

Лекция 5

Понятия пространственного объекта и пространственных данных. Типы пространственных объектов. Общее цифровое описание пространственного объекта.

Цель лекции: получить понятие о модели пространственных данных, о цифровых и аналоговых данных, о пространственном объекте и пространственных данных.

Краткое содержание лекции. Информационную основу ГИС образуют цифровые представления (модели) реальности. С появлением компьютера все множество данных разделилось на два типа: цифровые и аналоговые данные. Последними стали именовать данные на традиционных «бумажных» носителях, используя этот термин как антоним цифровым данным. В отличие от аналоговой, цифровая форма представления, хранения и передачи данных реализуется в виде цифровых кодов или цифровых сигналов. Рассматривая данные по отношению к описываемым ими объектам, говорят о цифровых моделях объектов, а применительно к пространственным объектам в ГИС — о цифровых моделях пространственных объектов. Термин «цифровая модель» нельзя признать удачным — он отражает внешнюю форму их представления, а не его суть как набор логических правил построения системы из слагающих ее элементов — в данном случае элементарных (атомарных) пространственных объектов, имеющих аналогии в компьютерной графике и называемых там графическими примитивами. Цифровые по форме, по своей сути модели пространственных данных относятся к типу информационных моделей, отличных от реальных (например, физических), математических, мысленных или моделей особого типа, например, картографических. Объектом информационного моделирования в ГИС является пространственный объект. Это одно из ключевых понятий геоинформатики. Он может быть определен как цифровое представление (модель) объекта реальности (местности), содержащее его местоуказание и набор свойств (характеристик, атрибутов), или сам этот объект. Некоторое множество цифровых данных о пространственных объектах образует пространственные данные. Они состоят из двух взаимосвязанных частей: позиционной (тополого-геометрической) и непозиционной (атрибутивной) составляющих, которые образуют описание пространственного положения и тематического содержания данных соответственно.

Пространственные объекты как абстрактные представления реальных объектов и предмет информационного моделирования (цифрового описания) в ГИС разнообразны и традиционно классифицируются сообразно характеру пространственной локализации отображаемых ими объектов реальности, мерности пространства, которое они образуют, модели данных, используемой для их описания, и по другим основаниям. В рамках объектно-ориентированных моделей данные могут конструироваться в новые классы объектов, отличные от базовых или созданных ранее. Базовыми (элементарными) типами пространственных объектов, которыми оперируют современные ГИС, обычно считаются (в скобках приведены их синонимы) следующие:

- точка (точечный объект) — 0-мерный объект, характеризуемый плановыми координатами;
- линия (линейный объект, полилиния) — 1-мерный объект, образованный последовательностью не менее двух точек с известными плановыми координатами (линейными сегментами или дугами);
- область (полигон, полигональный объект, контур, контурный объект) — 2-мерный (площадной) объект, внутренняя область, ограниченная замкнутой последовательностью линий (дуг в векторных топологических моделях (данных) или сегментов в модели «спагетти») и идентифицируемая внутренней точкой (меткой);
- пиксел (пиксель, пэл) — 2-мерный объект, элемент цифрового изображения, наименьшая из его составляющих, получаемая в результате дискретизации изображения (разбиения на

далее неделимые элементы раstra); элемент дискретизации координатной плоскости в растровой модели (данных) ГИС;

- ячейка (регулярная ячейка) — 2-мерный объект, элемент разбиения земной поверхности линиями регулярной сети;
- поверхность (рельеф) — 2-мерный объект, определяемый не только плановыми координатами, но и аппликатой Z , которая входит в число атрибутов образующих ее объектов; оболочка тела;
- тело — 3-мерный (объемный) объект, описываемый тройкой (триплетом) координат, включающей аппликату Z , и ограниченный поверхностями.

Общее цифровое описание пространственного объекта включает:

- наименование;
- указание местоположения (местонахождения, локализации);
- набор свойств;
- отношения с иными объектами;
- пространственное «поведение».

Два последних элемента описания пространственного объекта факультативны.

Наименованием объекта служит его географическое наименование (имя собственное, если оно есть), его условный код и/или идентификатор, присваиваемый пользователем или назначаемый системой. В зависимости от типа объекта его местоположение определяется парой (триплетом) координат (для точечного объекта) или набором координат, организованным определенным образом в рамках некоторой модели данных, о которых речь пойдет ниже. Это геометрическая часть описания данных, геометрия (метрика) рассматриваемых пространственных объектов, отличная от их семантики (непозиционных свойств).

