



Posudek oponenta závěrečné práce

Student: Thanh Quang Mai
Oponent práce: Ing. Daniel Langr, Ph.D.
Název práce: Efektivní násobení řídkých matic
Obor: Teoretická informatika

Datum vytvoření: 28. 5. 2019

<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – následující škálou 1 až 4:</i>
1. Splnění zadání	<u>1=zadání splněno,</u> 2=zadání splněno s menšími výhradami, 3=zadání splněno s většími výhradami, 4=zadání nesplněno
<i>Popis kritéria:</i> Posuďte, 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, posuďte 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.	
<i>Komentář:</i> Práce dostatečně splňuje cíle formulované v zadání, které je spíše jednoduššího charakteru. Je vyžadovna implementace existujících algoritmů pro násobení řídkých matic a jejich paralelizace.	
<i>Hodnotící kritérium:</i>	<i>Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):</i>
2. Písemná část práce	70 (C)
<i>Popis kritéria:</i> Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posuďte, 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. Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3. Posuďte, 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.	

Komentář:

Práce je poměrně dobře čitelná a logicky členěná, jednotlivé kapitoly na sebe navazují. K obsahové a formální stránce mám následující připomínky:

- Abstrakt je velmi stručný. Měl by shrnovat celou práci včetně představení hlavních získaných výsledků a poznatků.
- Abstract - "...implementation of said formats..." - místo "said" radši např. "mentioned" apod.
- str. 1 - "Práce s s takto..." -> "Práce s takto".
- sekce 1.1 - Připadá mi, že zde jsou zbytečné definice. K čemu je např. definovaná jednotková, regulární, či inverzní matice? Nenašel jsem žádné využití těchto definic/matic v dalším textu.
- sekce 1.2.4 - Zápis vzorce působí, jako když se jedná o jednotkovou matici umocněnou na dvanáctou. Vhonější by bylo přesunout indexy pro poznámky pod čarou do textu (např. za poslední slovo před vzorcem).
- Na mnoha místech v textu chybí mezery, většinou mezi slovem a referencí (např. "regulární[2]" -> "regulární [2]" apod.).
- str. 5 - "Řídkost matice se vyjadřuje pomocí nnz..." - nnz samo o sobě o řídkosti nic neříká, spíše se používá "hustota" vyjádřená jako $nnz/(n*m)$.
- str. 9 - Poslední řádek obsahující vzorec je špatně čitelný díky absenci mezer za čárkami. Ty by bylo vhodné doplnit manuálně.
- str. 11, poslední řádek - Opět matoucí zápis díky poznámce pod čarou. Vypadá to, že se jedná o nnz umocněno na třetí.
- str. 12 - $O(3\ nnz)$ je to samé jako $O(nnz)$. Paměťová složitost každého formátu pro řídké matice je minimálně $O(nnz)$. Pokud by chtěl autor zdůraznit rozdíly mezi formáty, je nutno paměťové nároky vyčíslit přesně a ne asymptoticky.
- str. 15 - opět díky poznámce pod čarou zápis odpovídá n umocněno na šestou.
- str. 18 - přetečení šířky řádku.
- str. 18 - Rovnost v matematickém vzorci vyjadřuje jednoduché rovnítko, nikoliv dvojité. Dvojitě rovnítko značí zápis operátoru porovnání z jistých programovacích jazyků, navíc výsledek této operace může být true/false, takže neznačí rovnost.
- kapitola 4 - Poměrně vážný nedostatek vidím v tom, že autor neuvádí, v jakém formátu jsou u jednotlivých knihoven ukládány výsledné matice operace násobení. To totiž má zásadní dopad na použité algoritmy, protože u hustého formátu (který používá i autor v jeho implementaci) lze indexovat jednotlivé prvky výstupní matice v konstantním čase, což u řídkých formátů není možné.
- str. 28 - "Data by měla být v paměti vhodně umístěny..." -> "umístěna".
- str. 30 - "...hlavní vlákno(master thread)..." - chybí mezera.
- str. 33 - Kód je zbytečně "roztahaný" s příliš velkým odzazením (v praxi se používá odzazení většinou do 4 znaků). Jsou zde zbytečné prázdné řádky, jednoprvkové tělo cyklu není nutné obalovat složenými závorkami, příkaz "return 0;" je v C++11 u funkce main zbytečný.
- Asi nejvýznamnější výtka mám k tomu, že se autor nepokusil o žádné efektivní řešení paralelizace v případech, kdy dochází k současnému zápisu více vláken na stejný prvek výsledné matice (viz sekce 5.3). Jediným řešením byla paralelizace vnitřní smyčky, což je ale obecně velmi neefektivní řešení (už jen proto, že každá paralelizace má určitou režii-overhead a vyplatí se spíše pro smyčky s větším počtem iterací). Zde by šlo snadno navrhnout a otestovat alespoň dva základní přístupy - první založený na atomických inkrementech hodnoty maticového elementu (ty nemusí architektura přímo podporovat, ale dají se řešit pomocí atomických operací compare-and-swap, což využívají i implementace OpenMP). Druhý přístup by spočíval v tom, že každé vlákno bude mít svoji kopii výsledné matice, nad kterými se na konci provede redukce. Určitě by bylo zajímavé vidět výsledky experimentů takovýchto řešení a porovnat je s triviálním řešením paralelizace vnitřního cyklu.
- str 43. - "Algoritmy jsem změřil na vlastních vygenerovaných maticích..." - Zcela chybí popis metody generování těchto matic.
- Velmi oceňuji verifikaci výsledků násobení pomocí porovnání s výsledky násobení hustých matic. Rovněž to, že násobení hustých matic bylo součástí měření, což je velmi dobré pro vyhodnocení efektivity násobení v řídkých formátech.
- Veškeré grafy jsou prakticky nečitelné, přitom místa na větší font je dostatek. Navíc je velmi obtížné odlišit některé barvy výsledných naměřených křivek. Pro přibuzené barvy by měly mít čáry i odlišnou strukturu (např. plná, čárkovaná, tečkovaná, apod.).
- U paralelních experimentů se mi líbí uvedení grafů zrychlení, ale měly by být uvedeny i naměřené časy pro "optimální" počet vláken (stejný jako počet jader). Takto je obtížné posoudit, který algoritmus bych v paralelních verzích nejrychlejší.
- Pro paralelní experimenty by bylo vhodnější použít nějakou knihovnu, která operaci násobení řídkých matic paralelizuje. Např. Intel MKL tuším tuto operaci podporuje.

