Częstochowa, dn. 06.02.2018

Prof. dr hab. inż. Roman Wyrzykowski
Instytut Informatyki Teoretycznej i Stosowanej
Politechnika Częstochowska
ul. Dąbrowskiego 69
42-201 Częstochowa

RECENZJA
ROZPRAWY DOKTORSKIEJ
mgr inż. Pawła Jarzębskiego
„Zastosowanie algorytmów wielowątkowych i rozproszonych do zwiększenia efektywności Metody Elementów Skończonych"

Promotor: Prof. dr hab. inż. Krzysztof Wiśniewski
Instytut Podstawowych Problemów Techniki PAN w Warszawie

1. Obszar problemowy rozprawy

Tematyka przedłożonej do recenzji rozprawy doktorskiej mgr inż. Pawła Jarzębskiego dotyczy ważnego obszaru badawczego w dziedzinie informatyki oraz współczesnych nauk obliczeniowych. Rozpatrywana problematyka ma nie tylko znaczenie poznawcze, lecz również wiąże się z implikacjami aplikacyjnymi, jakie dla rozwoju nauki i techniki niesie szerokie zastosowania najnowszych metod i narzędzi informatycznych do modelowania i symulacji różnego rodzaju zjawisk fizycznych, konstrukcji i procesów technologicznych czy też całych złożonych systemów.

Metoda elementów skończonych (MES) stanowi jedną z podstawowych metod obliczeniowych stosowanych w problematyce „Wielkich Wyzwań Nauki" (*Grand Challenges*), głównie do modelowania zjawisk opisywanych równaniami różniczkowymi cząstkowymi. Charakteryzuje się ona najczęściej wielką złożonością obliczeń oraz dużą objętością przetwarzanych danych. Jest ona dynamicznie rozwijana od wielu już lat przez badaczy z całego świata. Rozwój ten dotyczy praktycznie wszystkich jej aspektów - począwszy od podstaw matematycznych z uwzględnieniem specyfiki konkretnej dziedziny aplikacyjnej, poprzez metodologię wyznaczania rozwiązań numerycznych, a skończywszy na aspektach algorytmicznych i programistycznych związanych z jej efektywną implementacją we współczesnych systemach komputerowych.

Dla najbardziej interesujących z punktu widzenia poznawczego i aplikacyjnego zagadnień charakteryzujących się bardzo dużą złożonością czasową i/lub pamięciową wykorzystanie pojedynczego zasobu (wątku) obliczeniowego staje się niewystarczające czy to ze względu na czas obliczeń, czy też ograniczoność zasobów komputera, np. pamięci operacyjnej. Aby umożliwić skuteczne prowadzenie eksperymentów numerycznych w takich przypadkach, niezbędne jest zastosowanie systemów z wieloma jednostkami przetwarzającymi, funkcjonujących w trybie przetwarzania równoległego. Obliczenia równoległe umożliwiają rozwiązanie wspomnianych wyżej problemów dzięki podziałowi zadania na mniejsze fragmenty i ich równoległą implementację w systemie z wieloma rdzeniami czy na klastrze komputerowym.

Jednakże skuteczne zastosowanie obliczeń równoległych w przypadku MES związane jest z koniecznością rozwiązania szeregu złożonych problemów, w tym również efektywnej organizacji obliczeń z uwzględnieniem właściwości architektury systemu komputerowego. Dotyczy to praktycznie wszystkich etapów składających się implementację analizy MES, takich jak generacja samej siatki, budowa macierzy dla poszczególnych elementów i tworzenie macierzy globalnej, a także rozwiązywanie wynikowego układu równań liniowych, co w praktyce prowadzi do konieczności wykonywania obliczeń na macierzach rzadkich o dużych rozmiarach. Wynika stąd szczególne znaczenie ostatniego spośród wymienionych etapów, gdyż dla rzeczywistych zagadnień charakteryzuje się on bardzo dużą złożoność obliczeniową i pamięciową oraz skomplikowaną i nieregularną strukturą wykonywanych obliczeń.

W nakreślonym kontekście wybór tematyki rozprawy uważam za trafny, gdyż Autor w recenzowanej rozprawie, nie zaniedbując etapu generacji macierzy dla poszczególnych elementów i tworzenia macierzy globalnej, skoncentrował swoją uwagę na algorytmach rozwiązywania dużych układów równań liniowych występujących w rzeczywistych zagadnieniach MES, potwierdzając w rzeczywistych zastosowaniach przydatność opracowywanych rozwiązań algorytmicznych i tworzonych na ich podstawie kodów równoległych. Uważam więc, że rozważane w dysertacji problemy są ważne i aktualne, zarówno dla teorii, jak i przede wszystkim praktyki zastosowania MES do rozwiązywania złożonych zagadnień naukowych i technicznych.

