SOCIAL NETWORK ANALYSIS

Wouter de Nooy
University of Amsterdam, Netherlands

Keywords: Social networks, structuralism, cohesion, brokerage, stratification, network analysis, methods, graph theory, statistical models

Contents

Summary

Social network analysis (SNA) focuses on the structure of ties within a set of social actors, e.g., persons, groups, organizations, and nations, or the products of human activity or cognition such as web sites, semantic concepts, and so on. It is linked to structuralism in sociology stressing the significance of relations among social actors to their behavior, opinions, and attitudes. Social network analysis is felt to be appropriate for analyzing social cohesion, brokerage and exchange, as well as social ranking within or among social groups.

Two perspectives dominate SNA: the socio-centered and ego-centered perspective. The socio-centered perspective analyses overall network structure. It looks for patterns of ties that indicate cohesive social groups, central actors that may be paramount to the integration of the social network, and asymmetries that may reflect social prestige or social stratification. Recent advances are found primarily in the technique of blockmodelling. The ego-centered perspective focuses on the composition of local network structure. Do actors influence one another through their network ties (social influence model) and/or do actors adjust their ties to the characteristics of their peers and to their ties with them (social selection model)? Recent advances in this area include new types of statistical models.

The development and interest in SNA has increased sharply over the last few decades due to the application of mathematics – notably graph theory and statistical models – and the wide availability of software for network analysis - both commercial and freely available through the internet. In addition to the formal, quantitative approach to social network analysis, a qualitative approach to social networks is developing.

1. Introduction

Social network analysis (SNA) focuses on the structure of ties within a set of social actors, e.g., persons, groups, organizations, and nations, or the products of human activity or cognition such as web sites, semantic concepts, and so on. Any social process or system that can be conceptualized as a set of units and a set of lines connecting pairs of units can be studied as a social network. Examples of social structures that have been studied as networks are friendship among children in a school, family relations among members of a social elite, shared board members of corporations, trade relations between countries, and hyperlinks between websites.

This is not to say that it is always useful or necessary to apply social network analysis to data that can be conceptualized as networks. For instance, if a researcher is just interested in knowing the number of people that a person can turn to for help, the number of ties instead of the structure of ties is relevant and network analysis is not needed. For network analysis to be applicable, theory from sociology or other social and behavioral sciences should give reasons to believe that the structure of ties is linked to behavior, opinions, or social position of the members of the network. Three types of sociological concepts appear repeatedly in most applications of social network analysis: cohesion or solidarity, brokerage or influence, and ranking, prestige or status. They are discussed in Sections **Error! Reference source not found.**, **Error! Reference source not found.**, and **Error! Reference source not found.** respectively.

2. Definition of a network

There are several ways of formally defining a network, depending on the branch of mathematics used. The most usual and flexible definition is derived from graph theory, which conceptualizes a social network is conceptualized as a graph, that is, a set of vertices (or nodes, units, points) representing social actors or objects and a set of lines representing one or more social relations among them.

A network, however, is more than a graph because it contains additional information on the vertices and lines. Characteristics of the social actors, for instance a person's sex, age, or income, are represented by discrete or continuous attributes of the vertices in the network, and the intensity, frequency, valence, or type of social relation are represented by line weights, line values, line signs, or line type.

Formally, a network $\mathbf{N}$ can be defined as $\mathbf{N} = (U, L, F_U, F_L)$ containing a graph $\mathbf{G} = (U, L)$, which is an ordered pair of a unit or vertex set U and a line set L, extended with a function F_U specifying a vector of properties of the units ($f : U \rightarrow X$) and a function F_L specifying a vector of properties of the lines ($f : L \rightarrow Y$). The set of lines L may be regarded as the union of a set of undirected edges E and a set of directed arcs A

($L = E \cup A$). Each element *e* of *E* (each edge) is an unordered pair of units *u* and *v* (vertices) from *U*, that is, *e* (*u*: *v*), and each element *a* of *A* (each arc) is an ordered pair of units *u* and *v* (vertices) from *U*, that is, *a*(*u*: *v*).

Figure 1.illustrates the graph theoretical definition of a network. It shows a sociogram of marriage and business relations among 16 families in Florence (Italy) around 1400 AD. The data are part of a larger dataset collected and analyzed by John F. Padgett. In a sociogram, vertices are represented by circles, so each family is represented by a circle. Lines, connecting two circles in the sociogram, can be of two types: directed or undirected. Here, directed lines represent business relations among families. They are drawn as black arcs in Figure 1, pointing towards the more prosperous family. Marriage relations between families are undirected lines or edges in the network. They are represented by grey lines in the sociogram.

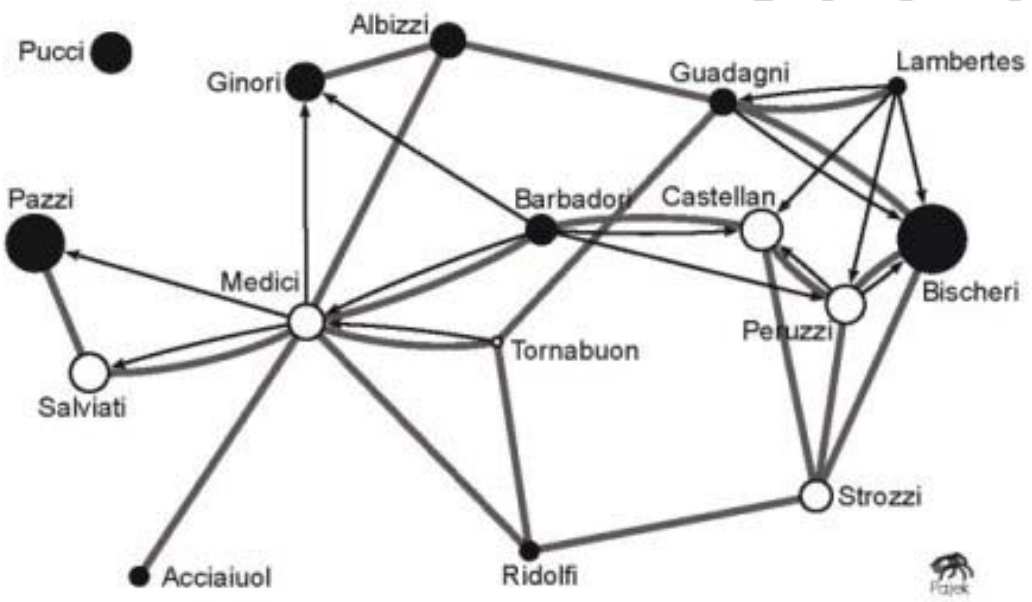

Figure 1 - Sociogram of marriage (grey edges) and business (black arcs) relations between 16 Florentine families (circa 1400 AD).

As stated before, a network is more than a graph, it contains additional information on the vertices or the lines. In the sociogram, the family names link the vertices to historical people. This information goes beyond the pure structure of the network, which is represented by the graph. In addition, the sociogram shows two properties of the families: (1) whether the family had seats in the local civic council in a previous period (AD 1282-1344) and (2) the net wealth of the family in 1427 (measured in thousands of lira). Vertex color shows the first property: black vertices represent families that used to be members of the civic counsel. Vertex size presents the wealth of the family: the bigger the circle, the more affluent the family in 1427. Finally, if information had been available on the number of marriages or business transactions between families, the network would also have included additional information on the lines. Then the width or color of the line would be used to visualize properties of lines.

The network of Florentine families contains two different relations. As a consequence, two families are sometimes connected by two lines: both a marriage edge and a business arc. Loops do not occur, although it is not impossible that members of the same family trade among themselves. In graph theory, a simple graph is a graph without multiple lines and, in the case of an undirected graph, without loops. A simple graph can easily be stored in a matrix, which is called an adjacency matrix or sociomatrix. So if we split the two relations in this network, we would obtain two simple graphs that can be represented by sociomatrices (Figure 2).

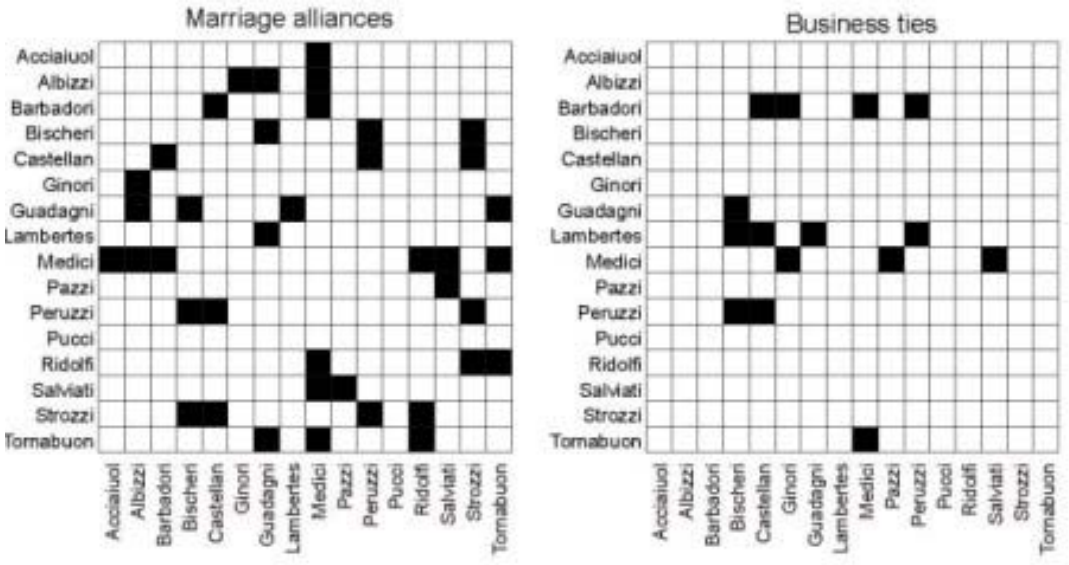

Figure 2 - Sociomatrices: marriage ties and business ties.

In the sociomatrices of the Florentine families network, the families define both the rows and columns. By convention, the entity in the row is the sender or tail of an arc and the entity in the column is the receiver or head of the arc. In these matrices, a black cell indicates the presence of a line. Black cells indicate the presence of lines, white cells the absence. Presence and absence re usually coded as 0 and 1 respectively, so the matrix may also contain numbers. We can see that the Acciaiuol family (1st row) is married to the Medici family (9th column) because the corresponding cell is black. Because the marriage relation is undirected, the latter family is also married to the former family, so the cell in the 9th row and 1st column is also black. The sociomatrix of an undirected relation (edges) is symmetric with respect to the diagonal running from the top left to the bottom right of the matrix. This is not the case with an asymmetric relation (arcs) such as the business ties.

In the example network, every family can be related to every other family by marriage or business ties. This is called a 1-mode network. There are many social relations, however, for which this is not the case. For instance, if persons are the vertices in the network, a marriage relation can only connect a man and a woman in many cultures. Then, we can split the vertices into two groups (men and women) such that ties can only

exist between the two groups, not within one group. These groups are called modes and a network with two modes is called a bipartite or 2-mode network. A 2-mode network without multiple lines can be represented by a rectangular matrix in which the vertices of the first mode are found in the rows and the columns contain the vertices of the second mode. A very popular example of a 2-mode network is a network of affiliations between people and organizations, e.g., people (first mode) who are board members of large corporations (second mode).

3. Two perspectives: overall and local network structure

The conceptualization of social systems as graphs and networks offers the opportunity for systematic investigation and theorizing on the structure of ties among social actors beyond the pair. Whereas classical sociology tended to make a quantum leap from the individual and the pair to the triple, group, or society, e.g., Georg Simmel, graph theory offers the tools to formally describe and visualize social structure consisting of three and more actors. This has led to a new awareness of social structure as a system of ties that is both the product and the context and condition for human action.

Scientists that regard social structure as the product of ties and interaction between persons or other social objects are primarily interested in examining the overall network structure. What is the structure of ties within a social group, among social groups, within or among organizations, etcetera? Network analysis offers tools for describing overall network structure, disclosing, for instance, cohesive subgroups and ranked layers, or bottlenecks in exchange networks. This approach to social networks is known as the socio-centered approach. The substantive interest, then, is to find out whether cohesive subgroups identified in the network actually represent communities and whether ranked layers identify social strata. In a constructivist perspective, it may even be hypothesized that social identities and classifications are derived from network structure, e.g. groups are labeled as different because they occupy clearly different positions within a network of ties.

