

RECENZJA

rozprawy doktorskiej mgr Michała Bałchanowskiego
„Ewolucyjna agregacja rang w systemach rekomendacyjnych”
Promotor: prof. dr hab. Urszula Boryczka

Tematyka rozprawy

Recenzowana dysertacja dotyczy systemów rekomendujących, które dzisiaj są nieodzownym narzędziem, za pomocą którego człowiek porusza się na—rosnącym w tempie dwucyfrowym w skali roku—ryнку danych. Znalezienie odpowiedniej informacji, zgodnej z szeroko rozumianymi preferencjami jest zadaniem systemów rekomendujących, zadaniem z jednej strony nieodzownym i koniecznym, biorąc pod uwagę ilość (już nie nawet liczbę) możliwych odpowiedzi, a z drugiej strony niezwykle trudnym, jeżeli uwzględnimy oczekiwaną dokładność rekomendacji. Literatura przedmiotu jest bogata, lista zaproponowanych rozwiązań długa, oczekiwania uczestników szeroko rozumianej cyfrowej gospodarki wobec systemów rekomendujących coraz większe, ale niestety trudno dzisiaj wskazać dominujące rozwiązania. Te istniejące i szeroko stosowane potrafią generować rekomendacje istotnie od siebie się różniące, nierzadko też zniekształcone danymi o niskiej jakości. Wśród wielu proponowanych dzisiaj metod, próba fuzji czy też agregacji rekomendacji generowanych jednocześnie przez kilka algorytmów zdaje się być zasadna.

Celem recenzowanej rozprawy było wykazanie tezy postawionej na stronie 5. dysertacji, iż zaproponowany przez Doktoranta algorytm ewolucyjnej agregacji rang poprawi jakość agregacji w porównaniu do wybranych metod. Autora rozprawy interesuje automatyczny dobór listy (czy też wektora) preferencji przedmiotów dla pojedynczego użytkownika, gdzie właściwa czy też suboptymalna agregacja poszukiwana jest przez algorytm metaheurystyczny (a dokładnie ewolucję różnicową) optymalizujący jedną z miar oceny jakości rekomendacji (a dokładnie średnią precyzję AP z parametrem określającym rozmiar listy).

O ile tematyka agregowania rang w systemach rekomendacyjnych jest stosunkowo nowa, to ewolucyjne uczenie się rangowania jest nowym i mało przebadanym podejściem, a zatem podjęte zadanie i proponowany sposób jego rozwiązania jest nietrywialne i wieloaspektowe.

W świetle powyższych uwag można stwierdzić, że **problematyka rozprawy jest ważna i w pełni uzasadniona.**

Kompozycja i redakcja rozprawy

Recenzowana rozprawa składa się z aż (!) dziesięciu rozdziałów, bibliografii, kilku spisów, wykazu ważniejszych oznaczeń i skrótów oraz jednego dodatku, i obejmuje 132 numerowanych stron maszynopisu. Mniej więcej połowa rozprawy, czyli sześć pierwszych rozdziałów, obejmuje wstęp teoretyczny, oryginalne treści zawierają rozdziały 7-9 wraz z podsumowaniem.

We Wstępie Autor zarysowuje obszar badawczy związany z systemami rekomendującymi, wyjaśnia ideę agregacji rang i wskazuje na oryginalność zastosowania metaheurystyk w podejmowanym problemie. W tym samym rozdziale zdefiniowana jest teza pracy, pięć celów cząstkowych i wreszcie opis zawartości poszczególnych rozdziałów dysertacji. Nie do końca przekonujące było umieszczenie stosunkowo szczegółowego diagramu prezentującego własne rozwiązanie (rys.

1.2), zwłaszcza, że po pierwsze, nie został ten diagram skomentowany, po drugie, Wstęp nie jest najlepszym miejscem na obciążanie czytelnika niepotrzebnymi jeszcze szczegółami, i po trzecie, na str. 57 można znaleźć jego inną wersję (rys. 7.1).

Rozdział drugi zawiera wprowadzenie w systemy rekomendacyjne wraz z ich klasyfikacją. W rozdziale można znaleźć listę głównych problemów, przed którymi stoją badacze zajmujący się tematyką (ale problemów zdefiniowanych na podstawie nie najnowszej pozycji bibliograficznej z 2015 r.). Rozdział kończy bardzo autorski (to jest subiektywny) przegląd wybranych algorytmów.

