

Warszawa, 7 lutego 2015

dr hab. inż. Jarosław Arabas, prof. nzw. PW
Politechnika Warszawska
Wydział Elektroniki i Technik Informacyjnych

RECENZJA

rozprawy doktorskiej dla
Rady Naukowej Instytutu Podstaw Informatyki
Polskiej Akademii Nauk

Tytuł rozprawy: **Równoległe i rozproszone algorytmy ewolucyjne**

Autor rozprawy: **mgr Tadeusz Puźniakowski**

1. Jakie zagadnienie naukowe jest rozpatrzone w pracy (teza rozprawy) i czy zostało ono dostatecznie jasno sformułowane przez Autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?

Praca jest poświęcona abstrakcyjnemu sformułowaniu zagadnienia poszukiwania rozbitków na morzu. Technika poszukiwania jest zdefiniowana w postaci grafowej, przy czym Autor rozważa dwie definicje: listę lub drzewo. Reprezentacja techniki poszukiwania jest poddawana ocenie poprzez symulację sposobu poruszania się rozbitków, spowodowanego ruchem wody, i trasy jednostki poszukującej.

Autor podjął próbę rozwiązania tego problemu poprzez konstrukcję specjalizowanego algorytmu ewolucyjnego. Ze względu na długi czas symulacji, koniecznej do oceny strategii poszukiwania, Autor przyspieszył działanie algorytmu wykorzystując m.in. programowalne karty graficzne.

Tytuł nie odpowiada treści pracy – jest zbyt ogólny i zupełnie nie wskazuje na wykorzystanie w zagadnieniu poszukiwania rozbitków..

Układ pracy jest następujący. We wstępie Autor przedstawia najpierw zadanie poszukiwania wielu rozbitków, zapowiadając, że będzie to przykład wiodący. Rozdział 2 jest poświęcony krótkiemu omówieniu standardowych technik zrównoleglania obliczeń: OpenCL, MPI oraz OpenMP. W rozdziale 3 Autor przedstawia wyniki porównania czasu wykonania operacji mnożenia macierzy kwadratowych na podstawie procedur zaimplementowanych w OpenCL oraz OpenMP, porównując czasy wykonania na CPU i GPU. Rozdział 4 prezentuje implementację prostego algorytmu genetycznego w OpenCL. Rozdział 5 stanowi rozszerzenie dyskusji zadania poszukiwania wielu rozbitków, w którym podany jest szereg właściwości teoretycznych reprezentacji strategii poszukiwania. W rozdziałach 6 i 7 prezentowany jest ponownie algorytm genetyczny, tym razem

z wieloma wariantami operatorów genetycznych specjalizowanych dla zadania poszukiwania. Rozdział 8 opisuje wyniki symulacji opracowanego algorytmu ewolucyjnego dla zadania poszukiwania rozbitków. Tekst jest krótko podsumowany w rozdziale 9.

Autor deklaruje, że rozprawa jest poświęcona weryfikacji hipotezy badawczej mówiącej o tym, że równoległa implementacja algorytmu ewolucyjnego, zastosowanego do zadania poszukiwania rozbitków, może działać szybciej niż bez wykorzystania równoległości. W hipotezie tej Autor wspomina również, że poprzez zrównoleglenie obliczeń można przyspieszyć proces generowania efektywnych (podkr. J.Arabas) rozwiązań w/w problemu.

Zatem do pozytywnej weryfikacji tej hipotezy potrzebne będzie wykazanie, że:

1. zaproponowany algorytm daje rozwiązania efektywne,
2. zrównoleglenie przyspiesza proces generowania rozwiązań.

Praca ma charakter studium wykonalności dla problemu mogącego znaleźć praktyczne zastosowanie. Zawarte są w niej elementy konstrukcyjne (własna implementacja) oraz eksperymentalne.

