

dr hab. Barbara Pękała, prof. UR  
Instytut Informatyki  
Kolegium Nauk Przyrodniczych  
Uniwersytetu Rzeszowskiego  
ul. Pigoń 1, 35-310 Rzeszów

Rzeszów, 9 maja 2022 r.

## **Recenzja rozprawy doktorskiej**

**mgr. Pawła Ładyżyńskiego**

### *Modelowanie niepewnych preferencji w systemach rekomendacyjnych*

Rozprawa napisana w Instytucie Badań Systemowych Polskiej Akademii Nauk z siedzibą w Warszawie pod kierunkiem: prof. dr hab. Przemysława Grzegorzewskiego

#### **1 Ogólna charakterystyka pracy, obszar problemowy pracy i uzyskane wyniki**

Systemy rekomendacyjne to systemy informatyczne, których głównym celem jest proponowanie zawartości serwisu lub produktów w jak najlepszym stopniu odpowiadających zainteresowaniom lub aktualnym potrzebom użytkowników danej platformy. W systemach rekomendacyjnych modelowanie preferencji pełni bardzo ważną funkcję. W praktycznych zastosowaniach zadanie to jest utrudnione w wyniku szeroko rozumianej niepewności dostępnych danych. W praktyce sytuacja, gdy mamy do czynienia w systemach rekomendacyjnych z pełnym rankingiem, występuje bardzo rzadko. W pracy doktorskiej zbadano, czy reprezentacja niepewnych systemów preferencji w formie intuicjonistycznych zbiorów rozmytych Atanassova pozwala modelować niepewne systemy preferencji w wielu typach systemów rekomendacyjnych i usprawniać różne aspekty ich działania. W związku z podstawowym celem systemów rekomendacyjnych polegających na proponowaniu produktów

w jak najlepszym stopniu odpowiadających zainteresowaniom lub aktualnym potrzebom użytkowników. Autor rozprawy przyjął że jednym z kluczowych zagadnień związanych z tworzeniem rekomendacji jest określenie podobieństwa pomiędzy preferencjami poszczególnych użytkowników. W pracy poddano analizie dostępne w literaturze miary podobieństwa zdefiniowane dla intuicjonistycznych zbiorów rozmytych Atanassova pod kątem ich zastosowania w systemach rekomendacyjnych oraz zaproponowano nową miarę podobieństwa posiadającą oczekiwane własności. Zaproponowano nowatorską graficzną metodę oceny rekomendacji pozwalającą na analizę typów rekomendacji generowanych przez różne systemy rekomendacyjne. Przedstawiono modyfikacje dostępnych w literaturze algorytmów rangowania etykiet wykorzystujących modele probabilistyczne. A także zaproponowano system rekomendacyjny łączący metody automatycznego przetwarzania tekstu z analizowanym modelem reprezentacji systemów preferencji, który rekomenduje produkty bazując na zagregowanej wiedzy płynącej z opinii o produktach w formacie tekstowym, uwzględniając jednocześnie indywidualne preferencje danego użytkownika. W rozprawie badania przeprowadzono zarówno pod kątem teoretycznym jak i metodami eksperymentalnymi, wykorzystującymi rzeczywiste dane.

