

Д. Е. АНДРИАНОВ, М. С. СОКОЛОВ

ИСПОЛЬЗОВАНИЕ ТОПОЛОГИЧЕСКИХ ПРАВИЛ ПРИ ПРОСТРАНСТВЕННОМ АНАЛИЗЕ КАРТОГРАФИЧЕСКИХ ОБЪЕКТОВ

Рассматривается подход к применению топологических отношений при групповом пространственном анализе картографических объектов. Введено понятие массива из матриц топологических отношений, хранящих взаимосвязи между объектами, что позволяет произвести их выборку в соответствии с заданными правилами.

Ключевые слова: картографический объект, топологическое отношение, геоинформационные системы, матрица топологических отношений, правило отбора.

Основой эффективной работы геоинформационной системы (ГИС) являются топологические отношения, позволяющие структурировать пространственно-распределенные данные.

Анализ топологических отношений позволяет произвести наиболее естественную выборку графических объектов, представленных на карте в векторной форме [1]. На основе пространственных отношений можно сформировать условия соответствия одного объекта или группы объектов некоторому набору правил.

Типичными пространственными запросами в ГИС являются, например, такие, как „найти все города, которые лежат в пределах 5 км от дороги N^c “ или „найти все дороги в городах, смежных с городом X^c “.

Существующие коммерческие языки запросов к базам данных (например, SQL) не обеспечивают достаточную функциональность, так как предлагают лишь средства для определения равенства или порядка простых типов данных (целого числа или строки) [2, 3].

Пространственные запросы могут быть реализованы, если все топологические отношения между исследуемыми объектами явно сохранены. Однако такой сценарий нереалистичен, поскольку требует большого объема оперативной памяти для каждого вида пространственных отношений между объектами [4]. Вместо хранения всех пространственных отношений, на практике они определяются на основе геометрического или пространственного расположения объектов.

Задача алгоритма топологического отбора — быстрое выполнение пространственного запроса пользователя. Это возможно на основе определения пространственных отношений до начала обработки запроса, кроме того информация о них должна быть размещена в оперативной памяти для быстрой обработки [4].

Введем понятие многомерного массива из матриц, хранящих информацию о межобъектных топологических отношениях; каждая из матриц — это кортеж на множестве, представляющем пространственные объекты, для которых определяется конкретная пространственная взаимосвязь.

Таким образом, решение задачи быстрого выполнения пространственных запросов — это построение указанных матриц топологических правил на первоначальном этапе работы алгоритма; дальнейшая работа с матрицами осуществляется с минимальными задержками. Фактически, формируется куб со слоями, представленными в виде топологических двумерных массивов, изображенных на рис. 1.

Элементами матриц являются бинарные значения, указывающие наличие или отсутствие одного из отношений между объектами, определяемыми порядковыми номерами столбца и строки.

Представление матриц в виде слоев куба дает возможность применения технологии OLAP — обработки данных в режиме реального времени. При этом куб, который может храниться на выделенном сервере, представляет собой хранилище фактов (межобъектных топологических отношений), каждому факту ставится в соответствие уникальное сочетание элементов измерений (наличие или отсутствие топологии).

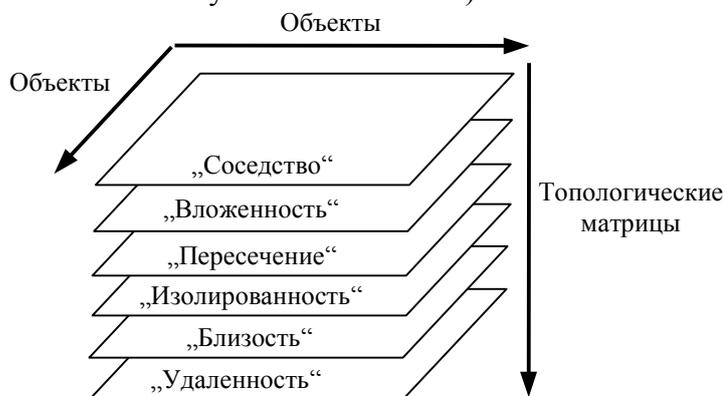


Рис. 1

Основной проблемой при организации пространственных запросов является высокая трудоемкость формирования топологических матриц на первоначальном этапе. Для снижения трудоемкости предлагается использовать возможность конкретного указания слоев и стилей карты, где отражены объекты, подлежащие анализу при отборе. Таким образом значительное уменьшение объема оперативной памяти и временных затрат на обработку данных будет достигнуто благодаря следующим факторам:

- конкретизации класса объектов, подлежащих поиску;
- использованию матриц, содержащих межобъектные топологические отношения и хранящихся в оперативной памяти, при этом сканирование матриц будет осуществляться с высокой скоростью.

