

Лекция №5. СТРУКТУРЫ И МОДЕЛИ ДАННЫХ

1.1 Отображение объектов реального мира в ГИС

1.2 Структуры данных

1.3 Модели данных

1.4 Форматы данных

5.1 Отображение объектов реального мира в ГИС

Объекты реального мира, рассматриваемые в геоинформатике, отличаются *пространственными, временными и тематическими характеристиками*.

Пространственные характеристики определяют положение объекта в заранее определенной системе координат, основное требование к таким данным – точность.

Временные характеристики фиксируют время исследования объекта и важны для оценки изменений свойств объекта с течением времени. Основное требование к таким данным – актуальность, что означает возможность их использования для обработки, неактуальные данные – это устаревшие данные.

Тематические характеристики описывают разные свойства объекта, включая экономические, статистические, технические и другие свойства, основное требование – полнота.

Для представления пространственных объектов в ГИС используют *пространственные и атрибутивные типы данных*.

Пространственные данные – сведения, которые характеризуют местоположение объектов в пространстве относительно друг друга и их геометрию.

Пространственные объекты представляют с помощью следующих графических объектов: точки, линии, области и поверхности.

Описание объектов осуществляется путем указания координат объектов и составляющих их частей.

Точечные объекты – это такие объекты, каждый из которых расположен только в одной точке пространства, представленной парой координат X, Y . В зависимости от масштаба картографирования, в качестве таких объектов могут рассматриваться дерево, дом или город.

Линейные объекты, представлены как одномерные, имеющие одну размерность – длину, ширина объекта не выражается в данном масштабе или не существенна. Примеры таких объектов: реки, границы муниципальных округов, горизонтали рельефа.

Области (полигоны) – площадные объекты, представляются набором пар координат (X, Y) или набором объектов типа линия, представляющих собой за-

мкнутый контур. Такими объектами могут быть представлены территории, занимаемые определенным ландшафтом, городом или целым континентом.

Поверхность – при ее описании требуется добавление к площадным объектам значений высоты. Восстановление поверхностей осуществляется с помощью использования математических алгоритмов (интерполяции и аппроксимации) по исходному набору координат X, Y, Z .

Дополнительные непространственные данные об объектах образуют набор атрибутов.

Атрибутивные данные – это качественные или количественные характеристики пространственных объектов, выражающиеся, как правило, в алфавитно-цифровом виде.

Примеры таких данных: географическое название, видовой состав растительности, характеристики почв и т.п.

Природа пространственных и атрибутивных данных различна, соответственно различны и методы манипулирования (хранения, ввода, редактирования, поиска и анализа) для двух этих составляющих геоинформационной системы. Одна из основных идей, воплощенных в традиционных ГИС – это сохранение связи между пространственными и атрибутивными данными, при раздельном их хранении и, частично, раздельной обработке.

Общее цифровое описание пространственного объекта включает: наименование; указание местоположения; набор свойств; отношения с другими объектами. Наименованием объекта служит его географическое название (если оно есть), его условный код или **идентификатор**, присваиваемый пользователем или системой.

Однотипные объекты по пространственному и тематическому признакам объединяются в **слои цифровой карты**, которые рассматриваются как отдельные информационные единицы, при этом существует возможность совмещения всей имеющейся информации.

5.2 Структуры данных

Для представления пространственных данных в ГИС применяют **векторные и растровые структуры данных**.

Векторная структура – это представление пространственных объектов в виде набора координатных пар (векторов), описывающих геометрию объектов (рисунок 4).

