ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 1998
Seria: INFORMATYKA z. 34 Nr kol. 1381

Katarzyna STĄPOR, Piotr BAJERSKI
Politechnika Śląska, Instytut Informatyki

WYBRANE METODY INDEKSACJI PRZESTRZENNEJ W GEOGRAFICZNYCH SYSTEMACH INFORMACYJNYCH

Streszczenie. Efektywne przetwarzanie zapytań w przestrzennych bazach danych wymaga metod indeksacji grupujących obiekty według ich położenia w przestrzeni. W artykule opisano dwie najczęściej wykorzystywane do tego celu struktury danych: R-drzewa i drzewa czwórkowe. Omówione zostały odmiany tych struktur oraz wykorzystujące je algorytmy realizacji zapytań przestrzennych.

ON SOME SPATIAL INDEXING METHODS IN GEOGRAPHICAL INFORMATION SYSTEMS

Summary. Effective query execution against spatial databases demands for indexing methods based on objects location in space. This article describes two data structures most often used for this purpose: R-trees i quadtrees. Some variations of these two structures and algorithms using them for spatial query execution are presented.

1. Wstęp

W ciągu ostatnich dwóch dziesięcioleci widać wyraźny i stały wzrost liczby systemów informacyjnych, które przy przetwarzaniu danych uwzględniają ich położenie w przestrzeni. Do grupy tej z jednej strony należą geograficzne systemy informacyjne (Geographical Information System, GIS) będące wyspecjalizowanymi narzędziami o bardzo dużych możliwościach w zakresie modelowania i analizy zależności przestrzennych, z drugiej strony do wielu aplikacji dołączane są moduły dostarczające najprostszych funkcji w tym zakresie. Specyfiką baz przechowujących informacje o położeniu, kształcie i wzajemnych relacjach przestrzennych obiektów jest ich duży rozmiar, złożone struktury danych i długi czas

wykonywania zapytań. Rodzi to wiele problemów, nad którymi trwają intensywne badania. Jednym z najbardziej istotnych jest przyśpieszenie dostępu do obiektów na podstawie ich położenia w przestrzeni.

Dane opisujące położenie obiektów w przestrzeni oraz ich kształt (geometrię) będą w dalszej części nazywane danymi przestrzennymi. Podstawowymi typami danych przestrzennych, nazywanych prymitywami, są: punkt, linia, i wielokąt (ang. polygon). Bardziej skomplikowane figury są tworzone z wymienionych prymitywów. Dane przestrzenne dotyczące danego obiektu są nazywane jego atrybutami przestrzennymi. Na przykład dla obiektu las atrybutem przestrzennym może być granica, będąca reprezentacją linii brzegowej lasu. Zbiór obiektów przestrzennych posiadających ten sam zestaw atrybutów definiuje relację przestrzenną (np. relacja lasy z atrybutem granice opisującym jej geometrię).

Główną częścią każdej aplikacji przetwarzającej duże ilości danych przestrzennych jest system przestrzennej bazy danych (ang. Spatial Database System), odpowiadający za ich efektywne gromadzenie i przetwarzanie. System przestrzennej bazy danych powinien [5]:

- dostarczać mechanizmów charakterystycznych dla klasycznego systemu bazy danych, takich jak: gromadzenie i przetwarzanie danych alfanumerycznych (nieprzestrzennych), zarządzanie współbieżnym dostępem do danych, odtwarzanie po awarii, autoryzacja dostępu,
- posiadać wbudowane w model danych i język zapytań tzw. przestrzenne typy danych (ang. spatial data type), np. *point, line, region* wraz z przestrzennymi operacjami na tych typach, np. *p intersects q, area(p) > x*,
- zapewniać efektywną realizację zapytań przestrzennych (np. selection, join), co wymaga przynajmniej jednego indeksu dla danych przestrzennych.

Zapytania dotyczące atrybutów przestrzennych nazywane są zapytaniami przestrzennymi. Ze względu na złożoność dostępu do danych są one dzielone na dwie grupy [2]:

- wymagające pojedynczego dostępu do danych w trakcie wykonywania zapytania (ang. single-scan). Charakteryzują się liniową lub lepszą zależnością czasu wykonania zapytania od liczby obiektów w tablicy. Najbardziej znanymi przykładami tej grupy zapytań są:
  - zapytania typu położenie punktu względem wielokąta (ang. point-in-polygon query): mając dany punkt P znajdź wszystkie prostokąty R, zawierające P,
  - zapytania typu okno (ang. window query): mając dany prostokąt S znajdź wszystkie prostokąty R przecinające (zawierające, należące do) S,
- wymagające wielokrotnego dostępu do niektórych obiektów w trakcie wykonywania zapytania (ang. multiple-scan). Najbardziej znanym przykładem takiego zapytania jest złączenie przestrzenne (ang. spatial join), które znajduje wszystkie pary obiektów wskazanych typów, spełniające zadaną własność. Na przykład, dla tablic *lasy* i *miasta*,

gdzie atrybut przestrzenny reprezentuje granice lasów i miast, złączenie przestrzenne pozwala odpowiedzieć na pytanie: „Które lasy leżą na terenach miejskich?". Wyznaczenie odpowiedzi w pesymistycznym przypadku wymaga porównania wszystkich obiektów z pierwszego zbioru ze wszystkimi obiektami z drugiego zbioru.

