Association for Information Systems
AIS Electronic Library (AISeL)

PACIS 2010 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

2010

Towards A Scale Free Network Approach to Study Organizational Communication Network

Mohammed Uddin
University of Sydney, shahadat.uddin@sydney.edu.au

Shahriar Tanvir Hasan Murshed
University of Sydney, tanvir.murshed@sydney.edu.au

Liaquat Hossain
University of Sydney, liaquat.hossain@sydney.edu.au

TOWARDS A SCALE FREE NETWORK APPROACH TO STUDY ORGANIZATIONAL COMMUNICATION NETWORK

Mohammed Uddin, Project Management Graduate Program, School of Civil Engineering, University Of Sydney, Sydney, Australia, shahadat.uddin@sydney.edu.au

Shahriar Tanvir Hasan Murshed, Project Management Graduate Program, School of Civil Engineering, University Of Sydney, Sydney, Australia, tanvir.murshed@sydney.edu.au

Liaquat Hossain, Project Management Graduate Program, School of Civil Engineering, University Of Sydney, Sydney, Australia, liaquat.hossain@sydney.edu.au

Abstract

In this experiment, we study the scale free network property of an organizational communication network. We used social network analysis methods during organizational crisis period that captures the dynamics of communication networks. We did not find any significant fluctuation between the actor prominence in daily and aggregated networks. We found that email communication network displays a high degree of scale free behavior described by power law.

Key words: Communication network, Social network analysis, Scale free network, Power law distribution, Centrality

1 INTRODUCTION

Communication network is described as patterns of contacts which are created due to the flow of messages among the participating actors (or communicators). The word 'message' encompasses everything that can flow from one point of contact to another within and between the networks, including data, information, knowledge, image and symbol. These communication networks could take various forms, such as, personal contact networks, work related contact networks, strategic alliances among various firms, global network of organizations etc. (Monge & Contractor, 2003).

Organizations are commonly viewed as dynamic systems of adaptation and evolution that contain several parts and these parts interact with one another and the internal and external environment. In fact, this representation is so common that it has been described as self-evident by researchers (Morel & Ramanujam, 1999). This 'self-evident' representation of organizations as dynamic and adaptive system also implicitly assumes that organizations are 'Complex Systems'. Complex systems change inputs to outputs in a nonlinear way because their components interact with one another via a web of feedback loops (Anderson, 1999). Some of the characteristics of the complex systems which could be related to organizational phenomena are: like organizations, complex systems contain large number of interacting agents and associated emerging properties; organizations are complex systems because they are comprised of many individuals, groups and departments that interact with each other and constantly provide feedback; and finally, like complexity systems, organizations also show emergent properties or behavior which evolves due to the collective behavior of the various interacting agents (Morel & Ramanujam, 1999; Anderson, 1999).

In recent years, our understanding of complex networks have changed significantly due to the availability of 'real world' networks (Barabási, 2009) coupled with the advances in the knowledge base of analytic techniques employed by social network researchers in the area as diverse as Physics and Biology; Mathematics and Sociology; Organizational science and Psychology. Interestingly, the most commonly cited link between these diverse areas of research is self organising behavior of complex system. Two of the most frequently mentioned properties of real world complex systems are:

clustering behavior and the existence of scale free network (Barabási, 2009). Research indicates that most network display a high degree of clustering; and many scientific, technical and organizational network, ranging from biological network (Jeong et al. 2000) to WorldWideWeb (Albert et al. 1999) have been found to be scale free. Scale free network displays the characteristics of power law distribution. Power Law states that the probability that a randomly selected node has *k* links (i.e. degree *k*) follows $P(k) \sim k^{-\gamma}$, where γ is the degree exponent (Ravasz & Barabási, 2003). In this experiment, we concentrate on the scale free network property of the communication network.

The aim of this research is to study the scale free network property of Enron email communication network or Enron corpus released by Federal Energy Regulatory commission (FERC) in May, 2002 (Shetty & Adibi, 2004). We adopt social network analysis measure of centrality to study the network. We discuss the relevance of studying communication network structure and some relevant studies in the next section. Then, we briefly introduced the idea of scale free network followed by data analyses methodologies of our experiment. We, then, presented our findings in the results section and discuss its relevance to 'rich-gets-richer' phenomena in the scale network of email communications. Finally, we discuss the implications of our study in concluding remarks.