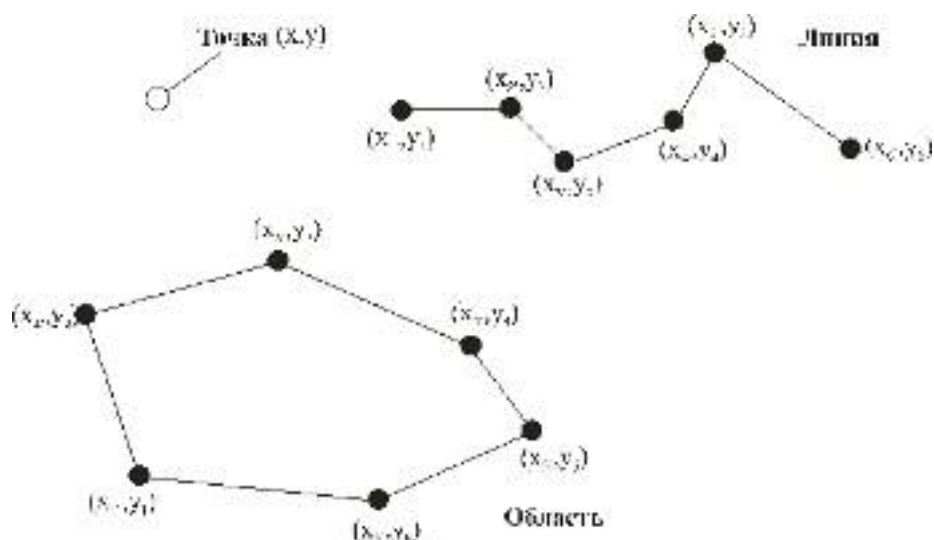


Рисунок 4 – Векторное представление пространственных данных

Растровая структура данных предполагает представления данных в виде двумерной сетки, каждая ячейка которой содержит только одно значение, характеризующее объект, соответствующий ячейке раstra на местности или на изображении. В качестве такой характеристики может быть код объекта (лес, луг и т.д.) высота или оптическая плотность (рисунок 5).

Точность растровых данных ограничивается размером ячейки. Такие структуры являются удобным средством анализа и визуализации разного рода информации.

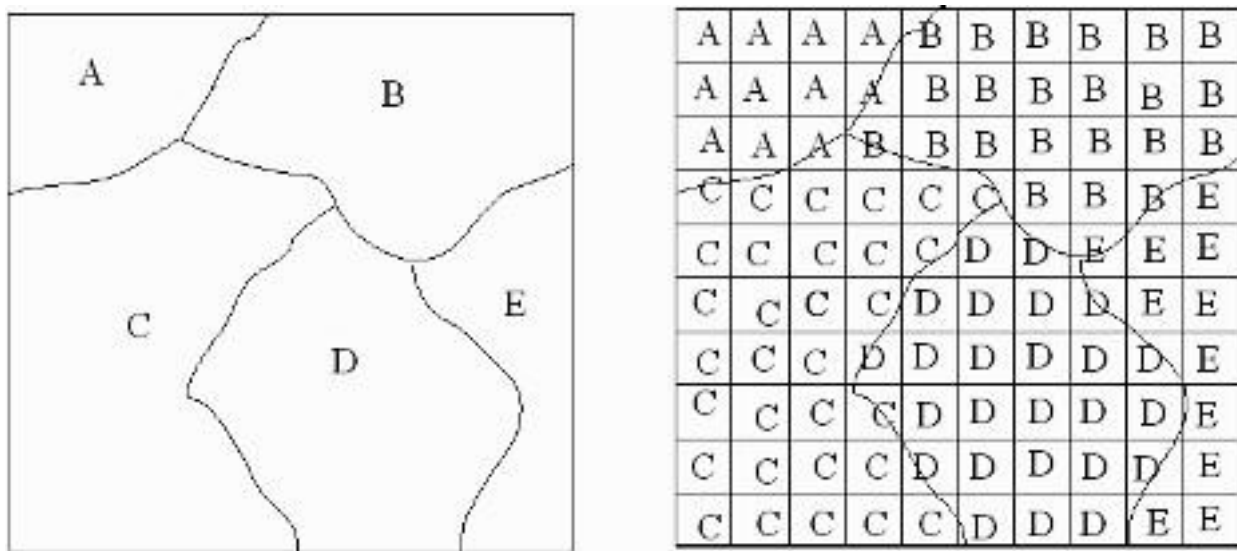


Рисунок 5 – Растровая структура данных.

Для реализации растровых и векторных структур разработаны различные модели данных.

5.3 Модели данных

Модели пространственных данных – логические правила для формализованного цифрового описания пространственных объектов.

