

ОСНОВЫ ГЕОИНФОРМАТИКИ И ГИС-ТЕХНОЛОГИЙ

Становление и этапы развития геоинформатики

Сегодня проблема получения, хранения, обработки и использования информации о территориях выделилась в отдельную научно-технологическую дисциплину – геоинформатику. На этой науке базируются и совершенствуются картографические и новые геоинформационные методы исследований.

Сегодня геоинформатика предстает в виде системы, охватывающей науку, технику и производство. Учитывая особенности геоинформатики с точки зрения этих трех систем, трактовка геоинформатики и самих геоинформационных систем сводится к следующим дефинициям¹.

1) *Научно-познавательный подход.* Геоинформатика – научная дисциплина, изучающая природные и социально-экономические системы (их структуру, связи, динамику, функционирование в пространстве и во времени) посредством компьютерного моделирования на основе баз данных и географических знаний. Основная цель геоинформатики как науки — это управление подобными системами в широком понимании, включая их инвентаризацию, оценку, прогнозирование, оптимизацию. ГИС – средство моделирования и познания таких систем.

2) *Технологический подход.* Геоинформатика – это технология сбора, хранения, преобразования, отображения и распространения пространственно-координированной информации, имеющая цель обеспечить решение задач инвентаризации, оптимизации, управления геосистемами. ГИС - техническое средство накопления и анализа информации в процессе принятия решений.

3) *Производственный подход.* Геоинформатика – производство (геоинформационная индустрия), имеющее цель изготовления аппаратных средств и программных продуктов, включая создание баз и банков данных, систем управления, стандартных (коммерческих) ГИС разного целевого назначения и проблемной ориентации, формирование ГИС-инфраструктуры и организация маркетинга. ГИС – программная оболочка, реализующая геоинформационные технологии.

Основным назначением ГИС считается формирование знаний о процессах и явлениях на земной поверхности и применение этих знаний для решения практических задач во всех сферах человеческой деятельности.

Геоинформатика как область деятельности появилась во второй половине XX века в связи с развитием электронно-вычислительной техники и появлением первых геоинформационных систем.

В истории геоинформационных систем выделяются четыре периода:

1. Пионерный период (конец 1950-х – начало 1970-х гг.):

- исследование принципиальных возможностей, пограничных областей знаний и технологий, наработка эмпирического опыта, первые крупные проекты и теоретические работы.

2. Период государственных инициатив (начало 1970-х – начало 1980-х гг.):

- развитие крупных геоинформационных проектов, поддерживаемых государством, формирование государственных институтов в области ГИС, снижение роли и влияния отдельных исследователей и небольших групп.

3. Период коммерческого развития (ранние 1980-е – настоящее время):

- широкий рынок разнообразных программных средств, развитие настольных ГИС, расширение области их применения за счет интеграции с базами непространственных данных, появление сетевых приложений, появление значительного числа непрофессиональных пользователей, системы, поддерживающие индивидуальные наборы данных на отдельных компьютерах, открывают путь системам, поддерживающим корпоративные и распределенные базы геоданных.

4. Пользовательский период (с конца 1980-х по настоящее время):

- повышенная конкуренция среди коммерческих производителей геоинформационных технологий услуг дает преимущества пользователям ГИС, доступность и «открытость» программных средств позволяет пользователям самим адаптировать, использовать и даже модифицировать программы, появление пользовательских клубов, телеконференций, территориально разобщенных, но связанных единой тематикой пользовательских групп, возросшая потребность в геоданных, начало формирования мировой геоинформационной инфраструктуры.

Обзор базовых концепций геоинформатики, ее задач и основных понятий

По С.Н. Сербенюку¹, **геоинформатика** – область деятельности в географии, геологии и др. науках о Земле, в рамках которой решаются задачи сбора, хранения и обработки информации (геоданных) о природных и социально-экономических системах. Это понятие, обозначающее автоматическую переработку пространственно-временной информации о геосистемах различного иерархического уровня и территориального охвата.

Берлянт А.М. увязал задачи геоинформатики с моделированием геосистем. По его мнению, **геоинформатика** – научная дисциплина, изучающая природные и социально-экономические геосистемы (их структуру, связи, динамику, функционирование в пространстве-времени) посредством компьютерного моделирования на основе баз данных и географических знаний.