Перечень свойств соответствует атрибутам объекта, качественным и количественным его характеристикам, которые приписываются ему в цифровом виде пользователем, могут быть получены в ходе обработки данных или генерируются системой автоматически (к последнему типу атрибутов принадлежат, например, значения площадей и периметров полигональных объектов). Существует расширенное толкование понятия атрибута объекта; последнему могут быть поставлены в соответствие любые типы данных: текст, цифровое изображение, видео- или аудиозапись, графика (включая карту), что, по существу, реализуется на практике в мультимедийных электронных атласах. Под атрибутами понимаются именно содержательные, тематические (непозиционные, непространственные) свойства объектов.

Под отношениями понимают прежде всего топологические свойства (топологию). К топологическим свойствам пространственного объекта принято относить его размерность (мерность, пространственную размерность), сообразно которой выше были выделены 0, 1, 2 и 3-мерные объекты; замкнутость, если речь идет о линейных объектах в широком смысле слова; связность; простота (отсутствие самопересечения линейных объектов и «островов» в полигоне); нахождение на границе, внутри или вне полигона; признак точечного объекта, указывающий, является ли он конечным для некоторой линии. Примерами топологических отношений объектов являются их свойства «пересекаться» (или «не пересекаться»), «касаться», «быть внутри», «содержать», «совпадать»

Топология вместе с геометрией образует тополого-геометрическую часть описания данных, его позиционную часть.

Таким образом, в самом общем виде в пространственных данных следует различать и выделять три составные части: топологическую, геометрическую и атрибутивную — «геометрию», «топологию» и «атрибутику» цифровой модели пространственного объекта. Четкое разделение позиционных и непозиционных данных — историческая традиция, имеющая определенные технологические корни. Управление атрибутивной частью данных обычно возлагается на средства систем управления базами данных (СУБД), встроенных в программные средства ГИС или внешних по отношению к ним. В наиболее яркой форме

оно реализовано в рассмотренной ниже векторной модели данных, атрибуты которой представлены таблицей, хранятся и управляются СУБД, поддерживающей реляционную модель данных, а их позиционная часть, связанная с атрибутивной через идентификаторы пространственных объектов, управляется другими средствами. Модели пространственных данных такого типа получили широкое распространение и наименование геореляционных. Будучи еще недавно практически единственной и став классической, геореляционная модель не выглядит достаточно изящной. Современной альтернативой этой модели является интегрированный подход, когда и атрибутивная, и тополого-геометрическая части данных хранятся и управляются в единой среде СУБД, а также объектный и объектно-реляционный подходы (и одноименные им типы моделей данных).

Объектно-ориентированный подход к моделированию пространственных объектов вводит также понятие их «пространственного поведения».

Способы организации цифровых описаний пространственных данных принято называть моделями данных по традиции, унаследованной из теоретических обобщений проектирования систем управления базами данных. Они называются также цифровыми представлениями или просто представлениями пространственных данных.

На концептуальном уровне все множество моделей пространственных данных можно разделить на три типа: модели дискретных объектов, модели непрерывных полей и модели сетей.

Типами (классами) моделей именуют также модели, различающиеся по своему внутреннему устройству. В литературе существует множество классификаций моделей и наименований конкретных моделей. Построить исчерпывающую классификацию моделей пространственных данных вряд ли возможно: чуть ниже будет показано, что особенности моделируемой предметной области и специфические требования к функциональности ГИС могут потребовать разработки и использования весьма специальных моделей данных. Кроме того, как справедливо заметил Ю. К. Королев, «их нельзя расклассифицировать по одной оси, они различаются как бы «в разные стороны»

Тем не менее в практике геоинформатики уже достаточно давно определился набор базовых моделей (представлений) пространственных данных, используемых для описания объектов размерности не более двух (планиметрических объектов):

- растровая модель;
- регулярно-ячеистая (матричная) модель;
- квадратоминая модель (квадродерево, дерево квадратов, квадрантное дерево, Q-дерево, 4-дерево);
- векторная модель;
- векторная топологическая (линейно-узловая) модель;
- векторная нетопологическая модель (модель «спагетти»).

