

Częstochowa, dn. 16.07.2020

Prof. dr hab. inż. Roman Wyrzykowski  
Katedra Informatyki  
Politechnika Częstochowska  
ul. Dąbrowskiego 69  
42-201 Częstochowa  
roman@icis.pcz.pl



**RECENZJA**  
**ROZPRAWY DOKTORSKIEJ**  
**mgr inż. Jędrzeja Marszałkowskiego**  
**“Scheduling divisible computations with energy constraints”**

**Promotor: Prof. dr hab. inż. Maciej Drozdowski**  
**Wydział Informatyki**  
**Politechnika Poznańska**

## **1. Problem badawczy i jego znaczenie**

Przez dziesięciolecia wydajność była siłą napędową obliczeń wielkiej skali (HPC). Jednak w ostatnich latach pobór mocy i energii w środowiskach komputerowych stał się bardzo istotnym ograniczeniem. Koszty operacyjne oraz utrzymania są obecnie stawiane na równi z kosztami nabycia systemów HPC. Efektywność energetyczna staje się poważnym problemem w projektowaniu nowoczesnych wysokowydajnych systemów obliczeniowych. W ciągu ostatnich kilku lat wykazano, że zwiększony pobór mocy ma negatywny wpływ na rozwój skalowalności systemów komputerowych, a co za tym idzie, również na ich wydajność uzyskiwaną dla aplikacji użytkowników. Nie taka już rzadka staje się sytuacja, gdy niektóre procesory nie są szybsze, ale jedynie bardziej energooszczędne niż ich poprzednicy. Trend ten rozpoczął się na rynku procesorów mobilnych, ale wkrótce pojawił się również w platformach obliczeniowych wielkiej skali.

Środowiska naukowe starają się rozwiązać ten problem na wiele sposobów, czego przykładem może być stosowanie rozmaitych technik poprawiających poziom zużycia energii w aplikacjach. Zastosowanie dynamicznego skalowania napięcia i częstotliwości (DVFS), selektywne wyłączanie serwerów znajdujących się w stanach bezczynności lub nie w pełni wykorzystywanych, a także wykorzystanie energooszczędnych dyspozytorów zadań, to obecnie bardzo popularne tematy wielu prac naukowych. Głównym warunkiem niezbędnym do realizacji tych technik jest

przede wszystkim dokładne monitorowanie zużycia energii i zarządzanie różnymi podzespołami serwerów, a szczególnie procesorami.

Znaczna część literatury wykazuje, że zużycie energii oraz wykorzystanie zasobów sprzętowych są silnie ze sobą powiązane. W związku z tym niektóre z prac mają na celu zmniejszenie wykorzystania zasobów w celu zaoszczędzenia energii, podczas gdy inne próbują osiągnąć równowagę pomiędzy wykorzystaniem zasobów i zużyciem energii. W rezultacie powszechnie przyjętym stwierdzeniem jest, że zużycie energii i wydajność aplikacji są nie tylko ściśle powiązane, lecz ich relacje są często sprzeczne i złożone, a wzajemna zależność jest nieliniowa. Znany jest fakt, iż nawet przyjęty model programowania równoległego ma wpływ na oszczędność energii w systemie wielordzeniowym.

Zrozumienie zjawiska korelacji między wydajnością i zużyciem energii wymaga głębokiej analizy aplikacji z uwzględnieniem ich właściwości (skalowalności, intensywności obliczeniowej, operacji na pamięci, itp.). To sprawia, że cały proces optymalizacji efektywności energetycznej staje się bardzo skomplikowany, szczególnie biorąc pod uwagę fakt, że nowoczesne architektury procesorów wydają się być z natury niezbalansowane pod względem teoretycznej maksymalnej wydajności i przepustowości pamięci. Każdy skuteczny proces optymalizacji wymaga przede wszystkim głębokiego zrozumienia analizowanych zjawisk oraz dokładnego opisu matematycznego samego zjawiska, tj. modelu.