2 SOCIAL NETWORK ANALYSIS AND ORGANIZATIONAL COMMUNICATION NETWORK: ITS RELEVANCE AND RELATED STUDIES

The framework of studying communication networks can be traced back to a broad school of thought: group behaviour. The conceptual framework for studying groups starts with two givens: individuals within the groups and the surrounding environment where these individuals are embedded (McGrath, 1984). In one of his influential research works, during the early development of communication network analysis, Bavelas (1950) gave an example where five actors communicate with each other and varied their communication links in various ways which produced different communication patterns (Figure 1).

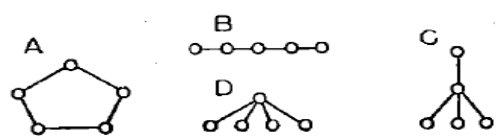

Figure 1: Some illustrative communication patterns among five individuals; from Bavelas (1950)

Another early researcher of communication networks, Leavitt (1951), found that communication patterns within which groups performed their tasks affected their behavior; the positions which individual occupied in a communication pattern affected their behavior while occupying those positions; and the characteristics of centrality were most clearly correlated with behavioral differences. Later on, Shaw (1964) demonstrated that groups with decentralised communication structure performed complex tasks in efficient manner compared to centralised communication network. Freeman (1978) also examined and elaborated the concept and measures of structural centrality in his influential work.

Traditionally, organizations were bounded by time and space. However, in this era we have seen that distance and time have become irrelevant and communication technologies have merged to generate a new kind of organization, so called, 'virtual organizations'. One of the key features of virtual organizations is a high degree of informal communication. As many organizations are moving towards decentralised, geographically dispersed structures, the lack of formal procedures, rules and norms are now more evident than ever. Researchers in the area of social networks have found that if the interactions of informal groups are tracked over a period of time, it may exhibit a pattern of communication and reveal what has been referred to as network structure (Ahuja & Carley, 1999).

In this research, we start with the premise that email networks constitute a useful proxy for the underlying communication networks within organizations. With the rapid advancement of information and communication technology, many organizations have been working in the virtual environments.

Technology has enabled organizational members to work collaboratively even though they are being geographically and spatially separated from each other. A study by Smith et al (2003) investigated how different age groups managed their personal networks and what types of technology-mediated communication tools they used. They found that people around their 30s (25-35 years) used email with most of their social network contacts (81%). The 60% of older age groups (50-60 years) also tended to keep in touch with their personal contacts primarily by using the email. As a modern and technologically advanced organization, we know that Enron employees used email as a significant medium of communication. Wellman (1997) reported that work groups in organizations using computer mediated communications tend to achieve higher levels of communication than those who do not, although this may reduce the use of face-to-face communication. Guimera et al (2002) argued that the email network provides an inexpensive but powerful alternative to the traditional approach of survey which is expensive and time consuming. Indeed, the exchange of email between individuals in organizations reveals how people interact and facilitates mapping the informal networks in a non-intrusive, objective, and quantitative way.

3 SCALE FREE NETWORKS IN SELF ORGANIZING COMPLEX SYSTEM

Barabási & Albert (1999) proposed that, independent of the system and the identity of its constituents, the probability *P(k)* that a vertex in the network interacts with *k* other vertices decays as a power law, following:

$$P(k) \sim k^{-\gamma} \qquad (1)$$

They have called this scale-free state. They proposed a model incorporating growth and preferential attachment, two key features of real life networks, and showed that these features are associated with the power-law distribution properties observed in many real networks. The following figure demonstrates a formation of scale free network, based on preferential attachment.

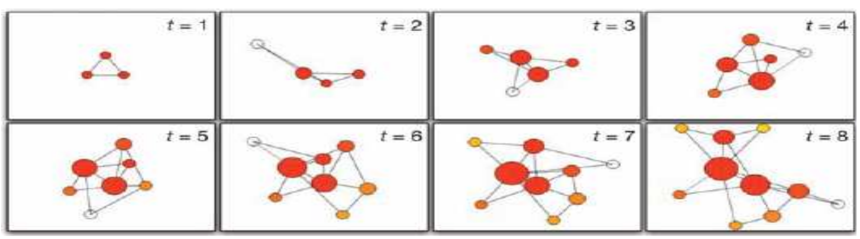

Figure 2: Formation of a Scale-free network. Adopted from (Barabási, 2009)

