

РАСПРЕДЕЛЁННАЯ СУБД ДЛЯ ИНТЕГРИРОВАННОЙ ОБРАБОТКИ ПРОСТРАНСТВЕННЫХ ДАННЫХ В ГИС

Ю.Г. Васин¹, Ю.В. Ясаков¹

¹ Центр информатики и интеллектуальных информационных технологий
Института информационных технологий, математики и механики

Нижегородского государственного университета им. Н.И. Лобачевского, Нижний Новгород, Россия

Аннотация

В статье рассмотрены проблемы, возникающие при создании и обработке данных, содержащих сложноструктурированную геопространственную информацию, и связанные с большими объемами данных, их размерностью и увеличением количества пространственно-логических и топологических связей между элементами данных, разнородностью данных. Приведены наиболее популярные средства обработки подобной информации. Рассмотрены вопросы организации хранения, обработки и анализа пространственно-распределенных топологических данных в среде ГИС Терра. Изложены основные особенности (модели, структуры и форматы данных) разработанной структурированной топологической модели объекта. Рассмотрены варианты организации системы управления базами пространственно-распределенных данных для различных программно-аппаратных платформ (Windows и Android) как в локальном, так и клиент-серверном варианте. Представлены схемы миграции данных при интегрированном использовании различных программно-аппаратных платформ. Приведены рекомендации по разработке новых приложений по анализу и обработке пространственно-распределенных данных с использованием разработанных средств хранения и обработки данных.

Ключевые слова: СУБД, пространственно-распределенные данные, модели данных, пространственно-логические отношения, алгоритмы, обработка данных, топология.

Цитирование: Васин, Ю.Г. Распределённая СУБД для интегрированной обработки пространственных данных в ГИС / Ю.Г. Васин, Ю.В. Ясаков // Компьютерная оптика. – 2016. – Т. 40, № 6. – С. 919-928. – DOI: 10.18287/2412-6179-2016-40-6-919-928.

Введение

При создании и обработке баз данных, содержащих сложноструктурированную геопространственную информацию, по сравнению с традиционными базами данных возникают дополнительные проблемы, связанные с большими объемами данных, их размерностью и увеличением количества пространственно-логических и топологических связей между элементами данных, разнородностью данных. Это приводит к необходимости разработки моделей, структур и форматов данных, существенно отличающихся от традиционных способов представления [1–4].

В то время как традиционные СУБД могут хранить и обрабатывать числовую и текстовую информацию, пространственные данные обладают расширенной функциональностью, позволяющей хранить целостный пространственный объект, объединяющий как традиционные виды данных в виде описательной или атрибутивной информации, так и геопространственные данные о положении объекта в пространстве, внутриобъектных и межобъектных связях.

Наиболее распространенными и популярными в настоящее время являются СУБД, основанные на реляционной модели данных. Реляционное представление данных – двумерные таблицы данных с изменяющимися взаимосвязями между различными столбцами. Это напоминает представление данных самим человеком.

К настоящему времени все большее число разработчиков пользовательских приложений для СУБД

сталкиваются с тем, что применение традиционных реляционных моделей не обеспечивает требований, предъявляемых к возможностям хранения и обработки сложных неструктурированных пространственно-распределенных данных (ПРД), а также скорости обработки запросов для них. Поэтому крупнейшие разработчики СУБД стали встраивать в свои продукты поддержку объектной ориентации, а также дополнительные средства представления и обработки пространственных данных. Кроме того, более эффективными становятся средства поддержки неинтерпретируемых (битовых) данных.

С этой точки зрения представляют интерес следующие разработки.

СУБД Oracle [5] с расширением Oracle Spatial, представляющий собой компонент СУБД, обеспечивающий хранение, получение и изменение пространственных данных.

СУБД PostgreSQL [6] – это свободно распространяемая объектно-реляционная система управления базами данных, одна из наиболее развитых из открытых СУБД в мире, являющаяся реальной альтернативой коммерческим базам данных. Необходимо еще отметить расширение для PostgreSQL – PostGIS [7]. PostGIS был выпущен в 2001 году компанией Refrations Research и составляет конкуренцию коммерческим решениям, являясь при этом свободным программным продуктом с открытым исходным кодом. Основным достоинством PostGIS является возможность использования языка SQL совместно с пространственными операторами и функциями. Кроме простого хранения

такого типа данных, PostGIS позволяет осуществлять широкий набор операций над ними.

СУБД ГИС Terra [3, 8, 9, 10] разработана в Нижегородском государственном университете им. Н.И. Лобачевского. Она обеспечивает обработку пространственно-распределенных данных на основе специально разработанной структурированной топологической модели данных.

Данная статья посвящена вопросам дальнейшего развития СУБД ГИС Terra на базе современных средств хранения и обработки данных и программно-аппаратных платформ, а также с учетом особенностей актуальных задач по представлению, хранению и обработке сложноструктурированной топологически связанной геопространственной информации.

1. Форматы и структуры данных для моделирования ПРД

Для моделирования сложноструктурированных пространственно-распределенных данных в ГИС Terra разработана специальная объектно-ориентированная структурированная топологическая модель объекта данных (рис. 1).

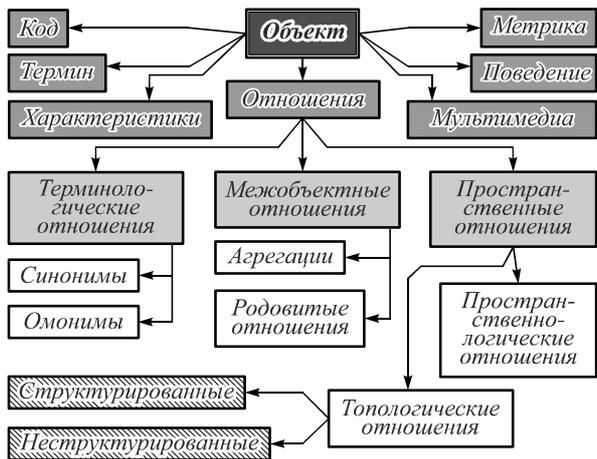


Рис. 1. Инфологическая модель ПРД

Для физической реализации этой модели разработана специальная структура данных – формат интегрального файла (ИФ) (рис. 2).

