

dr hab. inż. Mariusz Kamola
Instytut Automatyki i Informatyki Stosowanej
Politechnika Warszawska
mariusz.kamola@pw.edu.pl

Recenzja rozprawy doktorskiej
opracowana na wniosek
Rady Dyscypliny Naukowej
Informatyka Techniczna i Telekomunikacja
Uniwersytetu Jagiellońskiego

Tytuł rozprawy: Designing Urban Transit Network using Memetic Algorithm

Autor rozprawy: Hanan Ali Mohammed Ba Ali

Przedmiotem rozprawy jest problem konstrukcji sieci transportu publicznego. W pracy został on podzielony na trzy etapy: poszukiwania optymalnej struktury połączeń (rozdz. 3), kompresji wyjściowego grafu infrastruktury transportowej (rozdz. 4) oraz planowania procesu przejścia od struktury połączeń istniejących do postulowanej struktury optymalnej (rozdz. 5). Autorka formułuje przy tym stosowne pytania badawcze, natomiast większość eksperymentów dotyczy sieci tramwajowej m. Krakowa lub sieci jej podobnych.

Ocenę pracy należy wykonać zgodnie z art. 187 ustawy Prawo o szkolnictwie wyższym, wg którego rozprawa doktorska ma

- 1) prezentować ogólną wiedzę teoretyczną kandydata w dyscyplinie,
- 2) prezentować umiejętność samodzielnego prowadzenia pracy naukowej,
- 3) przedstawiać oryginalne rozwiązanie problemu naukowego *lub* rozwiązanie w zakresie zastosowania wyników własnych badań naukowych w sferze gospodarczej lub społecznej.

W odniesieniu do warunku 3., w opinii recenzenta charakter przedstawionych prac jest w przeważającej części badawczy. Dojrzałość technologiczna wykonanego oprogramowania i zakres jego zastosowania nie pozwalają na ocenę dokonań autora pod kątem aplikacyjnym.

Opinia o rozprawie została wyrażona poniżej względem ww. kryteriów cząstkowych.

1. Przedstawienie w rozprawie ogólnej wiedzy teoretycznej kandydata w dyscyplinie

Autorka, powołując się w rozprawie na ponad 100 publikacji, prezentuje należyłą wiedzę zarówno w prezentowanej dyscyplinie, jak i w dziedzinie, w której praca może mieć zastosowanie praktyczne. Przedstawione zestawienie algorytmów genetycznych i pokrewnych jest kompletne i czytelne. Przegląd zagadnień planowania i optymalizacji sieci transportu publicznego jest szeroki i kompleksowy, a przez to atrakcyjny. W ciągu rozprawy autorka wielokrotnie dowodzi szerokiej znajomości elementów algorytmiki, jak również teorii informacji czy algebry.

W tym kontekście niepokojąca jest i niezrozumiała maniera pisarska nieposługiwania się sformułowaniami matematycznymi, i ogólniej, środkami wyrazu przyjętymi w pracach naukowych, w zamian za rozbudowane opisy słowne i wydruki kodu źródłowego. Teoria grafów posługuje się od dawna ustalonym zapisem algebraicznym i terminologią, która tutaj albo nie jest jawnie przywołana w ogóle, albo wzmiankowana mimochodem (np. pojęcie stopnia wierzchołka). Posługując się uznaną konwencją, można byloby zapisać wiele przedstawionych pojęć w sposób zwięzły i czytelny. Podobnie, prezentacja algorytmów w postaci kodu źródłowego systemu MATLAB, zamiast pseudokodu, nie ma uzasadnienia. Jaka jest argumentacja takiego podejścia? Cechy szczególne składni języka MATLAB nie stanowią przecież przedmiotu prezentowanych badań, a wręcz utrudniają lekturę niezaznajomionym czytelnikom.

W tym miejscu należy jednocześnie nadmienić o dysproporcji pomiędzy faktycznie szerokim omówieniem zagadnień optymalizacji globalnej w rozdz. 3, a niedoborem lub wręcz brakiem systematycznego przedstawienia kontekstu badawczego dot. rozdz. 4 i 5. W pracy trafiają się tam jedynie nieliczne cytowania na bieżąco. Jaki jest faktyczny stan badań dotyczący meritum tych rozdziałów i jaki (wybiegając do p. 3 oceny) jest oryginalny wkład Autorki?

Mimo uwag dot. deficytu formalizmów i przy założeniu, że stan badań cudzych dot. rozdz. 4 i 5 jest nikły, w opinii recenzenta rozprawa dostatecznie prezentuje ogólną wiedzę kandydata w dziedzinie.

