

Uniwersytet Ekonomiczny w Poznaniu
Wydział Informatyki i Gospodarki Elektronicznej

Streszczenie pracy doktorskiej

*Modelowanie siły powiązań pomiędzy
jednostkami w telekomunikacyjnej
sieci społecznej*

Autor

Bartosz Perkowski

Promotor: Prof. dr hab. Witold Abramowicz
Promotor pomocniczy: dr Agata Filipowska

Poznań 2016

Istotą niniejszej pracy jest opracowanie i przedstawienie metody, której celem jest wyznaczenie siły powiązania pomiędzy jednostkami telekomunikacyjnej sieci społecznej na podstawie danych pochodzących z dzienników połączeń. Przeprowadzone badania empiryczne wykazują, że dzięki opracowanej metodzie możliwe jest wyznaczenie powiązań par jednostek oraz odwzorowanie relacji zachodzących pomiędzy nimi za pomocą siły relacji.

Sieć społeczna jest strukturą społeczną składającą się z jednostek, które są połączone ze sobą różnymi rodzajami powiązań (relacji), i pomiędzy którymi dochodzi do wymiany informacji na zasadach wspólnie wyznawanych norm i wartości [Krupski2003]. Sieć społecznościowa definiowana jest jako szczególny przypadek sieci społecznej, realizowany z wykorzystaniem narzędzi informatycznych. W telekomunikacji, sieć społecznościowa wynika z powiązań między jednostkami, powstających jako efekt wykorzystania usług telekomunikacyjnych. Powiązania te opisują relacje międzyludzkie, które mogą przyjmować różny charakter. Przykładowymi charakterami relacji definiowanymi w psychologii są: przyjaźń, miłość, małżeństwo, pokrewieństwo, współpracownicy czy więzi sąsiedzkie [Argyle2002]. Przedmiotem badań w pracy są relacje społeczne wynikające z interakcji pomiędzy wzajemnie na siebie oddziałującymi jednostkami lub grupami ludzkimi (społecznościami), zachodzące zgodnie z przyjętymi normami i wzorcami. Wyróżnić można dwa rodzaje relacji społecznych: współdziałanie społeczne i konflikt społeczny [Znaniński2011].

Historia sieci społecznych sięga lat 90. XIX wieku, kiedy zaczęto analizować proces kształtowania się społeczności na podstawie różnych rodzajów powiązań zachodzących między jednostkami. Wykazano, że społeczności wynikają zarówno relacji osobistych, jak i formalnych [Tönnies1887]. Dalsze prace obejmowały wpływ różnych cech społecznych i psychologicznych na budowanie relacji międzyludzkich: płeć jednostek [Moreno1943], teorię wymiany opisującą sieć społeczną jako układ wymiany dóbr pomiędzy jednostkami i społecznościami [Blau1964], specyfikę pozycji w organizacji [White1943, Magee2008], czy stereotypy jednostek [McPherson2001].

Badania związane z analizą sieci społecznych wykazują, że jakość modeli dotyczących modelowania sieci społecznej jest wyższa, gdy uwzględniany jest charakter powiązań [Liben-Nowell2003]. Wiele prac dotyczących sieci

społecznych opisuje relacje między jednostkami w sposób binarny, wskutek czego charakter relacji społecznej pomiędzy jednostkami zanika lub jest ukazany w sposób szcążkowy. Traktowanie wszystkich relacji binarnie skutecznie ogranicza możliwości analizy istniejących sieci społecznych [Xiang2010], a w szczególności wyróżnianie par użytkowników będących w silnej relacji (np. rodzina czy przyjaciele) czy par posiadających słabsze powiązanie (np. znajomi). Wprowadzenie miary siły relacji, polegającej na wartościowaniu powiązań w sieciach społecznych, umożliwiło lepsze odwzorowanie relacji zachodzących pomiędzy jednostkami i zapoczątkowało badania nad znaczeniem siły relacji w strukturze sieci społecznej. Wykazano, że nie tylko silne powiązania między jednostkami są istotne, ale także słabe relacje mogą mieć kluczowe znaczenie, zwłaszcza w przypadku powiązań występujących między społecznościami [Granovetter1973]. Obecne prace dotyczące analizy sieci społecznych skoncentrowane są na wyznaczeniu siły powiązań między jednostkami z uwzględnieniem społecznego charakteru tych powiązań. Na podstawie analizy mediów społecznościowych takich jak Facebook¹ czy Twitter² wyznaczone są zmienne niezależne mogące mieć wpływ na charakter powiązań: intensywność (liczba wymienionych słów, wymienionych wiadomości, liczba komentarzy do zdjęć), intymność (liczba znajomych, intymność słów, liczba wspólnych wystąpień na zdjęciach, status relacji), czas trwania (liczba dni od pierwszego kontaktu), wzajemność (liczba wymienionych odnośników, korzystanie z tych samych aplikacji), strukturalność (liczba wspólnych znajomych i grup), emocjonalność (związana z analizą pozytywnych i negatywnych emocji w wiadomościach) oraz odległość społeczna (różnica wieku, poziomu edukacji czy poglądów politycznych) [Gilbert2009].

