



Wroclaw, 28.05.2012

dr hab. inż. Przemysław Kazienko, prof. nadzw.
Instytut Informatyki, Politechnika Wroclawska
Wyb. Wyspiańskiego 27, 50-370 Wroclaw
Email: kazienko@pwr.wroc.pl, tel: 71 3203609
WWW: <http://www.ii.pwr.wroc.pl/~kazienko/>

RECENZJA

Rozprawy doktorskiej mgr Radosława Nieleka pt. „*Designing Algorithms for Realising Social Goals*”

Rozprawę napisano pod kierunkiem dr hab. Adama Wierzbickiego w Polsko-Japońskiej Wyższej Szkole Technik Komputerowych i ukończono w 2012 roku. Rozprawa jest napisana w języku angielskim.

Recenzja została wykonana na zlecenie Dziekana Wydziału Informatyki Polsko-Japońskiej Wyższej Szkoły Technik Komputerowych w Warszawie.

I. Przedmiot i problematyka rozprawy

Rozprawa doktorska mgr Radosława Nieleka jest poświęcona analizie systemów społecznych, w których osoby (ściślej jednostki społeczne – *social entities*) wchodzi wzajemnie w interakcje. Tego typu sytuacja ma miejsce zarówno w świecie rzeczywistym, dla którego trudno o duże próbki rzeczywistych danych, ale także w świecie wirtualnym, w którym interakcja następuje za pośrednictwem komputera. Na szczególne wyróżnienie zasługują tutaj zarówno wszelkiego rodzaju serwisy społecznościowe (Facebook, Nasza Klasa), jak i systemy aukcyjne (Allegro) czy serwisy umożliwiające publikowanie publicznie dostępnych wiadomości i komentarzy. Wszystkie one dostarczają olbrzymich ilości informacji. Szczególnie te dwa ostatnie obszary ludzkiej aktywności były analizowane przez Doktoranta w ramach recenzowanej rozprawy.

Należy, przy tym podkreślić, że analizy systemów społecznych są obecnie dziedziną bardzo intensywnie eksplorowaną przez wiele różnych specjalności naukowych: informatykę, fizykę, antropologię i socjologię. Świadectwem owego wzmożonego zainteresowania jest duża i rosnąca liczba publikacji, nowe konferencje i warsztaty czy nawet nowe czasopisma naukowe. Recenzowana praca doktorska ma charakter mocno interdyscyplinarny i w szczególności łączy informatykę z socjologią. Dowodem na to jest nawet tytuł, w którym wspomniano zarówno o celach społecznych (*social goals*) – socjologia, jak i algorytmach – informatyka. Obecnie w nauce światowej można zauważyć wyraźną tendencję do badań interdyscyplinarnych i recenzowana praca doskonale się w nią wpisuje. Stwarza to jednocześnie pewne problemy, gdyż każdy doktorat jest broniący w konkretnej dyscyplinie. W tym przypadku praca należy do nauk technicznych -



Instytut Informatyki

Politechnika Wroclawska
Wydział Informatyki i Zarządzania
Wybrzeże Wyspiańskiego 27
50-370 Wroclaw

tel.: 071 320 3516
fax: 071 321 1018

email: sekretariat.ii@pwr.wroc.pl



informatyka, i w związku z tym powinna mieć silnie zarysowany charakter tej dziedziny. Doktorant wywiązał się z tego zobowiązania i informatyczny charakter pracy nie budzi wątpliwości.

Szczegółowy obszar badawczy pracy związany jest z dwoma dziedzinami: (i) systemami aukcyjnymi, w których zaufanie i reputacja pomiędzy sprzedającym i kupującym jest bardzo istotnym elementem zarówno ekonomicznego jak i społecznego procesu oraz (ii) analizą odczuć i emocji (*sentiment analysis*) zawartych w komunikatach (zwykle krótkich tekstach) tworzonych przez ludzi - użytkowników. Z pierwszą dziedziną łączy się pkt. 3.1, podczas gdy z drugą - rozdział 2 i pkt. 3.2.

