



Kraków 13.08.2017

dr hab. inż. Aleksander Byrski, prof. AGH  
Grupa Inteligentnych Systemów Informacyjnych  
Katedra Informatyki, Wydział Informatyki, Elektroniki i Telekomunikacji  
Akademia Górniczo-Hutnicza im. St. Staszica w Krakowie  
Al. Mickiewicza 30, 30-059 Kraków  
e-mail: olekb@agh.edu.pl  
Tel: 012 328 33 00  
Fax: 012 617 51 72

Recenzja rozprawy doktorskiej pt.:

**Model formalny, metody i algorytmy wykrywania kontrowersji  
artykułów Wikipedii oraz zapytań do wyszukiwarek**

autorstwa:

**mgr inż. Kazimierz Zieliński**

Zakres badań prowadzony w ramach dziedziny „Informatyka” w od początku bieżącego wieku, musiał zostać znacznie rozszerzony w związku z błyskawicznym rozwojem wszelkiego rodzaju mediów komunikacyjnych, a w szczególności tzw. mediów społecznościowych, z których najpopularniejsze to oczywiście Facebook czy Twitter. Swoistym fenomenem, starszym od wspomnianych serwisów, jest encyklopedia edytowana przez szerokie społeczeństwo, czyli Wikipedia. Tzw. Informatyka Społeczna zajmuje się badaniem wszelkiego typu zależności między użytkownikami tworzącymi wspomniane media, pozwala na opracowanie nowych strategii marketingowych, badanie wydźwięku opinii nt. produktów i in. Tego typu badania były nie do wyobrażenia jeszcze 20 lat temu, teraz raczej wydają się już koniecznością.

Przedstawiona do recenzji rozprawa doktorska wpisuje się w dziedzinę Informatyki Społecznej i koncentruje się na wykrywaniu kontrowersyjności opinii przedstawianych przez użytkowników sieci Internet. Niewątpliwie anonimowość użytkowników w sieci umożliwia publikację praktycznie dowolnych treści, a same kontrowersje nie wynikają bynajmniej zawsze z przekonań twórcy, a także z faktu że wiele z treści wpisuje się w zakres tzw. trollingu, który z założenia prowadzony jest aby zdyskredytować pewną osobę czy fakt. Aby efektywnie orientować się w treściach publikowanych przez tak wiele osób z tak różnymi motywacjami, niewątpliwie warto byłoby mieć możliwość oceny kontrowersji, by nie przyjmować bezkrytycznie i natychmiastowo przedstawianych punktów widzenia, zamiast tego świadomie podejmować próbę wypracowania sobie własnego zdania na dany temat.

Wykrywając i oceniając treści kontrowersyjne, Doktorant jako źródło danych i główny punkt odniesienia obrał serwis Wikipedia a za cel postawił sobie opracowanie modelu formalnego porządkującego tematykę wykrywania kontrowersji, opracowanie algorytmu predykcji kontrowersji artykułów w serwisie Wikipedia a także opracowanie algorytmu oceny kontrowersji dla zapytań do wyszukiwarek (wraz z dedykowaną implementacją w postaci serwera gromadzącego niezbędną wiedzę wraz z algorytmami klasyfikacji oraz klienta/wtyczki do przeglądarki internetowej oraz dedykowanego serwera obsługującego zapytania przesyłane przez wtyczkę).

Manuskrypt pracy liczy 117 stron. Po zarysowaniu celu badań we wprowadzeniu następuje 8 rozdziałów (przegląd literatury, opis modelu, identyfikacja przetwarzanych danych, wyniki badań i podsumowanie), następnie zamieszczono podsumowanie oraz bibliografię liczącą 117 pozycji.

Rozdział 1: Zawiera wprowadzenie do tematyki poruszanej w pracy oraz przedstawia cele założone przez Doktoranta a także podsumowanie wyników naukowych osiągniętych w pracy.

Rozdział 2: Zawiera przegląd stanu wiedzy wychodząc od definicji kontrowersji na gruncie psychologii i socjologii oraz odnosząc się do innych elementów współczesnej kultury medialnej, mediów społecznościowych, wiarygodności, przedstawiając jednocześnie stan badań innych naukowców w dziedzinie wykrywania kontrowersji. Rozdział kończy się przeglądem metod analitycznych, które mogą być wykorzystane (i zostały wykorzystane przez Doktoranta) do wykrywania kontrowersji w badanych tekstach.