Берлянт А.М., кроме того, отмечает триединство геоинформатики как науки, техники и производства. С его точки зрения, **геоинформатика** – это наука, технология и производственная деятельность:

- по научному обоснованию, проектированию, созданию, эксплуатации и использованию географических информационных систем;
- по разработке геоинформационных технологий;
- по прикладным аспектам или приложениям ГИС для практических или геонаучных целей.

Предмет геоинформатики – пространственно-временные информационные потоки естественно-географической среды.

Метод геоинформатики – пространственно-временное моделирование территориально-распределенных (географических) эмпирических (объективно существующих) систем любой природы с использованием соответствующих ГИС-технологий в различных научных и практических целях. Как наука геоинформатика рассматривает управление геосистемами, включая их инвентаризацию, оценку, прогнозирование, оптимизацию и т.п. Как производство и технология геоинформатика (геоинформационная индустрия) рассматривает процессы изготовления аппаратуры, создания коммерческих программных продуктов и ГИС-оболочек, баз данных, систем управления, компьютерных систем.

Геоинформатика тесно связана с картографией (рис. 2.1.). Их взаи-

мосьвязь проявляется в следующих аспектах:

1) тематические и топографические карты – главный источник пространственно-временной информации;

2) системы географических и прямоугольных координат и картографическая разграфка служат основой для координатной привязки всей информации, поступающей и хранящейся в ГИС;

3) карты – основное средство географической интерпретации и организации данных дистанционного зондирования и другой используемой в ГИС информации;

4) картографический анализ – один из наиболее эффективных способов выявления географических закономерностей, связей, зависимостей при формировании баз знаний, входящих в ГИС;

5) математико-картографическое и ЭВМ-картографическое моделирование – главное средство преобразования информации в процессе обеспечения принятия решений, управления, проведения экспертиз, составления прогнозов развития геосистем и т.п.;

6) картографическое изображение – целесообразная форма представления информации потребителям, а автоматическое изготовление оперативных и базовых карт, трехмерных картографических моделей, дисплей-фильмов – одна из главных функций ГИС.

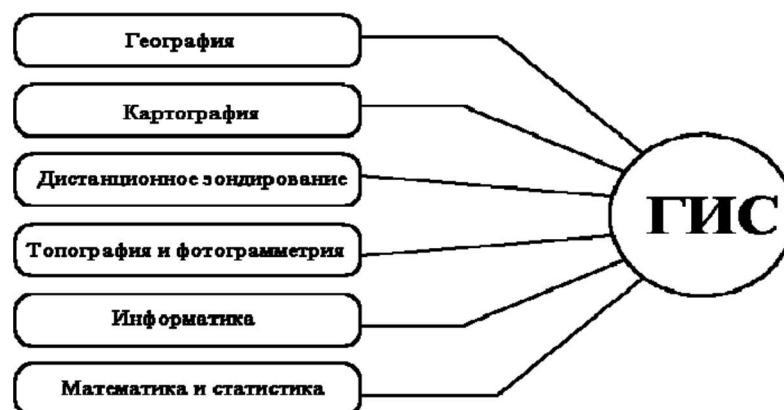


Рис. 2.1. Связь ГИС с научными дисциплинами и технологиями

Геоинформатика изучает и разрабатывает принципы, методы и технологии сбора, накопления, передачи, обработки и представления данных для получения на их основе новой информации и знаний о пространственно-временных явлениях в геосистемах. Тесно взаимосвязанные понятия: **данные, информация и знания** – имеют основополагающее значение для

геоинформатики.

Данные описывают явления реального мира или идей, которые представляются достаточно ценными для того, чтобы их сформулировать и точно зафиксировать.

Совокупность сведений о фактических данных, представленных в формализованном виде (цифровом, графическом и др.), пригодны для хранения, пересылки, интерпретации человеком и обработки компьютером, рассматриваются как объект обработки и получения **информации**, в которую заложена совокупность знаний о фактических данных и зависимостях между ними.

Сами же **знания** представляют собой отражение семантических аспектов реального мира человеком или технической системы (система искусственного интеллекта), интерпретацию информации об окружающих объектах и явлениях.

