



Prof. dr hab. inż. Witold Abramowicz  
Kierownik Katedry Informatyki Ekonomicznej  
Uniwersytet Ekonomiczny w Poznaniu

Poznań, 5 stycznia 2023

Recenzja rozprawy doktorskiej mgr inż. Michała  
Malinowskiego *Algorytm rekomendacji bazujący na sesjach  
rekomendacji działających na podstawie zachowań  
użytkowników oraz atrybutów*

Przedmiotem recenzji jest praca doktorska Pana mgr inż. Michała Malinowskiego przedłożona Radzie Dyscypliny Naukowej *Informatyka Techniczna i Telekomunikacja* Wojskowej Akademii Technicznej im. Jarosława Dąbrowskiego. Recenzja została sporządzona na podstawie Uchwały Rady nr 35/RDN ITiT/2022 z dnia 8 listopada 2022 roku.

### Przedmiot badań

Praca poświęcona jest badaniom prowadzącym do opracowania nowego algorytmu rekomendacji wykorzystującego analizę zachowań użytkowników oraz statyczne atrybuty obiektów, będące przedmiotem rekomendacji.

Systemy rekomendacyjne są jedną z klas systemów filtrowania informacji<sup>1</sup>. Mają szerokie zastosowanie w nauce, na przykład w zakresie wskazywania publikacji przydatnych dla realizacji celów badawczych spośród publikacji gromadzonych w bibliograficznych bazach danych. W zastosowaniach technicznych systemy rekomendacji wspomagają procesy projektowania zalecające dobór odpowiednich elementów tworzących złożone układy (np. układy pamięci do zastosowywania w konkretnym układzie sterowania uwzględniające ich atrybuty techniczne, niezawodnościowe, bezpieczeństwa oraz ekonomiczne). W procesach gospodarczych systemy rekomendacji mogą wskazywać przydatne w podejmowaniu decyzji źródła open data<sup>2</sup>.

<sup>1</sup> Witold Abramowicz, *Filtrowanie informacji*, Wydawnictwo UEP 2008, 626 pp.

<sup>2</sup> Krzysztof Węcel, Witold Abramowicz, Elżbieta Lewańska, Piotr Kałużny, Włodzimierz Lewoniewski, Szymon Wiczorek, *Nowy surowiec – otwarte zasoby danych dla polskiej gospodarki*, Ministerstwo rozwoju, Warszawa 2019, 357 pp.

Szerokie zastosowanie mają systemy rekomendacji w zakresie obronności. Pomagają one w eliminacji fake news<sup>3</sup> z przestrzeni informacyjnej podejmujących decyzje. Opisywana klasa systemów informacyjnych wspomaga nawigację morską w zakresie rekomendowania nawigacji w zależności od bezpieczeństwa żeglugi (np. czynników pogodowych, prawdopodobieństwa nieuprawnionego przejęcia ładunku), czy też analizy ekonomicznej oraz wpływu na zanieczyszczenie środowiska<sup>4</sup>.

Najwięcej użytkowników, spośród systemów rekomendacji, mają zapewne wyszukiwarki internetowe. Ich celem jest rekomendowanie stron internetowych relewantnych dla użytkowników formułujących zapytanie odpowiadające ich potrzebom informacyjnym. Miara relewancji określająca relacje pomiędzy stroną internetową a zapytaniem nie jest miarą boolowską. Dodatkowo jest miarą subiektywną. Oznacza to, że dla tak samo sformułowanego zapytania w odczuciu różnych użytkowników wyszukiwarka powinna wskazywać na inne strony internetowe. Mało tego, nawet dla tego samego użytkownika wartość miary relewancji zmienia się w czasie. W konsekwencji systemy rekomendacji muszą w coraz większym stopniu oceniać aspekty socjologiczne oceny relewancji informacji, które wpływają na oceny behawioralne poszczególnych odbiorców rekomendowanych informacji.

Innym przykładem zastosowania systemów rekomendacji jest handel. Obszar takich zastosowań jest bardzo szeroki. W zakresie wspomagania zarządzających organizacjami handlowymi mogą to być na przykład: systemy rekomendacji produktów, które przez daną organizację powinny być sprzedawane; systemy rekomendacji dostawców produktów dla organizacji; wielokryterialne systemy rekomendacji lokalizacji punktów sprzedaży. W obszarze handlu elektronicznego systemy rekomendacji znajdują także różnorodne inne zastosowania dla potrzeb systemów SCM i CRM.