Векторные модели данных. Существует несколько способов объединения векторных структур данных в векторную модель данных, позволяющую исследовать взаимосвязи между объектами одного слоя или между объектами разных слоев. Простейшей векторной моделью данных является «спагетти»- модель (рисунок 6). В этом случае переводится «один в один» графическое изображение карты.

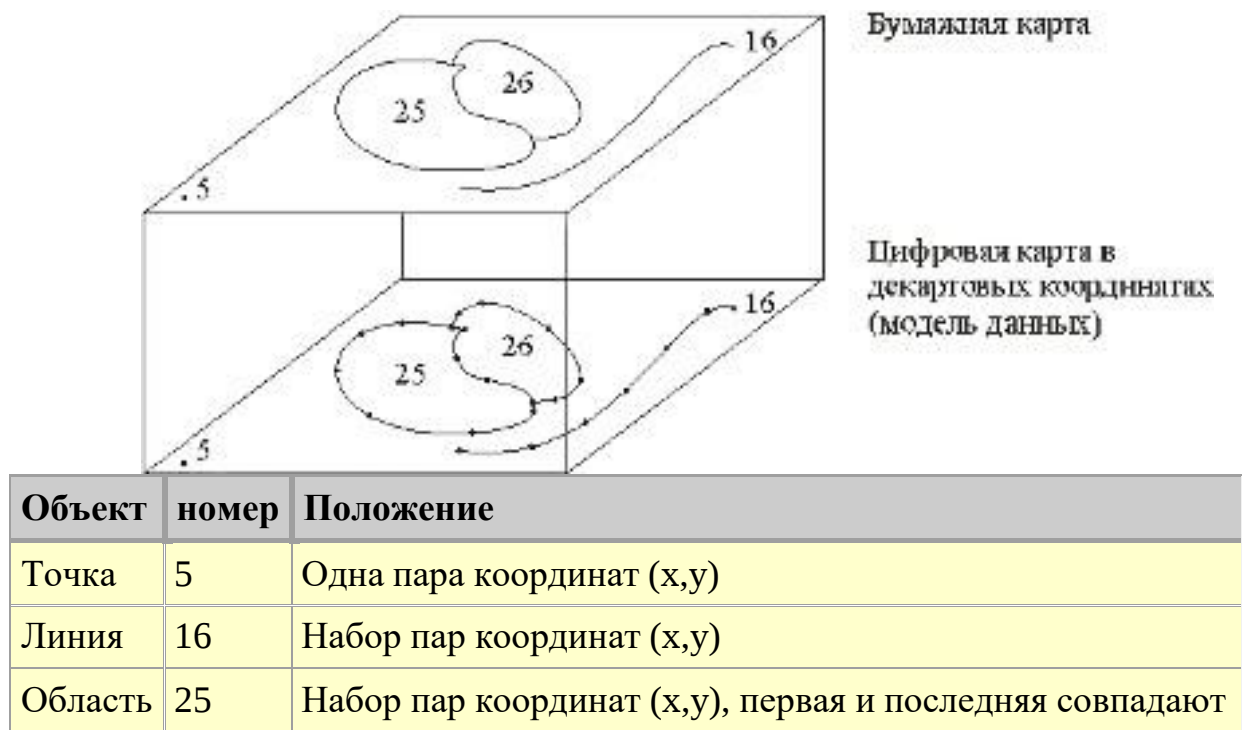


Рисунок 6 – «Спагетти»-модель

В этой модели не содержится описания отношений между объектами, каждый геометрический объект хранится отдельно и не связан с другими, например общая граница объектов 25 и 26 записывается дважды, хотя с помощью одинакового набора координат. Все отношения между объектами должны вычисляться независимо, что затрудняет анализ данных и увеличивает объем хранимой информации.

Векторные *топологические модели* (рисунок 7) содержат сведения о соседстве, близости объектов и другие, характеристики взаимного расположения векторных объектов.

