



STRUKTURY DRZEWIASTE W GEOMETRII OBLICZENIOWEJ

Rafał Piwowar
Jan Smółka



Motywacja

- Efektywne przechowywanie informacji o położeniu punktów na płaszczyźnie
- Szybkie wyszukiwanie punktów należących do ograniczonych obszarów

Rozwiązanie

- Wprowadzenie struktur analogicznych do drzew wyszukiwania binarnego

Oczekiwany rezultat

- Informacja o lokalizacji punktów w prostokątnych podobszarach płaszczyzny dostępna w czasie logarytmicznym

Propozycja – k-d drzewo

- Struktura danych oparta na rekurencyjnym podziale k -wymiarowej przestrzeni na dwie podprzestrzenie (tzw. „półprzestrzenie”) oddzielone $(k-1)$ -wymiarową hiperpłaszczyzną
- W każdym węźle przechowywane są punkty należące do danej półprzestrzeni oraz odesłania do dwóch podziałów tejże.
- Każda półprzestrzeń reprezentowana jest przez hiperpłaszczyznę podziału, poprzez wskazanie jej wektora normalnego. Zwykle są to kolejne wektory bazowe w danej przestrzeni, przy ustalonej bazie.
- W przypadku dwuwymiarowym informacja ta sprowadza się do wskazania osi (X lub Y), wzdłuż której przebiega podział w danym węźle.
- Określenie położenia hiperpłaszczyzny podziału, poza wektorem normalnym wymaga również podania punktu. Zwykle jest to mediana dzielonego zbioru, liczona w kierunku normalnym danego podziału.

Propozycja – drzewo czwórkowe

- Struktura danych oparta na rekurencyjnym podziale dwuwymiarowej przestrzeni na cztery podprzestrzenie.
- W każdym węźle przechowywane są punkty należące do danego obszaru przestrzeni oraz odesłania do jego podziałów. Dodatkowo, dla ułatwionego wyszukiwania, warto przechowywać w węzłach informacje o rozmiarach obszarów.
- Każdy obszar reprezentowany jest przez środkowy punkt podziału. Jest to jednoznaczna definicja podziału, ponieważ kolejne „ćwiartki” składają się z punktów, których każda współrzędna pozostaje w odpowiedniej relacji nierówności z punktem podziału.
- Zrównoważony podział przestrzeni zachodzi przy wyborze jako punktu podziału środka ciężkości każdego obszaru.

Wyszukiwanie

- Przy wykorzystaniu zaproponowanych struktur, możliwe jest wyszukiwanie punktów należących do obszarów normalnych w danej przestrzeni, zdefiniowanych poprzez podanie górnych i dolnych ograniczeń dla współrzędnych.
- W obu przypadkach wyszukiwanie odbywa się zgodnie z porządkiem przeglądania DFS, analogicznie jak ma to miejsce w drzewie przedziałowym czy drzewie wyszukiwania binarnego.
- Do węzła drzewa kierowane jest zapytanie o punkty spełniające nierówności. W węźle wybierane są podziały spełniające zapytanie. W wyszukiwaniu kumulowane są punkty z obszarów spełniających zapytanie. Jeśli jakiś obszar w całości spełnia zapytanie, jest dodawany do wyniku bez analizowania dalszych jego podziałów.
- W przypadku obu struktur, wyszukiwanie punktów ma złożoność obliczeniową, przy założeniu o zrównoważeniu drzewa, $O(\log n)$, gdzie n jest licznnością przeszukiwanego zbioru.

Porównanie struktur

K-d drzewo

- Podział na dwa obszary
- Złożoność pamięciowa $O(n)$
- Średni czas wyszukiwania $O(\log n)$
- Łatwo rozszerzalne na przestrzenie o większej liczbie wymiarów, naturalne ze względu na konstrukcję drzewa
- Gorsze działanie przy większej liczbie wymiarów

Drzewo czwórkowe

- Podział na cztery obszary
- Złożoność pamięciowa $O(n)$
- Średni czas wyszukiwania $O(\log n)$
- Stosunkowo trudna rozszerzalność – dla przestrzeni trójwymiarowej mamy drzewo ósemkowe
- Lepsze działanie przy większej liczbie wymiarów

Bibliografia

- <https://en.wikipedia.org/wiki/Quadtree>
- https://en.wikipedia.org/wiki/K-d_tree