W szczególności, problemy dotyczące zbyt dużego zapotrzebowania na energię systemów komputerowych doprowadziły do wprowadzenia architektur wielo- i masywnie wielordzeniowych. Ważnym krokiem w tym kierunku stało się również coraz szersze wykorzystanie systemów obliczeniowych o architekturze heterogenicznej, w skład których wchodzi kilkupoziomowa struktura łącząca urządzenia o różnych charakterystykach, np. procesory CPU ogólnego przeznaczenia i różnego typu akceleratory obliczeniowe. Hybrydowe aplikacje zdolne do optymalnej utylizacji pełnej mocy obliczeniowej takich systemów muszą jednocześnie wykorzystywać równoległość na wielu poziomach. W ostatnim okresie coraz bardziej popularne stają się też architektury specjalizowane, dedykowane do określonej dziedziny zastosowań (ang. Domain Specific Accelerators – DSA). Jest to szczególnie widoczne w obszarze uczenia maszynowego i sztucznej inteligencji, gdzie swoistym przełomem było opracowanie w 2015 r. przez firmę Google układu TPU (Tensor Processing Unit), przeznaczonego do realizacji fazy inferencji procesu głębokiego uczenia maszynowego.

Recenzowana praca mgr inż. Jędrzeja Marszałkowskiego dotyczy wszechstronnego zbadania możliwości opracowania metodologii efektywnego zarządzania wykonaniem ważnej klasy obliczeń/aplikacji równoległych, wykorzystywanych do przetwarzania dużych ilości danych, z uwzględnieniem dwóch podstawowych kryteriów jakości – nie tylko czasu wykonania obliczeń, lecz również zużycia energii przez wykorzystywane zasoby obliczeniowe. Zaproponowane w pracy synergiczne podejście do problemu

oparte na zastosowaniu zaproponowanej koncepcji map izoenergetycznych i uwzględniające architekturę systemu komputerowego oraz kluczowe parametry systemu i parametry wykonania aplikacji, a także charakteryzujące się wykorzystaniem odpowiednio dobranych metod heurystycznych i dokładnych algorytmów optymalizacji do wyznaczania rozwiązań formułowanych zagadnień, korzystnie świadczy o przygotowaniu Autora do działalności badawczej. Bardzo pozytywnie należy również ocenić dążenie kandydata do przyczynienia się do lepszego zrozumienia samej natury złożonych relacji pomiędzy wydajnością obliczeń i efektywnością energetyczną. Tak jasno sformułowany aspekt poznawczy pracy nie występował niestety zbyt często w recenzowanych przez mnie rozprawach.

Podsumowując, kierunek badań, wybór problematyki rozprawy oraz jej tezy i celów zaproponowane i zrealizowane przez mgra inż. Jędrzeja Marszałkowskiego oceniam zdecydowanie pozytywnie. Lokują się one korzystnie w nakreślonej wyżej tematyce współczesnej informatyki, dotyczącej aspektów teoretycznych efektywnej wydajnościowo i energetycznie organizacji obliczeń równoległych oraz ich zastosowań praktycznych do przetwarzania dużej ilości danych, definiując przede wszystkim teoretyczny, ale jednocześnie praktyczny charakter rozprawy.

## **2. Wkład autora**

Recenzowana praca doktorska obejmuje formalnie 7 rozdziałów, spis zastosowanych oznaczeń, streszczenie w języku polskim oraz stosowną bibliografię zawierającą 101 pozycji. Rozprawa liczy łącznie 160 stron.

W *rozdziale pierwszym* Autor zawarł motywację i cele badań oraz próbę szkicowego opisu zastosowanej metodologii badawczej, a także przedstawił treść pracy zawartą w kolejnych jej rozdziałach. Chciałbym zwrócić uwagę, iż opis celów pracy faktycznie wiąże się z wypunktowaniem najważniejszych oryginalnych osiągnięć doktoranta i wypunktowanie to rysuje się już tutaj w sposób dostatecznie interesujący dla specjalistów zajmujących się obszarem badań związanych z efektywnością energetyczną współczesnych systemów i aplikacji teleinformatycznych.