Rozdział 3: Zawiera model formalny proponowany przez Doktoranta oraz szereg definicji np. związanych z obserwacją kontrowersyjnego obiektu przez pewną społeczność w wybranym oknie czasowym.

Rozdział 4: Zawiera opis danych jako punkt wyjścia dla badań eksperymentalnych prowadzonych przez Doktoranta.

Rozdział 5: Koncentruje się na przedstawieniu algorytmów wykrywania kontrowersji w artykułach w szczególności opisuje klasyfikator „random forest” oraz podstawowe informacje nt. oceny klasyfikatorów binarnych. Następnie Doktorant opisuje konstrukcję własnego klasyfikatora kontrowersji i prezentuje wyniki oceny jego wiarygodności w porównaniu do klasyfikatora opracowanego przez Kittura, bazując na obserwacji krzywej ROC-AUC wykazuje że zaproponowany model hybrydowy przewyższa pod względem efektywności algorytm referencyjny.

Rozdział 6: Poświęcony jest w całości tematyce kontrowersyjnych artykułów na Wikipedii. Doktorant testuje opracowany klasyfikator korzystając z porównania wyników z referencyjnym zbiorem opracowanym przez Doktoranta na bazie ręcznie wybranej listy kategorii kontrowersyjnych a także przeprowadzono walidację krzyżową.

Rozdział 7: Traktuje o predykcji kontrowersji artykułów Wikipedii bazując na porównaniu podobieństwa na poziomie tekstu z artykułami o znanym stopniu kontrowersji. Doktorant starannie opisuje swoje przemyślenia i podjęte kroki, w szczególności dyskutując dosyć niskie rezultaty (niską korelację podobieństwa tekstowego) otrzymane w trakcie badań.

Rozdział 8: Poświęcony jest predykcji kontrowersji zapytań do przeglądarek. Doktorant

opisuje opracowaną metodę porównania zapytania zadawanego przez internautę przeglądarce z kategoriami Wikipedii i na bazie wcześniejszych obserwacji odnośnie kontrowersyjności ostrzega użytkownika, że zadanie potencjalnie kontrowersyjne pytanie. Dedykowana wtyczka zaimplementowana przez Doktoranta pozwala na instalację rozszerzenia popularnej przeglądarki Chrome i komunikuje się z „serwerem kontrowersji” umożliwiającym klasyfikację zadanego zapytania. Doktorant przedstawia wyniki walidacji działania zaproponowanego mechanizmu na podstawie ręcznie opracowanych danych.

Rozdział 9: Zawiera wnioski i podsumowanie a także szeroki opis potencjalnych zastosowań proponowanych rozwiązań wraz z planem przyszłych badań.

Doktorant przygotowując pracę nie uniknął pewnych błędów i niedociągnięć, które chciałbym teraz przedstawić. Poza tym czytając pracę nasunęło mi się kilka uwag ogólnych, o odniesienie się do których uprzejmie proszę Doktoranta.

#### **Uwagi o charakterze ogólnym i dyskusyjnym:**

- Głównym mankamentem recenzowanej pracy jest przedstawiony w 3. Rozdziale model formalny kontrowersji. Sam podrozdział 3.1. o tytule Model formalny kontrowersji zawiera dwa równania – 3.1 oraz 3.2, które tak naprawdę opisują funkcję klasyfikatora binarnego. Model formalny powinien dostarczać pewnego usystematyzowanego opisu (np. jakiegoś zjawiska czy algorytmu), precyzyjnie go definiować a przede wszystkim pozwalać na wyciągnięcie jakichś konstruktywnych wniosków, przeprowadzenie jakichś rozważań formalnych, z których będą wynikać obserwacje (będące wynikiem rozważań formalnych, wynikiem np. pewnego dowodzenia) ważne dla wspomnianego zjawiska i algorytmu. Takie obserwacje można potem np. weryfikować eksperymentalnie (por. np. model algorytmu genetycznego opracowany przez Michaela Vose’a). Przedstawiony model stanowi raczej nieskomplikowaną, choć precyzyjną definicję, ale trudno mówić o wnioskowaniu nt. własności samego zjawiska na niej bazując.
- W podrozdziale 3.2 Doktorant przeprowadza pewnego rodzaju wywód formalny, wykorzystując wspomnianą wyżej definicję klasyfikatora, który sprowadza się do definicji podobieństwa obiektów dla których obliczana jest kontrowersyjność i relacji kontrowersyjności do opinii pewnej społeczności. Należy podkreślić, że tego typu badania powinny stanowić jedną z newralgicznych części rozprawy, skoro „Model formalny” został wspomniany bezpośrednio w jej tytule. Niestety w samych rozważaniach (głównie w zapisie) można znaleźć wyraźne niedociągnięcia w sposobie oznaczeń co utrudnia postrzeganie całości. Przykładowo zmiana rozkładu  $\Delta r \text{ ob}(t)$  jest zdefiniowana przy pomocy dość zaskakujących sformułowań,  $\text{ob}=\text{ob}(t)$  – z lewej strony równania nie ma zależności od czasu a z prawej już jest? Z lewej jest zmienna a z prawej funkcja o tej samej nazwie co zmienna? Potem jeszcze Doktorant porównuje  $\Delta r \text{ ob}(t)$  z pewną wartością „threshold”, przy czym znowuż  $\text{ob}(t)$  sugeruje funkcję, a symbol  $\Delta r$  zwykle jest stosowany przy zmianie pewnej wartości, co do funkcji nie przystaje. Domyślam się, że chodziło tutaj o uzależnienie wartości  $\text{ob}$  od czasu i istnieje pewna wartość w momencie  $t$  oraz zupełnie inna w momencie  $t_0$  – niestety znowuż nie jest to zdefiniowane w sposób precyzyjny i poprawny co utrudnia postrzeganie pracy. We wzorze 3.10  $\text{int}0'$  jest zdefiniowane na dwa

