

# Использование OLAP-технологий для эффективного анализа данных ЛПУ

З.М. Аминова

Факультет информатики и робототехники

Уфимский государственный авиационный технический университет

Уфа, Россия

zukhrusha1994@gmail.com

## Аннотация<sup>1</sup>

В статье проводится обзор и анализ деятельности лечебно-профилактического учреждения (ЛПУ). Описываются методы и технологии создания системы поддержки принятия решения для решения задачи анализа данных ЛПУ. Проводится обзор концепции хранилища данных, OLAP-технологий и OLAP-кубов, использование Power BI в качестве инструмента визуализации и анализа данных ЛПУ.

## 1. Введение

Компьютерные и информационные технологии уже давно стали неотъемлемой частью самых разных сфер жизни, и медицина не стала исключением. Многие частные и государственные лечебно-профилактические учреждения (ЛПУ) уже используют или активно внедряют медицинские информационные системы (МИС) для автоматизации деятельности. Российский рынок МИС в настоящее время представлен немалым количеством участников и решений. Наиболее востребованными являются МИС, направленные на комплексную автоматизацию ЛПУ и объединяющих несколько взаимосвязанных подсистем, таких, как электронная история болезни (ЭИБ) или амбулаторная карта пациента, расписание рабочего времени персонала, бухгалтерия, статистика и др. Использование МИС решает проблему производства и накопления больших объемов информации, но одновременно с этим создает новые задачи, связанные с анализом накопленной информации для извлечения новых знаний [1].

## 2. Обзор предметной области

В работе под ЛПУ будем подразумевать многопрофильное медицинское учреждение, в котором оказываются услуги по направлениям:

---

Труды Шестой всероссийской научной конференции "Информационные технологии интеллектуальной поддержки принятия решений", 28-31 мая, Уфа-Ставрополь, Россия, 2018

педиатрия, терапия, урология, физиотерапия, функциональная и лучевая диагностика, хирургия и пр. Более 5 лет в ЛПУ используется МИС Медиалог, предназначенная для ведения электронной медицинской карты (ЭМК). В результате, на сегодняшний день вся работа сотрудников клиники ведется в единой МИС: работа с пациентами, ведение медицинских карт, проведение лабораторных исследований, работа со страховыми компаниями, оформление больничных листов, ведение кассового учета, работа Call-центра и др.

Организационная структура рассматриваемого учреждения не имеет отдельного подразделения по анализу деятельности организации, соответственно получение и анализ данных выполняется непосредственно заведующими и начальниками отделов. Для каждого из руководителей подразделений настроены рабочие места в МИС МЕДИАЛОГ для осуществления контроля за работой своего подразделения.

Создание отчетов производится во встроенном модуле «Статистика», который предоставляет пользователю возможность создать любое количество параметризуемых отчетов различной сложности, от простых списков до сложных таблиц с подсуммами и многоуровневой группировкой данных, и затем использовать их в разных модулях программы или экспортировать результаты в *Microsoft Excel* и другие форматы [2].

Наиболее актуальными и часто выполняемыми являются отчеты, в состав которых входят следующие показатели:

- Выручка по всей организации, по подразделениям, по врачам;
- Количество оказанных услуги, их состав и стоимость всей организации, по подразделениям, по врачам;
- Динамика посещаемости учреждения и др.

Наряду с вышеперечисленными преимуществами встроенный статистически-отчетный модуль МИС Медиалог имеет существенные ограничения.

Создание и редактирование какого-либо отчета возможно лишь программистами IT-отдела, отсутствует возможность графического представления данных, а для получения различных срезов и уровней детализации данных необходимо создавать новые отчеты.

Целью работы является создание системы поддержки принятия решения (СППР), которая предоставит возможность комплексного анализа *финансовых* и *медицинских* показателей деятельности ЛПУ, их динамики, тенденции и т.д.

## 2.1. Обзор и анализ предлагаемого решения

Основная задача СППР — предоставить аналитикам инструмент для выполнения анализа данных. Система не генерирует правильные решения, а только предоставляет аналитику данные в соответствующем виде (отчеты, таблицы, графики и т. п.) для изучения и анализа, именно поэтому такие системы обеспечивают выполнение функции поддержки принятия решений [3].

Задача разработки эффективной СППР также существует и в сфере здравоохранения. Большинство информационных потоков в данной сфере сконцентрировано вокруг МИС. Такие системы являются основным компонентом и обеспечивают сбор и накопление данных, среди которых как медицинская информация о состоянии пациента, так и финансовые показатели ЛПУ.