Teza rozprawy jest następująca: "zastosowanie teorii zbiorów rozmytych i intuicjonistycznych zbiorów rozmytych Atanassova (IFS, ang. Intuitionistic Fuzzy Sets), jako modelu reprezentacji niepewnych systemów preferencji pozwala modelować niepewne systemy preferencji w wielu typach systemów rekomendacyjnych i usprawniać różne aspekty ich działania". Została osiągnięta dzięki temu że zbiory rozmyte oraz IFS-y posiadają dobrze rozwinięty i zbadany aparat matematyczny radzący sobie w naturalny sposób z różnymi typami niedoskonałości czy niepewności przetwarzanych informacji. Ponadto dostępne są metody porównywania zbiorów rozmytych, ich agregowania, a także miary stopnia ich niepewności. Aspekty te są bardzo istotne z punktu widzenia analizy preferencji użytkowników i generowania rekomendacji. Zakres rozprawy obejmuje algorytmy rekomendacyjne bazujące na filtrowaniu zespołowym, rangowaniu etykiet (ang. label ranking) z wykorzystaniem modeli probabilistycznych oraz systemy bazujące na zawartości (ang. content-based) z wykorzystaniem automatycznych metod przetwarzania tekstu. Teza rozprawy była badana zarówno pod kątem teoretycznym jak i metodami eksperymentalnymi, wykorzystującymi rzeczywiste dane pochodzące z popularnych serwisów internetowych. Jednym z elementów rozprawy jest zdefiniowanie autorskich metod tworzenia rekomen-

cji wykorzystujących model reprezentacji preferencji w formie IFS-ów i porównanie ich z rozwiązaniami dostępnymi w literaturze. Ponieważ w procesie generowania rekomendacji bardzo istotnym elementem jest porównywanie obiektów i użytkowników między sobą, a dokładniej badanie ich podobieństwa, rozprawa obejmuje również analizę dostępnych miar porównujących IFS-y w kwestii ich przydatności do porównywania systemów preferencji w systemach rekomendacyjnych. Oprócz analizy algorytmów pod kątem takich kryteriów jak dokładność rekomendacji i czas obliczeń, znaczna część rozważań poświęcona została w rozprawie różnym strategiom decyzyjnym (np. związanym z poziomem ryzyka). Ważnym elementem tych rozważań było wykazanie, jak duże znaczenie na strategię rekomendacyjną danego algorytmu ma wybór funkcji agregującej. Dodatkowo, zaproponowana została autorska metoda graficznej ewaluacji generowanych rekomendacji, pozwalająca na dostosowanie strategii rekomendacyjnej do typu użytkownika i jego upodobań w kwestii podejmowania decyzji.

Podsumowując tę część stwierdzam, że Autor podjął ważny, aktualny a zarazem trudny problem naukowy. Przeprowadził analizę istniejących podejść, zaproponował autorskie rozwiązania, przedstawił ich zastosowania w wybranych aspektach badawczych oraz zwerifikował ich skuteczność, co pozwoliło ocenić opracowane rozwiązania.

## **2 Kompozycja i zawartość pracy**

W rozdziale 1 opisano najpopularniejsze grupy stosowanych obecnie systemów rekomendacyjnych. W rozdziale 2 poddano analizie miary podobieństwa i rozbieżności mogące służyć do porównywania preferencji użytkowników. Przeanalizowano w pracy wiele dostępnych w literaturze miar porównujących IFS-y i zaproponowano miarę podobieństwa posiadającą nie tylko niezbędne własności, by mogła być stosowana do porównywania niepewnych systemów preferencji oraz pozwalającą również na kontrolę poziomu wiedzy użytkowników określonych jako podobnych do nowego użytkownika. Ponadto, w rozdziale 2.3.3 zaproponowano graficzną metodę oceny rekomendacji pozwalającą na badanie typów rekomendacji generowanych przez różne systemy rekomendacyjne w celu dostosowywania ich do indywidualnych cech podejmowania decyzji. Reprezentacja niepewnych systemów preferencji w formie IFS-ów posłużyła w rozdziale 3 do istotnego skrócenia czasu obliczeń algorytmu rekomendacyjnego. W rozdziale 4 pokazano natomiast jak łącząc metody au-

tomatycznego przetwarzania tekstu z IFS-owym modelem reprezentacji można zbudować system rekomendujący produkty na podstawie zagregowanej wiedzy płynącej z opinii o produktach, uwzględniający jednocześnie indywidualne preferencje danego użytkownika.

