

Модели данных: формальные абстракции действительности (лекция 2)

Когда исследователь изучает объект, они упрощают свойственную ей сложность, фиксируя главные особенности, чтобы создать модель объекта. Мы рассмотрим различные модели, которые используются в процессе создания географических данных и которые могут использоваться в графической или цифровой форме.

Основные модели основаны на точках, линиях, многоугольниках или мозаичных модулях, описывающих различные особенности в ландшафте. Принятие специфической модели влияет на тип данных, которые могут использоваться, чтобы описать явления и пространственный анализ, который может быть предпринят.

Вообразите, что Вы говорите по телефону с кем-то, и вас просят описать вид из вашего окна. Как Вы бы это сделали? Вероятно, что Вы разбили бы ландшафт на детали типа: здания, дороги, поля, точки высот и использовали бы географическую ссылки типа «налево от» или «в ста метрах». Вы фактически разработали концептуальную модель ландшафта. Ваша интерпретация особенностей, которые Вы наблюдали, будет определяться вашим опытом и уровнем подготовки человека которому Вы описываете сцену.

Когда надо обмениваться информацией по большому региону, становится необходимым формализовать модели, чтобы гарантировать эффективную интерпретацию без двусмысленности (однозначно).

Концептуальные модели реальных мировых географических явлений

Географические объекты нуждаются в двух описаниях: что присутствует, и где это. Для первой, феноменологической категории типа «города», «реки», «экотоп», «ассоциации почв» используются фундаментальные стандартные описательные блоки. Эти объекты и явления известны и описаны в стандартных терминах. Однако термины и описания будут различны для различных уровней разрешающей способности.

Дорога, отображенная на космоснимке, может быть линией, но на плане строительного участка это должна быть полигон, чтобы показать ее различные структуры. Явления также очень часто сгруппируются или разделяются в модули при разных масштабах, согласно иерархически определенных таксономий; например иерархия модулей администрирования страны - область - городской район и т.д.

Позиционирование в пространстве объектов может быть геометрически точным или относительным местоположением. Изначально используются системы местных или мировых координат, базирующихся на стандартных системах сфероидов, проекций и координат, которые проецируют участки земли на плоскую поверхность. Система координат

может быть локальной, измеренной в десятках метров, или это может быть единая градусная сетка или международная принятая проекция, которая использует геометрические координаты широты и долготы. Некоторые карты обеспечивают относительное географическое позиционирование (адресное), как, например схема метро.

Концептуальные модели пространства: объекты или поля

Из концептуальных идей географических объектов возможно формализовать представление пространства и пространственных свойств. При рассмотрении любого пространства комнаты, ландшафта, или континента - мы можем применять несколько различных способов для описания того, что происходит на поверхности земли. Есть два варианта:

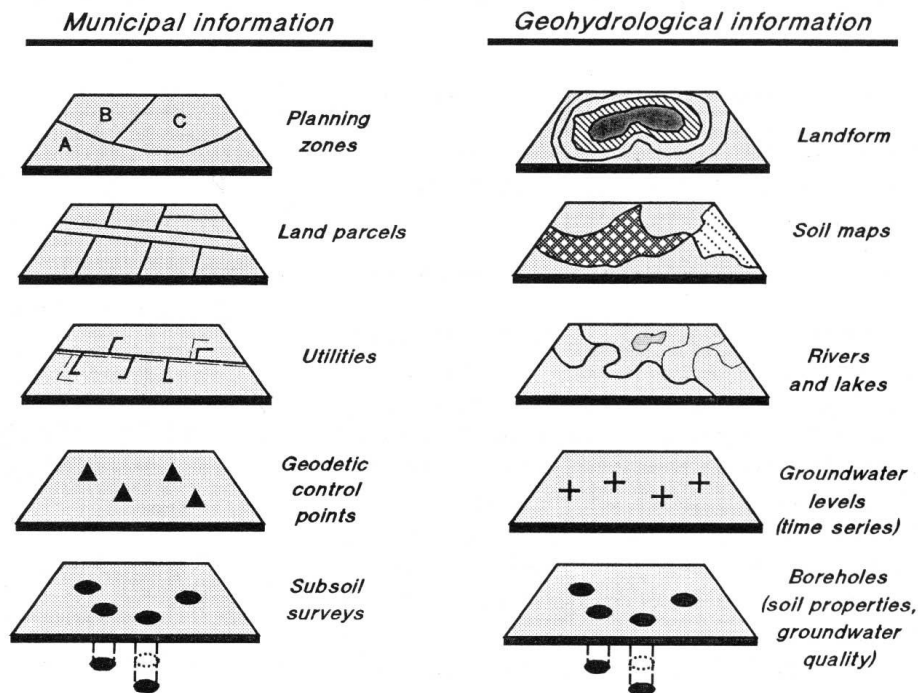
1. Представить пространство занятое дискретными объектами, которые описаны их атрибутами или свойствами, и чья позиция может быть определена, используя геометрическую систему координат.
2. Представить что изменение свойств объекта, представляет непрерывную математическую функцию или поле.