Bardzo interesującym obszarem zastosowań może być handel elektroniczny. W tym zakresie Doktorant skupił się na systemach rekomendacji dla użytkowników końcowych będących osobami fizycznymi dokonującymi zakupów przez Internet.

Zastosowania systemów rekomendacji w zakresie obsługi klientów końcowych są bardzo szerokie. Zakładając, że około 50% ludzkości jest wykluczona cyfrowo, a niewykluczeni korzystają dwa razy dziennie z jakiejś formy rekomendacji elektronicznej, to rocznie przeprowadza się przynajmniej kilka tysięcy miliardów (czyli kilka bilionów) transakcji w systemach rekomendacyjnych. Badania usprawniające działanie systemów rekomendacji

---

<sup>3</sup> Witold Abramowicz, Krzysztof Węcel, Milena Stróżyńska, Piotr Stolarski, Włodzimierz Lewoniewski, Marcin Sawiński, Ewelina Księżniak, OpenFact – narzędzia weryfikacji wiarygodności źródeł informacji i detekcji fałszywych informacji z wykorzystaniem metod sztucznej inteligencji – raport, Warszawa 2022, 97 pp.

<sup>4</sup> Witold Abramowicz, Krzysztof Węcel, Jacek Małyżko, Dominik Filipiak, Data analysis in the maritime domain, Wydawnictwo UEP 2022, 355 pp.



mają kluczowe znaczenie dla większości aspektów życia społeczeństw a także każdego człowieka z osobna.

**Temat i zakres badań został wybrany przez Doktoranta poprawnie i ma głębokie uzasadnienie społeczne i gospodarcze.**

### Struktura dysertacji

Praca składa się z siedmiu głównych części. W część monograficznej Doktorant przedstawia szkice dotyczące: handlu elektronicznego (rozdział 2), obszaru rekomendacji (rozdział 3) oraz algorytmów wykorzystujących analizy asocjacyjne (rozdział 4). W drugiej części dysertacji, stanowiącej prezentację dorobku badawczego Doktoranta, opisano algorytm rekomendacji sesji (rozdział 5), implementację tegoż algorytmu (rozdział 6) oraz jego ocenę (rozdział 7).

**Kompozycja dysertacji jest poprawna.**

### Część monograficzna

Daleki jestem od poglądu Doktoranta, że *Systemy rekomendacji stanowią część praktycznie każdego nowoczesnego urzędnienia i aplikacji*<sup>5</sup>, mimo że w pierwszej części recenzji wskazują na ich szerokie znaczenie. Mimo tak reprezentowanego poglądu, Doktorant bez specjalnego uzasadnienia swojej decyzji skupia się na zastosowaniach systemów rekomendacyjnych w handlu elektronicznym, a właściwie tylko handlu elektronicznym wykorzystywanym przez klientów będących osobami fizycznymi. Zawężenie przedmiotu jest konieczne, ale korzystnym byłoby, gdyby Badacz uzasadnił wybór przedmiotu swoich zainteresowań naukowych.

Czytając tę część widzę z całą ostrością wyraźny dysonans pomiędzy rozważaniami teoretycznymi a praktycznymi zastosowaniami systemów rekomendacji w gospodarce. Twórcy rozwiązań biznesowych nie mają interesu w informowaniu, w jaki sposób działają systemy rekomendacji. Nie mają takiego interesu także podmioty gospodarcze, które produkują lub użytkują systemy rekomendacji, ponieważ nie są zainteresowane skróceniem czasu, w jakim ich rozwiązania dają im przewagę konkurencyjną na rynku. Nie mają takiego interesu osoby pracujące w takich podmiotach, dlatego że ich wiedza o sposobach działania systemów rekomendacji nie może być upowszechniana ze względu na sankcje, jakie im grożą na podstawie umów o poufności. Trudno znaleźć publikacje naukowe wykorzystujące dane z systemów rekomendacji stosowane przez duże podmioty gospodarcze. Dlatego decyzja Doktoranta o skupieniu się na systemach rekomendacji stosowanych przez małe sklepy internetowe była słuszną.

---

<sup>5</sup> Dysertacja str. 5

Oceniam pozytywnie prezentacje tradycyjnych a także czysto akademickich podejść stosowanych w systemach rekomendacji przedstawioną w dostępnej literaturze. Jednak rozwiązania oferowane komercyjnie są zapewne dość odległe od wiedzy wynikającej z dostępnej literatury. Działalność biznesowa w zakresie systemów rekomendacji jest bardzo konkurencyjna. Nakazuje to firmom chcącym utrzymać się na rynku gromadzić ogromne zasoby danych z wykorzystaniem technologii big data. Tak zgromadzone dane są obecnie przetwarzane z wykorzystaniem narzędzi sztucznej inteligencji.