Zaproponowana metoda umożliwia wyznaczenie siły powiązania między jednostkami na podstawie jakościowych atrybutów relacji wynikających z parametrów dotyczących komunikacji międzyludzkiej. Zaprezentowana w rozprawie metoda związana jest z szeregiem wyzwań badawczych, dlatego jej opracowanie musiało zostać poprzedzone szczegółową analizą literaturową. Rezultaty tej analizy zostały zawarte w pierwszej części rozprawy.

Uwzględniając prace bezpośrednio związane z proponowaną koncepcją, konieczne jest zwrócenie uwagi na szczegółowe zagadnienia dotyczące

¹ <https://www.facebook.com/>

² <https://twitter.com/>

opracowanej metody: tworzenie sieci społecznych i ich analizę, modelowanie relacji zachodzących pomiędzy jednostkami, a także odwzorowanie sieci społecznych z wykorzystaniem systemów informatycznych.

Tworzenie sieci społecznych związane jest z koniecznością identyfikacji elementów mających wpływ na postać sieci społecznej oraz wykorzystanie sieci w praktyce. W ramach badań nad postacią sieci przeanalizowane zostały różnorodne z przykłady mające zastosowanie w różnych dziedzinach obejmujących biologię, medycynę [Barabási2011], informatykę [Easley2010] oraz telekomunikację [Filipowska2015]. Każdy rodzaj sieci posiada określoną strukturę wraz z atrybutami wykorzystywanymi do przedstawienia tej struktury, dotyczącymi zarówno jednostek jak i relacji, oraz wartościami, jakie te atrybuty mogą przyjmować. Analiza sieci społecznych jest natomiast podejściem mającym na celu badanie struktur społecznych, koncentrującym się na kontekście społecznym jednostki oraz relacjach pomiędzy jednostkami, a nie na indywidualnych własnościach jednostek [Wasserman1994]. Dlatego też w pracy przeanalizowane zostały metody mające zastosowanie w analizie sieci społecznych, na podstawie której zidentyfikowano atrybuty stosowane do opisu elementów sieci społecznych oraz sposoby obliczenia wartości tych atrybutów.

Modelowanie powiązań pomiędzy jednostkami wymagało szczegółowej analizy dostępnych metod służących do opisu relacji. Identyfikacja mechanizmów, dzięki którym sieci społeczne ewoluują jest wyzwaniem badawczym, umożliwiającym zrozumienie sytuacji, w których powstają nowe powiązania pomiędzy jednostkami, a w konsekwencji zmienia się struktura sieci w czasie [Liben-Nowell2003]. W ramach przygotowanej rozprawy przeanalizowano prace dotyczące homofilii jednostek oraz jej wpływu na kształtowanie relacji [McPherson2001], a także zidentyfikowano czynniki wpływające na opis relacji, takie jak złożoność relacji i sposoby mierzenia powiązań [Ansari2011]. Wyprowadzone zostało pojęcie siły relacji, polegające na wartościowaniu powiązań w sieciach społecznych, która umożliwia lepsze odwzorowanie relacji zachodzących pomiędzy jednostkami. Posługując się przykładami badań związanych z analizą powiązań wśród użytkowników takich portali społecznościowych jak Facebook [Kumar2012] czy Flickr³ [Zhuang2011] zidentyfikowane zostały ilościowe atrybuty wykorzystywane do obliczenia siły powiązania, które były podstawą dalszych prac mających na celu identyfikację jakościowych atrybutów siły relacji [Gilbert2009].

3 <https://www.flickr.com/>

Ponadto, ze względu na ukierunkowanie pracy na telekomunikacyjną sieć społeczną, przeanalizowane zostały dostępne metody umożliwiające wyznaczenie siły relacji między jednostkami na podstawie dzienników połączeń oraz wskazano ograniczenia wynikające ze stosowanych, ilościowych atrybutów relacji [Kazienko2009].

Przeprowadzanie analiz danych telekomunikacyjnych wymaga odpowiedniej klasy narzędzi. Wyzwanie stanowi wybór takiego narzędzia, które umożliwi zamodelowanie i efektywne przetwarzanie grafowych struktur danych stosowanych w telekomunikacyjnych sieciach społecznych. Eksperymenty przeprowadzone w ramach przedstawionej pracy umożliwiły selekcję narzędzi dostosowanych do grafowej struktury sieci oraz pozwalających na wdrożenie modelu odzwierciedlającego połączenia w sieciach telekomunikacyjnych przy zapewnieniu wysokiej efektywności przetwarzania danych.

