

ФОРМИРОВАНИЕ МЕДИЦИНСКОЙ СТАТИСТИЧЕСКОЙ ОТЧЕТНОСТИ С ПОМОЩЬЮ OLAP СИСТЕМ

Агеева С. А.

Пономарев А. А.

НИ ТПУ, ageevasophia@gmail.com

Для медицинских организаций существуют единые формы статистической отчетности, которые утверждает Росстат – Федеральная служба государственной статистики, а также единые формы учета и инструкции по их заполнению. Сбор статистической отчетности позволяет сравнивать результаты по субъектам федерации, городам и районам. Статистическая отчетность в большей степени нужна для самого медицинского учреждения и его руководителя. Годовой отчет представляет сводку данных о работе учреждения и условиях, в которых протекала работа за отчетный период. Анализ данного отчета дает возможность вскрыть причины отрицательных показателей работы учреждения в целом или его отдельных структурных подразделений. Также такая статистика необходима для истории работы учреждения, которая используется для справок и докладов.

Существуют Общие принципы построения и функционирования информационных систем и порядок информационного взаимодействия в сфере обязательного медицинского страхования, утвержденные приказом ФФОМС от 7 апреля 2011 № 79 (далее – Общие принципы). Настоящий документ регламентирует обмен информацией в электронном виде, в том числе в виде информационных сообщений и файлов, при информационном взаимодействии между участниками обязательного медицинского страхования – страховыми медицинскими организациями (далее – СМО), медицинскими организациями (далее – МО). Информационное взаимодействие между участниками ОМС предусмотрено частью 4 ст. 43 Федерального закона от 29.11.2010 № 326-ФЗ «Об обязательном медицинском страховании в Российской Федерации» (далее – Закон об ОМС), Правилами обязательного медицинского страхования, утвержденными приказом Минздравсоцразвития от 28.02.2011 № 158н (далее – Правила ОМС), Порядком ведения персонифицированного учета в сфере обязательного медицинского страхования, утвержденным приказом Минздравсоцразвития от 25 января 2011 № 29н (далее – Порядок персонифицированного учета).

Требованиями к предоставляемой статистической отчетности является:

- формат предоставляемого документа - xml, со специальным названием;
- для обмена информацией используется кодировка CP1251;

- соблюдение предоставленной структуры (№ п/п, код элемента, содержание элемента, тип элемента, размер, обязательность, наименование, дополнительная информация).

Предположим, что сбор данных осуществляется по средствам хранимой процедуры. Выгрузка статистической отчетности происходит каждый месяц. Требования к выгружаемым данным каждый месяц меняются, что приводит к затрате ресурсов на доработку хранимой процедуры.

Решения данной проблемы лежит в использовании OLAP системы. OLAP (англ. online analytical processing, аналитическая обработка в реальном времени) — технология обработки информации, которая используется для динамической публикации отчетов и документов. OLAP системы выполняют следующие задачи: обобщение, агрегация, гиперкубическое представление информации и многомерный анализ. OLAP системы могут быть представлены в виде многомерных СУБД или же реляционных баз с предварительной агрегацией данных. Эта технология помогает принимать решение при работе с многомерными данными. Но нельзя говорить об использовании только OLAP систем. Необходимо сказать, что OLAP необходимо использовать в совокупности с Data Mining. Data Mining – технология, которая анализирует информацию с целью нахождения в накопленных данных ранее неизвестных, практически полезных знаний, необходимых для принятия оптимального решения в разных областях деятельности. Интеграция технологий OLAP и Data Mining функциональнее одной и другой технологии, используемых раздельно. Результат этой интеграции называют OLAP Mining. J. Han предлагает следующие варианты интеграции этих технологий:

1. **"Cubing then mining"**. Возможность выполнения интеллектуального анализа должна обеспечиваться над любым результатом запроса к многомерному концептуальному представлению, то есть над любым фрагментом любой проекции гиперкуба показателей.
2. **"Mining then cubing"**. Подобно данным, извлеченным из хранилища, результаты интеллектуального анализа должны представляться в гиперкубической форме для последующего многомерного анализа.
3. **"Cubing while mining"**. Этот гибкий способ интеграции позволяет автоматически активизировать однотипные механизмы

интеллектуальной обработки над результатом каждого шага многомерного анализа (перехода между уровнями обобщения, извлечения нового фрагмента гиперкуба и т.д.).