W zapytaniach przestrzennych najczęściej stosuje się przetwarzanie dwuetapowe. W pierwszy etapie wykorzystuje się specjalne struktury danych oraz proste i szybkie algorytmy pozwalające na znalezienie obiektów, które mogą spełniać warunki zapytania. W drugim etapie wykonuje się dokładne, bardziej czasochłonne obliczenia, aby określić, które obiekty rzeczywiście spełniają kryteria zapytania. Artykuł jest poświęcony metodom dostępu do danych pozwalającym przyśpieszyć pierwszy etap.

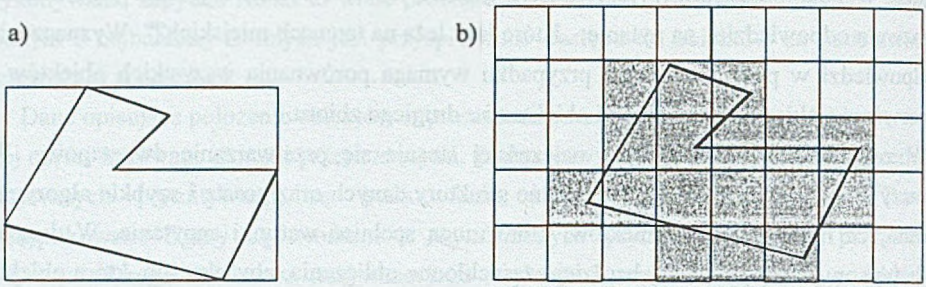
2. Charakterystyka metod indeksacji danych przestrzennych

Aby móc efektywnie realizować zapytania przestrzenne, obiekty muszą zostać posortowane w oparciu o atrybut wyrażający ich położenie w przestrzeni. Struktury danych umożliwiające efektywny dostęp do obiektów na podstawie ich położenia w przestrzeni nazywane są indeksami przestrzennymi (ang. spatial index) lub przestrzennymi metodami dostępu (ang. spatial access method, SAM).

Wiele metod indeksacji przestrzennej wykorzystuje przybliżenia (aproksymacje) rzeczywistych obiektów prostszymi figurami geometrycznymi. Aproksymacja umożliwia znaczne zmniejszenie ilości danych, które muszą być przechowywane i przetwarzane oraz użycie szybszych algorytmów przy zachowaniu podstawowych charakterystyk obiektów, takich jak położenie oraz rozciągłość wzdłuż wszystkich osi. Stosowane są dwa rodzaje aproksymacji obiektów:

- ciągła, której przykładem jest minimalny prostokąt ograniczający (ang. Minimum Bounding Rectangle, MBR), tzn. najmniejszy zawierający dany obiekt prostokąt, o osiach równoległych do osi układu współrzędnych (rys. 1a),
- siatkowa, w której przestrzeń jest podzielona za pomocą siatki na komórki, a obiekt jest reprezentowany poprzez zbiór komórek, które przecina (rys.1b).

Z punktu widzenia możliwości wykorzystania klasycznych serwerów baz danych metody indeksacji przestrzennej można podzielić na metody wykorzystujące dedykowaną zewnętrzną strukturę danych oraz metody odwzorowujące obiekty przestrzenne w 1-wymiarową przestrzeń, tak że możliwe jest użycie standardowej struktury indeksu, np. B-drzewa.



Rys. 1. Aproksymacje obiektu przestrzennego: MBR (a) oraz siatkowa (b)
Fig. 1. MBR (a) and grid (a) approximation of a spatial object

Metody indeksacji można podzielić ogólnie na dwie rodziny: metody dokonujące transformacji danych w przestrzeń o innej liczbie wymiarów i metody operujące na danych przestrzennych nie transformowanych, czyli działające w przestrzeni, w której obiekty zostały narysowane. Metody drugiej rodziny można z kolei podzielić na cztery grupy na podstawie sposobu dekompozycji przestrzeni:

1. metody wykorzystujące aproksymacje MBR obiektów przestrzennych, które dokonują ich grupowania w hierarchie. Zachowana zostaje własność bliskości położenia. Przykładem jest np. R-drzewo [4]. Wyróżnione fragmenty przestrzeni nie są rozłączne, co wymaga nieraz przeszukania znaczącej części indeksu, jeśli obiekt leży na obszarze pokrywanym przez kilka aproksymacji,
2. metody dokonujące dekompozycji aproksymacji obiektów na rozłączne podobiekty i grupowania otrzymanych w ten sposób podobiektów w hierarchie. Wyróżnione fragmenty przestrzeni są rozłączne. Aby określić obszar pokryty przez jakiś obiekt, należy przeszukać wszystkie komórki, które go zawierają. Przykładem jest R+-drzewo[12],
3. metody dokonujące rozłącznej dekompozycji przestrzeni na regularne bloki o jednakowym rozmiarze. Metoda nadaje się dla danych rozmieszczonych równomiernie,
4. metody dokonujące adaptacyjnej dekompozycji przestrzeni w zależności od danych. Przykładem może być drzewo czwórkowe (ang. quadtree) [11]. Rozmiar bloku jest potęgą liczby 2, a jego pozycja jest ograniczona. Metodę można stosować dla danych o dowolnym rozkładzie. W przypadku danych równomiernie rozłożonych degeneruje się ona do jednolitej siatki (grupa 3).

