

Posudek oponenta diplomové práce Pavly Marešové s názvem

Analýza sítí

Předložena diplomová práce se zabývá analýzou zranitelnosti silniční sítě, metodami, jak tuto zranitelnost kvantifikovat a algoritmicky nalézt a dále návrhem, jak přidávat hrany, aby zranitelnost co nejvíce klesla. Téma je velmi zajímavé, komplexní síť se poslední dobou ocitly ve středu pozornosti výzkumníků z mnoha různých oblastí a zřejmě o nich ještě mnohé uslyšíme. Problematika komplexních sítí se pohybuje na pomezí teorie grafů, statistické fyziky a informatiky, což jsou disciplíny, které na našem oboru příliš nepěstujeme, takže pro studenty není úplně jednoduché do nich proniknout.

Hlavním výsledkem diplomové práce je soubor programových kódů v MatLabu, které Pavla sama navrhla a implementovala, které jednak počítají zranitelnost silniční sítě a jednak síť optimalisují pomocí přidávání jedné nebo více hran tak, aby zranitelnost klesla. To je rozhodně velice dobrý a cenný výsledek. Algoritmy jsou v textu výborně popsány a je jim dobře rozumět. Navíc oceňuji, že práce je napsaná poměrně profesionálně (v první části možná až příliš didakticky), a tudíž se velmi dobře čte. K tématu ani k dosaženým výsledkům nemám žádné vážnější připomínky.

Přece jen je však několik bodů, na které bych se rád zeptal, případně upozornil:

1. Anglický abstrakt na straně 3 by nutně potřeboval korekturu. Jako experiment jsem zkusil, jestli by to google translate náhodou nezvládl lépe. Skončilo to remízou.
2. U povinného prohlášení chybí povinný podpis.
3. Smyslem úvodu a závěru diplomové práce je toto: Když si přečtu *jen* úvod a závěr, měl bych pochopit co je problém, co je o něm známo, proč je zajímavý, jak jej budeme řešit, proč zrovna takto, jak to řešení dopadlo, jak se tohle řešení srovnává s jinými a co by se dalo dělat dál. V průběhu času pozoruju, že úvody a závěry našich diplomových prací jsou čím dál kratší a méně informativní, tak si tímto začínám stěžovat.
4. Poslední odstavec na straně 16 je podivný. V konkrétních biologických sítích (food webs, gene regulation networks, metabolomic networks, ...) je samozřejmě vždy jasné dáno, co je uzel a co hrana, čímž se tedy biologické sítě nijak neliší od sítí v jiných disciplínách.
5. Perkolační práh je zapeklitý pojem. Já perkolačnímu prahu rozumím spíše jako fázovému přechodu, tedy takovému bodu na ose x , ve kterém sledovaná veličina (např. velikost největší komponenty) vykazuje nespojitost, když se velikost grafu blíží k nekonečnu. Vaše definice je poněkud vágní, protože bude-li graf dost velký, bod, od kterého se bude veličina škálovat *přesně* lineárně, bude typicky daleko doprava za perkolačním prahem. Pokud chcete *přibližně* lineární škálování, pak zase není jasné, jaká je tolerance. Doporučuju podívat se na http://en.wikipedia.org/wiki/Percolation_threshold pro srovnání vaší a standardní definice. Zajímalo by mě, jestli je vaše definice vlastní nebo odněkud přejatá.
6. Ačkoliv je práce psaná velmi jasné, přiznám se, že jsem nepochopil strukturu dat ze sekce 4.1. Prosím o krátké vysvětlení.
7. První odstavec na straně 30 popisuje zastavovací kritérium. Dle mého názoru byste měla brát zřetel na variabilitu průměru sledovaných matic. Z prvních 500 opakování máte představu, jaká je variabilita a podle toho byste měla navrhnout zastavovací kritérium. Abyste

algoritmus nezastavila jen proto, že došlo k normální náhodné odchylce. Mohl bych vidět graf průběhu prvních 500 průměrů?

8. Při pohledu na řešení na straně 40 nahoře mě napadá, jestli neexistuje nějaká rychlá a chytrá heuristika, která by zvládla totéž. Spočítat nějakou míru centrality vrcholů (closeness, betweenness, ...) a spojit hranou ty, co ji mají největší? Takový algoritmus proběhne okamžitě i na obrovských sítích a může alespoň sloužit jako prediktor pro složitější optimalizační nástroje.
9. Zajímaly by mě časy výpočtů pro jednotlivé algoritmy.
10. Největší výhradu mám k přiloženým kódům. Jednak bych ocenil, kdyby byly kódy lépe okomentované, aby bylo uživateli jasné, co která funkce dělá, na jaké vstupy se volá a z jakého programu je volána. Dále bych rád, aby CD obsahovalo nějaký readme soubor, ze kterého by bylo jasné, jak mohu načíst data (případně je zobrazit) a jak mohu na těchto datech vyzkoušet všechny algoritmy, které jsou v práci popsány. Prosil bych, abyste tuhle úpravu ještě dodělala, ať je CD (které prosím přiložte znovu) skutečně použitelné bez informaticko-detektivní práce. Navíc bych doporučoval podívat se alespoň na upozornění, která vám při zobrazení kódu nabízí matlabovský m-file editor, většinou mívá pravdu.

Celkově soudím, že práce je napsaná dobře, téma je zajímavé a výsledky dobré. Práci rozhodně **doporučuju k obhajobě** a zejména vzhledem k poslednímu bodu připomínku **navrhuju hodnocení B**.

V Olomouci 12. května 2015

Tomáš Fürst