

Recenzja rozprawy doktorskiej

mgra inż. Michała Wójcika

zatytułowanej:

Experimental analysis of the properties of models and decision support methods in the context of the use of holistic preferences

1. Problem badawczy i jego znaczenie

Jaki jest najważniejszy problem rozważany w rozprawie? Czy ma on charakter naukowy? Czy ma on znaczenie praktyczne?

Przedstawiona rozprawa doktorska ma formę cyklu czterech publikacji opatrzonych wprowadzeniem po polsku i po angielsku. Do materiałów dołączone zostały dwa oświadczenia autorów trzecich tj. dra Krzysztofa Ciomka oraz Prof. Mohammada Ghaderiego.

Pierwszy z artykułów pod tytułem „*Review and experimental comparison of ranking and choice procedures for constructing a univocal recommendation in a preference disaggregation setting*”, autorstwa M. Kadzińskiego, M. Wójcika i K. Ciomka, dotyczy metody UTA, jednego z popularniejszych podejść w wielokryterialnym wspomaganie decyzji. Metoda ta opiera się na addytywnej funkcji wartości (użyteczności). UTA jest popularna ze względu na swoją intuicyjność i przejrzystość, oraz wykorzystanie popularnego programowania liniowego do obliczania wyniku. W pracy autorzy zauważają, że względu na niekompletność informacji preferencyjnej (np. częściowe rankingi), metoda UTA może generować nieskończenie wiele rozwiązań, co utrudnia interpretację wyników. Aby temu zaradzić, stosuje się dodatkowe procedury, takie jak odporna regresja porządkowa lub metody agregujące, które tworzą jednoznaczne rekomendacje. W ramach badań przedstawionych w tej pracy autorzy przeprowadzili eksperymentalną analizę 35 procedur, oceniając je pod kątem trafności i odporności rekomendacji. Wykorzystano 7 miar jakości, które uwzględniały zarówno problemy wyboru, jak i rankingu. Badania objęły różne parametry problemów, takie jak liczba alternatyw, kryteriów i porównań. Dla problemów wyboru najlepsze wyniki osiągnęła procedura oparta na analizie stochastycznej, wskazująca alternatywę najczęściej wybieraną wśród wszystkich możliwych rozwiązań. Dla problemów tworzenia rankingu najlepsze były metody, które odzwierciedlały najpopularniejsze zależności między alternatywami w całej przestrzeni rozwiązań.

Drugi z artykułów zatytułowany „*Selection of a representative sorting model in a preference disaggregation setting: A review of existing procedures, new proposals, and experimental comparison*” autorstwa M. Wójcika, M. Kadzińskiego i K. Ciomka, porusza tematykę wielokryterialnego sortowania. W pracy znajdziemy porównanie różnych metod służących do przypisywania alternatyw do klas decyzyjnych. Wiele z tych metod opiera się na metodzie UTADIS, która jest adaptacją metody UTA i wykorzystuje pośrednią informację preferencyjną (np. przykładowe przypisanie alternatyw do klas) oraz wartości progowe do jednoznacznego sortowania alternatyw.

W artykule zostało przebadane 14 procedur sortowania, w czym trzy z nich to oryginalne propozycje zaproponowane w artykule. Przeprowadzone eksperymenty numeryczne wskazały metodę charakteryzującą się największą trafnością klasyfikacji oraz metodę a w zasadzie metody o największej odporności. Wyniki były statystycznie istotne (test Wilcoxon'a) i niezależne od parametrów problemu.

