Vysoké učení technické v Brně

Fakulta informačních technologií

Minimálna kostra grafu

Dokumentácia náhradného projektu IAL – zadanie č. 7

Autori :

xpospi95 - Michal Pospíšil (vedúci)

xtimko01 - Nikola Timková

xbalaz08 - František Balázsy

Brno, 3.12.2018

Obsah

Úvod

Táto dokumentácia popisuje riešenie náhradného projektu – zadania č. 7 – minimálna kostra grafu. Aplikáciu, ktorá je hlavným bodom tohto zadania sme sa snažili vytvoriť tak, aby bola jednoducho použiteľná.

V ďalších častiach dokumentu budeme popisovať aké stratégie sme zvolili pri riešení projektu, a prečo sme sa pre ne rozhodli. Popíšeme aj na aké problémy sme pri riešení narazili.

Zadanie

- Kostra grafu je podgraf, ktorý je stromom, a ktorý obsahuje všetky vrcholy pôvodného grafu. Z tejto definície vyplývajú nasledujúce tvrdenia:
    - Kostra grafu je súvislý graf.
    - Kostra grafu má o jednu hranu menej než má vrcholov.
    - Nesúvislý graf nemá kostru.
    - Kostra grafu neobsahuje kružnicu.
- Vytvorte program pre nájdenie kostry grafu s minimálnym ohodnotením pre ohodnotené neorientované grafy.
- Ak existuje viac riešení, nájdite všetky. Výsledky prezentujte vhodným spôsobom. Súčasťou projektu bude načítanie grafov zo súboru a vhodné testovacie grafy. V dokumentácií uveďte teoretickú zložitosť úlohy a porovnajte ju s experimentálnymi výsledkami.

Výber vhodného algoritmu

Pre riešenie zadanej problematiky existuje rada algoritmov, preto sme museli vybrať najvhodnejší pre pokrytie všetkých požiadaviek vyplývajúcich zo zadania s prihliadnutím na ich technické parametre a jednoduchosť implementácie. Pre prehľad uvádzame krátky prehľad najznámejších algoritmov a zdôvodnením, prečo sme ich použili alebo nepoužili:

Borůvkov algoritmus

Pažravý algoritmus vynájdený pri riešení problematiky konštrukcie elektrickej siete na Morave. Značne obmedzujúci, pretože predpokladá grafy s hranami kladných a rôznych hodnôt. Môže však byť paralelizovaný, čo mu dáva výhodu oproti ostatným v prípade dostatku výpočtových jednotiek. Tvorí základ náhodnostného algoritmu bežiaceho v lineárnom čase, ktorý však funguje len pre grafy s neizolovanými vrcholmi.

Prečo sme nevybrali tento algoritmus...

Primov algoritmus

Pažravý algoritmus pre neorientované grafy vynájdený Vojtechom Jarníkom. Svoje najväčšie využitie nachádza v prípadoch veľmi hustých grafov s veľa vrcholmi, kde beží s O(E + V log V). Táto časová zložitosť však platí len v prípade, že spracováva vstupné dáta v podobe Fibonacciho haldy a zoznamu susednosti. Pokiaľ by sme algoritmus spustili nad maticou susednosti v najhoršom prípade by dosiahol časovú zložitosť $O(|V|^2)$.

Prečo sme nevybrali tento algoritmus...

Kruskalov algoritmus

Pažravý algoritmus vhodný na typické, redšie grafy. Vyznačuje sa veľkou implementačnou výhodou oproti Primovmu algoritmu pretože dokáže spracovať priamo maticu susednosti. Taktiež má nižšiu časovú zložitosť nad bežnými grafmi, a to O(E * log V).

Pre riešenie nášho zadania sme si vybrali Kruskalov algoritmus, pretože vyhovuje povahe našich vstupných dát, vyznačuje sa najjednoduchšou implementovateľnosťou a dosahuje najlepšiu časovú zložitosť nad bežnými grafmi. Taktiež je možné pomocou neho jednoducho nájsť viacero riešení.