W kolejnym rozdziale zawarty jest opis najbardziej popularnych miar oceniających tworzone listy preferencyjne. Ewolucji różnicowej, czyli wykorzystanej przez Doktoranta we własnym rozwiązaniu metaheurystyce, poświęcony jest rozdział czwarty. Wcześniej jednak znajdujemy w rozdziale informacje o algorytmach ewolucyjnych stosowanych w systemach rekomendacyjnych.

Problemowi agregacji rang poświęcony jest rozdział piąty. W rozdziale tym w sposób niejawni (mimo deklaracji ze str. 76!) znaleźć można opis pięciu metod (a w istocie dwóch, bo cztery z owych metod są odmianami tego samego algorytmu agregacji *Comb*), z którymi będzie porównywane rozwiązanie Doktoranta. Rozdział kończy istotna informacja o stosowanej w badaniach eksperymentalnych mierze odległości Kendalla K_d .

Ostatni z rozdziałów wprowadzających w obszar rozprawy, rozdział szósty, podejmuje temat uczenia się rangowania. Poza formalną definicją możemy w rozdziale znaleźć klasyfikację algorytmów, w tym algorytmów rozpatrujących pełne listy rankingowe (nota bene szkoda, że nie podjęto próby porównania uzyskanych wyników z jednym z tych algorytmów).

Kolejne trzy rozdziały, nie licząc podsumowania, zawierają już oryginalne treści. I tak, w rozdziale 7 zawarty jest opis własnego rozwiązania. Opis został podany w formie słownej, diagramów i pseudokodów. Rozwiązanie jest dwuwariantowe, wariant drugi do rozpatrywanego zbioru pojedynczych list rekomendacji generowanych przez wybrane algorytmy, dodaje podzbiór najbliższych sąsiadów użytkownika. O ile opis rozwiązania jest w miarę czytelny, to pewną trudność sprawiła piszącemu te słowa interpretacja zawartości poszukiwanego przez algorytm heurystyczny wektora preferencji w_{u_A} . Rysunek 7.2 nie jest w tym pomocny i trzeba sięgnąć poza dysertację, do autorskiej publikacji [36].

Rozdział 8 zawiera opis przetwarzanego zbioru danych, zastosowanych bibliotek programistycznych oraz dostrojonych wartości parametrów pojedynczych algorytmów rekomendacyjnych, tworzących później agregacje. W ramach tzw. „metodyki” przeprowadzonych eksperymentów otrzymujemy informację o wielokrotnym uruchomieniu algorytmu ewolucji różnicowej, wartościach jego parametrów (niestety już niestrojonych), podziale zbioru danych na podzbiory treningowy, walidacyjny i uczący, protokołu ewaluacyjnego oraz zastosowanym do statystycznej weryfikacji danych teście Wilcozona (ale tylko do części danych).

Obszerne badania eksperymentalne zostały przedstawione w rozdziale dziewiątym. Badania podzielono dosyć nieoczekiwanie na trzy etapy, tj. badania „podstawowe”, których celem było zaprezentowanie wyników dla podstawowych algorytmów rekomendacyjnych i metod agregujących tylko dla 25 użytkowników wyciągniętych z bazy MovieLens100k; badania „eksperymentalne” (czyżby badania „podstawowe” nie były także badaniami „eksperymentalnymi”), w których badano różne warianty funkcji oceny, analizowano wpływ poszczególnych algorytmów cząstkowych na agregację w zależności od funkcji oceny, uwzględniano rekomendacje niskiej jakości (czy metody *Random_1*, *_2* i *_3* różnią się między sobą?) oraz badano wariant algorytmu rozszerzonego o rankingi najbliższych sąsiadów, wszystko tym razem dla 300 użytkowników; wreszcie badania „końcowe”, przeprowadzone dla wszystkich użytkowników w bazie (w liczbie 943) i zastosowanym (wreszcie!) testem statystycznym.