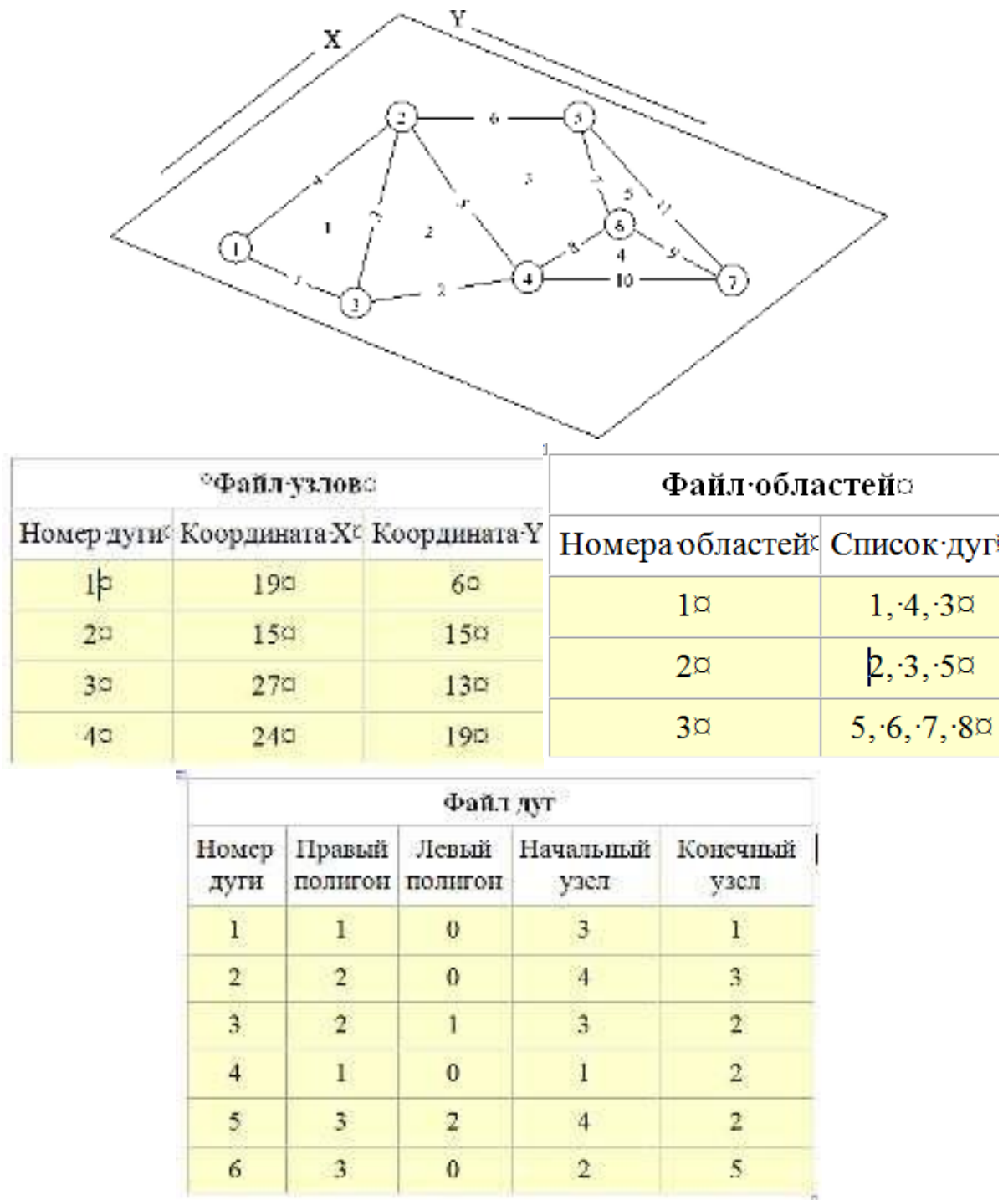


Рисунок 7. Векторная топологическая модель данных.

Топологическая информация описывается набором узлов и дуг. Узел – это пересечение двух или более дуг, и его номер используется для ссылки на любую дугу, которой он принадлежит. Каждая дуга начинается и заканчивается либо в точке пересечения с другой дугой, либо в узле, не принадлежащем другим дугам. Дуги образуются последовательностью отрезков, соединённых промежуточными точками. В этом случае каждая линия имеет два набора чисел: пары координат промежуточных точек и номера узлов. Кроме того, каждая дуга имеет свой идентификационный номер, который используется для указания того, какие узлы представляют её начало и конец.

