

Prof. dr hab. inż. Zbigniew Michalewicz  
Instytut Podstaw Informatyki  
Polska Akademia Nauk  
ul. Jana Kazimierza 5  
01-248 Warszawa

Warszawa, 12.12.2014

## **RECENZJA**

*rozprawy doktorskiej mgr Tadeusza Puźniakowskiego pt. „Równoległe i Rozproszone Obliczenia Ewolucyjne”*

### **1. Problematyka naukowa oraz przedmiot rozprawy**

Doktorant w swojej rozprawie zaproponował oryginalny algorytm genetyczny (do problemu wyszukiwania rozbitków na morzu) ze szczególną reprezentacją osobników; eksperymenty zostały przeprowadzone przy użyciu równoległych i rozproszonych obliczeń.

Tematyka badawcza pracy dotyczyła wykorzystania dostępnych metod zrównoleglania obliczeń do przyspieszenia procesu generacji efektywnych strategii poszukiwań wielu rozbitków za pomocą algorytmów genetycznych prowadząc obliczenia równoległe i/lub rozproszone na powszechnie dostępnym sprzęcie. Hipoteza badawcza zakładała, że takie obliczenia można znacznie przyspieszyć.

Podstawowy cel pracy, którym jest weryfikacja przedstawionej hipotezy badawczej, został zrealizowany poprzez wykonanie dwóch następujących zadań:

1. Porównanie wydajności oferowanych przez rozwiązania wykorzystujące heterogeniczną architekturę sprzętowa z rozwiązaniami tradycyjnymi (opartymi na CPU), oraz
2. Określenie kosztu prac programistycznych niezbędnych do adaptacji oprogramowania do efektywnej pracy z heterogenicznymi architekturami sprzętowymi.

Postawiony w rozprawie cel został ściśle określony oraz w pełni osiągnięty. Rozważane w rozprawie zagadnienia są aktualne i w istotny sposób wpisują się w obszar informatyki. Rozprawa może zatem być przedstawiona jako monografia doktorska w dziedzinie informatyki.

### **2. Analiza treści rozprawy oraz uzyskanych wyników**

Rozprawa została przygotowana w języku polskim i składa się z 8 rozdziałów, podsumowania, bibliografii obejmującej 78 pozycji literaturowych, oraz jednego dodatku.

Całość pracy obejmuje 185 stron i ma ona charakter eksperymentalny (z teoretycznym dodatkiem).

Rozdział 1 stanowi wprowadzenie do problematyki rozprawy. Doktorant przedstawia dwa architekuralnie odmienne warianty realizacji równoległości: CPU oraz GPU. Rozdział ten omawia również problem praktyczny (problem poszukiwania wielu rozbitków), hipotezę badawczą, oraz organizację pracy. Sekcja związana z problemem poszukiwania wielu rozbitków przedstawia podstawy ratownictwa morskiego jak również metody poszukiwań rozbitka na morzu.

Rozdział 2 omawia podstawowe technologie pozwalające na zrównoleglenie obliczeń: OpenMP (który udostępnia programistom interfejs programowania aplikacji, umożliwiając tworzenie programów komputerowych zarządzających wątkami dla systemów wieloprocesorowych z pamięcią współdzieloną), MPI (wieloplatformowy protokół komunikacyjny służący do zrównoleglenia obliczeń na klastrach i superkomputerach), oraz OpenCL (narzędzie pozwalające tworzenie oprogramowania dającego się wykonać na różnym sprzęcie).

Rozdział 3 opisuje eksperyment porównujący wydajność różnych metod zrównoleglenia obliczeń. W szczególności autor porównuje OpenCL z OpenMP pod względem wydajności na tym samym CPU oraz bada wielkość straty wydajności obliczeń przy wykorzystaniu pamięci lokalnej.

W rozdziale 4 przedstawiona jest pierwsza implementacja algorytmu genetycznego za pomocą OpenCL. Dla problemów z prostą funkcją oceny i dla takiej implementacji autor przedstawił bardzo dobre wyniki, ale również ta sama implementacja dla wiodącego przykładu (poszukiwanie rozbitków) z trudną funkcją oceny pokazała szereg problemów związanych z tym podejściem. Autor więc zrównoleglił funkcję oceny oraz generowanie tras rozbitków.

Rozdział 5 opisuje proponowane podejście do kwestii poszukiwania wielu obiektów z wykorzystaniem strategii generowanych offline, gdzie strategie poszukiwań są opisywane jako drzewa. Rozdział ten zawiera również analizę złożoności pamięciowej zaproponowanej reprezentacji oraz opis metody zapisu strategii jako ciągłej struktury danych.

Rozdział 6 omawia dokładnie program który został potem użyty do ostatecznych testów. Autor podaje opis działania programu z punktu widzenia użytkownika, szczegóły implementacyjne, oraz organizację kodu. Cały program został załączony do rozprawy na nośniku danych.