*Kolejny rozdział* poświęcono scharakteryzowania stanu prac w zakresie tematyki badań. W szczególności, w rozdziale tym Autor stara się wprowadzić czytelników do zagadnień zużycia energii w obliczeniach równoległych, w tym relacji pomiędzy wydajnością czasową obliczeń a ich efektywnością energetyczną. Relacje te mają dostatecznie złożony charakter i są przedmiotem intensywnych badań od ponad 10 lat. Literatura przedmiotu jest w związku z tym niesłychanie obszerna i trudno wymagać od Autora kompleksowego jej przedstawienia. Tym niemniej wydaje się, że potraktowano ją w sposób zbyt pobieżny, o czym jeszcze wspomnę w rozdziale recenzji poświęconym uwagom krytycznym. Natomiast nie mam zarzutów do sposobu

przedstawienia pozostałych kluczowych zagadnień, takich jak teoria obliczeń jednorodnie podzielnych (ang. divisible load theory) wraz ze strukturami komunikacyjnymi typu gwiazda, odpowiadającymi faktycznie powszechnie znanemu schematowi przetwarzania typu master-worker, czy też koncepcja izolinii oraz izomap jako wizualnych narzędzi wspierających analizę wydajności obliczeń równoległych.

**Rozdział trzeci** rozpoczyna w pełni oryginalną część dysertacji, prezentując w pierwszej kolejności zbudowane przez doktoranta stanowisko badawcze służące do pomiaru zużycia energii oparte na wykorzystaniu watomierza o deklarowanej rozdzielczości 1W w pomiarze mocy i rozdzielczości czasowej 1s. Pozwoliło ono na przeprowadzenie testów dla pięciu różnych aplikacji (quicksort, wyszukiwanie ciągu w tekście, obliczanie skrótu md5, wykrywanie krawędzi na bitmapach oraz transponowanie macierzy) uruchamianych na czterech różnych komputerach, co prawda dosyć chyba już wiekowych. Przeprowadzone testy pozwoliły sformułować stosowane w pracy modele dla czasu obliczeń i zużycia energii w zależności od wielkości porcji danych (obciążenia) z uwzględnieniem budowy pamięci systemu równoległego, tj. nie tylko dla pamięci jednorodnej typu flat, lecz również pamięci hierarchicznej. W pierwszym przypadku zużycie energii jest stałe i nie zależy od wielkości porcji, zaś w drugim zależność czasu obliczeń i zużycia energii od wielkości porcji może być opisana przez funkcje kawałkami liniowe.

Sformułowane w ten sposób założenia pozwoliły na zaproponowanie i wykorzystanie w kolejnych trzech rozdziałach autorskiej koncepcji map izoenergetycznych jako naturalnego rozszerzenia znanych wcześniej z literatury map izoefektywności związanych z metrykami wydajności obliczeń. Koncepcja ta pozwala na analizę relacji pomiędzy parametrami determinującymi zużycie energii w obliczeniach jednorodnie podzielnych, wykrycie stopnia ważności poszczególnych parametrów, a także ukierunkowane poszukiwanie możliwości zmniejszenia tego zużycia. W szczególności, w **rozdziale czwartym** skupiono się na ich zastosowaniu do badania zagadnienia zużycia energii w obliczeniach równoległych w modelu z nieograniczoną pamięcią jednorodną, co najczęściej odpowiada sytuacji, gdy wielkość porcji pozwala nie uwzględniać hierarchii pamięci systemu. Wykorzystanie map izoenergetycznych nawet w tym najprostszym przypadku wymagało uwzględnienia w modelu takich kluczowych elementów, jak np. koszty komunikacji oraz inicjalizacji przetwarzania poszczególnych porcji danych. Wobec bardzo dużej złożoności zagadnienia jednoczesnego zbadania wpływu wszystkich parametrów charakteryzujących obliczenia na zużycie energii, skupiono się na wyznaczeniu map izoenergetycznych dla par parametrów jako efektywnego narzędzia analizy nie tylko relacji pomiędzy tymi parami, lecz również uzyskania bardziej całościowego obrazu zagadnienia zużycia energii. Wśród tych parametrów należy wymienić rozmiar rozwiązywanego problemu i rozmiar porcji danych, liczbę procesorów, wydajność procesorów i

wydajność sieci komunikacyjnej, moc zużywaną zarówno przez procesor, jak i przez sieć komunikacyjną, czas inicjalizacji przetwarzania.