Dwie tezy pracy zostały umieszczone na str. 9 a w wersji polskiej na str. 6. Dotyczą one (i) badania wpływu budowy systemu informatycznego na zachowania użytkowników i zjawiska społeczne oraz (ii) projektowania algorytmów, które by umożliwiały osiągnięcie określonych celów społecznych.

II. Układ i zawartość rozprawy

Recenzowana rozprawa doktorska liczy 107 stron i składa się z 4 rozdziałów, bibliografii, wykazu rysunków i tabel. Jest napisana w bardzo ładnym języku angielskim z bardzo małą liczbą błędów językowych. Z tego powodu łatwo się ją czyta.

Pierwszy rozdział to obszerne wprowadzenie do problematyki pracy, w tym kontekst analizy systemów społecznych, analizy zachowań użytkowników czy symulacji społecznych.

Zasadnicze elementy pracy to rozdział drugi i trzeci. W drugim Autor rozważa różnorakie zależności pomiędzy komentarzami umieszczanymi przez sprzedających i kupujących w serwisie aukcyjnym (Allegro) a także próbuje odnaleźć korelację między wiadomościami dotyczącymi spółek giełdowych, a ściślej ich negatywnym lub pozytywnym wydźwiękiem, a rzeczywistym kursem odpowiednich akcji na giełdzie.

Symulacyjne badania nad systemami reputacji, w szczególności wpływ różnych ich parametrów, w tym posiadanej informacji zwrotnej są zawarte w rozdziale trzecim. W nim także umieszczono opis systemu i skróconą analizę mobilnego systemu monitorującego emocje użytkownika wyrażone poprzez teksty przez niego tworzone.

Podsumowanie pracy to rozdział 4, który zawiera także kierunki dalszych badań. Po nim następuje spis literatury oraz rysunków i tabel. W bibliografii umieszczono cztery prace Doktoranta. Kolejne cztery, z niezrozumiałych powodów, zawarto w dolnych przypisach, co daje łącznie osiem publikacji.





III. Oryginalne osiągnięcia

Oryginalne osiągnięcia zawarte w rozprawie mieszają się w rozdziałach 2 oraz 3. Należą do nich w szczególności:

1. Analiza spirali nienawiści a bardziej ogólnie zależności pomiędzy charakterem komentarzy w systemach aukcyjnych na przykładzie systemu Allegro, w szczególności dotyczących ich kolejności i długości.
2. Analizy treści komentarzy w systemie aukcyjnym, w tym wykazanie, że komentarze pozytywne mają niską zawartość użytecznej informacji, w przeciwieństwie do komentarzy negatywnych i neutralnych.
3. Propozycja zmiany sposobu oceniania w serwisie aukcyjnym.
4. Analiza zależności pomiędzy negatywnym lub pozytywnym charakterem wiadomości dotyczących konkretnych spółek giełdowych a kursami ich akcji na zagregowanym przykładzie spółek z indeksu WIG20. W efekcie nie stwierdzono silnych zależności, które umożliwiłyby np. predykcję zmian kursów na podstawie dużej liczby negatywnych lub pozytywnych wiadomości.
5. Analiza wpływu konstrukcji różnych metod (nieco na wyrost nazwanymi algorytmami) na osiąganie lub nie wybranych celów społecznych, na przykładzie systemów reputacyjnych w serwisach aukcyjnych, w szczególności ocena odporności całego systemu na różne strategie osób wchodzących w interakcje z zagwarantowaniem niezmiennego poziomu liczby dokonywanych transakcji. W tym celu przeprowadzono badania symulacyjne w środowisku agentowym.
6. Propozycja systemu do analizy stanu emocjonalnego użytkowników systemów mobilnych, który umożliwia także klasyfikowanie a nawet filtrowanie wiadomości albo negatywnych albo pozytywnych.

Należy zauważyć, że kierunek badań zawartych w rozprawie jest ważny zarówno z poznawczego jak i z praktycznego punktu widzenia. Uzyskane przez Autora rezultaty wskazują, że (i) systemy reputacji mogą mieć różną charakterystykę i właściwości, (ii) użyteczne są treści tylko negatywnych i neutralnych komentarzy w systemach aukcyjnych, (iii) negatywny lub pozytywny charakter wiadomości publikowanych w sieci nie jest silnie skorelowany z kursem akcji oraz (iv) można sensownie analizować nastrój użytkownika mobilnego na podstawie tworzonych przez niego tekstów.

