

УДК 681.518.(075.32)

**СТРУКТУРА БАЗ ДАННЫХ ДЛЯ ГЕОИНФОРМАЦИОННЫХ СИСТЕМ***Лавникович Д.С., Чеменцова А.В., Городная Ю.С.**Белорусский государственный университет информатики и радиоэлектроники,  
г. Минск, Республика Беларусь**Научный руководитель: Шалькевич П.К. – канд.техн.наук, доцент, доцент кафедры ЭИ*

**Аннотация.** В результате работы над данной статьей, был произведен анализ существующих структур баз данных. В статье приведены аргументы о том, какая структура организации информации в базе данных наиболее подходящая для геоинформационных систем. В статье описаны типы пространственных объектов, являющиеся основаниями, на базе которых построены требования к структуризации информации в базе данных.

**Ключевые слова:** база данных, пространственные данные, атрибутивные данные, реляционная база данных, геоинформационные системы

**Введение.** Область применения геоинформационных систем охватывает широкий спектр важных задач — от определения оптимального маршрута до анализа проблем экологии и перенаселения. Основным элементом функционирования геоинформационных систем являются базы данных. Правильно спроектированная база данных позволяет сохранять всю информацию, необходимую для работы с геоинформационной системой, и подготовить эту информацию для дальнейшей обработки, будь то, системный анализ или математическое моделирование.

**Основная часть.** Современные геоинформационные системы позволяют пользователю работать с пространственными объектами. Под пространственными объектами понимаются формализованные цифровые модели реальных географических объектов. Геоинформационные системы, в основном, оперируют 4 видами данных: пространственными данными, которые содержат информацию о пространственном положении объектов, атрибутивными данными, описывающими качественные и количественные свойства, библиотекой условных знаков, где содержатся наборы стандартных условных данных, и метаданными, которые, как правило, содержат информацию об источниках данных, методах получения и конкретных исполнителях, получивших данные.

Первый тип, пространственные данные, разделяются на множественные элементарные объекты-примитивы (точки, линии, контуры и т.д.). Представление таких данных подразумевает способы цифрового описания объектов, наиболее универсальными из которых являются: растровое представление, векторное, *GRID*-представление и *TIN*.

Растровая модель данных предполагает позиционирование объектов указанием их положения в соответствующей растру прямоугольной матрице. Данные для такой модели очень просто собирать и обрабатывать, а также для таких представлений возможны процессы rasterизации. Но один из недостатков модели – ее значительный объем. Поэтому, чтобы хранить представления такого рода в базе данных необходимо использовать способы сжатия растровых объектов (групповое кодирование, метод цепочного кодирования, блочное кодирование, метод квадродерева) [1].

Следующий вид – это векторная модель, которая является цифровым представлением точечных, линейных и полигональных объектов в виде набора координатных пар. Для представления объекта в векторной нетопологической модели в базе должна содержаться информация об идентификаторе линейного отрезка, координатных парах шести точек, а также элемент, позволяющий выделить заданный отрезок в общей совокупности записей линейных объектов слоя, которому обычно соответствует файл данных.

*GRID*-представление необходимо для трёхмерных объектов (рельефов, определяемых не только координатами  $x$  и  $y$ , но еще и аппликатой  $z$ ). Такое название характеризует регу-

лярная сетка, в основу которой положена сеть точек, каждой из которых сопоставлено значение уровня поля в этой точке. При хранении регулярной сетки можно хранить только значения уровня  $z$ , т.к. значения координат  $x$  и  $y$  вычисляются по геометрии сетки, зная номер ячейки. Для того, чтобы создать непрерывную поверхность из точечно полученных данных необходимо наличие у данных следующих форм: положение точек наблюдения на сетках, полурегулярное положение точек, расположенных на изолиниях, случайное расположение точек.

*TIN*-представление – векторный способ отображения рельефов, который основан на триангуляции, т.е. создании смежных непересекающихся треугольников, вершинами которых являются точки [1].