*Rozdział piaty* dotyczy bardziej rozpowszechnionego i jednocześnie złożonego przypadku obliczeń w systemach z pamięcią hierarchiczną, co pozwala uwzględniać obecność pamięci podręcznej, pamięci operacyjnej czy też pamięci zewnętrznej. Dla systemów homogenicznych (tj. składających się z takich samych komputerów) rozpatrywane są dwa przypadki: z jednoetapową dystrybucją pracy do wykonania oraz dystrybucją wieloetapową. W pierwszym z nich rozwiązanie problemu szeregowania jest formułowane jako zadanie mieszane programowania liniowego (ang. mixed integer linear programming), zaś w drugim przypadku rozwiązanie otrzymywane jest bądź to dostatecznie prostymi metodami heurystycznymi (zachłannymi), bądź to przez sprowadzenie do zagadnienia mieszane programowania liniowego pozwalającego uzyskać rozwiązanie optymalne. Nie negując wartości analizy porównawczej wyników uzyskiwanych za pośrednictwem tych metod, najbardziej interesującym aspektem badań prowadzonych w tym rozdziale wydaje się być wynikający z nich wniosek o istnieniu przetargu (czy też kompromisu - ang. trade-off) pomiędzy zużyciem energii a czasem uszeregowania. W związku z tym, wykorzystując aparat map izoefektywności, zbadano wpływ szeregu parametrów wykonywanych obliczeń na ich wydajność czasową i energetyczną. W szczególności, wykazano iż w wielu przypadkach dzięki zwiększeniu stopnia zrównoleglenia obliczeń możliwe jest zmniejszenie zarówno zużycia energii, jak i czasu uszeregowania

W *rozdziale szóstym* uzyskane wyniki dla obliczeń z pamięcią hierarchiczną rozszerzono na przypadek systemów heterogenicznych, ograniczając się do rozpatrzenia tylko wieloetapowej dystrybucji danych jako otwierającej większe możliwości poprawy efektywności. Do konstrukcji uszeregowania zaproponowane zostały zarówno algorytmy heurystyczne, jak i algorytmy oparte na zastosowaniu mieszane programowania liniowego. Zaproponowane algorytmy zachłanne są adaptacjami wykorzystanych w poprzednim rozdziale algorytmów Simple Static Chunk (SSC) oraz Guided Self-Scheduling (GSS). Chociaż zastosowanie mieszane programowania liniowego (MIP) wiąże się z wykładniczymi czasami działania algorytmów jako funkcji liczby procesorów oraz liczby paczek danych, co związane jest z NP-trudnością samego zagadnienia, to Autor wykazał, iż dla praktycznie ważnych rozmiarów problemów optymalne rozwiązania mogą być znalezione z wykorzystaniem współczesnych solverów MIP. I tutaj również nie negując wartości analizy porównawczej jak różne typy algorytmów szeregowania wymieniają czas swojego działania na jakość otrzymywanych rozwiązań (wyniki są tutaj intuicyjnie dosyć oczywiste, chociaż rzeczywiste bardzo wartościowe z punktu widzenia teoretycznego), najbardziej interesującym dla mnie aspektem badań zrealizowanych w danym rozdziale było wykazanie, iż nawet w tym skomplikowanym przypadku możliwe jest efektywne zastosowanie zaproponowanych metod i algorytmów do

ukierunkowanej analizy przetargu pomiędzy zużyciem energii a wydajnością obliczeń. Charakter tego przetargu jest jednak dosyć złożony, z częstymi nieregularnościami.

W *rozdziale siódmym* dokonano zwięzłego podsumowania dysertacji, w tym jej oryginalnych elementów, a także wskazano na obiecujące perspektywy rozwoju uzyskanych rezultatów w ramach dalszych prac Autora czy też innych badaczy. Nie ulega bowiem wątpliwości, iż koncepcja izomap może być zastosowana do budowy modeli dla innych przypadków zrównoleglenia obliczeń i związanych z nimi zagadnień wyznaczania uszeregowania.

