Proceedings of the South African Information Security Multi-Conference (SAISMC 2010)

Organizational Risk using Network Analysis

H.L. Armstrong[1] and I. McCulloh[2]

[1]School of Information Systems, Curtin University, Western Australia
[2]Adjunct Professor, School of Information Systems, Curtin University, Western Australia and Dept of Behavioral Sciences & Leadership, USMA West Point, NY
e-mail: h.armstrong@curtin.edu.au; ian.mcculloh@usma.edu

Abstract

Business organizations are held together not only by formal reporting and authority networks but also by informal networks that connect people across numerous layers of hierarchical organizational structures. People form networks of contacts and communications and through these networks they 'get things done'. Although extensive research has been carried out on social networks the application of these methods to organizational risk has not been widely published. However, network analysis does provide a source of information on potential risks to aid decision-makers within an organization. The application of network analyses in identifying and measuring potential risks based upon the analyses of people, knowledge, tasks and resources is presented in this paper.

Keywords

Organizational risk, network science, network analysis

1. Introduction

Determining risk in IT driven organizations has predominantly focussed on technological rather than social approaches. The risk analysis process is much easier where decision makers are dealing with measurable variables provided by quantitative analysis. There are, however, many variables in an organization that are difficult to capture, and these relate mainly to the human element, and how humans use their connections to get things done. Hence, social factors also have a place in investigating risk in organizational settings as strengths and vulnerabilities emanate from human interactions within organizations.

It has been well-recognised in past research the importance of a balanced organizational approach, one that includes both social and technological considerations. Kilduff and Tsai (2007) highlight the neglect to consider social factors and the relationships between individuals not reflected in organizational sociology, using sets of predefined individual characteristics to define generalized sets. The social environment can be expressed as patterns or regularities in relationships among interacting units and the study of these relationships entails the utilization of a set of methods and analytic concepts that are different from methods of traditional statistics and data analysis (Wasserman & Faust, 1994).

Network Science, also known as the science of networks, is "the study of the theoretical foundations of network structure/dynamic behaviour and the application

of networks to many subfields" (Lewis, 2009). Network Science is an overarching paradigm concerned with "the study of networks, be they biological, technological, or scholarly networks. It contrasts, compares, and integrates techniques and algorithms developed in disciplines as diverse as mathematics, statistics, physics, social network analysis, information science and computer science" (Börner et al., 2007:537). Social Network Analysis (SNA) uses statistical methods to analyze and predict social behaviour through the study of relationships. This field is highly advanced and SNA techniques are now being deployed in the field of network science to study meta-networks (the overlaying of simplex networks into integrated complex networks) in many disciplines. In order to manage problems effectively, it is necessary to identify and understand the stakeholders, the major connections and patterns in data collected from the problem domain. We need to be able to measure, model, manage, and understand the structure and function of large networked physical and information systems (Börner et al., 2007:537) and this can be achieved through SNA techniques. Although network science is the over-arching paradigm, the techniques of network analyses utilize tools and approaches that have been used within SNA for more than 8 decades (see for example Moreno, 1946). SNA employs a set of statistical techniques designed specifically for the study of networks and has been used to study relationships between people in numerous disciplines. However, the use of meta-networks that analyze the links between not only people in an organization, but also networks of other important factors, such as resources, tasks, knowledge and the like, have the potential to deliver valuable information to decision makers.

Little work utilizing this approach has been reported in this field. Although there is a wealth of publications reporting research into the application of SNA techniques to organizational success and learning (see the work of Steve Borgatti, for example), organizations are predominantly complex, integrated wholes and analysis of social relationships presents only a partial view of the total situation. The PCANS model was an early research project using broader network analyses to develop a structural model in organizations (Kilduff & Krackhardt 2008; Krackhardt and Carley 1998). The PCANS model was a 'meta-model' approach, studying the interrelationships between simplex models of Precedence (including sequence and dependence), Commitment of resources, Assignment of personnel to tasks, Network of people relationships (both formal and informal) and Skills. The meta-network approach is the study of not only each simplex network but also the interrelationships between each of the component networks in order to get a more comprehensive glimpse of the total situation.

