ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Social Network Analysis by Using Graphs to Represent Social Relations

Vandana Janjirala
M.Tech,
Department of CSE,
AVNIET, JNTUH,
Hyderabad.

SK Abdul Nabi
Professor and HOD,
Department of CSE,
AVNIET, JNTUH,
Hyderabad.

Abstract:

When a computer network connects people or organizations, it is a social network. Yet the study of such computer-supported social networks has not received as much attention as studies of human-computer interaction, online person-to-person interaction, and computer-supported communication within small groups. We argue the usefulness of a social network approach for the study of computer-mediated communication.

We review some basic concepts of social network analysis, describe how to collect and analyze social network data, and demonstrate where social network data can be, and have been, used to study computer-mediated communication. Throughout, we show the utility of the social network approach for studying computer-mediated communication, be it in computer-supported cooperative work, in virtual community, or in more diffuse interactions over less bounded systems such as the Internet.

A social network is a set of people (or organizations or other social entities) connected by a set of social relationships,such as friendship, co-working or information exchange.Social network analysis focuses on the analysis of patterns of relationships among people, organizations, states and such social entities. Social network analysis provides both a visual and amathematical analysis of human relationships. Web can alsobe considered as a social network. Social networks are formed between Web pages by hyperlinking to other Web pages. In this paper a state of the art survey of the works done on social network analysis ranging from pure mathematical analyses in graphs to analyzing the social networks in Semantic Web is given. The main goal is to provide a road map for researchers working on different aspects of Social Network Analysis.

Keywords:

Pattern analysis,Semantic Web,Social network services,Web pages.

INTRODUCTION:

N recent years, online social networks has become very popular, many web sites have sprung up where one can meet their offline friends in virtual world of the internet. Services like Facebook, Orkut, and MySpace etc allow people to host their online social networks, people create their profiles in such social networks and share this information with their friends and a vast amount of strangers on these social sites. And the concern about users' privacy that can be used for different purposes such as identity theft, or for advertisement markets and other negative intentions and the threat is more likely to be for those developing countries that have a young society and can be the main target for these concerns.

A social network is a social structure between actors, mostly individuals or organizations. It indicates the ways in which they are connected through various social familiarities ranging from casual acquaintance to close familiar bonds. Email traffic, disease transmission, and criminal activity can all be modeled as social networks. Socialnetwork analysis is the mapping and measuring of relationships and flows between people, groups, organizations,animals, computers or other information/knowledge processing entities. The nodes in the network are the people and groups, while the links show relationships or flows between the nodes. Social network analysis provides both a visual and a mathematical analysis of human relationships. Management consultants use this methodology with their business clients and call it Network Analysis.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Social network data consist of various elements. Following the definition by Wasserman and , social network data can be viewed as a social relational system characterized by a set of actors and their social ties. Additional information in the form of actor attribute variables or multiple relations can be part of the social relational system.

What is Social Network Analysis?

The Social Network Approach:

When a computer network connects people or organizations, it is a social network. Just as a computer network is a set of machines connected by a set of cables, a social network is a set of people (or organizations or other social entities) connected by a set of social relationships, such as friendship, co-working or information exchange. Much research into how people use computer-mediated communication (CMC) has concentrated on how individual users interface with their computers, how two persons interact online, or how small groups function online.

As widespread communication via computer networks develops, analysts need to go beyond studying single users, two-person ties, and small groups to examining the computer-supported social networks (CSSNs) that flourish in areas as diverse as the workplace and virtual communities. This paper describes the use of the social network approach for understanding the interplay between computer networks, CMC, and social processes.

Social network analysis focuses on patterns of relations among people, organizations, states, etc. This research approach has rapidly developed in the past twenty years, principally in sociology and communication science. The International Network for Social Network Analysis (INSNA) is a multidisciplinary scholarly organization, which publishes a refereed journal, Social Networks, and an informal journal, Connections.

Social network analysts seek to describe networks of relations as fully as possible, tease out the prominent patterns in such networks, trace the flow of information (and other resources) through them, and discover what effects these relations and networks have on people and organizations.

They treat the description of relational patterns as interesting in its own right -- e.g., is there a core and periphery? -- and examine how involvement in such social networks helps to explain the behavior and attitudes of network members -- e.g., do peripheral people send more email and do they feel more involved? They use a variety of techniques to discover a network's densely-knit clusters and to look for similar role relations. When social network analysts study two-person ties, they interpret their functioning in the light of the two persons' relations with other network members. This is a quite different approach than the standard CMC assumption that relations can be studied as totally separate units of analysis. "To discover how A, who is in touch with B and C, is affected by the relation between B and C . . . demands the use of the [social] network concept".