In the top left column of first row, at $t = 1$, three nodes are connected to each other for the initial network. At $t = 2$, a new node is connected to the existing network. At $t = 3$, another node is connected to the network. Preferential attachment predicts that, in deciding which node to connect to, the new node will prefer the node which is more connected node. Due to the phenomena of preferential attachment, a 'rich-gets-richer' network behavior is observed. It also implies that highly connected nodes acquire more links than those that are less connected, leading to the emergence of a few highly connected nodes which is referred to as hubs (Barabási & Albert, 1999; Barabási, 2009) or highly prominent nodes which play a vital role in shaping up the network. The resulting degree distribution of the network follows the power law, described in Equation 1 with exponent $\gamma = 3$. In recent years, we have observed new advances in the area of network analysis which have demonstrated the scale free network behavior in many large scale real world networks including: Telephone Call network (Abello et al., 1998); Worldwide Web (Albert et al. 1999); Internet (Faulstos et al. 1999); metabolic reaction networks (Jeong et al. 2000); software architecture (Valverde et al. 2002); e-mail communication network (Braha & Bar-Yam, 2006); and Distributed product development network in organization (Braha & Bar-Yam, 2007).

In the next section, we discuss the dataset being used for our analyses. Social network measure of centrality that is used to describe scale free network is also discussed in the next section.

4 DATASET AND MEASURE OF COMMUNICATION NETWORKS

We analyzed the modified Enron corpus which has been corrected and cleaned by Shetty and Adibi (2004). This modified corpus contains 252,759 email messages from 151 employees whereas the original Enron corpus, which was released by Federal Energy Regulatory commission (FERC) in May 2002, has 619,446 email messages from 158 users. As Enron collapsed in December 2001, we consider email messages only prior to six months; from July 2001 to December 2001. Moreover, most organizational crisis started to emerge for Enron during this period (Healy and Palepu 2003). After excluding weekends and public holidays, we used an observation period of 131 days for analysis purpose. We have also concealed the names of the email communicators and refer to them as Node 1, Node 2, Node 12, Node 58 etc.

One of the important and primary uses of graph theory and network analysis is the identification of the most important actor(s) within a social network. Various researchers used words like 'importance' or 'prominence' to describe this important network measure. Social network literature has established definitions of many of these measures that are based on degree, closeness, betweenness, information, and simply the differential status or rank of actors. Prominent actors are described as extensively involved in relationships with others (Wasserman & Faust, 1993). Hence, degree centrality has been used to describe the prominence of an actor in our email communication network. Centrality measure for an individual actor should be the degree of the node, $d(n_i)$. Thus, centrality is defined as an actor-level degree index (Wasserman & Faust, 2003):

$$C_D = d(n_i) = X_{i+} = \sum_j X_{ij} = \sum_j X_{ji} \qquad (2)$$

Also, relationships are defined as communication linkages between the actors. Number of emails sent by the employees to the actors within their respective communication networks is regarded as the degree centrality measures.

5 RESULTS & ANALYSES

In our first experiment, we analysed the correlation of all actors' out-degree centrality values between two consecutive days. Figure 3 shows the graphical representation of 130 correlation coefficient values for the network observation period of 131 days. The range of these correlation coefficient values is between 0.034 and 0.807, having mean and standard deviation of 0.543 and 0.138. Without few exceptions, there is a strong correlation for the actors' out-degree values between two consecutive days. The high correlation coefficient value implies that a high out-degree value for an actor in a particular day makes the same actor highly probable to have high out-degree values in the next consecutive day and vice-versa.

