

Dr hab. inż. Wojciech Turek, prof. Uczelni  
Instytut Informatyki  
Wydział Informatyki, Elektroniki i Telekomunikacji  
Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie  
Al. Mickiewicza 30, 30-059 Kraków  
wojciech.turek@agh.edu.pl

Kraków, 22 września 2021r.

Recenzja rozprawy doktorskiej  
dla  
Rady Dyscypliny Informatyka Techniczna i Telekomunikacja  
Politechniki Wrocławskiej

Autor rozprawy: mgr inż. Piotr Nadybski

Tytuł rozprawy: Algorytmy szeregowania zadań i rozdziału zasobów w systemach wieloprocesorowych

Promotor: prof. dr hab. Mieczysław Wodecki

Promotor Pomocniczy: dr inż. Mateusz Uchroński

## Wstęp

Praca została napisana w języku polskim i zawiera: spis treści, listy oznaczeń i skrótów, streszczenie w języku polskim i angielskim, sześć rozdziałów, indeks oraz spis literatury. Jest napisana poprawnym językiem, nie zawiera poważniejszych błędów językowych czy edytorskich. Praca jest wydana w atrakcyjnej formie w ramach serii wydawniczej Wydziału Elektroniki Politechniki Wrocławskiej. Pomimo widocznego, znacznego wysiłku edytorskiego, autor nie ustrzegł się drobnych błędów, jak np. brakująca litera w tytule rozdziału 5.2 czy wystąpienie rozdziału o numerze 2.5.0.1. Te nieliczne usterki nie wpływają na bardzo pozytywny odbiór językowej i edytorskiej strony rozprawy.

## Podjęmowany problem badawczy

Podjęmowane w rozprawie zagadnienia szeregowania zadań i rozdziału zasobów w rozproszonych systemach komputerowych to ważne problemy abstrakcyjne, których liczne przypadki szczególne ciągle stanowią wyzwanie dla projektantów sprzętu i oprogramowania zarządzającego nowoczesnymi

WPLYNĘŁO

29-09-2021

RDN-LIT/158/2021

infrastrukturami obliczeniowymi. Są to problemy badane od dziesięcioleci, które doczekały się wielu teoretycznych analiz i praktycznych rozwiązań. Jednak, jako problemy klasy NP, niezmiennie stanowią otwarte wyzwania i pozwalają na proponowanie coraz lepszych algorytmów poszukiwania rozwiązań, dostosowanych do własności specyficznych problemów.

Autor pracy skupia się na dwóch konkretnych przypadkach szczególnych, dodatkowo rozważając ich wybrane warianty. Są to:

1. Problem rozsyłania danych przez kanały komunikacyjne o ograniczonej przepustowości, który był rozważany w wariacie z najpóźniejszymi momentami zakończenia oraz z oknami czasowymi momentów zakończenia.
2. Problem szeregowania zadań w klastrze obliczeniowym, który był rozważany w wariacie węzłów sąsiednich oraz węzłów dowolnych.

Ogólnym celem prowadzonych prac było wykazanie, że do rozwiązywania powyższych problemów można stosować popularne metody optymalizacji dyskretnej. Postawione tezy skupiają się na zagadnieniu identyfikowania i dowodzenia specyficznych własności problemów oraz wykorzystaniu tych własności do poprawy skuteczności algorytmów optymalizacji. Tezy są sformułowane dość ogólnie i opisowo, nie wszystkie są jednoznacznie weryfikowalne. Zabrakło też w formie ilościowej oraz porównawczej, pozwalających na obiektywne odniesienie uzyskanych wyników do istniejącego stanu wiedzy w dziedzinie.

## Zawartość pracy

W pierwszym rozdziale znajduje się wprowadzenie w tematykę pracy, które wyraźnie podkreśla praktyczne zastosowania rozważanych problemów i potwierdza zasadność prowadzenia badań w tej dziedzinie. Uzasadnienie jest przekonujące. Autor prezentuje ogólnie znane cechy współczesnych architektur sprzętu i oprogramowania systemów komputerowych dużej skali i poprawnie identyfikuje liczne problemy w nich występujące. Następnie prezentuje wspomniane wcześniej tezy pracy i przedstawia krótko zawartość kolejnych jej rozdziałów.

