

Лекция № 3
СТРУКТУРЫ И МОДЕЛИ ДАННЫХ

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 замкнутый 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.

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 (рис.1).

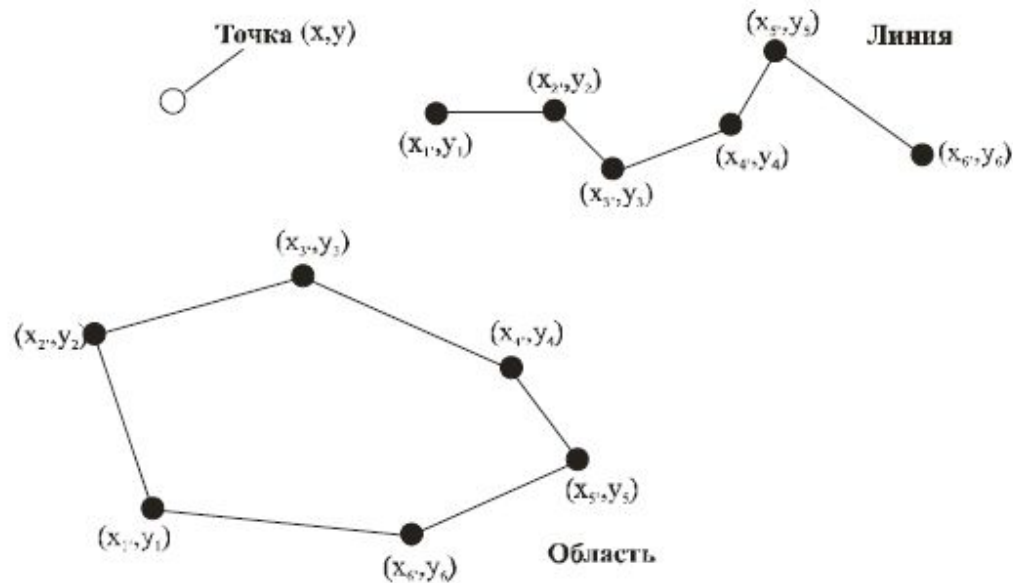
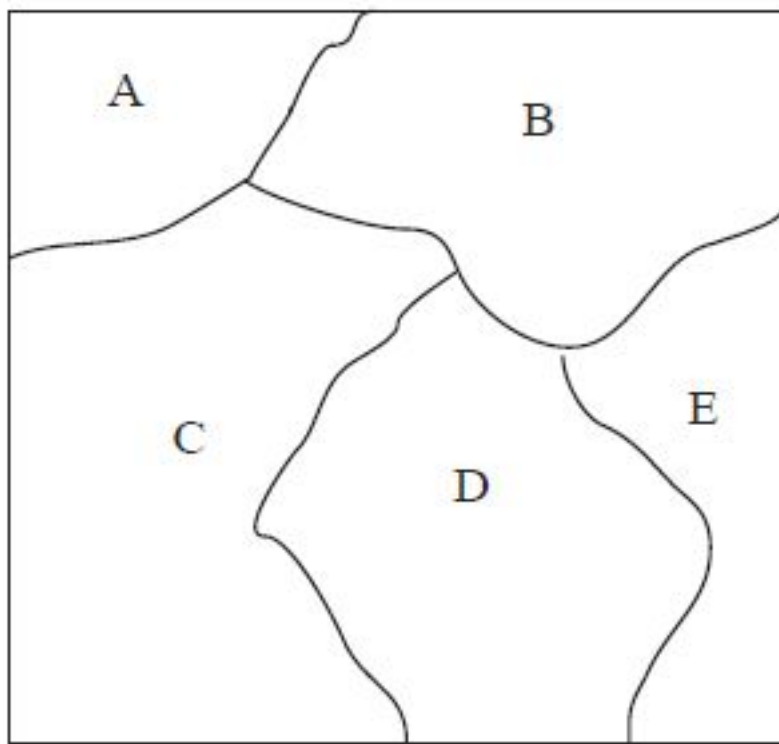


Рис. 1. Векторное представление пространственных данных



Растровая структура данных предполагает представления данных в виде двумерной сетки, каждая ячейка которой содержит только одно значение, характеризующее объект, соответствующий ячейке растра на местности или на изображении.

В качестве такой характеристики может быть код объекта (лес, луг и т.д.) высота или оптическая плотность. Точность растровых данных ограничивается размером ячейки. Такие структуры являются удобным средством анализа и визуализации разного рода информации.



A	A	A	A	B	B	B	B	B	B
A	A	A	A	B	B	B	B	B	B
A	A	A	B	B	B	B	B	B	B
C	C	C	C	C	C	B	B	B	E
C	C	C	C	C	D	D	E	E	E
C	C	C	C	D	D	D	D	E	E
C	C	C	D	D	D	D	D	D	E
C	C	C	C	D	D	D	D	D	E
C	C	C	C	C	D	D	D	E	E

Рис. 2. Растровая структура данных

Для реализации растровых и векторных структур разработаны различные модели данных.

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, Geo-TIFF, 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, Eoo, GEN (ESRI), VEC (IDRISI), MIF (MapInfo Corp.), DWG, DXF (Autodesk), WMF (Microsoft), DGN (Bentley). Только некото-рые, в основном отечественные системы, поддерживают российские об-менные 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.

Базы данных

Совокупность цифровых данных о пространственных объектах образует множество пространственных данных и составляет содержание баз данных.

База данных (БД) – совокупность данных организованных по определенным правилам, устанавливающим общие принципы описания, хранения и манипулирования данными.

Создание БД и обращение к ней (по запросам) осуществляется с помощью системы управления базами данных (СУБД).

Логическая структура элементов базы данных определяется выбранной моделью БД. Наиболее распространенными моделями БД являются иерархические, сетевые и реляционные и объектно-ориентированные.



Иерархические модели представляют древовидную структуру, в этом случае каждая запись связана только с одной записью, находящейся на более высоком уровне.

Сетевые модели были призваны устранить некоторые из недостатков иерархических моделей. В сетевой модели каждая запись в каждом узле сети может быть связана с несколькими другими узлами. Записи, входящие в состав сетевой структуры, содержат в себе указатели, определяющие местоположение других записей, связанных с ними. Такая модель позволяет ускорить доступ к данным, но изменение структуры базы требует значительных усилий и времени

Реляционные модели собирают данные в унифицированные таблицы. Таблице присваивается уникальное имя внутри БД. Каждый столбец – это поле, имеющее имя, соответствующее содержащемуся в нем атрибуту. Каждая строка в таблице соответствует записи в файле. Одно и тоже поле может присутствовать в нескольких таблицах. Так как строки в таблице не упорядочены, то определяется один или несколько столбцов, значения которых однозначно идентифицируют каждую строку. Такой столбец называется первичным ключом. Взаимосвязь таблиц поддерживается внешними ключами. Манипулирование данными осуществляется при помощи операций, порождающих таблицы. Пользователь может легко вносить в базу новые данные, комбинировать таблицы, выбирая отдельные поля и записи, и формировать новые таблицы для отображения на экране.



Объектно-ориентированные модели применяют, если геометрия определенного объекта способна охватывать несколько слоев, атрибуты таких объектов могут наследоваться, для их обработки применяют специфические методы.

Для обработки данных, размещенных в таблицах необходимы дополнительные сведения о данных, их называют метаданными.

Метаданные – данные о данных: каталоги, справочники, реестры и иные формы описания наборов цифровых данных.