2. Przedstawienie w rozprawie umiejętności samodzielnego prowadzenia pracy naukowej

Praca badawcza powinna być prowadzona rzetelnie, w obrębie własnej wiedzy i umiejętności, które to zostały ocenione pozytywnie powyżej. Kluczowym jej elementem powinien być krytycyzm, wyrażający się we właściwym tj. kompletnym odniesieniu do

zastanego stanu techniki, właściwym tj. systematycznym rozplanowaniu własnych eksperymentów oraz właściwym tj. dogłębnym i uczciwym oceniu ich rezultatów.

Autorka w sposób wzorowy odniosła się w rozdz. 3 do wyników swoich poprzedników implementując, obok własnych, cudze metody i weryfikując je na zbiorach danych uznanych w środowisku jako referencyjne. Natomiast w rozdz. 4 włożyła znaczący wysiłek w opracowanie mechanizmu generowania grafów o strukturze podobnej do sieci tramwajowej w Krakowie, w celu przedstawienia wyników kompresji w ujęciu statystycznym, czyli ogólniejszym. Eksperymenty numeryczne zostały zaplanowane i przeprowadzone zgodnie z regułami sztuki; sprawdzono działanie algorytmów dla zdefiniowanych zakresów parametrów, a przyjęte metody oceny uzyskiwanych wyników mają charakter wyłącznie ilościowy, więc obiektywne. Poszczególne wyniki zostały skomentowane i zinterpretowane.

Zastrzeżenia budzi jednakże pewna powierzchowność interpretacji wyników. Rozprawa doktorska ma prezentować oryginalne *rozwiązanie* problemu, dlatego też, komentując wyniki, należy oceniać ich użyteczność w zakresie zadania wyjściowego, a takim, co wynika z proporcji narracji, jest planowanie transportu tramwajowego w Krakowie. Tymczasem w rozdz. 4 widzimy kompleksowe badanie wpływu kompresji grafu na jakość rozwiązania – z pominięciem istotnej konkluzji, że zabieg ten jest nie do przyjęcia w praktyce. Czy planista pozwoliłby sobie na pogorszenie jakości rozwiązania o 1 punkt procentowy w zamian za przyspieszenie obliczeń o pół godziny?

Należy rzetelnie przeprowadzone badania uzupełniać rzetelną dyskusją. W przypadku rozdz. 4 niewykluczone, że kompresja grafu byłaby użyteczna w innym scenariuszu, który jednak nie został nakreślony. W podobnym tonie dokonano w rozdz. 3 porównania uzyskanego własnego rozwiązania z rzeczywistą siecią linii tramwajowych. Takie porównanie jest nieupoważnione choćby z powodu nieuwzględnienia w definicji problemu ograniczeń zasobowych, np. pojemności i częstotliwości kursowania pojazdów, przy czym ta ostatnia wynika m.in. z osiągnięcia limitu przepustowości na tzw. pierwszej obwodnicy. Uwzględnienie ograniczeń zasobowych powinno też stać się priorytetem dalszych prac badawczych, ale nie zostało wymienione w podsumowaniu.

Z pozostałych szczegółowych uchybień warsztatowych należy wskazać sposób porównywania rezultatów oparty o najlepszy uzyskany wynik (s. 84, „choosing the most favourable outcome”) oraz generalizację wnioskowania w oparciu o małą próbkę losową wyników charakteryzujących się dużym rozrzutem (Tab. 4.3, s. 129). Ponadto, razi i dezorientuje zamieszczenie definicji problemu w rozdz. 3, gdyż nie jest ono częścią prezentowanego algorytmu. Ostatnia istotna uwaga dotyczy sposobu doboru liczby rozważanych linii tramwajowych i ich maksymalnej długości – nie jest ona niczym uzasadniona; zresztą wiele przedstawionych decyzji projektowych pozostaje w pracy bez uzasadnienia ani dyskusji, które z pewnością wiele by jej przydały jakości.

Osobną kwestię stanowi w kilku miejscach stosowanie terminologii nieodpowiadającej standardowemu rozumieniu słowa. Przykładem może być mutacja – o ile stanowi niekwestionowany element algorytmu genetycznego, o tyle jej natura jest immanentnie losowa, co skutkuje (akceptowalnym i wręcz pożądanym) ryzykiem uzyskania rozwiązania gorszego niż wyjściowe, lecz jednocześnie o naturze eksploracyjnej. Tymczasem Autorka w obrębie mechanizmu mutacji wykonuje algorytm najkrótszej ścieżki. Czy została ostatecznie zachowana, a jeśli tak, to z czego wynika eksploracyjność tak zdefiniowanej operacji?