Cennym dokonaniem Autora jest jego osiem publikacji cytowanych w rozprawie w większości ściśle związanych z samą rozprawą, które świadczą o międzynarodowej weryfikacji zaprezentowanych badań i tez. W szczególności należy tutaj wymienić prace w czasopismach: (i) *Polish Journal of Environmental Studies* (JCR) – związaną z pkt. 2.2 rozprawy oraz (ii) *Journal of Artificial Societies and Social Simulation* (SSCI) – wykorzystaną w pkt. 3.1.





IV. Uwagi krytyczne i dyskusyjne

Uwagi ogólne

1. Praca jest głównie kompilacją różnych artykułów Doktoranta, przy czym nie do końca skutecznie przeprowadzono proces ich integracji w spójną całość. Przeglądy literaturowe są miejscami rozproszone, oznaczenia są niespójne (np. ρ oznacza etykietę komentarza, ale także reputację), czy nieco inne koncepcje opisu. W środku pracy (pkt. 3.1.2) umieszczono przegląd literaturowy dotyczący części rozdz. 3, a powinien on być zintegrowany ze wstępem; podobnego punktu nie ma natomiast w rozdz. 2.
2. Tytuł pracy sugeruje, że w treści czytelnik odnajdzie analizy algorytmów do osiągnięcia określonych celów społecznych lub chociażby opisy samych algorytmów. Drugi rozdział niezupełnie mieści się w tym obszarze – nie zawiera algorytmów, a jedynie badania danych związanych z zachowaniami społecznymi. Należy zauważyć, że wzór to nie algorytm, chociaż często algorytm to często praktyczny sposób obliczania wzoru. W rozdziale 3 także nie ma algorytmów opisanych wprost. W informatyce algorytmy są opisywane i analizowane za pomocą przyjętych metodologii, np. poprzez opis poszczególnych kroków, opis wejścia i wyjścia, pseudokod, schemat blokowy, analiza złożoności, itd. Rozdziały 2 i 3 (główna część pracy) nie zawierają w zasadzie tego typu opisów czy analiz lecz dyskusje nt. wyników obliczeń. Skoro np. chciano rozważać wpływ różnych parametrów na inne wartości odzwierciedlające użyteczność lub zachowania użytkowników i nawet wykorzystano to tego iteracji (*turn*), to można było opisać konkretne algorytmy do przeliczania odpowiednich wartości. Nielicznym przykładem wzoru, który mógłby być podstawą algorytmu jest nieoznaczony i nieco kontrowersyjny wzór na str. 81.
3. Chcąc badać wpływ tekstów umieszczanych w sieci Internet na kursy akcji (pkt. 2.2) można by bardziej uwzględnić czynnik czasowy, tj. następstwo zdarzeń. Autor w istocie porównuje to, co się mówi w danym dniu o spółkach z tym jaki jest kurs tych spółek (nie wiadomo kiedy – na zamknięciu sesji?). Warto zwrócić uwagę, że wpływ tekstów na kurs może być bardziej skomplikowany i zależeć od mniejszych odstępów czasowych, np. w trakcie dnia dużo dobrych wiadomości może wpływać pozytywnie na kurs spółki, po czym jedna negatywna wiadomość spowoduje załamanie jej kursu. Patrząc na kursy zamknięcia mamy wtedy dużo pozytywnych wiadomości przeważających nad negatywnymi i spadek kursu na zamknięciu. Dodatkowo, zamknięcie jest w godz. popołudniowych zaś wiadomości są publikowane przez cały czas. Wydaje się także, że należałoby usunąć ogólny trend i przyjrzeć się konkretnym spółkom i ich notowaniom z usuniętym wpływem całej giełdy lub spółek z danego sektora gospodarki. Ponadto, zastosowany współczynnik korelacji liniowej Pearsona bada siłę zależności pomiędzy dwoma zmiennymi (seriami) a nie to która na którą wpływa. Wydaje się, że lepiej zamiast ciągłych wartości (kursów) należało zacząć od wpływu na kierunek zmian (klasy zamiast wartości ciągłych) i sprawdzić, czy np. pozytywne wiadomości wpływają





pozytywnie na kursy, a nie jak zbadano czy dużo pozytywnych wiadomości wpływa na duże pozytywne zmiany kursów a nie tylko na niewielkie pozytywne zmiany. Wymagałoby to oczywiście innego przygotowania danych.