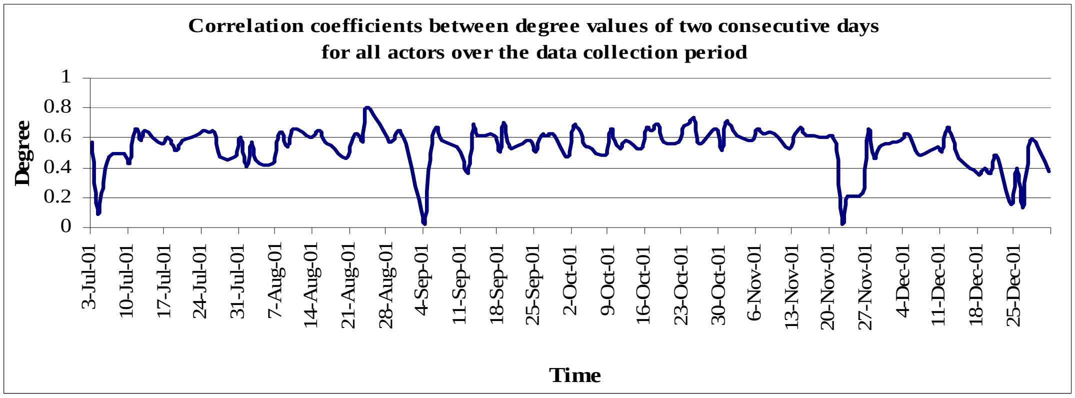

Figure 3: Correlation coefficient values for the out-degrees of two consecutive days over the data collection period.

Each of these daily networks shows a distribution for out-degree centrality scores well described by power-law. Thus, a small number of highly connected nodes have greater importance in the connectivity of the entire communication network. Moreover, we found that there is a repetition of highly connected nodes in each daily network. In Figure 4 [**a-b**], we plot the time series of degrees for randomly selected dates of 21 August 2001 and 14 November 2001. We also plot the corresponding log-log plot for the time series of degrees in Figure 4 [**c-d**].

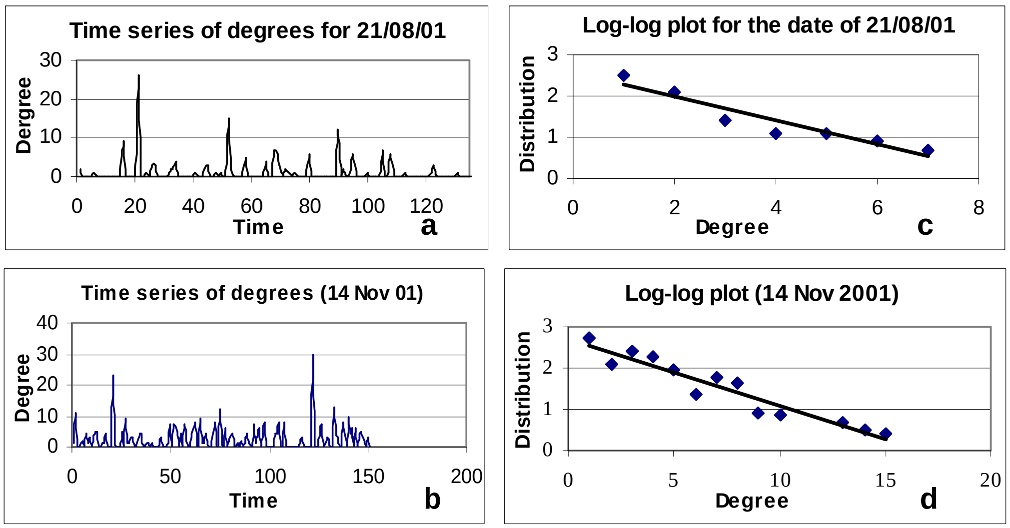

*Figure 4: [**a-b**] Time series of degrees for randomly selected dates of 21 August 2001 and 14 November 2001. [**c-d**] are the corresponding log-log plot for the time series of degrees.*

From Figure 4 [**c-d**], we clearly see that the distribution of degree follow power-law distribution as they produce straight lines in the log-log plot. After analysing the daily network, we measure the out-degree centrality scores for each of the identified prominent actor from everyday network, over the duration of observation period. We found that most of the prominent actors exhibit stable time series. Figure 5 plots the graph of degree variations for a local hub node (Node 58). This node is found most of the times (85 times) in the top-ten-rank list. Degree values are high in general with having low values in few occurrences.