sposoby (podane są dwa przedziały), poza tym występuje tu symbol ' (prim) który oznacza zmieniającą się wartość w czasie. Doktorant powinien wybrać jeden sposób określania następstwa czasowego – bo wcześniej stosuje  $ob(t)$  i  $ob(t_0)$ . Podsumowując, sam opis formalny przedstawiony przez Doktoranta po rozszerzeniu oraz gruntownej korekcie znacznie ułatwiłyby rozumienie całej pracy – **dlatego też uprzejmie proszę o gruntowną poprawę strony formalnej pracy przed prezentacją jej w trakcie obrony.**

- Proponowany przez Doktoranta algorytm predykcji kontrowersji bazujący na podobieństwie tekstowym nie dał zadowalających efektów (niska korelacja między podobieństwem tekstowym a podobieństwem EMD obliczonym dla atrybutów opinii społeczności). Pomysł z wykorzystaniem podobieństwa tekstowego był interesujący i warty podjęcia próby, ale wydaje mi się, że bez wejścia w semantykę tekstu nie będzie możliwa automatyczna ocena tego typu charakterystyk jak kontrowersja – czy choćby wydzwięk (którego badanie opiera się na prostych regułach statystycznych – niemniej wymagane jest istnienie dedykowanego słownika). Proszę przedyskutować możliwość wykorzystania semantyki do porównania treści tekstu w przypadku problemu wykrywania kontrowersji.

#### **Pomniejsze usterki:**

- Przeglądając pracę a w szczególności spoglądając na spis treści dostrzec można wyraźne niezbalansowanie zawartości pracy – poszczególne części pracy są mocno zróżnicowane jeśli chodzi o liczbę stron, np. Wprowadzenie ma 5 stron, przegląd stanu wiedzy kilkanaście, opis zbiorów badawczych 3 strony, rozdział 7 w ogóle nie ma podrozdziałów, niektóre podrozdziały są krótsze od jednej strony (np. 8.1). Struktura pracy powinna być bardziej zbalansowana.
- Rozdział z przeglądem literatury powinien mieć nazwę odpowiadającą bardziej jego treści – zamiast bardzo ogólnego „Przeglądu literatury” lepiej brzmiałby tytuł taki jak „Kontrowersja jako element kultury medialnej” lub podobny. W ogólności nie jestem zwolennikiem nadawania zbyt ogólnych nazw rozdziałom.
- Wszystkie wzory przytoczone przez Doktoranta powinny zostać precyzyjnie zdefiniowane i opisane, w szczególności dotyczy to podawania przestrzeni do których należą poszczególne symbole, nawet jeśli byłyby to zbiór liczb naturalnych czy rzeczywistych, nie podanie tego typu informacji prowadzi do konieczności domyślenia się czytelnika, co oczywiście może prowadzić do przeoczeń lub nadinterpretacji.
- Rozdział 4 prezentujący zbiory badawcze powinien być albo wstępem do części eksperymentalnej (jednym z pierwszych podrozdziałów) albo być przeniesiony do dodatku.
- Doktorant kilkakrotnie używa słowa „anotator” – bardziej poprawną formą byłby „adnotator” (annotation tłumaczy się jako adnotacja a nie anotacja), a najlepsze byłoby użycie „osoba oceniająca”, „tester” czy ew. „ewaluator”.
- Zdarza się kilkakrotnie w tekście odwołanie do rysunku zamieszczonego kilka rozdziałów później – np. str 41, w rozdziale 2 jest odwołanie do rysunku 4.1 (czyli umieszczonego w rozdziale 4). Zamiast tego warto by się odwołać do odpowiedniego podrozdziału – zgodnie z regułami typografii rysunek powinien znajdować się możliwie blisko zaraz po odwołaniu do niego.
- Podsumowanie pracy nie zawiera wyraźnego odniesienia do przedstawionych na początku celów wraz z oceną, na ile zostały zrealizowane.