Разработаны и другие модификации векторных моделей, в частности, существуют специальные векторные модели для представления моделей поверхностей, которые будут рассмотрены далее.

Растровые модели используются в двух случаях. В первом случае – для хранения исходных изображений местности. Во втором случае, для хранения тематических слоев, когда пользователей интересуют не отдельные пространственные объекты, а набор точек пространства, имеющих различные характеристики (высотные отметки или глубины, влажность почв и т.д.), для оперативного анализа или визуализации.

В растровых моделях данных в качестве атомарной модели используется двухмерный элемент пространства – ячейка, или пиксель. Упорядоченная совокупность атомарных моделей образует растр, который, в свою очередь, является моделью карты или геообъекта.

Для растровых моделей существует ряд характеристик.

1 Разрешение – минимальный линейный размер наименьшего участка пространства (поверхности), отображаемый одним пикселем. Более высоким разрешением обладает растр с меньшим размером ячеек.

2 Значение – элемент информации, хранящийся в элементе растра (пикселе). Тип значения может быть целым, действительным, комплексным, символьным.

3 Зона – соседствующие друг с другом ячейки, имеющие одинаковые значения.

4 Положение – упорядоченная пара координат (номер строки и номер столбца), которые однозначно определяют положение каждого элемента отображаемого пространства в растре.

Достоинствами растровых моделей является следующее:

- растр не требует предварительного знакомства с явлениями, данные собираются с равномерно расположенной сети точек, что позволяет в дальнейшем на основе статистических методов обработки получать объективные характеристики исследуемых объектов;

- растровые данные проще обрабатывать по параллельным алгоритмам;

- некоторые задачи, например создание буферной зоны, проще решать в растровом виде.

Наиболее часто растровые модели применяют при обработке данных дистанционного зондирования.

Недостатки растровых моделей:

- наряду с полезной информацией может попадать и избыточная (в том числе и бесполезная) информация;
- большой объём данных.

Для уменьшения объёма данных применяют алгоритмы сжатия. Наибольшую степень сжатия дают алгоритмы сжатия с потерями, но эти алгоритмы можно использовать только при визуальном анализе данных.

Растровые модели делятся на:

- регулярные;
- нерегулярные;
- вложенные (рекурсивные или иерархические) мозаики.

Существует несколько способов хранения и адресации значений отдельных ячеек растра, и их атрибутов, названий слоев и легенд.

При использовании растровых моделей актуальным является вопрос сжатия растровых данных, для которого разработаны методы группового кодирования, блочного кодирования, цепочного кодирования и представления в виде квадродерева.

5.4 Форматы данных

Форматы данных определяют способ хранения информации на жестком диске, а также механизм ее обработки. Модели данных и форматы данных определенным способом взаимосвязаны.

Существует большое количество форматов данных. Можно отметить, что во многих ГИС поддерживаются основные форматы хранения растровых данных (TIFF, JPEG, GIF, BMP, WMF, PCX), а также GeoSpot, GeoTIFF, позволяющие передавать информацию о привязке растрового изображения к реальным географическим координатам, и MrSID - для сжатия информации. Наиболее распространенным среди векторных форматов является - DXF.

Все системы поддерживают обмен пространственной информацией (экспорт и импорт) со многими ГИС и САПР через основные обменные форматы: SHP, E00, GEN (ESRI), VEC (IDRISI), MIF (MapInfo Corp.), DWG, DXF (Autodesk), WMF (Microsoft), DGN (Bentley). Только некоторые, в основном отечественные системы, поддерживают российские обменные форматы – F1M (Роскартография), SXF (Военно-топографическая служба).

Довольно часто для эффективной реализации одних компьютерных операций предпочитают векторный формат, а для других растровый. Поэтому, в некоторых системах реализуются возможности манипулирования данными в том и в другом формате, и функции преобразования векторного в растровый, и наоборот, растрового в векторный форматы.

Вопросы для повторения и самоконтроля

1. Назовите основные характеристики объектов реального мира.

2. Что такое пространственные данные?
3. Что такое атрибутивные данные?
4. Назовите основные структуры данных в ГИС.
5. Какие существуют модели данных в ГИС?
6. Какие существуют форматы данных в ГИС?