Атрибутивные данные хранят информацию о свойствах самих пространственных объектов. Чтобы полностью отобразить свойства объектов реального мира, необходимо записать эти свойства в базу данных.

Для составления тематических карт в геоинформационных системах также используют библиотеку условных знаков.

Основываясь на типе пространственных объектов, можно определить принципы организации данных:

- принцип послойной организации. В нем используется некоторое деление объектов на тематические слои. Объекты, отнесенные к одному слою, образуют некоторую логически отдельную единицу данных;
- объектно-ориентированный. Делает акцент не столько на общих свойствах объектов, сколько на их положении в какой-либо сложной схеме классификаций.

Разумеется, для хранения информации о пространственных объектах используются базы данных. При этом, важно учесть, что связи между данными одного объекта должны быть тесными и избыточными. При большом объеме координат, которые необходимо сохранить, производится их кодирование [2].

На основании типа хранимых объектов и их представления, к оптимальной структуре базы данных выдвигаются следующие требования:

- полнота представленной информации. Модель должна быть достаточно подробной для предполагаемых целей анализа и проектирования;
- избыточность и внутренняя непротиворечивость. Данные и связи между ними должны быть корректными и неповторяющимися;
- актуализация хранимых данных;
- позиционная точность. Информация в базе данных должна быть строго совместима с теми данными, которые могут добавляться в неё в процессе работы;
- достоверность. Все явления должны быть правильно отражены посредством атрибутов;
- доступность данных. Информация должна быть легко извлекаемой для выбранной категории пользователей [3].

На сегодняшний день выделяют три основных типа структур баз данных: иерархическая, сетевая и реляционная.

При иерархической структуре взаимосвязь между данными устанавливается отношением «один к одному» (1:1) (один объект базы данных содержит одну группу пространственных данных) или, гораздо чаще, «один ко многим» (1:m) (одной группе координат соответствует иерархически подчинённая группа объектов). Так как связь между элементами такой базы данных прямая, то доступ к данным прост и эффективен. Однако древовидная структура требует явного определения каждого элемента при проектировании базы данных. Также, при раздельном хранении пространственных и атрибутивных данных, потребуется установление большого числа связей между графической и атрибутивной частями, что может сделать структуру базы данных довольно громоздкой. Из-за иерархической соподчинённости составляющих обновление базы данных также представляется сложным и трудоёмким процессом.

Сетевые структуры обычно рассматриваются как усовершенствованные иерархические. Они используют отношение «многие ко многим» (m:n) [4]. За счёт установления двусторонних связей между элементами уменьшается избыточность данных. Однако увеличившееся количество связей в данной структуре имеет негативную сторону в виде затрат памяти на хранение, возможности потери или некорректного установления взаимосвязей между элементами.

Проблему большого числа указателей (связей) можно решить, используя реляционную структуру базы данных. Все данные такой структуры рассматриваются как хранимые в таблицах, в которых каждая строка имеет один формат. Основной принцип реляционной структуры – это возможность свести любые многомерные связи к двумерным таблицам. Данные в ячейках такой таблицы соответствуют принципу информационной неделимости (каждая ячейка представляет только одну порцию структурно неделимых данных). Реляционная структура обладает теми преимуществами, которых были лишены структуры, описанные выше, а именно: однородностью, наличием строгой математической теории построения модели (языком реляционной алгебры), полнотой модели данных, возможностью представления всех типов связей от 1:1 до m:n. Из недостатков у реляционной структуры выделяют лишь большие затраты на реализацию, однако они, как правило, полностью окупаются лёгкостью поддержки базы данных и её обновления [5]. Важным достоинством в контексте геоинформационных систем является простота связывания данных при раздельном хранении их графической и атрибутивной составляющей посредством идентификаторов.

При использовании реляционной структуры базы данных мы также располагаем широким диапазоном систем управления базами данных. СУБД *dBase*, *FoxBASE*, *R:Base*, *ORACLE*, *Informix*, *Clipper*, *Paradox*, *MS Access* и *SQL* предоставляют все необходимые средства для работы с реляционными базами данных.

