



Posudek oponenta závěrečné práce

Oponent práce:	Ing. Václav Blažej
Student:	Bc. Kamil Červený
Název práce:	Paralelní algoritmy pro maximální toky v sítích na architektuře CUDA
Obor / specializace:	Teoretická informatika
Vytvořeno dne:	3. června 2021

Hodnotící kritéria

1. Splnění zadání

- [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- ▶ [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Zadání po studentovi chce splnit tři body.

Prvním je nastudovat si a provést rešerši algoritmů pro hledání maximálních toků v síti; odkazuje se na článek, kde byl představen Goldbergův algoritmus, a článek s Ahuja-Orlin algoritmem. Popis obou algoritmů je v práci přítomen na stranách 18-26, ovšem část textu, kde se oba nachází je celá převzatá (detailně rozeberu v sekci "Písemná část práce"), nejedná se tedy o originální práci studenta.

Druhá část chce po studentovi navrhnout efektivní paralelizaci těchto algoritmů pomocí arch. CUDA -- tento bod je splněn bez výhrad.

Třetí vyžaduje implementaci, otestování, měření zrychlení, a porovnání s dalšími publikacemi a open-source knihovnamí. Implementace je v pořádku s drobnými výhradami; otestování je nedostatečné; měření zrychlení je nedostatečné, porovnání s dalšími publikacemi a open-source knihovnamí je nedostatečné (detailně rozebráno v sekci "Nepísemná část, přílohy"). Kapitola "Testování a měření algoritmů" je necitovaně přebrána z cizí práce -- provádí pouze drobné změny textu a nahrazuje data původní práce daty ze studentových měření (opět, detailně rozebráno v sekci "Písemná část práce"). Proto jsem nucen vyslovit důvodné podezření, že značná část této práce je plagiát.

2. Písemná část práce

20/100 (F)

Práce je členěna na pět sekcí. První popisuje naprosté základy teorie, druhá sekvenční algoritmy pro maximální toky, třetí popisuje architekturu CUDA a existující paralelní algoritmy pro maximální tok, čtvrtá se rozebírá detaily implementace a pátá obsahuje testování a měření.

Kapitola 1 "Základní pojmy a poznatky" (mimořadně začínající větou, která je totožná s BP K. Červeného) začíná uvedením toho, že definice a věty této kapitoly jsou převzaty z knihy [M. Mareš, T. Valla. Průvodce labyrintem algoritmů, 2017] a handoutů předmětu AG2 [O. Suchý, T. Valla. Síť, toky v sítích, Ford-Fulkersonův algoritmus, CVUT, FIT, KTI, 2019]. Kromě úvodního odstavce a prvních dvou definic, je pět a půl strany z šesti stran opsáno ze zmíněných dvou zdrojů s drobnými úpravami textu a značení (tj. záměna za synonyma, permutace slov, přeskládání vět). Spojovací texty mezi definicemi a větami jsou pouze parafráze spojovacích textů ze zdrojů.

Kapitola 2 "Sekvenční algoritmy" obashuje 16 stran, ze kterých je 11 převzato z Průvodce labyrintem algoritmů stejným stylem jako v Kapitole 1, což student zmiňuje začátkem kapitoly. Poslední dvě strany 25-26 ovšem v Průvodci nejsou; je zde podsekcce 2.3.2 "Vylepšení škálování přebytků", ve které se část definice delty, Lemma 2.3.14, Lemma 2.3.15 a Lemma 2.3.16 (včetně důkazů) jsou v podstatě stejné jako podsekcce 2.2.4 ve zde neodcitované Bakalářské práci J. Groschafta [J. Groschaft, Paralelizace algoritmů pro toky v síti, bakalářská práce, FIT ČVUT, 2019], konkrétně definice $\delta(u,v)$, následné pozorování, Lemma 14 a Věta 10. Diskuse v podsekcce 2.3.4 připomíná rozsáhlejší verzi diskuse v podsekcce 2.2.2 zmíněné BP.

Kapitola 3 "Paralelní algoritmy na architektuře CUDA" začíná podsekcí 3.1 "Základní pojmy složitosti paralelních algoritmů", která zavádí pojmy "doba výpočtu", "paralelní zrychlení", "paralelní cena" apod. Celá je převzatá z handoutů NI-PDP Přednášky 1 (která je citována), ovšem zavedené pojmy se v textu dále nevyužívají. Sekce 3.2, 3.3, a 3.4 (celkem 2 strany) jsou v pořádku. Sekce 3.5 (3 a půl strany) sestává z překladu a přejatých obrázků citovaného zdroje programovací příručky CUDA od NVIDIA [NVIDIA Copr. CUDA C++ Programming Guide, v11.0, online].

Pěkná část práce je Sekce 3.6, která popisuje BFS na CUDA (1 a půl strany); Sekce 3.7, popisující existující algoritmy pro maximální tok (včetně pseudokódů; 8 stran); a Sekce 3.8, popisující nový algoritmus pro řešení problému (3 strany). Tyto sekce mají rozumnou strukturu, dobře pracují se zdroji. Popis nového algoritmu by si zasloužil rozšířit, rozepsat do detailu.

Kapitola 4 "Detaily implementace" obsahuje 3 a půl strany popisu použité technologie, datových struktur, a poznámky k implementaci. Sekce 4.4 zmiňuje jaké bylo rozhodnuto nastavení parametrů jednotlivých algoritmů -- data podkládající tento výběr by měla být součástí práce.