Объекты. Наиболее общее мнение, что пространство населено объектами. Определение и распознавание объекта (будь то дом, кабель, лес, река, гора) - это первый шаг; затем перечисляются их атрибуты, определяются границы и местоположение. В этой книге мы используем термин объект для тех вещей, что большинство людей назвало бы объект (физический), потому что термин "объектная ориентация" приобрел очень распространенное применение в технологии баз данных и программировании.

Непрерывные поля. В категории непрерывного поля, самая простая концептуальная модель представляет географическое пространство в непрерывных Декартовых координатах в двух или трех измерениях (или четыре, если время включено). Атрибут обычно изменяется плавно и непрерывно по пространству. Сначала рассматривается атрибут (например, воздушное давление, температура, содержание глины в почве) и его пространственное изменение; лишь когда присутствуют заметные проявления атрибута в пространстве или времени, он признается как объект (например, ураган Эндрю, течение Гольфстрим и др.

Выбор объектной или непрерывной модели может быть затруднен, когда объекты могут также рассматриваться как наборы экстремальных значений атрибутов, сгруппированных в географическом пространстве. Швейцарию, например, можно рассматривать как область индивидуальных объектов (гор) или как область, на которой атрибут "высота" демонстрирует экстремальные значения. Практическое решение, основывается на целях пользователя ГИС. Выбор концептуальной модели определяет, как информация может позже быть получена. Выбор объектной модели

целесообразен для системы, которая делает запись, кто поднялся на гору и когда, но она не будет обеспечивать информацию для вычисления крутизны склонов. Непрерывная модель позволяет вычисление крутизны как первой производной поверхности, но не дает названия для тех частей поверхности, где производная равна 0.



Географические модели данных и данные географических примитивов

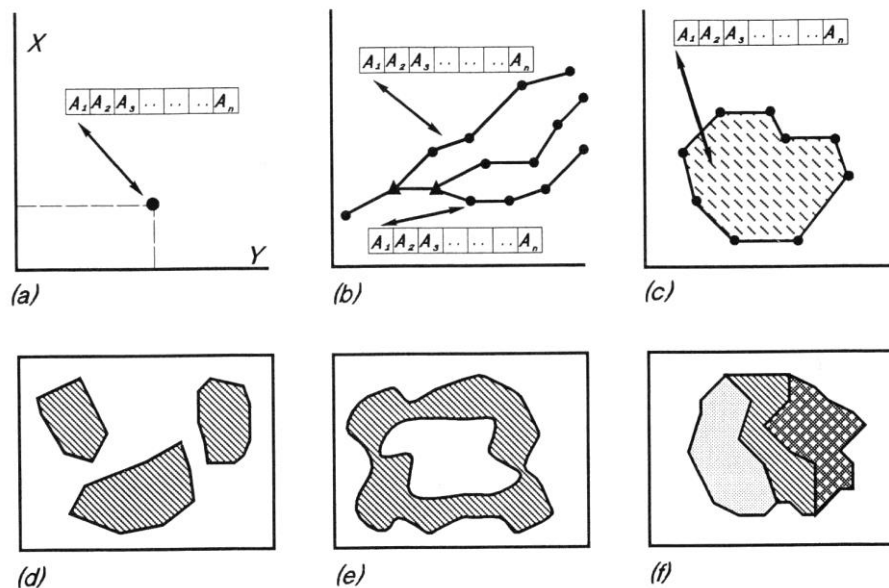
Географические модели данных - формализованные эквиваленты концептуальных моделей, используемых людьми, чтобы изучать географические объекты и явления. Они формализуют - как пространство разделяется на отдельные части для анализа и предполагают, что частные объекты могут быть уникально идентифицированы, что их атрибуты могут быть измерены и определены географические координаты. Поскольку данные могут быть собраны разными путями, то информация, касающаяся метода сбора также является важной частью модели данных.

Большинством антропогенных объектов (здания, земельные участки, административные районы, дороги, кабели, трубопроводы, сельскохозяйственные поля и др.) лучше всего оперировать, используя объектную модель. Самая простая и наиболее часто используемая модель данных действительности - пространственный объект, который далее определен атрибутами и географическим местоположением. Фундаментальные географические примитивы точек, линий, и многоугольников подразделяются по трем типам основных географических примитивов, а именно "точка", "линия" или "область" (полигон в ГИС). Это - фундаментальные модули векторной модели данных.