Разрабатываемая СППР для рассматриваемого в статье ЛПУ должна предоставлять конечному пользователю возможность настройки отображения отчетов, иметь возможность графической формы представления данных. Также важно наличие инструмента анализа в СППР, который позволит вскрывать причины отрицательных показателей работы учреждения в целом или его отдельных структурных подразделений, планировать деятельность учреждения на следующих период.

Сбор и хранение информации в МИС Медиалог реализуется средствами СУБД MS SQL Server, в основе которой OLTP-подсистема обработки данных [2]. Непосредственно OLTP-системы не подходят для полноценного анализа информации в силу противоречивости требований, предъявляемых к OLTP-системам и СППР. Поэтому для объединения в одной системе OLTP и СППР для реализации подсистемы хранения используются концепция хранилищ данных (ХД). В основе концепции ХД лежит идея разделения данных, используемых для оперативной обработки и для решения задач анализа, что позволяет оптимизировать структуры хранения [3].

ХД - предметно-ориентированный, интегрированный, редко меняющийся, поддерживающий хронологию набор данных, организованный для целей поддержки

принятия решений. Предметная ориентация означает, что ХД интегрируют информацию, отражающую различные точки зрения на предметную область. Интеграция предполагает, что данные, хранящиеся в ХД, приводятся к единому формату. Поддержка хронологии означает, что все данные в ХД соответствуют последовательным интервалам времени.

Кроме возможности работать с единым источником информации, руководители и аналитики должны иметь удобные средства визуализации данных, агрегирования, поиска тенденций, прогнозирования. Несмотря на многообразие аналитической деятельности можно выделить типовые технологии анализа данных, каждой из которых соответствует определенный набор инструментальных средств. Вместе с хранилищем данных эти средства обеспечивают полное решение для автоматизации аналитической деятельности и создания корпоративной информационно-аналитической системы [3].

Анализ накопленных данных в ХД реализуется при помощи OLAP. Данная технология интерактивной аналитической обработки была выбрана по причине того, что может выполнять сложные и разнообразные функции, включающие сбор данных из различных источников, их согласование, преобразование и загрузку в хранилище, хранение аналитической информации, регламентную отчетность, поддержку произвольных запросов, многомерный анализ и др. OLAP позволяет извлекать, визуализировать информацию и получать данные в различных срезах. Традиционные отчеты лишены такой «гибкости», поскольку их нельзя «покрутить», «развернуть» или «свернуть», чтобы получить необходимое представление и детализацию данных [4].

В общем виде архитектура корпоративной OLAP-системы описывается схемой с тремя выделенными слоями (рис. 1).

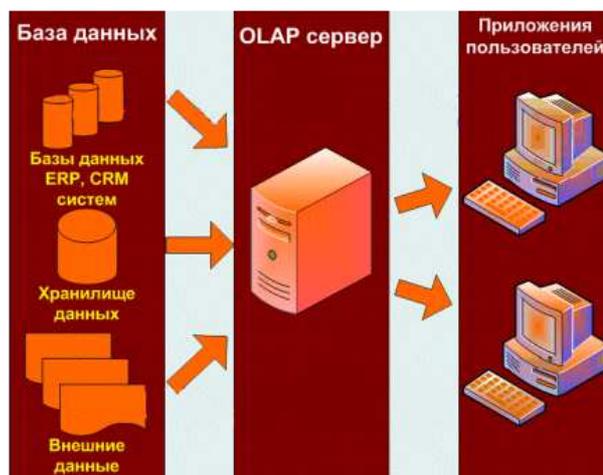


Рис. 1. Архитектура OLAP-системы

Она состоит из следующих элементов:

- база данных. База данных является источником информации для работы OLAP системы. Вид базы данных зависит от вида OLAP системы и алгоритмов работы OLAP сервера. Как правило, используются реляционные базы данных, многомерные базы данных, хранилища данных и т.п.
- OLAP сервер. Он обеспечивает управление многомерной структурой данных и взаимосвязь между базой данных и пользователями OLAP системы.
- пользовательские приложения (OLAP-клиент). Этот элемент структуры OLAP системы осуществляет управление запросами пользователей и формирует результаты обращения к базе данных (отчеты, графики, таблицы и пр.)

OLAP организует данные в виде многомерных кубов (OLAP-кубов), где осями служат основные атрибуты анализируемого бизнес-процесса, а на пересечении осей – данные, количественно характеризующие процесс (меры).

Таблица фактов - является основной таблицей хранилища данных. Как правило, она содержит

сведения об объектах или событиях, совокупность которых будет в дальнейшем анализироваться. Для многомерного анализа пригодны таблицы фактов, содержащие как можно более подробные данные (то есть соответствующие членам нижних уровней иерархии соответствующих измерений).