There are times when the social network itself is the focus of attention. If we term network members egos and alters, then each tie not only gives egos direct access to their alters but also indirect access to all those network members to whom their alters are connected. Indirect ties link in compound relations (e.g., friend of a friend) that fit network members into larger social systems The social network approach facilitates the study of how information flows through direct and indirect network ties, how people acquire resources, and how coalitions and cleavages operate.

Although a good deal of CMC research has investigated group interaction online, a group is only one kind of social network, one that is tightly-bound and densely-knit. Not all relations fit neatly into tightly-bounded solidarities. Indeed, limiting descriptions to groups and hierarchies oversimplifies the complex social networks that computer networks support. If Novell had not trademarked it already, we would more properly speak of "netware" and not "groupware" to describe the software, hardware, and peopleware combination that supports computer-mediated communication.

Comparisons with Other Approaches to the Study of CMC:

Much CMC research concentrates on how the technical attributes of different communication media might affect what can be conveyed via each medium. These characteristics include the richness of cues a medium conveys (for example, whether a medium conveys

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

text, or whether it includes visual and auditory cues), the visibility or anonymity of the participants (e.g., video-mail versus voice mail; whether communications identify the sender by name, gender, title), and the timing of exchanges (e.g., synchronous or asynchronous communication). A reduction in cues has been cited as responsible for uninhibited exchanges (e.g., flaming), more egalitarian participation across gender and status, increased participation of peripheral workers, decreased status effects and lengthier decision processes.

Studies of group communication are somewhat closer to the social network approach because they recognize that the use of CMC is subject to group and organizational The group communication approach includes CMC theories such as social influence, social information processing, symbolic interactionism, critical mass, and adaptive structuration. These theoretical approaches recognize that group norms contribute to the development of a critical mass and influence the particular form of local usage. Yet this focus on the group leads analysts away from some of the most powerful social implications of CMC in computer networks: its potential to support interaction in unbounded, sparsely-knit social networks .

SOCIAL NETWORK MODELS:

Using formal methods to show Social Networks One reason for using mathematical and graphical techniques in social network analysis is to represent the descriptions of networks compactly and systematically. A related reason for using (particularly mathematical) formal methods for representing social networks is that mathematical representations allow us to apply computers to the analysis of network data. The third, and final reason for using "formal" methods (mathematics and graphs) for representing social network data is that the techniques for graph processing and the rules of mathematics themselves suggest things that we might look for in our data. In the analysis of complete networks, a distinction can be made between.

- descriptive methods, also through graphical representations.
- analysis procedures, often based on a decomposition of the adjacency matrix.
- statistical models based on probability distributions.

Using Graphs to Represent Social Relations:

Network analysis uses (primarily) one kind of graphic display that consists of points (or nodes) to represent actors and lines (or edges) to represent ties or relations. When sociologists borrowed this way of graphing things from the mathematicians, they renamed their graphs as "sociograms". There are a number of variations on the theme of sociograms, but they all share the common feature of using a labeled circle for each actor in the population we are describing, and line segments between pairs of actors to represent the observation that a tie exists between the two. Visualization by displaying a sociogram as well as a summary.

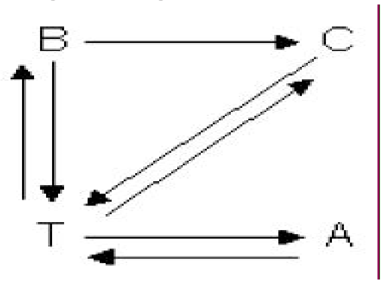

Fig. Using Graphs to Represent Social Relations

of graph theoretical concepts provides a first description of social network data. For a small graph this may suffice, but usually the data and/or research questions are too complex for this relatively simple approach.