Rozdział 7 omawia natomiast implementację algorytmu genetycznego dla proponowanej reprezentacji strategii poszukiwań, ze szczególnym uwzględnieniem opisu operatorów genetycznych użytych w algorytmie. Przedyskutowane jest również działanie funkcji oceny, które jest oparte na symulacji.

Rozdział 8 zawiera opis wszystkich eksperymentów (i ich wyników) przeprowadzonych przez autora. Praca kończy się (rozdział 9) krótkim podsumowaniem wyników osiągniętych w rozprawie.

### **3. Najistotniejsze osiągnięcia przedstawione w rozprawie**

Rozprawa doktorska mgr Tadeusza Puźniakowskiego zawiera wyniki dotyczące przyspieszania procesu generowania strategii do problemu znajdowania rozbitków przy implementacji algorytmu genetycznego i przy użyciu równoległych i rozproszonych obliczeń. Do najistotniejszych osiągnięć rozprawy zaliczyć należy:

- a) Porównanie technik zrównoleglania, z którego wynika, że wykorzystanie OpenMP pozwala na trzykrotne przyspieszenie obliczeń ale nie daje wzrostu wydajności obliczeń na kartach graficznych.
- b) OpenCL pozwala na siedmiokrotne przyspieszenie obliczeń. A wykonanie tych samych obliczeń na kartach graficznych daje przyspieszenie 40-krotne (karta graficzna AMD) i prawie 80-krotne (Tesla – dedykowana karta GPGPU).
- c) W przykładzie szukania rozbitków wyodrębniono nową klasę wielomianowej wielkości drzewiastych struktur danych opisujących rozwiązania, otrzymując drzewiaste chromozomy.
- d) Udowodniono, że w teorii struktury drzewiaste powinny dawać kilkunasto-procentową poprawę jakości rozwiązań przy tej samej długości strategii.

### **4. Uwagi merytoryczne**

W trakcie czytania rozprawy doktorskiej nasuwają się pewne uwagi o charakterze dyskusyjnym. Są to:

- Tutuł pracy jest zbyt ogólny – sugeruje on systematyczne badania nt. równoległych i rozproszonych obliczeń ewolucyjnych. Tymczasem doktorant koncentruje się bardziej na reprezentacjach strategii czy zaprojektowaniu algorytmu genetycznego dla wiodącego przykładu. Badania nt. równoległych i rozproszonych obliczeń ewolucyjnych są przeprowadzone tylko albo na prostej funkcji celu albo na wiodącym przykładzie – co uniemożliwia rozsądne uogólnienie otrzymanych wniosków.
- Prezentowane wyniki są oparte na eksperymentach z systemami o dużej liczbie parametrów (np. liczby węzłów klastrów, liczby migrantów, częstotliwość migracji, wielkość populacji, dobór operatorów genetycznych i ich współczynników) – jak również inne parametry (np. związane z pojęciem “znalezienia rozbitka”). Być może, że dla innego doboru parametrów wyniki byłyby inne. Praca przedstawia metodologię doboru parametrów algorytmu genetycznego (sekcja 8.1) i wszystkie prezentowane wyniki są oparte na wartościach tych parametrów, tzn. wyniki porównania wydajności implementacji CPU, OpenMP oraz GPU są podane dla ustalonych wartości parametrów.

- Tak naprawdę to ta praca doktorska opisuje nowe podejście do problemu znajdowania rozbitków – a podejście to jest oparte na pewnej implementacji algorytmu genetycznego ze szczególną reprezentacją osobników i liczone jest przy użyciu równoległych i rozproszonych obliczeń.

## 5. Uwagi redakcyjne i edytorskie

Rozprawa napisana jest niezbyt precyzyjnym językiem; redakcja pracy też budzi pewien niedosyt. Podaję poniej *kilka* przykładów:

- Pierwszy paragraf na stronie 9 wprowadza pojęcie długości trasy – definicja nie jest jasna, nie uwzględnia odległości pomiędzy węzłami, i nie wyjaśnia, dlaczego?
- Drugi paragraf na tej samej stronie wprowadza pojęcie “bieżącego akwenu” bez żadnych wyjaśnień, co to może być? W tym samym paragrafie autor pisze o “znalezieniu” rozbitka, bez żadnego wyjaśnienia, jak się rozbitka znajduje (np. czy trzeba być w jakimś *epsilon* otoczeniu od rozbitka? Jaki jest ten *epsilon*?)
- Strona 10, zdanie poprzedzające opis zadań 1 oraz 2 jest nieczytelne; początek tego zadania powinien zostać przepisany: “Podstawowy cel pracy, którym jest weryfikacja przedstawionej hipotezy badawczej, odbywać się będzie poprzez realizację następujących...”
- Strona 20: w dwóch różnych miejscach na tej stronie mamy wyrazy: “zrównoleglania” i “zrównoleglenia”. Czy tak ma być?
- Strona 20, drugi paragraf: co autor rozumie poprzez wyrażenie: “na bardzo różnym sprzęcie”? Jaka jest różnica pomiędzy “różnym sprzętem” a “bardzo różnym sprzętem”?
- Wiele zdań w pracy jest zbyt nieformalnych, np. “Cały czas jednak tworzyłem biorąc pod uwagę...” (strona 21), “Nie byłoby w tym nic złego...” (strona 23), “Chciałbym także...” (strona 71), “prawie nigdy”, “prawie zawsze” (strona 75), i wiele innych.
- Strona 21 – początek zdania: “Rozdział 7 poprzedzający eksperymenty...” nie jest zbyt dobry; rozdział nie może poprzedzać eksperymentów, tylko inny rozdział, który te eksperymenty opisuje.
- Jest w pracy sporo literówek: Pierwsze zdanie sekcji 1.1: “zajmmiemy się”... “Cluds” (strona 34), “crossoverPorb” (strona 133), “przyspieszyć” (strona 162), itd.
- Co to jest “niniejsza część pracy”? (strona 35) – paragraf, sekcja, rozdział?
- Tytuł rozdziału 4 wprowadza skrót “AG” bez żadnego wyjaśnienia.
- Funkcja fitness osobnika (strona 56, wzór 4.2) powinna być:  $f(a)$  a nie  $f a$ .
- Strona 61 – wyrażenie “zasada ruletki” nie jest wyjaśnione i jest bez sensu w kontekście zdania.
- Strona 61 – wyraz “Manual” został pominięty w tytule pracy [47].
- Strona 70, drugi paragraf od dołu: wyrażenie “jego konca” nie ma sensu – czyjego końca?
- Strona 71 – dlaczego liczba węzłów, które mogą zostać odwiedzone podczas realizacji strategii jest z założenia stała? Czyżby odległości pomiędzy węzłami nie miały żadnego znaczenia?

- p) Strona 73 – wektor  $c$  powinien być zapisany inaczej (jak na końcu strony 71). W odniesieniu do Twierdzenia 2 warto dodać, gdzie to Twierdzenie można znaleźć. Wyraz “Konstrukcyjnie” nie pasuje w zdaniu.
- q) Strona 91 – brakuje jednej ze składowych algorytmu genetycznego: inicjalizacji. (Pierwsze zdanie sekcji 6.4 omawia generowanie strategii początkowej).
- r) Strona 92 – algorytm genetyczny jest pisany raz z dużych, raz z małych liter. Autor nie wyjaśnił dobrze różnic pomiędzy funkcjami celu, kosztu, zysku, przystosowania, fitness (wszystkie te nazwy występują na stronie 93).
- s) Pierwsze zdanie sekcji 6.2.1. “Złożoną” czy “Kosztowną”?
- t) Pierwsze zdanie drugiego paragrafu sekcji 6.4 jest niejasne: mówi o funkcji która przyjmuje jako jeden z parametrów inną funkcję – nie jest to dobrze wyjaśnione.
- u) Strona 105 – co to znaczy “środek trasy”? Co to znaczy, że on musi być w obszarze największego prawdopodobieństwa odnalezienia rozbitka: Gdziekolwiek w takim obszarze czy też w środku (cokolwiek to znaczy) tego obszaru?
- v) Sekcja 6.4.2. Zdanie przed rysunkiem 6.5 nic nie wnosi – każdy iteracyjny algorytm poprawiający rozwiązania poprawi (albo zostawi) rozwiązanie początkowe – nie więc dziwnego, że algorytm genetyczny poprawia jakość rozwiązań tradycyjnych, które są obecne w populacji początkowej.
- w) Sekcja 6.6, pierwsze zdanie – nie rozdział 7.5, ale sekcja 7.5.
- x) Pierwsza linia strony 126 – pusta “bullet”.
- y) Parę linii przed sekcją 7.5.1 – nie rozdział 6.6, ale sekcja 6.6.
- z) Strony 146 oraz 162 – dodatek jest nazywany dodatkiem (146) albo aneksem (162).
- aa) Przedostatni jedno-zdaniowy paragraf na stronie 163 nie jest zdaniem.
- bb) Prezentacja bibliografii jest niedbała:
  - i. Pozycja [1] jest opisana niedokładnie.
  - ii. Pozycje [10] – [12] powinny mieć adresy stron internetowych.
  - iii. Pozycja [13] jest opisana niedokładnie (wyraz “pages”).
  - iv. Pozycja [15] jest opisana niedokładnie.
  - v. Opis większości referencji konferencyjnych (np. [18] czy [19]) jest niedokładny – gdzie odbyła się konferencja? Kto wydał materiały konferencyjne?
  - vi. Pozycja [21] jest opisana niedokładnie.
  - vii. Pozycja [25] jest opisana niedokładnie.
  - viii. Pozycja [37] jest opisana niedokładnie.
  - ix. Pozycja [38] jest opisana niedokładnie.
  - x. Pozycja [48] jest opisana niedokładnie.
  - xi. Pozycja [49] jest opisana niedokładnie – jakie to czasopismo?
  - xii. Pozycja [52] jest opisana niedokładnie.
  - xiii. Pozycja [53] jest opisana niedokładnie.
  - xiv. Pozycja [54] jest opisana niedokładnie.
  - xv. Pozycja [55] jest opisana niedokładnie. Bałtyk pisze się z dużej litery.
  - xvi. Pozycja [59] jest opisana niedokładnie.
  - xvii. Jaka jest różnica pomiędzy [62] oraz [63]?