Istotny element oceny umiejętności badawczej stanowi również własny dorobek publikacyjny. Doktorantka jest autorką referatu pod tym samym tytułem co praca, przedstawionego na konferencji IEEE CEC w Krakowie w 2021, we współautorstwie z promotorem. Pracę udostępnił portal IEEE Xplore. Jest to co prawda jedyna opublikowana praca; natomiast doczekała się ona trzech różnorodnych cytowań. Można więc przedstawioną publikację uznać w sensie ilościowym za spełniającą zaledwie w stopniu minimalnym wymagania formalne; natomiast w sensie jakościowym za docenioną w środowisku.

W opinii recenzenta, elementy warsztatu badawczego obecne w rozdz. 3 i 4 świadczą o umiejętności samodzielnego prowadzenia pracy naukowej przez autorkę.

3. Przedstawienie w rozprawie oryginalnego rozwiązania problemu naukowego

Oryginalne rozwiązanie problemu naukowego powinno stanowić oś narracji w pracy – jednak w tej rozprawie nie jest ono akcentowane, co utrudnia wyodrębnienie dokonanych faktycznie nowatorskich. Nie oznacza to, że praca nie zawiera oryginalnych pomysłów, albo że uzyskane wyniki nie są użyteczne. Należałoby jednak oba aspekty eksponować i precyzyjnie zakreślać ich granice, aby wykonane badania z jak największym pożytkiem służyły otoczeniu. Autorskie i specyficzne dla zastosowania modyfikacje algorytmu genetycznego, równie specyficzna kompresja grafu, wreszcie inżynieria ścieżki transformacji sieci połączeń – wszystkie te inicjatywy z pewnością zawierają istotny i oryginalny wkład autorki; problem polega na tym, iż trudno określić, gdzie się on zaczyna.

Jak już wskazano, bezpośrednia przewaga lub użyteczność niektórych uzyskanych rezultatów w odniesieniu do krakowskiej sieci połączeń tramwajowych może być dyskusyjna. Wynika to z licznych uwarunkowań, których przyjęty model nie uwzględnia. Nie oznacza to bynajmniej, iż opracowane algorytmy są bezużyteczne, lecz że nie powinny być stosowane bezkrytycznie, a jedynie w określonych sytuacjach. Na przykład, w celu weryfikacji siatki połączeń w ściśle ustalonych warunkach – porze dnia lub tygodnia przypadającej poza szczytem komunikacyjnym, czyli przy zredukowanym ryzyku naruszenia istniejących a nieuwzględnianych w modelu ograniczeń zasobowych.

Podobną uwagę można by sformułować w odniesieniu do algorytmu kompresji, gdzie agregacja czasów przejazdu odbywa się wg scenariusza pesymistycznego tj. w normie maksimum. Rys. 4.8-10 stanowią tu empiryczne potwierdzenie dość oczywistego

zjawiska, że skompresowana sieć będzie sprawdzać się tylko w aspekcie analizy korespondującym z zasadą przyjętą przy agregacji (tu: maksymalnego czasu podróży). Nadal więc zaproponowana metoda, choć ograniczona, może być skuteczna – należałoby jedynie wskazać jej praktyczne zastosowanie w rozszerzonej dyskusji wyników, której tutaj zabrakło.

Ostatnia kwestia dotyczy nieodzownie wielokryterialnej natury rozważanego problemu – i tak też jest on traktowany w licznych pracach cytowanych i wykorzystywanych przez autorkę. Natomiast w implementacji ta wielokryterialność zostaje zneutralizowana przez prostą skalaryzację. Dlaczego Autorka nie zdecydowała się na jej zachowanie i przedstawienie wyników np. w formie zbioru rozwiązań niezdominowanych?

Mimo utrudnionej oceny rzeczowego zakresu wkładu oryginalnego, w opinii recenzenta jest on z pewnością wystarczająco istotny w odniesieniu do wymagań przepisowych.

Uważam, że rozprawa mgr Hanan Ali Mohammed Ba Ali spełnia wymagania Ustawy z dn. 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. 2023, poz. 742) w odniesieniu do rozpraw doktorskich. W związku z tym wnoszę o przyjęcie rozprawy i dopuszczenie autorki do dalszych, przewidzianych przepisami, etapów postępowania awansowego.

Maniur Kamola