2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł (w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle) świadczącej o dostatecznej wiedzy Autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

W zakresie równoległych algorytmów ewolucyjnych, analiza źródeł nie została przedstawiona w sposób przekonujący. Optyka Autora jest związana przede wszystkim z implementacją równoległą algorytmu ewolucyjnego i skupia się na mechanizmach programistycznych. O istnieniu równoległych algorytmów ewolucyjnych (np. model wyspowy, komórkowy) wspomina się, natomiast nie jest to poparte dogłębną analizą literatury. W szczególności, nie przeprowadzono na podstawie literatury przekonującej analizy topologii równoległe działających AE, nie zanalizowano kwestii nastaw parametrów, intensywności i sposobu wymiany rozwiązań pomiędzy wyspami. Zamiast tego, Autor dokonuje próby samodzielnej weryfikacji tych parametrów, nie odnosząc się do opracowań innych Autorów.

Próby równoległej implementacji algorytmu ewolucyjnego (genetycznego), przedstawione w rozdziale 4, nie pozwalają dostrzec celowości wyboru zaimplementowanych elementów algorytmu – mam wrażenie, że zostały wybrane pierwsze lepsze.

Daje się zauważyć nieznaną Autora metodyki testowania algorytmów ewolucyjnych jako technik optymalizacji. Od co najmniej 25 lat praktykuje się testowanie algorytmów ewolucyjnych za pomocą zadań benchmarkowych, zaś od dziesięciu lat spopularyzowana została technika analizy oparta na wielokrotnych uruchomieniach i analizie statystyki uzyskiwanych rozwiązań (por. CEC 2005-CEC 2014 oraz BBOB 2009-2013). Badania Autora dotyczące losowych metod optymalizacji sprowadzają się do analizy wartości średniej wyników uzyskiwanych z kilku uruchomień metody, co nie spełnia standardów metodologicznych uznanych w literaturze za referencyjne

Praca nie zawiera wyodrębnionej części związanej z przeglądem źródeł. Z tego powodu, przegląd ten jest szczątkowy w odniesieniu do algorytmów ewolucyjnych. W odniesieniu do zadania poszukiwania rozbitków na morzu, przegląd literatury również nie jest imponujący, lecz przynajmniej ilościowo wydaje się bogatszy.

3. Czy Autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Autor podał przykłady związane z rozwiązywaniem zadania poszukiwania rozbitków, w których zrównoleglenie algorytmu ewolucyjnego prowadzi do przyspieszenia obliczeń. Przedstawione wyniki

wydają się świadczyć o obecności tego przyspieszenia, mimo że nie została przeprowadzona ich analiza statystyczna.

Autor nie wykazał, że rozwiązania generowane przez algorytm są efektywnymi rozwiązaniami zadania poszukiwania rozbitków. W szczególności, nie została określona miara pozwalająca na uznanie rozwiązania za efektywne, a także uzyskane wyniki nie zostały odniesione do rozwiązań generowanych przez inne metody optymalizacji.

Autor zapowiedział we wstępie, że wiodącym przykładem jest poszukiwanie rozbitków na morzu, jednak nie stosował się konsekwentnie do tej zapowiedzi. Gdyby rozwiązanie problemu miało być tematem wiodącym, wówczas należałoby rozważyć pewną liczbę starannie wybranych technik optymalizacji, zdalnych do przetwarzania przyjętej reprezentacji rozwiązania. Jeśli jednak akcent miałyby być położony na równoległe i rozproszone algorytmy ewolucyjne, to wypadałoby najpierw doprecyzować rozważania, wskazując na zakres pracy (czy dotyczy zapisu algorytmu, czy jego realizacji, czy możliwości sprzętu obliczeniowego) i nie zawężać jej praktycznej części wyłącznie do przykładu wiodącego. Niepokojące jest także to, że Autor przyznaje w podsumowaniu, że „testowany tutaj Algorytm Genetyczny” [z drzewiastą reprezentacją rozwiązania] „nie spełnia pokładanych w nim nadziei najprawdopodobniej ze względu na niezbyt szczęśliwie dobrane operatory genetyczne oraz duży rozmiar chromosomów”. Dobrze to świadczy o uczciwości Autora, jednak stawia pod znakiem zapytania część pracy związaną z reprezentacją drzewiastą.