Растровая модель данных. Модель данных, именуемая растровой взамен устаревшего наименования матричной модели данных, имеет аналогии в компьютерной графике, где растр — прямоугольная решетка — разбивает изображение на составные однородные (гомогенные) далее неделимые части, называемые пикселями (от английского pixel, сокращение от «picture element» — элемент изображения), каждому из которых поставлен в соответствие некоторый код, обычно идентифицирующий цвет в той или иной системе цветов (цветовой модели). Из множества значений логических пикселей складывается цифровое изображение. Растровая модель данных в ГИС предполагает разбиение пространства (координатной плоскости) с вмещающими ее пространственными объектами на аналогичные пикселям дискретные элементы, упорядоченные в виде прямоугольной матрицы. Для цифрового описания (позиционирования) точечного объекта при этом будет достаточно указать его принадлежность к тому или иному элементу дискретизации, учитывая, что его положение однозначно определено номерами столбца и строки матрицы (при необходимости координаты пикселя, либо его центроида или любого угла могут быть вычислены). Пикселу присваивается цифровое значение, определяющее имя или семантику

(атрибут) объекта. Аналогичным образом описываются линейные и полигональные объекты: каждый элемент матрицы получает значение, соответствующее принадлежности или непринадлежности к нему того или иного объекта.

Регулярно-ячеистая модель данных. Описанная выше растровая модель данных пригодна для цифрового представления не только пространственных объектов в ГИС, но и изображений. Примерами могут служить цифровые фотоизображения, снятые непосредственно цифровой фотокамерой или полученные путем цифрования аналоговых негативов или фотоотпечатков насканере хорошего разрешения и далее превращенные (возвращенные) в графику на страницах иллюстрированных журналов или в семейном фотоальбоме. Данные дистанционного зондирования Земли — аэроснимки и космические снимки, получаемые с борта космических платформ и других летательных аппаратов и представляющие собой, как отмечалось ранее, один из основных источников данных для ГИС, в настоящее время в существенной своей части по форме тоже цифровые, образуют класс растровых цифровых изображений, обрабатываемых программными средствами цифровой обработки изображений. Растровой цифровой копией можно назвать оцифрованную на том же сканере бумажную карту, используемую в качестве графической подложки (растровой цифровой карты-основы) в малозатратных геоинформационных проектах.

Во всех перечисленных выше случаях речь идет о цифровых растровых изображениях, образованных множеством его элементов — пикселей, каждому из которых ставится в соответствие значение (код) цвета или спектрального коэффициента яркости объекта съемки. На эти далее неделимые элементы растра «разбивается» и координатная плоскость с пространственными объектами в их растровом представлении.

Если атомарной единицей данных при их описании служит элемент «разбиения» территории — регулярная пространственная ячейка (территориальная ячейка) правильной геометрической формы — речь идет о другой, отличной от растровой, хотя и формально с нею схожей, регулярно-ячеистой модели данных. Формальное сходство абсолютно в случае квадратной (прямоугольной) формы ячеек, хотя известны примеры регулярных (или квазирегулярных) сетей (решеток) с ячейками правильной треугольной, гексагональной или трапециевидной формы, равновеликих или квазиравновеликих. При этом сеть может строиться (разумеется мысленно) на плоскости в условных прямоугольных координатах некоторой картографической проекции или на поверхности шара или эллипсоида; в последнем случае регулярными ячейками обычно являются сферические трапеции фиксированного или переменного углового размера.

Квадратомическая модель данных. Обычно описание этой своеобразной модели начинают с того, что главный мотив ее использования и поддержки программными средствами ГИС — компактность по сравнению с растровой моделью, расточительной по объемам машинной памяти, требуемой для хранения растровых данных. Рассматривая растровые данные, мы упоминали о возможностях их значительного сжатия. Не меньшей эффективностью при сохранении быстрого доступа к элементам описания пространственных объектов обладает квадратомическая модель данных, основанная на подходе, известном как квадратомическое дерево (квадро-дерево). В его основе лежит разбиение территории или изображения на вложенные друг в друга пиксели или регулярные ячейки с образованием иерархической древовидной структуры — декомпозиции пространства на квадратные участки (квадраты, квадратные блоки, квадранты), каждый из которых делится рекурсивно на четыре вложенных до достижения некоторого уровня пространственного разрешения. Линейная квадратомическая модель данных практически реализована в немногих из известных программных средств ГИС. Ее не следует путать с так называемыми пирамидальными моделями, также рекурсивно организуемыми и используемыми при обработке цифровых изображений, включая данные дистанционного зондирования. В отличие от квадродерева, они представляют собой набор растровых слоев изображений с механически двукратно уменьшаемым разрешением, а

потому более компактных, искусственно «затрубленных» в целях их быстрого воспроизведения. Известны трехмерные расширения линейной квадротомической модели — это так называемая октотомическая модель (модель октарного дерева) как результат рекурсивного деления трехмерного пространства на восемь октантов, используемая для цифрового описания тел. Предлагалась также модель трихотомического дерева с аналогичным квадротомическому делению треугольных элементов.