Фундаментальными понятиями геоинформатики являются **пространственные данные** и **пространственный объект**, с которыми неразрывно связано понятие «модель».

Геопространственные данные (геоданные) – это понятие означает информацию, которая идентифицирует географическое местоположение и свойства естественных или искусственно созданных объектов, а также их границ на земле. Эта информация может быть получена с помощью (помимо иных путей) дистанционного зондирования, картографирования и различных видов съемок.

Географические данные содержат четыре интегрированных компонента: местоположение, свойства и характеристики, пространственные отношения, время.

«Данные», «информация», «знания» в геоинформационных системах. Конкретизируя термины "данные", "информация", "знания" применительно к оперированию ими в информационной системе, можно отметить, что, имея много общего, эти понятия различаются по своей сути.

Под данными понимается совокупность фактов, известных об объектах, либо результаты измерения этих объектов. Данные, используемые в ГИС, отличаются высокой степенью формализации. Данные по Берлянту А.М - это как бы строительный элемент в процессе создания информации, поскольку она получается в процессе обработки данных.

Применительно к ГИС под информацией понимается совокупность сведений, определяющих меру наших знаний об объекте.

В таком контексте знания можно рассматривать как результат интер-

претации информации. Наиболее общее определение: знание – результат познания действительности, получивший подтверждение в практике. Научное знание отличается своей систематичностью, обоснованностью и высокой степенью структуризации.

Информационные системы можно рассматривать как эффективный инструмент получения знаний.

Различия между терминами «данные», «информация» и «знания» прослеживаются в истории развития технических систем, так вначале появились банки данных, позднее – информационные системы, затем появились системы, основанные на знаниях Берлянт А.М.¹ – интеллектуальные системы (экспертные системы).

В настоящее время на рынке программных продуктов представлено несколько видов систем, работающих с пространственно распределенной информацией, к ним, в частности, относятся системы автоматизированного проектирования, автоматизированного картографирования и ГИС. ГИС по сравнению с другими автоматизированными системами обладают развитыми средствами анализа пространственных данных, открывают принципиально новые аналитические возможности.

Типовые вопросы, на которые способна ответить ГИС:

- Где находится А?
- Как расположено А по отношению к В?
- Сколько А расположено в пределах расстояния D от В?
- Каково значение функции Z в точке X?
- Как велико по размерам В?
- Каков результат пересечения А и В?
- Каков оптимальный маршрут от X до Y?
- Что находится в X1, X2, ..., Xn?
- Какие объекты следуют за теми, у которых наблюдается определенное сочетание определенных свойств?
 - Как изменится пространственное распределение объектов, если изменить существующую классификацию?

Что может случиться с А, если изменить В и его расположение относительно А?

Концептуальная схема организации данных в ГИС представлена на рис 2.2.

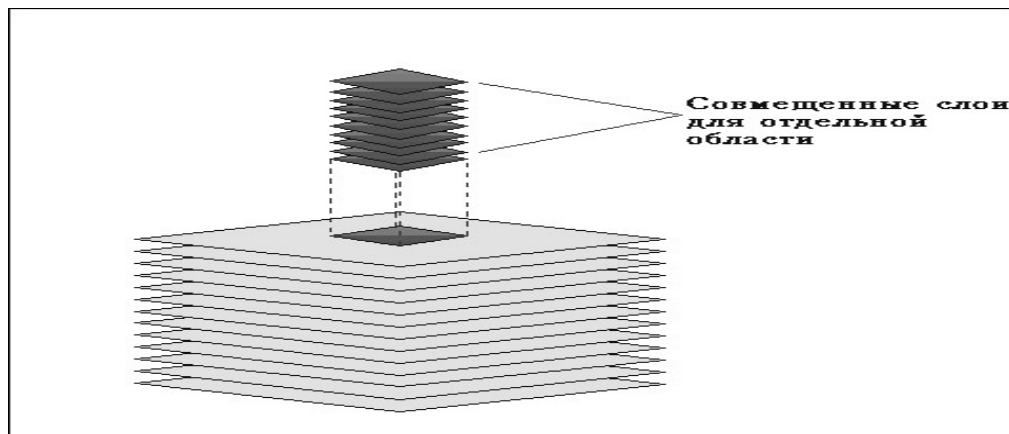


Рис.2.2. Концептуальная схема организации данных в ГИС

Пространственная выборка (уточнение территории) (рис.2.3.)