Using Matrices to Represent Social Relations:

The most common form of matrix in social network analysis is a very simple one composed of as many rows and columns as there are actors in our data set, and where the elements represent the ties between the actors. The simplest and most common matrix is binary. That is, if a tie is present, a one is entered in a cell; if there is no tie, a zero is entered. This kind of a matrix is the starting point for almost all network analysis, and is called an "adjacency matrix" because it represents who is next to, or adjacent to whom in the "social space" mapped by the relations that we have measured. By convention, in a directed graph, the sender of a tie is the row and the target of the tie is the column. Let's look at a simple example. The directed graph of friendship choices among Bob, Carol, Ted, and Alice looks like figure 1. Since the ties are measured at the nominal level (that is, the data are binary choice data), we can represent the same information in a matrix that looks like

	Bob	Carol	Ted	Alice
Bob	----	1	0	0
Carol	1	----	1	0
Ted	1	1	----	1
Alice	0	0	1	----

Fig.. Using Matrices to Represent Social Relations

Statistical Models for Social Network Analysis:

Statistical analysis of social networks spans over 60 years. Since the 1970s, one of the major directions in the field was to model probabilities of relational ties between interacting units (social actors), though in the beginning only very small groups of actors were considered. Extensive introduction to earlier methods is provided by Wasserman and Faust . Two of the most prominent current directions are Markov Random Fields (MRFs) introduced by Frank and Strauss and Exponential Random Graphical Models (ERGMs), also known as p! . The ERGM have been recently extended by Snijders et al in order to achieve robustness in the estimated parameters.

The statistical literature on modeling Social Networks assumes that there are n entities called actors and information about binary relations between them. Binary relations are represented as an $n \times n$ matrix Y , where Yij is 1, if actor I is somehow related to j and 0 otherwise. For example, Yij = 1 if i considers j to be friend. The entities are usually represented as nodes and the relations as arrows between the nodes. If matrix Y is symmetric, then the relations are represented as undirected arrows. More generally Yij can be valued and not just binary, representing the strength (or value) of the relationship between actors i and j .There are several useful properties of the stochastic models.

Some of them are:

• The ability to explain important properties between entities that often occur in real life such as reciprocity, if i is related to j then j is more likely to be somehow related to i; and transitivity, if i knows j and j knows k, it is likely that i knows k.

• Inference methods for handling systematic errors in the measurement of links .

• General approaches for parameter estimation and model comparison using Markov Chain Monte Carlo methods.

• Taking into account individual variability and properties of actors.

• Ability to handle groups of nodes with equivalent statistical properties

SOCIAL NETWORK PROPERTIES:

There are some properties of social networks that are very important such as size, density, degree, reachability, distance, diameter, geodesic distance. Here we describe some more complicated properties which may be used in social network analysis. The following properties are taken from

Maximum flow:

One notion of how totally connected two actors are, asks how many different actors in the neighborhood of a source lead to pathways to a target. If I need to get a message to you, and there is only one other person to whom I can send this for retransmission, my connection is weak - even if the person I send it to may have many ways of reaching you.

If, on the other hand, there are four people to whom I can send my message, each of whom has one or more ways of retransmitting my message to you, then my connection is stronger. This ”flow” approach suggests that the strength of my tie to you is no stronger than the weakest link in the chain of connections, where weakness means a lack of alternatives.

Centrality and Power:

All sociologists would agree that power is a fundamental property of social structures. There is much less agreement about what power is, and how we can describe and analyze

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Power Aspect Name	Definition	Influences
Degree	Number of ties for an actor	Having more opportunities and alternatives
Closeness	Length of paths to other actors	Direct bargaining and exchange with other actors
Between's	Lying between each other pairs of actors	Brokering contacts among actors to isolate them or prevent connections.

GROUPS AND SUBSTRUCTURES IN SOCIAL NETWORKS:

One of the most common interests of structural analysts is in the "sub-structures" that may be present in a network. Many of the approaches to understanding the structure of a network emphasize how dense connections are compounded and extended to develop larger cliques or sub-groupings. Network analysts have developed a number of useful definitions for algorithms that identify how larger structures are compounded from smaller ones. Divisions of actors into cliques or "sub-groups" can be a very important aspect of social structure.

It can be important in understanding how the network as a whole is likely to behave. For example, suppose the actors in one network form two non overlapping cliques; and, suppose that the actors in another network also form two cliques, but that the memberships overlap (some people are members of both cliques). Where the groups overlap, we might expect that conflict between them is less likely than when the groups don't overlap. Where the groups overlap, mobilization and diffusion may spread rapidly across the entire network; where the groups don't overlap, traits may occur in one group and not diffuse to the other. The main features of a graph, in terms of its cliques or sub-graphs, may be apparent from inspection:

• How separate are the sub-graphs (do they overlap and share members, or do they divide or factionalize the network)?

• How large are the connected sub-graphs? Are there a few big groups, or a larger number of small groups?