Badania nad opracowaniem metody wyznaczania siły powiązań stanowią dodatkowy wkład do nauki. Dzięki przeprowadzeniu analizy literatury wskazane zostało znaczenie poprawnego odwzorowania charakteru relacji zachodzących pomiędzy jednostkami za pomocą ich siły. Ponadto, zidentyfikowany został obszar badań związany z jakościowymi atrybutami powiązań dla relacji występujących na portalach społecznościowych. Problem badawczy związany z atrybutami jakościowymi w telekomunikacji nie został dotychczas sformułowany, a w konsekwencji nie ma doniesień o próbach stworzenia podobnej metody. Osobny wkład stanowi część związana z analizą narzędzi, dzięki której dokonano porównania dostępnych rozwiązań, a którego nie można było dokonać tylko i wyłącznie na podstawie analizy literatury.

Proces testowania metody wyznaczania siły powiązań pomiędzy jednostkami telekomunikacyjnej sieci społecznościowej wymaga przygotowania odpowiedniego zbioru testowego. W dalszej części pracy przedstawione zostały dane oraz procedury, które posłużyły do utworzenia zbioru testowego. Opisana została struktura plików źródłowych, zawierających dane o zdarzeniach telekomunikacyjnych oraz przedstawiono statystyki dla całego zbioru zdarzeń, uwzględniając zarówno parametry sieci dotyczące węzłów jak liczba nadawców i odbiorców, a także związane z krawędziami, do których zaliczyć można liczbę zdarzeń, relacji oraz charakterystyk tych relacji. Ponadto, opracowano model danych wykorzystywany w grafowej bazie danych oraz przeprowadzono import danych do odpowiednich narzędzi, w celu przeprowadzenia dalszych eksperymentów dotyczących metody.

Opis metody został zaprezentowany w trzeciej części pracy. Pierwszym etapem była identyfikacja jakościowych atrybutów relacji w telekomunikacyjnych sieciach społecznych. Wzorując się na przykładach badań prowadzonych dla innych sieci oraz uwzględniając analizę literatury z dziedziny socjologii [Morreale2007, Granovetter 1973] oraz psychologii [Argyle2002][Wellman1990], wyszczególnione zostały trzy atrybuty jakościowe relacji telekomunikacyjnych, uwzględnione w proponowanej metodzie: wzajemność, intymność oraz bliskość. Dla każdego z atrybutów zidentyfikowano parametry zdarzeń telekomunikacyjnych mające wpływ na dany atrybut. Wśród zidentyfikowanych parametrów należy wyróżnić: rodzaj usługi, liczbę zdarzeń, czas trwania zdarzeń, porę dnia wystąpienia zdarzeń oraz czas aktywności jednostki. Ponadto opisane zostało badanie kwestionariuszowe, które posłużyło do weryfikacji hipotez związanych z wpływem poszczególnych parametrów zdarzeń na bliskość relacji międzyludzkich. Na podstawie otrzymanych rezultatów badania oraz statystyk zbioru testowego, przygotowane zostały matematyczne reprezentacje atrybutów. Posłużyły one do wyprowadzenia modelu metody wykorzystywanej do wyznaczania siły powiązania pomiędzy jednostkami.

Ostatnim etapem była weryfikacja i walidacja opracowanej metody z wykorzystaniem dwóch eksperymentów, dzięki dokonaniu porównania rezultatów wybranej, ogólnie stosowanej metody z rezultatami opracowanej metody. Pierwszy eksperyment obejmował wyznaczenie siły relacji dla danych testowych, w wyniku którego możliwe było porównanie, jak opracowane atrybuty jakościowe wpływają na kształtowanie się siły powiązań między jednostkami. Wykazano znaczne różnice pomiędzy osiąganymi wartościami obu miar, podkreślając istotność wpływu poszczególnych atrybutów na kształt siły relacji. Przeprowadzony eksperyment pozwolił udowodnić tezę pracy dotyczącą modelowania siły powiązań na podstawie jakościowych atrybutów relacji. Na potrzeby drugiego eksperymentu została przygotowana autorska aplikacja mobilna, umożliwiająca walidację przygotowanej metody. Celem eksperymentu było: potwierdzenie lepszego odwzorowania siły relacji z wykorzystaniem opracowanej metody w porównaniu do klasycznego podejścia oraz wskazanie potencjału wykorzystania tej metody w praktyce. Osiągnięte rezultaty pokazały, że metoda wyznaczania społecznej siły powiązania lepiej odwzorowuje relacjehachodzące z wykorzystaniem usług telekomunikacyjnych niż metoda klasyczna. W 240 przypadkach lepsza okazała się opracowana, społeczna metoda, natomiast jedynie w 12 przypadkach klasyczna. Ponadto, udzielone

przez respondentów odpowiedzi wskazały na znaczenie właściwego odwzorowania siły relacji i potencjalne zastosowanie wypracowanej metody.