Rozdział drugi ma na celu zarysowanie tła dla rozważanych problemów i ich rozwiązań. Tytuł rozdziału jest dość zaskakujący w kontekście jego zawartości – autor prezentuje tu szerokie tło swoich prac i badań, obejmujące pobieżne wprowadzenie do problematyki optymalizacji dyskretnej, flagowe problemy dla tego typu optymalizacji, wybrane algorytmy metaheurystyczne czy nawet zagadnienia złożoności obliczeniowej. Te ogólne zagadnienia są umiejętnie omówione w kontekście badanych problemów, co wskazuje na szeroki zakres wiedzy autora. W rozdziale nieco dokładniej wprowadzone są podstawy teorii szeregowania zadań, która w późniejszych fragmentach pracy jest wykorzystywana jako baza modeli formalnych badanych problemów szczegółowych.

Trzeci rozdział pracy poświęcony jest pierwszemu z badanych problemów, czyli problemowi rozsyłania danych przez kanały komunikacyjne o ograniczonej przepustowości. W rozdziale bardzo krótko przedstawione jest tło problemu oraz istniejące prace badawcze poświęcone tej tematyce. Następnie autor przechodzi do formalnej specyfikacji konkretnych wariantów rozważanego problemu – jego wersji z najpóźniejszymi momentami zakończenia transmisji i oknami czasowymi zakończenia transmisji. Problem jest specyfikowany i rozwiązywany z wykorzystaniem elementów teorii

szeregowania poprzez zidentyfikowane analogie do problemów wielomaszynowych. Autor proponuje mechanizmy konstruowania rozwiązań problemu w oparciu o modyfikacje rozwiązań wcześniejszych, stosując szereg obserwacji dotyczących charakterystyki problemu. Obserwacje te zapisane są formalnie w formie lematów i twierdzeń z dowodami, które pozwalają ograniczyć liczbę potencjalnych rozwiązań do zbadania. Na ich podstawie zaprojektowane zostały różne warianty algorytmów przeszukiwania przestrzeni rozwiązań, które następnie eksperymentalnie przebadano i porównano pod kątem skuteczności. Zaprezentowane eksperymenty obliczeniowe zostały wykonane na bardzo niewielkiej próbie losowo wygenerowanych przypadków. Potwierdzają one tezy dotyczące możliwości poprawiania skuteczności algorytmów poprzez wykorzystanie specyficznych cech problemów dla rozważanego zagadnienia rozsyłania danych przez kanały komunikacyjne o ograniczonej przepustowości.

Kolejny, czwarty rozdział pracy, skupia się na drugim rozważanym problemie, czyli szeregowaniu zadań w klastrze obliczeniowym. Rozdział ma strukturę bardzo podobną do rozdziału trzeciego – po bardzo krótkim wprowadzeniu w zagadnienie i pobieżnym przeglądzie istniejących rozwiązań, autor podaje formalną definicję rozważanego problemu. Następnie identyfikuje analogię rozważanego problemu do zadania dwuwymiarowego pakowania, proponuje strategię generowania kolejnych rozwiązań oraz algorytm przeszukiwania z tabu. Skuteczność algorytmów jest następnie porównana poprzez wykonanie eksperymentów obliczeniowych. Również w tym wypadku wykonano je na bardzo niewielkiej próbie losowo wygenerowanych instancji problemu.