• Are there particular actors that appear to play network roles? For example, act as nodes that connect the graph, or who are isolated from groups?

SEMANTIC WEB AND SOCIAL NETWORKS:

There's a revolution occurring and it's all about making the Web meaningful, understandable, and machine-processable, whether it's based in an intranet, extranet, or Internet. This is called the Semantic Web, and it will transition us toward a knowledge-centric viewpoint of 'everything'. The Semantic Web (SW) is an emerging concept that launches the idea of having data on the Web defined and linked in a way that it can be used by people and processed by machines in a "wide variety of new and exciting applications" .

It develops "languages for expressing information in a machine process able form" , so to enable the machine to be able to participate and help inside the information space.The Semantic Web and social network models support one another. On one hand, the Semantic Web enables online and explicitly represented social information; on the other hand, social networks, especially trust networks , provide a new paradigm for knowledge management in which users "outsource" knowledge and beliefs via their social networks . In order to turn these objectives into reality, many challenging issues need to be addressed as the following.

• Knowledge representation:Although various ontologies capture the rich social concepts, there is no need to have hundreds of "dialectic" ontologies defining the same concept. How can we move toward having a small number of common and comprehensive ontologies?

• Knowledge management: The Semantic Web is, relative the entire Web, fairly connected at the RDF graph level but poorly connected at the RDF document level. The open and distributed nature of the Semantic Web also introduces issues. How do we provide efficient and effective mechanisms for accessing knowledge, especially social networks, on the Semantic Web?

• Social network extraction: integration and analysis.
Even with well-defined ontologies for social concepts, extracting social networks correctly from the noisy and incomplete knowledge on the (Semantic) Web is very difficult.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

What are the heuristics for integrating and fusing social information and the metrics for the credibility and utility of the results?

• Provenance and trust aware distributed inference: Provenance associates facts with social entities which are inter-connected in social network, and trust among social entities can be derived from social networks. How to manage and reduce the complexity of distributed inference by utilizing provenance of knowledge in the context of a given trust model?

Despite their early popularity, users have later discovered a number of drawbacks to centralized social networking services. First, the information is under the control of the database owner who has an interest in keeping the information bound to the site. The profiles stored in these systems cannot be exported in machine process able formats, and therefore the data cannot be transferred from one system to the next. Second, centralized systems do not allow users to control the information they provide on their own terms. These problems have been addressed with the use of Semantic Web technology.

The friend-of-a-friend(FOAF) project3 is a first attempt at a formal, machine process able representation of user profiles and friendship networks. show that the Friend of a Friend (FOAF) ontology is among the most used semantic Web ontologies.3http://www.foaf-project.org/The woogle Ontology Dictionary shows that the class foaf:Person4 currently has nearly one million instances spread over about 45,000 Web documents. The FOAF ontology is not the only one used to publish social information on the Web. For example, Swoogle identifies more than 360 RDFS or OWL classes defined with the local name "person". Extracting social network from noisy, real world data is a challenging task, even if the information is already encoded in RDF using well defined ontologies. The process consists of three steps: discovering instances of foaf:Person, merging information about unique individuals, and linking person through various social relation properties.

CONCLUSIONS:

In this paper we've reviewed social networks, formal methods to show them, and social networks' properties.

Social network analysis methods provide some useful tools for addressing many aspects of social structure. The Web itself can be considered as a social network. In the Web's social network, documens are node of the sociogram and links between documents are the edges of the sociogram.Weblogs, which are a special subset of Web could also be considered as social networks. We have described special link structure for Weblogs which contains comments other than explicit links. The Semantic Web (SW) is an emerging concept that launches the idea of having data on the Web defined and linked in a way that it can be used by people and processed by machines.The Semantic Web and social network models support one another. Basic properties of kinds of social networks described in this paper, and shows differences in their formations. As future works, we intend to mine the social networks of Persian Weblogs using the methods surveyed in this paper and find new interesting models. Also we're going to use semantics of those Weblogs and their link structure (their social network) to cluster the Weblogs using Semantic Web concepts.

REFERENCES:

[1] A. Hanneman and M. Riddle, "Introduction to social network methods," online at http://www.faculty.ucr.edu/ hanneman/nettext/, 2005.

[2] S. Wasserman and K. Faust, Social Network Analysis: Methods and Applications. Cambridge University Press, Novemver 1994.

[3] R. L. Breiger, The Analysis of Social Networks. London: Sage Publications Ltd, 2004, pages 505-526 in Handbook of Data Analysis, edited by Melissa Hardy and Alan Bryman.