**Część monograficzna pracy jest interesującym zaprezentowaniem na podstawie dociekań literaturowych problematyki badań będących przedmiotem rozważań Doktoranta.**

### Problem badawczy

W podrozdziale 3.4 opisano problem badawczy. Opis jest jednoznaczny. Doktorant, na tym etapie rozważań w dysertacji, przyjął zapewne świadomie dwa założenia upraszczające.

Po pierwsze, rekomendacje dla startu zimnego  $R_m^*$  oraz rekomendacje dla znanego użytkownika  $R_c^*$  są podzbiorami zbioru obiektów  $O$ . Zbiory nie przewidują porządku obiektów należących do zbioru. Użytkownicy chcą mieć rekomendację uporządkowaną według przyjętego kryterium (nazywanego przez Doktoranta funkcją użyteczności). Prezentowanie rekomendowanych obiektów jako elementów pewnego zbioru pozbawia możliwości zaprezentowania różnego stopnia znaczenia obiektów dla użytkownika. W konsekwencji system rekomendacji spełniający oczekiwania problemu badawczego z założenia będzie spełniał oczekiwania użytkowników w sposób połowiczny, pozbawiając użytkownika informacji o stopniu relewancji poszczególnych rekomendowanych obiektów dla użytkownika. Waga stopnia relewancji nie może być pomijana w systemach rekomendacji, co pokazano w części recenzji nazwanej *Przedmiot badań*.

Po drugie, daleko idącym uproszczeniem jest założenie, że system rekomendacji może działać w taki sposób, aby funkcja użyteczności  $u: C \times O \rightarrow W$ . Systemy rekomendacji są jeszcze dość odległe od rozpoznania funkcji użyteczności w rozumieniu ekonomicznym. *Użyteczność wyraża subiektywny stosunek konsumenta do oferowanych na rynku koszyków produktów i usług*<sup>6</sup>. Dlatego rozumiem, że, na tym etapie rozważań, intencją Doktoranta było znalezienie  $R_m^*$  na podstawie obiektu  $m$  wskazanego np. przez użytkownika, a  $R_c^*$  na zapytania użytkownika oraz danych behawioralnych zgromadzonych o nim przez system rekomendacji.

Widać, że dywagacje o koszcie pieniężnym i koszcie poznawczym wprowadziły Doktoranta w sferę ograniczonego komfortu kompetencyjnego w zakresie ekonomii<sup>7</sup>. Należy przyjąć to z pełnym zrozumieniem, ponieważ Doktorant ubiega się o stopień doktora w dziedzinie *nauk*

---

<sup>6</sup> Emil Panek, *Ekonomia Matematyczna*, Wydawnictwo UEP 2003, 904pp.

<sup>7</sup> Dysertacja str. 38.

inżynieryjno-technicznych w zakresie dyscypliny *informatyka techniczna i telekomunikacja*, a nie w zakresie ekonomii.

Opisując użytkownika nie zaliczałbym danych geograficznych do parametrów statycznych<sup>8</sup>, tak jak to zrobił Doktorant.

**Problem badawczy został sformułowany jednoznacznie i opisany poprawnie.**

### Rozwiązanie problemu badawczego

Doktorant pozycjonuje swoje rozwiązanie jako grafowy algorytm wykorzystujący analizę sesji. Takie zakwalifikowanie algorytmu jest poprawne.

Mimo że Doktorant odwołuje się do opracowanego przez siebie i opisanego w części 5.2 algorytmu rekomendacji wykorzystującego atrybuty obiektów, to niestety tej części brakuje w pracy<sup>9</sup>. Niestety także w części 3.2.4 nie znajdujemy szerokiego opisu wykorzystania grafów heterogenicznych dla budowy algorytmów rekomendacji.

W modelu danych określono, że obiekt „może stanowić” i dalej wymieniono listę odległych od siebie znaczeniowo klas obiektów, np. „artykuł prasowy w serwisie informacyjnym”, „osoba w serwisie społecznościowym”<sup>10</sup>. W podobny sposób szeroko opisano koncepcję jądra. Dzięki takiemu podejściu zdecydowanie uproszczono modelowanie grafu ograniczając się do grafów dwudzielnych. W przypadku przyporządkowania do wybranych grup obiektów i jąder osobnych podzbiorów zbioru węzłów, można byłoby wykorzystać modelowanie heterogonii węzłów w większym stopniu niż dwudzielność. W konsekwencji proponowany algorytm miałby większą ekspresywność semantyczną.