Для конкретного класса объектов задается бинарная структура отношений, поэтому в основу пространственного запроса заложены бинарные операции сравнения.

Описание алгоритма. Точечные, линейные и полигональные объекты — это пространственные элементы, которые называют примитивами или базовыми элементами [2, 3].

Элементом x_i картографического объекта X назовем примитив, который можно отнести к точечному, линейному или полигональному типу. Картографическим объектом X будем называть объект, который состоит из примитивов. Простым картографическим объектом называется объект, состоящий из множества базовых элементов, так что

$$X = \{x_i\}.$$

Сложным называется картографический объект, который состоит не менее чем из двух базовых элементов:

$$X = \{x_1, x_2, \dots, x_n\}.$$

Как показано в работе [3], в зависимости от набора сущностей, участвующих в пространственном взаимодействии, можно выделить три класса топологических отношений (внутриобъектные, межобъектные и концептуальные). Каждый класс определяет большую группу топологических отношений, которые отражают элементы внутренней или внешней топологии. Указанные классы могут перекрываться и дополнять друг друга.

Для эффективной работы средств отбора данных в ГИС необходимо найти сбалансированную схему порядка определения топологических отношений. Таким образом, указанные выше классы могут характеризоваться следующим образом:

1) внутриобъектные — топологические отношения, определяющие целостность объекта как совокупности элементарных геометрических частей и семантического контекста (компонентами топологического отношения являются геометрические примитивы, составляющие объект);

2) концептуальные — топологические отношения, устанавливающие наиболее общие правила расположения объектов, относящихся к разным классам;

3) межобъектные — топологические отношения, устанавливаемые между парой объектов.

Разделение топологических отношений на классы будем производить в соответствии с совокупностью объектов, участвующих в отношении [3].

Согласно постановке задачи по выборке картографических объектов необходимость определения сложных топологических отношений отсутствует — для формирования топологических матриц достаточно определить межобъектную топологию. В этой связи выделим следующие типы межобъектных топологических отношений: изолированность, пересечение, вложенность, соседство, близость, удаленность.

Однако, говоря о межобъектной топологии, необходимо отметить фактор влияния на нее внутренних отношений между составными частями картографического объекта (рис. 2). Фактически каждое установленное межобъектное топологическое отношение основывается на анализе взаимосвязей базовых элементов — точек, линий, полигонов. Важность этих взаимосвязей можно продемонстрировать на следующем примере: при выборке всех строений, расположенных на определенном расстоянии от заданной улицы, необязательно анализировать все примитивы, составляющие строение (обычно это сложные контуры, состоящие из нескольких форм), а целесообразно проводить вычисления только для внешних (по отношению к остальным) примитивов, что позволит сократить время расчета. При анализе любых отношений, кроме изолированности, выполнение правила для одного из примитивов означает, что оно выполняется и для объекта в целом.

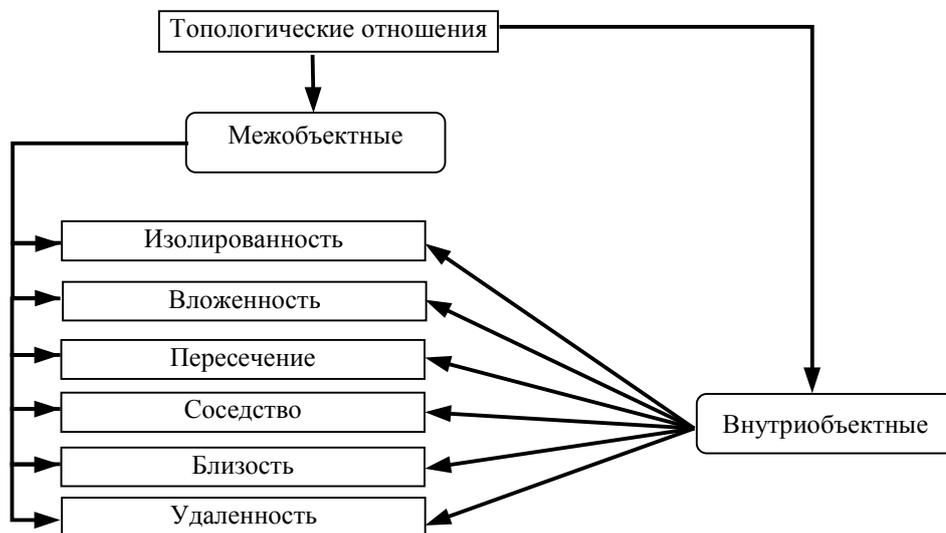


Рис. 2

Пусть имеется совокупность картографических объектов

$$K = \{X_n\},$$

где n — количество объектов, а также совокупность сформированных правил P отбора

$$L = \{P_m\},$$

где m — количество правил.

