

УДК:656.71:504.43

## **ОРГАНИЗАЦИЯ ГЕОПОРТАЛОВ И ГИС СИСТЕМ НА БАЗЕ ХРАНИЛИЩА ПРОСТРАНСТВЕННЫХ ДАННЫХ**

*С.М. Андреев, канд. техн. наук  
(Национальный аэрокосмический университет  
им. Н.Э. Жуковского "ХАИ");  
В.Н. Лысенко, канд. физ.-мат. наук,  
Вит.В. Радчук,  
И.В. Радчук  
(Институт телекоммуникаций и глобального  
информационного пространства НАН Украины)*

*Запропоновано підхід щодо організації ГИС і геопорталів на базі сховища просторових даних і OLAP – технологій багатовимірного аналізу даних, який комплексно вирішує питання збору, зберігання й аналізу геоданих.*

*Предложен подход по организации ГИС и геопорталов на базе хранилища пространственных данных и OLAP - технологий многомерного анализа данных, который комплексно решает вопросы сбора, хранения и анализа геоданных.*

*The approach to the organization of GIS and geoportals based on storage of spatial data and OLAP-technology of multivariate data analysis that deals with geodata collection, storing and analyzing is offered.*

В настоящее время системы цифрового картографирования широко используют карты автоматизированного производства в файлах формата Autocad или Microstation. Данный подход имеет ряд существенных недостатков, таких как:

- избыточность объектного состава (кодов объектов, атрибутивной информации);
- частичное дублирование информации на картах разных масштабов;
- сложность интеграции, разобщенность и разнородность информационных ресурсов существенно затрудняет обмен пространственными данными и их совместное использование;
- низкая эффективность картографического мониторинга вследствие дублирования объектов на картах разных масштабов и нестандартизированного представления информации.

Современное техническое развитие геоинформатики знаменуется переходом

от устоявшейся, но изолированной и автономной технологии картографирования, к концепции интегрированных ГИС и геопорталов. Эта концепция, как демонстрирует платформа ESRI ArcGIS Server, четко направлена на создание решений на основе распределенных, обновляемых, объектно-ориентированных баз пространственных данных. Продуктом такой концепции является не карта, как прежде, а данные из баз данных ГИС и геопорталов.

Интегрированные ГИС и геопорталы на основе хранилища пространственных данных (ХПД) - это сравнительно новое технологическое решение, которое стало широко использоваться только в начале 90-х годов XX века. Его автором стал Билл Инмон [1]. Основная идея заключается в постепенном переходе от полистного хранения пространственной информации к единому массиву данных, в котором основной единицей хранения является пространственный объект. Основные задачи, решаемые при помощи ХПД:

- проведение всестороннего анализа пространственных данных;
- регламентирование и разработка методов подготовки различных типов пространственной информации для ввода всистему;
- оптимизация мониторинга информации о пространственных объектах;
- исключение дублирования пространственных данных;
- облегчение актуализации данных;
- упрощение разработки новой картографической продукции (цифровых и бумажных карт, Web-публикаций и т. п.).

Преимущества, получаемые от внедрения интегрированной ГИС и геопорталов на основе ХПД:

- появляется единая ГИС хранения пространственных данных, в которой используется единая справочная информация;
- появляется возможность регламентирования и разработки методов подготовки различных типов пространственной информации для ввода в систему;
- появляется возможность проведения анализа с использованием исторических данных;
- исключается дублирование пространственных данных;
- облегчается актуализация данных;
- упрощается разработка новой картографической продукции (цифровых и бумажных карт, Web-публикаций, собственных онлайн сервисов и т. п.);
- возникает возможность проведения всестороннего анализа информации;
- появляется возможность соединения и анализа информации, ранее хранившейся в разных информационных системах.

Главную роль при организации ХПД имеют транзакционные OLTP системы (англ. OnlineTransactionProcessing, обработка транзакций в реальном времени) – это системы, предназначенные для ввода, структурированного хранения и обработки информации в режиме реального времени. OLTP - системы

## **Розділ 2. Основи природокористування та безпека життєдіяльності**

проектируются, настраиваются и оптимизируются для выполнения максимального количества транзакций за короткие промежутки времени. Как правило, большой гибкости здесь не требуется, и чаще всего используется фиксированный набор надёжных и безопасных методов ввода, модификации, удаления данных и выпуска оперативной отчётности.