[4] T. Berners-Lee, "Semantic web road map," online at http://www.w3.org/DesignIssues/Semantic.html, 1998.

[5] A. Swartz and J. Hendler, "The semantic web: A network of content for the digital city," in Proceedings of Second Annual Digital Cities Workshop, Kyoto, Japan, October 2001.

[6] J. Hendler, T. Berners-Lee, and E. Miller, "Integrating applications on the semantic web," Journal of the Institute of Electrical Engineers of Japan, vol. 122, no. 10, pp. 676–680, 2002.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

[7] L. Freeman, Graphic techniques for exploring social network data. New York: Cambridge University Press, 2005, pages 248-269 in 'Models and methods in social network analysis' edited by P.J. Carrington, J. Scott, and S. Wasserman.

[8] O. Frank and D. Strauss, "Markov graphs," Journal of the American Statistical Association, vol. 81, pp. 832–842, 1986.

[9] S. Wasserman and P. Pattison, "Logit models and logistic regression for social networks: I. an introduction to markov graphs and p!," Psychometrika, vol. 61, pp. 401–425, 1996.

[10] C. Anderson, S. Wasserman, and B. Crouch, "A p! primer: logit models for social networks," Social Networks, vol. 21, pp. 37–66, 1999.

[11] T. A. Snijders, P. E. Pattison, G. L. Robins, and M. S. Handcock, "New specifications for exponential random graph models," 2004.

[12] G. Robins, P. Pattison, and S. Wasserman, "Logit models and logistic regressions for social networks," Psychometrika, vol. 64, pp. 371–394, November 1999.

[13] P. Hoff, A. Raftery, and M. Handcock, "Latent space approaches to social network analysis," Journal of the American Statistical Association, vol. 97, pp. 1090–1098, 2002.

[14] C. Butts, "Network inference, error, and informant (in)accuracy: a bayesian approach," Social Networks, vol. 25, no. 2, pp. 103–140, 2003.

[15] T. Snijders, "Markov chain monte carlo estimation of exponential random graph models," Journal of Social Structure, 2002.

[16] P. Hoff, "Random effects models for network data," Irvine, California, November 2003.

17] Y. Wang and G. Wong, "Stochastic blockmodels for directed graphs," vol. 82, no. 8-19, 1987.

[18] M. Handcock, "Assessing degeneracy in statistical models of social networks," December 2003, working paper, University of Washington.

[19] P. Smyth, "Statistical modeling of graph and network data," in Proceedings of IJCAI Workshop on Learning Statistical Models from Relational Data, Acapulco, Mexico, August 2003.

[20] D. S. Johnson and M. A. Trick, Cliques, Coloring, and Satisfiability: Second Dimacs Implementation Challenge. American Mathematical Society, October 1996.

[21] "Connectivity and generalized cliques in sociometric group structure," Psychometrika, vol. 15, pp. 169–190, 1950.

[22] M. Everett, "Graph theoretic blockings, k–plexes and k–cutpoints," Journal of Mathematical Sociology, vol. 9, pp. 75–84, 1982.

[23] P. Mika, "Flink: Semantic web technology for the extraction and analysis of social networks," Journal of Web Semantics, vol. 3, pp. 211–223, October 2005.

[24] L. Kahney, "Making friendsters in high places," July 2003, the Wired News.

[25] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," in Proceedings of the 7th World-Wide Web Conference (WWW7), Bris-Bane, Australia, April 1998.

[26] J. Kleinberg, "Authoritative sources in a hyperlinked environment," in Proceedings of Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, San Francisco, California, January 1998.

[27] J. Dean and M. R. Henzinger, "Finding related pages in the world wide web," in In Proceedings of the 8th World Wide Web Conference, Toronto, Canada, May 1999.

[28] K. Bharat, A. Broder, M. Henzinger, P. Kumar, and S. Venkatasubramanian, "The connectivity server: Fast access to linkage information on the web," Brisbane, Australia, April 1998.

[29] P. Reddy and M. Kitsuregawa, "Inferring web communities through relaxed cocitation and dense bipartite graphs," in Proceedings of the Second International Conference on Web Information Systems Engineering (WISE2001), Kyoto, Japan, December 2001.

[30] R. A. S. Chakrabarti, B. Dom and P. Raghavan, "Using taxonomy, discriminants, and signatures to navigate in text databases," in Proceedings of 23rd International Conference on Very Large Data Bases, Athens, Greece, August 1997.