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

1. W mojej opinii, podstawowym wynikiem ogólnym, o charakterze zarówno poznawczym, jak i aplikacyjnym, jest wykazanie możliwości poprawy efektywności energetycznej dla ważnej klasy obliczeń/aplikacji równoległych przetwarzających duże ilości danych dzięki opracowaniu efektywnych metod zarządzania zarówno wykonaniem samych aplikacji, jak i zasobami systemu komputerowego, co pozwala uwzględniać nie tylko kryterium zużycia energii, lecz również drugie podstawowe kryterium jakości obliczeń jaki stanowi czas wykonania obliczeń.
2. Weryfikacja powyższej tezy została zrealizowana poprzez opracowanie autorskiej metodologii szeregowania zadań jednolicie podzielnych, opartej na zastosowaniu zaproponowanej koncepcji map izoenergetycznych i uwzględniającej architekturę systemu komputerowego oraz kluczowe parametry systemu i parametry wykonania aplikacji, takie jak np. liczbę procesorów czy koszty komunikacji, a także charakteryzującej się wykorzystaniem odpowiednio dobranych metod heurystycznych i dokładnych algorytmów optymalizacji do wyznaczania rozwiązań formułowanych zagadnień poszukiwania optymalnych lub zbliżonych do optymalnych rozwiązań dla uszeregowania.
3. W mojej opinii, najbardziej innowacyjnym elementem opracowanej metodologii jest autorska koncepcja map izoenergetycznych jako naturalnego rozszerzenia znanych wcześniej z literatury map izoefektywności związanych z metrykami wydajności obliczeń. Koncepcja ta pozwala na analizę relacji pomiędzy parametrami determinującymi zużycie energii w obliczeniach jednorodnie podzielnych, wykrycie stopnia ważności poszczególnych parametrów oraz ich wzajemnych relacji, a w rezultacie uzyskanie całościowego obrazu zagadnienia zużycia energii i na tej podstawie prowadzenie ukierunkowanych poszukiwań możliwości zmniejszenia tego zużycia.
4. Istotnym elementem, który zdecydował o efektywności opracowanej metodologii, jest także umiejętne sformułowanie problemów szeregowania dla

zarówno homogenicznych, jak i heterogenicznych systemów komputerowych z pamięcią hierarchiczną jako problemów optymalizacji matematycznej. Do ich rozwiązania kandydat zaproponował wykorzystanie nie tylko heurystycznych metod zachłanych opartych w szczególności na zastosowaniu zasady preferencji procesorów zgodnie z ich parametrami wydajnościowymi, lecz również dokładnych algorytmów optymalizacji, takich jak mieszane programowanie liniowe, które może być rozwiązane dość efektywnie przez współczesne solwery przy ograniczonych rozmiarach problemów.

5. Kontynuując wątek z poprzedniego punktu, pragnę zwrócić uwagę na wyczerpującą analizę porównawczą wspomnianych alternatywnych podejść do rozwiązania problemów szeregowania, która pozwoliła na przekonywujący dowód występowania przetargu pomiędzy złożonością obliczeniową a jakością rozwiązania. W szczególności, pokazano, że rozwiązania uzyskane w wyniku zastosowania algorytmu mieszanego programowania liniowego są zawsze najlepsze co do jakości rozwiązania. Jednakże rozwiązania te są uzyskiwane kosztem czasu obliczeń, najwyższego spośród wszystkich badanych metod.
6. Niejako w podsumowaniu mojej analizy wkładu Autora przytoczę rezultat łączącym aspekt poznawczy i praktyczny (aplikacyjny), a mianowicie nie tylko wykazanie istnienia, lecz również wyczerpującą analizę przetargu (ang. tradeoff) pomiędzy kryterium czasu wykonania i zużycia energii. Do analizy wpływu parametrów różnych składników platformy obliczeniowej na czas i wydajność przetwarzania zadań jednolicie podzielnych wykorzystano przy tym nie tylko modelowanie, lecz również wsparto ją zastosowaniem klasycznych map izoefektywności i zaproponowanych w pracy map izoenergetycznych. Jak można było oczekiwać, charakter tego przetargu charakteryzuje się znaczną złożonością i nieregularnością, co jest szczególnie widoczne w przypadku systemów heterogenicznych, przy czym, co zasługuje na uwagę, w szeregu przypadkach kombinacji parametrów zrównoleglenie obliczeń możliwa jest redukcja zarówno zużycia energii, jak i czasu uszeregowania.