Объект является основной информационной единицей базы данных. Вся информация об отдельном объекте содержится в одной структурированной записи. Этим обеспечивается быстрое получение любого объекта как единого целого (например, дороги со всеми придорожными сооружениями). Объект может находиться в базе данных или выдаваться пользователю в любой комбинации его составных частей.

В базе данных поддерживаются несколько типов объектов:

- элементарный объект, состоящий из одной записи базы данных в формате интегрального файла (дискретный, линейный, площадной объект);
- сложный или составной объект, состоящий из нескольких записей базы данных;
- структурированный объект, состоящий из одной записи базы данных и содержащий описание как ос-

новного объекта (например, дороги), так и ему сопутствующих (например, мосты и придорожные сооружения);

- распределенный объект, объединяющий в себя несколько объектов предыдущих типов из разных баз данных, покрывающих некоторый район электронной карты.

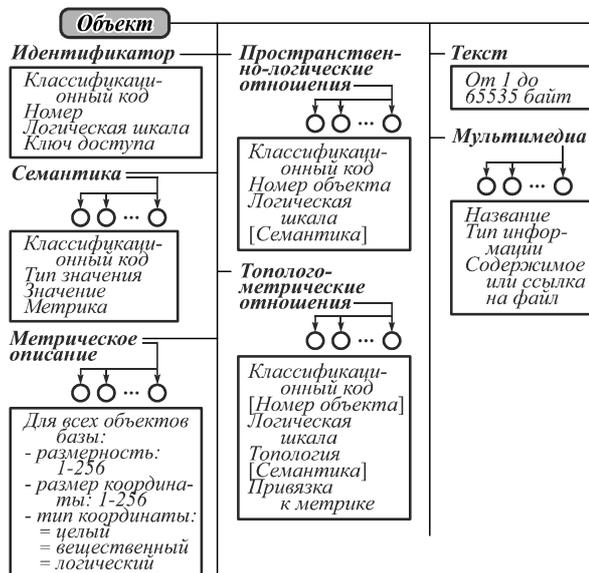


Рис. 2. Структура объекта в формате интегрального файла ГИС Terra

База данных представляет собой дерево объектов, построенное в соответствии с иерархией многопозиционного классификационного кода объекта предметной области (рис. 3).

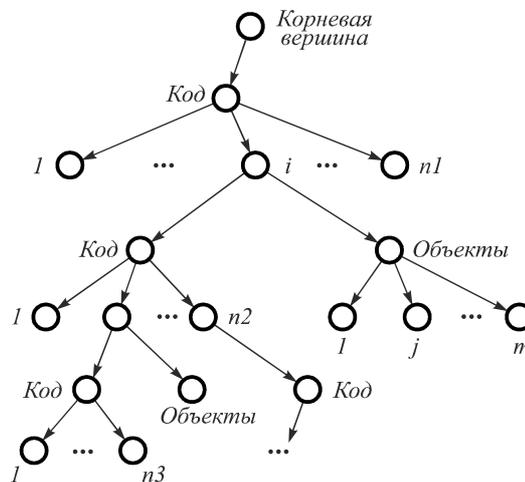


Рис. 3. Дерево объектов базы данных

Кроме того, допускается установление (и автоматическая поддержка) прямых (сетевых) связей между объектами базы данных. В результате подобного подхода обеспечивается:

- быстрый индивидуальный доступ к отдельному объекту;
- быстрая выборка совокупности объектов, объединенных одним обобщенным кодом;
- выборка совокупности объектов, объединенных сетевыми ссылками.

2. Распределенная СУБД ГИС Terra

Система управления базами пространственно-распределенных топологических данных ГИС Terra реализована на нескольких программно-аппаратных платформах с использованием единой информационной модели данных – модели объекта в формате интегрального файла.

2.1. Платформа Windows

На платформе операционной системы Windows реализовано два варианта СУБД:

- локальная (встраиваемая) СУБД (СУБД ГИС Terra) [3, 8, 9];
- клиент-серверная СУБД (СУБД ГИС Terra Plus).

СУБД ГИС Terra является первой разработанной СУБД для поддержки формата интегрального файла. Ее основные особенности:

- база данных представляет собой один файл специальной структуры;
- поддерживается формат обработки и формат обмена (бинарный и текстовый);
- имеется ограничение на размер файла базы данных (до 2 Гбайт);
- в качестве механизма доступа к данным в СУБД использован древовидный метод доступа, что обеспечивает высокие временные характеристики поиска информации;
- обладает развитой системой запросов. Запросы кодируются на специально разработанном языке.

СУБД ГИС Terra Plus является развитием предыдущей версии СУБД на новой программной платформе. В качестве базовой СУБД используется СУБД PostgreSQL с расширением PostGIS. Соответственно поддерживается клиент-серверная архитектура.

Модифицирована модель объекта в формате интегрального файла (рис. 4).

Поддерживается многоверсионное представление всех полей объекта.

Обеспечивается установление связи объектов с метаданными источника их поступления в базу данных.

В базе данных ведется журнал работ, выполненных с объектом.