4. Brak porównania wyników dla systemu zamkniętego i otwartego (pkt. 3.1).
5. Metodyka badań w pkt. 3.2 (klasyfikacja) budzi pewne wątpliwości.

Uwagi szczegółowe

6. Pierwszy rozdział zajmujący ponad 30 stron (niemal 1/3 pracy) wydaje się zbyt długi w stosunku do reszty. Zawiera on, co jest zrozumiałe, ciekawy przegląd problematyki, ale spora część tego przeglądu ma dość luźny związek z resztą pracy i z tego powodu nie zasługuje na tak obszerne omówienie. Dodatkowo, osobne przeglądy literaturowe można odnaleźć w kolejnych rozdziałach.
7. W rozdziale drugim zabrakło ogólnej informacji o badanym środowisku (Allegro), np. w postaci blokowego schematu procesów, które były badane.
8. Brak osobnej listy symboli użytych w pracy, zwykle umieszczanej na początku. Lista taka ułatwiłaby zrozumienie niektórych fragmentów. Jest ona zdecydowanie bardziej przydatna niż lista rysunków lub tabel. Dodatkowo, Autor uniknąłby niepotrzebnych konfliktów oznaczeń, np. symbol t to w zależności od miejsca zarówno transakcja, termin jak i czas, zaś komentarz to zarówno c – wzór (1), jak i d – wzór (5).
9. Pewne zamieszanie wprowadzają niespójne oznaczenia z pkt. 2.1.3.2, np. rola r , która na liście symboli oznacza ‘kupującego’ lub ‘sprzedawcę’ (str. 41), we wzorach (2), (3) przyjmuje dziwne wartości logiczne jako element koniunkcji, równej ‘neg’ lub ‘neu’. Jest to niezrozumiałe.
10. Tabele 2.2 i 2.3 mogłyby mieć sumy jako dodatkowe kolumny, gdyż nie bardzo wiadomo względem jakich wartości podane są procenty (co jest całością) – po uważnej lekturze można wywnioskować, że chodzi tutaj o wiersze.
11. „The most popular comment appears 33k times” – szkoda, że Autor nie zacytował tego najpopularniejszego komentarza.
12. „similar to Harald Baayen’s degree of productivity P” – brak odpowiedniego odesłania do literatury.
13. Na początku 2.1.4.2 mamy liczbę 1,7 mln. komentarzy i takową można by się spodziewać w wynikach. Natomiast na rys. 2.3. mamy wartości ‘tylko’ ponad 11.000, które dodatkowo nie sumują się z tymi z tab. 2.2 i 2.3 łącznie. W efekcie nie bardzo wiadomo na jakim zbiorze jakie badania były wykonywane i jaka jest przyczyna tych niespójności.
14. Wzór 5 (str. 49) jest formalnie niepoprawny, jego mianownik powinien mieć postać $\{d:t_i \in d\}$. Dotyczy to także wyjaśnienia poniżej.
15. W pkt. 2.1.5.1 jest wymienione pojęcie emo_i oraz τ , przy czym nie bardzo wiadomo jak się je liczy, mamy tylko krótką informację że wyraża pozytywne lub negatywne emocje.
16. Liczności zbiorów ze str. 50 są niespójne z tab. 2.2 i 2.3. Nie jest jasne w jaki sposób zbiory I-IV powstały – przez losowanie? Dlaczego akurat takie





zbiory wylosowano? Dodatkowo, zastosowanie zbalansowanego zbioru jest kontrowersyjne, gdyż nie odpowiada to rzeczywistości. Jeszcze większe wątpliwości można nabrać, patrząc na wyniki z tab. 2.5: dokładność predykcji waha się od 70% do 90% w zależności od zbioru, co oznacza że sposób powstania tych zbiorów ma zasadnicze znaczenie dla wyników.