Непрерывные полевые данные в 2-х и 3-х мерном измерении слишком сложны, чтобы быть описанными простой математической функцией типа полиномиального уравнения. В некоторых ситуациях могут использоваться простые уравнения регрессии (трендовые поверхности), чтобы представить крупномасштабные вариации в виде мозаики (грид и растр).

Векторные модели данных объектов

Векторная модель данных представляет пространство как ряд дискретных объектов: точек, линий и полигонов, которые привязаны к Декартовым координатам.



Простые точки, линии и полигоны - по существу статичные представления объектов в XY координатах. Полагается, что они неизменны и не содержат никакую информацию о временной или пространственной изменчивости.

Объект «точка» подразумевает, что местоположением может быть определено одним набором координат на уровне разрешающей способности карты. Город может быть представлен точкой на континентальном уровне разрешающей способности, но как полигон на региональном уровне. Увеличение масштаба показывает внутреннюю структуру объекта (в случае города - кварталов, предместий, улиц, зданий, фонарных столбов и др.

Объекты «линии» подразумевает, что географическое протяжение линии может быть адекватно представлено набором пар координат, которые определяют конкретный путь в пространстве, но не имеющий никакой ширины, если она не определена атрибутом объекта. Дорога на мелкомасштабной карте представляется линией; в масштабе городской схемы она становится полигоном с определенным покрытием и в данном случае ее нецелесообразно отображать линией. Телефонный кабель может быть представлен линией при любых масштабах карты.

Самое простое определение полигона подразумевает, что это - однородный участок 2-мерного пространства. Это также зависит от масштаба представления. Полигон можно задать через координаты его границ, или координаты точки внутри границы (грид). Полигоны могут включать отверстия, они имеют прямых соседей, полигоны с одинаковыми параметрами могут находиться в различных местах карты.

Если граница может быть четко определена, то она описывается конечным числом связанных координатных пар, причем количество координатных пар определяется масштабом карты. Обычно уровень разрешающей способности задается числом десятичных знаков, определяемых масштабом.

Сложные объекты могут иметь внутреннюю структуру и функциональность. Город включает улицы, здания, и парки, они имеют различные функции. Топологические связи могут показывать - как линии связаны в полигоны или сети.

Мозаика непрерывных полей

Непрерывные поверхности могут быть дискретны в отдельных модулях, типа квадрата, треугольных, или шестиугольных ячеек, или в неправильных треугольниках, или многоугольниках которые являются элементами мозаики, формирующих географические представления. Использование неправильных треугольников, давно известно в геодезии, оно основано на принципе триангуляции, посредством чего непрерывная поверхность земли аппроксимирована сетью треугольников, чьи вершины или узлы тщательно измеряются. Главное преимущество этого подхода состоит в том, что плотность сети может быть легко откорректирована к степени, в зависимости от степени точности (размер треугольника). Треугольная поверхность может легко приспособлена для трехмерного представления поверхности. Эти модели данных - по существу статические представления гипсометрической поверхности, которая не изменяется во времени.

Треугольные сети также используются, для моделирования динамических полей (поверхность грунтовых вод, ветра и др.). Треугольную неправильную сеть или TIN можно рассматривать как упрощенную полигональную векторную схему.

Наиболее используемая альтернатива триангуляции - регулярная мозаика или грид. 2-мерная геометрическая поверхность разделена на квадратные ячейки (пиксели) чей размер определен разрешающей способностью (масштабом). Такое представление пространства называется растровая модель данных.

Когда ячейки сетки используются, чтобы представить изменение непрерывно варьирующего атрибута, то каждая ячейка будет иметь различное значение атрибута, а изменения между ячейками будут таковы,

чтобы отображать непрерывную поверхность. Это позволит использование дифференциальных уравнений для подсчета локальных площадей, степени изменений и др. Каждую ячейку сетки можно считать отдельным объектом, который отличается от полигона лишь правильной формой.

Несмотря на то, что регулярная сетка (грид) наиболее часто используется, чтобы представлять статичные явления, она легко может быть адаптирована для изучения динамических изменений. Изменения во времени могут фиксироваться в отдельных слоях - на каждый временной интервал - отдельный слой. Время, как и пространство дискретно в этой модели.

Изменения в пространстве может легко рассчитать по регулярной сетке поскольку она аппроксимирует непрерывную математическую поверхность, которую можно дифференцировать. Поток вещества в пространстве вычисляется методом конечных разностей, так как производные первого и второго порядка находятся простым суммированием и вычитанием.