Metody 1 i 2 grupy są zależne od danych, a więc pojawiają się trudności przy złożeniu różnych operacji i zbiorów. Metody 3 i 4 grupy są niezależne od danych i umożliwiają składanie wielu różnych operacji na różnych zbiorach danych. Najbardziej elastyczna jest metoda 4, gdyż dane przestrzenne najczęściej są rozmieszczone nieregularnie.

Część metod może być stosowana lub jest efektywna tylko dla określonych typów geometrycznych. Dla danych punktowych są przeznaczone metody grid file i kd-tree [10], dla wielokątów metody oparte na R-drzewie. Odcinki można transformować do przestrzeni o dwukrotnie wyższym wymiarze i stosować metody dla punktów lub w przestrzeni obiektów stosować metody dla wielokątów. Metody oparte na drzewach czwórkowych mogą być stosowane dla wszystkich typów danych.

Wykorzystanie metod stosowanych w konwencjonalnych bazach danych do obiektów przestrzennych opiera się na ich przekształceniu do punktu wielowymiarowego (tzw. punktu reprezentatywnego). Wymiar przestrzeni jest najczęściej za wysoki. Jednym z rozwiązań tego problemu jest aproksymacja obiektu, tak aby zredukować wymiar punktu reprezentatywnego. Na tablicach zawierających przetransformowane dane w charakterze indeksów przestrzennych wykorzystywane są indeksy dla zapytań wieloatrybutowych, dostępne w tradycyjnych bazach danych.

Podejście oparte na transformacji danych do reprezentacji wielokątów polega na transformacji minimalnych prostokątów ograniczających obiekty przestrzenne w punkty w przestrzeni o wyższym wymiarze. Na przykład 2-wymiarowy prostokąt o osiach równoległych do osi układu współrzędnych jest reprezentowany za pomocą punktu 4-wymiarowego, którego współrzędne stanowią współrzędne środka prostokąta oraz odległości boków prostokąta od jego środka (reprezentacja środkowa) lub współrzędne lewego dolnego oraz prawego górnego wierzchołka (reprezentacja brzegowa).

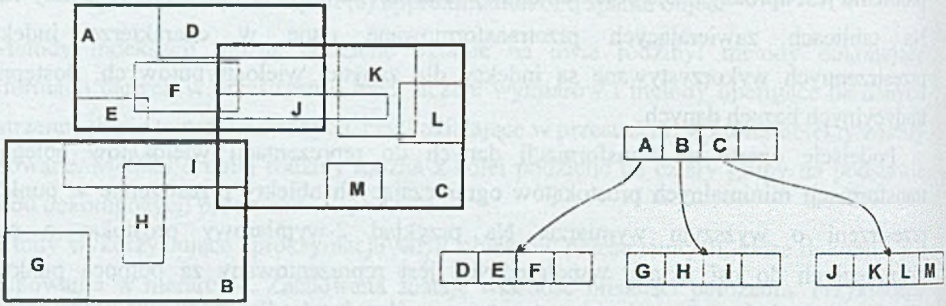
Struktura typu grid file dokonuje podziału przestrzeni na komórki za pomocą nieregularnie rozmieszczonych linii rozciągających się na całą przestrzeń. Wszystkie punkty leżące w obrębie komórki są zapisane w odpowiadającej jej komorze. Wiele komórek może być odwzorowanych w tę samą komorę.

3. Indeksacja przy użyciu R-drzewa

Zarządzanie strukturami danych dla prostokąta jest trudniejsze niż dla punktów, gdyż w ogólnym przypadku prostokąty nie mieszczą się w obrębie jednej komórki podziału przestrzeni danych, lecz przecinają granice podziału. Zasadniczo wyróżnia się trzy podejścia do budowy indeksu przestrzennego opartego na aproksymacji obiektów przestrzennych za pomocą minimalnych prostokątów ograniczających (MBR):

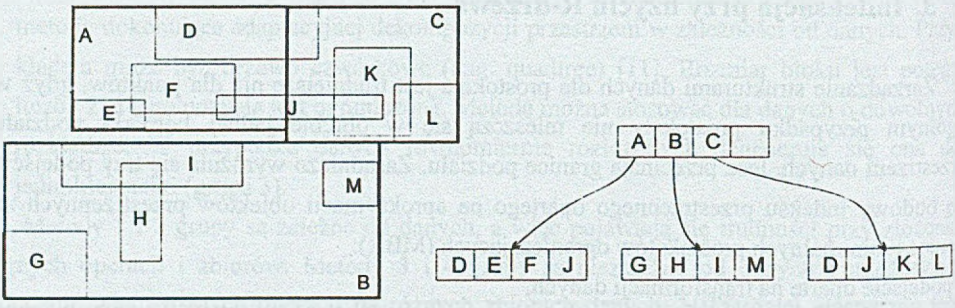
- podejście oparte na transformacji danych,
- podejście z nakładającymi się obszarami,
- podejście z obcinaniem.