Uzyskane wyniki zostały opublikowane w pięciu pracach w języku angielskim, z których trzy ukazała się w dobrych i bardzo dobrych czasopismach z listy JCR (Future Generation Computer System, Journal of Grid Computing oraz Journal of Parallel and Distributed Computing - JPDC), a dwa artykuły na międzynarodowych konferencjach z cieszącą się zasłużoną renomą serii PPAM, indeksowanej w Web of Science oraz Scopus (aktualnie Core C). Wobec faktu, iż dwa artykuły z czasopism ukazały się stosunkowo dawno (2014 i 2016), na szczególne podkreślenie zasługuje fakt opublikowania w roku bieżącym pracy pt. „Time–energy trade-offs in processing divisible loads on heterogeneous hierarchical memory systems” w czasopiśmie JPDC (100 punktów według aktualnej klasyfikacji Ministerstwa). W mojej opinii, świadczy

to bardzo pozytywnie o stopniu weryfikacji uzyskanych rezultatów przez międzynarodową społeczność badawczą.

#### 4. Poprawność pracy i uwagi krytyczne

Poprawność treści pracy nie wzbudza moich zasadniczych zastrzeżeń, a stwierdzenia w niej zawarte wydają się generalnie być godne zaufania, co nie wyklucza uwag o charakterze krytycznym czy dyskusyjnym. Powyższe zaufanie wynika w szczególności z dosyć szczegółowych uzasadnień, popartych wynikami przeprowadzonych badań eksperymentalnych. Na przykład, część eksperymentalna obejmuje nie tylko symulacje komputerowe, lecz również testy wykonane dla pięciu różnych aplikacji uruchamianych na czterech platformach komputerowych, a także dla benchmarku o nazwie POLY. Należy podkreślić przy tym wkład pracy Autora włożony w własnoręczne zbudowanie i uruchomienie stanowiska badawczego do pomiarów zużycia energii przez aplikacje. Zastosowany sposób pomiaru z wykorzystaniem sprzętowego watomierza wydaje się pozwalać na uzyskanie dostatecznie wysokiej wiarygodności pomiarów – potencjalnie wyższej niż szeroko stosowane w wielu pracach rozwiązania softwarowe oparte na predykcji. Generalnie pozytywna ocena zastosowanego podejścia do części eksperymentalnej pracy nie wyklucza jednak szeregu uwag krytycznych, o których będzie mowa poniżej.

Wśród uwag o charakterze krytycznym, a częściowo dyskusyjnym, wymienić należy:

1. Pierwsza z nich dotyczy zbyt pobieżnego uwzględnienia i przedstawienia zagadnień zapewnienia wiarygodności, w tym powtarzalności pomiarów zużycia energii z wykorzystaniem zbudowanego stanowiska badawczego. Autor nie przytacza jakichkolwiek szczegółów na temat zapewnienia statystycznej wiarygodności uzyskanych wyników pomiarów, ograniczając się do stwierdzenia, że poszczególne pomiary powtarzano 3 razy. Generalnie jest to zbyt mała liczba, aby można w pełni zaufać wynikom pomiarów otrzymanym w ten sposób, nie mówiąc już o tym, iż nie pozwala ona na zastosowanie jakichkolwiek systematycznych metod statystycznej obróbki wyników.
2. Podobny zarzut pobieżności dotyczy możliwości wydzielenia w całkowitej energii (mocy) zużywanej przez system komputerowy poszczególnych części odpowiadających procesorom przetwarzającym porcje danych oraz sieci komunikacyjnej realizującej przesyłanie danych pomiędzy procesorami. Z mojego doświadczenia i literatury przedmiotu wynika, iż jest to zagadnienie dalece nietrywialne. Autora usprawiedliwia fakt, iż w pracy wielkości te nie są otrzymywane eksperymentalnie za pośrednictwem rzeczywistych pomiarów, powinien jednak chociażby zasygnalizować trudności takiego wydzielenia.
3. Ostatnia uwaga krytyczna w obszarze badań eksperymentalnych dotyczy faktu, iż zostały one zrealizowane tylko dla walidacji podstawowego modelu opisującego zależność zużycia energii od wielkości porcji danych w obliczeniach jednorodnie podzielnych (rozdział 3 pracy). Rozumiem, że przeprowadzenie takich badań w pozostałych rozdziałach w celu porównania



wyników teoretycznych w zakresie zużycia energii i czasu wykonania z wynikami pomiarów wymagałoby bardzo dużego nakładu pracy, np. związanego z koniecznością wyłącznego dysponowania systemem równoległym zawierającym wiele procesorów. Tym niemniej szkoda, że Autor nie zdecydował się na przeprowadzenie takich badań chociażby w niektórych mniej wymagających przypadkach. Moim zdaniem, wpłynęłoby to m.in. na większe zainteresowanie innych badaczy z danego obszaru wynikami uzyskanymi w dysertacji.