### Zagadnienie 1

Dobrze byłoby w czasie obrony poznać stosunek Doktoranta do wykorzystania przykładowo  $E_n \subseteq \{\text{lista życzeń klienta}\} \times \{\text{towar w sklepie internetowym}\}$  dla różnych  $n$  (par podzbiorów obiektów i jąder, czyli  $E_n \subseteq O_k \times J_l$ , gdzie  $O_k \subseteq O$  i  $J_l \subseteq J$ ).

Ze stwierdzeń „każde jądro musi być związane z co najmniej jednym obiektem” i „każdy obiekt musi być związany z co najmniej jednym jądrem” wnoszę, że

- Nie można wprowadzić do grafu żadnego  $o \in O$ , jeżeli nie można mu przyporządkować uprzednio istniejącego  $j \in J$ .

---

<sup>8</sup> Dysertacja str. 38.

<sup>9</sup> Dysertacja str. 53.

<sup>10</sup> Dysertacja str. 55.



- Nie można wprowadzić do grafu żadnego  $j \in J$ , jeżeli nie można mu przyporządkować uprzednio istniejącego  $o \in O$ .
- W konsekwencji zbiory  $J$  i  $O$  są puste.

Ograniczenia zapewniające dwudzielność grafu opisane na str. 57 są implikacją odpowiednich ograniczeń ze str. 56.

## Zagadnienie 2

Z punktu widzenia modelu grafu takie rozumowanie jest oczywiste. Jakie scenariusze wykorzystania modelu danych proponuje Doktorant dla poszczególnych  $E_n \subseteq O_k \times J_l$ . Na str. 57 i 58 wprowadza się co prawda pojęcie typów jąder, ale nie wprowadza się symetrycznej koncepcji typów obiektów, mimo że na str. 55 i 56 opisuje się przykłady takich typów. Dlaczego Doktorant zrezygnował z zamodelowania takiego symetrycznego podejścia do typizacji obiektów i jąder? Jak takie podejście wpłynęłoby na złożoność obliczeniową tak zaprojektowanego algorytmu?

Wysoko oceniam część nazwą *Badania algorytmu*. Dobrze, że Doktorant miał nieograniczony dostęp do rzeczywistych danych transakcyjnych oraz dokonał implementacji weryfikującej zaproponowane w dysertacji rozwiązania. Uwiarygadnia to proces walidacji algorytmu.

**Oceniam pozytywnie rozwiązanie problemu badawczego.**

## Aspekty redakcyjne dysertacji

W tej części wskazuję na kilka przykładów możliwości poprawienia redakcyjnego pracy.

Nie zgadam się z Doktoratem, że *Współczesne systemy rekomendacji bazują na algorytmach, które w swej idei badają podobieństwo pomiędzy użytkownikami (klientami) lub obiektami (przedmiotami)*<sup>11</sup>. Jeżeli Jan Kowalski szuka z pomocą systemu rekomendacji odpowiedniego samochodu, to nie znaczy, że właściwym samochodem dla Jana Kowalskiego, mającego prócz serca i wątroby także skrzynię biegów oraz akumulator, będzie pojazd wyposażony także w nerki i układ krwionośny. Proponuję Doktorantowi zawarcie kompromisu: *Współczesne systemy rekomendacji wykorzystują algorytmy, starające się znaleźć wysokie podobieństwo pomiędzy wyobrażonymi przez użytkownika atrybutami przedmiotów jego zainteresowania, a konkretnymi przedmiotami posiadającymi pożądane atrybuty, których danymi zarządza system rekomendacji.*

Niestety ten trudny do zaakceptowania skrót myślowy stosowany jest w pracy części<sup>12</sup>.

<sup>11</sup> Dysertacja str. 5

<sup>12</sup> Np. Dysertacja str. 6

Trudno zrozumieć, dlaczego Doktorant posługuje się pojęciem e-commerce w pracy napisanej po polsku. *Handel elektroniczny* jest od wielu lat używanym, niekontrowersyjnym odpowiednikiem.

Kontrowersyjne jest tworzenie liczby mnogiej rzeczowników w sposób, którego przykładem jest *e-Book'i*<sup>13</sup>.