Трехмерные эквиваленты пикселей называются элементами объема - это основные модули пространственных изменений в трехмерном пространстве. Все операции, возможные в 2-мерных регулярных сетках могут применяться к данным 3-мерных.

Пиксели и элементы объема как "объекты"

Основные единицы дискретизации в правильной мозаике непрерывного поля могут использоваться, чтобы обеспечить геометрическую привязку для простых наборов данных (точек, линий и полигонов). Векторная точка может быть представлена одной ячейкой; векторная линия набором смежных с одинаковым атрибутом; векторный полигон - набором ячеек с одинаковым атрибутом. Векторное представление предпочтительнее, потому что правильные ячейки сетки могут терять пространственные подробности, хотя это становится меньшей проблемой с увеличением мощности компьютеров и блоков памяти.

Это сходство между векторными и растровыми моделями пространства часто вызывает затруднение в понимании характера представляемых явлений. Для примера, непрерывные поля могут быть представлены изолиниями или контурами, которые являются наборами XY координат, связывающих точки с одинаковым значением атрибута. Контурные полезны для представления величин атрибутов на бумажных картах и компьютерных экранах, но они менее эффективны для обработки непрерывного пространственного изменения в численных моделях пространственных взаимодействий. Контурные могут быть обработаны как простые полигоны или как замкнутые линии в векторных моделях, но они лишь представления, а не границы реальных объектов.

Векторный и растровый подходы перекрываются, когда мы имеем дело с такими вещами как почвенное картографирование и использование земель. Классический подход состоит в том, чтобы определить классы почвы, использование земель и т.д., а затем идентифицировать участки земли, которые соответствуют этим классам. Эти области могут быть представлены векторными границами - замкнутыми полигонами или наборами растровых ячеек, имеющих равное значение.

Другой подход к представлению искусственных объектов типа землепользования или классов почв состоит в том, чтобы постулировать, что землепользование или почвы - непрерывные переменные, не объекты, но географическая поверхность, составленная из зон, где атрибуты имеют то же самое значение (полигоны) и зоны, где атрибуты резко изменяются (границы). Этот метод необходим, чтобы извлекать объекты и гомогенные зоны из растровых данных, типа космоснимков. *Обратите внимание, что хотя об изменении этих видов атрибутов можно думать как о непрерывных в пространстве (потому что каждая ячейка имеет значение атрибута почвы, включая классы "не почва" или "заброшенная земля", поверхность - не непрерывна в терминах дифференциального исчисления. Это означает, что математические операции подобно вычислению уклонов не должны применяться к данным, которые не могут быть аппроксимированы непрерывной математической функцией.*

В 2-мерных и 3-мерных моделях мы можем рассматривать пиксели (элементы объема) или полигоны как модули, которые могут быть обработаны как ряд открытых систем с правильной или неправильной формой. Состояние каждой ячейки (локальная система) определено значением ее атрибутов; значения атрибутов могут быть изменены операциями, которые относятся только к рассматриваемой ячейки или используют информацию соседних ячеек, которые являются частью окружения ячейки.

Таблица 2.1. Дискретные модели данных для пространственных объектов

Векторное представление точных объектов	Мозаики непрерывных полей
<p>Нетопологические структуры (свободные точки и линии "спагетти")</p> <p>Простая топология со связанными линиями - например сеть дренажа или сервисная инфраструктура</p> <p>Сложная топология со связанными линиями и вложенными</p>	<p>Регулярная треугольная, квадратная или шестиугольная сетка (квадратные пиксели = растр)</p> <p>Нерегулярная мозаика: многоугольники Тиссена Треугольные нерегулярные сети (TIN)</p>

<p>структурами - например связанные полигоны</p> <p>Сложная топология объектной ориентации со внутренними структурами и отношениями</p>	
---	--