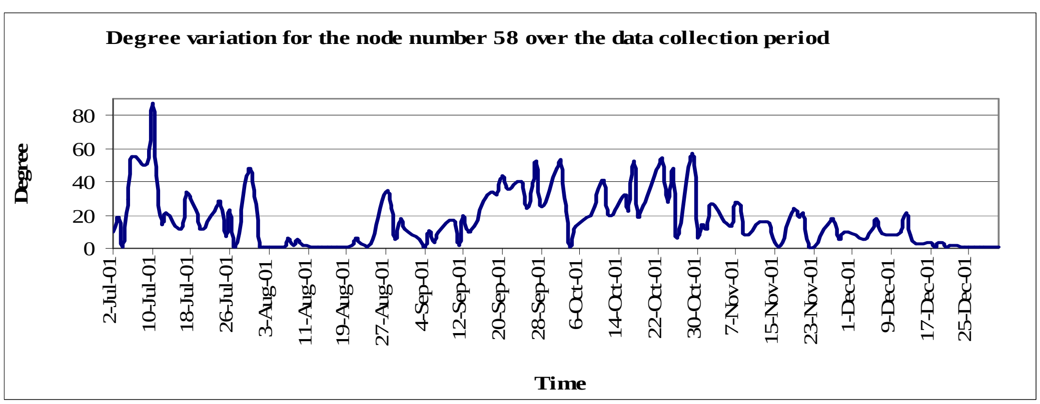

Figure 5: Example of degree variation for the local hub (Node 58). Distribution of Node 58 does not follow the power-law theory.

From the graph of Figure 5, we found that this distribution for Node 58 does not follow the power-law distribution. However, we also found that a few nodes also exhibit a highly fluctuating time series. One such node is Node 12. Figure 6[**a**] demonstrates the fluctuating time series graph of the degree variations of Node 12. This node was found once acting as a prominent hub within the experiment period of 131 days of years 2001.

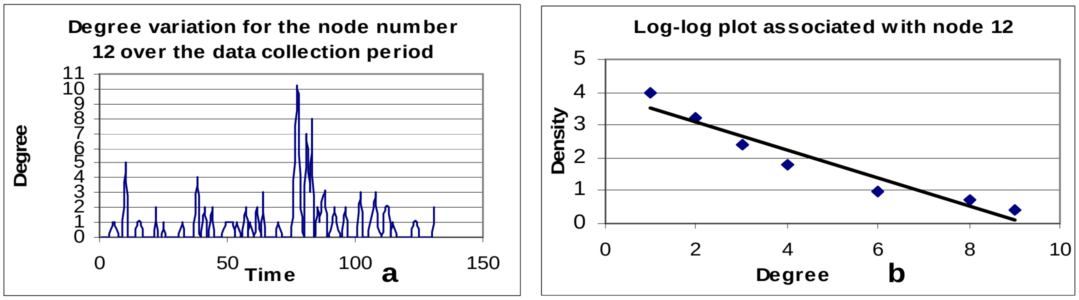

*Figure 6: [**a**] Example of fluctuating time series associated with the local hub (ID 12). [**b**] The log-log plot for the degree distribution of this node also follows the power-law theory.*

In the Figure 6 [**b**], we can also see that the log-log plot for the degree distribution of this node also follows the power-law theory. This implies that some of the actors became very prominent in the network over the time while some others became isolates. By summarising the results of figure 4, 5 and 6 we can conclude that the daily communication network of Enron employees (included in the dataset) followed the power-law distribution. However, when we look at the network structure of individual prominent actor, we find that some of them follow power-law distribution whereas some do not.

Finally, we checked the percentage of degree values showed by top ten actors in rank-list and compared it with the overall degree value showed by all actors in the network over the data collection period of 131 days. The average percentage values by top ten actors in the rank-list are 60.92 (Figure 7).

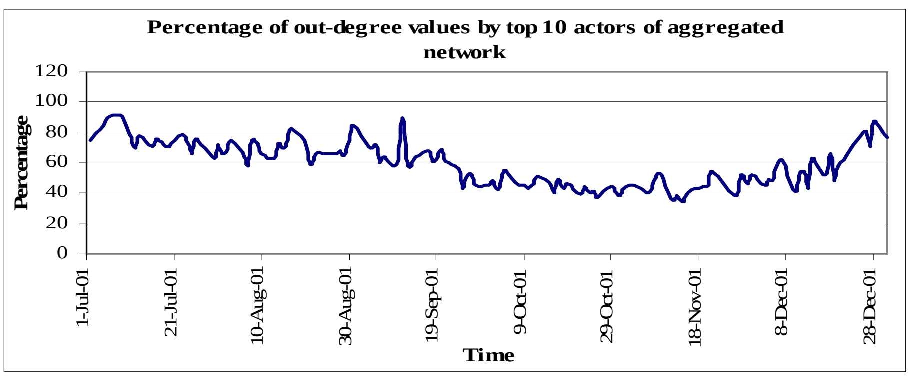

Figure 7: Percentage of degree values exhibit by the top-ten actors of the daily network compare to the overall daily network degree values.