4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek Autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Oryginalność rozprawy jest trudna do zdefiniowania, gdyż zależy od przyjętej miary. Z pewnością oryginalne są próby Autora wykorzystania algorytmu ewolucyjnego do zadania poszukiwania rozbitków. Zaprezentowana realizacja równoległa algorytmu genetycznego jest wprawdzie własną, ale typową implementacją typowego algorytmu w typowym środowisku. Wyniki porównania wydajności mechanizmów zrównoleglania są pożyteczne z inżynierskiego punktu widzenia, ale nie nazwałbym ich oryginalnymi, są ponadto dość luźno związane z tematyką rozprawy.

Ze względu na uchybienia warsztatowe związane z testowaniem, a także ze względu na bardzo słabą część związaną z przeglądem literatury, praca jest poniżej poziomu reprezentowanego przez literaturę światową.

5. Czy Autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników (zwięzłość, jasność, poprawność redakcyjna rozprawy)?

Praca nie prezentuje logicznego ciągu, w którym treści poprzedzające są niezbędne do zrozumienia kolejnych. Bez większej szkody dla całości można by usunąć rozdziały 3 i 4. Rozdział 5 począwszy od 5.4 nie wnosi treści wykorzystywanych w pracy. Podrozdział 7.2 pt. „Obsługa załączonego programu” zawiera wielką, nieczytelną i niepotrzebną w tym miejscu tabelę oraz zrzut ekranu interfejsu użytkownika dla programu symulatora. Dalej w rozdziale 7 jest sporo szczegółów technicznych symulatora akcji ratowniczej, których zasadność opisu jest dla mnie niezrozumiała.

Nie rozumiem potrzeby umieszczania fragmentów kodu realizujących poszczególne elementy algorytmu – ich implementacja jest poprawna, ale w pracy doktorskiej powinno być to rzeczą oczywistą. Jeśli Autorowi zależało na wykazaniu, że praca jest poparta implementacją, powinien był ją zawrzeć w formie dodatku do pracy, najlepiej w formie elektronicznej.

Ponadto, praca jest zredagowana w nietypowej formie jak na rozprawę doktorską – Autor czasem wpada w ton pogadankowy.

6. Jakie są słabe strony rozprawy i jej główne wady?

Podstawową słabością formalną pracy jest to, że zawiera liczne i długie wątki poboczne, przez co do końca nie pozwala czytelnikowi stwierdzić, jakiego tematu praca dotyczy.

Podstawową wadą metodologiczną jest sposób testowania algorytmów i analizy ich wyników prezentowany w pracy, który nie pozwala na wyciągnięcie wiarygodnych wniosków. Badane metody optymalizacji są losowe, co wymaga analizy statystycznej, która praktycznie nie została przeprowadzona.

Lektura pracy pozostawia wrażenie, że Autor rozumie i biegle włada warsztatem programistycznym, w tym technikami uwzględniającymi zrównoleglenie procesów obliczeniowych. Jednak uchybienia warsztatu naukowego spowodowały, że recenzowany tekst nie w pełni odpowiada wymogom stawianym rozprawom doktorskim.

7. Jaka jest przydatność rozprawy dla nauk technicznych?

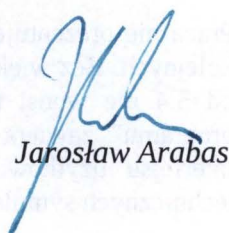
W obecnej chwili trudno ocenić przydatność tego tekstu. Są w nim elementy ciekawe, jak np. sposób rozwiązania zadania poszukiwania za pomocą algorytmu ewolucyjnego, czy też sposób organizacji obliczeń równoległych. Jednak forma pracy tak zaciemnia jej treść, że rzetelna ocena meritum jest bardzo utrudniona.