Таблицы измерений содержат неизменяемые либо редко изменяемые данные. В подавляющем большинстве случаев эти данные представляют собой по одной записи для каждого члена нижнего уровня иерархии в измерении. Таблицы измерений также содержат как минимум одно описательное поле (обычно с именем члена измерения) и, как правило, целочисленное ключевое поле (обычно это суррогатный ключ) для однозначной идентификации члена измерения.

Создание хранилища данных является одной из подзадач выполняемой работы, так как в ЛПУ аналогичной технологии накопления данных нет. В процессе создания ХД были использованы службы DTS MS SQL Server, при помощи которых была спроектирована архитектура ХД, организован импорт, экспорт и преобразование данных. ХД спроектировано по схеме «Снежинка», которая позволяет минимизировать избыточность данных и более эффективно выполнять запросы, связанные со структурой значений измерений (рис. 2).

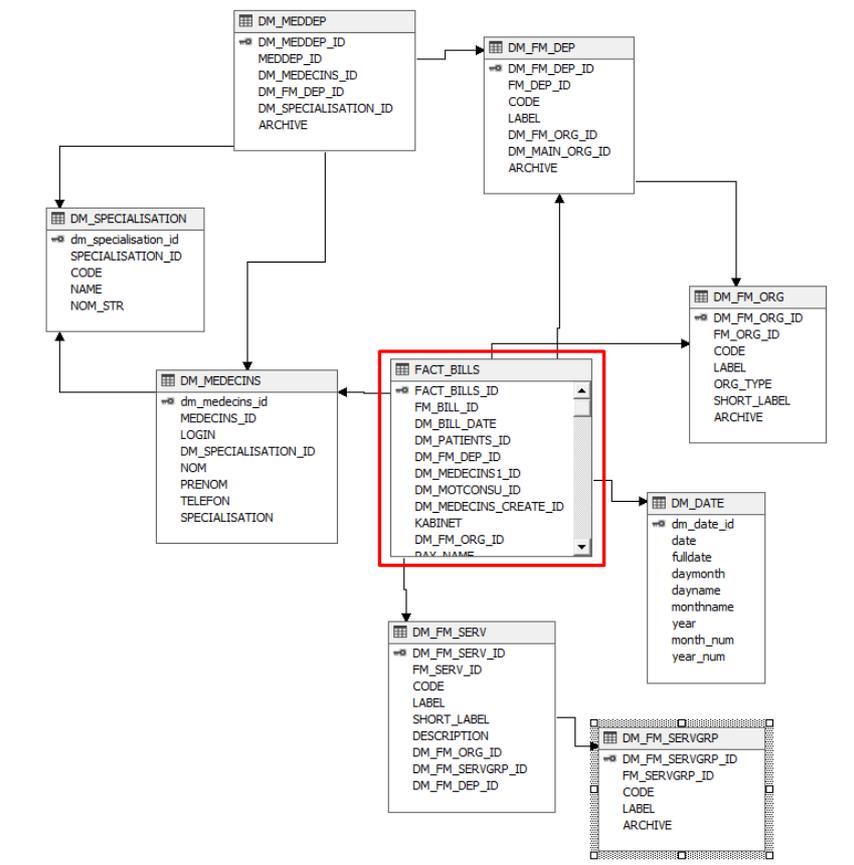


Рис. 2. Схема "Снежинка"

В качестве таблиц-фактов были выбраны таблицы, отражающие финансовые и медицинские показатели ЛПУ, такие как оказанные услуги, количество и частота посещений, статистика обращений пациентов по нозологиям, возвращаемость пациентов и др.

Для работы с OLAP используется служба Microsoft SQL Server Analysis Services, которая обеспечивает высокую производительность работы приложений и масштабируемость на уровне миллионов записей и тысяч пользователей, а также предоставляет возможность построения OLAP- кубов и решений.

За счет высокой масштабируемости службы SSAS позволяют работать с терабайтными базами данных, и с тысячами пользователей. Кубы служб SSAS - это многомерные структуры, обеспечивающие высокоскоростной доступ к большим объемам предварительно объединенных данных, и позволяющие конечным пользователям получать интересующие их бизнес-данные в реальном времени.

В службах SSAS хранятся бизнес-данные в формате MOLAP, предоставляющем возможность высокой степени оптимизации и сжатия. Присущая этому формату гибкость дает также возможность частично или полностью хранить данные в реляционной базе данных в режиме реляционного OLAP (ROLAP) [5].