Об определённом объекте должна храниться следующая информация: идентификатор, атрибутивные данные и пространственные данные. Существует несколько способов связать атрибутивную и пространственную информацию.

Первый способ, геореляционный заключается в том, что два вида данных хранятся раздельно (например, пространственные данные - в файлах, атрибутивные – в таблицах реляционных баз данных).

Второй способ предполагает совместное хранение и пространственных, и атрибутивных данных. В случае интегрированного хранения обоих типов предусматривается использование средств обработки данных, предоставляемых СУБД. Однако, как правило, инструментов стандартных реляционных СУБД недостаточно для эффективной работы с пространственными объектами. Средства, встроенные в систему управления базами данных, предназначены, в основном, для описания точечных объектов, что не соответствует современным требованиям, предъявляемым к геоинформационным системам [5].

Третий способ – объектный подход. Данный способ позволяет легко описывать сложные структуры, однако этот подход крайне тяжело внедрять в геоинформационные системы, поэтому он практически не применяется в данной сфере.

Таким образом, в контексте стандартных СУБД, первый способ является наиболее оптимальным, так как мы можем сохранять геометрические, пространственные данные отдельно от атрибутивных и использовать для работы с ними эффективные и удобные специализированные инструменты. Однако, появляются и внедряются новые инструменты для работы с пространственными данными в виде расширений, существующих и использующихся СУБД.

Примером такого расширения служит надстройка над системой управления базами данных *PostgreSQL - PostGIS*. *PostGIS* добавляет поддержку географических объектов в объектно-реляционную базу данных *PostgreSQL*. По сути, *PostGIS* «пространственно расширяет» сервер *PostgreSQL*, позволяя ему использоваться в качестве промежуточной пространственной базы данных для географических информационных систем.

*H2GIS* – ещё одно пространственное расширение СУБД H2 в духе *PostGIS*. Он добавляет поддержку для управления пространственными функциями и операциями, включая ограничения геометрии, простые функции открытого геопространственного консорциума (*OGC*) для *SQL* (*SFSQL*) и дополнительные пространственные функции.

**Заключение.** Выполнен анализ структур баз данных (иерархической, сетевой и реляционной) и, на основании приведённых типов данных, аргументирован вывод о том, что в сфере геоинформационных систем наиболее подходящей является реляционная структура базы данных. Также рассмотрены технологии организации хранения геометрических и атрибутивных данных о пространственном объекте и установлено, что наиболее эффективным подходом является раздельное хранение этих данных либо использование специализированных СУБД.

### **Список литературы**

1. Ананьев, Ю.С. Геоинформационные системы: Учебное пособие / Ю.С. Ананьев. – Томск: Изд. ТПУ, 2003. – 70 с.
2. Ципилева, Т.А. Геоинформационные системы: Учебное пособие / Т.А. Ципилева. – 2-е изд. – Томск: Томский межвузовский центр дистанционного образования, 2004. – 162 с;
3. Самардак, А.С. Геоинформационные системы / А.С. Самардак. – Дальневосточный Федеральный Университет, Владивосток, 2005 г. – 122 с.
4. Карпова, И.П. Базы данных. Курс лекций и материалы для практических заданий / И.П. Карпова. – М.: Питер, 2013. – 240 с.
5. Тарасов, С.В. СУБД для программиста. Базы данных изнутри / С.В. Тарасов. – М.: СОЛОН-Пресс, 2015. – 320 с.

UDC 681.518.(075.32)

## **THE STRUCTURE OF DATABASES IN GEOGRAPHIC INFORMATION SYSTEMS**

*Lavnikovich D.S., Chementsova A.V., Gorodnaya Yu.S.*

*Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus*

*Shalkevich P.K. – PhD, associate professor, associate professor of the Department of EI*

**Annotation.** As a result of work on this article, an analysis of existing database structures was made. The article presents arguments about what structure of information organization in the database is the most suitable for geoinformation systems. The article describes the types of spatial objects that are the basis on which the requirements for structuring information in the database are built.

**Keywords:** database, spatial data, attribute data, relational database, geoinformation systems