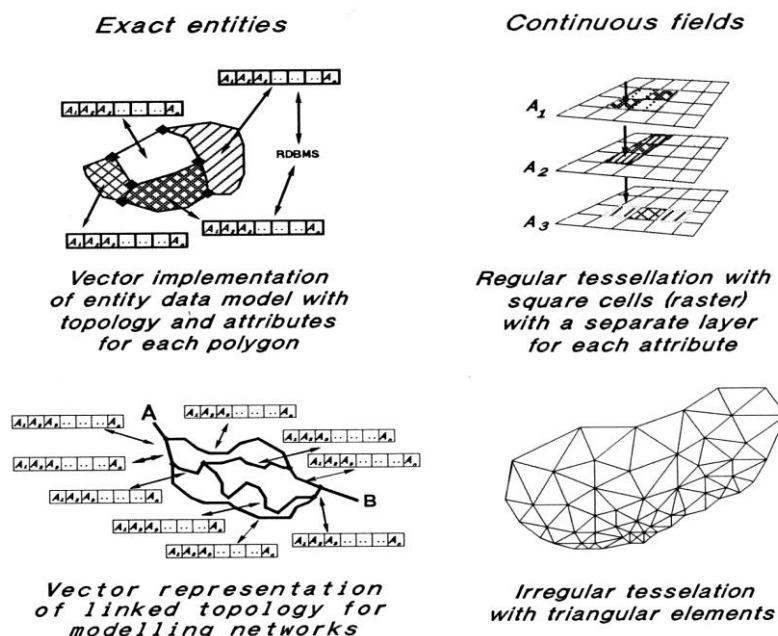


Рис. 2.4. Кодирование точных объектов и непрерывных полей в различных моделях данных, (а) верхний левый: векторное представление четких многоугольников; (б) верхний правый - растровая модель непрерывных полей; (с) левый нижний - векторное представление связанных линий; (д) правый нижний - триангуляция непрерывного поля методом Дилейни.

Отображение географических примитивов, используя векторные и растровые модели.

Люди привыкли видеть пространственную информацию как линиями так и разноцветными точками (например см. картины Сера, фотографию в газете под увеличением). На географической карте мы видим и векторные объекты и непрерывные поверхности в характеристике местности. Реальный мир изображается цветными и стилизованными точками, линиями или заливками разного тона (рельеф) и изолиниями. В компьютере это они формализованы в векторные и растровые методы представления. Рисунок 2.6 представляет главные пути, которыми простые географические модели данных могут визуализироваться в векторном или растровом виде. Рисунок

также передает влияние картографической семиотики в векторном представлении.

VECTOR	Points	Lines	Areas	RASTER	Points	Lines	Areas
Feature data				Feature data			
Areal units				Areal units		-	
Networks				Networks	-	-	-
Sampling records				Sampling records		-	
Surface data				Surface data		-	
Label/text	Utrecht Arnhem	A12	Lake	Label/text	-	-	-
Symbols				Symbols			
Relations	attributes and pointers	attributes and pointers		Relations	attributes and relations	attributes and relations	

Рис. 2.6. Различные пути графически отображающихся данных, (а) - векторные модели объектов, и (б) - растровые модели

Типы данных

В обычной речи мы отличаем качественные или номинальные атрибуты от количественных данных или чисел, и признаем, что различные виды операций требуют различные виды данных. Точно так же и при описании географических явлений, использующих формализованную модель данных. Информация может быть записана (и сохранена в компьютере) при использовании различных типов данных.

Тип	Допустимые величины	Допустимые операции
Булевы	0 или 1	Логические операции: правда или ложь
Номинальные	Любые имена	Логические операции, классификация и идентификация
Порядковые	Номера от 0 до бесконечности	Логические и ранжирующие операции, сравнения величин

Целые	Целые числа от 0 до бесконечности	Логические операции, арифметика целых чисел
Реальные	Реальные числа с десятичными от 0 до бесконечности	Все логические числовые операции
Топологические	Целые числа	Показывают связи между объектами

Атрибуты объектов могут быть выражены булевыми, номинальными, порядковыми, целочисленными, или реальными типами данных. Реальные типы данных включают десятичные числа; целые числа и реальные - это скалярные данные. Географические координаты иногда выражаются как целые числа, но главным образом как реальные типы данных, и топологические связи основаны на целых числах. Дифференцируемые непрерывные поверхности требуют реальных типов данных, хотя иногда используются как приближение целые числа. Не дифференцируемые непрерывные поверхности и их дискретные формы (ячейки сетки или пиксели) могут принимать тот же самый диапазон типов данных, как и объекты.

Логические операции могут быть выполнены со всеми типами данных, но арифметические операции ограничены реальными и целочисленными типами данных. Следовательно, вид анализа данных определяется типами данных, используемыми в модели данных. Пределы точности арифметических операций ограничены длиной компьютерного слова, используемого для записи чисел.

Аксиомы и процедуры для обработки данных в информационных системах

Исследовав проблемы определения моделей данных как представлений явлений в реальном мире и путях, которыми они могут быть сконструированы из географических примитивов как простые или сложные объекты или дискретизированные непрерывные поверхности, мы можем теперь определить логические основные правила и аксиомы, которые управляют способом обработки этих моделей данных. Хотя кое-что из нижеследующего может казаться самоочевидной, следующие утверждения обеспечивают формальное основание для пространственной обработки данных.