Refernce jsou využity správně, nenašel jsem zde žádný nedostatek.

Hodnotící kritérium:

Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):

Popis kritéria:

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

Komentář:

Kód je přehledně členěn do hlavičkových a zdrojových souborů. Oceňuji i přítomnost souboru Makefile pro možnost automatického překladu. Nechápu přítomnost zdrojového souboru "matrix.cpp", který neobsahuje žádný kód (definice funkcí apod.). Dále bych trochu více doporučil využívat možnosti C++, např. proudový výstup pomocí `std::cout` namísto funkce jazyka C `printf`.

Hodnotící kritérium:

Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):

4. Hodnocení výsledků, jejich využitelnost

75 (C)

Popis kritéria:

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.

Komentář:

Výsledky práce přináší určitý náhled na algoritmy násobení matic v různých řídkých formátech. Líbí se mi porovnání s násobením v hustém formátu, což umožňuje vyhodnotit základní přínos a efektivitu algoritmů pro řídké formáty. Na druhou stranu výsledná matice byla vždy ukládána v hustém formátu, což znemožňuje násobení matic, pro něž by se výsledná matice v hustém formátu nevešla do paměti. A takové případy v praxi nastávají často. Je ale pravdou, že takové řešení by pravděpodobně přesahovalo hranice jedné bakalářské práce.

Co se týče získaných poznatků z paralelizace algoritmů, ty jsou pouze základní. Chybí mi porovnání pokročilejších metod paralelizace, např. těch založených na atomických modifikacích maticových prvků či na lokálních kopiích výsledných matic pro jednotlivá vlákna. Takové přístupy ale tvoří dobrý námět na rozšíření prezentovaného výzkumu.

Hodnotící kritérium:

Způsob hodnocení – nehodnotí se

5. Otázky k obhajobě

Popis kritéria:

Uveďte případné dotazy, které by měl student zodpovědět při obhajobě ZP před komisí (body oddělte odrážkami).

Otázky:

- 1) V práci je definována operace transpozice matice (sekce 1.3.4) a dále je tato operace i implementována. Rád bych se zeptal k čemu? Nenašel jsem žádnou souvislost s tématem práce a jejím zadáním? Byla transpozice někde využita?
- 2) Autor předpokládá lexikografické uspořádání prvků ve formátu COO (klíč řádek-sloupec). Nebylo by pro násobení matic $A \times B$ vhodnější u matice B seřadit prvky podle klíče sloupec-řádek?
- 3) Jaké formáty (hustý/řídké) podporují knihovny prezentované v sekci 4.1 pro výslednou matici operace násobení matic?
- 4) Str. 37 - "Prvky matice jsou nejdříve načteny do STL kontejneru `map<int, double>...`" - Co značí onen klíč typu `int`? Kde je uložen řádkový a sloupcový index prvku matice?

Hodnotící kritérium:

Způsob hodnocení – bodové hodnocení 0 až 100 bodů (známka A až F):

6. Celkové hodnocení

75 (C)

Popis kritéria:

Shrňte stránky ZP, které nejméně 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.

Text hodnocení:

Jedná se o zajímavé výzkumné téma, v tomto případě spíše s jednodušším zadáním. Víceméně autor převzal existující algoritmy a provedl jejich implementaci, optimalizaci a následnou paralelizaci. Právě od paralelizace bych v práci čekal trochu více, ta byla provedena jen na velmi triviální úrovni, tj. doplnění direktiv OpenMP pro paralelizaci smyčky či vytváření OpenMP úloh. Tvzení "Algoritmy se mi nepodařilo upravit tak, aby tzv. race conditions nevznikaly, a proto jsem se nakonec rozhodl k paralelizaci vnitřního cyklu..." mi připadá poněkud slabé a určitě by zde šlo vymyslet více přístupů než jen ten nejjednodušší zvolený. Jinak ale práce zadání splňuje a umožňuje náhled a porovnání na problém násobení řídkých matic v případech uložení vstupních matic v různých formátech.

Podpis oponenta práce: