

Jakub Swacha

Analiza sieci społecznych jako źródło wiedzy o czynnikach ryzyka dla funkcjonowania organizacji

Ekonomiczne Problemy Usług nr 68, 565-572

2011

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

JAKUB SWACHA

Państwowa Wyższa Szkoła Zawodowa w Gorzowie Wlkp.

ANALIZA SIECI SPOŁECZNYCH JAKO ŹRÓDŁO WIEDZY O CZYNNIKACH RYZYKA DLA FUNKCJONOWANIA ORGANIZACJI

Wprowadzenie

Analiza sieci społecznych (*Social Network Analysis* – SNA) to narzędzie badawcze, umożliwiające mapowanie i analizę relacji zachodzących pomiędzy członkami organizacji¹. Celem niniejszego artykułu jest wykazanie, że może zostać ona wykorzystana jako źródło wiedzy o czynnikach ryzyka dla prawidłowego funkcjonowania organizacji.

Jakkolwiek powiązania w sieciach społecznych mogą opierać się na dowolnym kanale komunikacji, w niniejszym artykule, ze względu na łatwość gromadzenia i przetwarzania danych, skupiono się na analizie sieci opartej na wymianie poczty elektronicznej. Jest to podejście uzasadnione, gdy poczta elektroniczna jest podstawowym kanałem komunikacji w organizacji, jak ma to miejsce w konsorcjum opisanym w zamieszczonym w dalszej części niniejszego artykułu studium przypadku.

¹ P. Stęпка, K. Subda: *Wykorzystanie analizy sieci społecznych (SNA) do budowy organizacji opartej na wiedzy*, „E-mentor” 2009, nr 1(28), <http://www.e-mentor.edu.pl/artukul/index/numer/28/id/618>

1. Wykorzystanie poczty elektronicznej w analizie sieci społecznych

Na wielkie możliwości, jakie stwarza dla analizy sieci społecznych wykorzystanie poczty elektronicznej, pierwszy zwrócił uwagę B. Wellman w swoim artykule o sieciach społecznych opartych na sieciach komputerowych².

R. Guimera i in. w roku 2002 zbadali wymianę poczty elektronicznej w rzeczywistej organizacji, aby określić jej strukturę społeczną³. Rok później J.R. Tyler i in. opisali w swej pracy metodę pozwalającą na automatyczną identyfikację społeczności w organizacji na podstawie dzienników poczty elektronicznej⁴. Również w roku 2003 opracowany został EmailNet – system, który automatycznie badał wymianę poczty elektronicznej w organizacji i generował informacje o sieciach społecznych w celu ich późniejszej analizy⁵. W roku 2006 T. Mitchell i in. zaproponowali metodę automatycznego zbierania informacji o aktywności użytkownika w oparciu o analizę danych przechowywanych na ich stacjach roboczych, w tym poczty elektronicznej⁶. Z kolei Ch. Bird i in. opisali rezultaty swojej pracy nad analizą opartej na wymianie poczty elektronicznej sieci społecznej osób pracujących nad serwerem HTTP Apache⁷. Natomiast A. Perer i M.A. Smith zaproponowali trzy sposoby wizualizacji, ukazujące hierarchiczne, korelacyjne i czasowe zależności występujące w poczcie elektronicznej, zalecając ich wykorzystanie jako łatwego w odbiorze źródła wartościowej informacji o relacjach społecznych i historii komunikacji użytkownika⁸.

W roku 2007 M. Studer i in. opublikowali pracę opisującą wyniki wykorzystania archiwów poczty elektronicznej projektu KDE do budowy typologii tak zwa-

² B. Wellman: *Computer Networks As Social Networks*, „Science” 2001, 293, s. 2031–2034.

³ R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, A. Arenas: *Self-similar Community Structure in Organizations*, arXiv:cond-mat/0211498v1, 2002.

⁴ J.R. Tyler, D.M. Wilkinson, B.A. Huberman: *Email as spectroscopy: automated discovery of community structure within organizations*, *Communities and technologies*, Kluwer, Dordrecht 2003, s. 81–96.

⁵ M. van Alstyne, J. Zhang: *EmailNet: A System for Automatically Mining Social Networks from Organizational Email Communication*, NAACSOS Conference 2003 Proceedings, CASOS, Pittsburgh, 2003, <http://www.casos.cs.cmu.edu/events/conferences/2003/proceedings/day1.zip>

⁶ T. Mitchell, S. Wang, Y. Huang, A. Cheyer: *Extracting Knowledge about Users Activities from Raw Workstation Contents*, Proceedings of the 21st National Conference on Artificial Intelligence – vol. 1, AAAI, Boston 2006, s. 181–186.