4. Kolejna grupa uwag krytycznych wiąże się z poprzednimi i dotyczy literatury przedmiotu wykorzystanej w pracy, a także sposobu odniesienia się do niej w tekście rozprawy. Jak już wspomniałem w punkcie 2 mojej recenzji, literatura w zakresie zużycia energii w obliczeniach równoległych, w tym relacji pomiędzy wydajnością czasową obliczeń a ich efektywnością energetyczną, jest niezwykle bogata i trudno wymagać od Autora kompleksowego jej przedstawienia. Jednakże w mojej opinii przegląd literaturowy w tej dziedzinie został zrealizowany zbyt pobieżnie – po prostu poświęcono mu zbyt mało miejsca. W szczególności, pominięte zostały w ogóle prace zespołu prof. A. Lastovetsky z University College Dublin, który od lat zajmuje się daną tematyką, uzyskując niezwykle znaczące wyniki w zakresie m.in. modeli zużycia energii w obliczeniach równoległych, optymalizacji bi-kryterialnej czy też zagadnień zapewnienia wiarygodności pomiarów zużycia energii. Chciałbym też zwrócić uwagę na stosunkowo małą liczbę pozycji literaturowych pochodzących z ostatnich 5 lat – 16 publikacji spośród wszystkich 101 pozycji bibliografii. Przykładowo, taka ważna technika zarządzania zużyciem energii jaką jest DVFS poparta jest publikacjami z lat 2007 i 2010.
5. Należy zdawać sobie sprawę, iż przyjęty w pracy model obliczeń równoległych jednorodnie podzielnych, chociaż ważny z teoretycznego i praktycznego punktu widzenia, jest modelem, którego zakres zastosowania jest jednak ograniczony. W szczególności, przyjęty schemat komunikacyjny typu gwiazda zakłada, iż komunikacja odbywa się tylko pomiędzy węzłem centralnym (master) a węzłami roboczymi (typu worker). Może to spowodować nadmierne obciążenie zasobów węzła centralnego, w tym jego sieci komunikacyjnej, ograniczając w ten sposób skalowalność obliczeń. Inne ograniczenie modelu to np. brak komunikacji pomiędzy workerami. W związku z tym nasuwa się pytanie czy Autor rozpatrywał możliwość rozszerzenia zaproponowanego podejścia na inne modele obliczeń równoległych ?
6. Moim zdaniem, celowe byłoby zamieszczenie dodatkowych informacji na temat aplikacji testowych, o których mowa w rozdziale 3. W pracy Autor ograniczył się praktycznie do podania nazw tych aplikacji. Tymczasem sposób implementacji danej aplikacji i optymalizacji algorytmu/kodu ma często bardzo duży wpływ na efektywność obliczeń. W jednej z naszych ostatnich prac pokazaliśmy na przykład, że szereg ukierunkowanych transformacji algorytmu/kodu aplikacji służącej do modelowania zjawisk geofizycznych pozwala zmniejszyć zużycie energii oraz czas wykonania ponad 10-krotnie.

Recenzent rozumie, iż w danym przypadku modyfikacja kodu aplikacji, która spełnia założenia modelu jednolicie podzielnych obliczeń, nie naruszy słuszności założeń przyjętych do budowy modeli zużycia energii i czasu wykonania. Zmienia ona jednak parametry obliczeń, w tym również moc konsumowaną przez procesor czy sieć komunikacyjną, wpływając na wyniki uzyskiwane przez Autora dla różnych przypadków, szczególnie przy uwzględnieniu hierarchii pamięci oraz heterogeniczności procesorów.

7. Kontynuując wątek poruszony w poprzednim punkcie chciałbym również zapytać o możliwość uwzględnienia w przyszłości techniki DVFS dynamicznego skalowania napięcia-częstotliwości, która prowadzi do zmiany mocy konsumowanej przez procesory przy wykonywaniu aplikacji.