W podejściu z nakładającymi się obszarami (ang. overlapping regions) nie dokonuje się podziału przestrzeni danych, lecz grupuje się obiekty tworząc hierarchie komór. Komory na tym samym poziomie mogą na siebie zachodzić. Przykładem realizacji tego podejścia jest struktura typu R-drzewo [4]. Jest to podobnie jak B-drzewo drzewo wielokierunkowe. W każdym jego węźle zapisany jest zbiór prostokątów. W liściach zapisane są MBR obiektów, które mają być reprezentowane. Węzły pośrednie i korzeń powstają poprzez zgrupowanie węzłów na niższym poziomie. Każdy taki węzeł posiada ciąg elementów postaci: wskaźnik do potomka p i prostokąt będący minimalnym prostokątem ograniczającym dla wszystkich prostokątów stanowiących wejścia w węźle-potomku wskazywanym przez p.



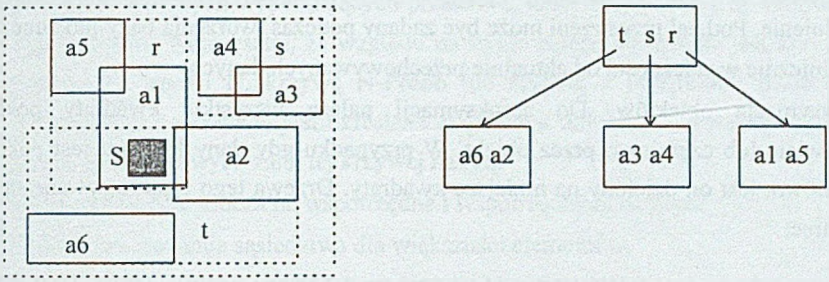
Rys. 2. Zbiór prostokątów oraz jego reprezentacja w postaci R-drzewa
Fig. 2. A set of rectangles and its R-tree representation

Na rys. 2 pokazany został przykład R-drzewa, którego korzeń zawiera prostokąt A, B, i C. Prostokąt A jest prostokątem ograniczającym prostokąty D, E i F zapisane w jego węźle potomnym. Komory mogą zachodzić na siebie. Dany prostokąt może przecinać wiele granic komór, ale będzie reprezentowany tylko w jednej z nich. Problemem staje się wyszukiwanie, które musi odbywać się wieloma ścieżkami, w przypadkach gdy interesujący nas prostokąt leży na przecięciu się obszarów wielu komór.



Rys. 3. Zbiór prostokątów oraz jego reprezentacja w postaci R+-drzewa
Fig. 3.A set of rectangles and its R+-tree representation

W podejściu z obcinaniem (ang. clipping) obszary komór nie mogą się nakładać. Jeśli prostokąt przecina granice podziału, to zostaje podzielony na kilka części, zapisywanych w komorach na obszarach, w których leżą. Przykładem może być struktura typu R+-drzewo [12]. Na rys. 3 pokazano R+-drzewo dla tego samego zbioru prostokątów jak na rys. 2. Prostokąty A, B i C w korzeniu zostały wybrane tak, aby obszary komór w potomku były rozłączne. Wadą podejścia jest konieczność zapisu podzielonych prostokątów w różnych komorach.



Rys. 4. Przykład R-drzewa z zaznaczonym oknem zapytania S
Fig. 4. An example of a R-tree with a query window S

Realizacja zapytania typu okno (z oknem S) rozpoczyna się w korzeniu drzewa, gdzie obliczane są wszystkie wejścia, które przecinają się z prostokątem S. Dla tych wejść pobierane są do pamięci operacyjnej węzły potomne i realizacja zapytania (jeśli węzeł nie jest liściem) przebiega dalej jak dla korzenia. Rys. 4 pokazuje R-drzewo składające się z trzech stron danych i jednej strony katalogu (jeden węzeł drzewa odpowiada jednej stronie w pamięci zewnętrznej). Prostokąty strony katalogu stanowią minimalne prostokąty ograniczające (MBR) prostokątów, które są zapisane w odpowiadającym węźle-potomku. Okno zapytania jest wyróżnione szarym kolorem. Najpierw zapytanie jest wykonywane na korzeniu, w którym prostokąty t i r przecinają się z prostokątem zapytań S. Odpowiednie dwie strony są wczytywane do pamięci i ich wejścia są sprawdzane na przecięcie się z prostokątem zapytań S. W końcu znaleziona zostaje odpowiedź zawierająca prostokąt a1 jako jedyny przecinający okno zapytania.