This result clearly shows that only about 7% (top 10 nodes compared to 151) of the nodes exhibit around 60% of the total network degree values. Only a handful of actors became very prominent in the network. From figure 7, we see that in the daily network, top-ten actors exhibit an average of 60% of the total nodal degrees. A few actors who are repeatedly located in the top-ten list also showed high degree values in the daily networks as well as aggregated network. These actors have become the prominent actors in the network over the time, which further reinforces the ‘rich-gets-richer’ phenomena observed in the scale free emerging communication network of Enron.

6 CONCLUSION

In this experiment, we did not find any significant fluctuation between the actor prominence in daily and aggregated networks of Enron. Our research showed similar outcomes with the works of Abello et al., (1998); Albert et al. 1999); Faulstos et al. (1999); Jeong et al. (2000); Valverde et al. (2002); and Barabási (1999 & 2009).

Although Barabási (1999) predicted a decade ago that scale variant state of a network could be generic property of many real life complex network, it wasn't until recently when researchers noted one of the most surprising discovery of modern network theory

"………the universality of the network topology: Many real networks, from the cell to the Internet, independent of their age, function and scope, converge to similar architectures" (Barabási, 2009)

We have also found from our experiment that email communication network displays a high degree of scale free behavior described by power law. However, Strogatz (2001) argued that scale-free property is common in real life network but not universal. For example, the co-authorship network of scientists shows power law distribution but with an exponential cutoff (Newman, 2001); the power grid network of western United States distribution is exponential (Amaral et al. 2000); and for the social network of Mormons in Utah the distribution is Gaussian (Amaral et al. 2000).

One of the many questions that have come to our mind during this research is that - is there any functional advantage of scale free network topology? Albert et al (2000) found that there are practical advantages and disadvantages to it. They found that this type of network displays high degree of tolerance against random failures as only a few prominent hubs dominate their topology. However, the flip side is that such networks are extremely vulnerable to the attack on their hub(s). It has also been confirmed numerically and analytically by examining how the average path length and size of the prominent hubs depend on the number and degree of the nodes removed (Strogatz, 2001). Even though, we haven't fully tested this hypothesis with Enron dataset, one can argue that (based on the literature review) it might well be true in the case of Enron. We know that several of Enron's prominent hubs, including 2 former CEOs, Chief Financial Officer, a number of vice presidents, and some senior management staff were involved (and subsequently implicated by court) in the defrauding process of the organization. As a corporation, Enron was widely acclaimed as a paragon of economic and organizational innovation only to be pilloried after its collapse.
The methodological contributions of this study are worthy of note. It builds on an emerging stream of network structural research that applies social network analysis to organizational email communication data in order to research important questions on organizational communication network. With the increasing popularity of electronic communications, the increasing popularity of social network analysis and the growing sophistication of SNA tools, it is to be expected we can develop deeper insights into a wide range of organizational phenomena.

References

Abello, J., Buchsbaum, A. & Westbrook, J. (1998). A functional approach to external graph algorithms. Lecture Notes in Computer Science, 1461, 332–343.

Ahuja, M. K. & Carley, K. M. (1999). Network Structure in Virtual Organizations. Organization Science, 10(6), 741-757.

Albert, R., Jeong, H. & Barabási, A. (1999). Diameter of the World Wide Web. Nature, 401,130-131.

Albert, R., Jeong, H. & Barabási, A. (2000). Error and attack tolerance of Complex networks. Nature, 406, 378 – 382.

Amaral, L. A. N., Scala, A., Barthélémy, M. & Stanley, H. E. (2000). Classes of behavior of small-world networks. Proc. Natl Acad. Sci. USA, 97, 11149–11152.

Anderson, P. (1999). Complexity Theory and Organization Science. Organization Science, Special Issue: Application of Complexity Theory to Organization Science, 10(3), 216-232.

Barabási, A. (2009). Scale-Free Networks: A Decade and Beyond. Science, 325, 412.

Barabási, A. and Albert, R. (1999). Emergence of scaling in random networks. Science, 286. 509-512.

Bavelas, A. (1950). Communication Patterns in Task-Oriented Groups. The Journal of Acoustical Society of America, 22(6).

