

Warszawa, 27 sierpnia 2018

dr hab. inż. Szymon Jaroszewicz
Instytut Podstaw Informatyki PAN
ul. Jana Kazimierza 5
Warszawa

Recenzja rozprawy doktorskiej mgr inż. Kazimierza Zielińskiego

Promotor: dr hab. Adam Wierzbicki, prof. PJATK
Promotor pomocniczy: dr inż. Radosław Nielek

Uwagi ogólne

Rozprawa doktorska Pana mgr inż. Kazimierza Zielińskiego ma formę zwartej publikacji. Dotyczy ona ważnego problemu jakim jest kontrowersyjność treści w Internecie i jest wpływ na jego użytkowników.

Pan Zieliński logicznie uporządkował cele badawcze, oraz strukturę pracy im podporządkowaną. Rozpoczął od formalnego zdefiniowania pojęcia kontrowersji, dostosowanego do zastosowań w informatyce i formalnego określenia metod jej pomiaru. Następnie autor podjął temat automatycznego wykrywania kontrowersji w stronach WWW i opracował odpowiednią wtyczkę do przeglądarki.

Autor skupił się na dwóch najważniejszych metodach, przy pomocy których użytkownicy poszukują treści w internecie: Wikipedii i wyszukiwarkach internetowych. Fakt, że te dwie metody pokrywają znaczny procent wszystkich wejść do internetu metody pozytywnie wpływa na praktyczną istotność uzyskanych wyników.

Przejdę teraz do oceny wkładu pracy w dziedzinę informatyki a następnie przedstawię merytoryczną oceną uzyskanych wyników.

Ocena wkładu w dziedzinę informatyki

Należy podkreślić interdyscyplinarny charakter pracy. Praca dotyczy dziedziny informatyki społecznej i występuje w niej znaczący element socjologiczny. Nie mam jednak wątpliwości, że można ją uznać za pracę przede wszystkim z dziedziny informatyki. Po pierwsze, praca jest ściśle związana z technologiami internetowymi np. zapytaniami do wyszukiwarek internetowych. Praca stosuje też metody typowe dla informatyki takie jak klasyfikacja przy pomocy lasów losowych. W wyniku pracy powstał spójny system *informatyczny* do oceny kontrowersji treści w Internecie. Należy też zwrócić uwagę na dobrą znajomość przez doktoranta technik informatycznych i sieciowych co pozwoliło mu na implementację i powszechne udostępnienie opracowanego systemu w formie wtyczki do przeglądarki komunikującej się z serwerem kontrowersji.

Ocena merytoryczna uzyskanych wyników

Przejdę teraz do omówienia kolejnych rozdziałów pracy i oceny ich wkładu w dziedzinę informatyki społecznej.

Autor rozpoczyna w rozdziale 2 od przeglądu literatury. Na początku przedstawione zostają najważniejsze koncepcje dotyczące kontrowersji z dziedzin psychologii, socjologii i politologii. Jako, że praca dotyczy dziedziny informatyki społecznej, przegląd taki jest bardzo potrzebny. Wydaje się, że autor uwzględnił najważniejsze idee dotyczące psychologicznego i socjologicznego kontekstu analizowanego pojęcia. Szczególną uwagę doktorant poświęcił pojęciu tzw. „bańki filtrującej”, czyli zjawisku gdy personalizacja wyników wyszukiwania powoduje, że użytkownik zostaje odcięty od poglądów innych niż jego własne. Zapobieganiu temu zjawisku autor poświęca w swojej pracy wiele uwagi.

Rozdział dotyczący kontrowersji w mediach społecznościowych jest nieco krótki, zwłaszcza w porównaniu z rozdziałami poprzednimi, nie jest on jednak kluczowy z punktu widzenia pracy. Istotniejsze rozdziały dotyczące kontrowersji w Wikipedii i wyszukiwarkach internetowych są obszerne i uwzględniają najważniejsze prace w tych dziedzinach. Przegląd literatury należy zatem uznać za w pełni zadowalający.

Ostatnia część rozdziału 2 prezentuje podstawowe pojęcia matematyczne użyteczne w analizie kontrowersji takie jak skala Likerta, Kappa Fleissa czy TF-IDF.

W rozdziale 3 autor przedstawia formalną definicję kontrowersji, podanie której było jednym z celów badawczych pracy. Definicja podana we wzorach 3.1 i 3.2 określa kontrowersję jako funkcję f rozkładu ocen dokonanych w

danym czasie przez daną społeczność. Ponieważ autor nie podaje żadnych warunków jakie ta funkcja powinna spełniać trudno uznać tę definicję za ścisłą. Jest to raczej ogólne stwierdzenie, że kontrowersja będzie modelowana na podstawie rozkładu ocen. Słabość definicji widać już w tym rozdziale, gdzie autor często posiłkuje się np. intuicją, że rozkład kontrowersyjny ma kształt litery U nie definiując tego formalnie. Problem nie rzutuje na szczęście na pozostałą część pracy, gdzie autor używa metod klasyfikacji do określenia funkcji f .