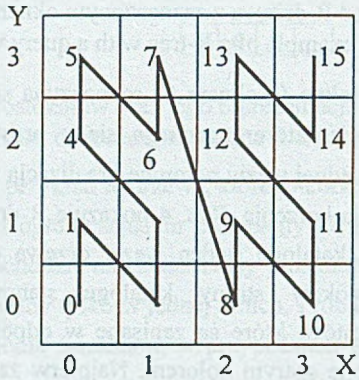
4. Drzewo czwórkowe

Drzewo czwórkowe (ang. quadtree) jest strukturą danych klasy drzewo, powstałą poprzez rekurencyjny podział przestrzeni 2-wymiarowej. Podstawowym elementem struktury jest kwadrat o boku będącym potęgą liczby 2. Przy zwiększaniu rozdzielczości kwadrat jest

dzielony na cztery jednakowe kwadraty. Proces podziałów kończy się gdy aproksymowana cecha znajduje się na obszarze całego kwadratu lub gdy osiągnięto założoną granicę rozdzielczości.

Drzewo czwórkowe można wykorzystać do przyśpieszenia dostępu do danych przestrzennych na dwa sposoby:

- podział przestrzeni, w której znajdują się obiekty. W geograficznych systemach informacyjnych na ogół mamy do czynienia z nierównomierną gęstością rozkładu obiektów. Zastosowanie komórek o różnych rozmiarach pozwala nam zrównoważyć ich zapełnienie. Podział przestrzeni może być zadany podczas tworzenia bazy lub zmieniać się dynamicznie w zależności od aktualnie przechowywanych danych,
- aproksymacja obiektów. Do aproksymacji należą wszystkie kwadraty pokrywane całkowicie lub częściowo przez obiekt. W przypadku gdy dany kwadrat jest pokrywany częściowo, jest on dzielony na mniejsze kwadraty. Drzewa tego typu nazywane są region quadtree.



Rys. 5. Przykład krzywej N-Peano
Fig. 5. An example of a N-Peano curve

W pierwszym przypadku drzewo ma znacznie mniejszą rozdzielczość niż w drugim. W dwuetapowym przetwarzaniu zapytań prowadzi to do powstania większego zbioru obiektów, które trzeba przetworzyć na drugim etapie, zmniejszając jednak liczbę wierszy przeszukiwanych na pierwszym etapie.

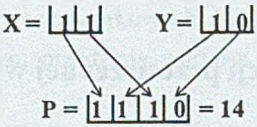
Drzewa czwórkowe łatwo jest uogólnić na przestrzeń 3-wymiarową, gdzie kwadrat jest zastępowany sześcianem, który dzieli się na 8 mniejszych sześcianów. Otrzymujemy w ten sposób drzewo ósemkowe. Można również reprezentować dane w przestrzeni N-wymiarowej, używając hipersześcianów N-wymiarowych.

Przechowanie drzewa czwórkowego w przedstawionej postaci wymagałoby wprowadzenia specjalnej struktury danych, uniemożliwiając ich wykorzystanie w klasycznych bazach danych. Istnieje jednak możliwość liniowego uporządkowania kwadratów tworzących drzewo czwórkowe za pomocą krzywej wypełniającej przestrzeń. Drzewo takie nazywane jest liniowym drzewem czwórkowym (ang. linear quadtree). Taką reprezentację drzewa można efektywnie zapamiętać w relacyjnej bazie danych i wykorzystać dla niej tradycyjne mechanizmy indeksacji.

Istnieje wiele krzywych wypełniających przestrzeń, które mogą służyć do liniowego uporządkowania drzewa. Najczęściej, ze względu na swoje zalety, stosowana jest krzywa nazywana krzywą Z, krzywą N, krzywą N-Peano lub krzywą z przeplotem bitów (rys. 5). Położenie kwadratu na krzywej jest określane za pomocą indeksu, nazywanego kluczem N-Peano [7]. Do podstawowych zalet tej krzywej należą:

- proste odwzorowanie klucza na współrzędne i współrzędnych na klucz,
- zapewnienie zachowania sąsiedztwa dla większości elementów,
- możliwość uogólnienia na przypadek przestrzeni N-wymiarowej.

Klucz N-Peano tworzony jest poprzez przeplot bitów współrzędnych, tak że parzyste bity klucza N-Peano odpowiadają współrzędnej X, a nieparzyste współrzędnej Y. Na rys. 6 przedstawiono przykład tworzenia klucza dla komórki o współrzędnej x = 3 i współrzędnej y = 2, której to komórce odpowiada klucz 14. Tworzenie kluczy w podany sposób wymaga określania przed budową drzewa maksymalnej rozdzielczości. Innym rozwiązaniem jest dodawanie do klucza kolejnych elemetów na kolejnych poziomach podziału, tak że długość klucza odpowiada poziomowi w drzewie. Rozwiązanie takie zastosowano systemie Oracle i jest ono opisane w następnym rozdziale.