Podsumowując tę część recenzji uważam, iż praca jest obszerna i pokazuje dużą wiedzę Doktoranta. Widoczne jest naukowe podejście do realizowanego zadania. Rozprawa przedstawia dużą zawartość merytoryczną. Autor pracy wykazał się także dobrą umiejętnością redakcyjną.

### **3 Opinia o rozprawie i oryginalność osiągnięć**

Przede wszystkim należy podkreślić, iż problem badawczy podejmowany w rozprawie jest bardzo ważny i złożony. Przeprowadzone w ramach niniejszej rozprawy doktorskiej badania potwierdziły, że posłużenie się reprezentacją niepewnych systemów preferencji w formie IFSów pozwala udoskonalić różne aspekty działania dostępnych w literaturze algorytmów rekomendacyjnych, jak również przydaje się do budowania nowych rozwiązań.

Za jedno z najważniejszych osiągnięć przedstawionych w pracy uważam metodę graficznej ewaluacji generowanych rekomendacji pozwalającą na dostosowanie strategii rekomendacyjnej do typu użytkownika i jego upodobań w kwestii podejmowania decyzji. Ta dodatkowa informacja na temat proponowanych rekomendacji nabiera jeszcze większego znaczenia w przypadku, gdy celem użytkownika jest zakup produktu o znacznej wartości.

Drugi w mojej opinii interesujący a zarazem złożony aspekt pracy to określenie warunków, które powinna spełniać miara podobieństwa porównująca systemy preferencji, aby mogła być stosowana w systemach rekomendacyjnych. Wykorzystując reprezentacje systemów preferencji w formie IFS-ów przeanalizowano własności różnych miar porównujących IFS-y pod kątem ich wykorzystania do tworzenia rekomendacji. Wykazano, że miara odległości pomiędzy IFS-ami zdefiniowana jako odległość euklidesowa jest odpowiednia do konstrukcji miary podobieństwa, która z powodzeniem została wykorzystana w systemach rekomendacyjnych. Dodatkowo, zaproponowano nową miarę podobieństwa pomiędzy systemami preferencji, pozwalającą na promowanie użytkowników posiadających szerszą wiedzę na temat różnych obiektów. Badania i symulacje przeprowadzone na rzeczywistym zbiorze danych wykazały, że zaproponowana miara podobieństwa jest efektywniejsza od innych badanych miar zarówno pod względem ewaluacji algorytmu jak

również w samym procesie tworzenia rekomendacji (algorytm filtracji zespołowej).

Kolejnym, ważnym elementem rozprawy jest aspekt praktyczny obrazujący różnorodność zastosowania zaproponowanych metod. W szczególności przedstawiono metodę, która potrafi automatycznie analizować i wydobywać informacje z tysięcy recenzji dostępnych w internecie (serwis Amazon). Następnie, zaprezentowano konstrukcję systemu rekomendującego produkty na podstawie zagregowanej wiedzy pochodzącej z opinii o produktach, uwzględniającego jednocześnie indywidualne preferencje danego użytkownika. Przeprowadzone w ramach rozprawy doktorskiej badania potwierdziły, że posłużenie się reprezentacją niepewnych systemów preferencji w formie IFSów pozwala udoskonalić różne aspekty działania dostępnych w literaturze algorytmów rekomendacyjnych, a także przydaje się do budowania nowych rozwiązań.

Podsumowując uważam, że wykonana w ramach doktoratu praca stanowi istotny wkład w rozwój metod dotyczących zastosowania teorii zbiorów rozmytych i intuicjonistycznych zbiorów rozmytych Atanassova do modelowania reprezentacji niepewnych systemów preferencji. Rozprawa prezentuje wysoki poziom naukowy oraz implementacyjny związany z przedstawionymi zastosowaniami i w pełni zasługuje na ocenę pozytywną. Praca napisana jest starannie, bez większych pomyłek językowych. Zawarte w rozprawie przykłady pozwalają lepiej zobrazować i wyjaśnić działanie proponowanych metod i algorytmów.