1. Необходимо определить типы дискретизации объектов, которые несут данные. В ГИС примитивные объекты - точки, линии, полигоны, и пиксели (элементы сетки). Сложные объекты, имеющие определенную внутреннюю структуру, могут быть сформированы из наборов точек, линий и полигонов.

2. Все фундаментальные объекты определены через их географическое положение (пространственные координаты или геометрия), их атрибуты (свойства) и отношения (топология). Эти отношения могут быть только геометрические (относительно пространственных отношений или соседей), или иерархические (относительно атрибутов) или то и другое.

3. Индивидуальные объекты различимы друг от друга их атрибутами, их расположением или их внутренними или внешними отношениями. В простом, статичном виде, объекты обычно внутренне гомогенны, если они не представляют математическую поверхность.

4. И объекты, и атрибуты могут быть классифицированы в соответствующие категории.

5. Над объектами, атрибутами, отношениями и группами можно исполнять логические операции (Булева алгебра).

6. В ГИС возможно подсчитать следующие параметры:

- расстояние;
- направление;
- связность (топология);
- смежность;
- близость;
- суперпозиция;
- членство в группе;
- принадлежность других объектов;

7. Новые объекты (или наборы объектов) могут быть созданы геометрическим объединением или пересечением существующих объектов (пересечение линий, перекрытие полигонов).

8. Новые сложные объекты могут быть созданы из набора основных точек, линий, площадей или пикселей.

9. Новые атрибуты могут быть получены из существующих атрибутов посредством логических и/или математических процедур.

Математические операции включают все действия арифметики (сложение, вычитание, умножение, деление, тригонометрия, дифференцирование и интегрирование и т.д.) в зависимости от типа данных. Новые атрибуты могут также быть получены из существующих топологических отношений и геометрических свойств (например связь с, или размер, форма, площадь, периметр) или интерполяцией.

10. Объекты, имеющие некоторые наборы именованных атрибутов, могут храниться в отдельных подблоках данных (планах и оверлеях).

11. Данные в координатах XYZt могут быть связаны со всеми подблоками данных.

12. Новые значения атрибутов в любой XYZt координатной системе могут быть получены как функция окружения (например вычисление уклона, экспозиции, связности).

Моделирование данных и пространственный анализ

Существует прямая связь между этими фундаментальными понятиями, моделью данных и типа данных используемых для представления географического явления, соответственно и типа анализа, который может быть применен. Следующие различные ситуации иллюстрируют это:

1. Если расположение и форма объекта неизменны и должны быть точно известны, но атрибуты могут изменяться, чтобы отразить различия в его состоянии, вызванные вводом новых данных, то используется векторное представление модели объекта. Это - обычная ситуация в ГИС.

2. Если атрибуты установлены, но объект может изменять форму, но не позицию (например, высыхание озера), то векторная модель требует переопределения границы через временной интервал. Растровая модель непрерывного поля, однако, обработала бы изменение водной поверхности как реакцию на процесс так, чтобы размер озера мог бы отображаться непрерывно.

3. Если никакие явные объекты не могут быть описаны, то предпочтительно обрабатывать явление как дискретизированное непрерывное поле.

Примеры использования моделей данных

Земельный кадастр

Главная цель кадастра или системного реестра земли состоит в том, чтобы обеспечить записи деления и использования земли. Важные моменты - местоположение, площадь рассматриваемой земли и ее атрибутов (типа наименования и адреса владельца), адрес рассматриваемого владения и информации о сделках и юридических вопросах. В этом случае точная векторная модель - наилучший выбор, используются номинальные, целочисленные и реальные типы данных для записи атрибутов и реальных типов данных координат. Границы четко определены. Важный аспект представления полигонов - то, что границы могут быть составлены смежными владениями. Это сохраняет двойное представление в базе данных и связывает границы в топологическую полигональную сеть, которая обрабатывает и смежность и включения.

Сервисные сети

Сервисная сеть - универсальный термин для совокупностей каналов и проводов, которые связывают здания потребителей к запасам воды, газа, электричества, телефонной связи, телеграфа, телевидения, а также к

системам распределения водопроводов и коллекторов. Эти коммуникации скрыты под землей и на поверхности. Три аспекта таких сетей важны для управления, а именно:

- (a) атрибуты данной сети (что несет, какой провод или канал, дополнительная информация относительно используемых материалов, возраста, название подрядчика, кто установил это, и так далее);
- (b) расположение сети (для ремонта и избежания повреждений);
- (c) информации относительно как, различные части сети связаны вместе.

Ясно все эти требования могут быть включены в модель данных топологически связанных линий, которые описаны атрибутами. Типы данных могут включать все формы.