На рисунке 1 изображен процесс многомерного интеллектуального анализа.



Рис. 14. Процесс многомерного интеллектуального анализа

Данные, находящиеся в оперативной памяти, взятые из разных источников, складываются в реляционное хранилище. В момент попадания данных в реляционное хранилище они уже доступны для анализа при помощи средств построения отчетов. Затем данные подготавливаются для OLAP анализа, загружаясь в БД OLAP или оставаясь в реляционном хранилище.

Критерии подбора инструментальных средств для подбора OLAP:

1. Удобство администрирования.
2. Интуитивное представление информации (понятны ли графические возможности, присутствуют ли механизмы экспорта результатов в стандартные форматы).
3. Богатство аналитических возможностей инструмента.
4. Скорость обработки больших хранилищ данных с нужной производительностью.
5. Совместимость OLAP инструментов с существующими в организации СУБД.

Классификация OLAP по месту размещения OLAP-машины:

1. OLAP-серверы.
2. OLAP-клиенты.

Классификация OLAP-серверов

1. MOLAP-многомерный (multivariate) OLAP. Для реализации многомерной модели используют многомерные БД. Представители: Cognos Powerplay, Oracle OLAP Option, Oracle Essbase, Microsoft Analysis Services, TM1, Palo, IdeaSoft O3.

- Преимущества: высокая скорость обработки запросов, так как данные хранятся в многомерных структурах;

- Недостатки: необходимы специальные инструменты для формирования кубов и их пересчёта в случае изменения базовых значений.

2. ROLAP-реляционный (relational) OLAP. Для реализации многомерной модели используют

реляционные БД. Представители: Metaphor компании Metaphor Computer Systems, DSS Suite фирмы MicroStrategy, MetaCube фирмы IBM Informix, Platinum Beacon от Platinum.

Преимущества: возможность использования ROLAP с хранилищами данных и различными OLTP-системами; безопасность и администрирование обеспечивается реляционными СУБД.

Недостатки: получение агрегатов и листовых данных происходит медленнее, чем, например, в MOLAP и HOLAP; сложно поддерживать таблицы агрегатов.

3. HOLAP-гибридный (hybrid) OLAP. Для реализации многомерной модели используют и многомерные, и реляционные БД. Microsoft Analysis Services, MicroStrategy, IBM DB2 OLAP Server, Sagent HoloS.

- Преимущества: комбинирование технологии ROLAP для разреженных данных и MOLAP для плотных областей.

- Недостатки: поддержания MOLAP и ROLAP.

Подводя итоги, следует отметить, что интеграция OLAP Data Mining требует средств и времени. Возникнет вопрос необходимости перехода от одной технологии к другой. Актуальность OLAP-технологий обусловлена их практической значимостью для анализа больших объемов данных, что необходимо медицинской организации при выгрузке реестров. При выборе OLAP появляется проблема выбора оптимальной схемы хранения и обработки OLAP данных.

Список используемой литературы:

1. А.Н.Андреев, Классификация OLAP-систем вида xOLAP. 2010 г. - Режим доступа: http://citforum.ru/consulting/BI/xolap_classification/.
2. Российская Академия Естественных наук, OLAP И МНОГОМЕРНЫЕ БАЗЫ ДАННЫХ. 2010 г. - Режим доступа: <http://www.rae.ru/monographs/141-4638>.
3. Л.В. Щавелёв Оперативная аналитическая обработка данных: концепции и технологии. 2001 г. - Режим доступа: http://www.olap.ru/basic/olap_and_ida.asp.
4. И.А. Чубукова Базы данных: Data Mining Интернет-университет информационных технологий - ИНТУИТ.ру 2008.