



Хранение геометрических данных



Игорь ДОНЦОВ

Igor E. DONTSOV

Storage of Geometric Data
(текст статьи на англ. яз. –
English text of the article – p. 194)

Реализация программы цифровизации связана с развитием информационно-управляющих систем и телекоммуникаций на железнодорожном транспорте, с интенсификацией комплексной автоматизации управления.

В транспортной отрасли свой интерес представляет новый тип ИС – геоинформационные системы (ГИС), которые предназначены для принятия решений по оптимальному управлению транспортом. Получили распространение различные тренажёры и стенды для тренинга, имитирующие движения различных объектов и управление ими. Они применяются в авиации, на железных дорогах и в других сферах транспорта.

Развитие тренажёров связано с разработкой систем отображения визуальной обстановки на базе ЭВМ. В соответствующих операциях основная роль отведена пространственно распределённой геометрической информации. Делается обзор базовых подходов к способам её хранения и передачи через информационные сети.

Ключевые слова: транспорт, геоинформационные системы, геометрические данные, SQL, СУБД.

Донцов Игорь Евгеньевич – инженер ООО «Газпром трансгаз Сургут», Сургут, Россия.

Системы хранения данных очень востребованы в современном мире. С каждым годом объём обрабатываемой информации существенно возрастает. Этому способствуют рост производительности и прогресс электронных устройств и сетей связи.

Требования, предъявляемые к системам хранения, остаются неизменными:

- надёжность – записанные данные не должны быть испорчены;
- доступность – разрешённым пользователям должен быть обеспечен бесперебойный доступ к системе;
- производительность – возможность работы с большим количеством клиентов и большими объёмами данных (основными показателями производительности являются количество одновременных операций ввода/вывода и время их выполнения);
- масштабируемость – способность увеличивать объём памяти и количество поддерживаемых клиентов без ущерба для других характеристик системы.

При создании базы данных стремятся упорядочить информацию по различным признакам для быстрого её извлечения. Это возможно, если данные структурированы [1].

Структурирование – процесс группировки данных по определённым параметрам [2].

К неструктурированным данным можно отнести: информацию из социальных сетей, XML, видео/аудиофайлы, изображения, документы формата PDF [1].

В связи с хранением в базах данных стратегически важной информации и возникающими при этом вопросами информационной безопасности ключевым моментом становится использование системы управления базой данных (СУБД), основанной на оригинальных и надёжных разработках. Этим и определяется актуальность темы исследования [3].

Графическая база данных геометрических объектов (БД ГО) с опорой на метод MRO [4] будет содержать следующую информацию: структуру-матрицу ГО, R-descriptor, коды кусков, описанных известными способами (например, методом специального контура), описательные параметры [5].

Геометрический объект – базовый класс для всех остальных, единственный абстрактный (неинстанцируемый) класс, все прочие – инстанцируемые.

Согласованность и доступность при отсутствии устойчивости к разделению – к этому классу относятся все традиционные реляционные базы данных [6].

Иллюстрация к алгоритму обработки и передачи по сети составных ГО на основе MRO представлена на рис. 1.

Условно можно выделить два подхода для доступа к графическим данным (рис. 2) [4]: соответственно, record- и set-ориентированный.

Второй подход используется в SQL-серверах БД [2]. Язык БД обычно имеет две составляющие:

- язык определения схемы БД (Schema Definition Language – SDL) [7];

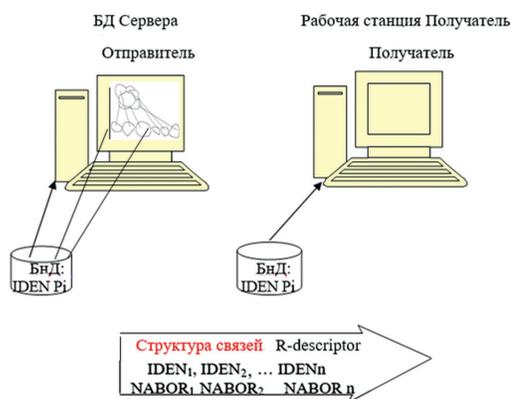


Рис. 1. Иллюстрация к алгоритму обработки и передачи по сети составных ГО на основе MRO.

- язык манипулирования данными (Data Manipulation Language – DML).