При разработке моделей, форматов и структур данных для эффективной реализации объектно-ориентированной топологической структурированной модели ИФ в рамках реляционной СУБД PostgreSQL необходимо было решить несколько принципиально важных задач:

- иерархическую структуру доступа к объектам базы данных;
- многоверсионное хранение всех элементов (полей) объекта;
- эффективное кодирование метрического описания атрибутов и объектов;
- ведение специального журнала работ с объектом, содержащего информацию о произведенных с ним (и его полями) действиях по обновлению;
- связь объекта с метаданными источника, с которого получен объект (топографические и морские карты, планшеты гидрографических съемок и т.п.);

- возможность выполнения SQL-запросов наряду с не SQL-запросами СУБД ГИС Terra;
- функционирование программных средств СУБД ГИС Terra для работы с объектом (его полями), полученным из базы данных;
- конвертацию информации из базы данных СУБД ГИС Terra в базу данных ГИС Terra Plus (PostgreSQL) и наоборот.

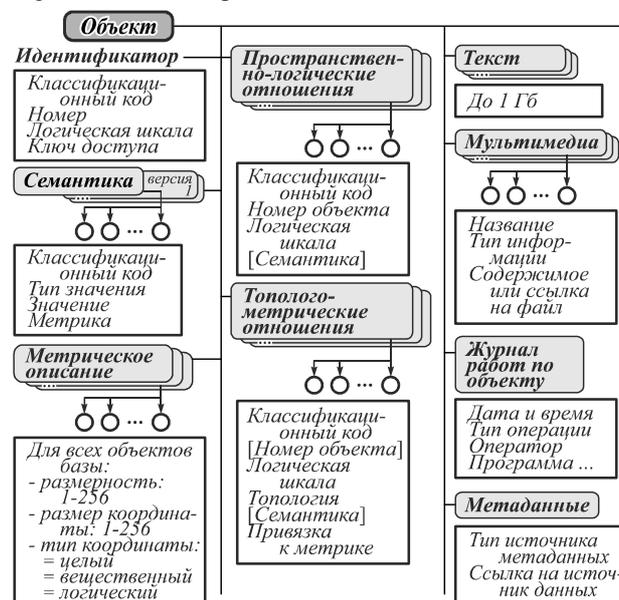


Рис. 4. Структура объекта в формате ИФ ГИС Terra Plus

Эти задачи были решены путем разработки специальной схемы базы данных, представляющей собой ряд взаимосвязанных таблиц СУБД PostgreSQL (рис. 5), а также использование типа данных 'binary' для сложноструктурированных полей объекта с разработкой специальных средств программной поддержки.

Для реализации иерархического доступа к объектам базы данных разработана специальная таблица классификационных кодов, хранящая многопозиционные классификационные коды объектов (на каждую позицию кода отдельная колонка) и соответствующие им уникальные идентификаторы. При выполнении запроса на отбор объектов по классификационному коду (обобщенному или полному) происходит выборка уникальных идентификаторов классификационных кодов по этой таблице, а затем происходит полная сборка объекта из таблиц атрибутов, связей, прерываний, метрического описания, мультимедийной информации с использованием полученных уникальных идентификаторов.

Каждый объект в базе данных идентифицируется уникальным идентификатором классификационного кода и номером в пределах этого классификационного кода.

Возможна выборка из базы данных как отдельного объекта, так и группы объектов, объединенных обобщенным или полным классификационным кодом.

Для каждого поля объекта (семантика, связи, прерывания, метрика, мультимедиа) в базе данных имеется отдельная таблица, в которой можно хранить за-

писи, соответствующие разным версиям объекта. Версия с номером 0 является актуальной версией. При выполнении запроса на выборку объектов имеется возможность для каждого отдельного поля объекта указать номер желаемой версии поля объекта.

Таблица классификационных кодов (код ↔ идентификатор кода (IDC))

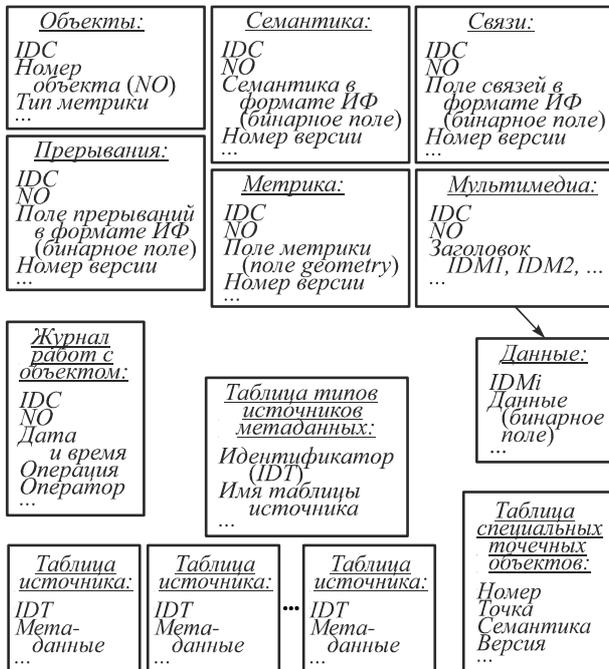


Рис. 5. Принципиальная схема базы данных в СУБД PostgreSQL

Для хранения метрического описания объектов в таблице метрического описания используется поле типа 'geometry'. В СУБД поддерживаются два варианта работы с этим полем. Первый подразумевает доступ к метрике объекта посредством функций PostGIS. Этот способ обеспечивает независимость от физической реализации поля типа 'geometry', но требует дополнительных временных ресурсов на преобразование координат метрического описания из внутреннего представления во внешнее и наоборот. Второй подразумевает непосредственный доступ к координатам метрического описания на физическом уровне, что существенно повышает быстродействие, но может привести к проблемам при изменении структуры метрического описания (например, при выпуске новой версии PostGIS). PostGIS позволяет использовать богатый набор функций для работы с метрическим описанием объектов (преобразование координат, анализ взаимного расположения объектов, вычислительная геометрия и т.п.).