Although network analysis provides great promise very few publications report the application of SNA techniques to meta-models of the relationships between not only people, but other factors that are fundamental to decisions regarding organizational risk and the ongoing survival of an organization. Krackhardt and Carley (Krackhardt and Carley 1998) write of the importance of the consideration of complex interdependent networks in their research on command and control, and their application of network analysis via meta-models of individuals, tasks and resources. Research has been undertaken in the development of theoretical models using organization simulation and complexity theory in organizations. Approaches to destabilizing adversarial organizations and the study of terrorist networks have also

been reported (for example, see McCulloh, 2009; Moon, 2008; Ressler, 2008; US Department of Defence 2007) but little is reported on the practical application into live organizational settings. If network analysis can provide methods for destabilizing adversary organizations, then surely the same analytical techniques can provide business organizations with a useful tool for determining risks.

The tools for managers to carry out risk analysis using SNA techniques in organizational settings are available. For example, the Center for Computational Analysis of Social and Organizational Systems (CASOS) at Carnegie Mellon University has undertaken numerous risk-related projects with defence organizations employing their custom-built tool the Organizational Risk Analyzer (ORA). ORA provides an automated tool to help carry out meta-network analysis and network visualization and produces useful standard reports. There are several pre-made parsing functions that allow a user to download structured data from databases, email communications, or from spreadsheets. There is also a suite of features for analyzing networks over time (McCulloh & Carley 2009; McCulloh et al. 2008), giving valuable information on trends and movement to decision-makers in organizations.

2. Network Analysis Measures

In order to use network analyses for organizational risk it is necessary to view the organization as a series of interconnected networks. The organizational network is made up of nodes (people, objects, tasks, etc.) and the links (relationships) between these nodes. For example nodes could be agents, tasks, knowledge, resources, roles, beliefs, events, locations or actions. Measurements that can aid decision-makers are calculated at both the node level and the network level. The importance of a node in a network is measured by its centrality. There are many ways in which a node might be considered "central." There are several basic centrality measures used as a foundation for network analyses: degree, closeness, betweenness, and eigenvector centrality. We also describe some less well known measures for meta-network analysis.

Degree centrality measures the number of connections a particular node has with other nodes, and is the sum of the lines in and out of the node, i.e. the total of connections the node has. Node degree is measured by three calculations: In-degree, the total number of connections into the node; Out-degree, the total number of outward connections or lines; and Total degree, the number of in and out connections. Nodes with a high total degree value have direct influence on the nodes that are connected to it. When nodes are people, these individuals are most likely to know and diffuse new information; and isolation of individuals with high degree centrality can be slightly crippling for a short time (Carley et al. 2007).

Closeness measures the sum of the geodesic distances to all other nodes in the graph, where a geodesic is the shortest path between two nodes. Nodes with a high closeness value are usually good sources of information as well as aiding the diffusion of information through the network as they are more closely connected to a greater numbers of nodes. People that are high in closeness centrality are good

individuals to survey for understanding attitudes, ideology, and potential sources of organizational improvement.

Betweenness centrality measures the extent to which a particular node lies on the geodesic path between other nodes in the graph. In a social network this measure reflects the number of people with whom a person is connecting indirectly through their direct links (Wasserman & Faust, 1994). Betweenness centrality measures a node's ability to broker knowledge and resources between other individuals in the network. As such, it tells you who or what holds power in the network. Nodes with a high betweenness value can lead to areas of vulnerability and disruption in the organization. Nodes high in betweenness centrality also have high levels of stress as they can become overburdened with their brokerage role. Thus betweenness is one of the major measures for identifying organizational risk in a network.

Eigenvector centrality measures the extent to which a node is connected to other highly connected or important nodes. If you took a random walk through the network you would come across nodes high in eigenvector centrality most often. This measure indicates emergent leadership in social networks. Nodes with a high eigenvector centrality value form valuable connectors between groups or cliques. Where the nodes are people, a high eigenvector centrality indicates individuals who are more likely to make things happen due to their knowledge and connections with the important 'who' and 'what' in the network.