W dalszej części rozdziału autor analizuje wpływ czasu na możliwość wykrycia nieistniejących kontrowersji. Kwestia ta jest bardzo istotna (co ilustrują przykłady) a zarazem często pomijana przez badaczy. Cieszy zatem fakt, że doktorant nadał jej duże znaczenie.

Niestety autor nie uniknął tu błędów formalnych. Po pierwsze we wzorze 3.3 autor odejmuje od siebie obiekty, które nie są zwykle liczbami rzeczywistymi. Należałoby tu raczej stosować jakąś miarę odległości/podobieństwa między obiektami. Po drugie, we wzorze 3.6 autor podaje inną definicję opartą na mierze kontrowersji f . Jednak podobieństwo obiektów nie gwarantuje stałości f , która zależy również od społeczności. Zauważmy również, że obiekt może się zmienić w bardzo znacznym stopniu, bez wpływu na f . Lepiej byłoby porzucić na definicji ze wzoru 3.6.

W kolejnym rozdziale autor opisuje zbiory danych użyte w badaniach. Ich dobór wydaje się odpowiedni, chociaż, z przyczyn obiektywnych jeden z nich (zbiór dotyczący kontrowersyjnych artykułów z Wikipedii) jest niewielki. Autorowi udaje się jednak ominąć ten problem w rozdziale 6.

W rozdziale 5 autor rozpoczyna część pracy zawierającą jego wkład w wykrywanie kontrowersji. Przedstawiono szereg modeli klasyfikacyjnych opartych o lasy losowe służących do przewidywania czy dany artykuł Wikipedii jest kontrowersyjny. Zaproponowano trzy modele: model oparty o definicję kontrowersji z rozdziału 3 (wzbogacony jednak o kilka dodatkowych zmiennych), model oparty o polaryzację emocji oraz klasyfikator zaproponowany wcześniej przez Kittura, oparty o historię edycji artykułu. Dodatkowo autor przedstawił model łączony, budowany wspólnie na zmiennych wszystkich trzech modeli. Ogólnie model opierający się o rozkład opinii (definicja autora) działał prawie tak samo dobrze jak model Kittura, a model łączony go przewyższył. Należy zatem stwierdzić, że doktorantowi udało się zrealizować cel badawczy i uzyskać model wykrywający kontrowersję w treściach Wikipedii oparty o oceny użytkowników, a także poprawę dokładności względem dostępnych do tej pory metod.

W kolejnym rozdziale autor rozszerza metodę na kategorie Wikipedii, skutecznie identyfikując kategorie zawierające wiele kontrowersyjnych artykułów. Wyniki przedstawione w rozdziale stanowią dodatkowy, przekonujący

argument na rzecz skuteczności proponowanej metody.

W rozdziale 7 doktorant podjął próbę przewidywania kontrowersji na podstawie tekstowej zawartości artykułów. Okazało się, że podobieństwo treści mierzone miarą kosinusową nie jest dobrym predyktorem kontrowersji. Mimo iż jest to wynik negatywny, przedstawione analizy są z pewnością interesujące.

Rozdział 8 przenosi opracowaną wcześniej metodologię na problem wykrywania kontrowersji w zapytaniach do wyszukiwarek internetowych. Zapytanie jest najpierw dopasowywane do kategorii artykułów Wikipedii przy pomocy słów kluczowych i miary TF-IDF. Autor proponuje odpowiednią miarę agregującą kontrowersyjność kategorii, do których zostało przypisane zapytanie. Dodatkowo doktorant opracował wtyczkę do przeglądarki Chrome, która automatycznie sygnalizuje, czy wpisywane przez użytkownika zapytania są kontrowersyjne.

Rozdział 8 uważam za największe osiągnięcie doktoranta a zarazem za logiczną kontynuację rozdziałów 5 i 6 rozszerzającą aplikowalność przedstawionych tam metod.

Podsumowanie

Rozprawa przedstawia szereg ciekawych rozwiązań wychodzących poza state-of-the-art w wykrywaniu kontrowersji w treściach internetowych i stanowi istotny wkład w dziedzinę informatyki społecznej. Zwłaszcza rozdziały 5–8 tworzą spójną całość pokazującą stopniowy rozwój metody proponowanej przez autora i rozszerzanie zakresu jej stosowalności.

Należy też podkreślić bardzo staranne przygotowanie pracy pod względem językowym i redakcyjnym. Autor nie ustrzegł się pewnych błędów formalnych i matematycznych jednak nie rzutują one na całość pracy.

Według bazy DBLP doktorant jest współautorem 3 związanych z tematem doktoratu publikacji, w tym jednej w czasopiśmie z listy A MNiSW (Information Processing and Management, 30 pkt), jednej w czasopiśmie z listy B i jednej na międzynarodowej konferencji. Dorobek publikacyjny należy zatem uznać za w pełni wystarczający.

Podsumowując, stwierdzam że praca spełnia z wymagania ustawowe stawiane dysertacjom doktorskim w dziedzinie informatyki i stawiam wniosek o dopuszczenie doktoranta do publicznej obrony.


Szymon Jaroszewicz