В СУБД ГИС Terra поля, содержащие семантическое описание объектов, пространственно-логические и тополого-метрические отношений между объектами, имеют сложную структуру и представляют собой произвольный (не всегда упорядоченный) набор элементов, содержащий классификационные коды объектов, произвольные наборы атрибутов, координат-

ную привязку и т.п. Эффективно реализовать подобную конструкцию, используя реляционную таблицу, не представляется возможным. Поэтому данные части объекта хранятся в полях типа 'binary'. Средства СУБД ГИС Terra Plus обеспечивают запись и чтение этих полей целиком. При получении данных полей из базы данных их обработка ведется обычными средствами СУБД ГИС Terra.

В СУБД ГИС Terra Plus поддерживаются два режима работы системы запросов:

- запросы СУБД ГИС Terra (не SQL-запросы);
- дополнительные SQL-запросы.

Под запросами СУБД ГИС Terra понимаются запросы на отбор информации по полям атрибутов, связей и прерываний объекта, связанные со спецификой представления этих полей в базе данных. Для реализации подобных запросов используется технология хранимых процедур СУБД PostgreSQL. Разработана специальная динамически загружаемая библиотека функций (хранимых процедур), обеспечивающих отбор объектов по условиям на их поля, представленные в бинарном виде. При выполнении такого типа запросов после отбора подходящего объекта на сервере СУБД PostgreSQL дополнительно происходит его анализ (а именно, полей атрибутов, связей или прерываний) с помощью соответствующей хранимой процедуры.

Приведем пример реализации такого запроса, а именно запроса: Выбрать все объекты «Береговая линия достоверная», проходящие по объектам «Граница государства». Отбор объектов будет производиться по наличию у них отношения с объектами «Граница государства».

В СУБД ГИС Terra он будет выглядеть следующим образом:

Code=4.1.1 – Классификационный код объекта "Береговая линия достоверная";

Type=CODE – Только объекты данного уровня иерархии;

Filter=P((K(11.3))) – 11.3 – это классификационный код объекта "Территория государства".

В СУБД ГИС Terra Plus этот же запрос будет выглядеть следующим образом:

```
SELECT <необходимые поля объекта> FROM
<соответствующие таблицы базы данных> WHERE
N=3 AND C1=4 AND C2=1 AND C3=1 AND
QUERY_IF('P((K(11.3)))')=TRUE
```

Здесь:

N – количество позиций классификационного кода объекта (задает уровень иерархии классификационного кода);

C1, C2, C3 – соответствующие позиции кода объекта «Береговая линия достоверная»;

QUERY_IF – хранимая процедура СУБД ГИС Terra Plus, производящая на сервере СУБД PostgreSQL анализ объекта по поданному фильтру.

Для отбора объектов базы данных могут быть использованы дополнительные SQL-запросы. Ниже приведены основные из них:

- по метрическому описанию. В этом случае из не SQL-запроса исключается запрос на метрическое описание (если он там присутствует);
- по логической шкале объекта. В этом случае из не SQL-запроса исключается запрос на логическую шкалу объекта (если он там присутствует);
- по полю мультимедиа. В этом случае из не SQL-запроса исключается запрос на элементы поля мультимедиа (если он там присутствует);
- по типу источника поступления объектов;
- по названию источника поступления;
- по метаданным источника поступления;
- по масштабу;
- по алгебраическим условиям на значения отметок глубины.

Не SQL-запросы и дополнительные SQL-запросы могут применяться как по отдельности, так и одновременно, что существенно повышает эффективность запросной системы СУБД. При одновременном применении производится анализ не SQL-запроса и из него исключаются запросы на поля объекта, если таковые имеются и в дополнительном SQL-запросе.

СУБД ГИС Terra Plus имеет клиент-серверную архитектуру (рис. 6).

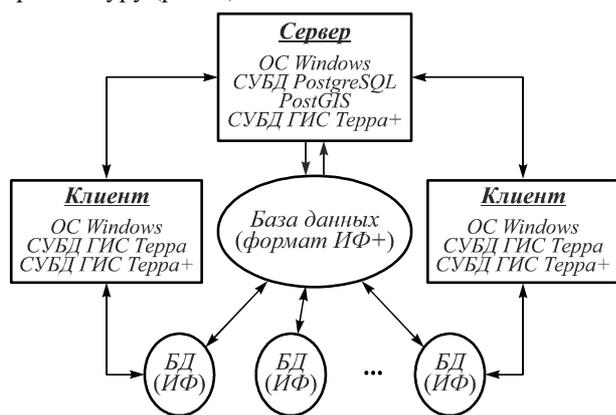


Рис. 6. Архитектура СУБД ГИС Terra Plus

Сервер СУБД может функционировать под управлением операционной системы Microsoft Windows или операционной системы Linux (например, Astra Linux Special Edition).

На сервере должна быть установлена СУБД PostgreSQL версии не ниже 9.3 и PostGIS версии не ниже 2.1 – расширение СУБД PostgreSQL, предназначенное для хранения и обработки пространственных данных.

Кроме того, на сервере должны находиться хранимые функции СУБД ГИС Terra Plus, обеспечивающие выполнение запросов на отбор объектов базы данных по некоторым полям объекта (прерываниям, связям и характеристикам).

На клиентских местах должна быть установлена СУБД ГИС Terra и клиентская часть СУБД ГИС Terra Plus. Клиентские места функционируют под управлением ОС Microsoft Windows.

2.2. Платформа Android

Для мобильных устройств на платформе операционной системы Android реализовано клиентское

программное обеспечение для удаленного доступа к базам данных СУБД ГИС Terra Plus.