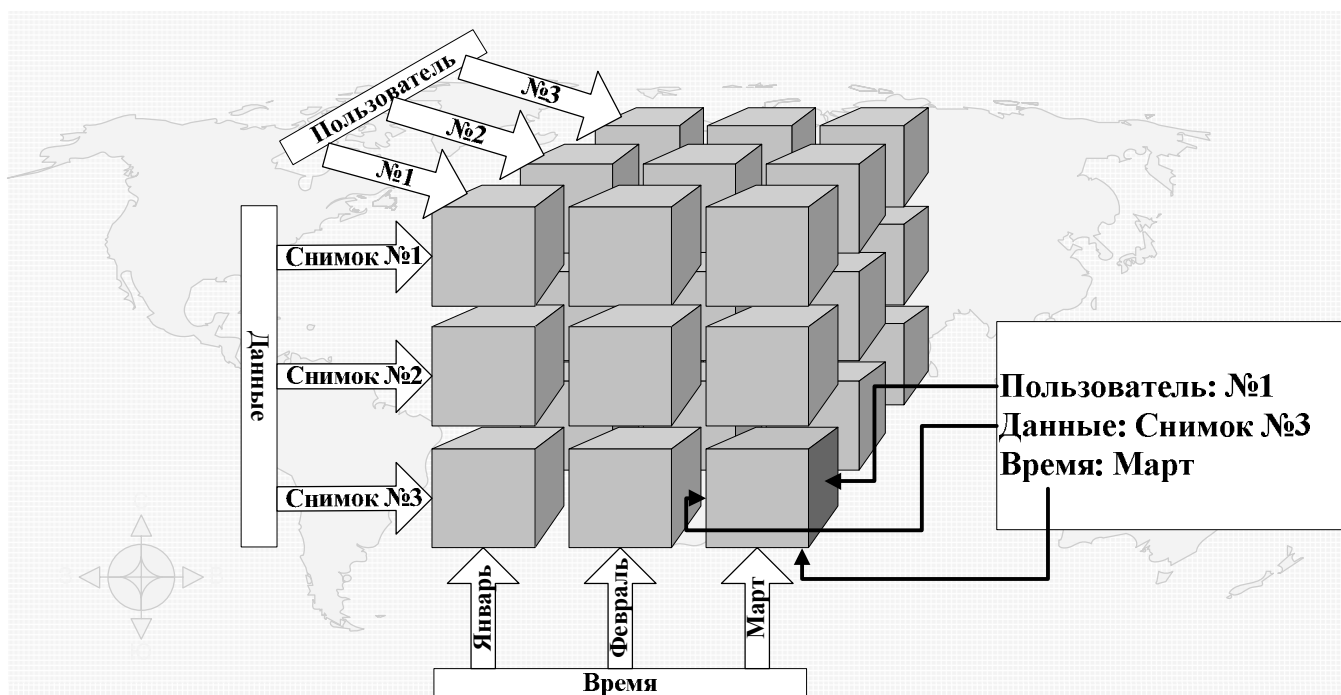
OLAP - технология комплексного многомерного анализа данных. OLAP - это ключевой компонент организации хранилищ данных. Концепция OLAP была описана в 1993 году Эдгаром Коддом, известным исследователем баз данных и автором реляционной модели данных [2]. В 1995 году на основе требований, изложенных Коддом, были сформулированы следующие требования к приложениям для многомерного анализа:

- предоставление пользователю результатов анализа за приемлемое время (обычно не более 5 с), пусть даже ценой менее детального анализа;
- возможность осуществления любого логического и статистического анализа, характерного для данного приложения, и его сохранения в доступном для конечного пользователя виде;
- многопользовательский доступ к данным с поддержкой соответствующих механизмов блокировок и средств авторизованного доступа;
- многомерное концептуальное представление данных, включая полную поддержку для иерархий и множественных иерархий (это - ключевое требование OLAP);
- возможность обращаться к любой нужной информации независимо от ее объема и места хранения.

Системы поддержки принятия решений обычно обладают средствами предоставления пользователю агрегатных данных для различных выборок из исходного набора в удобном для восприятия и анализа виде. Как правило, такие агрегатные функции образуют многомерный, не реляционный набор данных (OLAP-куб), оси которого содержат параметры, а ячейки - зависящие от них агрегатные данные (рис.1).

Вдоль каждой оси данные могут быть организованы в виде иерархии, представляющей различные уровни их детализации.

Благодаря такой модели данных пользователи могут формулировать сложные запросы, генерировать отчеты, получать подмножества данных. OLAP-куб может быть реализован на основе универсальных реляционных СУБД или специализированным программным обеспечением.



*Рис.1. Простейший (трехмерный) OLAP-куб для организации хранения данных ДЗЗ*

Существуют три типа OLAP:

– OLAP со многими измерениями (Multidimensional OLAP - MOLAP). Использует суммирующую БД, специальный вариант процессора пространственных БД и создаёт требуемую пространственную схему данных с сохранением базовых данных и агрегатов. MOLAP лучше всего подходит для небольших наборов данных, он быстро рассчитывает агрегаты и возвращает ответы, но при этом генерируются огромные объёмы данных;

– реляционный OLAP (Relational OLAP - ROLAP). ROLAP работает напрямую с реляционным хранилищем, факты и таблицы с измерениями хранятся в реляционных таблицах, и для хранения агрегатов создаются дополнительные реляционные таблицы. ROLAP оценивается как более масштабируемое решение, использующее к тому же наименьшее возможное пространство. При этом скорость обработки значительно снижается;

– гибридный OLAP (Hybrid OLAP - HOLAP). HOLAP использует реляционные таблицы для хранения базовых данных и многомерные таблицы для агрегатов. HOLAP находится посреди этих двух подходов, он достаточно хорошо масштабируется и быстро обрабатывается.