Наиболее распространённым для разных типов данных является декларативный язык SQL (Structured Query Language), который поддерживает средства SDL и DML реляционных СУБД. В компьютерных сетях SQL присутствует наряду с некоторыми протоколами и интерфейсами сеансового уровня:

- NFS (Network File System) – сетевая файловая система, используется на станциях UNIX вместе с TCP/IP;
- RPC (Remote Procedure Call) – вызов удалённых процедур, которые создаются на ПК клиента и выполняются на сервере;
- Xwindows – для связи с удалёнными ПК UNIX;
- DNASCP – протокол сеансового уровня в сетях DECnet, применяется широко.

Существуют три формы SQL: интерактивный (Interactive), статический (static) и динамический (dynamic). Функционируют они в основном одинаково, но используются по-разному. Интерактивный составляет основу языка [8].



Рис. 2. Подходы для доступа к графическим данным.

Рейтинг баз данных

Модели	Реляционная	Ключ значения	Распределённое хранилище	Документо-ориентированная	На основе графов
1	Oracle	Redis	Cassandra	MongoDB	Neo4j
2	MySQL	Memcached	HBase	Amazon DynamoDB	Titan
3	Microsoft SQL Server	Riak KV	Microsoft Azure Table Storage	Couchbase	Giraph
4	PostgreSQL	Hazelcast	Hypertable	CouchDB	InfiniteGraph
5	DB2	Enhache	Google Cloud Bigtable	RethinkDB	Dgraph

Источник информации: <https://db-engines.com/en/>.

Таблица 2

matr	OB1	OB2	OB3

Для некоторых специфических проектов больше подходят новые модели хранения данных. Они получили название нереляционных или NoSQL. На данный момент существуют уже сотни различных СУБД, каждая из которых имеет свои преимущества и недостатки и подходит для определённых задач. И чтобы иметь возможность подобрать наилучший инструмент для поставленной задачи, необходимо иметь инструменты тестирования и сравнительного анализа баз данных. Выбирать технологию хранения данных следует при наличии полной информации о каждом виде баз данных и испытав их для своей задачи. Из всех моделей наиболее востребованной является реляционная, которая строится на отношениях между хранимой информацией. Результаты сгруппированы по выбранным моделям и представлены в таблице 1 [9].

Стоит задача выбора СУБД для хранения геометрических объектов, заданных методом MRO [4]. Проведём анализ СУБД по образцу, представленному в таблице 2, содержащему матрицу связи ГО и простейшие геометрические объекты (в заполненном виде не приводится ввиду большого объёма и неформатируемой для публикации информации) [10].

В результате анализа как вариант для обмена и хранения геометрической информации предлагается Microsoft SQL Server [11].

Достоинства SQL – это прежде всего лёгкий для понимания язык и в то же время

универсальное программное средство управления данными.

Успех языку SQL принесли следующие его особенности:

- независимость от конкретных СУБД;
- переносимость с одной вычислительной системы на другую;
- наличие стандартов;
- одобрение компанией IBM (СУБД DB2);
- поддержка со стороны компании Microsoft (протокол ODBC);
- реляционная основа;
- высокоуровневая структура, напоминающая английский язык;
- возможность выполнения специальных интерактивных запросов:
- обеспечение программного доступа к базам данных;
- возможность различного представления данных;
- полноценность языка, предназначенного для работы с базами данных;
- возможность динамического определения данных;
- поддержка архитектуры клиент/сервер.

Все перечисленные факторы явились причиной того, что SQL стал стандартным инструментом для управления данными на персональных компьютерах, мини-компьютерах и больших ЭВМ [12].

В таблице 3 представлен общий вид заполняемой для конкретных случаев матрицы алгоритмизации операций, включающей коды кусков заданной поверхности и алгоритмы их реализации разными методами.

В результате реализации алгоритмов выбранным методом по сети передаётся R-скриптор – способ закраски параметров сетевого транспорта с указанием, в том числе, следующих данных:

Хранение геометрической информации

Коды кусков	Алгоритм реализации метод 1	Алгоритм реализации метод 2	Алгоритм реализации метод 3	Алгоритм реализации метод 4

• state – состояние объекта, сведения о блокировках фрагмента [13];

• UserID – идентификатор режима работы пользователя.