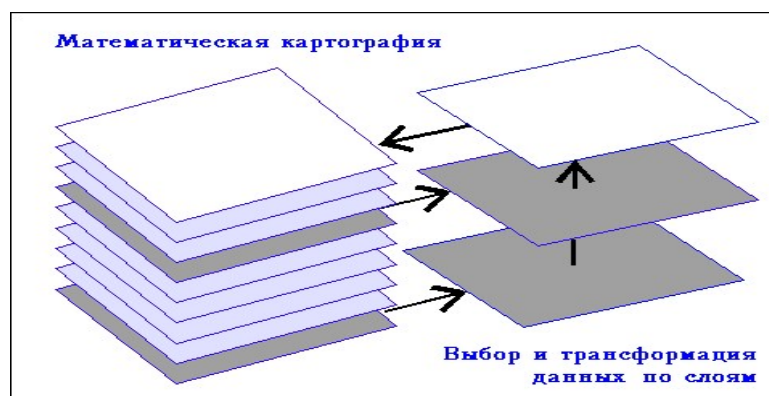


Рис. 2.3. Пространственная выборка (уточнение территории)

Тематическая проблемно-ориентированная выборка представлена на рис 2.4.



Рис. 2.4. Тематическая проблемно-ориентированная выборка

Существующие области использования ГИС показаны на рис. 2.5.



Рис. 2.5. Существующие области использования ГИС

Базовые структуры данных в ГИС

Природа географических данных:

- географическое положение (размещение) пространственных объектов представляется 2-х, 3-х или 2-х мерными координатами в географически соотнесенной системе координат (широта/долгота);
- свойства (атрибуты) являются описательной информацией определенных пространственных объектов. Они часто не имеют прямых указаний на пространственное размещение, поэтому нередко атрибуты называют непространственной информацией;
- пространственные отношения определяют внутренние взаимоотношения между пространственными объектами (например, направление объекта А в отношении объекта В, расстояние между объектами А и В, вложенность объекта А в объект В);
- временные характеристики представляются в виде сроков получения данных, они определяют их жизненный цикл, изменение местоположения или свойств пространственных объектов во времени.

Основополагающие элементы базы пространственных данных.

- Элементы действительности, смоделированные в базе данных ГИС, имеют два тождества: реальный объект и смоделированный объект (объект БД).
- Реальный объект – явление окружающего мира, представляющее интерес, которое не может быть более подразделено на явления того же самого типа.

- Объект БД – элемент в том виде, в каком он представлен в базе данных. Объект БД является «цифровым представлением целого или части реального объекта.

- Метод цифрового представления явления изменяется исходя из базового масштаба и ряда других факторов.

Модель базы пространственных данных.

- Каждый тип реального объекта представляется определенными пространственными объектами базы данных.

- Пространственные объекты могут быть сгруппированы в слои, также называемые оверлеями, покрытиями или темами.

- Один слой может представлять одиночный тип объекта или группу концептуально связанных типов.

Общие подходы к представлению пространственных объектов в БД.

1. Растровый способ: ячейки, сетки.

2. Векторный способ: точки, линии, полигоны.

Растровая модель данных:

- разбивает всю изучаемую территорию на элементы регулярной сетки или ячейки;

- каждая ячейка содержит только одно значение;

- является пространственно заполненной, поскольку каждое местоположение на изучаемой территории соответствует ячейке растра, иными словами – растровая модель оперирует элементарными местоположениями.

Соглашения, принятые для растровой ГИС.

1. *Разрешение.* Минимальная линейная размерность наименьшей единицы географического пространства, для которой могут быть приведены какие-либо данные. В растровой модели данных наименьшей единицей для большинства систем выступает квадрат или прямоугольник. Такие единицы известны как сетка, ячейка или пиксель. Множество ячеек образует решетку, растр, матрицу.

2. *Площадной контур (Зона).* Набор смежных местоположений одинакового свойства. Термин «класс» (или район) часто используется в отношении всех самостоятельных зон, которые имеют одинаковые свойства. Основными компонентами зоны являются ее значение и местоположение.