Tytuł i wprowadzenie do rozdziału piątego sugerują, że zaprezentowane we wcześniejszych rozdziałach rozwiązania zostały skutecznie zastosowane do rozwiązania realnego problemu praktycznego, co byłoby wartościową weryfikacją ich poprawności i przydatności. Rozdział prezentuje rzeczywisty problem inżynierski związany z zarządzaniem oprogramowaniem w uczelni wyższej. Potrzeba racjonalizacji wydatków na licencje specjalistycznego oprogramowania powoduje konieczność migrowania oprogramowania pomiędzy różnymi stanowiskami komputerowymi w określonym czasie, co tworzy potencjał dla zastosowania algorytmu optymalizacji rozsyłania danych przez dostępne kanały komunikacyjne. W rozdziale przedstawiono szczegółowe wymagania przykładowej uczelni, które wykraczają daleko poza sam problem dystrybucji oprogramowania. Następnie zaprezentowano komercyjne oprogramowanie, które potencjalnie jest w stanie te wymagania spełnić. Dopiero na samym końcu tego dość długiego rozdziału znajduje się odniesienie do rozwiązań prezentowanych w rozdziałach wcześniejszych, które podsumowane jest stwierdzeniem, iż rozwiązania te „mogą mieć w tym przypadku zastosowanie”, co budzi spory niedosyt.

Ostatni rozdział pracy to krótkie podsumowanie zaprezentowanych osiągnięć i ważne stwierdzenie dotyczące potwierdzenia postawionych tez.

## Osiągnięcia autora zaprezentowane w rozprawie

Zaprezentowany w pracy wywód stanowi z pewnością spójne i oryginalne rozwiązanie problemu naukowego. Autor poprawnie zidentyfikował i opisał ważny obszar badawczy oraz uzasadnił jego znaczenie teoretyczne i praktyczne. Zaprezentował też szerokie tło tego problemu i potencjalnych rozwiązań stosowanych do jego rozwiązywania. Następnie sformułował hipotezy badawcze, których

poprawność w kolejnych etapach badań wykazał. Najważniejsze hipotezy oparte są o uznane twierdzenie „No free lunch”, które stanowi punkt wyjścia dla dalszych rozważań.

Dobre problemy szczegółowe zostały poprawnie sformalizowane, co pozwoliło na opracowanie twierdzeń ograniczających przestrzeń poszukiwań rozwiązań. Na ich podstawie zdefiniowano algorytmy referencyjne oraz algorytmy korzystające ze specyficznych cech problemów. Skuteczność działania algorytmów została przebadana w eksperymentach obliczeniowych, co pozwoliło wykazać postawione tezy.

Na szczególną uwagę zasługuje interesujący sposób analizy problemów i proponowania ich rozwiązań. Formalizacja problemu rozsyłania danych przez kanały komunikacyjne o ograniczonej przepustowości w oparciu o teorię szeregowania dała podstawę do zdefiniowania wartościowych lematów i twierdzeń. Na ich podstawie zaproponowano metodę ograniczania przestrzeni rozwiązań, czyli dostrojono algorytm optymalizacji do badanego problemu. Bardzo interesującym i oryginalnym podejściem jest również zidentyfikowanie analogii między problemem harmonogramowania zadań w klastrze obliczeniowym oraz problemem dwuwymiarowego pakowania. Identyfikowanie analogii i powiązań badanych problemów z innymi dziedzinami czy zagadnieniami często prowadzi do wartościowych wyników, dlatego takie podejście należy uznać za bardzo dobre.

Autor pracy cytuje trzynaście opublikowanych prac naukowych, których jest współautorem, co jest znaczącym dorobkiem na tym etapie kariery naukowej. Znajdują się wśród nich publikacje z krajowych i międzynarodowych konferencji oraz czasopism, z których część indeksowana jest w bazie DBLP.

## Uwagi krytyczne do zawartości rozprawy

Pomimo niezaprzeczalnej wartości naukowej zaprezentowanej pracy, niektóre jej aspekty budzą wątpliwości. Głównym problemem, na który chciałbym zwrócić uwagę, jest sposób ulokowania proponowanych rozwiązań wśród istniejących badań naukowych. Chodzi tu zarówno o bardzo pobieżne przedstawienie istniejących badań, metod i wyników prac podejmujących podobną tematykę, jak również o brak porównania uzyskanych wyników ilościowych z wynikami istniejących algorytmów rozwiązywania zbliżonych problemów.