Дальнейшая обработка переданных и сохранённых изображений опирается на несколько последовательных уровней восходящей информационной линии («иерархическое представление объектов (растровое изображение, неструктурированная информация) – символическое представление (векторные и атрибутивные данные в структурированной форме, реляционные структуры)») и осуществляется посредством следующих этапов обработки:

- предобработка изображения;
- первичная сегментация изображения;
- выделение геометрической структуры видимого поля;
- определение относительной структуры и семантики видимой сцены [15].

Представленный обзор резюмирует шаги по реализации хранения ГО, заданных методом MRO [14].

В заключение следует отметить, что другие способы задания поверхностей геометрических объектов, кроме MRO, имеют довольно трудоёмкие алгоритмы обработки, малоприменимые для построения изображений сцен, так как не обеспечивают требуемого запаздывания в замкнутой системе управления. И это оставляет рассмотренные исходные вопросы актуальными для дальнейшей проработки в интересах транспортной отрасли.

ЛИТЕРАТУРА

1. Коптева Л. Г. Подходы к информационной обработке математических моделей геометрических объектов при формировании базы данных для ГИС // Высшее профессиональное образование на железнодорожном транспорте: настоящее и будущее: Сб. науч. трудов по материалам международной конференции. – М.: РГОТУПС, 2001. – С. 142–144.

2. Роб П., Коронел К. Системы баз данных: проектирование, реализация и управление: Пер. с англ. – 5-е изд. – СПб.: БХВ–Петербург, 2004. – 1040 с.

3. Борисенков Д. В. Специальное математическое и программное обеспечение манипулирования распределёнными объектами в реляционной СУБД на основе политомических представлений / Дис... канд. техн. наук. – Воронеж: ВГУ, 2007. – 145 с.

4. Коптева Л. Г. Построение моделей геометрических объектов, их обработка и передача в транспортных информационных сетях / Дис... док. техн. наук. – М.: РОАТ, 2005. – 189 с.

5. Рвачев В. Л. Теория R-функций и некоторые её приложения. – Киев: Наукова думка, 1982. – 552 с.

6. Сборник трудов конференции «Управление знаниями и технологии семантического веба-2010». – СПб.: СПбГУ ИТМО, 2010. – 187 с.

7. Лецкий Э. К., Панкратов В. И., Яковлев В. В. и др. Информационные технологии на железнодорожном транспорте: Учебник. – М.: УМК МПС России, 2001. – 668 с.

8. Скорбилина Т. С. Разработка программного модуля тестирования баз данных // Молодой учёный. – 2017. – № 10. – С. 34–37. [Электронный ресурс]: <https://moluch.ru/archive/144/40328/>. Доступ 12.12.2018.

9. Маличенко Д. А. Разработка и исследование методов хранения и передачи информации в распределённых системах / Дис... канд. техн. наук. – СПб.: ГУАП, 2017. – 125 с.

10. Рвачев В. Л. Геометрические приложения алгебры логики. – Киев: Техника, 1968. – 287 с.

11. Qin Yong, Liu Feng, Zhou Zhi-min, Lai Wen-bin, Jia Li-min. Исследование областей применения ГИС в Китайском управлении железными дорогами // Zhongguo tiedao kexue. – 2003. – № 24. – р. 7.

12. Тетеркина Н. И. Разработка базы данных для страхового бизнеса на примере ОАО «Ренессанс» / Дипломный проект. – М., 2018. – 50 с.

13. Райан Д. Инженерная графика в САПР: Пер с англ. – М.: Мир, 1989. – 392 с.

14. Смирнов С. Н. Разработка информационных моделей для ГИС (графических информационных систем) железнодорожного транспорта // Материалы XII международной конференции по компьютерной графике и машинному зрению «Графикон-2002». – Н. Новгород, 2002. – С. 403–405.

15. Техническое зрение в системах управления мобильными объектами-2010: Труды научно-технической конференции-семинара. Вып. 4 / Под ред. Р. Р. Назирова. – М.: КДУ, 2011. – 328 с.

16. Тимошенко В. В. Лабораторные работы по администрированию БД / Дипломный проект. – М., 2018. – 183 с.

17. Kopteva L. G. The Questions of Geometry and Programme-Information Supports for Computer-Aided Design Systems // Information Technology in Design. EWITD96. Proceedings of International Conference. – Moscow, Russia, 1996. – pp. 318–322. ●

Координаты автора: **Донцов И. Е.** – dontsov2011@mail.ru.

Статья поступила в редакцию 03.12.2018, принята к публикации 21.01.2019.