17. Metodologia polegająca na czterokrotnym losowaniu podziału na zbiór testowy i treningowy jest mało reprezentatywna. Zwykle w uczeniu maszynowym stosuje się w tym celu tzw. walidację krzyżową (*cross-validation*) i to 10-krotną (*10-fold*), przy czym możliwie dba się o jednakowy rozkład klas w każdej z 10 części.
18. Nie wiadomo na czym polegała klasyfikacja, tzn. co było wejściem klasyfikatora? Zbiór terminów? Wartości SI (różne POS, NEG i łączna)?
19. Wzór (8) budzi wątpliwości. Oznacza on, że dla dwóch niezwiązanych ze sobą zupełnie komentarzy (tylko wystawionych później), tj. dla innych transakcji i wystawionych przez innych użytkowników istnieje jakiś wzrost sentymentu. Innymi słowy z biegiem czasu ludzie ogólnie łagodnieją. Dodatkowo, nie bardzo wiadomo o jakie sumy chodzi po prawej stronie implikacji. Z dalszych opisów można jednak wywnioskować, że Autorowi chyba chodziło o coś nieco innego: wydźwięk (sentyment) kolejnych, niepozytywnych komentarzy dla tej samej transakcji jest coraz bardziej negatywny. Niestety wzór (8) tego nie wyraża.
20. W pkt. 2.2.6 podano fragment kodu w języku programowania zamiast zwykłej definicji funkcji. Dodatkowo, nie bardzo wiadomo jakie znaczenie mają funkcje `pos()` i `neg()` a także co to jest 'fragment'? Jak to się ma do oznaczeń z pkt. 2.1.3.2?
21. Nie bardzo wiadomo co obrazuje wykres 2.5. Indeks E jest to funkcja słowa kluczowego K – czyli firmy K, co oznacza, że jest on obliczany dla konkretnej spółki giełdowej. Dla której spółki mamy wykres 2.5? Wydaje się, że jest to średnia z 20 spółek, nie jest to jednak pewne. Zakładając, że tak jest, kontrowersyjnym jest badanie zależności pomiędzy tym co się mówi (ściślej pisze) o 20 spółkach **łącznie**, a indeksem WIG20. Oznacza to, że badamy to co się ogólnie mówi z tym co się ogólnie dzieje na giełdzie. Dużo lepszym byłoby sprawdzenie tego dla konkretnych spółek z osobna i to z usunięciem trendu (WIG20).
22. „any permutation τ of I ” (str. 67) – co to jest I ? Wzór (9) jest niejasny. Wzór (10) jest nieopisany, podobnie (11).
23. Szeroko używane pojęcie *utility* powinno być precyzyjnie zdefiniowane – nie wiadomo jak jest liczone. Potrzebny byłby co najmniej konkretny wzór. W efekcie nie wiadomo co oznacza sformułowanie „All agents start with a zero utility”, oraz jakie znaczenie mają osie Y na rys. 3.7, czy trzy kolumny w tab. 3.1.
24. „The coefficient α is used to control the importance of implicit negative feedbacks” (str. 81). Jest to niezrozumiałe, zwłaszcza biorąc pod uwagę, że α to współczynnik stojący przy transakcjach bez komentarza m^* a mamy $m^* = n - m$.





25. Nieoznaczony wzór na str. 81 jest kontrowersyjny. Powoduje on wzrost ρ_u z biegiem czasu, w skrajnym przypadku do nieskończoności. Należało np. użyć współczynnika 0,5 przed drugim składnikiem sumy, a jeszcze lepiej osobnego współczynnika β i $(1-\beta)$ odpowiednio przed oboma składnikami, $\beta \in [0;1]$. W efekcie nieprawdziwe jest także stwierdzenie: „Note that over time, the impact of the initial reputation decreases exponentially”.

Uwagi formalne:

Praca ogólnie poprawna formalnie (pomijając kilka uwag szczegółowych dotyczące formalizmów), aczkolwiek sprawia wrażenie połączonej z kilku fragmentów (artykułów), bez ich pełnej integracji. W efekcie mamy np. listę symboli w środku pracy na str. 41 dotyczącą tylko jednego rozdziału a brak jest ogólnej notacji (spisu symboli), w tym odnoszącej się do innych rozdziałów. Innym objawem są niejednorodne referencje do literatury (czasami wykorzystujące przypisy u dołu strony). Żadnych wątpliwości nie budzi bardzo ładny język angielski.