Разработка была выполнена в среде программирования Embarcadero RAD Studio (<https://www.embarcadero.com/ru/>) с использованием программного обеспечения UniDAC (<http://www.devart.com/unidac/>), представляющего собой набор невизуальных компонентов для межплатформенного доступа к базам данным. Использование средств UniDAC обеспечивает возможность полноценного удаленного доступа к базам данных СУБД ГИС Terra Plus, созданным в среде СУБД PostgreSQL.

Кроме того, реализована полнофункциональная локальная (встраиваемая) объектно-ориентированная топологическая СУБД – СУБД ГИС Terra M на базе программных средств СУБД ГИС Terra и встроенной СУБД SQLite (<http://www.sqlite.org/>). Она обеспечивает локальную работу с базами данных формата СУБД ГИС Terra, перенесенными на мобильное устройство.

Рассмотрим структуру базы данных в этой СУБД (рис. 7).

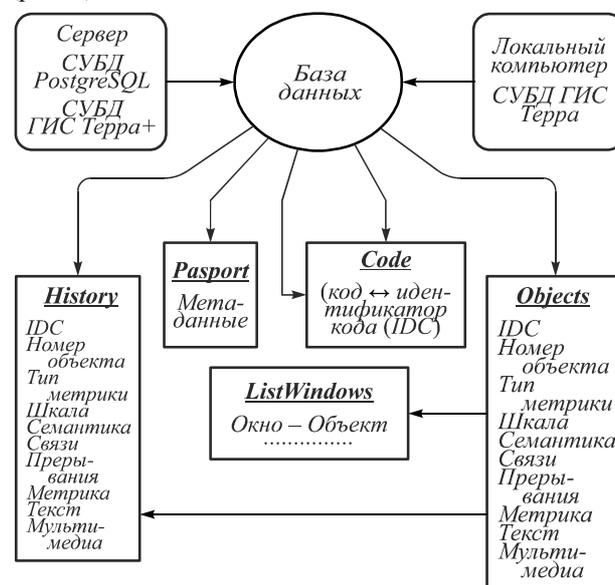


Рис. 7. Структура базы данных

База данных содержит пять таблиц.

1) Pasport – содержит метаданные для конкретной базы данных (имя классификатора предметной области, параметры метрического описания, координаты углов минимального охватывающего прямоугольника и т.п.).

2) Code – устанавливает соответствие между многопозиционным классификационным кодом объекта и уникальным числовым идентификатором. В дальнейшем эта таблица используется для организации иерархического доступа к объектам базы данных.

3) Object – основная таблица объектов. Содержит информацию объекта в формате интегрального файла (рис. 8).

Поля объекта в формате интегрального файла [1]: семантика (атрибутов), связи, прерывания, метрическое описание, мультимедиа. Эти поля являются данными типа BLOB (битовыми) и не интерпретируются СУБД SQLite. Их обработка производится после счи-

тывания из базы данных средствами, аналогичными средствам обработки объекта СУБД ГИС Terra. Поддерживается многоверсионное представление объекта.

4) History – таблица для сохранения истории изменения объекта. По своей структуре полностью аналогична таблице Object.

5) ListWindows – содержит двумерный пространственный индекс, обеспечивающий быстрый поиск объекта по метрическому описанию.

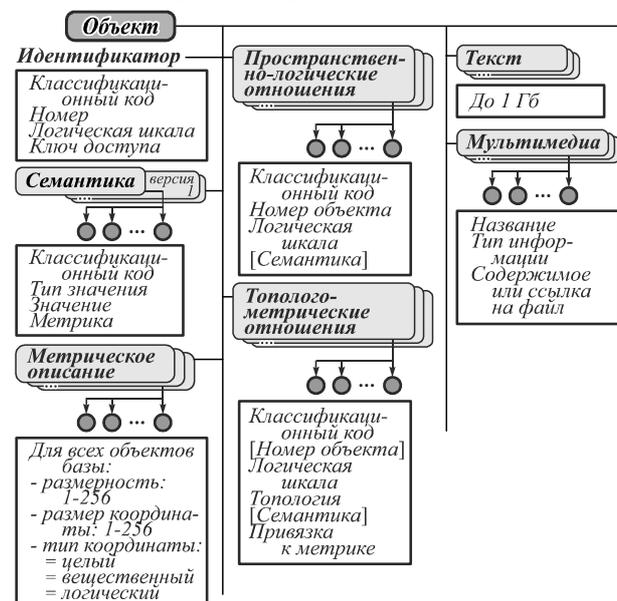


Рис. 8. Структура объекта в формате интегрального файла СУБД ГИС Terra M

3. Разработка новых приложений

Разработанная распределенная СУБД предназначена как для использования в рамках уже существующих приложений по обработке ПРД, так и для создания новых приложений и программных комплексов. Поэтому одним из основных их компонентов является API (интерфейс программирования приложений) – набор готовых классов, процедур, функций и структур для использования во внешних программных продуктах и оформленных в виде соответствующих библиотек.

Основным элементом разработанного API является реализованный на языке C набор функций, алгоритмы которых обеспечивают унифицированную работу со структурированной топологической моделью объекта, находящегося в оперативной памяти компьютера (мобильного устройства). Этот компонент API может быть использован при разработке приложений для разных платформ СУБД (рис. 9).

Разработку новых приложений предпочтительней вести в среде современной системы программирования Embarcadero RAD Studio – единой интегрированной среды разработки приложений для платформ MS Windows, iOS, Android с использованием компонентов FireDAC, UniDac – универсальных библиотек доступа к данным, предназначенных для разработки приложений для различных устройств, подключаемых к корпоративным базам данных и обеспечиваю-

щей универсальный и высокоэффективный прямой native-доступ из Delphi и C++Builder к InterBase, SQLite, MySQL, SQL Server, Oracle, PostgreSQL, DB2, SQL Anywhere, Advantage DB, Firebird, Access, Informix, DataSnap и т. д.