Kapitola 5 "Testování a měření algoritmů" je nejproblematickejší část této práce. Kapitola 5 je strukturou téměř totožná s Bakalářskou prací J. Groschafta. Konkrétně (porovnání této práce vůči BP J.G.): Vynechává BP 5.1 "Existující řešení"; Sekce 5.1 (str. 53) ~ Sekcí 5.2 (str. 33) BP J.G. (2 strany); Sekce 5.2 "Testování správnosti" ~ Sekce 5.3 "Testy správnosti" BP; Sekce 5.4 "Měření" ~ Sekce 5.4 "Testy rychlosti"; Podsekcce 5.3.2 "Výsledky měření" ~ Sekce 5.5 "Vyhodnocení výsledků".

Nejen struktura poslední kapitoly je stejná, ale Sekce 5.1 (str. 53) je na volbu slov a

konkrétních dat totožná se Sekcí 5.2 (str. 33) BP J.G. (2 strany), tj. vybrané datasety pro testování (instance DIMACS, KaHIP, Sítě z počítačového vidění) jsou totožné s BP J.G., z uvedených datasetů byly vybrány mírně jiné instance.

Dokonce na straně 54 je věta totožná s BP J.G. strany 35 "Počet hran nezahrnuje zpětné hrany s nulovou kapacitou.", která je použita ve stejném kontextu (jako druhá věta úvodu následné tabulky).

Testy správnosti jsou velmi odbyté -- porovnání stejné velikosti výsledku proběhlo pouze na osmi uvedených vstupních instancích, což je nedostatečné.

Měření je provedeno pro 1 sekvenční algoritmus a 5 paralelních algoritmů (na CUDA), zároveň je měřeno použití struktury "lattice" a data pro barvení v novém algoritmu.

Chybí porovnání implementovaných algoritmů s cizími implementacemi a chybí měření zrychlení paralelních algoritmů (části zadání).

Dodatečné poznámky: Na obrázek 3.3 a tabulky 2.1, 3.1 není odkázáno z textu. Strana 45 obsahuje značné množství překlepů (pebytku, proměnná, přebutku, Ve této).

3. Nepísemná část, přílohy

75 /100 (C)

Samotný kód je v pořádku.

Soubor README.md obsahuje pouze strohé informace a chybí popsat veškeré informace k požadavcích na instalované programy (nezmiňuje nutnost programu nvcc), nezmiňuje formát vstupních a výstupních souborů a jakým způsobem je načte (stdin či parametr?), jestli má další parametry při spuštění, jakou program vyžaduje platformu (tj. co udělat, když program píše `GPUassert: 'no CUDA-capable device is detected' in function 'init' in file src/utis/init.cpp: 63`), a co udělá program po spuštění. Základní informace lze zjistit vyzkoušením spuštění, protože program vypíše help (seznam parametrů), ovšem zbytek informací chybí.

4. Hodnocení výsledků, jejich využitelnost

80 /100 (B)

Oceňuji implementaci několika cizích algoritmů a jednoho studentova algoritmu v C++ za pomoci CUDA. Závěr práce je, že "paralelní implementace výrazně zaostávají za sekvenční verzí," proto se obávám, že implementace nebude mít vysokou využitelnost. Dalšími měřeními mohl student odhalit, jak a proč implementace reagují na různé nastavení parametrů, či vstupní data, což by dodalo práci velikou informační hodnotu, toto ovšem chybí.

Celkové hodnocení

20 /100 (F)

Student implementoval 4 cizí paralelní algoritmy a 1 svůj pro řešení problému maximálního toku na architektuře CUDA. Popis těchto algoritmů s prvky implementace jsou pěkně shrnuty v části kapitoly 3 a v kapitole 4. Potřebná teorie (kapitoly 1, 2 a začátek kapitoly 3) je převzata z dalších zdrojů (Průvodce labyrintem algoritmů, handouty přednášek předmětů BI-AG2 a NI-PDP, a bakalářské práce J. Groschafta, NVIDIA manuál pro CUDA). Převzaté části jsou pouze mírně upraveny (záměny slov za synonyma, úprava pořadí slov a odstavců, změna notace), parafrázovány, či přeloženy. Příprava a výsledky testování jsou prezentovány v kapitole 5.

Převzaté části jsou nevhodně citovány (pouze na začátku kapitoly) a z celé práce tvoří více jak 47% textu (25 z celkových 53 stran textu). Větší problém je v kapitole 5 a uvnitř

kapitoly 2 na stranách 25-26, kde student převzal text, upravil a neodcitoval práci jiného studenta [J. Groschaft, Paralelizace algoritmů pro toky v síti, bakalářská práce, FIT ČVUT, 2019].

Pokud vyloučíme převzaté části, práce by rozsahem odpovídala práci bakalářské. Algoritmy jsou sice rozumně implementované a popsány, ale testování a měření (hlavně proti dalším implementacím) je nedostatečné. V neposlední řadě nelze přehlížet fakt, že práce nedodrží prohlášení, které má na samém začátku. Vzhledem ke všem uvedeným výhradám nemohu doporučit tuto práci k obhajobě na FIT.

Otázky k obhajobě

Prostor pro otázky bych rád přenechal studentově reakci na tento posudek.

Instrukce

Splnění zadání

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Nepísemná část, přílohy

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Celkové hodnocení

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.