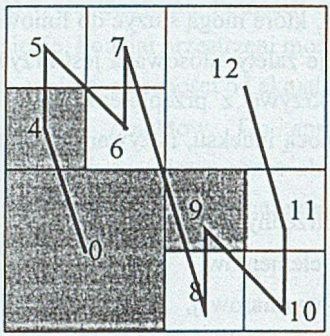


Rys. 6. Związek miedzy kluczem N-Peano i współrzędnymi X, Y
Fig. 6. The relation between the N-Peano key and the X, Y coordinates

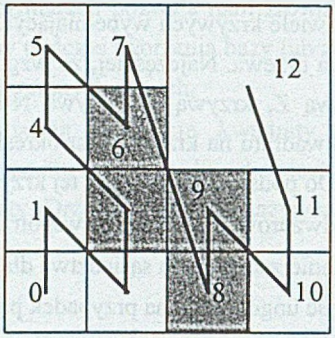
Rys. 7 przedstawia aproksymację dwóch obiektów liniowymi drzewami czwórkowymi. Drzewa takie można zapisać w relacji Peano [6, 7], której krotki opisują kwadraty tworzące drzewo czwórkowe. Schemat relacji Peano ma postać *(IdObiektu, KluczPeano, DługośćBoku)*, gdzie *IdObiektu* identyfikuje opisywany obiekt, *KluczPeano* określa położenie kwadratu na płaszczyźnie, a *DługośćBoku* podaje długość boku zapamiętanego kwadratu.

Na aproksymacjach obiektów zapisanych za pomocą liniowych drzew czwórkowych można efektywnie wykonywać zapytania przestrzenne bez użycia geometrii obliczeniowej

wykorzystując jedynie standardowy SQL. Wykonanie przestrzennego złączenia sprowadza się do znalezienia obiektów aproksymowanych przez kwadraty o tych samych kluczach, a następnie sprawdzenia przy użyciu metod geometrii obliczeniowej, czy obiekty rzeczywiście spełniają zadane kryterium. Ponieważ nie wszystkie kwadraty występują w sposób jawny, należy badać pokrywanie się odcinków na N krzywej. Dla przykładowych obiektów z rys. 7 częścią wspólną są kwadraty elementarne o kluczach 3 i 9. Do głównych wad tego podejścia należy wrażliwość na obrót i wybór początku układu współrzędnych.



Id obiektu	Klucz N	Dł. boku
A	0	2
A	4	1
A	9	1



Id obiektu	Klucz N	Dł. Boku
B	3	1
B	6	1
B	8	1
B	9	1

Rys. 7. Aproksymacja przykładowych obiektów przez liniowe drzewa czwórkowe oraz ich zapis w relacji Peano

Fig. 7. An exemplary approximation of an objects by linear quadtrees and their representation in Peano relation

5. Wykorzystania indeksacji przestrzennej w systemie Oracle

Idee opisane w punkcie poświęconym drzewom czwórkowym i krzywym wypełniającym przestrzeń zostały zastosowane w module Oracle Spatial Data Option (SDO), dodającym do relacyjnego serwera Oracle możliwość prostego i efektywnego gromadzenia i przetwarzania danych przestrzennych. Przedstawiony opis dotyczy wersji 7.3. Pakiet SDO zawiera dwa niezależne mechanizmy. Pierwszy służy do gromadzenia, indeksowania i przetwarzania dowolnych danych przestrzennych. Drugi nastawiony jest na optymalizację obsługi bardzo dużych zbiorów danych punktowych, poprzez przechowywanie ich w tablicy podzielonej na fragmenty, które przechowują dane pogrupowane według ich położenia w przestrzeni.

Model danych przestrzennych w systemie Oracle jest zorganizowany hierarchicznie. Podstawowym „budulcem" bardziej skomplikowanych struktur są elementy (ang. element) jednego z trzech podstawowych typów: punkt, linia łamana i wielokąt. Budowane z nich struktury są nazywane obiektami geometrycznymi (ang. geometric object) lub geometriami (ang. geometry). Złożone obiekty, takie jak wielokąty z dziurami, są przechowywane jako sekwencje elementów typu wielokąt. Kolekcja obiektów mających taki sam zbiór atrybutów jest nazywana warstwą (ang. layer).

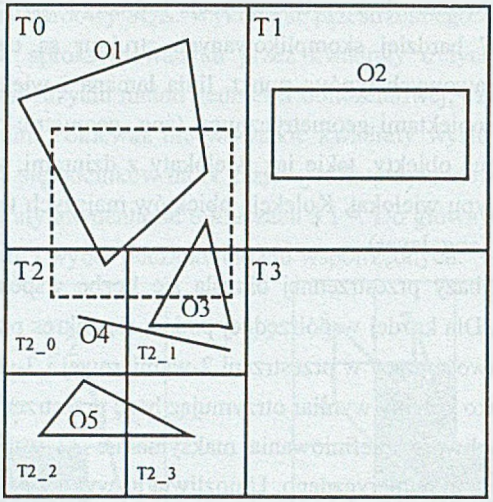
Podczas tworzenia bazy przestrzennej określa się liczbę współrzędnych służących do identyfikacji obiektów. Dla każdej współrzędnej podaje się zakres przyjmowanych przez nią wartości. Istnieje możliwość pracy w przestrzeni 2-wymiarowej i 3-wymiarowej. Dodatkowo można dołączyć czas jako kolejny wymiar otrzymując bazę przestrzenno-temporalną. Ogólnie w modelu istnieje możliwość zdefiniowania maksymalnie 32 współrzędnych, przy czym każda z nich musi być typu numerycznego. Umożliwia to wykorzystanie Oracle Spatial Data Option do gromadzenia i przetwarzania danych wielowymiarowych.