## 4 Uwagi i problemy do dyskusji

Praca napisana jest starannie, a wyniki w niej przedstawione świadczą o szerokiej wiedzy Autora z poruszanej tematyki. Wartość merytoryczną pracy potwierdza fakt opublikowania części przedstawionych wyników w indeksowanych materiałach konferencyjnych i rozdziałach w monografiach wydanych przez renomowane wydawnictwa, a także cenionych czasopismach wg. wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych opracowanego przez MEiN (1.12.2021) (7 prac w bazie Web of Science, H-Index 4).

W pracy autor wykazuje się dojrzałością badawczą poprzez dostrzeżenie luk w swoim rozumowaniu i wskazuje dalsze kierunki rozwoju rozpoczętych badań. Wiele kwestii poruszonych w niniejszej rozprawie jest warte dalszego badania. W szczególności, w aspekcie

wykorzystania miar podobieństwa w systemach rekomendacyjnych, które wykazują pozytywne własności w przypadku porównywania niepewnych preferencji. Zatem

- Interesującym będzie znalezienie dodatkowego warunku, który pozwoliłby uzyskać ogólną postać miary podobieństwa spełniającej warunki (C-1)–(C-3).
- W niniejszej rozprawie ograniczono rozważania do operacji na IFS-ach bazujących na operatorach min oraz max, natomiast z punktu widzenia tworzenia rekomendacji, wydaje się więc istotnym zastosowanie operatorów bazujących na innych operatorach, w szczególności różnych klas agregacji.
- Ważnym uzupełnieniem analiz poczynionych w pracy będzie zbadanie:
  1. różnych miar podobieństwa (innych niż rozważanych w pracy) o wartościach rzeczywistych (typ I);
  2. innych metod konstrukcji miar podobieństwa;
  3. miar podobieństwa o wartościach intuicjonistycznych (typ II) i porównanie ich efektywności z miarami typu I.

Powyższe uwagi i wskazówki oraz kilka literówek nie wpływają na wysoką wartość merytoryczną pracy.

## 5 Konkluzja

Reasumując rozprawa jest cennym wkładem do teorii zbiorów rozmytych i z zakresu aplikacyjnego, czyli zastosowania teorii zbiorów rozmytych i intuicjonistycznych zbiorów rozmytych Atanassova do modelowania niepewności w systemach preferencji. Pan Paweł Ładyżyński przedstawił rozprawę doktorską stanowiącą oryginalne rozwiązanie problemu naukowego z ww. zakresu. Opracowane metody zostały zastosowane i zweryfikowane dla różnych problemów/danych rzeczywistych. Autor wykazał się dużą wiedzą w zakresie tematyki rozprawy, umiejętnością pracy naukowej oraz znajomością metod badawczych. Osiągnięte wyniki świadczą o bardzo dobrym przygotowaniu Autora do pracy naukowej.

Rozprawa doktorska mgr Pawła Ładyżyńskiego spełnia wymagania art. 13 ustawy z dnia 14 marca 2003 r. *o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki* (Dz. U. Nr 65, poz. 595, stan prawny na dzień 30 września 2011 r.) i w

związku z art. 179 ust. 1 ustawy z dnia 3 lipca 2018 r. *Przepisy wprowadzające ustawę – Prawo o szkolnictwie wyższym i nauce* (Dz. U. z 2018 r. poz. 1669 z późn. zm.) wnoszę o dopuszczenie mgr Pawła Ładyżyńskiego do dalszych etapów przewodu doktorskiego.

Ponadto wnoszę o wyróżnienie tej pracy z racji na całościowe potraktowanie problemu badawczego, przejrzystość rozprawy i dobry poziom wyników.

A handwritten signature in black ink, appearing to read "Rafał Bala". The signature is written in a cursive style with a large initial 'R'.