3. DÚ Z ÚVODU DO APROXIMACÍ

Hešování, dolní odhady, paralelizace

Každý příklad je za dva body. Termín: **15. 1. 2016 17:19.** Řešení odevzdávejte emailem nebo osobně – buď přímo na cvičení, nebo ho přineste do označené krabice ve 3. patře naproti kanceláři S323.

Pokud by zadání bylo nejasné nebo podle vás obsahovalo chybu, nebojte se mi napsat email.

PŘÍKLAD PRVNÍ   Vymyslete 2-aproximační algoritmus pro ROZLOŽENÍ ZÁTĚŽE V SÍTI SONET.

Síť se skládá z kružnice na n vrcholech. K síti máme seznam požadavků, kde každý požadavek má určen zdrojový vrchol a cílový vrchol. Ke každému požadavku musí být přiřazena právě jedna z dvou možných cest mezi zdrojem a cílem (po směru nebo proti směru hodinových ručiček).

Cílem je minimalizovat zátěž (počet použití) na nejvíce zatížené hraně.

PŘÍKLAD DRUHÝ   **Dolní odhad pro silné 2-univerzální hešovací funkce pomocí LA.** Silně 2-univerzální hešovací funkce jsou definovány z přednášky takto: Rodina funkcí $\mathcal{H}$ (kde všechny funkce jsou hešovací, čili $h\colon U \to HT$) je silně 2-univerzální, pokud pro libovolnou čtveřici proměnných $x, x' \neq x, y, y'$ platí:

$$P_{h \text{ náhodná z } \mathcal{H}}[h(x) = y \wedge h(x') = y'] = \frac{1}{|HT|^2}.$$

Na cvičení jsme si ukazovali, že pro každé $|U| = 2^n, |HT| = 2^m$ existuje slabě 2-univerzální rodina, ze které lze vybrat náhodnou funkci pomocí $O(m + n)$ náhodných bitů. Stejný fakt platí i pro silně 2-univerzální rodiny.

Naším cílem je ukázat asymptotickou těsnost daného odhadu; bude nám k tomu sloužit lineární algebra. Předpokládejme tedy, že máme libovolnou silně 2-univerzální hešovací rodinu $\mathcal{H}$.

1. Dokažte, že je-li $|U| \geq 2$, tak $|\mathcal{H}| \geq |HT|^2$.
2. Dokažte, že je-li $|HT| = 2$, tak $|\mathcal{H}| \geq |U| + 1$.

   Tip: Pomocí funkcí z $\mathcal{H}$ sestavte pro každé x jeden vektor v_x z $\mathbb{R}^{|\mathcal{H}|}$ se souřadnicemi jen ± 1 tak, aby každá dvojice v_x a $v_{x'}$ byla na sebe ortogonální. Co tato množina vektorů říká o dimenzi $\mathbb{R}^{|\mathcal{H}|}$? A proč je pravá strana $|U| + 1$ a ne jen $|U|$?
3. Zobecněte předchozí postup a dokažte, že pro obecnou velikost $|HT|$ platí

   $$|\mathcal{H}| \geq |U|(|HT| - 1) + 1.$$

   Tip: Pomocí funkcí z $\mathcal{H}$ sestavte pro každé x celkem $|HT| - 1$ vektorů $v_{x,y} \in \mathbb{R}^{\mathcal{H}}$ takových, že pro jedno x jsou všechny $v_{x,y}$ lineárně nezávislé, a pro dvě různá $x \neq x'$ jsou vektory $v_{x,y}$ a $v_{x',y'}$ na sebe ortogonální. Co tato množina vektorů říká o dimenzi $\mathbb{R}^{|\mathcal{H}|}$?
4. Z předchozích bodů odvoďte, že pokud $|U| = 2^n$ a $|HT| = 2^m$, tak pro vygenerování náhodné funkce z $\mathcal{H}$ je potřeba alespoň $(\max(n, m) + m)$ bitů, což je asymptotický těsný dolní odhad.

PŘÍKLAD TŘETÍ   Uvažme opět problém BALANCOVÁNÍ GRAFU.

Pro připomenutí: na vstupu je neorientovaný graf G s váhami na hranách $p\colon E(G) \to \mathbb{R}^+$. Cílem algoritmu je zorientovat všechny hrany tak, že ten nejtěžší vrchol (vrchol s největší vahou hran orientovaných k němu) má co nejmenší možnou váhu. Formálně je náš cíl najít přiřazení hran, které minimalizuje účelovou funkci $u = \max_{v \in V} \sum_{e \in E; e \text{ orientována do } v} p(e)$.

Dokažte, že žádný aproximační algoritmus nemůže dosáhnout aproximačního poměru menšího než 3/2.

Tip: Jistě je mnoho způsobů, jak dostat tento odhad, tady je jeden možný: můžete zkusit namodelovat jeden slavný NP-úplný problém pomocí grafu, který obsahuje jen hrany o váze 1 a 1/2. Pak zkuste odvodit, že pokud aproximační algoritmus na tomto grafu najde řešení s poměrem menším než 3/2, tak už musí jít o řešení NP-úplného problému.

Příklad čtvrtý Uvažme grafy s maximálním stupněm omezeným konstantou, čili $\Delta(G) \leq \Delta$. Hladový algoritmus takové grafy snadno nabarví $\Delta + 1$ barvami, ale není vůbec jasné, jak ho paralelizovat. Navrhněte a zanalyzujte rychlý paralelní pravděpodobnostní algoritmus, který graf s velkou pravděpodobností správně obarví $\Delta + 1$ barvami.

Tip: Může se hodit stavět na základech z prednášky.

Příklad pátý *Bonusový.* Ještě jednou uvážíme metrický Problém k dodavatelů.

Dokažte, že žádný aproximační algoritmus nemůže dosáhnout poměru menšího než 3.

Pro jistotu připomeneme: na vstupu máme číslo k a také na vstupu $m + n$ bodů, kde m z nich jsou předem označeni jako *dodavatelé* a zbylí jsou *zákazníci.* Mezi všemi těmito body je metrika (splňující trojúhelníkovou nerovnost). Úkolem v tomto problému je vybrat k dodavatelů tak, že minimalizujeme nejdelší vzdálenost mezi zákazníkem a jeho nejbližším vybraným dodavatelem.

Tip: V nějakou chvíli můžete využít, že k je součástí vstupu – takže lze vyrobit instanci a zvolit k chytře podle dané instance, aby pro toto k byla instance těžká (ale pro jiné by být nemusela).