W mojej ocenie, praktyk zajmujący się zastosowaniami algorytmów ewolucyjnych w zadaniach innych niż poszukiwanie rozbitków, a także zajmujący się szeroko pojętą równoległością w algorytmach ewolucyjnych, nie znajdzie w pracy treści nowych. Przypuszczam natomiast, że w dziedzinie poszukiwań rozbitków praca ma potencjał aplikacyjny.

8. Do której z następujących kategorii Recenzent zalicza rozprawę:

- a) nie spełniająca wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy
- b) wymagająca wprowadzenia poprawek i ponownego recenzowania
- c) spełniająca wymagania
- d) spełniająca wymagania z wyraźnym nadmiarem
- e) wybitnie dobra, zasługująca na wyróżnienie

W mojej ocenie praca wymaga wprowadzenia poprawek i ponownego recenzowania. Wnoszę o niedopuszczenie rozprawy do publicznej obrony i jednoczesną zachętę Autora do dokonania niezbędnych zmian, poprawek i uzupełnień, tak aby przy ponownym recenzowaniu można było bez wątpliwości uznać tekst za rozprawę doktorską.


Jarosław Arabas

dr hab. inż. Jarosław Arabas, prof. nzw. PW
Politechnika Warszawska
Wydział Elektroniki i Technik Informacyjnych

R E C E N Z J A

rozprawy doktorskiej dla
Rady Naukowej Instytutu Podstaw Informatyki
Polskiej Akademii Nauk

Tytuł rozprawy: **Równoległe i rozproszone algorytmy ewolucyjne w poszukiwaniach morskich**

Autor rozprawy: **mgr Tadeusz Puźniakowski**

1. Podstawowe informacje o rozprawie

Opiniowana rozprawa bazuje na wcześniejszej pracy Autora pt. *Równoległe i rozproszone algorytmy ewolucyjne*, której byłem recenzentem w 2015 roku. W konkluzji recenzji sugerowałem wprowadzenie pewnych modyfikacji i uzupełnień. Przedłożona do recenzji rozprawa uwzględnia wszystkie wyrażone przeze mnie wówczas uwagi krytyczne. W szczególności, zmodyfikowany został tytuł rozprawy na bardziej adekwatny do treści, została wprowadzona metodycznie poprawna analiza eksperymentalna zaproponowanego rozwiązania oraz dokonano przeakcentowania wątków tak, aby rozprawa zyskała na spójności tekstu.

2. Jakie zagadnienie naukowe jest rozpatrzone w pracy (teza rozprawy) i czy zostało ono dostatecznie jasno sformułowane przez Autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?

Praca jest poświęcona abstrakcyjnemu sformułowaniu zagadnienia poszukiwania rozbitków na morzu. Technika poszukiwania jest zdefiniowana w postaci grafowej, przy czym Autor rozważa dwie definicje: listę lub drzewo. Reprezentacja techniki poszukiwania jest poddawana ocenie poprzez symulację sposobu poruszania się rozbitków, spowodowanego ruchem wody, i trasy jednostki poszukującej.

Autor podjął próbę rozwiązania tego problemu poprzez konstrukcję specjalizowanego algorytmu ewolucyjnego. Ze względu na długi czas symulacji, koniecznej do oceny strategii poszukiwania, Autor przyspieszył działanie algorytmu wykorzystując m.in. programowalne karty graficzne.

Układ pracy jest następujący. We wstępie Autor przedstawia zadanie poszukiwania wielu rozbitków i stawia hipotezę badawczą. Rozdział 2 jest poświęcony krótkiemu omówieniu standardowych technik zrównoleglania obliczeń: OpenCL, MPI, OpenMP i CUDA. W rozdziale 3 Autor przedstawia wyniki porównania czasu wykonania operacji mnożenia macierzy kwadratowych na podstawie procedur zaimplementowanych w OpenCL oraz OpenMP, porównując czasy wykonania na CPU i GPU. Rozdział 4 zawiera dyskusję zadania poszukiwania wielu rozbitków, w którym zadanie to jest analizowane jako rozszerzenie zadania poszukiwania jednego rozbitka. Szczególny nacisk jest położony na reprezentację rozwiązania w postaci drzewa. W rozdziale 5 prezentowany jest algorytm