Рис. 9. Структура взаимосвязи СУБД для разных платформ

4. Временные оценки

Была проведена оценка временных характеристик доступа к данным с использованием различных вариантов организации разработанной СУБД.

Оценка проводилась на созданной базе батиметрических данных, содержащей как точечные (отметки глубин, опасности и т.п.), так и протяженные (изобаты, береговые линии и т.п.) объекты.

База данных содержала около 100000 объектов.

Оценка проводилась на программно-аппаратном комплексе в следующей комплектации.

Персональный компьютер на базе процессора Intel Core 2CPU 6700 с тактовой частотой 2,67 ГГц. Объем оперативной памяти – 4 ГБ. Операционная система Microsoft Windows 7 (x64). На компьютере установлен сервер СУБД PostgreSQL версии 9.3 (32-разрядная версия) с надстройкой PostGIS версии 2.5. Сервер СУБД был установлен со значениями по умолчанию.

Смартфон на базе процессора Qualcomm Snapdragon 1.2 ГГц с архитектурой 4xARM Cortex-A5. Объем оперативной памяти 512 МБ. Операционная система Android 4.1.2. База данных находилась на внешней SD-карте (скорость чтения около 20 МБ/с).

Выполнялись как массовые операции выборки, так и операции с условиями на отбор по атрибутам объектов и их отношениями с другими объектами.

Полученные временные характеристики приведены ниже.

Табл. Временные характеристики

	в1	в2	в3
з1	91237/0,870	91237/5,838	91237/63
з2	76185/0,510	76158/2,468	76185/52
з3	139/0,120	139/0,079	139/38
з4	19/менее 0,01	19/0,105	19/0,225
з5	8/менее 0,01	8/0,110	8/2,4

Запросы на выборку информации:

з1 – все объекты;

- 32 – все отметки глубин;
- 33 – отметки глубин со значением глубины 47 метров;
- 34 – затонувшие суда с атрибутом «Выдающийся объект»;
- 35 – береговые линии, проходящие по границе государства.

Варианты СУБД ГИС Терра:

- v1 – СУБД ГИС Терра;
- v2 – СУБД ГИС Терра Plus;
- v3 – СУБД ГИС Терра М (локальный вариант).

В таблице указывается количество объектов, полученных по запросу/время выполнения запроса в секундах.

Как видно из полученных результатов, наибольшим быстродействием отличается СУБД ГИС Терра, что в значительной степени обусловлено ее специализацией под особенности сложноструктурированных иерархически организованных данных.

СУБД ГИС Терра Plus проигрывает в быстродействии СУБД ГИС Терра (особенно на массовых операциях выборки). Однако в отличие от СУБД ГИС Терра она обеспечивает полноценный клиент-серверный режим работы, практически снимает ограничения на объем данных и обладает усовершенствованной моделью данных.

Временные характеристики СУБД ГИС Терра М в значительной степени обусловлены особенностями технических средств мобильного устройства и возможностями СУБД SQLite, используемой в качестве базовой.

Полученные результаты показывают наиболее перспективные сферы использования представленных вариантов СУБД:

СУБД ГИС Терра наиболее эффективна для ведения локальных баз данных ПРД на ограниченные территории;

СУБД ГИС Терра Plus предназначена для многопользовательской работы с базами данных на большие территории;

СУБД ГИС Терра М в локальном варианте предназначена для автономной работы с базами данных на небольшие территории (например, план города), а в удаленном варианте для оперативного получения информации из базы данных на сервере и для решения задач интерактивного сбора данных и помещения их в единую базу данных (например, при геодезических работах).

Заключение

Использование единой информационной модели представления объектов ПРД, а также унифицированных средств управления данными обеспечивает эффективную миграцию данных между различными программно-аппаратными платформами и приложениями и повышает эффективность решения задач анализа и обработки ПРД с использованием средств ГИС Терра.

Общую схему миграции данных и доступа к данным (локального и удаленного) с использованием

распределенной СУБД для ПРД в среде ГИС Терра можно представить так, как изображено на рис. 10.

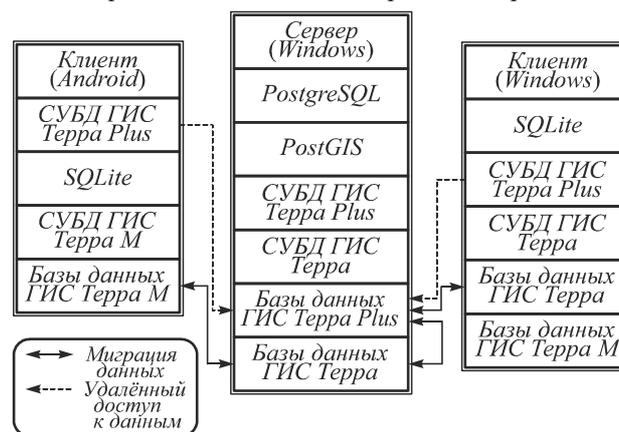


Рис. 10. Схема миграции и доступа к данным

Гибкая адаптируемая архитектура разработанной СУБД обеспечивает принятие эффективных решений при создании, разработке и эксплуатации сложных программно-аппаратных комплексов, обеспечивающих работу с ПРД.