- xviii. Pozycje [64] oraz [65] – autor powinien pisać swoje nazwisko polską czcionką jak w innych miejscach pracy (np. [19]).
- xix. Pozycja [67] jest opisana niedokładnie.
- xx. Pozycja [69] jest opisana niedokładnie – numer tomu?
- xxi. Pozycja [70] jest opisana niedokładnie.
- xxii. Pozycja [74] jest opisana niedokładnie.
- xxiii. Pozycja [75] jest opisana niedokładnie.
- xxiv. Pozycja [77] jest opisana niedokładnie – wydawca?

## 6. Podsumowanie

Powyżej przedstawione uwagi merytoryczne oraz redakcyjne nie mają istotnego wpływu na jakość i wagę przedstawionych rozwiązań i w żadnym stopniu nie obniżają wartości pracy. Przedstawione przez autora badane aspekty zostały ujęte wystarczająco szczegółowo i dokładnie. Reasumując można stwierdzić, że główne wyniki rozprawy potwierdzają osiągnięcie z powodzeniem założonego w rozprawie celu. Wyniki te były prezentowane na kilku krajowych i międzynarodowych konferencjach (publikowane w serii LNCS wydawnictwa Springer).

Podsumowując, stwierdzam, że przedstawiona do oceny rozprawa doktorska mgr Tadeusza Puźniakowskiego pt. *„Równoległe i Rozproszone Obliczenia Ewolucyjne”* spełnia wymagania stawiane rozprawom doktorskim przez obowiązującą ustawę o stopniach i tytule naukowym. W konsekwencji, może ona stać się przedmiotem publicznej obrony. Wnoszę zatem o dopuszczenie mgr Tadeusza Puźniakowskiego do dalszych etapów przewodu doktorskiego.

Zbigniew Michalewicz

Prof. dr hab. inż. Zbigniew Michalewicz  
Instytut Podstaw Informatyki  
Polska Akademia Nauk  
ul. Jana Kazimierza 5  
01-248 Warszawa

Warszawa, 30.08.2016

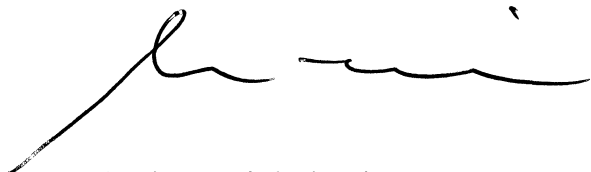
## RECENZJA

*poprawionej rozprawy doktorskiej mgr Tadeusza Puźniakowskiego pt.  
„Równoległe I Rozproszone Obliczenia Ewolucyjne w Poszukiwaniach  
Morskich”*

Wszystkie przedstawione uwagi merytoryczne oraz redakcyjne przedstawione w mojej poprzedniej recenzji tej pracy (z dnia 12 grudnia 2014) zostały poprawione przez autora w zadowalający sposób.

Przedstawione przez autora badane aspekty zostały ujęte wystarczająco szczegółowo i dokładnie. Reasumując można stwierdzić, że główne wyniki rozprawy potwierdzają osiągnięcie z powodzeniem założonego w rozprawie celu. Wyniki te były prezentowane na kilku krajowych i międzynarodowych konferencjach (publikowane w serii LNCS wydawnictwa Springer).

Podsumowując, stwierdzam, że przedstawiona do oceny nowa wersja rozprawy doktorskiej mgr Tadeusza Puźniakowskiego pt. „*Równoległe i Rozproszone Obliczenia Ewolucyjne w Poszukiwaniach Morskich*” spełnia wymagania stawiane rozprawom doktorskim przez obowiązującą ustawę o stopniach i tytule naukowym. W konsekwencji, może ona stać się przedmiotem publicznej obrony. Wnoszę zatem o dopuszczenie mgr Tadeusza Puźniakowskiego do dalszych etapów przewodu doktorskiego.



Zbigniew Michalewicz