ewolucyjny służący do rozwiązywania zadania poszukiwania rozbitków, przy czym szczególny nacisk jest położony na koncepcje operatorów genetycznych dopasowanych do reprezentacji. Szczegóły implementacyjne oraz związane ze środowiskiem symulacji ruchu rozbitków, służącym do oceny rozwiązań generowanych przez algorytm ewolucyjny, podane są w rozdziale 6. Rozdział 7 opisuje wyniki symulacji opracowanego algorytmu ewolucyjnego dla zadania poszukiwania rozbitków. Tekst jest krótko podsumowany w rozdziale 8.

Tekstowi głównemu rozprawy towarzyszą trzy dodatki. W dodatku A przedstawiono wyniki zastosowania algorytmu ewolucyjnego w wersji zrównoleglonej, służącego do rozwiązania problemu OneMax. Dodatek B zawiera instrukcję obsługi załączonego materiału, zaś dodatek C – analizę wartości oczekiwanych prawdopodobieństwa odnalezienia rozbitka dla dwóch strategii poszukiwań (niecierpliwa i bezrefleksyjna), różniących się momentem przejścia do poszukiwań kolejnego rozbitka; analiza ta jest przeprowadzona dla co najwyżej trzech rozbitków.

Rozprawa jest poświęcona weryfikacji hipotezy badawczej mówiącej o tym, że równoległa implementacja algorytmu ewolucyjnego, zastosowanego do zadania poszukiwania rozbitków, może działać szybciej niż bez wykorzystania równoległości. W hipotezie tej Autor wspomina również, że poprzez zrównoleglenie obliczeń można przyspieszyć proces generowania rozwiązań w/w problemu.

Praca ma charakter studium wykonalności dla problemu mogącego znaleźć praktyczne zastosowanie. Zawarte są w niej elementy konstrukcyjne (własna implementacja) oraz eksperymentalne.

3. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł (w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle) świadczącej o dostatecznej wiedzy Autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

Praca nie zawiera wyodrębnionej części związanej z przeglądem źródeł. W odniesieniu do zadania poszukiwania rozbitków na morzu, przegląd literatury wydaje się być wyczerpujący. W odniesieniu do algorytmów ewolucyjnych, Autor przyjął zapewne, że literatura jest zbyt bogata aby ją wyczerpująco przeanalizować. Skupił się zatem na dwóch pozycjach podstawowych – [Gol89] i [Mic96]. W tym „towarzystwie” zdecydowanie odstaje pozycja [Gwi07]. Więcej uwagi Autor poświęcił zagadnieniu równoległości w algorytmach ewolucyjnych, zarówno na poziomie algorytmicznym, jak i programistycznym. Przegląd literatury uzasadnia podjęcie tematyki pracy.

4. Czy Autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Autor podał przykłady związane z rozwiązywaniem zadania poszukiwania rozbitków, w których zrównoleglenie algorytmu ewolucyjnego prowadzi do przyspieszenia obliczeń. Przedstawione wyniki świadczą o obecności tego przyspieszenia.

Jako miarę efektywności poszukiwań Autor przyjął prawdopodobieństwo odnalezienia rozbitka, wyznaczone na drodze symulacji. Wyniki przeprowadzonych eksperymentów świadczą o tym, że rozwiązania generowane przez algorytm ewolucyjny z reprezentacją drzewiastą prowadzą do wyników lepszych pod względem przyjętej miary niż przy reprezentacji sekwencyjnej. Ponadto Autor dokonał prób użycia innych technik optymalizacji, w szczególności lokalnych metod wspinaczkowych, w celu ulepszenia rozwiązań wygenerowanych ewolucyjnie. Ulepszenia nie stwierdzono, co świadczy o tym, że ewolucyjnie uzyskane rozwiązania są przynajmniej lokalnie optymalne. Wnioski te są wyciągnięte na bazie analizy statystycznej wielu niezależnych uruchomień i przekonująco uzasadniają hipotezę badawczą

5. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek Autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

W mojej ocenie oryginalne jest ujęcie poszukiwania wielu rozbitków jako zadanie inne niż proste złożenie wielu zadań poszukiwania pojedynczego rozbitka. Ponadto oryginalne są próby Autora wykorzystania algorytmu ewolucyjnego do zadania poszukiwania rozbitków. Uzyskane wyniki eksperymentów świadczą o tym, że praca zawiera nie tylko wyniki oryginalne, ale też wartościowe.

Praca pod względem zawartości nie odstaje od poziomu reprezentowanego przez literaturę światową.

6. Czy Autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników (zwięzłość, jasność, poprawność redakcyjna rozprawy)?

Układ treści pracy jest logiczny. Pewnym zaburzeniem tej logiki jest rozdział 3, będący w dużej mierze rozbudowaną dygresją mającą przekonać czytelnika o potencjalnych korzyściach związanych z zastosowaniem GPU. Być może moja ocena niewielkiej przydatności tego rozdziału dla całości pracy wynika z mojego przekonania o korzyściach stosowania GPU, jednak brakuje mi jednocześnie w tym rozdziale krytycznego podejścia do doskonałych wyników przyspieszenia obliczeń, będących jednak pochodną faktu, że mnożenie macierzy jest zadaniem dość dobrze dopasowanym do specyfiki GPU ze względu na lokalność i identyczność operacji.

Styl narracji rozprawy jest niejednorodny. Z jednej strony, można doszukać się typowego stylu rozprawy doktorskiej, gdzie stawiana jest hipoteza, po czym przedstawiane są przesłanki przemawiające za jej prawdziwością, co kończy się pozytywną konkluzją. Z drugiej strony, na poziomie rozdziałów widać tendencję Autora do prowadzenia pamiętnika wykonanych prac. Wrażenie pozostałości takiego stylu sprawia rozdział 3, który merytorycznie niewiele wnosi do badanego zagadnienia, jednak robi wrażenie pozostawionego ze względów sentymentalnych.

Strona językowa pracy pozostawia pewien niedosyt. Autor opowiada o historii badań nierzadko w pierwszej osobie, co wzmaga wrażenie pamiętnika, również terminologia jest traktowana dość luźno (np. wymiennie stosowane terminy: „funkcja przystosowania”, „funkcja oceny”, „funkcja fitness”), zdarzają się również nieliczne błędy ortograficzne. Usterki te trochę uprzykrzają lekturę, lecz nie zaciemniają zbyt mocno treści rozprawy

7. Jakie są słabe strony rozprawy i jej główne wady?

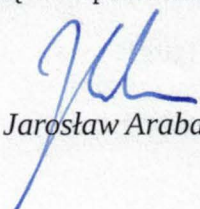
Do słabych stron pracy zaliczam skąpy przegląd literatury dotyczący zagadnień równoległych obliczeń w algorytmach ewolucyjnych oraz formę i język tekstu.

8. Jaka jest przydatność rozprawy dla nauk technicznych?

W obecnej chwili trudno ocenić jej przydatność. Uzyskane wyniki budzą nadzieję, że przyczynić się ona może do poprawy praktyki poszukiwań wielu rozbitków na morzu. Jednak należy się spodziewać konieczności długotrwałego testowania w różnych warunkach, a także przewyciężenia problemów natury zarówno etycznej, jak i formalnej, których rozwiązanie będzie niezbędne w celu upowszechnienia wyników pracy.

9. Konkluzja

Mimo usterek praca spełnia wymagania stawiane rozprawom doktorskim. Wnoszę o dopuszczenie Doktoranta do publicznej obrony.


Jarosław Arabas