Отметим основные особенности разработанной распределенной СУБД для обработки пространственных и/или распределенных данных:

- поддерживается объектно-ориентированная сложноструктурированная топологическая модель ПРД;
- вся информация об отдельном объекте содержится в одной структурированной записи. Этим обеспечивается быстрое получение любого объекта как единого целого (например, дороги со всеми придорожными сооружениями);
- объект базы данных имеет динамическую структуру, адаптируемую под потребности конкретных приложений;
- объект допускает хранение любой (в том числе и мультимедийной) информации;
- отсутствуют ограничения на структуру объектов в базах данных, что существенно облегчает процесс отображения объектов реального мира: нет необходимости разбивать их на части для того, чтобы поместить в различные отношения, как это, например, требуется в реляционных СУБД;
- ограничения на общее количество объектов в базе данных и на размеры этих объектов обуславливаются только размером базы;
- средства защиты как базы данных в целом, так и ее объектов (вплоть до одного отдельного объекта) от несанкционированного доступа, обеспечивают реализацию распределенного доступа к информации;
- используется ключевая форма задания атрибутов объекта. Это позволяет, во-первых, иметь в объекте любое переменное количество атрибутов и, во-вторых, иметь произвольный порядок следования атрибутов в объекте. В этом заключается отличие от моделей данных большинства других (ориентированных на запись) СУБД, где на структуру записи накладываются довольно жесткие ограничения;

- используется самоопределенная форма представления атрибутов. Все атрибуты содержат в себе свое полное описание, в том числе и тип ее значения;
- имеется возможность рекурсивного кодирования сложных атрибутов для структурированного описания многоаспектных атрибутов объекта;
- может использоваться аппарат пространственной координатной «привязки» любого атрибута, позволяющий эффективно описывать пространственно-распределенные атрибуты. Каждый атрибут может иметь свое метрическое описание;
- поддерживается тип характеристики «База данных». С ее помощью реализуются ситуации, когда объект в базе данных должен иметь не только семантическое и метрическое описание, но и описание некоторой совокупностью других объектов. Например, объект «Дом» на плане города может представлять собой целую базу данных, содержащую описание его элементов (этажи, комнаты, системы коммуникаций и т. п.). Эта база данных может храниться непосредственно в объекте, но допускается хранить ссылку на нее;
- имеется возможность установления пространственно-логических и тополого-метрических отношений между любыми объектами;
- действует инвариантность по отношению к предметным областям. Все взаимодействие с объектами реального мира в данной предметной области осуществляется через специальную базу знаний предметной области (классификатор).

Средства разработанной СУБД обеспечивают:

- иерархический доступ к объектам базы данных;
- многоверсионное хранение всех элементов (полей) объекта;
- эффективное кодирование метрического описания атрибутов и объектов;
- ведение специального журнала работ с объектом, содержащего информацию о произведенных с ним (и его полями) действиях по обновлению;
- связь объекта с метаданными источника, с которого получен объект (топографические и морские карты, планшеты гидрографических съемок и т.п.);
- возможность выполнения SQL-запросов наряду с не SQL-запросами (запросами на специальном языке);
- унифицированную поддержку программных средств для работы с объектами, полученными из базы данных;

- миграцию баз данных между различными программно-аппаратными платформами.

Благодарности

Работа выполнена за счёт гранта Российского научного фонда (РНФ) – проект №16-11-00068.

Литература

1. **Васин, Ю.Г.** Информационная модель изображений в автоматизированных системах обработки графической информации / Ю.Г. Васин, С.А. Дмитриев, Р.Ю. Кобрин // Доклады II-й Всесоюзной конференции «Автоматизированные системы обработки изображений (АСОИЗ-86)», г. Львов, сентябрь 1986. – 1986. – С. 6-8.
2. **Васин, Ю.Г.** Информационно-терминологическое обеспечение банка картографических данных / Ю.Г. Васин, С.А. Дмитриев, Р.Ю. Кобрин. – В кн.: Научно-техническая информация. – Серия 2. – М.: ВИНТИ, НТИ, 1989. – С. 20-28.
3. **Vasin, Yu.G.** GIS Terra: A graphic database management system / Yu.G. Vasin, Yu.V. Yasakov // Pattern Recognition and Image Analysis. – 2004. – Vol. 14(4). – P. 579-586.
4. **Васин, Ю.Г.** Система управления базами видеоданных / Ю.Г. Васин, Ю.В. Ясаков. – В кн.: Методы и средства обработки графической информации: Межвузовский сборник научных трудов. – Горький: ГГУ им. Н.И. Лобачевского, 1989. – С. 93-115.
5. **Эбби, М.** Oracle 8: Первое знакомство: Полное введение в базы данных Oracle; пер. с англ. / М. Эбби, М. Кори. – М.: Лори, 1998. – 470 с. – ISBN: 5-85582-039-4.
6. **Стоунз, Р.** PostgreSQL. Основы / Р. Стоунз, Н. Мэттью. – Символ-Плюс, 2002. – ISBN 5-93286-043-х.
7. PostGIS. Руководство пользователя. – URL: <http://gis-lab.info/docs/postgis/manual/>.
8. **Vasin, Yu.G.** Electronic archive of nautical navigation charts with remote access / Yu.G. Vasin, Yu.V. Yasakov // PRIA-8-2007. 8th International Conference on Pattern Recognition and Image Analysis: New Information Technologies, Yoshkar-Ola, October 8-13, 2007. – Vol. 2. – P. 211-212.
9. СУБД ГИС Терра. Руководство пользователя. – Книга 1 – Нижний Новгород: НИИ ПМК ННГУ им. Н.И. Лобачевского, 2004.
10. **Васин, Ю.Г.** Система управления базами пространственно-распределенных данных на основе объектно-ориентированной топологической модели / Ю.Г. Васин, Ю.В. Ясаков. – Свидетельство о государственной регистрации программы для ЭВМ в Реестре программ для ЭВМ № 2016616992.

Сведения об авторах

Васин Юрий Григорьевич, 1940 года рождения, в 1962 году окончил Горьковский государственный университет им. Н.И. Лобачевского. Получил степень доктора наук в 1988 году. Звание профессора – в 1994 году. Заслуженный деятель науки и техники РФ, Почетный работник Высшей школы, Лауреат премии Совета Министров СССР. Руководитель Центра информатики и интеллектуальных информационных технологий ИИТММ Нижегородского государственного университета им. Н.И. Лобачевского. Область научных интересов: теоретическая и прикладная информатика, распознавание образов и обработка изображений, информационные технологии. Автор более 150 публикаций. E-mail: ya.vasinyuri@yandex.ru.

