

Wyrocznie odległości (distance oracles)

Juliusz Kopczewski

18 marca 2010

1 Wstęp

Wyrocznie odległości (ang. distance oracles) to struktury danych potrafiące obliczać przybliżoną odległość między zadanymi wierzchołkami $u, v \in G$ w czasie podliniowym. Przy czym ich wielkość względem danych wejściowych jest rzędu $o(n^2)$. To eliminuje zwykle rozwiązanie typu All-Pairs-Shortest-Paths, którego wynikiem jest macierz wymiaru $O(n^2)$. Fazę budowania wyroczni nazywa się *preprocessingiem*.

2 Grafy skierowane, planarne

- Wejście algorytmu: graf planarny, **skierowany**.
- Współczynnik aproksymacji: $1 + \epsilon$, czyli $OPT \leq \delta(u, v) \leq (1 + \epsilon) \cdot OPT$.
- **Preprocessing**. Czas: $O(n(\log n)^3(\log(nN))/\epsilon^2)$, pamięć: $O(n(\log n)(\log(nN))/\epsilon)$, gdzie n - l. wierzchołków, N - maks. waga.
- **Zapytania**. Czas: $O(\log \log(nN) + 1/\epsilon)$.

W pierwszej kolejności zajmiemy się problemem osiągalności w grafach skierowanych. Konstrukcje, które opiszemy przy tej okazji okażą się pomocne również dla problemu znajdowania odległości.

Pierwszy krok algorytmu będzie polegał na wyznaczeniu ciągu grafów G_i o łącznym rozmiarze liniowym względem rozmiaru G , powstałych przez usuwanie wierzchołków oraz usuwanie i ściągnięcie części krawędzi. Otrzymane G_i będą tzw. grafami 2-warstwowymi. W dodatku dla każdego u będzie można wyznaczyć i , t.ż. $u \rightsquigarrow_G v$ wtw. gdy $u \rightsquigarrow_{G_i} v$.

Definicja. Graf skierowany G jest t -warstwowy wtw. gdy posiada t -warstwowe drzewo rozpinające T , czyli takie drzewo rozpinające nieskierowany graf indukowany przez G , że dowolna ścieżka w T rozkłada się na co najwyżej t skierowanych ścieżek w G .

Następnie pokażemy, że w grafie 2-warstwowym G można znaleźć w czasie liniowym **separator** S , t.ż. każda składowa $G-S$ (w sensie nieskierowanym) jest co najmniej 2-krotnie mniejsza niż G w sensie sumy wag. Pokażemy również jak

w wydajny sposób reprezentować informacje o osiągalności $u \rightsquigarrow v$, gdy wiadomo, że ścieżki te przecinają pewną ścieżkę skierowaną P . Obie te własności pozwolą nam na zastosowanie podejścia “dziel i rządź”. Rozstrzygnięcie osiągalności $u \rightsquigarrow v$ będzie polegało na znalezieniu separatora S oddzielającego u od v . Następnie dla każdej ścieżki $P \subseteq S$ stwierdzimy, czy istnieje $u \rightsquigarrow v$ przechodzące przez P .

W ten sposób otrzymamy algorytm rozstrzygający osiągalność w czasie $O(\log n)$. Następnie przedstawimy szereg usprawnień, które pozwolą nam zbić złożoność do $O(1)$. Usprawnienia będą polegały na sprytniejszej konstrukcji separatorów, tak aby były one zawsze stałej wielkości.

Bardzo zbliżone podejście zastosujemy do ogólniejszego problemu znajdowania odległości w grafie skierowanym, planarnym. Tym razem zastosujemy redukcję do grafów $(3, \alpha)$ -warstwowych.

Definicja. Graf skierowany G jest (t, α) -warstwowy wtw. gdy posiada (t, α) -warstwowe drzewo rozpinające T , czyli takie drzewo rozpinające nieskierowany graf indukowany przez G , że dowolna ścieżka dł. $\leq \alpha$ w T rozkłada się na co najwyżej t skierowanych ścieżek w G .

Podejście polegające na szukaniu separatorów zasadniczo nie ulegnie zmianie. Mając separator S oddzielający u od v oraz dowolne $P \subseteq S$ będziemy znajdowali $\text{dist}_P(u, v)$ - długość najkrótszej ścieżki między u i v z błędem addytywnym $\epsilon\alpha$. Tym samym otrzymamy konstrukcję wyroczni zwracającej wyniki z takim właśnie błędem.

W celu otrzymania błędu względnego $1 + \epsilon$, skonstruujemy ciąg wyroczni z błędem addytywnym dla wykładniczo rosnących α .

3 Grafy nieskierowane, dowolne

- Wejście algorytmu: graf nieskierowany
- Współczynnik aproksymacji: $2k - 1$.
- **Preprocessing.** Oczekiwany czas: $O(kmn^{1/k})$, pamięć: $O(kn^{1+1/k})$, gdzie n - l. wierzchołków, m - l. krawędzi.
- **Zapytania.** Czas: $O(k)$.

W pierwszej kolejności zajmiemy się uproszczonym zagadnieniem konstrukcji wyroczni dla jawnie zadanej metryki. Oczywiście taki algorytm nie ma większego praktycznego zastosowania i polega jedynie na “kompresji” metryki.

Skonstruujemy wstępujący łańcuch A_i podzbiorów wierzchołków G o malejącej wykładniczo wielkości. Dla każdego wierzchołka będziemy przechowywali jego odległość od zbiorów A_i oraz specjalny zbiór $B(v)$. Dokładniej, jeśli $w \in A_i - A_{i+1}$ to $w \in B(v)$ wtw. gdy $\delta(v, w) < \delta(v, A_{i+1})$.

Okaże się, że na podstawie tych informacji przechowywanych w formie *hash-table*, będziemy w stanie wyznaczać odległość $\text{dist}(u, v)$ z zadaniem współczynnikiem aproksymacji.

Następnie przeniesiemy konstrukcję na sytuację, w której metryka nie jest zadana jawnie, a jedynie indukowana przez pewien graf nieskierowany G .