Структурная схема отвечает всем требованиям к разработке аналитических систем и представлена на рис. 2.

Принцип работы предлагаемой информационной системы на базе ХПД заключается в следующем. Данные поступают из различных источников (так

## Розділ 2. Основи природокористування та безпека життєдіяльності

называемых операционных систем). Это могут быть данные из архивов, различных внешних приложений и т.д.

При помощи ETL-приложения (программного обеспечения, позволяющего извлекать, трансформировать и загружать данные), данные из систем-источников попадают в базу данных хранилища данных. В качестве ETL-приложений могут использоваться: Microsoft Integration Services, Informatica Power Center, IBM Data Stage, Oracle Data Integrator, Oracle Ware House Builder.

База данных хранилища пространственных данных не является абстрактной по своей структуре (набору таблиц, полей в них и взаимосвязей между таблицами), а создана на основе внутренней модели данных хранилища. В качестве баз данных в подавляющем большинстве используются Oracle, Microsoft или Teradata.

Далее, используя данные из хранилища данных, производится настройка областей анализа, отчетности и витрин данных. При реализации данной системы пользователи с довольно низким уровнем подготовки смогут вполне самостоятельно составлять необходимую отчетность и проводить многомерный анализ. В качестве инструментов анализа в основном используются Microsoft Analysis Services, Business Objects, Oracle Discoverer, IBM Alpha Blocks и другие программные продукты.

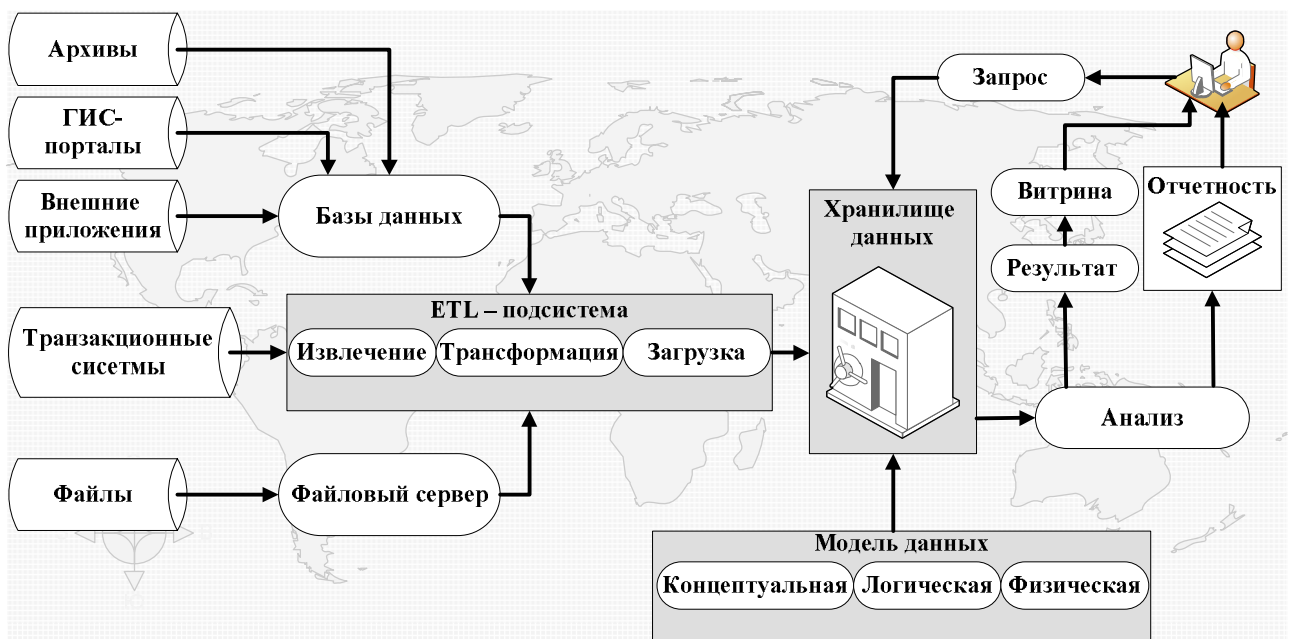


Рис. 2. Информационная система на базе ХПД

Таким образом, организация ГИС и геопорталов на базе ХПД реализует принципиально новый подход к представлению пространственной информации, соответствующий современным требованиям развития информационных

технологий и комплексно решает вопросы сбора, хранения и анализа пространственной информации.

\* \* \*

1. Inmon W.H. Building the Data Warehouse, QED/Wiley. - New-York: John Wiley & Sons, 1991. - 312 p.
2. Codd E.F., Codd S.B. and Salley C.T. Providing OLAP to User-Analysts: An IT Mandate, 1993.
3. Асадуллаев С.С. Архитектуры хранилищ данных. - [http://www.ibm.com/developerworks/ru/library/sabir/axd\\_1/index.html](http://www.ibm.com/developerworks/ru/library/sabir/axd_1/index.html)

**Отримано: 14.05.2012 р.**