Тогда задача алгоритма по выборке объектов состоит в проверке множества K на соответствие входящих в него элементов отношениям множества L .

Принцип работы алгоритма. Алгоритм основан на предварительном формировании шести матриц по числу базовых межобъектных топологических отношений. Строки и столбцы матриц соответствуют идентификаторам картографических объектов, на пересечении находится бинарный элемент, показывающий наличие или отсутствие топологического отношения.

Подробная схема работы алгоритма может быть представлена следующим образом.

1. Пользователь указывает объекты каких слоев и стилей карты должны подлежать анализу. Пользователь должен указать также меру близости для расчетов. В инструментальной ГИС стиль и слой определяются идентификаторами, которые обозначим как ID₁ и ID₂.

2. Осуществляется обработка множества K по следующему правилу:

$$M[i, j] = O(X_i, X_j), (X)_{\text{style}} = \text{ID}_1 \text{ и } (X)_{\text{layer}} = \text{ID}_2 \quad \forall X \in K, \forall P \in L,$$

где i, j — порядковые номера строки и столбца матрицы, соответствующие картографическим объектам; M — топологическая матрица; O — топологическое отношение; индексы „style“, „layer“ отражают свойства картографического объекта, определяющие его отношение к стилю и слою отображения соответственно.

3. Формируются простые (одно топологическое отношение) или сложные (комбинация топологических отношений) правила выборки объектов (множество L).

4. Выборка осуществляется посредством анализа каждой топологической матрицы, для которой формируется свой список S соответствующих объектов, т.е. для каждого i -го объекта определяются другие (с номерами j), находящиеся с ним в каком-либо пространственном отношении:

$$S \leftarrow \text{Get ID}(j), \text{ если } M[i, j]=1 \text{ и } i \neq j \quad \forall i, j \in (1 \dots n),$$

где n — количество объектов, Get ID — оператор определения идентификатора объекта по его номеру в матрице.

5. Для сложного (составного) правила производится объединение сформированных списков S в соответствии с компоновкой правила (конъюнкция, дизъюнкция, отрицание):

$$Q = S_1 \text{ (и/или/нет) } S_2 \dots S_P,$$

где P — число простых правил в сложном.

6. Результат поиска представляется пользователю в виде списка найденных объектов, возможен просмотр результатов на карте, удаление объектов, модификация.

Для оптимизации алгоритма предпринимаются следующие шаги:

1) при определении отношений „соседство“, „изолированность“, „близость и удаленность“, „пересечение“ возможно сокращение вычислительных затрат по меньшей мере в два раза, т.е. при формировании матрицы можно заполнять лишь ее половину, так как при выполнении этих правил для одного объекта пары они выполняются и для другого; кроме того, можно провести оптимизацию следующего характера: если определены отношения „соседство“ или „пересечение“, то отношения „изолированность/близость/удаленность“ для этих же объектов не выполняются и их определять не нужно, и наоборот; взаимоисключающими для пары объектов являются также отношения „соседство“ и „пересечение“;

2) при определении отношения „вложенность“ возможно некоторое снижение вычислительных затрат за счет исключения повторных проверок на вложенность для пары объектов, а также в случае если все остальные, рассчитанные ранее отношения не выполняются;

3) анализ отношения „изолированность“ возможен не поэлементно, а по столбцам ввиду исключительности данного правила: если в строке, соответствующей картографическому объекту, нет ни одного бинарного элемента со значением „0“, то объект изолирован.

Результаты исследования алгоритма приведены на рис. 3, где представлен график его работы при обычной выборке (кривая 1) и при выборке, произведенной на основе матриц (кривая 2).