Istniejący stan badań w zakresie zagadnienia rozsyłania danych przez kanały komunikacyjne o ograniczonej przepustowości autor pracy przedstawia na połowie strony nr 48, korzystając jedynie z kilku referencji. Z kolei omówienie stanu badań nad szeregowaniem zadań w klastrach obliczeniowych zmieściło się na mniej niż połowie strony nr 117. Cytując autora pracy, podejmowane problemy „stanowią od lat obszar intensywnie prowadzonych badań”, zatem z pewnością możliwe i zasadne było zaprezentowanie szerszego tła oraz, co ważniejsze, umiejscowienie proponowanych rozwiązań wśród innych badań.

Niedosyt budzi również sposób prezentacji osiągniętych wyników ilościowych w testach obliczeniowych. Autor prezentuje prawie wyłącznie wyniki swoich własnych algorytmów, nie porównując ich z wynikami algorytmów opracowanych przez innych badaczy. Pominięcie porównania wyników jest uzasadnione jedynie w przypadku rozważania problemu, który nie było wcześniej badany i dla którego nie istnieją działające rozwiązania. Jednak w przypadku szeregowania zadań w klastrach obliczeniowych autor przedstawia nawet szeroką listę wykorzystywanych w praktyce systemów

kolejkowania zadań, które rozwiązują analogiczny problem. Nie podejmuje jednak próby porównania jakości generowanych przez nie wyników z wynikami własnych metod.

Przy ewaluacji wyników algorytmów optymalizacji operujących na zadaniach, które można generować losowo, warto stosować podejście statystyczne. Zaprezentowanie metryk statystycznych z odpowiednio licznej próby testów dla zadania tego samego rozmiaru pozwala na lepszą ocenę skuteczności działania metody niż wynik pojedynczego uruchomienia.

Warto zwrócić też uwagę na sposób prezentowania wyników na wykresach. Łączenie linią wartości pomiarów dla poszczególnych przypadków testowych może nie być uzasadnione, szczególnie gdy wartości na osi poziomej nie są porównywalne (np. Rys 4.7).

Pomimo, że kwestie te nie wpływają na ogólną ocenę pracy, są one ważnym elementem badań w dyscyplinie. Uprzejmię proszę o staranne ustosunkowanie się do powyższych uwag w trakcie publicznej obrony. Proszę szczególnie o przygotowanie analizy innych algorytmów rozwiązujących analogiczne problemy oraz o porównanie ilościowe wyników tych algorytmów z wynikami metod prezentowanych w pracy lub uzasadnienie niemożliwości wykonania takiego porównania.

## Podsumowanie

Całokształt zaprezentowanego w pracy wywodu naukowego, jakość przeprowadzonych analiz, opracowanych modeli i algorytmów oraz sposób ich prezentacji stoją na wysokim poziomie. Autor wykazał się szeroką wiedzą teoretyczną w dyscyplinie oraz potwierdził umiejętność samodzielnego prowadzenia badań naukowych. Ponadto zaprezentował oryginalne rozwiązanie problemu naukowego. Biorąc powyższe pod uwagę stwierdzam, że w świetle Ustawy o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki z dn. 14.03.2003 (Dz. U. nr 65, poz. 595 z późn. zm.) w związku z §179 u1 ustawy Przepisy wprowadzające ustawę - Prawo o szkolnictwie wyższym i nauce (Dz. U. 2018, poz. 1669 z późn. zm.), rozprawa doktorska mgr. inż. Piotra Nadybskiego spełnia wymogi stawiane rozprawom doktorskim. W związku z tym wnoszę o dopuszczenie do dalszych etapów postępowania o nadania stopnia naukowego doktora, w tym o dopuszczenie do publicznej obrony.



Wojciech Turek