Behr, P. & Witt, A. (2002, 28 July). Visionary's dream led to risky business: opaque deals, accounting sleight of hand built and energy giant and ensured its demise. Washington Post, p. A01.

Braha, D. and Bar-Yam, Y. (2006). From Centrality to Temporary Fame: Dynamic Centrality in Complex Network. Complexity, 12(2), 59-63.

Braha, D. and Bar-Yam, Y. (2007). The statistical mechanics of Complex Product Development: Empirical and Analytical Results. Management Science, 53(7), 1127 – 1145.

Erdos, P. and Renyi, A. (1959). On random graphs, Publicationes Mathematicae (Debrecen), 6, 290.

Faloutsos, M., Faloutsos, P. & Faloutsos, C. (1999). On power-law relationships of the internet topology, Comp. Comm. Rev. 29, 251–262.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. Social Networks, 1(3), 215-239.

Guimera, R. Danon, L. Dfaz-Guilera, A. Giralt, F. & Arenas, A. (2002). Self-similar Community Structure in Organizations [Electronic Version]. retrieved on 20/03/2007 from http://arxiv.org/PS_cache/cond-mat/pdf/0211/0211498.pdf

Hamilton, E. A. (2006). An exploration of the relationship between loss of legitimacy and the sudden death of organizations. Group & Organization Management, 31(3), 327-358.

Jeong H., Tombor, B., Albert, R., Oltavi, Z, N. & Barabási, A.-L. (2000). The large-scale organization of metabolic networks. Nature, 407, 651–654.

Johnson, J. D. (1992). Approaches to Organizational Communication Structure. Journal of Business Research, 25, 99-113.

Knoke, D., and Burt, R. S. (1983). Prominence. In Burt, R. S., and Minor, M. J. (Eds.) of Applied Network Analysis, 195-222, Sage Publications, CA.

Leavitt, H. J. (1951). Some effects of certain communication patterns on group performance. Journal of Abnormal and Social Psychology, 46, 38-50.

McGrath, J. E. (1984). Groups: interaction and performance. Prentice-Hall, New Jersey.

Monge, P. R. and Contractor, N. S. (2003). Theories of Communication Network. Oxford University Press, NY.

Morel, B. & Ramanujam, R. (1999). Through the looking glass of complexity: The dynamics of organizations as adaptive and evolving systems. Special issue: Application of Complexity Theory to Organization Science. Organization Science, 10(3), 278-293.

Newman, M.E.J. (2001). The structure of scientific collaboration networks. Proceedings of National Academy of Science. USA 98, 404-409.

Ravasz, E. and Barabási, A. (2003). Hierarchical organization in complex networks. Physical Review E 67, 026112.

Roberts, J. & Armitage, J. (2006). From Organization to Hypermodern Organization: On the accelerated appearance and disappearance of Enron, Journal of Organizational Change Management, 19(5), 558-577.

Roberts, K. H. and O'Reilly III, C. A. (1978). Organizations as Communication Structures: An empirical approach. Human Communication Research, 4(4).

Shaw, M. (1964). Communication Network. In Berkowitz, L. (Eds.) of Advances in Experimental Social Psychology, 111-147, Academic Press, New York.

Smith, H., Rogers, Y., Brady, M. (2003). Managing one's social network: Does age make a difference? In Ninth IFIP TC13 International Conference on Human-Computer Interaction (INTERACT), Zurich, Switzerland, 551-558.

Shetty, J. S. & Adibi, J. (2004). The Enron Email Dataset Database Schema and Brief Statistical Report. Information Sciences Institute, Los Angeles, CA.

Strogatz, S. H. (2001). Exploring Complex Network. Nature, 410, 268-276.

Tyler, R. J. Wilkinson, D. M. & Huberman, B. A. (2003). Email as Spectroscopy: Automated Discovery of Community Structure within Organizations. In the International Conference on Communities and Technologies, Netherland, 81-96.

Valverde, S., Cancho, R. F. and Sole, R.V. (2002). Scale free networks from optimal design, Europhysics Lett. 60, 512 – 517.

Wasserman, S. & Faust, K. (1994). Social network analysis: methods and applications. Cambridge University Press, NewYork.

Wellman, B. (1997). An Electronic Group is virtually a Social Network. In S. Keisler (Eds.) of Culture of the Internet, 179-205, Lawrence Erlbaum, Hillsdale, NJ.