### 3. Wiedza kandydata

Z omówienia treści pracy, które przytoczono w punkcie 2 niniejszej recenzji, wynika, iż rozdział drugi rozprawy poświęcony został krytycznemu przedstawieniu stanu wiedzy w zakresie tematyki pracy, tj. w takich obszarach jak zagadnienia zużycia energii w systemach komputerowych, teoria systemów jednorodnie podzielnych czy też sama koncepcja izofunkcji i izomap jako wizualnego narzędzia badawczego używanego w różnych sferach zastosowań, w tym również w analizie wydajności/efektywności obliczeń równoległych. Pomimo pewnych uwag krytycznych sformułowanych w poprzednim punkcie recenzji, które mają w znacznej części charakter dyskusyjny, wynikając z moich osobistych doświadczeń, jakość tego rozdziału należy ocenić pozytywnie. Potwierdza on wiedzę kandydata w zakresie dyscypliny Informatyka, ze szczególnym uwzględnieniem zagadnień systemów i obliczeń równoległych oraz teorii szeregowania zadań. Elementy krytycznego odniesienia się do aktualnego stanu wiedzy można znaleźć także w końcowym fragmencie podrozdziału 3.3 rozprawy, gdzie Autor formułuje swoją polemiczną opinię o innych modelach zużycia energii w systemach z pamięcią hierarchiczną.

W pozostałych rozdziałach pracy, które w całości poświęcono badaniom własnym kandydata, również demonstruje on swoją niewątpliwą wiedzę we wspomnianych wyżej obszarach, rozszerzając ją o szczegółowe zagadnienia z zakresu architektury systemów komputerowych i ich modelowania, algorytmów szeregowania oraz klasycznych algorytmów optymalizacyjnych, takich jak mieszane programowanie liniowe (MIP). Należy stwierdzić, iż kandydat porusza się w tych zagadnieniach z niezbędną w tej dziedzinie ścisłością i nie daje powodów do uwag krytycznych.

Przy tej generalnie bardzo pozytywnej ocenie wiedzy kandydata, pozostaje więc kwestia wyważonej oceny bibliografii przytoczonej w rozprawie. W mojej opinii, w zakresie literatury poświęconej dotychczasowym badaniom nad zużyciem energii można by ją sporządzić w sposób bardziej kompletny, lepiej uwzględniający tę jakże

złożoną i co tu kryć – niesłuchanie obszerną tematykę. Faktycznie, kandydat podszedł do tej części bibliografii minimalistycznie. Natomiast nie budzi moich zastrzeżeń kompletność i jakość bibliografii związanej z szeregowaniem zadań, w tym szeregowaniem obliczeń jednorodnie podzielnych.

Podsumowując moje wywody w tym punkcie, mogę finalnie stwierdzić po przeczytaniu całej pracy, iż pomimo przytoczonych uwag krytycznych w zakresie części bibliografii, kandydat z pewnością posiada ogólną wiedzę teoretyczną w dyscyplinie Informatyka.

#### **4. Podsumowanie**

Biorąc pod uwagę opinie zaprezentowane w poprzednich punktach i wymagania aktualnie obowiązującej Ustawy o stopniach naukowych i tytule naukowym moja ocena pod względem trzech podstawowych kryteriów jest następująca:

A. Czy rozprawa zawiera oryginalne rozwiązanie problemu naukowego ?

Zdecydowanie TAK

B. Czy po przeczytaniu rozprawy zgadzasz się, że kandydat posiada ogólną wiedzę teoretyczną w dyscyplinie Informatyka ?


Zdecydowanie TAK

C. Czy kandydat posiada umiejętność samodzielnego prowadzenia pracy naukowej ?

Zdecydowanie TAK

Ponadto biorąc pod uwagę oryginalność przedstawionego w dysertacji wkładu kandydata do dyscypliny Informatyka, a także sumaryczną punktację trzech publikacji z listy JCR wynoszącą 340 punktów (według aktualnej punktacji), rekomenduję wyróżnienie rozprawy doktorskiej.

Prof. dr hab. inż. Roman Wyrzykowski



R. Wyrzykowski