Cognitive Demand is very similar to degree centrality, but it is defined for a meta-network. The cognitive demand is the number of connections a node has to other nodes inclusive of all node types in the network. For example if an individual node was connected to three people, two resources, and four knowledge nodes; their degree centrality would be three, however, their cognitive demand would be nine. Nodes with high cognitive demand are likely to become emergent leaders.

Exclusivity is similar to betweenness centrality, but it is defined for a meta-network. Exclusivity is the number of times a node lies on the shortest path from one node type to another. For example, a critical resource might be controlled by one particular individual. Anybody else in the organization who wants that resource must request it through the first individual. That individual would be high in Resource Exclusivity because he would be on the shortest path between the resource and everyone else in the social network. Similarly, someone high in Knowledge Exclusivity would control the organization's access to key knowledge.

Network Density is measured by the total number of links actually appearing in the network as a proportion of the total possible links. Lin and Zhang (2007) in their research into knowledge transfer using SNA reports that if the density is low relations between members become indifferent and intercourse is superficial, which is a disadvantage to tacit knowledge sharing and transfer; while high density may result in a negative influence to performance. In essence low density indicates little collaboration and communication within the network, indicating a low level of shared objectives and common goals.

A Hub is a node that points to other nodes that have a relatively large number of links. Hubs have a large number of out-degree links, and possess high degree and betweenness centrality. An Authority is a node with a large number of in-degree links, is knowledgeable within a domain and a respected source of information. Where the relationships or links are always reciprocated (undirected networks) eigenvector, hub and authority centralities are the same, as eigenvector centrality is being connected to highly connected nodes.

The application of the above basic measures to a network can produce indications of strengths and weaknesses in the network. Too much, or too little can indicate risk areas. The weaknesses identified can then be investigated further to determine whether or not they are risks that warrant further consideration.

3. Organizational Risk Indicators

Organizational risk analysis is the process of investigating a network and determining where the organization has vulnerabilities. The first step must be to determine whether the organization should be an efficient or a learning organization (Daft, 2009). Organizations that perform routine tasks repetitively with low likelihood of facing new challenges should be organized as efficient hierarchies with centralized decision making. On the other hand, organizations that continually face new challenges and must adapt to succeed require greater collaboration, sharing of knowledge and resources, and decentralized decision making. In an efficient organization, it is desirable to see well documented procedures and formal reporting. Senior leaders should be more central in the organization and centrality measures should be highly correlated. In learning organizations, all nodes should be more similar in their centrality values. For this type of organization high centrality may represent risk and vulnerability.

Simplex networks consist of a singular node type with a common relationship or link type. Multi-mode networks comprise nodes of different types, such as agents, knowledge and tasks. Multi-plex networks comprise relationships of different types, such as information flow, monetary flows, etc. Organizational networks tend to be multi-mode, multi-plex networks due to the variety of tasks and knowledge across the organization and the complexity of their relationships. Take, for example, a network representing an IT security applications development organization. This example organization is a simulation based upon a real organizational structure of an SME whose main task is development of turn-key applications. This is an efficient organization in Daft's terms. The network consists of nodes representing agents (individuals), knowledge, tasks and resources with 17 agents, 9 resources, 12 types of knowledge and 14 tasks. The function of this organization is to develop security products, a repetitious and structured activity requiring segmentation of tasks with specialized knowledge. Managers are the centre of most of the activities and relationships with little collaboration occurring outside the individual's immediate area. Figures 1 illustrates the decision making structure of the organization via the agent x agent network illustrating the relationships between the individuals. Centralized control and decision making is reflected in the hubs. Notice that these hubs are the senior managers; A1-SecManager, A2-BAManager, A3-DevManager,

and A4-QualManager have more links than other individuals. Figure 2 in contrast illustrates a learning organization structure for a simulated firm with the same number of agents, knowledge, tasks and resources. Within a learning organization more collaboration takes place between individuals hence a greater sharing of knowledge and more distributed decision-making. This means there are more links joining nodes across the network and the senior managers are no longer the only major hubs.