2. Koncepcja i redakcja rozprawy

Recenzowana praca doktorska obejmuje formalnie 7 rozdziałów, bibliografię zawierającą 119 pozycji, a także aż 9 dodatków o dostatecznie różnorodnym charakterze. Zasadnicza część rozprawy (bez dodatków) liczy łącznie 150 stron.

W ***rozdziale pierwszym*** Autor zawarł lakoniczne wprowadzenie do tematyki rozprawy, a także przedstawił jej cel oraz krótko scharakteryzował zawartość.

W rozdziale drugim przedstawiono architektury systemów komputerowych wykorzystywanych w pracy, a także dokonano wprowadzenia do programowania równoległego zarówno w modelu z pamięcią współdzielona (obliczenia wielowątkowe), jak i z pamięcią rozproszoną (zwane w pracy obliczeniami rozproszonymi). Istotny element tego rozdziału stanowi wprowadzenie do wykorzystywanego w pracy modelu wydajnościowego, powszechnie znanego pod nazwą Roofline.

R***ozdział trzeci*** rozpoczyna oryginalną część pracy i poświecony został zrównolegleniu obliczeń wielowątkowych w programie FEAP. Opracowany algorytm wielowątkowy wykorzystuje standard OpenMP i przeznaczony jest do przyspieszenia obliczeń odpowiedzialnych za generowanie macierzy dla poszczególnych elementów oraz tworzenie macierzy globalnej. Istotnym elementem tego rozdziału są testy opracowanego rozwiązania dla dwóch przypadków, w tym analizy powłoki nieliniowej, a także analiza wydajności w oparciu o model Roofline.

Rozdział czwarty stanowi kontynuacje poprzedniego rozdziału i dotyczy bezpośrednich solwerów wielowątkowych dedykowanych do rozwiązywania układów równań liniowych z macierzami rzadkimi na jednym węźle obliczeniowym. Na podstawie analizy m.in. skalowalności, zapotrzebowania na pamięć oraz wpływu przenumerowania kolumn i wierszy macierzy na czas faktoryzacji, przeprowadzonej dla pięciu różnych solwerów, wybrano potencjalnie najlepszy z nich, a następnie zaproponowano dla niego kilka stosunkowo drobnych ulepszeń dotyczących konfiguracji parametrów i ustawień solwera. Interesujący element tego rozdziału stanowią badania solwera w mieszanej precyzji z iteracyjnym poprawianiem rozwiązania. Jak wykazano, jego użycie pozwala w szeregu zastosowań nie tylko skrócić czas wyznaczania rozwiązania, lecz także zwiększyć rozmiar problemu, który można rozwiązać na jednym węźle obliczeniowym.

Kontynuując powyższe rozważania, w ***rozdziale piątym*** zawarto wyniki badań poświeconych rozproszonym solwerom, wykorzystującym standard programowania MPI z wymianą komunikatów do zwiększenia skalowalności przeprowadzanych obliczeń dzięki użyciu wielu węzłów obliczeniowych. W oparciu o analizę właściwości algorytmów dedykowanych do wyznaczania uzupełnień Schura opracowano rozproszony solwer wykorzystujący hierarchiczną faktoryzację macierzy, zarówno w podwójnej, jak i mieszanej precyzji, przy czym w celu zwiększenia efektywności solwera zaproponowano wielowątkowy algorytm do redukcji zapotrzebowania na pamięć w trakcie częściowej faktoryzacji macierzy na pojedynczym węźle. Na podkreślenie zasługuje zrealizowane w tym rozdziale porównanie opracowanych przez Autora solwerów z alternatywną metodą FETI-DP, która wykorzystuje metody bezpośrednie na węzłach oraz solwer iteracyjny PCG (*Preconditioned Conjugent Gradient*) dla równania interfejsów.

Rozdział szósty zawiera bardzo interesujące wyniki o charakterze aplikacyjnym, dotyczące testowania zaproponowanych algorytmów w wybranych zadaniach MES, w celu weryfikacji ich właściwości numerycznych, wydajności, skalowalności oraz zapotrzebowania na pamięć. Do zadań tych należy modelowanie właściwości kompozytu ceramicznego oraz pianki korundowej, przy czym największy obliczony model zawierał ponad 30 mln niewiadomych.

