

Magdalena Kowalska-Musiał

Wyższa Szkoła Zarządzania i Bankowości w Krakowie

METODY WIZUALIZACJI DANYCH SIECIOWYCH

1. Wstęp

Termin sieć społeczna (*social network*) funkcjonuje w teorii socjologii od lat 50. XX wieku. Jako pierwszy posłużył się nim J. Barnes w celu nazwania układu stosunków społecznych utrzymywanych przez członków społeczności lokalnej niezależnie od norm kulturowych [Barnes 1954]. Termin ten obecnie używany jest do oznaczenia dowolnego zbioru „obiektów społecznych” rozważanego wraz z pewnym układem „społecznych powiązań” między nimi. W analizie sieci społecznych szczególnie znaczenie mają graficzne metody wizualizacji dotyczące zarówno zależności między wchodzącymi w interakcje obiektami, jak i formalnych własności powstałych na ich podstawie sieci i subsieci [Wellman 1988].

Określenie analiza sieci społecznych (*Social Network Analysis* – SNA) zaczęto powszechnie stosować na przełomie lat 70. i 80. na oznaczenie perspektywy związanej z pojęciem sieci społecznej. Sam termin SNA został prawdopodobnie zapożyczony z badań operacyjnych [Ford, Fulkerson 1969; Korzan 1978]. W okresie prekursorskim do początku lat 50. za podstawowe źródło informacji o sieciach społecznych służyły: teoria grafów oraz monografie matematyczne Harrayego i jego współpracowników.

SNA dostarcza alternatywnych wyjaśnień zachowań aktorów, nie traktując procesów społecznych jako sumy atrybutów czy norm wewnętrznych, ale pozwalając na wyłonienie emergentnych zjawisk, powstających jako efekt wzajemnej interakcji. Modelowanie SNA pozwala badaczom na konceptualizację struktury sieci przez ustalenie wzorów relacji, dyfuzję informacji pomiędzy aktorami oraz zrozumienie, jak jednostki wpływają na środowisko strukturalne sieci i kształtują je.

Aplikacja podejścia sieciowego w marketingu obejmuje obszary badawcze o charakterze empirycznym, np. zmianę postaw względem marek [Iacobucci i in. 1996], określenie struktury sieci biznesowych [Cadeaux 1997], dalej obszar badawczy rozwijający nurt sieci, mający uniwersalny charakter dla nauk społecznych, tutaj przykład mogą stanowić prace z zakresu wymiany czy zależności w sieciach [Cook 1982; Emerson 1981; Granovetter 1985], roli i pozycji [Burt 1976; 1992] czy dyfuzji

innowacji [Rogers 1962; 1976]. Trzecim wymiarem wykorzystania sieci w dyscyplinie marketingu jest strategiczny obszar badawczy, jak np. modelowanie sieci [Iacobucci, Hopkins 1992], sieci traktowane w wymiarze strategicznym [Arabie, Wind 1994], sieci konsumenckie [Martin, Clark 1996] czy alianse strategiczne na rynku [Hakansson, Sharma 1996].

2. Rozwój technik wizualizacyjnych – perspektywa retrospektywna

Wizualizacja sieci od początku jest trwale związana z rozwojem metod i technik SNA. Metoda wizualizacyjna pozwala analitykom na odkrycie struktury zarówno sieci, jak i jednostek ją tworzących. Podstawową jednostką analityczną są relacje pomiędzy aktorami społecznymi. Jeśli liczba aktorów, a zatem relacji i powiązań pomiędzy nimi, wzrasta, pojawia się obiektywna trudność w zrozumieniu, interpretacji, nawigacji oraz samej wizualizacji sieci. Istota SNA zmierza do odnalezienia wzoru połączeń, czyli określenia sposobu, w jaki elementy są ze sobą powiązane.

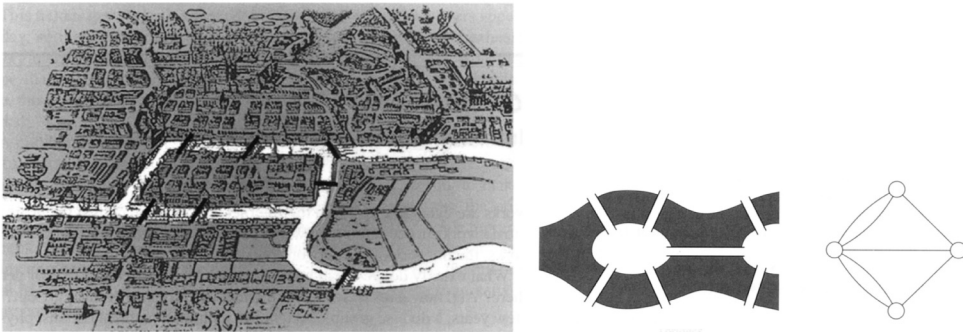
W rozwoju technik wizualizacyjnych wyróżnić można 6 etapów, są to okresy związane z ewolucją samych technik i metod graficznych, wynikających bezpośrednio z rozwoju technologicznego:

- lata 30. to odręczne rysunki sieci,
- lata 50. to pierwsze mechaniczne ilustracje sieci,
- lata 70. to pierwsze wydruki wizualizujące sieci, będące efektem oprogramowania komputerowego,
- lata 80. to wydruki wizualizujące sieć, o doskonalszej jakości obrazu oraz z zastosowaną techniką graficzną,
- 1990-2009 to przestrzenne obrazy graficzne sieci wykonywane w technologii 3D.

Według L.C. Freemana metodologia SNA zmierza do odkrycia dwóch typów wzorów w sieci: ujawnienia skupisk węzłów, które tworzą zwarte grupy społeczne, lub skupisk strukturalnie ekwiwalentnych pozycji węzłów w strukturze relacji. Od wczesnych lat 30. odręczne rysunki sieci uwzględniały obydwie te podejścia. Rysunki były skonstruowane w taki sposób, aby podkreślić wagę cech opisujących strukturę grupy lub też nacisk kładziony był na ukazanie podobieństw lub przeciwieństw pomiędzy pozycjami społecznymi zajmowanymi przez aktorów w sieci powiązań. Najwcześniejsze ilustracje sieci zostały tworzone przez J.L. Moreno, a następnie prace kontynuowali G.A. Lundberga i M. Steele oraz M.L. Northway.

Warto jednak podkreślić tutaj znaczenie i wpływ teorii grafów na inspirację dla badaczy, czego efektem jest powstanie metodologii analizy sieci. Już w 1736 r. L. Euler postawił problem rozwiązania zagadnienia mostów w Królewcu zilustrowany na rys. 1.

Problem postawiono w następujący sposób: czy można przejść kolejno przez wszystkie mosty w taki sposób, żeby przekroczyć je tylko raz. L. Euler wykazał, że



Rys. 1. Rycina obrazująca miasto Królewiec: 7 mostów na rzece Pregola, obok graf stanowiący jej odwzorowanie

Źródło: [Newman, Barabasi, Watts 2006].

jest to niemożliwe, a decyduje o tym nieparzysta liczba wylotów mostów zarówno na każdą z wysp, jak i na oba brzegi rzeki. Następcy Eulera rozwinęli teorię grafów jako dział w matematyce dyskretnej¹.

J.L. Moreno za pomocą odręcznych rysunków, nazwanych socjogramami², podjął próbę przedstawienia struktury powiązań pomiędzy aktorami. Wprowadził 5 istotnych założeń dotyczących właściwej konstrukcji obrazu sieci społecznych:

- określenia zasad odwzorowania sieci za pomocą grafów prostych,
- określenia zasad odwzorowania sieci za pomocą grafów skierowanych,
- wykorzystania kolorów w celu różnicowania relacji w sieci – multigrafy (np. odczucie sympatii i antypatii pomiędzy aktorami, ilustracja na rys. 2),
- zróżnicowania kształtów węzłów (np. różnicujące płeć aktorów),
- zróżnicowania lokacji węzłów w celu zaakcentowania ważności cech strukturalnych własności danych (węzeł na górze lub na dole symbolizował pozycję o wyższym statusie społecznym).

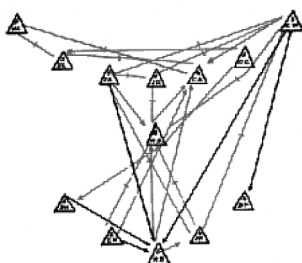
Zasady powyższe rozwijane były przez badaczy zainteresowanych zgłębieniem problemu odzwierciedlenia statusu socjometrycznego węzłów w sieci³.

¹ Teoria grafów pozwoliła badaczom na wizualizowanie dwóch różnych typów relacji pomiędzy punktami za pomocą podwójnych linii i strzałek; teoria obrazowała stany negatywne i pozytywne, dostarczyła lepszej procedury analizowania złożonych procedur społecznych, sprawiła, że struktury społeczne lepiej poddawały się opracowaniom matematycznym i statystycznym.

² Grupa punktów połączonych liniami służąca do odzwierciedlenia sieci relacji pomiędzy ludźmi w celu zidentyfikowania wzorów interakcji, klik oraz dynamiki małych grup. Socjogram przedstawia każdego członka grupy, zaszeregowanego przestrzennie w swoim położeniu względnym wraz z liniami łączącymi, które przedstawiały wzór jego wyborów.

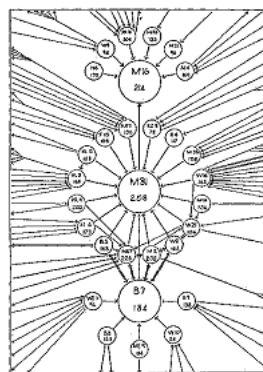
³ Status socjometryczny jest rozumiany jako liczba wyborów, siły wyborów otrzymywane przez każdy punkt. Badacze sądzili, że jest to dobry indeks siły lub wpływu społecznego.

W 1938 r. badacze G.A. Lundberga i M. Steele przedstawili koncepcję wizualizacji sieci społecznej, w której aktorzy zajmują różne pozycje społeczne. Węzeł o największej średnicy ulokowany jest w centrum i stanowi jądro sieci (*nuclei*) – zob. rys. 3. Interpretacja oznacza, że aktorzy społeczni osiągnęli wysoki status socjometryczny. Idea ta została rozwinięta w latach 40. przez M.L. Northwaya, który uogólnił tę zasadę przez ulokowanie węzłów obrazujących wysoki status socjometryczny w centrum socjogramu kołowego.



Rys. 2. Socjogram obrazujący rozkłady sympatii i antypatii w grupie sportowej

Źródło: [Freeman 2000].



Rys. 3. Ilustracja „lady bountiful” G.A. Lundberga i M. Steele

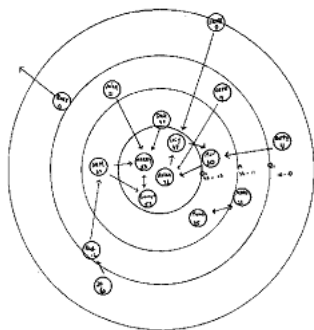
M.L. Northway przykładał również wagę do interpretacji okręgów symbolizujących różne poziomy statusu socjometrycznego (*target sociogram*). Węzły umieszczone w centralnym okręgu przyjmują bowiem centralną pozycję w sieci społecznej⁴. Linie pomiędzy węzłami odzwierciedlają wzajemne relacje. Ilustracja socjogramu kołowego została zamieszczona na rys. 4.

Począwszy od lat 50. badacze zaczęli poszukiwać metod oraz procedur obliczeniowych niezbędnych w procesie rozlokowania położenia węzłów oraz połączeń między nimi. W roku 1952 R.D. Bock i S.Z. Husain, a rok później C. Proctor, uzyskali metodę lokacji węzłów w sieci przy zastosowaniu metody analizy czynnikowej.

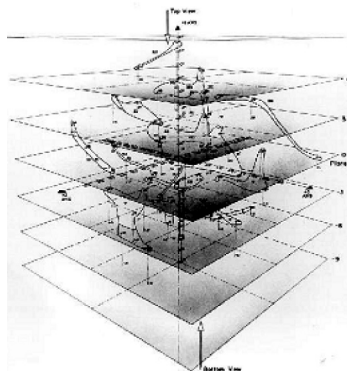
W 1966 r. S. Baumann i L. Guttman wykorzystali metodę skalowania wielowymiarowego (*multidimensional scaling*) w celu rozmieszczenia punktów pochodzących z analizy sieciowej. Jako pierwsi wizualizowali sieć w trzech wymiarach.

⁴ Centralność może być rozumiana jako własność sieci w wymiarze lokalnym lub globalnym. Pozycja jest centralna lokalnie wówczas, gdy istnieje relatywnie dużo powiązań z innymi węzłami. Węzeł jest centralny globalnie wtedy, gdy ma strategiczną pozycję w sieci. Centralność sieci określa się przez liczbę pozycji, z którymi połączona jest dana pozycja, lub też przez liczbę punktów, pomiędzy którymi znajduje się dana pozycja, lub jako bliskość danej pozycji do innych pozycji w obrębie sieci.

Rysunek 5 przedstawia obraz wyników pochodzących z badań nad kategoriami zawodowymi. Z grupy 55 zawodów uzyskano 17 skupisk o podobnym poziomie prestiżu. Zawody cechujące się wyższym prestiżem społecznym umieszczane były na górze wykresu, a stopniowo przechodząc w dół, umieszczano zawody o niższym prestiżu.



Rys. 4. Socjogram kołowy (tarczowy)



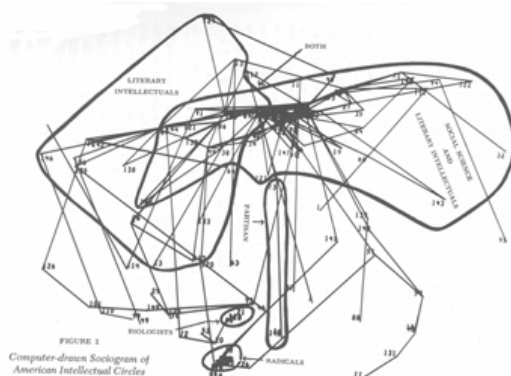
Rys. 5. Ilustracja 17 skupisk grup zawodowych
S. Baumana i L. Guttmana

Źródło: [Freeman 2000].

Lata 70. to okres, w którym następuje rozwój oprogramowania do analizy sieci społecznych. W 1972 r. R.D. Alba we współpracy z M.P. Gutmanem napisali oprogramowanie SOCK, które umożliwia odkrycia ukrytych struktur w sieciach. Wizualizacja sieci pochodząca z programu SOCK została zobrazowana na rys. 6. Następnie R.D. Alba kontynuował prace w tym zakresie, czego efektem jest oprogramowanie COMPLT.

W latach 1975-76 powstają trzy pakiety oprogramowania komputerowego rozwiązujące problem ekwiwalencji strukturalnej pozycji w sieciach: CONCOR (R.L. Breiger, S.A. Boorman i P. Arabie) BLOCKER (G.H. Heil i H.C. White), STRUCTURE (R.S. Burt). W 1975 r. W.D. Richards w NEGOPHY kontynuuje problematykę ukrytych grup w sieciach, a w 1978 r. S.B. Seidman i B.L. Foster przedstawiają SONET, zbiór technik graficznych do wizualizacji relacji. W 1979 r. L.C. Freeman opracowuje program CENTER przeznaczony do określania pozycji centralnych w sieci. W 1981 r. grupa badaczy z Niderlandów rozwija program graficzny GRADAP.