The other approach to social networks focuses on the individual actor and its immediate network neighborhood. This is known as the ego-centered approach, which is currently being developed as the actor-oriented approach (see Section 7.2). It is assumed that social behavior is orchestrated: actors adjust their behavior and attitudes, opinions, and beliefs to the behavior (etc.) of other members of the social system in which they participate. This is the social influence model of networks. As a network of ties, the system defines to whom an actor is exposed. The immediate contacts – the neighbors in graph theory – of an actor are usually most salient to its behavior, but indirect contacts such as their neighbors' neighbors may be taken into account as well. In other words, an actor's local context or ego-network is likely to affect its behavior.

At the same time, however, actors decide on which ties to establish, maintain, or end. This has been labeled the social selection model. Properties of alters usually play a role here, e.g., a preference for interaction with people that are similar to you, known as the homophily principle. Local network structure may also affect the creation or dissolution of ties, e.g., the often encountered phenomenon that people tend to become friends of their friends' friends. By ending ties or creating new ones, the individual changes both

local network structure and overall network structure, that is, the system. Overall network structure, then, is regarded as the outcome of individual action. To the actors, the change of network structure is not necessarily predictable, so the interplay between individual action and network structure may offer surprising results.

The subsequent sections present the three main theoretical approaches to social networks: cohesion, brokerage, and ranking. Whenever applicable, both the socio-centered and ego-centered approaches are discussed.

-
-
-

Bibliography

Barabási A.-L. (2002). *Linked: The New Science of Network*. Cambridge, MA: Perseus Publishing. [The seminal book on scale free networks and the power law.]

Carrington P.J., Scott J., Wasserman S. (2005). *Models and Methods in Social Network Analysis*, 328 pp. New York, USA: Cambridge University Press. [Overview of recent developments in social network analysis and software overview.]

De Nooy W., Mrvar A., Batagelj V. (2005). *Exploratory Social Network Analysis with Pajek*, 334 pp. New York, USA: Cambridge University Press. [This handbook offers an overview of network ideas and methods and guidance for doing network analysis with the free software package Pajek.]

Degenne A., Forsé M. (1999). *Introducing Social Networks*, 248 pp. London: Sage Publications. [Overview of social network analysis in the context of sociological theory.]

Doreian P., Batagelj V., & Ferligoj A. (2005). *Generalized Blockmodeling*, 384 pp. New York: Cambridge University Press. [Presents the latest developments in blockmodelling.]

Freeman L.C. (2004). *The Development of Social Network Analysis*, 205 pp. Vancouver: Empirical Press. [A history of social network analysis as a scientific endeavor.]

Huberman, Michael and Matthew Miles (2002) *The Qualitative Researcher's Companion*. Sage Publications: London {Handbook for Qualitative Research.]

Macy M.W., Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology*, 28, 143-166. [Describes the application of agent-based models to social network analysis.]

Pajek Large Network Analysis Program discussion website http://pajek.imfm.si/doku.php [Information, examples and discussion of network analyses based on one particular network analyis package.]

Robins G., Snijders T., Wang P., Handcock M., Pattison P. (2007). Recent developments in exponential random graph (p*) models for social networks. *Social Networks* 29, 192-215. [Special issue on new developments in statistical models for network analysis.]

Scott J. (1991, 2000[2]). *Social Network Analysis: A Handbook*, 208 pp. London: Sage Publications. [Overview of social network analysis at an introductory level.]

Wasserman S., Faust K. (1994). *Social Network Analysis: Methods and Applications*, 825 pp. New York,

USA: Cambridge University Press. [The standard handbook for social network analysts.]

Biographical Sketch

Wouter de Nooy (1962) is associate professor of research methods at the Department of Communication Science at the University of Amsterdam, The Netherlands. He studied language and literature and obtained his PhD in 1993. In his research, he has specialized in applying and developing social network analysis. In collaboration with Andrej Mrvar and Vladimir Batagelj (University of Ljubljana, Slovenia), he authored a handbook on social network analysis, which appeared at Cambridge University Press in 2005. A Japanese translation of the book is planned to appear in 2008. His applications of social network analysis focus on cultural fields (literature, visual arts) and policy networks