Векторные модели данных. Обобщенный класс векторных моделей включает два их типа: векторные топологические и векторные нетопологические модели. Они используются для цифрового представления точечных, линейных и площадных (полигональных) объектов, имея аналогии в картографии, где различаются объекты с точечным, линейным и площадным характером пространственной локализации, что определяет выбор графических средств их картографического отображения, и исторически связаны с технологиями цифрования карт, планов и другой графической документации с помощью устройств ввода векторного типа — дигитайзеров (цифрователей) с ручным обводом, генерирующих поток пар плановых координат (векторов) вслед за движением курсора (обводной головки) по планшету дигитайзера при отслеживании и записи графических объектов помещенного на него оригинала.

Множество точечных объектов, образующее слой однородных данных (например множество объектов, соответствующих населенным пунктам), может быть представлено в векторном формате в виде неупорядоченной (необязательно упорядоченной) последовательности записей (строк таблицы), каждая из которых содержит три числа: уникальный идентификационный номер объекта (идентификатор), значение координаты X и значение координаты Y в системе условных плановых прямоугольных декартовых координат, например, плоскости стола дигитайзера. Линейный объект (в общем случае кривая) или граница полигонального объекта могут быть представлены в виде последовательности образующих их точек (промежуточных точек), т.е. набором линейных отрезков прямых (сегментов), образующих полилинию.

Расположение образующих полилинию точек будет зависеть от структуры исходной кривой. Их привязка к характерным точкам кривой при достаточно мелком шаге цифрования позволяет дать достаточно точное ее приближение. Запись линейного объекта образована последовательностью координатных пар (в нашем случае пяти точек) и содержит элемент, позволяющий выделить его в общей совокупности записей линейных объектов слоя, которому соответствует обычно файл данных. В случае А это делается путем помещения вслед за идентификатором целого числа, указывающего число координатных пар, в варианте В линейные объекты отделяются друг от друга меткой END. Разумеется, запись должна быть снабжена идентификатором объекта.

Таким же образом может быть представлена граница полигонального объекта. При этом каждый именованный полигон (со своим идентификатором) представляется записью пар координат, образующих его границу в избранной последовательности (например, по часовой стрелке). При описании множества полигонов каждый отрезок границы, заключенный между двумя узловыми точками (за исключением внешней границы полигонов), будет описан в этом случае дважды (по и против часовой стрелки). Такая модель данных для описания точечных, линейных и полигональных объектов носит наименование модели «спагетти». Она не является эффективной с точки зрения избыточности хранимых данных и возможностей использования аналитических операций ТИС и поддерживается недорогими программными средствами настольного картографирования и ГИС. Другое ее наименование — векторная нетопологическая модель.

Векторная топологическая модель обязана своим происхождением задаче описания полигональных объектов. Ее называют еще линейно-узловой моделью. С ней связаны и особые

термины, отражающие ее структуру; главные ее элементы (примитивы):

- промежуточная точка;
- сегмент (линейный сегмент, отрезок (прямой));
- узел;
- дуга;
- полигон (область, полигональный объект, многоугольник, контур, контурный объект), в том числе:
 - простой полигон;
 - внутренний полигон («остров», анклав);
 - составной полигон;
 - универсальный полигон (внешняя область).

Описание полигона в векторной топологической модели — это множество трех элементов: узлов, дуг и собственно полигонов. Между этими объектами устанавливаются некоторые топологические отношения, необходимым элементом которых должна быть связь Дуг и узлов, полигонов и дуг. Последнее, что стоит отметить о векторной топологической модели данных — связь между позиционной частью и атрибутикой.

В классическом варианте их взаимодействия, когда атрибутивные данные управляются средствами реляционной СУБД и организованы в таблицы, а модель взаимодействия носит название геореляционной, как упоминалось выше, связь устанавливается и поддерживается через идентификатор объекта