3. *Значение.* Единица информации, хранящаяся в слое для каждого пикселя или ячейки. Ячейки одной зоны (или района) имеют одинаковое значение.

4. *Местоположение.* Наименьшая единица географического пространства, для которого могут быть приведены какие-либо характеристики или свойства (пиксель, ячейка). Такая частица картографического плана однозначно идентифицируется упорядоченной парой координат – номерами строки и столбца.

Векторная модель данных:

- основана на векторах (направленных отрезках прямых);
- базовым примитивом является точка;
- объекты создаются путем соединения точек прямыми линиями или дугами;
- площади определяются набором линий;
- представляет собой объектно-ориентированную систему.

Векторная структура – это представление пространственных объектов в виде набора координатных пар (векторов), описывающих геометрию объектов. Пример векторного представления пространственных объектов (рис.2.6.).

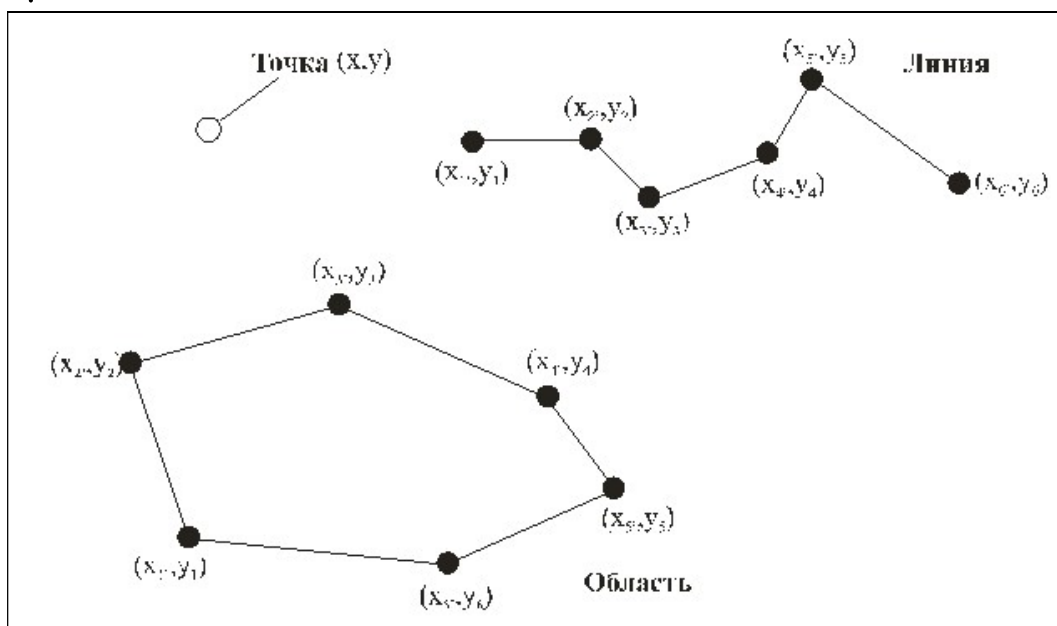


Рис.2.6. Векторное представление пространственных объектов

Типы векторных объектов, основанные на определении пространственных размеров

Безразмерные типы объектов:

- точка – определяет геометрическое местоположение;
- узел – топологический переход или конечная точка, также может определять местоположение.

Одномерные типы объектов:

- линия – одномерный объект;
- линейный сегмент – прямая линия между двумя точками;
- строка – последовательность линейных сегментов;
- дуга – геометрическое место точек, которые формируют определенную математическую функцию;
- связь – соединение между двумя узлами;
- направленная связь – связь с одним определенным направлением;
- цепочка – направленная последовательность непересекающихся линейных сегментов или дуг с узлами на их концах;
- кольцо – последовательность непересекающихся цепочек, строк, связей или замкнутых дуг.

Двумерные типы объектов:

- область – ограниченный непрерывный объект, который может включать или не включать в себя собственную границу;
- внутренняя область – которая не включает собственную границу;
- полигон – область, состоящая из внутренней области, одного внешнего кольца и нескольких непересекающихся, не вложенных внутренних колец;
- пиксель – элемент изображения, который является самым малым;

Пример слоев, составленных из пространственных объектов линейного типа (рис.2.7).