Lata 80. przynoszą intensywny rozwój oprogramowania do wizualizacji danych sieciowych związany z coraz większą powszechnością i dostępnością komputerów osobistych (PC). W tym czasie ukazują się pakiety do analizy danych sieciowych oraz ich wizualizacji. W latach 1993-95 powstają następujące programy: GLAD



Rys. 6. Wydruk z programu SOCK

Źródło: [Internet 1].

(*General Lattice Analysis Design*) napisany przez V. Duquenne'a i Krackplot napisany przez D. Krackhardta, J. Blythe'a i C. McGrath, który jest oprogramowaniem przeznaczonym do wizualizacji sieci. W roku 1996 V. Batagelj i A. Mrvar stworzyli pakiet oprogramowania Pajek, który pozwala na analizy zaawansowane statystycznie, jak również na wizualizację dużych sieci. Przykładowe ilustracje pochodzące z programu Pajek umieszczone są na rys. 7 i 8.

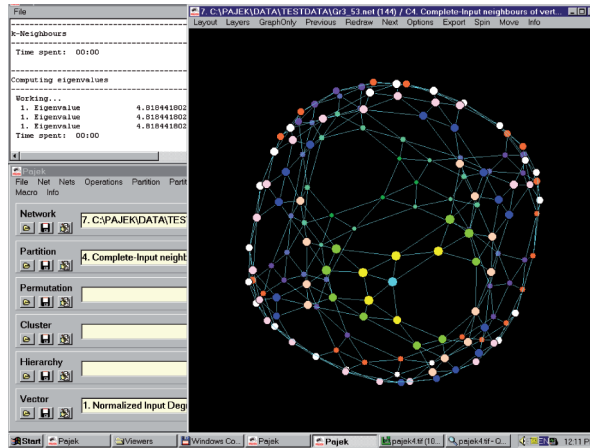
W latach 90. w związku z upowszechnieniem Internetu, a zwłaszcza globalnej sieci WWW, nastąpił kolejny etap rozwoju technik wizualizacji sieci. Stało się to możliwe poprzez wykorzystanie w pakietach oprogramowania komputerowego formatu graficznego rysunków *.gif* i *.jpeg* generowanego automatycznie, zastosowanie aplikacji *Java*, przeznaczonej do analizy i wizualizacji sieci społecznych, oraz użycie VRML (*Virtual Reality Modeling Language*), formatu pliku opisującego trójwymiarową grafikę 3D.

W 1992 r. J.S. Richardson i D.C. Richardson przedstawiają oprogramowanie MAGE przeznaczone wyłącznie do tworzenia graficznej wizualizacji sieci [Freeman 2004].

Współczesne oprogramowanie do analizy sieci społecznych można sklasyfikować według następujących kategorii:

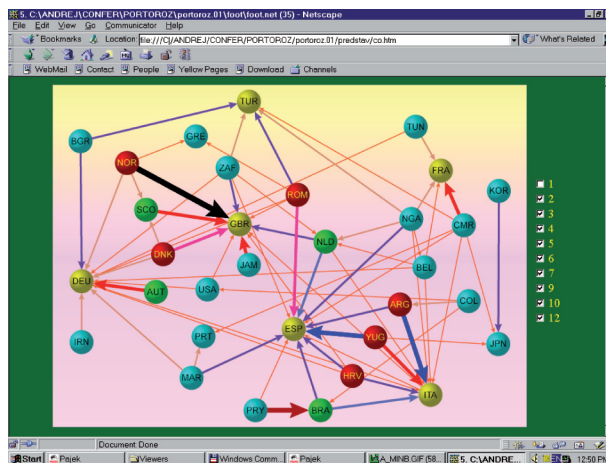
- ogólnego oprogramowania statystycznego z możliwością analizy sieci społecznej,
- specjalistycznego oprogramowania służącego wyłącznie do analizy sieciowej,
- oprogramowania służącego tylko do celów wizualizacji danych sieciowych (relacyjnych).

W pierwszej grupie można wyróżnić bezpłatne pakiety softwarowe, takie jak: Agna, DyNet, GUESS, MultiNet, NetVis, Network Workbench, ORA, Pajek, Sentinel Visualizer, SocNetV, UCINET 6, visione, igraph, JUNG, libSNA, NetworkX,



Rys. 7. Wizualizacja sieci w programie Pajek ver.1.25

Źródło: [Internet 1].



Rys. 8. Wizualizacja sieci w programie Pajek ver.1.25

Źródło: [Internet 1].

NodeXL, SNA, oraz komercyjne pakiety, takie jak: Blue Spider, InFlow, mdlogix solutions, NetMiner3, SNAP, yFiles. Warto również wymienić oprogramowania GRADAP, STRUCTURE, jednakże nie są one poddawane już aktualizacji, ponieważ działają w systemie operacyjnym DOS.

W ramach bezpłatnego specjalistycznego oprogramowania można wymienić programy, takie jak: Blanche, CFinder, CiteSpace, Commetrix, E-Net, EgoNet, Iknow, KeyPlayer, KliqFinder, Network Genie, PGRAPH, PNet, Puck, ReferralWeb,

SONIVIS, StOCNET, Unison, statnet suite, tnet, UrlNet. W grupie oprogramowania płatnego można wspomnieć o: MetaSight, Network Genie, ONA surveys, MatMan. Programy, takie jak: FATCAT, NEGOPHY, PermNet, Snowball, nie są już aktualizowane z powodu pracy w środowisku DOS.

Bezpłatne oprogramowania przeznaczone wyłącznie do wizualizacji danych sieciowych to: aiSee, Apacze Agora, Cytoscape, Graphviz, Jacob's Ladder, KracPlot, Mage, NetDraw, OGDF, Otter, SoNIA, Tulip, uDraw(Graph), Zoomgraph. Natomiast pakiety komercyjne to: KeyHubs i TouchGraph [Huisman, Duijn 2005].

3. Podsumowanie

Począwszy od lat 30., następuje rozwój metod wizualizacji danych sieciowych. Ewolucja ta przebiega od technik statycznych, ręcznych, mechanicznych do metod dynamicznych wielopoziomowych, algorytmicznych, komputerowych.

Analizując oprogramowanie do analizy SNA, można zauważyć, że coraz silniej kładzie się nacisk na integrację pomiędzy równorzędną jakością prezentacji związanej z grafiką, trójwymiarowością 3D, interaktywnością a uwzględnieniem w programach własności wskaźników statystycznych analizowanych związków. Niewątpliwie wizualizacja od początku jest silnie związana z analizą SNA, a sama umożliwia w analizie sieci odkrycie nowych, emergentnych, ponadjednostkowych, strukturalnych własności obiektów i relacji między nimi.

Warto podkreślić, że podejście takie pozwoli na przełamanie ograniczeń w badaniach relacji zogniskowanych jedynie na pomiarze oceny relacji z punktu widzenia tylko jednej ze stron.