- Использование OLAP в ЛПУ решает следующие задачи:
- Предоставление пользователю результатов анализа за приемлемое время;
- Многопользовательский доступ к данным;
- Графическое представление данных;
- Получение необходимых срезов данных конечным пользователем, без помощи программистов и др.

Спроектированные OLAP-кубы рассматриваемого ЛПУ имеют такие оси (измерения), как время, филиал, отделение, врачи и пользователи, специальности, наименования услуг, нозологии и пр. В качестве мер вычисляются объемы оказанных услуг, выручка, количество пациентов, издержки, количество обращений и т.п. [6]

BI-платформа Microsoft предоставляет возможности интеграции с продуктами семейства Microsoft Office System, в частности:

- Microsoft Office Excel дает возможность просматривать данные, хранящиеся в OLAP-кубах SSAS путем построения динамических представлений Microsoft PivotTable, что не требует установки дополнительного программного обеспечения;

- Microsoft Office Word, как и Microsoft Office Excel, позволяют просматривать отчеты, генерируемые при помощи Reporting Services;
- Microsoft Office Visio позволяет визуализировать деревья решений, деревья зависимостей, кластерные диаграммы и другие модели технологии data mining;
- Microsoft Office SharePoint Server позволяет создать единый пользовательский интерфейс для просмотра и управления отчетами, генерируемыми при помощи Reporting Services [4].

Такие инструменты позволяют просматривать данные, визуализировать деревья решений, деревья зависимостей, создавать единый интерфейс для просмотра и управления отчетами [6].

В качестве инструмента визуализации и просмотра данных был выбран относительно новый продукт Microsoft – облачная служба бизнес аналитики Power BI, использование которой обладает рядом преимуществ: бизнес-аналитика дает полное представление о деятельности компании в нужный момент; доступ к данным с любого устройства, так кроме конструктора отчетов существуют мобильная и веб-версии для просмотра информации; вместо статистических отчетов идет работа с визуализированной информацией; широкий выбор способов визуализации данных и возможность создания собственных; управление параметрами безопасности происходит централизованно, что облегчает работу пользователей и сотрудников ИТ-отдела [7]. Использование инструмента походит как для аналитиков, так и для конечных пользователей.

### 3. Заключение

Хочется отметить, что спроектирована СППР, которая решает следующие задачи:

- быстрый доступ к данным, посредством графического интерфейса пользователя;
- возможность работы с информацией многим пользователям одновременно;
- многомерное концептуальное представление данных;
- расчеты и вычисления по нескольким измерениям, иерархиям и/или членам;
- выборка подмножеств данных для просмотра на экране;
- углубление в данные для просмотра информации на более детализированном уровне;
- представление информации в графическом виде;
- получение отчетных документов;
- экспорт результатов анализа в MS Excel;

- проведение анализа агрегированных данных, отражающих:
- структуру распределения затрат за оказанную медицинскую помощь по ЛПУ, СМО, врачам, медицинским услугам и т.п.;
- нагрузку на врачей по количеству оказанных услуг;
- анализ высокозатратных услуг;
- динамику изменений структуры диагнозов по сумме затрат и т.д.

### Список используемых источников

1. Медицинские информационные технологии / Эльянов М.М. — М.: Третья медицина, 2005. — 320 с.
2. Решения Пост Модерн Текнолоджи. МИС Медиалог // ПОСТ МОДЕРН ТЕКНОЛОДЖИ. URL: [https://www.medialog.ru/products/mis\\_medialog/](https://www.medialog.ru/products/mis_medialog/) (дата обращения: 01.03.2018).
3. Комплексный подход к внедрению Data Mining, OLAP и хранилищ данных в СППР // Национальный открытый институт ИНТУИТ URL: <https://www.intuit.ru/studies/courses/6/6/lecture/190> (дата обращения: 01.12.2017).
4. Microsoft SQL Server Analysis Services. OLAP и многомерный анализ данных / Э. Меломед, В. Степаненко, В. Щербинин, И. Горбач, А. Бергер. — БХВ-Петербург, 2007. — 248 -360 с.
5. Введение в OLAP-технологии Microsoft / Н. Елманова, А. Федоров - Диалог МИФИ, 2002 г. – 113- 240 с.
6. Использование MS SQL Server Analysis Services для построения хранилищ данных./ НОУ ИНТУИТ. URL: <https://www.intuit.ru/studies/courses/568/424/lecture/9641?page=2> (дата обращения: 26.01.2018)
7. Службы Power BI/ Power BI. URL: <https://powerbi.microsoft.com/ru-ru/what-is-power-bi/> (дата обращения: 07.04.2018 г.)