Оценим потребность алгоритма в оперативной памяти: для хорошо детализированной карты среднестатистического города число объектов может достигать 1 млн., соответственно объем одной матрицы составит $10^6 \cdot 10^6$ байт или $\sim 0,91$ Тбайт, а для всех матриц — $0,91 \cdot 5 = 4,54$ Тбайт. Число объектов на одной карте может различаться в зависимости от текущих масштабов отображения и активных слоев, но общая тенденция, выражающаяся в экспоненциальном росте загрузки памяти с увеличением числа объектов, неизменна.

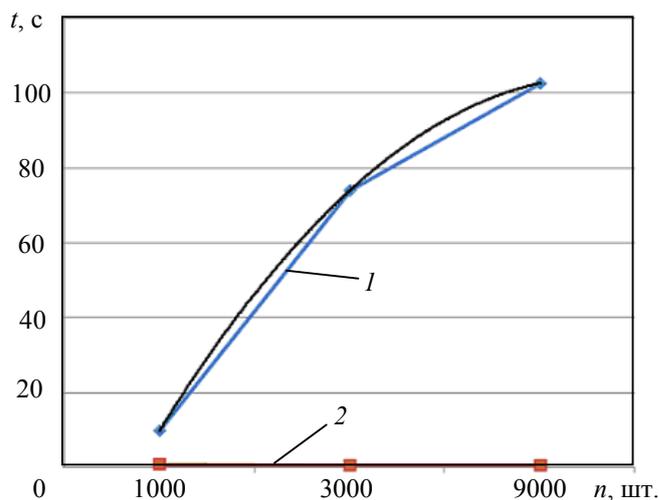


Рис. 3

Заключение. Рассмотренный алгоритм анализа картографических объектов позволяет организовать быструю выборку данных по топологическим правилам и упрощает работу пользователя при поиске объектов. Представленный в виде подсистемы расширения ГИС этот алгоритм может быть составной частью муниципальной геоинформационной системы.

Преимущества использования алгоритма очевидны при выполнении множества процедур отбора данных по различным правилам: используются однажды сформированные и сохраненные топологические матрицы, скорость сканирования которых не сопоставима со временем выборки объектов из ГИС. Несмотря на оптимальность расчетных топологических процедур, первоначальный этап формирования поисковых матриц весьма трудоемок, так как зависимость времени работы от количества объектов является параболической.

Трудоемкость выполнения алгоритма и высокая потребность в оперативной памяти обуславливают необходимость применения технологии, подобной OLAP. При этом можно обеспечить следующие преимущества:

- пользователь сам осуществляет детализацию матриц топологических отношений, выбирая интересующие его классы объектов, которые и формируют итоговые слои куба;
- факты (межобъектные топологические отношения), представленные в многомерном матричном виде, легко обрабатываются в соответствии как с простыми, так и сложными правилами отбора (пространственными запросами на основе бинарных операций сравнения), при этом анализируются бинарные структуры данных;
- учитывая многомерное послойное представление топологической информации, можно говорить о возможности распараллеливания обработки пространственных запросов.

СПИСОК ЛИТЕРАТУРЫ

1. Садыков С. С., Андрианов Д. Е., Еремеев С. В. Формальное определение топологических отношений между картографическими объектами в ГИС // Обработка информации: методы и системы: Сб. науч. статей. М.: Горячая линия — Телеком, 2003.
2. Андрианов Д. Е., Булаев А. В. Алгоритм представления сложных топологических отношений в геоинформационных системах // Тр. Рос. науч.-техн. общества радиотехники, электроники и связи им. А. С. Попова. Сер. Цифровая обработка сигналов и ее применение. М., 2006. Вып. VIII-2.
3. Egenhofer M. Spatial query languages // Ph. D. Thesis: University of Maine. Orono, ME. 1989.
4. Egenhofer M. A formal definition of binary topological relationships // 3rd Intern. Conf. on Foundations of Data Organization and Algorithms (FODO), Paris: France Lecture Notes in Computer Science, 367. Berlin & N.Y.: Springer-Verlag, 1989. P. 457—472.

Сведения об авторах

- Дмитрий Евгеньевич Андрианов** — д-р техн. наук, доцент; Муромский институт (филиал) Владимирского государственного университета, кафедра информационных систем;
E-mail: AndrianovDE@inbox.ru
- Михаил Сергеевич Соколов** — аспирант; Муромский институт (филиал) Владимирского государственного университета, кафедра информационных систем;
E-mail: acp.devel@googlemail.com

Рекомендована кафедрой
информационных систем

Поступила в редакцию
14.04.10 г.