Ясаков Юрий Васильевич, 1949 года рождения, в 1971 году окончил факультет вычислительной математики и кибернетики Горьковского государственного университета им. Н.И. Лобачевского по специальности «Вычисли-

тельная математика». Лауреат премии Совета Министров СССР. Работает старшим научным сотрудником в Центре информатики и интеллектуальных информационных технологий ИИТММ Нижегородского государственного университета им. Н.И. Лобачевского. Область научных интересов: системное программирование, системы управления базами данных, геоинформационные системы, цифровая картография. E-mail: yuri961@yandex.ru.

ГРТИ: 50.41.21

Поступила в редакцию 1 августа 2016 г. Окончательный вариант – 24 октября 2016 г.

DISTRIBUTED DATABASE MANAGEMENT SYSTEM FOR INTEGRATED PROCESSING OF SPATIAL DATA IN A GIS

Yu. G. Vasin¹, Yu. V. Yasakov¹

¹ Center for Informatics and Intellectual Information Technologies Institute of Information Technologies, Mathematics and Mechanics, Lobachevsky State University of Nizhni Novgorod, Nizhni Novgorod, Russia

Abstract

The article examines some challenges in the development and processing of data containing complex structured geospatial information that are associated with large amounts of data, their dimensionality and an increased number of spatial/logical and topological relations between data elements, as well as data heterogeneity. The most popular tools for processing such information are described. We discuss the organization of storage, processing and analysis of spatially distributed topology data in the environment of the GIS Terra. The main features (models, structures and data formats) of the developed topological structured object model are examined. We also consider options for organizing the management system for databases of spatially distributed data for different hardware and software platforms (Windows and Android) in the local and client-server variants. Data migration schemes are presented for the integrated use of different hardware and software platforms. Some recommendations are given concerning the development of new applications for the analysis and processing of spatially distributed data using the developed data storage and processing tools.

Keywords: database management system, spatially distributed data, data models, spatial and logical relations, algorithms, data processing, topology.

Citation: Vasin YuG, Yasakov YuV. Distributed database management system for integrated processing of spatial data in a GIS. Computer Optics 2016; 40(6): 919-928. DOI: 10.18287/2412-6179-2016-40-6-919-928.

Acknowledgements: The work was partially funded by the Russian Science Foundation (RSF), grant No. 16-11-00068.

References

- [1] Vasin YuG, Dmitriev SA, Kobrin RYu. Information model of images in automated systems for graphic information processing. Reports of the Second All-Union Conference "Automated image processing systems (ASOIZ-86)", Lvov, September 1986: 6-8.
- [2] Vasin YuG, Dmitriev SA, Kobrin RYu. Information and terminology support for the cartographic data bank. Collection of Scientific and Technical Information. Seria 2. Moscow: "VINITI, NTI" Publisher; 1989: 20-28.
- [3] Vasin YuG, Yasakov YuV. GIS Terra: A graphic database management system. Pattern Recognition and Image Analysis 2004; 14(4): 579-586.
- [4] Vasin YuG, Yasakov YuV. A database management system for video data [In Russian]. Inter-university collection "Methods and tools for graphic information processing". Gorky: "Lobachevsky Gorky State University" Publisher; 1989: 93-115.
- [5] Abbey M, Corey MJ. Oracle 8: A Beginner's Guide. Berkeley, California: Osborne McGraw-Hill, Oracle Press; 1997. ISBN: 0078823935.
- [6] Stones R, Matthew N. Beginning Databases with PostgreSQL. Wrox Press; 2001. ISBN: 9781861005151.
- [7] GisLAB. Manual PostGIS. Source: <http://gislab.info/docs/postgis/manual/>.
- [8] Vasin YuG, Yasakov YuV. Electronic archive of nautical navigation charts with remote access. 8th International Conference on Pattern Recognition and Image Analysis: New Information Technologies, Yoshkar-Ola, Oct 8-13, 2007. 2007; 2: 211-212.
- [9] GIS Terra database management system. User manual. Book 1. Nizhni Novgorod: "Research Institute of Applied Mathematics and Cybernetics of the Lobachevsky State University of Nizhni Novgorod" Publisher; 2004.
- [10] Vasin YuG, Yasakov YuV. A database management system for spatially distributed data based on an object-oriented topological model. Certificate of state registration of the computer program in the Computer Program Register No 2016616992.

Authors' information

Yurii Grigor'evich Vasin. Born 1940. Graduated from Gorky State University in 1962. Received a doctoral degree in 1988. Received a title of Professor in 1994. Honored Worker of Science and Technology of the Russian Federation, Honored Worker of Higher School, and Winner of the USSR Council of Ministers. Head of the Center for Informatics and Intellectual Information Technologies of the Institute of Information Technology, Mathematics and Mechanics of the State University of Nizhny Novgorod. Scientific interests: theoretical and applied computer science, pattern recognition and image processing, and information technology. Author of more than 150 papers. E-mail: ya.vasinyuri@yandex.ru.

Yurii Vasil'evich Yasakov. Born 1949. Graduated from the Faculty of Computational Mathematics and Cybernetics of the Gorky State University in 1971 in specialty «Computational Mathematics». Winner of the USSR Council of Ministers in the field of cybernetics. Works as a senior researcher at the Center for Informatics and Intellectual Information Technologies of the Institute of Information Technology, Mathematics and Mechanics of the State University of Nizhny Novgorod. Scientific interests: system programming, database management system, geographic information systems, digital mapping. E-mail: yuri961@yandex.ru.

Received August 1, 2016. The final version – October 24, 2016.