W ***rozdziale szóstym Autor*** dokonał zwięzłego omówienia rezultatów rozprawy, starając się scharakteryzować elementy oryginalne oraz podając listę opracowanych implementacji (kodów numerycznych).

3. Wkład Autora i zaprezentowana wiedza

Uwzględniając powyższe omówienie zawartości pracy oraz ogólną pozytywną ocenę jej zawartości merytorycznej, uważam, że za bezsporne osiągnięcia Autora należy uznać następujące rezultaty:

1. Podstawowym wynikiem ogólnym o charakterze zarówno poznawczym, jak i aplikacyjnym, jest wykazanie możliwości zwiększenia efektywności równoległych kodów numerycznych przeznaczonych do rozwiązywania rzeczywistych problemów analizy MES zarówno na platformach z pamięcią współdzieloną (wielowątkowych), jak i rozproszonych dzięki zastosowaniu i efektywnej implementacji odpowiednio zmodyfikowanych technik algorytmicznych uwzględniających zarówno właściwości współczesnych równoległych systemów obliczeniowych, jak i kluczowe aspekty praktycznej implementacji zagadnień MES, a w szczególności, metod rozwiązywania dużych i bardzo dużych układów równań liniowych z rzadkimi macierzami współczynników.
2. Wśród wyników cząstkowych jako najbardziej interesującym w mojej opinii chciałbym wymienić opracowanie rozproszonego solwera, bazującego na dekompozycji obszaru i wyznaczaniu uzupełnień Schura, który wykorzystuje technikę częściowej faktoryzacji na pojedynczym węźle, jak również hierarchiczną faktoryzację macierzy blokowej dla podwójnej i mieszanej precyzji obliczeń. Wydajność oraz skalowalność opracowanego solwera dla podwójnej precyzji obliczeń porównywalna jest z solwerem komercyjnym, umożliwiając uzyskanie ponad 25-krotnego przyspieszenia obliczeń dla badanego klastra komputerowego, przy uwzględnieniu wielowątkowości obliczeń na pojedynczym węźle. Dodatkowo, wykorzystanie mieszanej precyzji pozwoliło Autorowi zwiększyć istotnie wartość przyspieszenia do ponad 36-krotnego.
3. Wśród innych wyników cząstkowych wymienić należy opracowanie solwera z mieszaną precyzją dla pojedynczego węzła obliczeniowego, który wykorzystuje metodę iteracyjnego poprawiania rozwiązania układu równań, co umożliwia nie tylko stosunkowo nieznaczne zwiększenia przyspieszenia, lecz przede wszystkim dzięki

zmniejszeniu wykorzystania pamięci o ponad 40% pozwala na analizę znacznie większych zagadnień.

4. Istotnym osiągnięciem Autora jest także praktyczne wykazanie skuteczności i efektywności zaproponowanych algorytmów oraz opracowanych kodów równoległych uruchamianych zarówno na pojedynczym węźle, jak i klastrze komputerowym, w dalece nietrywialnych zagadnieniach analizy MES, związanych bezpośrednio z realnymi zastosowaniami. Należą do nich modelowanie właściwości pianki korundowej oraz kompozytu ceramicznego, przy czym największy obliczony model zawierał ponad 30 mln niewiadomych.

Uzyskane wyniki zostały opublikowane w 4 pracach w języku angielskim, przy czym jedna z nich ukazała się w wysokopunktowanym czasopiśmie z tzw. listy filadelfijskiej, kolejna w czasopiśmie z listy B, a jeszcze jedna w materiałach międzynarodowej konferencji o zasięgu ogólnoświatowym. Świadczy to generalnie w sposób pozytywny o stopniu weryfikacji uzyskanych rezultatów przez międzynarodową społeczność specjalistów zajmujących się rozpatrywaną dziedziną, ze szczególnym uwzględnieniem mechaniki komputerowej.