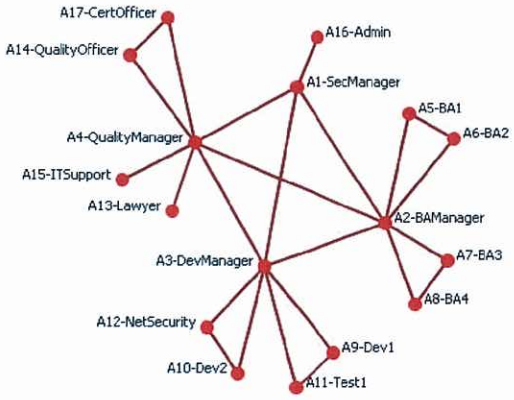

Figure 1: An efficient organization's agent x agent network illustrating high centralisation of control by managers (ORA generated image).

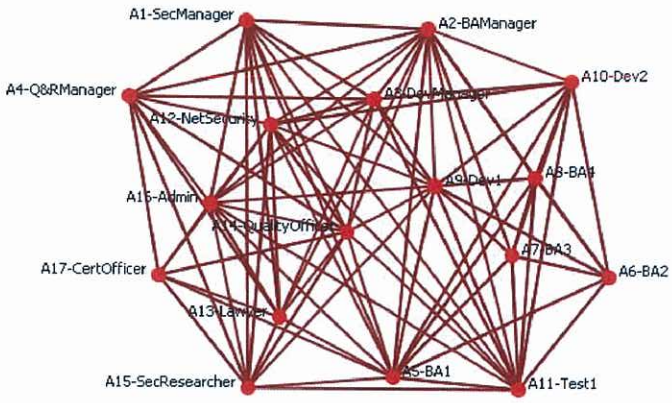

Figure 2: A learning organization's agent x agent network illustrating more distributed control (ORA generated image).

To aid our analyses we study the relationships between other types of nodes; for example, agent x agent (who interacts with who), agent x knowledge (who has what knowledge), agent x task (who does which tasks), agent x resource (who manages which resources), task x resources (what resources are required for a given task), etc. Visualization of these networks aids our understanding of the situation. Each node type is differentiated by color with agent nodes being red, knowledge are yellow,

resources are aqua and tasks are mid-blue. Figures 3 and 4 illustrate the total network for the two different types of organization. The efficient organization in Figure 3 has less links across the network, evidence of less interaction and integration and less sharing of knowledge and ideas. The density of the network for an efficient organization is smaller than that for a learning organization. The simulated learning organization in Figure 4 illustrates many hubs and a higher density (greater number of links overall) reflecting a greater level of communication, collaboration and knowledge sharing.

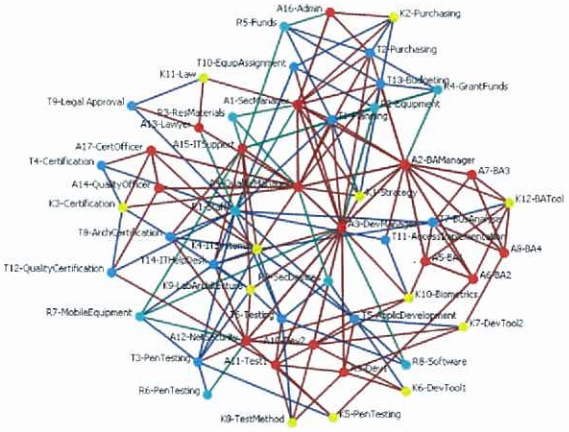

Figure 3: Total network for the efficient organisation (ORA generated image).

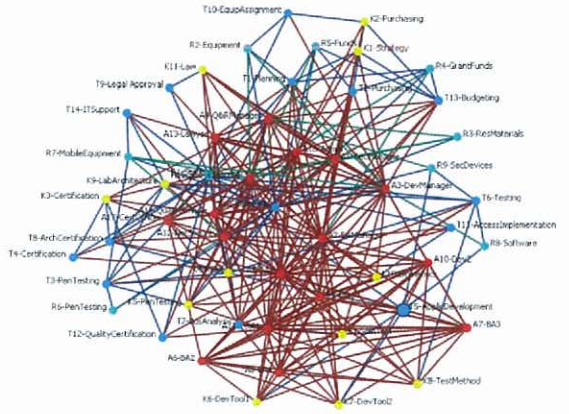

Figure 4: Total network for the learning organization (ORA generated image).