Literatura

- Arabie P., Wind Y., *Marketing and Social Networks*, [w:] *Advances in Social Network Analysis*, S. Wasserman, J. Galaskiewicz (red.), Sage, Thousand Oaks, CA, 1994.
- Barnes J.A., *Class and Committee in a Norwegian Island Parish*, „Human Relations” 1954 nr 7, s. 39-58.
- Burt R.S., *Position in Networks*, „Social Forces” 1976 nr 55, s. 93-122.
- Burt R.S., *Structural Holes: The Social Structure of Competition*, MA: Harvard University Press, Cambridge 1992.
- Cadeaux J., *A closer look at the Interface between the Product Lines of Manufacturers and the Assortments of Retailers*, „International Journal of Retail and Distribution Management” 1997 nr 25, s. 197-203.
- Cook K.S., Emerson R.M., *Power, Equity, and Commitment in Exchange Networks*, „American Sociological Review” 1978 nr 43, s. 721-739.
- Cook K.S., *Network Structures from an Exchange Perspective*, [w:] *Social Structure and Network Analysis*, P.V. Marsden, Lin N. (red.), Sage, Newbury Park, CA, 1982, s. 177-199.
- Emerson R.M., *Social Exchange Theory*, [w:] *Social Psychology: Sociological Perspectives*, M. Rosenberg, R.H. Turner (red.), Basic Books, Inc, New York 1981, s. 30-65.

- Ford L.R. Jr., Fulkerson D.R., *Przepływy w sieciach*, PWN, Warszawa 1969.
- Freeman L.C., *Visualizing Social Networks*, „Journal of Social Structure” 1, 1 (February 4, 2000).
- Freeman L.C., *The Development of Social Network Analysis. A Study in the Sociology of Science*, Empirical Press, Vancouver 2004.
- Granovetter M.S., *Economic Action and Social Structure: the Problem of Embeddedness*, „American Journal of Sociology” 1985 nr 91, s. 481-510.
- Hakansson H., Sharma D.D., *Strategic Alliances in a Network Perspective*, [w:] *Networks in Marketing*, D. Iacobucci (red.), Sage Publications, 1996.
- Huisman M., van Duijn M.A.J., *Software for Social Network Analysis*, [w:] *Models and Methods in Social Network Analysis*, P.J. Carrington, J. Scott, S. Wasserman (red.), Cambridge University Press, New York 2005, s. 270-316.
- Iacobucci D., Hopkins N., *Modeling Dyadic Interactions and Networks in Marketing*, „Journal of Marketing Research” 1992 nr 29, s. 5-17.
- Iacobucci D., Henderson G., Marcati A., Chang J., *Network Analyses of Brand Switching Behavior*, „International Journal of Research in Marketing” 1996 nr 13, s. 415-429.
- Iacobucci D., Henderson G., Marcati A., Chang Jennifer E., *Network Analyses and Consumer Brand Switching Behavior: the Ehrenberg Automobile Data*, [w:] *Networks in Marketing*, D. Iacobucci (red.), Thousand Oaks, CA: Sage Publications, Inc. 1996.
- Korzan B., *Elementy teorii grafów i sieci metody i zastosowania*, Wydawnictwa Naukowo-Techniczne, Warszawa 1978.
- Martin C.L., Clark T., *Networks of Customer-to-Customer Relationships in Marketing*, [w:] *Networks in Marketing*, D. Iacobucci (red.), Sage, London 1996, s. 342-366.
- Moody J., McFarland D., Bender-deMoll S., *Dynamic Network Visualization*, „American Journal of Sociology” nr 110, January, 2005, s. 1206-41.
- Newman M.E.J., Barabasi A.L., Watts D.J. (red.), *The Structure and Dynamics of Networks*, Princeton University Press, 2006.
- Rogers E.M., *Diffusion of Innovations*, The Free Press, New York 1962.
- Rogers E.M., *New Product Adoption and Diffusion*, „Journal of Consumer Research”, tom 2, nr 3, 1976, s. 290-301.
- Wellman B., *Structural Analysis from Method and Metaphor to Theory and Substance*, [w:] *Social Structures: A Network Approach*, B. Wellman, S.D. Berkowitz (red.), Cambridge 1988.

Źródła internetowe

[1] <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>.

METHODS OF NETWORK DATA VISUALIZATION

Summary

In social network analysis, graphic methods of visualization are of particular significance - both of relations among interacting objects, and of formal attributes of networks and sub-networks formed on the basis of those interactions.

The paper presents methods of network visualization in the retrospective picture, as well as problems of contemporary packages intended for SNA analysis.