26. “a NLP system” -> an NLP system.
27. Niektóre wzory mają oznaczenia (numerację) a niektóre nie, przy czym nie ma to żadnego związku ani z ich znaczeniem ani cytowaniem. Jedynym wyjaśnieniem wydaje się to, że pochodzą bezpośrednio z odpowiednich artykułów.
28. Praca zawiera odesłania do 4 artykułów autora umieszczonych w rozdziale *Bibliography*, oraz odwołania do 4 artykułów, których opisy bibliograficzne z niezrozumiałych powodów umieszczono jedynie w przypisach.

Bibliografia i cytowania

29. Niektóre pozycje bibliograficzne są niecytowane, w szczególności brak cytowań następujących prac (chyba, że użyto innego schematu odwołań):
- Breznitz D., Murphree M., Goodman S., Discussion about privacy, “Computer Innovative Technology for Computer Professionals 44” 2011, No. 6, 100–102.
 - Resnick M., Turtles, Termites, and Traffic James. Explorations in Massively Parallel Microworlds, The MIT Press 1997.
 - Yang B., Lugger M., Emotion recognition from speech signals using new harmony features, Signal Processing 2010, Vol. 90, No. 5, 1415–1423.
 - Tsai T., Lo J., Chen W., Design and Implementation of Mobile Personal Emotion, Proceedings of the 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, IEEE Computer Society, 2009, 430–435.
30. Bajari P., Hortacsu A., *The Winner’s Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions*, The RAND Journal of Economics, Vol. 34, No. 2 (Summer 2003), 329–355 – została zacytowana jako *Bajari and Hortacsu 1999* (niespójny rok)
31. Konwencja odesłań przyjęta przez Autora ma pewne wady. Otóż w przypadku tych samych 2-3 autorów publikujących w tym samym roku nie ma możliwości odróżnienia cytowań. W efekcie dwie pozycje:





- a) Berners-Lee T., Hall W., Hendler J.A., O'Hara K., Shadbolt N., Weitzner D.J., *A Framework for Web Science*, Foundations and Trends in Web Science (now Publishers Inc.) 1, No. 1 (2006), 1-134.
- b) oraz Berners-Lee T., Hall W., Hendler J., Shadbolt N., Weitzner J., *Computer Science: Enhanced: Creating a Science of the Web*, "Science" 2006, 769-771.

są cytowane tak samo jako (*Berners-Lee, Hall, Hendler et al. 2006*), chyba że któraś nie jest cytowana, str. 12.

V. Podsumowanie i ocena rozprawy

Podsumowując, należy stwierdzić, że rozprawa pomimo pewnych zastrzeżeń stanowi istotny wkład w trzy szczegółowe dziedziny: analizy systemów komentarzy w społecznych systemach aukcyjnych, budowę systemów reputacji oraz w analizę sentymentu. Autor uzyskał wartościowe i oryginalne z naukowego punktu widzenia rezultaty zawarte zwłaszcza w analizie wyników eksperymentalnych badań na danych rzeczywistych i symulacyjnych.

Doktorant posiada dość znaczący, przynajmniej w ujęciu liczbowym, dorobek naukowy wyrażony przez osiem anglojęzycznych artykułów i referatów konferencyjnych zacytowanych w rozprawie, w tym w jednym w czasopiśmie z listy JCR i jednym z listy SSCI. W uznanej bazie *Scopus* zarejestrowanych jest obecnie nieco więcej – 11 prac Doktoranta.

W związku z powyższym stwierdzam, że opiniowana rozprawa doktorska mgr Radosława Nieleka spełnia wymagania stawiane w obowiązujących przepisach ustawy o stopniu naukowym doktora i wnoszę o dopuszczenie jej Autora do publicznej obrony.



Instytut Informatyki

Politechnika Wroclawska
Wydział Informatyki i Zarządzania
Wybrzeże Wyspiańskiego 27
50-370 Wrocław

tel.: 071 320 3516
fax: 071 321 1018
email: sekretariat.ii@pwr.wroc.pl