Rozprawę zamyka Podsumowanie, zawierające uogólnione wnioski (trudno powstrzymać się od uwagi do wniosku pierwszego, który w świetle treści samego Wstępu wydaje się trywialny

i niewarty empirycznej weryfikacji), następnie konstatacji Autora, że postawiona teza została udowodniona, a cele szczegółowe osiągnięte. Rozdział zamyka wskazanie kierunków dalszych badań.

Kompozycja i struktura rozprawy nie budzą większych zastrzeżeń, może poza niepotrzebnym rozdrobnieniem rozdziałów części teoretycznej.

Na uwagę zasługuje staranna strona edytorska pracy i bogaty, swobodny oraz lekki język rozprawy. Napotkane w rozprawie błędy interpunkcyjne (w liczbie kilkudziesięciu) nie utrudniają lektury i zostały zaznaczone w otrzymanym elektronicznym egzemplarzu. Niezrozumiałe, a na pewno niepotrzebne, było podawanie angielskich odpowiedników polskich pojęć w nawiasach zaczynających się od przyimka „z” (np. „... algorytmie ewolucji różnicowej (z ang. differential evolution, DE)”). Jak już mowa o języku, to pewnym *curiosum* (na szczęście jednorazowym) było użycie neologizmu „odstańce” w miejsce „wartości odstających” (str. 82).

Bibliografia zawiera 193 pozycje, w tym 5 publikacji, których współautorem jest Doktorant (prace wspólne z promotorem). Dwie z tych publikacji umieszczono w czasopismach z tzw. listy filadelfijskiej (*Entropy* i *Electronics*, każda po 100 pkt. ministerialnych, odpowiednio w 2023 i 2022 roku), pozostałe są publikacjami pokonferencyjnymi. Dorobek publikacyjny Doktoranta jest wystarczający. Bibliografia odwołuje się do aktualnego piśmiennictwa.

Oryginalne osiągnięcia Autora

Do najważniejszych i oryginalnych wyników przedstawionych w recenzowanej rozprawie należy zaliczyć:

- Opracowanie i zweryfikowanie algorytmu ewolucyjnej agregacji rang w kilku wariantach:
 - dla trzech wersji funkcji oceny,
 - z losowymi rekomendacjami,
 - oraz z sąsiedztwem i bez sąsiedztwa.
- Porównanie jakości działania innych metod agregujących.
- Zbadanie wpływu rozmiaru sąsiedztwa na końcową jakość agregacji.
- Zbadanie wpływu rekomendacji o niskiej jakości na końcową jakość agregacji.
- Pogłębioną analizę zawartości wektora preferencji optymalizowanego przez algorytm różnicowy, również w funkcji czasu.

Zgodnie z powyższym można sformułować wniosek, że uzyskane wyniki eksperymentalne potwierdzają postawioną tezę.

Uwagi krytyczne i dyskusyjne

Podczas lektury recenzowanej dysertacji nasuwają się pewne uwagi natury ogólnej i szczegółowej. Do uwag natury ogólnej można zaliczyć poniższe kwestie:

- Badania eksperymentalne zostały przeprowadzone tylko na jednym i to niewielkim zbiorze danych. Co prawda Doktorant zauważa ten brak na str. 103 rozprawy, ale nie zmienia to faktu, że wyciągnięte wnioski o jakości działania proponowanego rozwiązania są w związku z tym (mocno) ograniczone.

- Brakuje analizy złożoności rozwiązania. Jest wzmianka o skutecznej próbie przyspieszenia działania algorytmu (patrz publikacja [36]), ale w samej rozprawie nie ma dyskusji na ten temat, poza stwierdzeniami o czasochłonności obliczeń.
- Algorytmy (meta)heurystyczne mają naturę stochastyczną i jako takie wymagają wielokrotnych powtórzeń i obliczania wartości oczekiwanych. Tylko w jednej tabeli (9.4) podano wartości średnie wraz z odchyleniami, w rozprawie są tylko 4 wykresy pudełkowe i tylko dla badań „końcowych” do weryfikacji wyników zastosowano testy statystyczne.
- Dlaczego nie strojono parametrów algorytmu ewolucji różnicowej. Tym bardziej pytanie jest aktualne, że zadano sobie trud strojenia algorytmów cząstkowych (dla pełności obrazu dodać należy, że Doktorant zauważa ten brak na str. 104 rozprawy).
- W analizie literatury brakuje ewidentnie odwołań do prężnie rozwijającego się algorytmów na styku systemów rekomendacyjnych i uczenia głębokiego (patrz chociażby praca przeglądowa „Deep Learning based Recommender System: A Survey and New Perspectives” z 2019 roku.). Zauważyć należy jednak, że Doktorant cytuje trzy prace z tego obszaru, ale nie podejmuje się ich głębszej analizy.
- Czy nie byłoby zasadne porównanie własnego algorytmu z innym algorytmem rangującym, również opartym na listach (patrz str. 51)?
- W opisie metod hybrydowych (rozdział 2.5.5 i tabela 2.1) brakuje odwołania się do przykładowych implementacji literaturowych.
- Wniosek czwarty ze str. 101 o zmniejszaniu się jakości rekomendacji przy losowych sąsiadach zdaje się być trywialny.