⁷ Ch. Bird, A. Gourley, P. Devanbu, M. Gertz, A. Swaminathan: *Mining email social networks*, Proceedings of the 2006 International Workshop on Mining Software Repositories, ACM, Shanghai, 2006, s. 137–143.

⁸ A. Perer, M. A. Smith: *Contrasting portraits of email practices: visual approaches to reflection and analysis*, AVI '06: Proceedings of the working conference on Advanced visual interfaces, ACM, New York 2006, s. 389–395.

nych **trajektorii uczestnictwa**, w celu lepszego zrozumienia struktury społeczności oprogramowania *open source*⁹.

W roku 2010 M. Laclavik i in. opublikowali dyskusję zalet analizy prywatnych i firmowych archiwów poczty elektronicznej, takich jak m.in. lepsze zrozumienie treści wiadomości oraz umożliwienie działania różnych typów aplikacji w kontekście wskazanej organizacji lub społeczności¹⁰.

Ostatnio D. MacLean i in. zaproponowali algorytm do tworzenia topologii społecznych (zbiorów potencjalnie zachodzących na siebie i zagnieżdżonych grup reprezentujących strukturę i zawartość sieci społecznej wskazanej osoby) w oparciu o analizę historii wymiany wiadomości i identyfikację prawdopodobnych grup poprzez wykrywanie wzorców¹¹.

2. Czynniki ryzyka

Do czynników ryzyka dla funkcjonowania organizacji, które można zidentyfikować w wyniku analizy sieci społecznych istniejących w tej organizacji, zaliczyć można:

1. Występowanie w sieci tak zwanych luk komunikacyjnych¹²: grup osób całkowicie odizolowanych od pozostałych, co oznacza ryzyko braku dostępu odizolowanych osób do głównego obiegu informacji w organizacji.
2. Występowanie w sieci tak zwanych brokerów informacji¹³: osób pośredniczących w kontaktach między wieloma osobami, między którymi nie ma bezpośrednich połączeń, co niesie ryzyko:
 - opóźnień i utrudnień w komunikacji pomiędzy osobami komunikującymi się poprzez pośredników;
 - poważnych zakłóceń w komunikacji w organizacji w sytuacji czasowej absencji takich osób lub ich odejścia z pracy.

⁹ M. Studer, N.S. Müller, G. Ritschard: *Understanding the KDE Social Structure through Mining of Email Archive*, 2nd Workshop on Public Data about Software Development, Third International Conference on Open Source Systems, IFIP, Limerick, 2007, http://mephisto.unige.ch/pub/publications/gr/wopdasd_studer_et_all_full.pdf

¹⁰ M. Laclavik, M. Kvassay, S. Dlugolinsky, L. Hluchy: *Use of Email Social Networks for Enterprise Benefit*, International Workshop on Computational Social Networks, International Conference on Web Intelligence and Intelligent Agent Technology, IEEE/WIC/ACM, Toronto, 2010, http://laclavik.net/publications/email_SN_final_poster4.pdf

¹¹ D. MacLean, S. Hangal, S.K. The, M. S. Lam, J. Heer: *Groups Without Tears: Mining Social Topologies from Email*, Proceeding of the 15th International Conference on Intelligent User Interfaces, ACM, Palo Alto, 2011, s. 83–93.

¹² P. Stęпка, K. Subda: *Wykorzystanie analizy...*, *op. cit.*

¹³ R. Cross, A. Parker: *The Hidden Power of Social Networks. Understanding How Work Really Gets Done in Organizations*, Harvard Business School Press, Boston 2004, s. 76–79.

3. W odniesieniu do połączeń pomiędzy pracownikami organizacji a osobami z zewnątrz (reprezentującymi klientów, kontrahentów lub podwykonawców):
 - wymiana wiadomości z osobami spoza organizacji prowadzona wyłącznie przez pojedynczą osobę należącą do organizacji rodzi ryzyka podobne do opisanego poprzednim punkcie, jednak o znacznie groźniejszych konsekwencjach, szczególnie w przypadku odejścia osoby kontaktującej się z partnerem spoza organizacji (brak możliwości odtworzenia przebiegu komunikacji, a w skrajnym przypadku – brak możliwości kontynuowania kontaktu); ponadto taka sytuacja oznacza brak bezpośredniej kontroli nad przebiegiem wymiany wiadomości, a zatem ryzyko: świadomego lub nieświadomego przekazania poufnych danych poza organizację, przekazywania informacji zniekształconych, prowadzenia komunikacji w sposób niezgodny z interesem organizacji;
 - wymiana wiadomości z osobami spoza organizacji prowadzona przez szeregowych pracowników z pominięciem przełożonych (i nie na ich polecenie) oznacza ryzyko zawierania uzgodnień niezgodnych z zamiarami kierownictwa organizacji, przyjmowania niepotrzebnych zobowiązań (w tym wykonywania ponadplanowych prac na rzecz podmiotów spoza organizacji) lub akceptowania nieuzasadnionych żądań wysuwanych przez podmiot zewnętrzny.
4. W odniesieniu do aktywności dla poszczególnych połączeń pomiędzy węzłami sieci wymieniłem należy takie czynniki ryzyka, jak:
 - generowanie niepotrzebnego szumu informacyjnego, nadmiaru wiadomości trudnego do ogarnięcia przez odbiorców, powodującego zmniejszenie czytelności innych, istotnych wiadomości (aż do kompletnego ich ignorowania przez odbiorców włącznie);
 - zbyt pochopne przesyłanie wiadomości o bardzo dużym rozmiarze, powodujące opóźnienia w komunikacji poprzez zapychanie łączы komunikacyjnych i szybki przyrost objętości archiwów wiadomości, utrudniający dotarcie do zgromadzonych w nich informacji;
 - ignorowanie badanego kanału komunikacji (w tym przypadku poczty elektronicznej), a w konsekwencji brak świadomości przekazanych tą drogą informacji i niewykonywanie zleconych tą drogą prac.

3. Wykorzystane narzędzia analityczne

Do przeprowadzenia badań wykorzystano autorski zestaw narzędzi analizy sieci społecznych oparty na wykorzystaniu poczty elektronicznej. Zestaw jest do-

stępny do bezpłatnego użytku, może być pobrany ze strony internetowej <http://uoo.univ.szczecin.pl/~jakubs/mxa>.

W skład zestawu wchodzi następujące narzędzia:

- *Mail Fetch* – narzędzie służące do ekstrakcji danych niezbędnych do przeprowadzenia analizy z serwera poczty elektronicznej;
- *Mail Graph* – narzędzie służące do wizualizacji sieci społecznej w postaci grafu w oparciu o algorytm Fruchtermana–Rheingolda;
- *Mail Table* – narzędzie służące do generowania raportów tabelarycznych ilustrujących aktywność komunikacyjną poszczególnych osób;
- *Mail History* – narzędzie służące do wizualizacji zmian aktywności komunikacyjnej w czasie w postaci histogramu.

4. Analiza studium przypadku

Jako przedmiot studium przypadku wybrano konsorcjum realizujące projekt międzynarodowy Baltic Museums 2.0. Członkami konsorcjum są dwie uczelnie oraz cztery muzea oceanograficzne z regionu południowego Bałtyku.

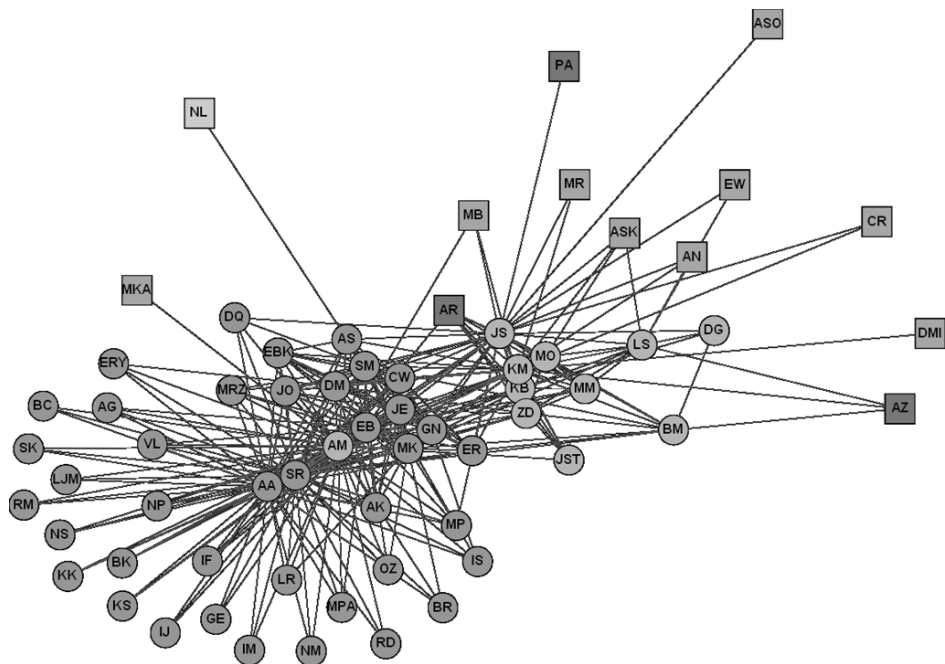
W pierwszym etapie analizy dokonano wizualizacji sieci społecznej konsorcjum z wykorzystaniem narzędzia *Mail Graph*. Wynik wizualizacji zamieszczono na rysunku 1.

Każdy węzeł grafu przedstawionego na rysunku 1 oznacza pojedynczą osobę uczestniczącą w pracach konsorcjum. Kółkami oznaczono pracowników instytucji będących uczestnikami konsorcjum, natomiast kwadratami – osoby spoza konsorcjum. Krawędź łączy dwa węzły grafu, jeżeli pomiędzy osobami, przez te węzły reprezentowanymi, miała miejsce bezpośrednia wymiana wiadomości przy użyciu poczty elektronicznej.

Już pobieżne zapoznanie się z grafem przedstawionym na rysunku 1 pozwala wskazać pewne czynniki ryzyka dla funkcjonowania konsorcjum.

Widoczne na grafie węzły z dużą liczbą krawędzi to wspomniani wcześniej brokerzy informacji. Ich obecność wskazuje na istnienie ryzyka opóźnień w komunikacji oraz wystąpienia poważnych zakłóceń w funkcjonowaniu konsorcjum w przypadku czasowej lub trwałej absencji takich osób.

Przyglądając się bliżej połączeniom z osobami spoza konsorcjum, zauważa się, że część z nich prowadzi wyłącznie do pojedynczej osoby z konsorcjum, co rodzi ryzyko zarówno zbliżone do opisanego powyżej, jak i wynikające z braku nadzoru nad komunikacją realizowaną w ten sposób. Szczegółowa analiza połączeń ujawnia także, że co najmniej w dwóch przypadkach miała miejsce bezpośrednia wymiana wiadomości szeregowych pracowników konsorcjum z osobami spoza konsorcjum.



Rys. 1. Wizualizacja sieci społecznej w postaci grafu

Źródło: opracowanie własne.

W drugim etapie analizy zbadano ilościowo aktywność komunikacyjną, związaną z realizacją projektu, z wykorzystaniem narzędzia *Mail Table*. Uzyskane tą drogą zestawienia aktywności poszczególnych osób są zbyt obszerne do bezpośredniego zamieszczenia w niniejszej pracy, nadmienić należy zatem, że zawierają:

- macierz wymiany wiadomości, ukazującą liczbę wiadomości przesłanych między każdą parą osób należących do analizowanej sieci, wraz z podaniem całkowitej liczby wiadomości wysłanych i otrzymanych przez każdą osobę;
- macierz wymiany danych, ukazującą ilość danych (w mebibajtach) przesłanych między każdą parą osób należących do analizowanej sieci, wraz z podaniem sumarycznej ilości danych (w mebibajtach) wysłanych i otrzymanych przez każdą osobę;
- listę czołowych nadawców, uporządkowaną według liczby wysłanych wiadomości, zawierającą tę liczbę, sumaryczną ilość danych (w mebibajtach) przez nich wysłanych oraz listę wszystkich ich odbiorców z podaniem liczby wiadomości adresowanych do każdego z nich i ich udziału procentowego w całkowitej liczbie wiadomości wysłanych przez danego nadawcę.

Analiza zestawień aktywności ujawniła następujące przesłanki wskazujące na istnienie czynników ryzyka dla funkcjonowania konsorcjum:

- jedna z osób w ciągu siedmiomiesięcznego okresu zatrudnienia wysłała blisko 1600 wiadomości, co wskazuje na generowanie niepotrzebnego szumu informacyjnego;
- jedna z osób w ciągu siedmiomiesięcznego okresu zatrudnienia wysłała blisko 2,4 gibibajta danych, co wskazuje na zbyt pochopne przesyłanie wiadomości o bardzo dużym rozmiarze;
- w trakcie półtorarocznego okresu poddanego analizie 16 osób należących do sieci nie wysłało ani jednej wiadomości przy użyciu poczty elektronicznej, co wskazuje na ignorowanie przez nie poczty elektronicznej jako kanału komunikacji.

Podsumowanie

Dzięki wykorzystaniu sieci społecznej jako źródła wiedzy o czynnikach ryzyka dla prawidłowego funkcjonowania organizacji możliwe jest także jej przekonstruowanie, by zminimalizować ryzyko wystąpienia zdarzeń niekorzystnych dla funkcjonowania organizacji. Zamieszczona w punkcie drugim artykułu lista czynników ryzyka, które można zidentyfikować dzięki analizie sieci społecznych, z pewnością nie jest wyczerpująca, jednak już w takiej postaci daje obraz korzyści, jakie może odnieść organizacja, stosując proponowane rozwiązanie. Należy również podkreślić łatwość przeprowadzenia analizy za pomocą oprogramowania opisanego w punkcie trzecim.

Literatura

1. Bird Ch., Gourley A., Devanbu P., Gertz M., Swaminathan A.: *Mining email social networks*, Proceedings of the 2006 International Workshop on Mining Software Repositories, ACM, Shanghai 2006.
2. Cross R., Parker A., *The Hidden Power of Social Networks. Understanding How Work Really Gets Done in Organizations*, Harvard Business School Press, Boston 2004.
3. Guimerà R., Danon L., Diaz-Guilera A., Giralt F., Arenas A.: *Self-similar Community Structure in Organizations*, arXiv:cond-mat/0211498v1, 2002.
4. Laclavik M., Kvassay M., Dlugolinsky S., Hluchy L.: *Use of Email Social Networks for Enterprise Benefit*, International Workshop on Computational Social Networks, International Conference on Web Intelligence and Intelligent Agent

- Technology, IEEE/WIC/ACM, Toronto 2010, http://laclavik.net/publications/email_SN_final_poster4.pdf
5. MacLean D., Hangal S., The S. K., Lam M. S., Heer J.: *Groups Without Tears: Mining Social Topologies from Email*, Proceeding of the 15th International Conference on Intelligent User Interfaces, ACM, Palo Alto 2011.
 6. Mitchell T., Wang S., Huang Y., Cheyer A.: *Extracting Knowledge about Users' Activities from Raw Workstation Contents*, Proceedings of the 21st National Conference on Artificial Intelligence – vol. 1, AAAI, Boston 2006.
 7. Perer A., Smith M. A.: *Contrasting portraits of email practices: visual approaches to reflection and analysis*, AVI '06: Proceedings of the working conference on Advanced visual interfaces, ACM, New York 2006.
 8. Stęпка P., Subda K.: *Wykorzystanie analizy sieci społecznych (SNA) do budowy organizacji opartej na wiedzy*, „E-mentor” 2009, nr 1(28), <http://www.e-mentor.edu.pl/artukul/index/numer/28/id/618>
 9. Studer M., Müller N.S., Ritschard G.: *Understanding the KDE Social Structure through Mining of Email Archive*, 2nd Workshop on Public Data about Software Development, Third International Conference on Open Source Systems, IFIP, Limerick 2007, http://mephisto.unige.ch/pub/publications/gr/wopdasd_studer_et_all_full.pdf
 10. Tyler J.R., Wilkinson D.M., Huberman B.A.: *Email as spectroscopy: automated discovery of community structure within organizations*, Communities and technologies, Kluwer, Deventer 2003.
 11. van Alstyne M., Zhang J.: *EmailNet: A System for Automatically Mining Social Networks from Organizational Email Communication*, NAACSOS Conference 2003 Proceedings, CASOS, Pittsburgh 2003, <http://www.casos.cs.cmu.edu/events/conferences/2003/proceedings/day1.zip>
 12. Wellman B.: *Computer Networks As Social Networks*, „Science” 2001, 293.

SOCIAL NETWORK ANALYSIS AS A SOURCE OF KNOWLEDGE ON RISK FACTORS FOR ORGANIZATION'S FUNCTIONING

Summary

In this paper we describe how Social Network Analysis can be a source of knowledge on risk factors for organization's functioning. We list eight typical risk factors and propose a practical solution, based on e-mail exchange analysis using our own software designed for this purpose, capable of identifying them in social networks.

Translated by Jakub Swacha