Trzecia praca zatytułowana „*Nature-inspired Preference Learning Algorithms Using the Choquet Integral*” autorów Michała Wójcika i Miłosza Kadzińskiego zajmuje się problematyką nauki preferencji rozumianych jako odpowiednia parametryzacja wybranych metod podejmowania decyzji. Głównym celem rozważanych metod jest klasyfikacja alternatyw do preferencyjnie uporządkowanych klas, przy jednoczesnym zapewnieniu, że rekomendacje są łatwe do interpretacji i uzasadnienia. Jedną z pomocnych, wykorzystywanych przez autorów, technik jest całka Choquet'a. Dzięki jej wykorzystaniu można modelować interakcje między kryteriami (zarówno pozytywne, jak i negatywne) a w rezultacie otrzymać bardziej ekspresywny model niż w podejściach tradycyjnych. W pracy autorzy wykorzystali szereg metod heurystycznych takich jak przeszukiwanie lokalne, zachłanne przeszukiwanie lokalne, programowanie genetyczne, przeszukiwanie fish-school czy optymalizacja oparta o rój cząstek.

Czwarta i ostatnia praca cyklu publikacji zatytułowana „*From investigation of expressiveness and robustness to a comprehensive value-based framework for multiple criteria sorting problems*”, autorstwa Miłosza Kadzińskiego, Michała Wójcika oraz Mohammada Ghaderiego zajmuje się problemem modyfikacji modeli decyzyjnych w kontekście wielokryterialnego sortowania. Autorzy zauważają, że tradycyjne modele, takie jak UTA i UTADIS, opierają się na addytywnej funkcji wartości, która zakłada monotoniczność preferencji i niezależność kryteriów. Jednak w praktyce często występują interakcje między kryteriami oraz niemonotoniczne preferencje, których te modele nie są w stanie odzwierciedlić. W odpowiedzi na tak zdefiniowany problem autorzy proponują rozszerzenie istniejących modeli o funkcje reprezentujące zależności między parami kryteriów. Rozważane modele zostały zaadaptowane do problemów sortowania, a ich skuteczność została oceniona w ramach analizy eksperymentalnej. Przeprowadzone badania skupiały się na dwóch kluczowych aspektach: ekspresywność modeli tj. zdolność do odzwierciedlenia złożonych preferencji decydenta oraz odporność rekomendacji tj. stabilność i spójność wyników w różnych warunkach.

Prezentowane artykuły mają charakter w większości analityczny tj. skupiają się przede wszystkim na analizie istniejących rozwiązań a proponowanie nowych ulepszonych rozszerzeń i podejść wydaje się celem drugoplanowym. Wyjątkiem tu może być praca trzecia, która nie stara się analizować istniejących rozwiązań a od razu zawiera propozycję nowego podejścia wykorzystującego całkę Choquet'a i do pewnego stopnia praca czwarta, która za tytułowy cel stawia sobie zaproponowanie pewnego rozwiązania służącego sortowaniu wielokryterialnemu. Ponadto łącznikiem pomiędzy pierwszą, drugą i czwartą pracą jest przedmiot analizy - metoda UTA wraz z różnymi jej wariantami.

Pierwsza, druga i czwarta praca zostały opublikowane w bardzo dobrych czasopiśmie z zakresu badań operacyjnych. Trzecia to praca konferencyjna opublikowana w materiałach uznanej konferencji GECCO w roku 2024. *Miejsca publikacji prac potwierdzają istotną wartość naukową proponowanych analiz i rozwiązań. Przeprowadzone analizy i otrzymane wyniki mają istotną wartość praktyczną i mogą posłużyć w konkretnych przypadkach do oceny zasadności wyboru takiego lub innego rozwiązania. Znaczenie otrzymanych wyników może być do pewnego stopnia ograniczone wzrostem skomplikowania rozważanych wariantów metod decyzyjnych. Np. wraz z wprowadzaniem*

modyfikacji do metody UTA komplikują się warunki modelu będącego przedmiotem optymalizacji co może wiązać się z jego mniejszą atrakcyjnością w oczach potencjalnych odbiorców.