Z omówienia treści pracy, które przytoczono w punkcie 2 niniejszej recenzji, wynika, iż szczególnie rozdział 2, ale także kolejne rozdziały 3-5 są w dużym stopniu poświecone krytycznemu przedstawieniu stlanu wiedzy w zakresie tematyki pracy, potwierdzając w ten sposób ogólny stan wiedzy w zakresie dyscypliny Informatyka, ze szczególnym uwzględnieniem zagadnień programowania równoległego, w tym: sprzętowych uwarunkowań wydajności aplikacji równoległych, standardów i środowisk programistycznymi, wykorzystania modeli i benchmarków do oceny wydajności aplikacji równoległych. Tę pozytywną ocenę wspiera również poziom wiedzy Autora w zakresie symulacji komputerowych MES oraz optymalizacji kodów numerycznych, ze szczególnym uwzględnieniem algorytmów wykorzystywanych w MES, w tym algorytmów przeznaczonych do rozwiązywania dużych układów równań układów liniowych z rzadkimi macierzami współczynników. Jakość tych rozdziałów nie budzi moich generalnych zastrzeżeń, co nie wyklucza szczegółowych uwag krytycznych, które zawarłem w kolejnym punkcie recenzji. Świadczą one o dużej wiedzy Autora w zakresie tematyki badań, popartej szerokim doświadczeniem praktycznym związanym z wdrożeniem i wykorzystaniem aplikacji równoległych implementujących metodę elementów skończonych. Również moja opinia o bibliografii wykorzystanej w pracy oraz jej kompletności jest generalnie pozytywna, co również nie wyklucza szczegółowych uwag krytycznych.

4. Poprawność pracy i uwagi krytyczne

Poprawność treści pracy nie wzbudza moich istotnych zastrzeżeń, a stwierdzenia w niej zawarte wydają się być godne zaufania, co wynika w szczególności z dosyć szczegółowych uzasadnień, popartych wynikami przeprowadzonych badań

eksperymentalnych, w tym również testów porównawczych. Te ostatnie wykonano dla różnych wariantów zaproponowanych i znanych algorytmów oraz implementacji na architektury równoległe, z uwzględnieniem weryfikacji efektywności opracowanych implementacji nie tylko w przykładach o charakterze typowo testowym (test kostki czy test nieliniowej powłoki), lecz również opisanych w rozdziale 6 przykładach reprezentujących niekwestionowany interes z punktu widzenia aplikacyjnego. Generalnie sposób i jakość przeprowadzenia badań eksperymentalnych stanowi bardzo wartościowy element pracy i zasługuje na podkreślenie.

Jednocześnie Autor nie ustrzegł się pewnych braków i słabości. Wśród uwag o charakterze krytycznym, a po trosze dyskusyjnym, wymienić należy:

1. Rozdział 2.6 Autor zatytułował zdecydowanie zbyt szumnie, gdyż jego zawartość porusza co najwyżej wybrane problemy programowania równoległego i to praktycznie tylko dla modelu obliczeń z pamięcią współdzieloną. Podstawowe problemy programowania równoległego w chwili obecnej mają zdecydowanie bardziej generalny charakter, np. dotyczą minimalizacji narzutów na komunikację, zwiększenie produktywność tworzenia kodów równoległych, uzyskanie przenośności oprogramowania równoległego z uwzględnieniem wydajności itd.
2. Kontynuując ten wątek, chciałem zwrócić uwagę na punkt 1 podsumowania oryginalnych wyników pracy (rozdział 7), gdzie Autor deklaruje, iż w rozdziale 2 **opracował** (sic!) podstawy teoretyczne obliczeń równoległych. Autorowi wypada życzyć większej skromności, gdyż podstawy te zostały w rzeczywistości opracowane przez szerokie grono naukowców z wielu krajów w wyniku badań prowadzonych przez kilka dekad, a Autor po prostu przytoczył rezultaty tych prac.
3. Wykorzystując model Roofline, Autor ograniczył się do zastosowania jego podstawowej wersji, która uwzględnia przepustowość dostępu tylko do pamięci RAM, nie próbując zbadać zachowanie swoich kodów z uwzględnieniem wydajności pamięci podręcznej cache. Ta ostatnia zwykle odgrywa rolę wręcz decydującą w uzyskaniu możliwie największej wydajności kodu. Możliwość jej uwzględnienia w sposób bardzo wygodny dla programisty pojawiła się w standardowych narzędziach programistycznych oferowanych przez firmę Intel i w ostatnim okresie przeprowadzenie takiej głębszej analizy stało się wręcz standardem. Dodam, że korzystanie z tych narzędzi nie tylko daje zdecydowanie większe możliwości, ale też uwalnia od konieczności wykonywania dosyć pracochłonnych i dosyć już rutynowych w chwili procedur i obliczeń niezbędnych do manualnej budowy modelu Roofline nawet w takim okrojonym wariancie, z jakim mamy do czynienia w pracy.
4. Pomimo powyższych uwag, rozdział 2 zawierający wprowadzenie do obliczeń równoległych należy zapisać Autorowi na plus. Niestety zabrakło w pracy podobnego wprowadzenia do algorytmów i oprogramowania wykorzystywanego do rozwiązywania zagadnień MES. Zamiast tego za rozdziałem 2 mamy do czynienia z