Detaily implementácie

Reprezentácia grafu

Graf sme sa rozhodli reprezentovať maticou susednosti. Naše rozhodnutie ovplyvnilo, že jej veľkosť je ľahko vypočítateľná z počtu uzlov a na rozdiel od ukladanie samotných hrán sa nevyskytuje redundancia názvov uzlov ak vychádza viac hrán z jedného vrcholu. Navyše tento modul môže byť použitý s minimálnymi zmenami na čítanie súborov, ktoré by reprezentovali orientované grafy v inom programe.

Špecifikácia súborového formátu

Kódovanie súboru je UTF-8 alebo kódovania s ním kompatibilné (ASCII).

Formát súboru je veľmi jednoduchý. Pri jeho návrhu sme vychádzali z formátu CSV, ktorý je veľmi jednoduchý na spracovanie a kontrolu.

Riadok 1: Ako prvý reťazec musí súbor obsahovať reťazec ".xIALAdjMatrix". Nasledovať musí znak ukončenia riadka (EOL znak) v akomkoľvek štýle (LF/CRLF...). Tento reťazec umožňuje kontrolovať, že vstupom je naozaj súbor podporovaného typu.

Riadok 2: Počet uzlov zadaný číselne (číslicami 0-9), bez akýchkoľvek znamienok, ukončený EOL znakom. Maximálny počet uzlov je 255. Slúži na zjednodušenie implementácie syntaktickej kontroly matice susednosti - je ľahšie čítať zo vstupu neznáme dáta a overiť ich správnosť, keď vieme, koľko ich má byť.

Riadok 3: Názvy uzlov v takom poradí, v akom sú zapísané v matici vodorovne zľava doprava a zvislo zvrchu nadol. Názvy uzlov sú reťazce obsahujúce ľubovoľné znaky. Reťazce sú oddelené čiarkou. V prípade, že názov uzlu obsahuje čiarku, musí byť escapovaná znakom '\\' => (takto\,). To znamená, že aj samotný znak '\' musí byť escapovaný (takto: \\niečoZaLomítkom). Riadok je ukončený EOL znakom.

Ďalšie riadky (max. 255 - max. počet uzlov): Obsahujú maticou uložený graf - prvkami matice sú váhy hrán (min. $-(2^{64}/2)$ max. $2^{64}/2-1$) spájajúca vrchol reprezentovaný indexom v matici, ktorý je naviazaný na názvy uzlov. Ak uzly nie sú spojené, váha hrany je 0. Všetky váhy hrán sú oddelené čiarkami, nie sú tolerované žiadne biele znaky. Každý (aj posledný) riadok končí EOL znakom a súbor musí byť ukončený hneď za posledným EOL znakom..

Limity

Graf

Veľkosť grafu - max. 2^8-1 uzlov a $2^{16}-1$ hrán. Maximálna váha hrany je $2^{64}/2-1$, povolené sú aj záporné váhy (min. $-(2^{64}/2)$).

Vstupný súbor

Súbor môže mať akúkoľvek dĺžku. Dĺžka jedného názvu by nemala byť dlhšia ako $2^{16}/2$ (neošetrené). Pri načítaní súboru je program ošetrený proti pretečeniu, ktoré hrozí pri grafoch nespĺňajúcich špecifikáciu a proti syntaktickým a lexikálnym chybám v súbore.

Porovnanie teoretickej zložitosti úlohy s experimentom

Použitie programu

Preklad programu

Preklad prebieha pomocou nástroja make. Pre zostavenie finálneho programu stačí v hlavnom adresári projektu spustiť príkaz make bez parametrov, výsledkom je binárny súbor **sptree** (skratka zo spanning tree).

Spustenie programu

./sptree <filename>

./sptree -h

Program očakáva len jeden argument a tým je buď cesta k súboru <filename>, alebo prepínač -h, ktorý vypíše krátku nápovedu.

Zdroje