Wszystkie prace zawierają wyniki eksperymentów numerycznych, których zadaniem jest ocena jakości osiągniętych rezultatów. W różnych pracach oceniane są różne parametry badanych metod od odporności i zaproponowanych miar jakości takich jak np. FRAI (first rank acceptability index), MRAI (mean rank acceptability index) w pracy pierwszej do odporności i siły ekspresji w pracy czwartej i ostatniej. Otrzymane wyniki są sprawdzone pod względem istotności statystycznej testem Wilcozona dla par obserwacji. Na tej podstawie w wybranych publikacjach jest generowany diagram Hassego będący reprezentacją pewnego częściowego porządku na zbiorze badanych metod indukowanego otrzymanymi wynikami. Po lekturze artykułów nie jest jednak dla mnie do końca jasne w jaki sposób w poszczególnych przypadkach te diagramy są budowane. Tj. co sprawia, że pomiędzy pewnymi węzłami istnieją krawędzie a pomiędzy innymi nie istnieją. I dalej, czy relacje reprezentowane przez te diagramy są przechodnie? Wykorzystanie testu statystycznego Wilcozona w celu potwierdzenia istotności obserwowanej zmiany rodzi też pytanie natury ogólnej. Otóż, w przypadku tego typu testów statystycznych do wyznaczenia poziomu istotności obserwacji służy tzw. wartość *p-value* wskazująca na prawdopodobieństwo otrzymania obliczonego wyniku przy założeniu prawdziwości hipotezy zerowej. Zwykle przyjmuje się, że obliczony wynik jest istotny statystycznie, gdy wartość *p* jest nie większa niż 0,05. W praktyce jednak bardzo często wartość *p* maleje wraz ze wzrostem wielkości badanej próby. Stąd pojawia się pytanie na ile zasadne jest korzystanie z testu Wilcozona (i testów statystycznych w ogóle) w przypadku eksperymentów Montecarlo, gdzie wielkość próby z hipotetycznej populacji może być dowolnie duża. Tj. nawet dość dużą początkowo wartość *p* można próbować zredukować poprzez wytrwałą generację przypadków testowych.

Choć w każdej z prac można znaleźć odpowiedni fragment opisujący przeprowadzone testy numeryczne to w prezentacji wyników trudno doszukać się odniesień do zmierzonych czasów wykonania poszczególnych eksperymentów. W szczególności, dla oceny praktycznej przydatności proponowanych rozwiązań dobrze byłoby poznać rzeczywistą zależność pomiędzy wielkością zadania (wielkością rozważanego modelu) a czasem wykonania obliczeń przy założonej pewnej platformie sprzętowej, na której wykonywane były testy. Odpowiedź na to pytanie jest o tyle istotna, że w niektórych przypadkach istnieje niebezpieczeństwo, że czas obliczeń może wzrastać wykładniczo względem wielkości modelu (ilości zmiennych, które trzeba uwzględnić w procesie optymalizacji). Na przykład w pracy nr. 3 jak się wydaje, ilość wartości miary rozmytej określonych na wszystkich elementach zbioru potęgowego nad zbiorem kryteriów będzie wzrastać wykładniczo względem ilości kryteriów. A jeśli to prawda, to powstaje pytanie jaka jest rzeczywista (praktyczna) granica ilości kryteriów w modelu dla użytkowników dysponujących różnymi platformami obliczeniowymi. Oprócz rozważań natury eksperymentalnej warto by również poznać teoretyczną, asymptotyczną złożoność obliczeniową poszczególnych rozwiązań, ze wskazaniem, które miejsca w rozważanych algorytmach mają znaczenie rozstrzygające dla takich a nie innych oszacowań.

2. Wkład autora

Jaki jest najważniejszy wkład autora opisywane w rozprawie?

Wkład autora rozprawy można próbować szacować na podstawie zamieszczonych oświadczeń oraz pozycji na liście autorów. W pracach nr. 2 i 3 autor rozprawy znajduje się na pierwszym miejscu co sugeruje jego istotny wkład w powstanie tych dwóch prac. Tj. można domniemywać, że jego udział był wiodący i kluczowy dla powstania co najmniej tych dwóch

publikacji. W przypadku prac 1 i 4 do rozprawy zostały dołączone oświadczenia autorów trzecich niebędących ani autorem rozprawy ani jego promotorem. Oświadczenia te wskazują na pewien, ale raczej niekluczowy, udział zarówno dra Ciomka (praca 1) jak i prof. Ghaderi'ego (praca 4). Pomimo tych oświadczeń nie jest jasny podział pracy pomiędzy autorem rozprawy oraz jego promotorem, prof. Miłoszem Kadzińskim. Stąd w świetle przedstawionych materiałów *sprawa wkładu autora rozprawy we wskazane osiągnięcia wymaga doprecyzowania w trakcie prezentacji na obronie pracy doktorskiej.*