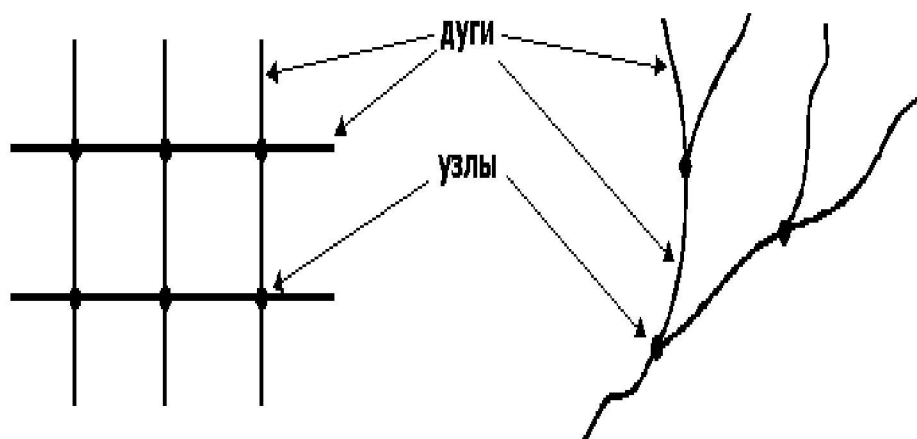


Рис.2.7. Пространственные объекты линейного типа

Примеры слоев, составленных из пространственных объектов полигонального типа (рис.2.8).



Рис. 2.8. Пространственные объекты полигонального типа

Формы векторной модели данных

- цельнополигональная структура (структура типа «спагетти»);
- линейно-узловая структура (топологическая структура);
- реляционная структура;
- нерегулярная триангуляционная сеть (TIN);

Топологическое представление векторных объектов см. на рис.2.9.

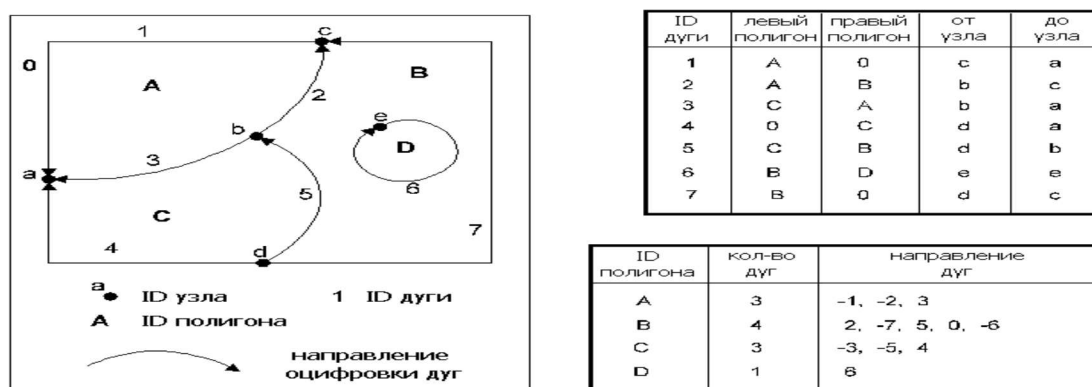


Рис. 2.9. Топологическое представление векторных объектов

Формирование топологии включает определение и кодирование взаимосвязей между точечными, линейными и площадными объектами. Сопоставление растровой и векторной моделей данных представлено на