Pewne fragmenty tekstu są dla mnie niezrozumiałe, np. *Klasyfikacja ta zapewnia relację pomiędzy różnymi grupami, np. produkty zakupione to produkty umieszczone w koszyku, a produkty umieszczone w koszyku to produkty, na które kliknięto (nie można umieścić w koszyku i następnie kupić produktu bez uprzedniego kliknięcia w niego)*<sup>14</sup>.

Szkoda, że w części nazwanej *Spis skrótów i oznaczeń* nie umieszczono oznaczeń stosowanych we wzorach.

**Praca jest napisana z odpowiednią troską o formę wypowiedzi.**

## Bibliografia

Mam kilka uwag do strony redakcyjnej pracy doktorskiej. Poszczególne referencje literaturowe przygotowano według różnych standardów cytowania. Niektóre są tak skrótowe, że nie odpowiadają zasadom referencji naukowych, np. referencje nr 69, 74, 75. Podobnie formatowanie referencji jest niekonsekwentne, np. kursywa stosowana jest do wskazania tytułu cytowanej pracy, dla innych kursywę zarezerwowano dla nazwy czasopisma naukowego, w którym opublikowano cytowaną pracę, w jeszcze innych nie zastosowano kursywy.

**Zaprezentowana bibliografia jest wystarczająca jakościowo oraz aktualna.**

## Sprawy, które powinny być przedyskutowane w czasie obrony

Prócz zagadnień wyróżnionych uprzednio w recenzji proszę o udzielenie w czasie obrony odpowiedzi na następujące pytania:

### Zagadnienie 3

Jak brzmi jednoznaczne sformułowanie tezy?

### Zagadnienie 4

Jak można krótko określić cele badawcze wykorzystane do weryfikacji tezy?

---

<sup>13</sup> Dysertacja str. 9

<sup>14</sup> Dysertacja str. 10



### Zagadnienie 5

Proszę podać wykorzystane bazy bibliograficzne oraz podejście do modyfikowania zapytań wykorzystywanych w badaniach literaturowych. Proszę uzasadnić przeprowadzone modyfikowania zapytań lub powody, dla których takich modyfikacji nie trzeba było przeprowadzać.

### Zagadnienie 6

W części recenzji *Przedmiot badań* naszkicowano przykładowe zastosowania systemów rekomendacji. Jakie zastosowania dla wyników swoich badań widzi Doktorant poza obszarem handlu elektronicznego?

## Wnioski

### Na podstawie konstatacji, że

- 1) Temat i zakres badań został wybrany przez Doktoranta poprawnie i ma głębokie uzasadnienie społeczne i gospodarcze.
- 2) Kompozycja dysertacji jest poprawna.
- 3) Część monograficzna pracy jest interesującym zaprezentowaniem na podstawie dociekań literaturowych problematyki badań będących przedmiotem rozważań Doktoranta.
- 4) Problem badawczy został sformułowany jednoznacznie i opisany poprawnie.
- 5) Oceniam pozytywnie rozwiązanie problemu badawczego.
- 6) Praca jest napisana z odpowiednią troską o formę wypowiedzi.
- 7) Zaprezentowana bibliografia jest wystarczająca jakościowo oraz aktualna.

### Stwierdzam, że:

- 1) Rozprawa doktorska mgr inż. Michała Malinowskiego *Algorytm rekomendacji bazujący na sesjach rekomendacji działających na podstawie zachowań użytkowników oraz atrybutów* stanowi rozwiązanie oryginalnego problemu naukowego.
- 2) Udowadnia opanowanie przez Doktoranta wiedzy teoretycznej w dziedzinie *nauk inżynieryjno-technicznych* w zakresie dyscypliny *informatyka techniczna i telekomunikacja*.
- 3) Wskazuje na umiejętność prowadzenia przez Doktoranta samodzielnej pracy naukowej wykorzystującej wiedzę teoretyczną dla rozwiązania konkretnych problemów badawczych.





Praca odpowiada wymaganiom stawianym pracom doktorskim. Stanowi dowód odpowiednich umiejętności naukowych Doktoranta. Rekomenduję pracę doktorską mgr inż. Michała Malinowskiego Radzie Dyscypliny Naukowej *Informatyka Techniczna i Telekomunikacja* Wojskowej Akademii Technicznej im. Jarosława Dąbrowskiego i wnoszę o dopuszczenie jej do publicznej obrony.

*Witold Abramowicz*