Zarówno zaprezentowana metoda jak i otrzymane wyniki eksperymentów mogą mieć różnorakie zastosowania. Do istotnych zastosowań zaliczyć należy:

- budowa i analiza sieci społecznościowych,
- analiza i ocena powiązań między jednostkami, tworzonych z wykorzystaniem różnorodnych środków komunikacji,
- marketing w mediach społecznościowych,
- analiza migracji w telekomunikacji.

[Ansari2011] Ansari, A., Koenigsberg, O. i Stahl, F., 2011, *Modeling Multiple Relationships in Social Networks*, Journal of Marketing Research.

[Argyle2002] Argyle, M., 2002, *Psychologia stosunków międzyludzkich*, Wydawnictwo Naukowe PWN.

[Blau1964] Blau, P.M., 1964, *Exchange and Power in Social Life*, Transaction Publishers.

[Barabási2011] Barabási, A.L., Gulbahce, N. i Loscalzo, J., 2011, *Network medicine: a network-based approach to human disease*, Nature Reviews Genetics, , nr 12, s. 56–68.

[Easley2010] Easley, D. i Kleinberg, J., 2010, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*, rozdz. The Structure of the Web, Cambridge University Press.

[Filipowska2015] Filipowska, A., Mucha, M., Perkowski, B., Szczekocka, E., Gromada, J. i Konarski, A., 2015, *Towards Social Telco Applications based on User Behaviour and Relations*, w: 18th International Conference on Intelligence in Next Generation Networks (ICIN), IEEE, s. 95–102.

[Gilbert2009] Gilbert, E. i Karahalios, K., 2009, *Predicting Tie Strength with Social Media*, w: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, New York, NY, USA, s. 211–220.

[Granovetter1973] Granovetter, M., 1973, *The Strength of Weak Ties*, The American Journal of Sociology, tom 78, nr 6, s. 1360–1380.

[Kazienko2009] Kazienko, P. i Ruta, D., 2009, *The Impact of Customer Churn on Social Value Dynamics*, International Journal of Virtual Communities and Social Networking, tom 1, nr 3.

[Krupski2003] R. Krupski, *Zarządzanie strategiczne. Koncepcje. Metody.*, Wrocław: Wydawnictwo Akademii Ekonomicznej we Wrocławiu, 2003.

[Kumar2012] Kumar, A., Rao, T. i Nagpal, S., 2012, *Contemporary Computing*, tom 306, rozdz. Using Strong, Acquaintance and Weak Tie Strengths for Modeling Relationships in Facebook Network, Springer, s. 188–200.

- [Liben-Nowell2003] Liben-Nowell, D. i Kleinberg, J., 2003, *The Link Prediction Problem for Social Networks*, w: International Conference on Information and Knowledge Management.
- [Magee2008] Magee, J. C. i Galinsky, A. D., 2008, *The Self-Reinforcing Nature of Social Hierarchy: Origins and Consequences of Power and Status*, w: IACM 21st Annual Conference Paper.
- [McPherson2001] McPherson, M., Smith-Lovin, L. i Cook, J.M., 2001, *Birds of a Feather: Homophily in Social Networks*, Annual Review of Sociology.
- [Moreno1943] Moreno, J.L., 1934, *Who Shall Survive?*, Beacon House Inc.
- [Morreale2007] Morreale, S., Spitzberg, B. i J.K., B., 2007, *Komunikacja między ludźmi. Motywacja, wiedza i umiejętności*, Wydawnictwo Naukowe PWN.
- [Tönnies1887] Tönnies, F., 1988, *Wspólnota i stowarzyszenie*.
- [Wasserman1994] Wasserman, S. i Faust, K., 1994, *Social Network Analysis: Methods and Applications*, Cambridge University Press.
- [Wellman1990] Wellman, B. i Wortley, S., 1990, *Different Strokes from Different Folks: Community Ties and Social Support*, The American Journal of Sociology, tom 96, nr 3, s. 558–588.
- [White1943] White, H. C., 1970, *Chains of opportunity: system models of mobility in organizations*, Harvard University Press.
- [Xiang2010] Xiang, R., Neville, J. i Rogati, M., 2010, *Modeling Relationship Strength in Online Social Networks*, w: Proceedings of the 19th International Conference on World Wide Web, ACM, s. 981–990.
- [Zhuang2011] Zhuang, J., Mei, T., Hoi, S.C.H., Hua, X.S. i Li, S., 2011, *Modeling Social Strength in Social Media Community via Kernel-based Learning*, w: ACM Multimedia, ACM Multimedia.
- [Znaniecki2011] Znaniecki, F. i Hałas, E., 2011, *Relacje społeczne i role społeczne: niedokończona socjologia systematyczna*, PWN.