W pakiecie SDO dodany został nowy typ o nazwie HHCODE. Koduje on wiele wymiarów w jedną unikalną wartość. Kolumna typu HHCODE odpowiada kluczowi N-Peano, kodowanemu przy użyciu przyrostków - każdy kolejny podział dodaje na końcu symbol określający lokalizację kwadratu. Posortowanie według HHCODE zapewnia, że dane są pogrupowane według ich położenia w przestrzeni. Typ HHCODE jest on przechowywany w bazie jako RAW(255).

HHCODE reprezentuje całkowity wzajemnie rozłączny podział na komórki nazywane płytkami (ang. tiles), które mogą być jednakowego lub różnego rozmiaru, wynikającego z rekurencyjnego podziału płytki na cztery mniejsze takiej samej wielkości Ze względu na różne zakresy wartości płytki nie muszą być kwadratami, co nie zmienia faktu, że stosują się do nich wszystkie algorytmy drzew czwórkowych. System zarządza specjalnym słownikiem opcji przestrzennej, zawierającym przedziały wartości wymiarów reprezentowanych przez każdą z płytek.

Punk n-wymiarowy jest reprezentowany jako N-wymiarowy HHCODE. Odcinek jest kodowany jako punkt początkowy i końcowy, tak więc odcinkowi na płaszczyźnie odpowiadają 4 wymiary w HHCODE.

Na rys. 8 przedstawiono przykładową przestrzeń z 5 obiektami (O1-05) oraz okno zapytania (zaznaczone linia przerywaną). Przestrzeń jest podzielona na płytki różnych rozmiarów. Kolejny poziom podziału dodaje przyrostek do kodu nadrzędnej płytki, określający położenie płytki w obszarze płytki wyższego poziomu.



Rys. 8. Przykład podziału warstwy w przestrzennym rozszerzeniu bazy Oracle
Fig. 8. An example of layer tessellation in Oracle Spatial Data Option

Tabela 1 przedstawia tablicę systemu Oracle o nazwie *PRZYKLAD_SDOINDEX* zawierającą indeks przestrzenny dla obiektów z rys. 8. Kolumna *SDO_GID* zawiera identyfikator obiektu, kolumna *SDO_CODE* kod płytki, na której leży obiekt. Na kolumnach tych jest zakładany klasyczny indeks dostępny w relacyjnym serwerze.

Tabela 1

PRZYKLAD_SDOINDEX zawierająca indeks przestrzenny dla obiektów z rys. 8

SDO_GID	SDO_CODE	SDO_MAXCODE
O1	T0	T0FF...F
O1	T2_0	T2_0FF...F
O2	T1	T1FF...F
O3	T0	T0FF...F
O3	T2_1	T2_1FF...F
O4	T2_0	T2_0FF...F
O4	T2_1	T2_1FF...F
O5	T2_2	T2_2FF...F
O5	T2_3	T2_3FF...F

Wykonanie zapytania znajdującego obiekty przecinające się z oknem realizowane jest w następujący sposób. Na podstawie słownika opcji przestrzennej znajdowane są płytki przecinające się z oknem zapytania. Dla przykładu są to płytki o kodach T0, T2_0 i T2_1. Następnie znajdowane są obiekty leżące na tych płytkach (O1, O3 i O4), co kończy pierwszy etap przetwarzania zapytania. W drugim etapie przetwarzania zapytania na podstawie dokładnych współrzędnych obiektów przy użyciu geometrii obliczeniowej wyznaczane jest rozwiązanie (O1 i O3).

Mechanizm tworzenia partycji danych punktowych znajduje zastosowanie w bazach danych, przechowujących bardzo duże ilości danych punktowych o postaci identyfikator obiektu, współrzędne określające jego położenie w przestrzeni oraz atrybuty opisowe. Przykładem może być baza banku, w której przechowywane są dane osobowe klienta, przybliżone położenie i zadeklarowane dochody. Zapytanie do niej kierowane może wyglądać następująco: „Podaj średnie dochody klientów banku zamieszkujących obszary otaczające specjalne strefy ekonomiczne". Zapytanie takie sprowadza się do wybrania z bazy danych leżących na zadanym obszarze i wykonaniu na nich zadanych obliczeń.