Poniżej wyszczególniono uwagi natury szczegółowej:

- Na str. 9 definiuje się nieformalnie pojęcie Big Data, które nie tylko—jak chce Doktorant—odnosi się do dużych zbiorów danych, ale również do całych ekosystemów, złożonych z danych masowych, technik, narzędzi i metodologii.
- Na str. 20 został użyty niestety powszechnie dzisiaj stosowany i rażący anglicyzm „zadresować problem”.
- Na str. 29 nieprawidłowo użyto czasownik „karać”.
- Na str. 35 Doktorant pisze o programowaniu genetycznym, jako metodzie wykorzystywanej do „automatycznego generowania programów komputerowych”. Nie tylko!
- Na str. 36 w tłumaczeniu heurystyki ACO użyto sformułowania „algorytm mrowiskowy”, w miejsce częściej jednak stosowanego „algorytmu mrówkowego” (ale uwaga do dyskusji).
- Na str. 38 nie zdefiniowano znaczenia symbolu „g” we wzorze (4.2).
- Na str. 45 jest „łańcuchach markowa” w miejsce „Markowa”.
- Na rys. 6.1 (str. 50) nie wyjaśniono symbolu „y”.
- Dwa pierwsze akapity ze str. 53 niezgrabnie stylistycznie (powtórzenia słów „zapropo-nowany” i „zaprezentowany”); podobnie zdanie z trzeciego akapitu ze str. 54.
- Na str. 55 następuje odwołanie do wzoru 7.4 podanego dopiero na str. 58.

- Nieprawdziwe jest zdanie ze str. 55 „a cały algorytm będzie dążyć do znalezienia optymalnego rozwiązania”; co najwyżej „suboptymalnego”.
- Algorytmy referencyjne z tabeli 8.7 (str. 69) powinny zostać krótko opisane.
- Na str. 79 w pierwszym akapicie pisze Doktorant, że po „uśrednieniu algorytmy generują rekomendacje o zbliżonej jakości”, gdy tymczasem algorytm MostPopular mocno odstaje od pozostałych.
- Trudno oczekiwać optymalnej oceny od algorytmu heurystycznego (patrz stwierdzenie o wartości, która „nie zbliżyła się do jedności” ze str. 88).
- Pierwsze zdanie drugiego akapitu ze str. 93 wymaga korekty stylistycznej.
- Na str. 100 w drugim akapicie jest „jako problemu”, powinno być „jako problem”.

Wszystkie wymienione wyżej krytyczne uwagi i spostrzeżenia nie podważają pozytywnej oceny pracy.

Wniosek końcowy

Recenzowana dysertacja niewątpliwie stanowi oryginalne rozwiązanie postawionego problemu naukowego, istotnie rozszerzając wiedzę o systemach rekomendacyjnych. Doktorant wykazał się właściwą wiedzą teoretyczną i praktyczną w dziedzinie nauk ścisłych i przyrodniczych, w dyscyplinie naukowej informatyka, dowiódł też opanowania umiejętności samodzielnego prowadzenia pracy naukowej.

Biorąc pod uwagę powyższe fakty, łącznie z zawartymi w recenzji uwagami stwierdzam, że rozprawa spełnia wymagania Ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2022 r. poz. 574, z późn. zm.) i **wniosuję o dopuszczenie mgra Michała Bałchanowskiego do dalszych etapów postępowania o nadanie stopnia doktora.**

Olgierd Uner