

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Efektivní výpočty vlastností neorientovaných grafů
Jméno autora:	Jana Zelenková
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Oponent práce:	Petr Ryšavý
Pracoviště oponenta práce:	IDA, Katedra počítačů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	průměrně náročné
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
<p>Úkolem práce bylo porovnat efektivitu dostupných programů pro počítání vlastností prostých neorientovaných grafů. Studentka měla za úkol vybrat množinu těchto vlastností, následně pomocí vlastních skriptů měřit čas běhu na vygenerované množině grafů a určit vliv typu grafu a použitého hardware. Zadání neobsahuje žádné body, které by představovaly riziko, že je nepůjde uskutečnit, náročnost je daná pouze pracností danou velkým množstvím zkoumaných metod.</p>	

Splnění zadání	splněno s menšími výhradami
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
<p>V bakalářské práci jsou zvoleny pouze dva balíky, pro které studentka porovnávala časy běhu, a to SageMath knihovna pro jazyk python a Wolfram Mathematica. Přehled vlastností, které jsou počítány, považuji za zcela dostatečný, jde dohromady o přibližně 12 vlastností zahrnujících spojitost grafu, hledání Eulerovských a Hamiltonovských kružnic, určování poloměru grafu a jiné. Naopak diskusi charakteristických vlastností považuji za nedostatečnou, studentka rozděluje grafy pouze na spojitě, cyklické a náhodné. Očekával bych například porovnání hustých a řídkých grafů, především u problémů typu hledání minimální kostry či hledání Hamiltonovské kružnice. Porovnání efektivnosti při využití superpočítače obsahuje pouze sdělení, že výkonnější stroj nemá na výsledky vliv, protože uvažované programy jsou spouštěny jednovláknově. Porovnání výsledků s teoretickými předpoklady je provedeno pouze neformálně, vždy pár vět pro každou vlastnost.</p>	

Zvolený postup řešení	správný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
<p>Studentka zvolila správný postup řešení. Pro evaluaci používá kombinaci kódů v jazycích Bash, Python, Wolfram a C++. Kód je dokumentovaný, přehledný a rozšířitelný.</p>	

Odborná úroveň	E - dostatečně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
<p>Zde mám největší výhrady. Přehled grafových vlastností je z velké části neformální, mnoho definic nejednoznačných a nepřesných. Jako příklad mohu uvést např. definici minimální kostry grafu (<i>Je podmnožina grafu, kde se počet vrcholů v této podmnožině rovná počtu vrcholů grafu. Zároveň je tato podmnožina stromem.</i>) či poloměru grafu (<i>Naopak poloměr grafu G je minimální excentricita, tedy minimální vzdálenost mezi dvěma vrcholy v G.</i>)</p> <p>Definice na sebe často navenazují, například:</p> <ul style="list-style-type: none"> v kapitole 4.2 je zaváděn souvislý graf jako graf, jehož každé dva vrcholy jsou spojené cestou. Nikde předem není řečeno, co je cesta. Stejně tak definice stromu v kapitole 4.3 se opírá o chybějící definici cesty/kružnice. Ve 4.4 je definovaná excentricita vrcholu, přičemž se tato definice opírá o nikde nevysvětlenou vzdálenost, ta je až 	

ve 4.18.

- Ve 4.12 u vysvětlení chromatického čísla chybí informace o tom, co je to diskrétní či bipartitní graf.
- V kapitole 4.13 se mluví o planárních grafech. Rovinné grafy jsou vysvětlené až v 4.15.

Některé neformální definice jsou nejednoznačné. Například popis NP-úplné úlohy (kapitola 3.3) sice ve dvou poskytnutých větách neobsahuje nic nepravdivého, ale autorka nevysvětluje, co to vlastně NP-úloha je nebo proč jde o důležitou množinu problémů. Po přečtení by pak uživatel mohl nabýt dojmu, že NP-úplné úlohy jsou obecně všechny, u kterých není znám polynomiální algoritmus

Formální a jazyková úroveň, rozsah práce

D - uspokojivě

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

Text je místy špatně srozumitelný. Příkladem může být vůbec první odstavec bakalářské práce či věta v kapitole 4.8 (Eulerovský graf), kde studentka píše: „*Je překvapivé, že řešení tohoto problému je poměrně rychlé a jedná se o polynomiální nebo i lineární obtížnost.*“

V práci stále zůstávají místy překlapy, např. drobné chyby v gramatice – velké písmeno uprostřed věty na straně 1 v druhém odstavci. Množství měřených dat je prezentováno ve formě velkého množství grafů, vždy jeden graf pro každou vlastnost a každý program. Studentka se snažila chyby v gramatice na základě připomínek komise opravit, nicméně ne vždy to bylo k lepšímu. Například v závěru věty „*Dále jsme se soustředily na porovnání, kolik času zabere výpočet nějaké vlastnosti pokud měníme velikost grafu,*“ nahradila věta „*Dále jsme se soustředila na porovnání, kolik času zabere výpočet nějaké vlastnosti, pokud měníme velikost grafu.*“

Mezi původním a novým textem se vyskytují nekonzistence, např. v kapitole 4.7 je střídavě počet vrcholů značen n a $|V|$. Celkově došlo ale ke zlepšení původního stavu.