Oceniając publikacje wskazane w ocenianym cyklu, w oderwaniu od precyzyjnie zdefiniowanych obszarów odpowiedzialności poszczególnych autorów, trzeba docenić zarówno zaprezentowane i opisane w nich eksperymenty numeryczne istniejących już w literaturze rozwiązań jak i propozycje nowych i zmodyfikowanych a następnie przetestowanych numerycznie metod. *W obu tych aspektach przedstawione prace wnoszą istotny wkład w rozwój metod podejmowania decyzji a co za tym idzie badań operacyjnych.* Wkład ten określiłbym jak przede wszystkim polegający na analizie własności dostępnych rozwiązań i dostarczenia wskazówek tym osobom, które chcą w praktyce wykorzystywać różne warianty metody UTA w procesach wspomaganego podejmowania decyzji. W tym miejscu warto zauważyć, że prace 1,2 i 4 ukazały się w renomowanych periodykach zajmujących się tematyką badań operacyjnych i metod podejmowania decyzji.

3. Poprawność

Czy stwierdzenia zawarte w rozprawie są godne zaufania? Czy uzasadnienia są poprawne? Wskaż zauważone słabości i błędy. Wskaż także te aspekty dotyczące poprawności, które są najbardziej wartościowe (elegancja dowodów, plan eksperymentów, analiza danych empirycznych, jakość prototypowego oprogramowania/sprzętu...).

Poprawność prezentowanych koncepcji w dużej mierze była przedmiotem dociekań recenzentów publikacji tworzących oceniany cykl artykułów. Dla tego dublowanie ich pracy w moim przekonaniu ma sens ograniczony. *Mogę jedynie potwierdzić, że po przeczytaniu wskazanych prac wyniki w nich zawarte uważam za wartościowe i godne zaufania.*

W przedstawionych materiałach, jednakże niewiele można znaleźć informacji dotyczących implementacji zarówno samych testowanych rozwiązań jak i przeprowadzonych testów. Stąd po przeglądnięciu cyklu niewiele można powiedzieć o poprawności koncepcji przyświecających części implementacyjnej wykonanych prac. W tym sensie uważam, że *wątek ten*, jako ciekawy z praktycznego punktu widzenia, *powinien zostać podjęty przez doktoranta* w ramach prezentacji w trakcie obrony pracy.

Pisząc o jakości przedstawionych wyników nie można nie wspomnieć o momentami imponującej liczbie testowanych metod i ich wariantów (35 rozważanych metod w pierwszej z rozważanych prac). *Stąd jedną z głównych i bardziej wartościowych cech prezentowanych rozwiązań jest bardzo przekrojowe i szerokie ujęcie dyskutowanych w pracach tematów.*

Dla porządku trzeba wspomnieć również (choć nie dotyczy to rozprawy jako takiej) o używaniu przez autora rozprawy niepotrzebnie długich i wielokrotnie złożonych zdań w streszczeniach dołączonych do rozprawy tj. cyklu czterech publikacji. W rezultacie jasność przekazu, a co za tym idzie, przydatność obu streszczeń (zarówno angielskiego jak i polskiego) w mojej ocenie ucierpiała.

4. Wiedza kandydata

Które z rozdziałów (lub sekcji w rozdziałach) rozprawy omawiają istniejący stan wiedzy i dzięki temu potwierdzają ogólny stan wiedzy kandydata w zakresie Informatyki? Jakie obszary tych dyscyplin zostały omówione w tych rozdziałach/sekcjach? Jaka jest opinia recenzenta o jakości tych rozdziałów/sekcji? Jaka jest opinia recenzenta o bibliografii? Na ile bibliografia jest kompletna? Prosimy o podanie innych argumentów za lub przeciw stwierdzeniu, że kandydat posiada ogólną wiedzę w dyscyplinie Informatyka techniczna i telekomunikacja.