By employing advanced statistical techniques to calculate the centrality measures described earlier risk analysis can be performed on organizational networks using a variety of indicators. Centralized networks are dominated by a small number of very central nodes having a high degree of direct links to other nodes. This structure reflects an efficient organizational network. If the central nodes are removed the

network is likely to fragment into unconnected sub-networks. Akin to the center of a star network topology central nodes run the risk of being single points of failure. They are susceptible to disruption and damage to a central node can negatively affect the entire network. Decentralized networks on the other hand possess numerous important hubs, with nodes indirectly connected to all other nodes giving the network more flexibility. Learning organizations tend to have a decentralized structure with a higher degree of links between nodes. Decentralized networks are able to more easily regenerate lost connections and replace damaged nodes due to their flexibility of structure.

Analysis of key entities (nodes) in a network uses total degree, closeness, betweenness, eigenvector centrality, cognitive demand, and exclusivity to identify the most important persons in the organization. People with high total degree centrality are linked to many others in the organization and have high direct influence. However, we can't just assume that the individual who is most connected to others (the hub, the person with high degree centrality) is the most critical. Many nodes have connections only to other nodes in their immediate cluster or clique, connecting solely to those nodes that are already connected to each other. We need to identify who is the right person to gain information from, to influence others, who is critical due to access to specialized resources or skills, and so on. People with high closeness centrality have greater access to informal information including ideas, rumours, happenings, beliefs, etc. These individuals have great influence over the attitudes and organizational culture of the group. When looking for organizational risk, the opinions and ideas for improvement of these people should be viewed carefully. People with high betweenness centrality hold power in the organization. Ideally, many nodes should have equivalent betweenness centrality values. If there are a few nodes with unusually high betweenness, they may hold too much power or be under high stress. If these individuals are removed from the network, the organization can suffer greater harm. Boundary spanners are those nodes that connect their group to others, having connections outside their immediate cluster. These nodes are reflected by high betweenness but not necessarily high degree and act as bridges connecting groups. Boundary spanners have access to information and ideas which give them a more comprehensive picture of the situation not usually seen by local members of the group. Eigenvector centrality provides another measure of influence in an organization. People high in eigenvector centrality may form a social-elite within the group. This type of structure may significantly affect the organizational climate and culture, as these individuals will form social norms and mores specific to the group that will likely influence others. Understanding who these individuals are will allow senior leaders to shape the organizational culture.

One agent node can have high task exclusivity where very few other agents can perform a particular task. The impact of isolating an agent with task exclusivity for critical tasks can be severe. Due to the specialization of tasks efficient organizations tend to display greater task exclusivity. When individuals are not available to perform their specialist tasks then the organization's performance can be negatively affected. An agent node can also possess knowledge exclusivity where it holds knowledge not held by other agents. Core knowledge or knowledge key to achievement of missions and objectives should not be held exclusively as isolation of key agents can result in poor decisions and performance. People with high

knowledge exclusivity are not only core to the operations of the network but are also targets for the supply or blocking of crucial information to business processes and decisions. The same logic applies to exclusive sources of information or data within an IT network.

Other potential problems that can be identified within a network include those nodes that are under-performing or unconnected from the organization. These potential issues include the least integrated agents in the network, the least socially connected agents, clueless agents who have the lowest shared situation awareness, weak knowledge points, weak resource points, weak task points, and isolates (nodes without links). Many managers would like to remove key nodes to see the effect on the network. This is particularly appropriate for organizations undertaking projects where key staff or resources need to be moved between projects, or limited resources need to be shared. It is also useful to provide information for restructuring, diversification or merger activities. Linking unconnected nodes answer other 'what if?" questions posed by decision makers in organizations. Where risks are evident due to lack of information flow, isolation of nodes, etc. the introduction of links or new paths can alleviate potential problems and reduce risk.