Земная поверхность (типизация)

Ландшафты делятся согласно классам земель: городские покрытия, пахотные зерновые культуры, поле, лес, водные объекты, побережья, горы, и т.д. Создание такой модели данных для решения научных и прикладных задач требует выполнения следующих шагов. Сначала, необходимо определить точно, что предполагается классами. Во-вторых, как их распознавать, и третье, нужно выбрать методологию съемки, чтобы собрать данные, которые могут быть идентифицированы по классам, чтобы получить желательный результат.

Самая простая модель данных предполагает, что классы четки и взаимоисключающие и что есть прямая связь между классом и его расположением на земле. Если это приемлемо, тогда можно использовать простые примитивы полигоны как модель для каждого местонахождения каждого класса. Объекты, включенные в базу данных определяются классами, идентифицируя границы и закрепляя атрибуты к каждому классу способом, который является эквивалентным модели сети полигонов. Используемые типы данных - от целых (для регистрации названий классов) к реальным (для вычисления и регистрации областей).

Рассмотрим теперь случай, если есть разногласия о том, как классифицировать покрытие земли. Различные науки или организации пользуются различными типами классификации; даже при одинаковом количестве классов и сходных концепциях классов. Ясно, что любые исследования, основанные на данных из различных источников дали бы различные результаты, плохо сопоставимые между собой.

Теперь рассмотрим, как это отразится на результатах. Если мы собираем данные о земле, отбирая пробы, то полученные данные будут экстраполироваться на территории, где не отбирались пробы. Отнесение неизученных территорий к данному классу - функция качества и плотности данных и качества методики, используемой для интерполяции. Обратите внимание, что, если мы решаем интерполировать к резко очерченным

областям земли, мы используем модель, основанную на географическом примитиве - полигоне.

Если мы интерполируем поверхность как регулярный растр, тогда карта покрытия земли состоит из наборов пикселей с атрибутами, указывающими класс покрытия земли к которому они принадлежат. Если мы используем космоснимки, для определения типа покрытия земной поверхности, мы автоматически работаем с растром, потому что именно так работает спутниковый сканер. В отличие от случая отбора проб, мы имеем полную поверхность территории (исключая проблемы с облаками на изображении), так что информация представлена в каждом пикселе с равным качеством, которая отличается от интерполяции.

Главная проблема с идентификацией типа земли на снимке состоит в том, чтобы конвертировать измеренную отраженную радиацию для каждого пикселя в доминирующий класс покрытия земли. Очевидно, успех качества данных зависит от качества процесса классификации.

Этот пример показывает, что две различные, но дополняющие друг друга модели данных могут использоваться для отображения покрытия земли. Представление этих моделей как векторы или растры зависит частично от того, как данные были собраны и, частично, от того, как они будут использоваться.

Почвенные карты

Большинство изданных почвенных карт используют модель данных, основанную на векторном полигоне как географическом примитиве. Полигоны определены классами почвы, значение которых гомогенно. Границы представлены как бесконечно тонкие линии, подразумевающие резкие изменения почвы по очень коротким расстояниям. Обратите внимание, что внутренние полигоны - важный аспект почвенных и геологических карт.

Эта практичная модель данных, находя место на карте и определяя единицы карты можно отыскивать информацию относительно свойств почвы. Однако, эта парадигма с научной точки зрения неадекватна, потому что это игнорирует пространственные вариации почвы.

Типы данных, используемые для записи атрибутов, включают номинальные, целочисленные и реальные.

Критический аспект выделения классов почвы в географическом пространстве касается интерпретации границ на земле. Важные различия почвы, могут обозначаться резкими, ясно заметными изменениями в литологии, дренаже или перегибах уклонов, которые можно назвать "первичными границами". Однако границы почвы могут также просто отражать интерпретируемые различия в классификации почвы. Это "вторичные границы".

Главная проблема с почвой, растительностью и другими подобными естественными явлениями, состоит в том, что они изменяются пространственно во всех масштабах от миллиметров до целых континентов. Хотя ученые почвоведы признают это, картографы почвы все еще используют полигонально-векторную модель для отображения почвы на различных уровнях разрешающей способности. Это генерирует серьезную логическую ошибку, потому что принимается однородность в пределах модуля, в то время как существуют различия в почвах. Кроме того, принимая во внимание особенности землепользования, реальные границы почвы могут быть резкими, постепенными или диффузными.

Альтернатива к дискретной полигональной модели данных для почв должна предположить, что свойства почвы изменяются постепенно по ландшафту. Почва изучается полевыми методами в ряде пунктов и для них устанавливаются атрибуты. Самая простая модель данных - тогда географический пункт (точка) со значениями связанных атрибутов. От этой простой модели данных новые модели данных непрерывного пространственного изменения могут быть созданы интерполяцией.