Artykuły składające się na cykl publikacji zostały opublikowane albo w czasopismach przypisanych do dyscypliny Informatyka Techniczna i Telekomunikacja albo na konferencjach (artykuł nr. 3) o charakterze informatycznym. Przeprowadzone badania we wszystkich tych pracach wymagały dużego nakładu czasu przeznaczanego na implementację badanych metod oraz przeprowadzenie niezbędnych eksperymentów numerycznych. Na tej podstawie można wnioskować, zakładając, że główny ciężar niezbędnej implementacji spoczywał na autorze rozprawy, że posiada on niezbędną wiedzę programistyczną potrzebną do wykonania takiego zadania. Część prac, szczególnie zaś praca trzecia, dokumentuje znajomość szeregu algorytmów heurystycznego poszukiwania rozwiązań w przestrzeni rozwiązań dopuszczalnych. Implementacja i wykorzystanie tych algorytmów wskazuje zarówno na posiadanie przez doktoranta pewnej wiedzy zakresu algorytmiki jak i sztucznej inteligencji. W tym kontekście warto przypomnieć, że algorytmy takie są klasyfikowane jako rozwiązanie mieszczące się w granicach tego co jest nazywane sztuczną inteligencją właśnie (rozdział 3, Solving By Searching, S. Russel & P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. 2021, Pearson).

Ze względu na zakres tematyczny prezentowanych artykułów można zakładać, że doktorant posiada szeroką wiedzę w zakresie metod i systemów wspomagania decyzji opartych o addytywną funkcję użyteczności (UTA), wiedzę z zakresu programowania liniowego, optymalizacji oraz wykorzystania całki Choquet'a w modelach, które muszą uwzględniać istnienie zależności (np. efektu synergii) pomiędzy rozważanymi kryteriami bądź wariantami decyzyjnymi. Wszystkie wskazane tu elementy stanowią wiedzę z zakresu badań operacyjnych, a stąd pośrednio przynależą do informatyki, w zakresie w jakim jest ona używana w systemach wspomagania podejmowania decyzji. *Konkludując uważam, że przedstawiona praca doktorska będąca w swojej istocie cyklem czterech publikacji wskazuje na posiadanie przez doktoranta wiedzy zarówno z zakresu informatyki jak i badań operacyjnych.*

5. Inne uwagi¹

Wprowadzenia w swojej istocie nie próbują dokonać syntezy osiągniętych wyników z poszczególnych publikacji a są raczej rozszerzonym streszczeniem materiału, który można znaleźć w poszczególnych publikacjach. Szkoda, bo streszczenia mogłyby posłużyć za miejsce, w którym można by umieścić treści niekoniecznie obecne w artykułach. Na przykład z przyjemnością powitałbym w streszczeniu krótkie, ale wyczerpujące wprowadzenie do metody UTA i jej odmian takich jak UTADIS opatrzone prostym przykładem studialnym.

¹ Opcjonalnie

Na podstawie streszczeń można jednak wyrobić sobie zdanie o cechach wspólnych prac składających się na prezentowany cykl takich jak np. duży nacisk na numeryczne badania eksperymentalne czy wykorzystanie metody UTA.

6. Podsumowanie

Biorąc pod uwagę opinie zaprezentowane w poprzednich punktach i wymagania zdefiniowane przez art. 187 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (z późniejszymi zmianami)² moja ocena rozprawy pod względem trzech podstawowych kryteriów jest następująca:

A. Czy rozprawa zawiera oryginalne rozwiązanie problemu naukowego? (wybierz jedną opcję stawiając znak **X**)


<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE

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

<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE

C. Czy kandydat posiada umiejętność samodzielnego prowadzenia pracy naukowej?

<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE



Podpis

² <http://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20190000276>