- Brak informacji dot. osiągnięć naukowych Doktoranta utrudnia odniesienie się do nich w recenzji.

Podsumowując moją recenzję chciałbym stwierdzić, że przedstawiona praca przedstawia drogę od koncepcji opisanej formalnie za pomocą nieskomplikowanego modelu, który choć wywołuje znaczny niedosyt poznawczy, zostaje potem uzupełniony bardziej precyzyjnym przedstawieniem problemu detekcji kontrowersji oraz zastosowanie opracowanych przez Doktoranta metod do oceny kontrowersji różnych bytów (artykuły w Wikipedii, zapytania wyszukiwarek internetowych). Doktorant opracował też dedykowane rozwiązania programistyczne w postaci wtyczki do przeglądarki oraz serwera dostarczającego implementację metody ustalającej kontrowersyjność zapytania.

Chciałbym stwierdzić, że Doktorant podjął się rozwiązania ciekawego przedsięwzięcia i konsekwentnie realizował założone cele, choć nie wszystkie zakończyły się sukcesem (wspomnieć należy tu model obarczony niedociągnięciami oraz zakończoną niepowodzeniem predykcją kontrowersji artykułów). Pozostała część pracy zawiera jednak w zasadzie prawidłowo prowadzone wywoły (wyłączając problemy z symboliką używaną w równaniach) wskazujące na poprawność realizacji założonych celów, a nawet w przypadku porażki (predykcja kontrowersji) Doktorant przedstawia szczegółową dyskusję jej przyczyn, co samo w sobie również jest wartościowe z punktu widzenia poznawczego.

Doktorant posiada wystarczającą wiedzę teoretyczną w dziedzinie Informatyki (a w szczególności bardzo sprawnie porusza się po tematyce Informatyki Społecznej). Opracowane przez Doktoranta oryginalne rozwiązania postawionego problemu naukowego dowodzą posiadanych przez Niego umiejętności efektywnego prowadzenia badań naukowych.

Jednocześnie stwierdzam, że przytoczone przeze mnie w tekście recenzji uwagi krytyczne i dyskusyjne nie wpływają w znacznym stopniu na poziom prezentowanej pracy i powinny być traktowane jako propozycje ewentualnych modyfikacji strategii prowadzenia dalszych badań przez Doktoranta. W szczególności proszę o bardziej precyzyjne konstrukcje wszelkich formalizmów w ramach przyszłej Jego pracy.

W uzupełnieniu do wniosków merytorycznych płynących z pracy chciałbym przedstawić ocenę dorobku naukowego Doktoranta. W popularnej bazie publikacji informatycznych DBLP posiada On 3 publikacje. W bazie Web of Science posiada jedną publikację (Q1, wysoki IF=3,4), która jeszcze nie posiada cytowań z pewnością dlatego, iż została niedawno opublikowana (styczeń 2018). W bazie SCOPUS posiada On 2 artykuły. Parametry te spełniają wymagania stosownej ustawy, jednak jeśli porównać je ilościowo z innymi przewodami doktorskimi, nie plasują się w czołówce. Obserwacja ta jest łagodzona faktem wysokiego IF najważniejszej publikacji Doktoranta, tym bardziej że jest On jej pierwszym autorem.

**Zgodnie z moim rozeznaniem przedstawiona rozprawa spełnia warunki określone przez Ustawę z dnia 14 marca 2003 (wraz z późniejszymi zmianami) o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz.U. 2003 Nr 65 poz. 595), dlatego też wnoszę o dopuszczenie Pana mgr inż. Kazimierza Zielińskiego do dalszych etapów przewodu doktorskiego.**