4. Conclusion

Centrality measures may indicate areas of organizational risk for those managers open to viewing their organizations as complex socio-technical multi-mode networks. Via statistical analyses and relational algebra, network analyses provide the capability to analyse the structure and characteristics of the organizational network and identify potential risks. For decision makers not trained in network analysis, the theory is not difficult to apply using computerized tools. The measures used in SNA that easily transfer into considerations of organizational risk have been incorporated into automated analysis and mapping tools. Many computerized tools are available for carrying out network analysis and visualization, and ORA is one such tool. It is recommended that the chosen tool provide specific management reports relating to organizational risk, thus aiding managers to generate findings that are relevant to their specific situation.

The quality of analyses will be limited by the quality of the data, and managers rely on good quality information for decision-making. To get best results, data needs to be gathered from multiple sources including emails, surveys, observation and formal documented sources. The selection of node and link types is another important decision when studying organizational risk. This decision must be made prior to collecting data. In addition to the common measures of centrality to obtain information about the network, relational algebra (Wasserman and Faust, 1994) can be used to construct new network relations from existing data, which may provide greater insight into the organization. We must be cognizant, however, not to rely on network analyses as the only source of information as it is not a magic ball showing all the vulnerabilities in a network. Network analysis is, however, a highly useful tool to guide individuals as they study organizational risk.

5. References

Börner, K., Sanyal, S. & Vespignani, A. (2007), Network Science. In Cronin (Ed) Annual Review of Information Science & Technology. (pp 537-607) Medford, NJ

Carley, K., Diesner, J., Reminga, J. & Tsvetovat, M. (2007), Toward an interoperable dynamic network analysis toolkit. Decision Support Systems. Vol 43. pp 1324-1347

Daft, R. (2009), Organizational Theory and Design. South-Western Cengage Learning, Mason OH, USA

Kilduff, M. & Krackhardt, D. (2008), Interpersonal Networks in Organizations: Cognition, personality, dynamics and culture. Cambridge University Press, USA.

Kilduff, M. & Tsai, W. (2007), Social Networks and Organizations. Sage Publications, LA, USA.

Krackhardt, D. & Carley, K. (1998), A PCANS Model of Structure in Organizations. In Proceedings of the 1998 International Symposium on Command and Control Research and Technology. Monterey, CA.

Lewis, T.G. (2009), Network Science: Theory and Applications. New Jersey, Wiley

Lin, X. & Zhang, Q. (2007), Optimization of Knowledge Sharing & Transfer Network, Wireless Communications, Networking and Mobile Computing, WiCom, pp 5613-5616

McCulloh, I. (2009), Detecting Changes in a Dynamic Social Network. Carnegie Mellon University, School of Computer Science, Doctor of Philosophy, CMU-ISR-09-104

McCulloh, I. & Carley, K. (2009), Longitudinal Dynamic Network Analysis: Using the Over Time Viewer Feature in ORA. Carnegie Mellon University, CMU-ISR-09-118.

McCulloh, I., Webb, M., Carley, K. & Horn, D. (2008), Change Detection in Social Networks. U.S. Army Research Institute for the Behavioral and Social Sciences, Technical Report No. 1235. Arlington, VA., http://www.casos.cs.cmu.edu/publications (Accessed 30 Jan 2010)

Moon, I. (2008), Destabilization of Adversarial Organizations with Strategic Interventions. Carnegie Mellon University, Doctor of Philosophy, CMU-ISR-08-124

Moreno, J. (1946), Psychodrama and Group Psychotherapy, Sociometry, Vol. 9, No. 2/3, pp. 249-253

Ressler, S. (2006), Social network analysis as an approach to combat terrorism: past, present and future research, Homeland Security Affairs, Vol. II, No. 2

US Department of Defence. (2007), US Army Counterinsurgency Handbook, Department of the Army; Skyhorse Publishing, New York

Wasserman, S. & Faust, K. (1994), Social Network Analysis: Methods and Applications, Cambridge University Press, New York, USA.