rozdziałem dotyczącym udoskonaleniu pewnego konkretnego oprogramowania, praktycznie bez przedstawienia żadnego szerszego kontekstu, analizy możliwych wariantów itd. Musi nam wystarczyć fraza o FEAP jako popularnym oprogramowaniu. Takich popularnych pakietów oprogramowania, które można wykorzystać w różnym stopniu, istnieje naprawdę duża liczba i wybór najbardziej odpowiedniego z nich nie jest wcale prosty. Nawiasem mówiąc, nie znalazłem w pracy wzmianki o takich popularnych i bardzo rozbudowanych kolekcjach oprogramowania dla platform HPC, jak np. Trilinos czy PETSc, które są rozpowszechniane na zasadzie Open Source.

5. Za wyjątkiem prac Fiałko, Autor nie cytuje niestety prac żadnego innego polskiego badacza zajmującego się solwerami dla zagadnień MES. Mam tu w szczególności na myśli grupę prof. Paszyńskiego z AGH, która od lat zajmuje się intensywnie solwerami bezpośrednimi i ma w tym zakresie duże osiągnięcia. Podobna sytuacja występuje z pracami zespołu prof. Schaefera z AGH, dotyczącymi zrównoleglenie solwerów wykorzystujących dekompozycję obszaru oraz wyznaczanie obliczeń Schura.
6. Dodatkowo, w przypadku metody FETI-DP zabrakło mi odwołania do najnowszych prac, np. z ośrodka w Ostrawie, w których metoda ta została znacznie rozwinięta.
7. Autor w swojej pracy wykorzystuje klaster, którego procesory zawierają tylko 6 rdzeni, podczas gdy współczesne procesory serwerowe wykorzystują już ponad 20 rdzeni. Byłoby interesujące zbadanie zachowania opracowanych kodów na takich właśnie platformach.
8. Ostatnia moja uwaga ma charakter zdecydowanie dyskusyjny. Autor w sposób zbyt pobieżny potraktował temat wyboru typu solwera – bezpośredniego czy iteracyjnego, koncentrując się w swoich twórczych poszukiwaniach na zwiększeniu efektywności solwerów bezpośrednich jako preferowanych. Tymczasem sam Autor na koniec rozdziału 5 implementuje znana metodę FETI-DP z solwerem iteracyjnym wykorzystywanym do rozwiązywania równania dla interfejsów, uzyskując znacznie lepsze przyspieszenie w porównaniu z opracowanym przez siebie oryginalnym solwerem bezpośrednim. Już ten wynik wskazuje na znacznie większą perspektywę zwiększenia skalowalności obliczeń, niezbędną dla efektywnego wykorzystania możliwości współczesnych systemów HPC, którą stwarza użycie metod iteracyjnych. Oczywiście ich praktyczne zastosowanie wiąże się z problemami innego rodzaju, np. koniecznością dostosowania preconditionera do właściwości zagadnienia. Tym niemniej, zabrakło mi w pracy takiej uczciwej analizy „za i przeciw".

5. Podsumowanie

Przytoczone wyżej uwagi krytyczne nie umniejszają wysokiej wartości merytorycznej pracy, która stanowi istotny i oryginalny wkład Autora w rozwój metod i algorytmów

organizacji obliczeń równoległych z wykorzystaniem MES, co pozwala ją jednoznacznie zakwalifikować do **dyscypliny Informatyka.**

Podsumowując recenzję, stwierdzam więc, że moja generalna opinia o pracy **„Zastosowanie algorytmów wielowątkowych i rozproszonych do zwiększenia efektywności Metody Elementów Skończonych"** jest zdecydowanie pozytywna. Uważam, że przedstawiona mi do recenzji praca zawiera samodzielne rozwiązanie przez doktoranta ważnego i trudnego problemu naukowego, co w pełni odpowiada wymaganiom stawianym rozprawom doktorskim przez odnośną ustawę o tytule i stopniach naukowych. Na tej podstawie wnioskuję o dopuszczenie pracy do publicznej obrony celem uzyskania przez Autora stopnia doktora nauk technicznych w zakresie informatyki.

Wyrzykowski R.

Prof. dr hab. inż. Roman Wyrzykowski