Rozwiązanie firmy Oracle opiera się na rekurencyjnym podziale przestrzeni i kodowaniu współrzędnych punktów przy użyciu atrybutu typu HHCODE, który jest kluczem tablicy. Tablica jest dzielona według tego klucza, tak że w danym fragmencie znajdują się tylko punkty znajdujące się w tym samym rejonie przestrzeni. Administrator ustala liczbę wierszy, przy której partycja będzie dzielona.

6. Podsumowanie

Aby móc efektywnie realizować zapytania przestrzenne, obiekty muszą zostać posortowane w oparciu o atrybut wyrażający ich położenie w przestrzeni. W pracy przedstawiono przegląd metod indeksacji wykorzystywanych w przestrzennych bazach danych. Metody indeksacji stosowane do danych nie transferowanych zostały podzielone na cztery grupy na podstawie metody dekompozycji przestrzeni, w której znajdują się obiekty. Omówiono dwie najczęściej wykorzystywane struktury indeksów przestrzennych R-drzewa i drzewa czwórkowe. Dla każdej z nich przedstawiono podstawowe algorytmy oraz najczęściej spotykane odmiany. Omówiono krzywą z przeplotem bitów wypełniającą przestrzeń oraz jej wykorzystanie do zapisu drzewa czwórkowego. Powstałe w ten sposób liniowe drzewo czwórkowe można zapisać w relacyjnej bazie danych wykorzystując tradycyjne metody indeksacji do efektywnego przetwarzania zapytań przestrzennych.

R-drzewa są wykorzystywane w programie ArcInfo, będącym jednym z liderów rynku oprogramowania narzędziowego dla geograficznych systemów informacyjnych. Podejście oparte na liniowym drzewie czwórkowym zostało wykorzystane przez firmę Oracle do rozszerzenia Uniwersalnego Serwera o możliwość przechowywania i przetwarzania danych przestrzennych (Oracle Spatial Data Option).

Ciekawe wydaje się rozszerzenie serwera Oracle, będącego nie kwestionowanym liderem na rynku relacyjnych baz danych o mechanizmy poprawiające efektywność przetwarzania danych przestrzennych. Umożliwia to programistom włączanie do tworzonych w tym systemie aplikacji modułów prezentujących rozmieszczenie obiektów w przestrzeni oraz

wykonujących analizy według kryterium zawierających warunki na położenie obiektów. Podejście takie jest znacznie tańsze niż kupowanie narzędzi do tworzenia geograficznego systemu informacyjnego i integrowanie stworzonego produktu z resztą aplikacji tworzonej w systemie Oracle. Z drugiej strony oprogramowanie specjalizowane dostarcza wielu narzędzi, których pisanie na potrzeby jednej aplikacji może się okazać nieefektywne.

LITERATURA

[1] Aboulnaga A.: An Introduction to Spatial Access Methods, Proceedings of the Fourth IASTED Intern. Conference on Computer Applications in Industry. Kair 4-6 grudnia 1995r.

[2] Brinhhoff T., Kriegel H., Seeger B.: Efficient Processing of Spatial Joins Using R-trees. ACM SIGMOD, 1993.

[3] Greene D.: An implementation and performance analysis of spatial data access methods. Proc. 5th Int. Conf. on Data Engineering, Los Angeles, 1989.

[4] Guttman A.:R-trees: a dynamic index structure for spatial searching. ACM SIGMOD June, 1984.

[5] Güting R.H.: An introduction to Spatial Database Systems. Special issue of spatial database systems of VLDB , Journal No 4 October, 1994.

[6] Laurini R., Francoise M.: Spatial database queries: relational algebra versus computational geometry. Proceedings of the Fourth International Conference on Statistical and Scientific Database Management, Rome, Italy 1988, M. Rafamelli et al.,(eds) Berlin;Germeny: Springer Verlag. pp. 291-313.

[7] Laurini R., Thompson D.: Understanding GIS, Academic Press Limited, third printing 1994.

[8] Oracle7 Spatial Data Option™ Overiew: Oracle April 1996.

[9] Oracle7 Spatial Data Option™ Application Developer's Guide: Oracle 1996.

[10] Samet H.: The design and Analysis of Spatial Data Structures, Addison-Wesley, 1990.

[11] Samet H.: The Quad-tree and related hierarchical data structures ACM Computing Surveys, 1984, 16(2).

[12] Sellis T., Roussopoulos N., Faloutsos C.: The R*-tree: a dynamic index for multi-dimensional objects. Proceedings of the 13th VLDB Conferenc,e September, 1987.

Recenzent: Dr inż. Andrzej Krupiczka

Wpłynęło do Redakcji 23 grudnia 1997 r.

Abstract

Effective query execution against spatial databases demands for indexing methods based on objects location in space. This article describes two data structures most often used for this purpose: R-trees i quadtrees. Some variations of these two structures and algorithms using them for spatial query execution are presented. The space filling curve based on bit interleaving and its usage for ordering the square cells of a quadtree is described. A quadtree represented in such way is called linear quadtree. Oracle Spatial Data Option based on linear quadtree approach is described. The SDO enables spatial data to be stored, accessed and analyzed in the Oracle Universal Server.