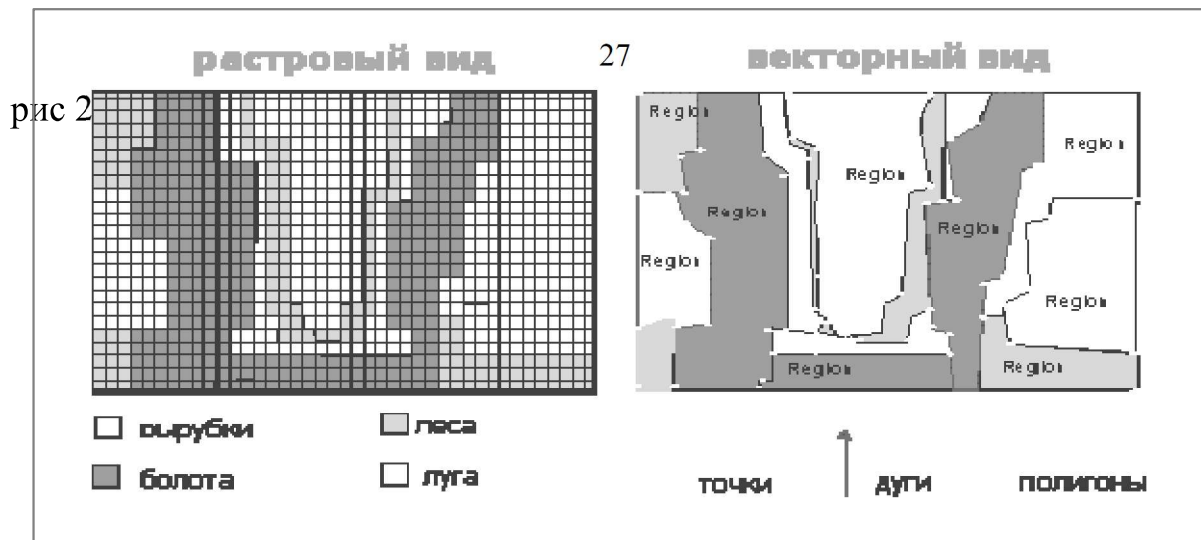


Рис. 2.10. Сопоставление растровой и векторной моделей данных

Преимущества растровой и векторной моделей данных

Растровая модель:

- 1) простая структура данных;
- 2) эффективные оверлейные операции;
- 3) работа со сложными структурами;
- 4) работа со снимками.

Векторная модель:

- 1) компактная структура;
- 2) топология;
- 3) качественная графика.

Представление пространственных объектов в ГИС

Представление пространственных объектов реальной действительности основано на следующих допущениях.

- Пространственные данные состоят из цифровых представлений реально существующих дискретных пространственных объектов.
- Свойства, показанные на карте, например, озера, здания, контуры должны пониматься как дискретные объекты.
- Содержание карты может быть зафиксировано в базе данных путем превращения свойств карты в пространственные объекты.
- Многие свойства, которые показаны на карте, на самом деле виртуальны. Например, контуры или границы реально не существуют, но здания и озера – реальные объекты.

Содержание базы пространственных данных включает:

- 1) цифровые версии реально существующих объектов (например, зданий);

2) цифровые версии искусственно выделенных свойств карты (например, контуры);

3) искусственные объекты, созданные специально для целей построения базы данных (например, пиксели).

Разновидность непрерывных свойств:

1) некоторые свойства пространственных объектов существуют повсеместно, изменяются непрерывно над земной поверхностью (высота, температура, атмосферное давление) и не имеют реально представленных границ.

Непрерывная изменчивость может быть представлена в базе данных несколькими способами:

- посредством величин измерений в некоторых характерных пунктах (точках), например, метеостанции и посты;

- посредством описаний трансектов (профилей);

- посредством разделения площади на контуры, зоны, принимая при этом, что некоторое значение свойства внутри контура (зоны) есть величина постоянная;

- посредством построения изолиний, например, горизонталей для отображения рельефа.

Каждый из этих способов создает дискретные объекты, которые могут быть зафиксированы в виде точек, линий, площадей.

Компоненты пространственных данных:

- расположение: пространственные данные вообще часто называются данными о размещении;

- пространственные отношения: взаимосвязи между пространственными объектами описываются как пространственные отношения между ними (например, А содержит В; смежен с С, находится к северу от D);

- атрибуты: атрибуты фиксируют тематические описания, определяя различные характеристики объектов;

- время: временная изменчивость фиксируется разными способами:

1) интервалом времени, в течение которого существует объект;

2) скоростью изменчивости объектов;

3) временем получения значений свойств.

Источники пространственных данных:

совокупность первичных данных (измерений и съемок) по выборкам:

- произвольная выборка, каждое место или время одинаково вероятно, чтобы быть выбранным;

- систематическая выборка, проводится согласно правилу, например, через каждый 1 км;

- упорядоченная (стратифицированная) выборка, когда известно, что генеральная совокупность содержит существенно различные подсовкупности.

Совокупности вторичных данных, полученные из существующих карт, таблиц, или других баз данных.