Гидрология

Гидрологические прикладные программы нужны для моделирования транспорта воды и материалов в пространстве и времени. Кроме этого, изменение уровней воды на реках и озерах приводит к изменению геометрии и расположения водных объектов. Наводнение может приводить к появлению новых водотоков и исчезновению других, таким образом, меняется и положение и топология.

Простая векторная модель данных не очень хорошо подходит для манипуляции с гидрологическими явлениями, потому что изменения в геометрии означают изменение координат и топологических данных в сетях, которые требуют значительных пересчетов. Лучше использовать объектно-ориентированную модель, в которой первичные объекты связаны вместе в функциональные группы. Внутренняя структура модели содержит не только географическое расположение, геометрию, топологию, но также и информацию относительно того, как все компоненты реагируют на изменения.

Транспорт веществ может также быть легко зафиксирован моделями данных непрерывного поля. Использование треугольной или квадратной сети элементов обычно в гидрологических моделях. Транспорт вещества по поверхности можно также рассчитывать, используя растровую модель данных.

Резюме: объекты или поля?

В некоторых случаях решение о выборе подхода вполне ясно. В других - это возможно вопрос определяется в зависимости от целей пользователя.

Взаимное преобразование между векторным или растровым представлением технически возможно, если первоначальные объекты были ясно идентифицированы.

Никакая технология, однако, не может соединить различия в интерпретации, которая сделана прежде, чем явления зарегистрированы. Если один ученый чувствует ландшафт как набор четких объектов, представленных полигонами, его взгляд мира функционально отличается от мнения ученого, предпочитающего думать категориями непрерывного поля. Оба подхода могут исказить сложную действительность, которая не может быть описана полностью любой моделью.

Предпочтительный выбор объектов или непрерывных полей может измениться между прикладными программами и в пределах дисциплин. Вообще говоря, те дисциплины, которые имеют дело с описью и регистрацией статичных аспектов ландшафта выбирают векторный подход; дисциплины, имеющие дело с изучением процессов, выбирают модель непрерывного поля.

В большинстве ГИС все локальные данные и значения атрибутов считаются точными. Полагается, что нет никакой неопределенности. Очень часто это бывает не потому, что мы неспособны справиться со статистической неопределенностью, а скорее потому что это стоит слишком дорого, чтобы собирать или обработать все данные о полосах ошибки, которые должны быть связаны с каждым атрибутом для каждого модуля данных. В сущности, интеллектуальный уровень простых четких объектных моделей пространственных явлений немного отличается от детского конструктора. Эти игрушки повинуются основным аксиомам информационных систем, включая ту, которая говорит о возможности создать широкое разнообразие полученных объектов, объединяя различные блоки различными способами. Логически нет никакой разницы от комбинирования наборов точек, линий и полигонов из ГИС, чтобы сделать новую карту. Из достаточного количества кирпичей, можно формировать здания, воссоздавать ландшафты или даже создавать модели в натуральную величину животных подобно жирафам и слонам.

Жираф, сформированный из пластмассовых блоков, может иметь размер, цвет и форму реального жирафа, но модель не может иметь функций жирафа. Она не может ходить, есть, или дышать, потому что основные модули из которых она сформирована (блоки или элементы базы данных) - не способны к поддержке этих функций. Это тривиальный пример, одинаковые точки могут использоваться для многих модулей базы данных, которые используются, чтобы обеспечить географические данные для управления анализом или процессом модели. Никакое количество обрабатываемых данных не может обеспечивать истинные функциональные

возможности, если основные модули моделей данных не были должным образом выбраны.

Девять факторов при разработке и анализа в ГИС

Следующие девять вопросов относительно пространственных данных имеют фундаментальную важность при выборе моделей данных:

1. Реальная изучаемая ситуация (явление) - простая или сложная?
2. Типы объектов используемых для описания ситуации (явления), детальные или обобщенные?
3. Тип данных, используемый для записи атрибутов - булевы, номинальные, порядковые, целочисленные, реальные, или топологические?
4. Представляют ли объекты в базе данных реальные объекты, которые могут быть описаны точно, или этот комплекс объектов и возможно? Действительно ли их свойства точны, детерминированы или стохастические?
5. Объекты базы данных представляют дискретные физические вещи или непрерывные поля?
6. Атрибуты объектов, получены вычислением или наблюдением?
7. База данных будет использоваться для презентационных, административных или аналитических целей?
8. Требуется ли пользователю логические, эмпирические модели для получения новой информации из базы данных и следовательно делать выводы о реальном мире?
9. Процесс статический или динамический?