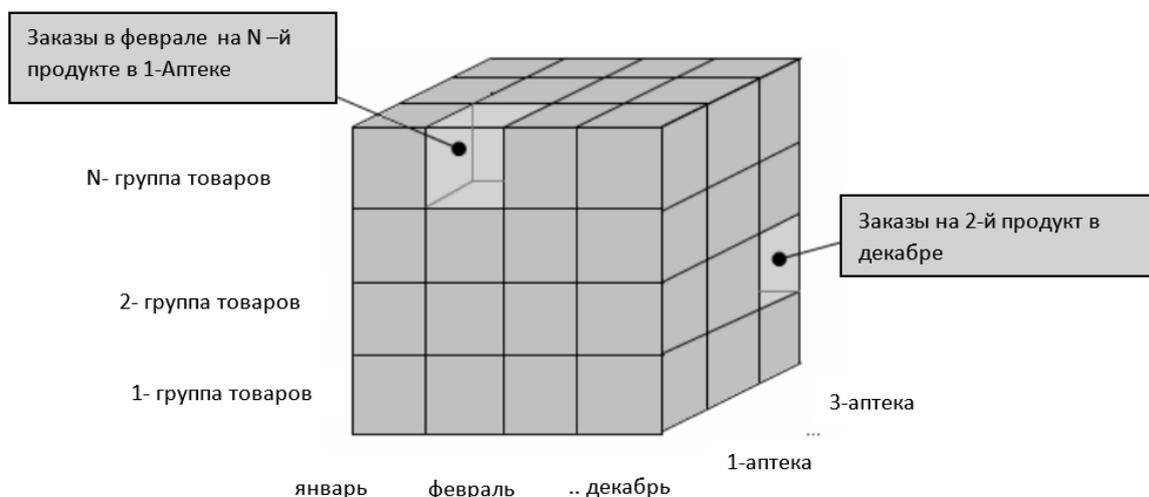


Рисунок 2. Схематическое изображение OLAP-куба

На рисунке 2 показаны следующие измерения: «Месяцы» на одной оси, «Продукты» на второй оси и «Точки продаж» на третьей оси. Ячейка, соответствующая точке пересечения осей, является измеренным значением.

OLAP-куб позволяет агрегировать не только хранилище данных, но также и сортировать необходимые данные. С помощью куба выполняются следующие операции:

1. Разделение части - небольшой набор кубов в соответствии с необходимыми размерами.
2. Преобразование позволяет изменить визуальный вид данных путем изменения размеров.
3. Детализация - представление конкретных данных.
4. Консолидация - предоставление обобщенной информации [2].

Качество анализа данных на основе OLAP-куба в основном определяется наличием исходных данных и прозрачностью аналитической модели предметной сферы. Аналитическая модель - это набор локальных OLAP-моделей, имеющих соответствующий маршрут данных для решения определенных проблем [1, 2].

Заключение. Многомерные базы данных и их средства являющиеся хранилищем данных образуют OLAP-структуру, которая должна оперативно анализировать имеющиеся у нее данные. Таким образом, OLAP-анализ должен отвечать следующим требованиям:

- 1) Анализ должен осуществляться быстро независимо от количества данных.
- 2) Способность выполнять любой анализ (статистический, математический).
- 3) Доступ к нескольким компьютерам, то есть различные пользователи могут одновременно работать с одной базой данных.
- 4) Возможность обращения к необходимой информации в любое время.
- 5) Наличие бесконечного количества измерений для работы с большими объемами данных.

Перечисленные требования соответствуют принципу онлайн-обработки многомерных баз данных. Качество анализа данных на основе OLAP-куба в основном определяется наличием исходных данных и прозрачностью аналитической модели в области изучаемого предмета. Аналитическая модель - это набор локальных OLAP-моделей с соответствующей витриной данных (Data Mart) для решения определенных проблем [7].

Список литературы:

1. Методы и модели анализа данных: OLAP и Data Mining / А.А. Барсегян [и др.]. - СПб. : БХВ-Петербург, 2004.
2. Методы интеллектуального анализа данных при создании баз знаний / В.С. Абриков [и др.] // Вестн. Чуваш. ун-та. – 2015. – № 1. – С. 140–146.
3. Демченко А.А. Использование OLAP-технологий при обработке данных // Решетневские чтения. Информационно-управляющие системы. - 2014. - С. 185-186.
4. Демченко А.А., Молоков В.В. OLAP-технология анализа данных // Секция «Информационно-управляющие системы». - 2014. - № 2/2. - С. 332-333.
5. Кокоулин А.Н., Южанинов Р.И. Многомерный анализ данных по обращаемости в лечебные учреждения с помощью средств Oracle Olap // Вестник ПНПУ. Электротехника, информационные технологии, системы управления. 2015. - № 13. - С. 5-14.
6. Коробко А.В., Пенькова Т.Г. Представление и применение знаний о кубах –концептах для поддержки адаптации манипулирования объектами анализа OLAP// Вестник СибГАУ. - № 30 (49). - 2013. - С. 51-57.
7. Терещенко О.В., Курилович Н.В., Князева Е.И. Многомерный статистический анализ данных в социальных науках : учеб. пособие. – Минск : БГУ, 2012. 239 с. : ил.
8. Шешукова Т.Г. Многомерный анализ данных: теория и практика // Экономический анализ: теория и практика. – № 17 (75). – 2006. – С. 8–13
9. Эргашев А.А. Bigdata: Бугунги салмоқли маълумотлар таҳлили. //«Иновацион ғоялар, ишланмалар ва уларни ишлаб чиқариш ҳамда таълимда қўллашнинг замонавий муаммолари» халқаро илмий-амалий конференция. 2019 йил 15 апрель, Андижон.
10. Эргашев А.А. Выбор паттерна проектирования автоматизированной// информационной системы. // Журнал Проблемы науки. -2021 - 6`65. -С 17
11. Codd E.F. Providing OLAP to user-analysts: An IT mandate : technical report. San Jose : Codd and Date, 1993.
12. Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab, and Sub-Totals / J. Gray, A. Bosworth, A. Layman, H. Priahesh // Proc. of the 12th Intern. Conf. on Data Engineering. IEEE, 1995. P. 152–159.
13. Ergashev A.A., Eshankulov H.I. Bilimlarni tasvirlashda freymli modellardan foydalanish.// Buxoro davlat universiteti Ilmiy Axboroti jurnali. - 2019/4. 92-b
14. Korobko A., Penkova T. OLAP-modeling of municipal procurement automation support problem // Proc. Intern. Conf. on Conceptual Structures (ICCS'09). 2009. P. 87–91.