Výběr zdrojů, korektnost citací

E - dostatečně

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Určité výhrady mám jak k volbě zdrojů, tak ke způsobu citování. Studentka v přehledu důležitých grafových vlastností přednostně cituje vždy obecné knihy o teorii grafů a k nim vždy zvolenou stránku z online dokumentace poskytnuté společností Wolfram na adrese <https://mathworld.wolfram.com/>. Jde sice o zdroj s uvedeným autorstvím a studentka navíc správně uvádí datum navštívení stránek, nicméně některé věty v textu v rešeršní části jsou pouhým doslovným překladem těchto stránek, který by měl být označen jako doslovná citace. Například:

- *Jinými slovy, lze říct, že průměr grafu je nejvyšší počet vrcholů, které musí být navštíveny, pokud cestujeme z jednoho vrcholu do jiného, ale to pouze za předpokladu, že jsou vyloučeny cesty, které vedou zpět, jsou objížďkami nebo smyčkami. Jedná se tedy o maximum všech hodnot v matici grafové vzdálenosti.* [kapitola 4.5]
- *In other words, a graph's diameter is the largest number of vertices which must be traversed in order to travel from one vertex to another when paths which backtrack, detour, or loop are excluded from consideration. It is therefore equal to the maximum of all values in the graph distance matrix.* [<https://mathworld.wolfram.com/GraphDiameter.html>]

Porovnání s citovanými knihami jsem neprováděl.

Musím nicméně konstatovat, že v tomto ohledu došlo ke zlepšení oproti první verzi práce. V několika případech studentka doplnila citace původních článků, které se zmíněnými algoritmy přišly. Naopak v případě Fleuryho algoritmu studentka nevhodně cituje webové stránky místo původního článku, dále například vyjmenované balíky jazyka Python v kapitole 2.2 by také bylo vhodné citovat. Citace jsou často nevhodně umístěny na konec odstavce či podkapitoly místo na první výskyt citovaného pojmu.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a

funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Nemám další komentář. Studentkou vytvořené řešení je funkční, avšak prezentace ve formě velkého množství grafů je nevhodná. Očekával bych například navíc agregované výsledky, které by ukazovaly tabulku s porovnáním, na kterých problémech byla rychlejší knihovna SageMath, a na kterých Wolfram Mathematica.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Studentka měla za úkol provést porovnání efektivity dostupných knihoven při výpočtu vlastností prostých neorientovaných grafů. Studentka si vybrala dvě knihovny, a to SageMath a Wolfram Mathematica. Dále měla porovnat vliv typu grafu a použitého hardware na čas výpočtu. Šlo o pracné, ale ne příliš náročné zadání.

Studentka zadání splnila, nicméně s určitými výhradami. Například vliv použití superpočítače na čas běhu byl označen za zanedbatelný s tím, že byly použity jedno-vláknové knihovny. Podobně knihovna SageMath není uvažována na zmiňovaném superpočítači, i když je možnost ji spustit a nainstalovat i pro uživatele bez administrátorských práv (testováno na xUbuntu 20.04). Rozdělení grafů na spojité, cyklické a náhodné také není příliš vhodné. Práce obsahuje často nepřehledné definice, z popisu NP-úplných úloh čtenář může nabýt dojmu, že jde z definice o úlohy, které nemají polynomiální řešení, některé myšlenky na sebe příliš nenavazují, kapitoly jsou krátké, definice využívají pojmů, které nejsou definované vůbec nebo až v následujících kapitolách. Na druhou stranu kód, který studentka odevzdala je čitelný, s komentáři a rozšiřitelný.

Na studentku bych měl následující otázky.

- Jaký matematický model byl použit pro generování náhodných grafů?
- V kapitole 5.3 mluvíte o „superpočítačích“? Je něco, co Vám bránilo sama nainstalovat SageMath na tento stroj?
- Obrázky 5.21 a 5.22 ukazují jeden bod, který neodpovídá trendu. Byl odpovídající graf něčím výjimečný nebo jde z nějakého důvodu o outliera?

Předloženou práci jsem měl možnost číst podruhé. Komise studentce vytýkala malý rozsah práce, velké množství chyb a způsob provedení práce. Studentka v přepracované verzi mírně změnila pořadí kapitol, v rešeršní části přidala do přehledu okolo jednoho odstavce ke každé grafové úloze, v experimentech pak podobně přidala k některým měřeným vlastnostem odstavce popisující naměřené výsledky. Poslední změna pak spočívá v drobných opravách gramatiky a překlepů. Nicméně některé z mých připomínek zůstaly nezpracované, např. v definicích zůstává místy doslovný překlad ze serveru MathWorld bez označení jako doslovná citace či používání nedefinovaných pojmů v textu. V metodice či kódu jsem neobjevil žádné změny. Pokud beru jako referenci pro ohodnocení práce známkou rozhodnutí komise při prvním předložení práce, navrhuji známku E.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **E - dostatečně**.

Datum: 18.1.2021

Podpis: