5-2013

Reliability Analysis of Social Networks

Kellie R. Schneider
University of Arkansas, Fayetteville

RELIABILITY ANALYSIS OF SOCIAL NETWORKS

RELIABILITY ANALYSIS OF SOCIAL NETWORKS

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Industrial Engineering

By

Kellie R. Schneider
University of Arkansas
Bachelor of Science in Industrial Engineering, 2002
University of Arkansas
Master of Science in Industrial Engineering, 2006

May 2013
University of Arkansas

Abstract

The primary focus of this dissertation is on the quantification of actor interaction and the dissemination of information through social networks. Social networks have long been used to model the interactions between people in various social and professional contexts. These networks allow for the explicit modeling of the complex interrelations between relevant individuals within an organization and the role they play in the decision making process. This dissertation considers social networks represented as network flow models in which actors have the ability to provide some level of influence over other actors within the network. The models developed incorporate performance metrics and reliability analysis established in the multi-state reliability literature to gain insights into organizational behavior.

After a brief introduction, Chapter 2 provides a survey of the relevant literature on several topics of interest within this dissertation. In Chapter 3, actor criticality findings using traditional social network analysis are compared to those obtained via multi-state reliability importance measures. Chapter 4 extends the model developed in Chapter 3 to consider that an actors social interaction and level of influence within the organization are not only multi-valued and stochastic in nature but also a function of the interactions with its neighbors. A Monte Carlo simulation model is presented to evaluate the reliability of the network, and network reliability is evaluated under various influence communication rules. In Chapter 5, a hierarchical network structure is investigated where actors are arranged in layers and communication exists between layers. A probability mass function is developed to compute the expected level of influence at the target nodes as a function of the existing communication paths within the network. An illustrative example is used to demonstrate the effects on expected influence at the target as connections are either added or removed and when the uncertainty associated with an actor's influence level is removed. Finally, in Chapter 6, a methodology is developed for eliciting the probabilities associated with the influence levels used in the network analysis of Chapters 3 - 5.

This dissertation is approved for recommendation
to the Graduate Council

Dissertation Co-Directors:

Edward A. Pohl, Ph.D.

Chase Rainwater, Ph.D.

Dissertation Committee:

Justin R. Chimka, Ph.D.

Jose Emmanuel Ramirez-Marquez, Ph.D.

Dissertation Duplication Release

I hereby authorize the University of Arkansas Libraries to duplicate this dissertation when needed for research and/or scholarship.

Agreed ______________________________________

Kellie R. Schneider

Refused ______________________________________

Kellie R. Schneider

Acknowledgments

"How lucky I am to have something that makes saying goodbye so hard."

- A.A. Milne, Winnie-the-Pooh

I am lucky.

I am lucky to have exceptional advisors and committee members. I would like express my sincerest gratitude to Dr. Ed Pohl and Dr. Chase Rainwater. Without your guidance and dedication, this may have never happened. I truly appreciate how freely you gave of your time during the past two years. I would also like to thank my committee members Dr. Justin Chimka and Dr. Jose Ramirez-Marquez for your support and insightful comments that contributed to the quality of my work.

I am lucky to have supportive co-workers. I would like to thank Dr. Richard Cassady for employing me for the past twelve years. I am very grateful for the various opportunities you provided me for professional and personal growth. Also, I would like to thank Dr. Heath Schluterman and Dr. Candace Rainwater for allowing me the flexibility in my work schedule to complete my degree. It has been a pleasure to work with you, and I will miss you both tremendously.

I am lucky to have beautiful friends. It's interesting how your circle of friends ebbs and flows over the years, and I can say that I am blessed to have a number of friends who have been constants in my life. You are all very dear to me, and I love you very much.

I am lucky to have an amazing family. We've been through a lot together, and I believe we are closer now than we have ever been. Words cannot describe how grateful I am for the unwavering support of my parents, Mark and Gail Garner, and my grandparents, Ruby and Earl Howard. Thank you so much for always believing in me especially during the times when I didn't believe in myself.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Social networks have long been used to model the interactions between people in various social and professional contexts. These networks allow for the explicit modeling of the complex interrelations between relevant individuals within an organization and the role they play in the decision making process. Recently, operations researchers have shown increased interest in the quantitative study of social networks particularly in the area of clandestine organizations and terrorist networks. Network flow models have been of particular interest [21, 22, 42]. The primary focus of this dissertation is on the quantification of actor interaction and the dissemination of information through social networks. In this work, social networks are represented as network flow models in which actors have the ability to provide some level of influence over other actors within the network. The models developed incorporate performance metrics and reliability analysis established in the multi-state reliability literature to gain insights into organizational behavior. In addition, a methodology is adopted for eliciting the probabilities associated with the influence levels used in the network analysis.

After a brief introduction, Chapter 2 provides a survey of the relevant literature on several topics of interest within this dissertation. First, a brief overview of graph theory and its application to social networks is presented followed by a review of the literature in which clandestine organizations are modeled using social networks. Particular emphasis is placed on the mapping between network flow models and the terminology associated with social networks. Considering network flow models of social networks along with varying levels of influence associated with actors in the network lends itself to established multi-state reliability models. Therefore, a review of the multi-state reliability literature is provided. Particular focus is given to the development of multi-state system models, computation of multi-state performance metrics, and optimization techniques associated with multi-state systems.

In Chapter 3, a social network in which actors have the ability to provide some level of influence over other actors in the network is considered. The network is examined using the traditional social network analysis metrics of actor centrality to identify the most critical members. The network is then evaluated using a multi-state network flow formulation. The actors within the network are represented as multi-state components and the level of influence associated with an actor is represented as a component state. Treating influence as a commodity that flows through the network allows the network reliability to be defined as the probability that a required demand level (amount of influence) reaches an intended target. In the reliability literature, this quantity is referred to as multi-state two-terminal reliability, $M2TR$. The computation of $M2TR$ facilitates the computation of reliability importance measures for multi-state systems [35]. These importance measures are used to identify the actors that have the most impact on network reliability. The actor criticality findings obtained through traditional social network analysis are compared to those obtained via multi-state reliability importance measures. Finally, a sensitivity analysis is performed on the probability distributions associated with the influence levels of the most important actors identified via the multi-state reliability importance measures.

In Chapter 4, the modeling paradigms to evaluate the exchange and propagation of influence within a social network are presented. The models in this chapter differ from traditional social networks in that actor influence is considered to be multi-valued, stochastic, and a function of an actors interactions with others in the network. A set of influence rules are developed to describe the manner in which a particular actor will pass influence to other actors in the network. An influence campaign is initiated to a set of initially accessible actors and influence is treated as commodity that flows through the network. The success of the influence campaign is determined by the amount of influence that reaches a set of target actors. Exhaustive enumeration is used to evaluate the reliability of a small, illustrative example, and a Monte Carlo simulation model is developed to efficiently approximate the network reliability of larger networks. Two additional performance metrics are defined to further evaluate the propagation of influence through a network. These metrics focus on the extent to which individual actors within the network are influenced and allow decision-makers the ability to assess the impacts on target nodes when influence is considered as a limited resource.

In Chapter 5, a hierarchical social network is investigated where actors are arranged in levels and communication occurs between actors in subsequent levels. A closed-form expression for the probability mass function (PMF) for the various actor influence levels is presented as a function of the number of preceding actors. The PMFs are then used to compute the expected influence level of target actors. An illustrative example is presented, and the expected influence values under two of the influence rules developed in Chapter 4 are investigated to gain insights into the effects of altering communication paths or eliminating the uncertainty associated with actors within the network.

In Chapters 3 - 5, levels of actor influence are defined, and the probabilities associated with an actor being in each influence level is given. It is assumed that both the influence levels and the probability associated with an actor having a particular level of influence have been determined by experts familiar with the organization under study. However, there is a lack of relevant literature for identifying influence levels and/or their probability distributions for social networks. In Chapter 6, a method for eliciting influence probabilities within a college classroom environment is developed. It is worthwhile to note that the methodology developed in this chapter would not apply directly to a clandestine or terrorist organization since those organizations, by definition, attempt to hide their social structure. However, this methodology does provide a baseline for quantifying perceived influence levels and the probabilities associated with actors being in each level.

Chapter 2

Literature Review

This chapter provides a survey of the relevant literature on several topics of interest within this dissertation. First, a brief overview of graph theory and its application to social networks is presented followed by a review of the literature in which clandestine organizations are modeled using social networks. Particular emphasis is placed on the mapping between network flow models and the terminology associated with social networks. Considering network flow models of social networks along with varying levels of influence associated with actors in the network lends itself to established multi-state reliability models. Therefore, a review of the multi-state reliability literature is provided. Particular focus is given to the development of multi-state system models, computation of multi-state performance metrics, and optimization techniques associated with multi-state systems.

2.1. Social Networks and Graph Theory

Social networks have long been used to model the interactions between people in various social and professional contexts. The sociogram is a visual representation of a social network where the nodes of a graph represent the actors in the network and the edges represent the relationships between the actors. This representation allows for the use of graph theory to analyze the social network [23].

Below is a brief review of basic graph theory. For a more complete review of the graph theory literature see Hamill [22] and the included references. Consider a social network modeled as an undirected graph. Let $V = (v_1, v_2, \ldots, v_n)$ denote the set of actors (or nodes) in the network, and let $E = (e_1, e_2, \ldots, e_m)$ denote the number of relationships (or edges) between the actors. The size of the network is defined as the number of actors within the network, n. A fully connected graph contains $(n * (n - 1))$ connections, and the density of a network is defined as the proportion of actual connections between actors to the number of possible connections between the actors in the network. The structural position of an actor and his relationships with other actors within the social network allows for an analysis based on measures of power and centrality. The relationships depicted in a sociogram can be represented in matrix form. An adjacency matrix represents the connectivity of actors in the relationship of interest, and based on the adjacency matrix, the degree of each actor is computed as the number of edges incident to a node. An actor with a higher degree is more central to the network structure and is generally viewed as being in a more favored position and having more power [23]. Measures of centrality and betweenness describe the location of an actor based on his closeness to the center of the social network. One of the most popular metrics in social network analysis (SNA) is actor centrality. In general, an actor's centrality is a measure of his visibility or importance within the social network. The concept of centrality was formally defined by Freeman [18]. Specifically, Freeman identified three primary centrality measures: degree, closeness, and betweenness. Degree centrality measures an actor's direct connectedness with other actors. Let $deg(v)$ denote the degree of actor v, and the degree centrality of an actor, denoted by $C_D(v)$, is given by

$$C_D(v) = \frac{deg(v)}{n-1} \tag{2.1}$$

Closeness centrality provides a more global network prospective than degree centrality. Specifically, closeness centrality is a measure that indicates the extent to which an actor is near all the other actors in the network, not just those adjacent to them. Let $d_G(b, c)$ denote the length of shortest path connecting actor b with actor c, so that the closeness centrality of an actor, denoted by $C_C(v)$, is given by

$$C_C(v) = \frac{1}{\sum_{c \in V} d_G(b, c)} \tag{2.2}$$

Betweenness centrality is a measure of the strategic location of an actor along a potential communication path. Let σ_{bc} denote the number of shortest paths from actor b to actor c, and let $\sigma_{bc}(v)$ denote the number of shortest paths from actor b to actor c that contain actor v. The betweenness centrality of an actor, denoted by $C_B(v)$, is given by

$$C_B(v) = \sum_{b \neq v \neq c \in V} \frac{\sigma_{bc}(v)}{\sigma_{bc}} \tag{2.3}$$

Other traditional SNA techniques include hierarchical clustering, multidimensional scaling and correspondence analysis. Hierarchical clustering can be used to identify the different types of individuals within a network. This technique attempts to seek out subsets of actors within a network. These subsets are formed by either cohesion or equivalence. Cohesion results in the forming of cliques (or groups with common interests) in which actors are completely interconnected. Equivalence results in grouping actors that share common relationships even in the absence of direct connections (for example, students and teachers). These groups are then formed into clusters such that groups are internally homogeneous and as heterogeneous as possible between the groups [13]. Multidimensional scaling and correspondence analysis are techniques that can be used to create a visual image of a social network. Multidimensional scaling is used to reveal patterns within a social network's structure. A matrix of social proximity values are scaled and used to determine the optimal location of actors (points) on a graph (generally in two or three dimensions) [8]. Correspondence analysis is a multivariate statistical technique that is used to provide a visual display of affiliation networks. Affiliation networks involve not only individuals (actors) but also various types of social interactions (events) [8].

2.2. Social Networks and Clandestine Organizations

Since the tragic events of September 11, 2001, operations researchers and government entities have shown increased interest in using social network analysis to understand various terrorist networks [9, 12, 19, 21, 22, 30, 42]. McCormick [30] provides a review of the terrorist decision making literature where terrorism is defined as "the deliberate use of symbolic violence or the threat of violence against non-combatants for political purposes." Theories associated with terrorist decision making are categorized as strategic, organizational, and psychological. Strategic theories consider the parties involved to be in a game where the objective is to win given the circumstances and strategies of other players. Organizational theories describe the ways in which groups become clandestine in nature. In general, groups are impatient for results, have a deliberate intention of pursuing violence, and separate from mainstream society to maintain their clandestine existence. Psychological theories attempt to investigate the psychological traits of terrorists. Unfortunately no specific profile of a terrorist exists, and most terrorists are described as "disturbingly normal" [30]. More recently, the RAND corporation provides a review of the counterterrorism literature [12]. Of particular interest to this dissertation is the chapter that focuses on the organizational structure of clandestine organizations and the analytical efforts to map the various functions of a terrorist organization. In general, the analytic efforts to map these organizations has shifted to a more "network centric" approach. The use of social networks allows researchers to focus on the *actual* interpersonal connections between actors in the network as opposed to connections dictated by an organizational chart. However clandestine and terrorist organizations, by nature, do not lend themselves to traditional SNA techniques. Most of these traditional techniques rely on members within the organization of interest to provide information regarding both the actors of interest and their various affiliations [12]. Therefore, researchers are developing new techniques to describe and analyze organizations that attempt to hide their membership and/or organizational structure.

2.2.1 Network Flow Models and Influence

Network flow models are of particular interest in the quantitative analysis of social networks. Renfro and Deckro [42] provide a mapping of the relationships between traditional social network terminology and that associated with network flow problems. The actors within the organization are represented graphically as nodes, and their associations and interactions correspond to capacitated arcs between the nodes. In general, the number of arcs incident to a node relates to an actor's importance within the network, and the capacity of the arcs between actors is a measure of the 'social closeness' between two individuals. Social closeness quantifies the potential influence one actor within the organization has on another actor. Influence is defined as the ability of an actor "to induce a change in behavior of another that conforms to the influencing actor's desires... [22]." In the network flow formulation of a social network, influence is modeled as a commodity that travels from the network source (or actor initiating the influence) to the network sink (or the target of the influence). A summary of this mapping is provided in Table 2.1.

Social Network Terminology	**Network Flow Terminology**
Actors (people)	Nodes (sinks, sources, or transshipments)
Affinity (connectivity)	Capacitated arcs between nodes
Social Closeness	Capacity
Influence	Commodity
Potential Influence	Magnitude of Flow
Initiators of Influence	Source(s)
Targets to be influenced	Sinks(s)
Intermediaries involved	Transshipment node(s)

Table 2.1: Social Network and Network Flow Relationships

Hamill models the scenario in which a distinct communication path exists in the network and target members within the network are not easily accessible [21, 22]. In his work, Hamill considers the amount of 'influence' passed between actors within the network is modeled as a commodity. In general, Hamill assumes the most important actors and thus the intended targets of influence within a network are the organizational leaders, and those leaders are generally inaccessible to influences outside of the organizational network (for example, Osama bin Laden of Al Qaeda). However, some members of the organization are, in fact, accessible and can be used to initiate a course of action to influence the group leaders. In the network flow representation of this scenario, influence is initiated to accessible members of the network via a source node and has reached its intended target when influence flows into the sink node. Hamill [21] presents a review of the generalized network flow problem (GNF) presented by Ahuja et al. [1]. The GNF model is given below.

Let N denote the set of nodes within a given network, and let A denote the set of arcs. The cost per unit flow from node i to node j is given by $c_{(i,j)}$, and the amount of flow is given by $x_{(i,j)}$ with a flow capacity on arc (i,j) of $u_{(i,j)}$. Let b_i be a demand variable such that $b_i = 0$ indicates a transshipment node, $b_i < 0$ indicates demand is required by node i, and $b_i > 0$ indicates supply is provided by node i. Finally, let $g_{(i,j)}$ quantify the gains or losses associated with an arc where $g_{(i,j)} > 1$ indicates a gain on an arc, $g_{(i,j)} < 1$ indicates loss on an arc, and $g_{(i,j)} = 1$ indicates the arc is neither. The mathematical formulation of the model is given by:

$$\text{minimize} \sum_{(i,j)\in A} c_{(i,j)} x_{(i,j)}$$

subject to (GNF)

$$\sum_{j:(i,j)\in A} x_{(i,j)} - \sum_{j:(j,i)\in A} g_{(j,i)} x_{(j,i)} \geq b_i \quad \forall i \in N$$

$$0 \leq x_{(i,j)} \leq u_{(i,j)} \quad \forall (i,j) \in A$$

Hamill is able to quantify many aspects of social network behavior using the GNF model. For example, it is possible for actors within an organization to exhibit a high degree of social closeness, and that does not accurately reflect the amount of influence one actor may have over the other. An example used by Hamill is that of a father and son. While they are socially close, the father generally has more influence over the son. Hamill considers the use of gains and losses to describe these types of situations in which influence is inequitable between actors within the network by using an arc multiplier, $g_{(i,j)}$. Each arc within the network is assigned an arc multiplier. A multiplier greater than one indicates a positive influence from one actor to another, while a multiplier between zero and one indicates a degradation or loss of influence [21].

Hamill also considers that each actor may have a threshold value that must be satisfied before the actor disseminates the information to other actors within the network. Some actors do not hesitate to accept influence and pass information along to others, i.e. $b_i = 0$, while some individuals must receive some given amount of influence, i.e. $b_i < 0$. Others must authenticate a message using some number of independent reports. Hamill refers to these individuals as conditional gatekeepers. An example of a conditional gatekeeper is shown in Figure 2.1 [21].

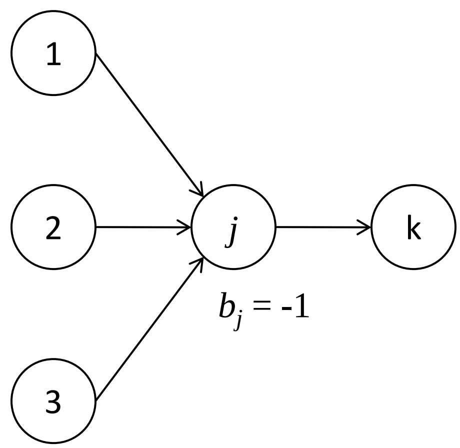


Figure 2.1: Conditional Gatekeeper [21]

Finally, Hamill quantifies the costs associated with communicating influence. These costs include not only actual costs associated with transmitting information but also the assessment of an actor's own risks as well as potential penalties associated with spreading rumors or propaganda throughout the network. Hamill provides a notional example and uses the LP formulation of the GNF with post-optimality analysis to investigate the uncertainty in the input data [21].

In this work, social networks are represented as network flow models in which actors have the ability to provide some level of influence over other actors in the network. In general "influence is exerted by some members over others and that scope of influence varies" [12]. Therefore, multiple levels of influence are considered. From a modeling perspective, the actors in the network are analogous to system components and the level of influence of each actor are comparable to component states. Therefore, the established multi-state reliability models are used to further model social networks. A review of the multi-state reliability literature is provided below.

2.3. Multi-State Reliability

In traditional reliability theory, systems are modeled as having binary state components that are either functioning or failed. When this assumption reasonably describes the system of interest, a well-developed theory exists for evaluating the reliability of such systems [14]. However, most systems are more complex in nature, and system components are often in a state somewhere between fully functioning and completely failed. For these systems, the theory of multi-state reliability must be considered for modeling and evaluating the effectiveness of system function.

2.3.1 Modeling Multi-State Systems

Work in the field of modeling and computing reliability for systems with components that exhibit multi-state behavior began appearing in the literature in the late 1970s and early 1980s. The seminal works in the area are presented by Barlow and Wu [3], El-Neweihi et al. [15] and Griffith [20]. Soon after, Natvig provides suggestions for defining multi-state systems [31] and presents a case study of an electric power generation application that can be described by the model [32]. Hudson and Kapur [24] develop a generalized model to quantify the behavior of multi-state systems. The model is given below.

Let n denote the number of components in a multi-state system, and let M_i denote the state of component i. A component state of M_i indicates a perfectly functioning component while a state of 0 indicates a component is completely failed. The probabilities associated with a component occupying a particular state is given by $p_{ij} = \Pr(x_i = j)$, and $P_{ij} = \Pr(x_i \leq j)$. The component state vector is given by $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, and the system structure function is given by $\phi(\mathbf{x})$ [24].

Aven [2] presents two algorithms using minimum cut sets and minimum path sets to compute an exact multi-state system reliability. Boedigheimer and Kapur [4] extend the model of Hudson and Kapur [24] by defining a customer-driven structure function to quantify the possible system states. They suggest two reliability measures for evaluating system performance. One measure is used to compute the expected value of the utility (assuming it can be quantified) of a particular state vector, and the other involves computing the probability that a particular state vector results in at least some threshold value of demand. The second measure leads to the development of multi-state two-terminal reliability models where the system reliability is based on the probability that system capacity can satisfy a given demand level.

2.3.2 Multi-State Two-Terminal Reliability

Ramirez-Marquez et al. [37] extend traditional two-terminal reliability models by incorporating multi-state components within a system. In multi-state two-terminal reliability models, a system is represented as a network with capacitated arcs which represent the system states of the components within the system. System reliability is defined as the probability that the system capacity in a given state is greater than or equal to some demand value, d, i.e. $M2TR_d = \Pr(\phi(\mathbf{x}) \geq d)$.

2.3.3 Solution Procedures

Several authors have developed methods for computing performance metrics associated multi-state two-terminal reliability. Fishman [17] develops a Monte-Carlo sampling plan to estimate the distribution on the maximum flow of a system that is subject to random amounts of degradation at random points in time. Lin et al. [29] provide an implicit enumeration method to generate all minimum paths that satisfy a given demand level (d-MPs). However, the implicit enumeration method requires a priori knowledge of all binary minimal paths. Ramirez et al. [37] present an algorithm to generate all multi-state minimum path vectors without requiring a priori knowledge of the system binary minimum paths. Once multi-state minimum path vectors (or d-MPs) have been identified, system reliability can be computed using traditional multi-state reliability approaches including inclusion-exclusion [15] or disjoint subsets [25]. Zuo et al. [48] develop an algorithm known as Recursive Sum of Disjoint Products to evaluate network reliability. The algorithm makes use of a special maximum operator when evaluating the various d-MPs which significantly reduces the computations required for obtaining $M2TR_d$ as long as system state vectors tend to dominate one another. Ramirez-Marquez and Coit [35] present a Monte-Carlo simulation approach for approximating the reliability of a multi-state network. Satitsatian and Kapur [43] define an algorithm to determine a subset of lower boundary points that can be used to compute a lower bound on system reliability. Results from their numerical examples show that it is possible, in some cases, to obtain exact system reliabilities without a priori knowledge of all binary minimum cuts. Jane and Laih [26, 27] present direct decomposition methods to determine the exact multi-state system reliability. The first method allows for the computation of $M2TR_d$ exactly without a priori knowledge of all minimum paths, while the second method allows for the exact computation of $1 - M2TR_d$ without a priori knowledge of all minimum cuts. Finally, Jane and Laih [28] modify their original method [26] by incorporating bounds so the user of the algorithm may determine an appropriate trade-off between the accuracy of the reliability computation and algorithm runtime.

2.3.4 Importance Measures

Ramirez-Marquez and Coit [35] present general composite importance measures (CIM) by reformulating traditional reliability importance measures to include component states. Traditional importance measures include Birnbaum, Reliability Achievement Worth, Fussell-Vesley, and Reliability Reduction Worth. Birnbaum importance (or Average of the Sum of Absolute Deviations SAD) provides the probability that a component is critical to the functioning of the overall system. In the social network context, this is analogous to identifying the actor that is most critical in influencing the intended target and is given by:

$$MI_I^{SAD} = \frac{\sum_{j=1}^{\omega_i} |\Pr(\phi(x) \geq d | x_i = b_{ij}) - \Pr\phi(x) \geq d|}{\omega_i - 1} \tag{2.4}$$

Multi-State Reliability Achievement Worth ($MRAW$) provides the maximum percent increase in system reliability generated by a particular component. That is, $MRAW$ is a measure of an actors worth in achieving the current system reliability and indicates the importance of maintaining the current level of influence for the actor [40]. The importance measure is given by:

$$MRAW_i = 1 + \frac{1}{\omega_i - 1} \sum_{j=1}^{\omega_i} \max(0, \beta_{ij}) \tag{2.5}$$

where

$$\beta_{ij} = \frac{\Pr(\phi(x) \geq d | x_i = b_{ij}) - \Pr(\phi(x) \geq d)}{\Pr(\phi(x) \geq d)} \tag{2.6}$$

Multi-State Fussell-Vesley (MFV) provides the maximum decrement in system reliability that can be attributed to a particular component. This measure takes into account that an actor may contribute to system failure without being critical [40] and is given by:

$$MFV_i = \frac{1}{\omega_i - 1} \sum_{j=1}^{\omega_i} \max(0, -\beta_{ij}) \tag{2.7}$$

Multi-State Reliability Reduction Worth ($MRRW$) quantifies the potential damage caused to the system by a particular component and is given by

$$MRRW_i = 1 + \frac{1}{1 - MFV_i} \tag{2.8}$$

The CIM above are applicable to multistate systems, but they only involve the actor states and not the probability, p_{ij}, of actor i being in state j. Ramirez-Marquez and Coit [35] also present a set of alternative CIM that incorporate state probabilities. Mean Absolute Deviation is analogous to Birnbaum CIM and measures the expected absolute deviation in system reliability caused by an actor's probability of possessing a particular level of influence. Mean Multi-State Reliability Achievement Worth and Mean Multi-State Fussell-Vesley are presented using the same logic. The alternative CIM are given by

Birnbaum (MAD)

$$MAD_i = \sum_j p_{ij} |\Pr(\phi(x) \geq d | x_i = b_{ij}) - \Pr(\phi(x) \geq d)| \tag{2.9}$$

Reliability Achievement Worth ($MMAW$)

$$MMAW_i = 1 + \sum_{j=1}^{\omega_i} p_{ij} \max(0, \beta_{ij}) \tag{2.10}$$

Fussell-Vesley ($MMFV$)

$$MMFV_i = \sum_{j=1}^{\omega_i} p_{ij} \max(0, -\beta_{ij}) \tag{2.11}$$

Chapter 3

Investigating Actor Importance in a Multi-State Social Network

3.1. Introduction

Traditional social network analysis (SNA) allows researchers to investigate the interactions of actors within a social network to determine an actor's centrality, or relative power, within a network. Recently, operations researchers have shown increased interest in the quantitative study of social networks to determine the most critical members of terrorist networks. Renfro and Deckro [42] provide a mapping of the relationships between clandestine social networks and network flow models. These relationships are shown in Table 2.1 and discussed in detail in section 2.2.2. Recall that actors within the organization are represented graphically as nodes, and influence is modeled as a commodity that flows along capacitated arcs between the nodes. Many aspects of social network behavior may be quantified using the generalized network flow (GNF) model. The GNF model allows for gains and losses within the network which is analogous to communication problems or actors within a network having different levels of persuasion [22, 21].

In this chapter, a social network model is analyzed using the traditional SNA technique of actor centrality. Specifically, degree, closeness, and betweenness centrality are considered. Degree centrality measures an actor's direct connectedness with other actors, while closeness centrality provides a more global perspective. Specifically, closeness centrality is a measure that indicates the extent to which an actor is near all other actors in the network not just those adjacent to them. Betweenness centrality is a measure of the strategic location of an actor along a potential communication path [18]. Unfortunately, traditional SNA metrics do not take into account that, despite their position within the network, some actors may be more influential or important than other actors.

In his work, Hamill [21, 22] considers "influence" as a commodity that flows through a social network. Influence is initiated to accessible members of the network via a source node and has reached its intended target when influence flows into the sink node. In this work, social networks are represented as network flow models in which actors have the ability to provide some level of influence over other actors in the network. From a modeling perspective, the actors in the network are analogous to system components and the levels of influence of each actor are comparable to component states. Therefore, established multi-state reliability models are used to further model and analyze social networks. Treating influence as a commodity that flows through the network allows the network reliability to be defined as the probability that a required level of demand (amount of influence) reaches an intended target.

In the reliability literature, this quantity is known as multi-state two-terminal reliability ($M2TR_d$). To solve these types of problems, all minimal path vectors that satisfy the flow demand (often referred to as d-MPs) must be identified. Lin et al. [29] provide a framework for identifying all d-MPs. Their method requires the identification of all binary minimum paths and uses a problem-specific implicit enumeration method to elicit the d-MPs. Recently, Ramirez-Marquez et al. [37] developed an algorithm that allows for the computation of all multistate minimum cut vectors but does not require a priori knowledge of the binary minimum paths.

Generally, the $M2TR_d$ computation involves the use of an inclusion/exclusion formula, and Zuo et al. [48] provide an algorithm known as Recursive Sum of Disjoint Products to evaluate the network reliability. Zuo implements a special maximum operator when evaluating the various d-MPs which significantly reduces the computations required for obtaining $M2TR_d$ as long as state vectors have a tendency to dominate one another. When this is not the case, however, the computations associated with $M2TR_d$ are non-trivial due to the probability calculations associated with the union of dependent events. Ramirez-Marquez and Coit [36] develop a Monte-Carlo simulation approach that can be used to compute $M2TR_d$

and other quantities associated with multi-state reliability importance measures. Specifically, the simulation approach allows for the computation of the conditional $M2TR_d$ of the system given an actor is in a particular state within the network, i.e. $\Pr(\phi(x) \geq d | x_i = b_{ij})$.

Once the system reliability has been determined, multi-state reliability importance measures provide additional insights into actor criticality. Ramirez-Marquez and Coit [36] present general composite importance measures (CIM) by reformulating traditional reliability importance measures. A complete review of the metrics along with the corresponding equations is provided in Chapter 2. Birnbaum importance (MI_i^{SAD}) provides the probability that a component is critical to the functioning of the overall system. In the social network context, this is analogous to identifying the actor that is most critical in influencing the intended target. Multi-State Reliability Achievement Worth ($MRAW$) provides a measure of an actor's worth in achieving the current system reliability and indicates the importance of maintaining the current level of influence for the actor. Multi-State Fussell-Vesley (MFV) provides the maximum decrement in system reliability that can be attributed to a particular component, and Multi-State Reliability Reduction Worth ($MRRW$) quantifies the potential damage caused to the system by a particular component. Ramirez-Marquez and Coit [36] also develop a set of alternative CIM that incorporate the probabilities associated with a component (actor) being in a particular state. These include Mean Absolute Deviation (MAD), Mean Multi-State Reliability Achievement Worth ($MMAW$) and Mean Multi-State Fussell-Vesley ($MMFV$).

In the remainder of the chapter, a notional network is presented and an illustrative example is used to compare actor criticality findings using traditional SNA and those found using multi-state importance measures.

3.2. Model Development

Consider a social network modeled as graph, $G(V, E)$. The size of the network is defined as the number of actors within the network, n. Let $V = (v_1, v_2, \ldots, v_n)$ denote the set of actors in the network, and let $E = (e_1, e_2, \ldots, e_m)$ denote the number of relationships (or edges) between the actors. Let ω_i denote the number of states of influence for actor i and M_i denote the maximum influence of actor i. Let $\mathbf{A}$ denote a connectivity matrix of a social network where element a_{ij} of $\mathbf{A}$ is binary valued as follows:

$$a_{ij} = \begin{cases} 1 & \text{if actor } i \text{ is connected to actor } j \\ 0 & \text{otherwise} \end{cases} \qquad \forall i = 1, 2, \ldots, n; j = 1, 2, \ldots, n; i \neq j$$

Let $\mathbf{B}$ denote the influence vector for the network where element b_{ij} denotes the j^{th} influence state for actor i, and let p_{ij}, denote the probability that actor i is in state j, $p_{ij} = \Pr(x_i = b_{ij})$. The current state of actor i is defined as x_i, and the system state vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ denotes the state of all the actors in the network. Let ϕ denote the system structure function which is defined as the network flow given a particular state vector, $\mathbf{x}$. In this case, it is interesting to evaluate the probability that the flow of influence that reaches the target actor(s) is greater than or equal to some threshold value, d, i.e. $\Pr(\phi(x) \geq d)$. In the reliability literature, this quantity is known as multi-state two-terminal reliability ($M2TR_d$).

To solve this type of problem, all the minimal path vectors that satisfy the flow demand (often referred to as d-MPs) must be identified and used compute to $M2TR_d$. The framework provided by Lin et al. [29] is used in this chapter to identify all d-MPs for a given network. This method requires the identification of all binary minimum paths and uses a problem-specific implicit enumeration method to elicit the d-MPs. Let $P_s = 1, 2, \ldots, k$ denote the number of minimum paths in graph $G(V, E)$ where the maximum capacity of each path is given by L_s such that $L_s = \min(M_i - 1 | v_i \in P_s)$. The set of feasible solutions $(f_1, f_2, \ldots, f_k)$ satisfying these three constraints

$$\sum_{s=1}^{k} f_s = d$$

$$f_s \leq L_s \quad \forall s = 1, 2, \ldots, k$$

$$\sum_{s=1}^{k} (f_s | v_i \in P_s) \leq M_i \quad \forall i = 1, 2, \ldots, n$$

are determined using implicit enumeration. Lin et al. [29] prove that if the network is acyclic, then all candidate solutions are d-MPs. Additional post processing is required for cyclic networks to ensure that a candidate solution does not dominate any of the other candidate solutions [29]. Although the technique is relatively easy to implement, it does require a priori knowledge of all binary minimum paths, and the enumeration method is problem specific. Therefore, any change in network topology requires that the implicit enumeration constraints be updated to reflect the change in network structure, and the entire problem must be solved again. Once all d-MPs have been identified, the Monte-Carlo simulation approach developed by Ramirez-Marquez and Coit [36] is used to compute $M2TR_d$ and conditional $M2TR_d$ for each influence level of each actor. These values are then used to compute the CIM and alternative CIM.

3.3. Illustrative Example

3.3.1 Traditional Social Network Analysis

Consider the modified (undirected) network from Hamill [21, 22] with $n = 11$ actors having the social relationships (or paths of communication) depicted in Figure 3.1. The values associated with the three centrality metrics (degree, closeness, and betweenness) for the network are given below in Table 3.1. The column Betweenness* provides the normalized betweenness values for the actors. The rankings of the actors by performance metric are also shown below. Actors divided with a "/" indicate that the value of the performance metric is the same for these actors.

Degree: 4 – 6/7 – 5/11 – 1/8/9/10 – 2/3

Closeness: $4 - 7 - 5/6 - 11 - 1 - 10 - 2/3/8/9$

Betweenness: $4 - 7 - 6 - 11 - 5 - 1/2/3/8/9/10$

Actor 4 displays the largest centrality values for each measure. Actors 6 and 7 exhibit the second highest degree centrality, and actor 7 exhibits the second highest closeness and betweenness centrality. These rankings provide some insights into the most critical members of the network. However, these metrics treat each actor in the network identically and assume a perfect communication chain. In reality, certain actors within the network may be more persuasive, and communication between actors in the network may not be perfect. In these instances, the centrality metrics may not adequately quantify the criticality of actors within the network.

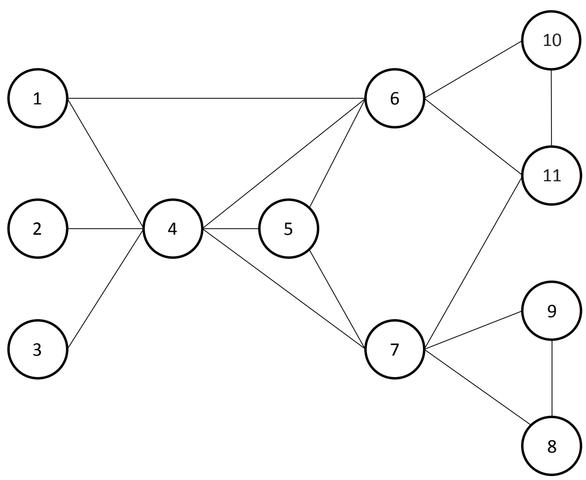


Figure 3.1: Modified Notional Network for the Illustrative Example

Vertex	Degree	Closeness	Betweenness	Betweenness*
1	0.20	0.500	0	0
2	0.10	0.0435	0	0
3	0.10	0.0435	0	0
4	**0.60**	**0.0714**	**43**	**1**
5	0.30	0.0588	2	0.0465
6	**0.50**	0.0588	17	0.3953
7	**0.50**	**0.0667**	**36**	**0.8372**
8	0.20	0.0435	0	0
9	0.20	0.0435	0	0
10	0.20	0.0455	0	0
11	0.30	0.0526	8	0.1860

Table 3.1: Centrality Measures for the Illustrative Example

3.3.2 Reliability Importance Analysis

Let's reconsider consider the original notional network from Hamill [22, 21] with $n = 11$ actors and the directed social relationships depicted in Figure 3.2. Suppose an influence campaign is launched, and the goal of the campaign is to influence actors 9 and 11. However, the initiators of the influence campaign only have direct access to actors 1, 2, and 3. Note that in this scenario, influence will emanate from a source node to the accessible actors (actors 1, 2, and 3) and flow through the network until it reaches the intended targets (actors 9 and 11).

The framework provided by Lin et al. [29] is used to identify all the minimum paths satisfying the system demand or d-MPs. The use of this framework requires the use of the binary minimum paths. Therefore, the focus for this portion of the analysis is only on actors falling on those paths. Figure 3.3 shows the modified network containing only actors that fall on the minimum paths.

For this example, each actor in the network may be in one of four different states detailing his current level of influence within the network (i.e $\omega_i = 4 \ \ \forall i = 1, 2, \dots, n$). The levels of influence and associated probabilities are shown below in Table 3.2. State 0 indicates the actor will not pass influence to its successor, and state 3 indicates the influence is passed

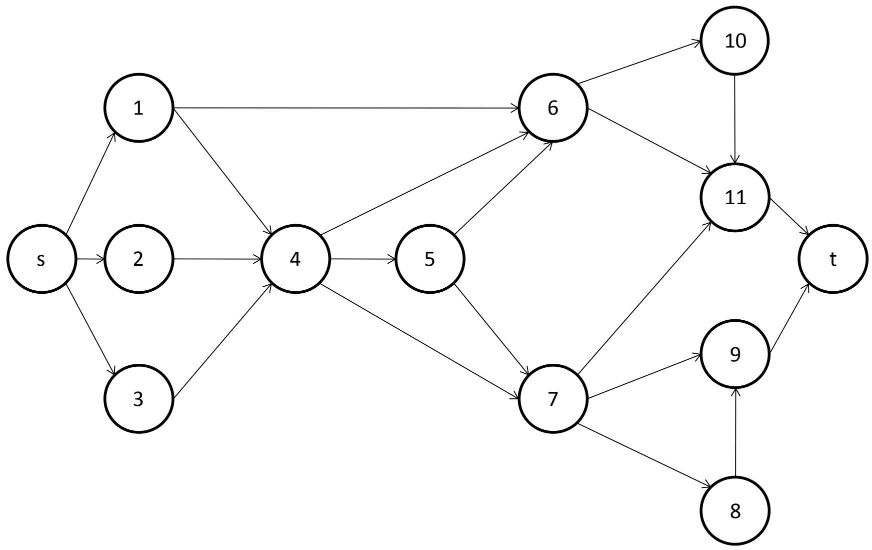


Figure 3.2: Notional Network as a Network Flow Problem for Illustrative Example

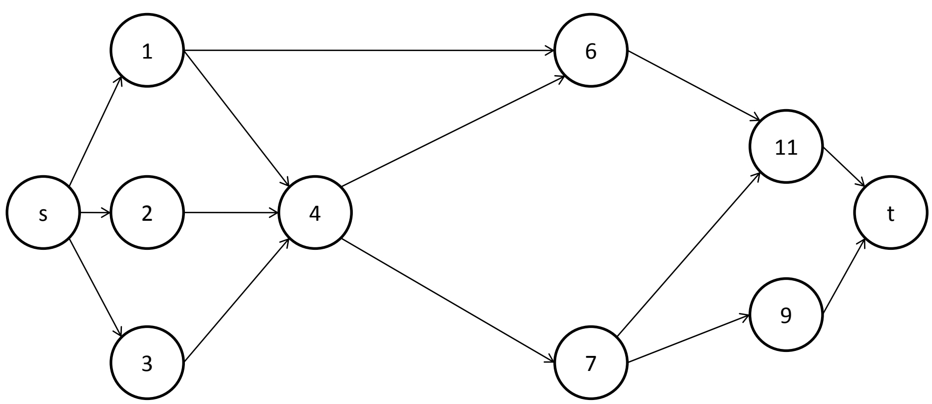


Figure 3.3: Modified Network Depicting Actors on Minimum Paths

exactly as provided. States 1 and 2 provide the possibility for communication problems or actors within the network having different levels of persuasion within the organization. It is assumed that the levels of influence (and the probability associated with an actor having a particular level of influence) have been determined by experts familiar with this organization.

	Level of Influence			
Actor	0	1	2	3
1	0.10	0.10	0.45	0.35
2	0.10	0.10	0.40	0.40
3	0.15	0.20	0.20	0.45
4	0.20	0.05	0.20	0.55
6	0.15	0.10	0.25	0.50
7	0.20	0.05	0.20	0.55

Table 3.2: Actor Influence Probabilities for Illustrative Example

Recall that d-MPs are the minimum paths that satisfy the system demand. For the illustrative example, there are seven 1-MPs, twenty-eight 2-MPs, and eighty-four 3-MPs. Once the d-MPs are identified, the Monte Carlo simulation model developed by Ramirez-Marquez and Coit [36] is used to determine the system reliability given the demand level, d. In the simulation model, random state vectors are generated based on the actor influence probabilities. The state vector is then checked to see if it dominates at least one of the d-MPs. If it does not, then the system is considered failed, and another state vector is generated. For this example, $L = 100{,}000$ state vectors are generated for each demand level. Let Q denote the number of state vectors within the simulation model that result in a failed system. The overall system reliability is then estimated by:

$$\widehat{M2TR}_d = 1 - \frac{Q}{L}$$

The simulation model is also used to compute the conditional $M2TR_d$ of the system given an actor is in a particular state within the network, i.e. $\Pr(\phi(x) \geq d | x_i = b_{ij})$. Point estimates for the overall and conditional system reliabilities are shown below in Table 3.3.

The results for the general CIM are shown in Table 3.4, and the results for the alternative CIM are shown in Table 3.5. To facilitate discussion of the insights obtained from the Illustrative Example, a table of the actor rankings by reliability performance metric and demand level are provided in Table 3.6. Recall that the framework implemented to determine the d-MPs required the identification of the binary minimum paths, so only actors falling on those paths are included in the network analysis.

	Level of Influence			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 1 \mid x_i = b_{ij})$			
1	0.7648	0.9455	0.9443	0.9440
2	0.9065	0.9296	0.9293	0.9293
3	0.9120	0.9274	0.9282	0.9295
4	0.7640	0.9668	0.9672	0.9682
6	0.6282	0.9795	0.9781	0.9797
7	0.8323	0.9495	0.9507	0.9496
	$\Pr(\phi(\mathbf{x}) \geq 1) = 0.9256$			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 2 \mid x_i = b_{ij})$			
1	0.6717	0.7335	0.8931	0.8941
2	0.7929	0.8417	0.8647	0.8646
3	0.8238	0.8493	0.8647	0.8639
4	0.6009	0.7448	0.9304	0.9316
6	0.5367	0.6253	0.9502	0.9503
7	0.7096	0.7903	0.8981	0.8973
	$\Pr(\phi(\mathbf{x}) \geq 2) = 0.8557$			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 3 \mid x_i = b_{ij})$			
1	0.3992	0.5920	0.6428	0.7509
2	0.5439	0.6253	0.6593	0.6734
3	0.5698	0.6463	0.6648	0.6728
4	0.1744	0.5561	0.6867	0.8205
6	0.2551	0.5154	0.5973	0.8238
7	0.4107	0.5971	0.6663	0.7381
	$\Pr(\phi(\mathbf{x}) \geq 3) = 0.6496$			

Table 3.3: System Reliability under Various Levels of Demand

	$d = 1$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	3	0.0726	4	1.0205	3	0.0579	3	1.0615
2	5	0.0102	5	1.0041	5	0.0069	5	1.0069
3	6	0.0073	6	1.0030	6	0.0049	6	1.0049
4	2	0.0957	2	1.0452	2	0.0582	2	1.0618
6	1	0.1526	1	1.0578	1	0.1071	1	1.1199
7	4	0.0554	3	1.0263	4	0.0336	4	1.0348
	$d = 2$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	3	0.1273	4	1.0295	3	0.1193	3	1.1354
2	5	0.0316	5	1.0070	5	0.0299	5	1.0308
3	6	0.0185	6	1.0067	6	0.0149	6	1.0151
4	2	0.1721	2	1.0587	2	0.1425	2	1.1661
6	1	0.2462	1	1.0737	1	0.2140	1	1.2723
7	4	0.0985	3	1.0327	4	0.0824	4	1.0898
	$d = 3$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	3	0.1387	4	1.05200	3	0.1615	3	1.1927
2	5	0.0545	6	1.0172	5	0.0667	5	1.0715
3	6	0.0405	5	1.0197	6	0.0426	6	1.0445
4	**1**	**0.2589**	**1**	**1.1067**	2	0.2918	2	1.4121
6	2	0.2517	2	1.0894	**1**	**0.2981**	**1**	**1.4248**
7	4	0.1322	3	1.0540	4	0.1495	4	1.1758

Table 3.4: General CIM Results

	$d = 1$					
	MAD		*MMAW*		*MMFV*	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	4	0.0329	4	1.0182	4	0.0174
2	5	0.0053	5	1.0036	6	0.0021
3	6	0.0047	6	1.0028	5	0.0022
4	2	0.0661	2	1.0365	2	0.0349
6	1	0.0902	1	1.0492	1	0.0482
7	3	0.0381	3	1.0210	3	0.0202
	$d = 2$					
	MAD		*MMAW*		*MMFV*	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	4	0.0609	4	1.0354	4	0.0358
2	5	0.0148	5	1.0084	5	0.0090
3	6	0.0116	6	1.0064	6	0.0071
4	2	0.1132	2	1.0662	2	0.0660
6	1	0.1418	1	1.0829	1	0.0828
7	3	0.0639	3	1.0366	3	0.0380
	$d = 3$					
	MAD		*MMAW*		*MMFV*	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	4	0.0693	4	1.0546	4	0.0521
2	5	0.0264	6	1.0206	5	0.0200
3	6	0.0261	5	1.0208	6	0.0194
4	1	0.2011	1	1.1561	1	0.1535
6	2	0.1728	2	1.1341	2	0.1319
7	3	0.1024	3	1.0801	4	0.0776

Table 3.5: Alternative CIM Results by Demand Level

Metric	$d = 1$	$d = 2$	$d = 3$
SAD	6-4-1-7-2-3	6-4-1-7-2-3	4-6-1-7-2-3
MRAW	6-4-7-1-2-3	6-4-7-1-2-3	4-6-7-1-3-2
MFV	6-4-1-7-2-3	6-4-1-7-2-3	6-4-1-7-2-3
MMRW	6-4-1-7-2-3	6-4-1-7-2-3	6-4-1-7-2-3
MAD	6-4-7-1-2-3	6-4-7-1-2-3	6-4-1-7-2-3
MMAW	6-4-7-1-3-2	6-4-1-7-3-2	4-6-1-7-3-2
MMFV	6-4-7-1-2-3	6-4-7-1-2-3	6-4-7-1-2-3

Table 3.6: Actor Rankings by Importance Metric and Demand Level

3.3.3 Illustrative Example Insights

Traditional social network analysis focuses primarily on the centrality and connectedness of actors within a social network. Based on the results from Table 3.1, actor 4 received the highest ranking for each centrality metric followed by actor 7 and then, perhaps actors 6 and 5. From a traditional analysis standpoint, actors 4 and 7 are the two most important actors within the network. Therefore, if the initiators of the influence campaign have the ability to allocate resources for improving the influence level of actors within the network, they may choose, based on this analysis to focus their efforts on these actors. By incorporating multi-state reliability importance measures in this type of analysis, decision makers may also consider both the various levels of influence for an actor as well as the probabilities associated with being at the various levels of influence. In this example, using the reliability importance measures to rank the actors does, indeed, provide different results from traditional social network centrality measures.

For scenarios in which the demand level at the target is relatively low, i.e. $d = 1$, and $d = 2$, all of the reliability importance measures identify actor 6 as the most important actor. This is a result of the location of this actor within the network (directly connected to a target of influence) as well as the probability distribution associated with its levels of influence. Actor 6 is more likely to be able to provide one or two units of influence to a target actor than actor 7. At the higher demand level, $d = 3$, actor 4 is identified as the most critical actor for the measures related to overall system function and achieving the current system state (SAD and $MRAW$). This result is expected because 67% of the 3-MPs require actor 4 to be in influence level 3. For the measures quantifying potential for system decrement (MFV and $MMRW$), actor 6 is identified as the most critical. It is also interesting to note that actor 1 is occasionally ranked third in this analysis because in the traditional analysis it was never ranked higher than fourth.

3.3.4 Sensitivity Analysis of Actor Influence Probabilities

To explore the impact of an actor's influence level on identifying the most critical actors in a social network, a sensitivity analysis is performed on the probabilities associated with the influence levels for actors 4 and 6. For the "-low" experiments the actor is most likely to be in influence level 0, and for the "-high" experiments, the actor is most likely in influence level 3. These experiments and their associated influence probabilities are shown below in Table 3.7. The remaining actor influence probabilities for each experiment are from Table 3.2. For each experiment, the system reliabilities under the various demand levels, the general CIM results, the alternative CIM results, and importance rankings are available in the Appendix. To facilitate the discussion of these results, the actor rankings by reliability performance metric and demand level are provided in Table 3.8.

		Influence Level			
Experiment	Actor	0	1	2	3
4-low	4	0.85	0.05	0.05	0.05
4-high	4	0.05	0.05	0.05	0.85
6-low	6	0.85	0.05	0.05	0.05
6-high	6	0.05	0.05	0.05	0.85

Table 3.7: Sensitivity Analysis for Critical Actors

Since the experiments involved a wide swing across the influence probabilities of the two most important actors, a more dramatic shift in the rankings of the actors was expected. However, due to the size and structure of the network, actors 4 and 6 are still repeatedly identified as the most important actors. In the 4-low experiment, actor 4 (who is highly central in the network) has a high probability of being in the lowest influence state of 0. At demand levels $d = 1$ and $d = 2$, actor 6 is identified as the most important actor for each metric except for $MRAW$ which identifies actor 4 as most important. It is also interesting to note that actor 1 is identified as the second most important actor for each metric except $MRAW$. In the high demand case, $d = 3$, actor 1 is identified as the most important actor for all alternative CIM results. Similar results are obtained with the 6-high experiments. In

Metric	**4-high**	**4-low**	**6-high**	**6-low**
	$d = 1$			
SAD	6-4-7-1-2-3	6-1-4-7-2-3	6-1-4-7-2-3	4-6-7-2-3-1
MRAW	6-7-4-1-2-3	4-6-1-7-3-2	4-1-6-7-3-2	6-4-7-3-2-1
MFV	6-4-7-1-2-3	6-1-4-7-2-3	6-1-4-7-2-3	4-7-2-6-3-1
MMRW	6-4-7-1-2-3	6-1-4-7-2-3	6-1-4-7-2-3	4-7-2-6-3-1
MAD	6-7-4-1-2-3	6-1-4-7-2-3	4-1-6-7-3-2	4-7-6-3-2-1
MMAW	6-7-4-1-2-3	6-1-4-7-3-2	4-1-6-7-3-2	4-7-6-3-2-1
MMFV	6-7-4-1-2-3	6-4-1-7-2-3	4-1-6-7-2-3	4-7-6-2-3-1
	$d = 2$			
SAD	6-4-7-1-2-3	6-1-4-2-7-3	6-1-4-7-2-3	4-7-6-2-3-1
MRAW	6-7-4-1-3-2	4-6-1-7-3-2	4-1-6-7-2-3	6-4-7-2-3-1
MFV	4-6-7-1-2-3	6-1-4-7-2-3	6-1-4-7-2-3	4-7-2-3-6-1
MMRW	4-6-7-1-2-3	6-1-4-7-2-3	6-1-4-7-2-3	4-2-7-3-6-1
MAD	6-7-4-1-2-3	6-1-4-7-2-3	4-1-6-7-2-3	4-7-6-2-3-1
MMAW	6-7-4-1-2-3	6-1-4-7-3-2	4-1-6-7-2-3	4-7-6-2-3-1
MMFV	6-7-4-1-2-3	6-1-4-7-2-3	4-1-6-7-2-3	4-7-6-2-3-1
	$d = 3$			
SAD	4-6-7-1-2-3	4-1-6-7-2-3	6-4-1-2-7-3	6-4-7-2-3-1
MRAW	6-7-4-1-3-2	4-1-6-7-3-2	4-1-6-3-2-7	6-4-7-3-2-1
MFV	4-6-7-1-2-3	6-1-4-7-2-3	6-4-1-2-7-3	4-7-2-3-6-1
MMRW	4-6-7-1-2-3	6-1-4-7-2-3	6-4-1-2-7-3	4-7-2-3-6-1
MAD	6-7-4-1-3-2	1-6-4-7-3-2	4-1-6-7-3-2	4-7-6-3-2-1
MMAW	6-7-4-3-1-2	1-6-4-7-3-2	4-1-6-7-3-2	4-7-6-3-2-1
MMFV	6-7-4-1-3-2	1-6-4-7-2-3	4-1-6-7-3-2	4-7-6-3-2-1

Table 3.8: Actor Rankings for Sensitivity Analysis by Importance Metric and Demand Level

this experiment, an actor 6 which immediately precedes the target of the influence campaign has a high probability of being in influence level 3. For each demand level, either actor 4 or 6 is identified as the most important actor, and actor 1 is the second most important for all metrics at demand levels $d = 1$ and $d = 2$. For $d = 3$, actor 1 is the second most important actor for the alternative CIM as well as $MRAW$. Recall from Table 3.2 and Figure 3.3 that of the actors to which influence is initiated, actor 1 has the lowest probability of having an influence level of 3.

In the 4-high experiments either actor 4 or 6 is the most important actor for each metric, but actor 7 is identified as the second most important actor for $MRAW$ and all the alternative CIM. Similar results are achieved with the 6-low experiments. For all demand levels, actors 4 and 6 are identified as the most important actors, and actor 7 is the second most important actor under the alternative CIM and MFV.

3.4. Conclusions and Future Work

In this chapter, it is shown that the use of reliability importance measures for a multi-state social network can lead to the identification of different critical members than those identified by traditional social network analysis. A sensitivity analysis further demonstrates that the use of reliability importance measures can provide insights on actor importance that are not available through traditional SNA. Furthermore, Yang and Knoke [47] have proposed that for valued graphs "a direct tie between two actors may not be as optimal as an indirect path through one or more intermediaries." This comment, along with the results from the analysis presented in this chapter, raises two important research questions: (i) what metric should be chosen to *quantify* the value of a communication path in a social network and (ii) if traditional centrality measures, driven explicitly by the structure of the network, fail to fully capture the communications in a *social* network, then what alternative approaches should be considered? In the next chapter, the importance of each network path is quantified from a reliability perspective. More specifically, the flow of information in a social network is assessed by framing the communication between connected actors as a multi-state reliability problem. This addresses the concerns of Yang and Knoke [47] by considering alternative metrics that expand standard centrality measures to account for characteristics that are unique to social networks, such as the levels of influence that actors have on one another.

Chapter 4

Modeling Social Networks under Conditional Influence

4.1. Introduction

The previous chapter introduced a model which took into account not only the interconnections of actors within a social network but also the strength of influence between actors. Influence was defined as the ability of an actor "to induce a change in behavior of another that conforms to the influencing actor's desires [21, 22]," and influence was treated as a commodity (or service) that was distributed across the network. An influence campaign was initiated to a set of accessible actors and was deemed successful when influence reached a set of intended targets. The levels of influence for each actor corresponded to their ability/willingness to continue the influence campaign, and the reliability of the network was given as the probability that a specified level of influence reached the intended targets.

In this chapter, the model developed in Chapter 3 is extended to take into account that an actor's social interaction and level of influence are not only multi-valued and stochastic in nature but also a function of the interactions with its neighbors. To analyze the influence interaction, three different metrics are proposed. To obtain computational results a multi-state network reliability model [37] is presented that provides the probability of influence level for a given actor in the network. Illustrative examples are presented, and the network reliability under the various influence levels is computed using exhaustive enumeration for a small example and Monte Carlo simulation for a larger, more realistic sized example. Finally, additional performance metrics to assess the propagation of influence through a network are defined and analyzed. The essence of the models, performance metrics, and results for this chapter have been accepted for publication in *Reliability Engineering & System Safety* [44] and in the proceedings of the *ESREL 2012 Annual Conference* [38]

4.2. Model Development

Consider a social network modeled as graph, $G(V, E)$. The size of the network is defined as the number of actors within the network, n. Let $V = (v_1, v_2, \ldots, v_n)$ denote the set of actors in the network, and let $E = (e_1, e_2, \ldots, e_m)$ denote the number of relationships (or edges) between the actors. Let $\mathbf{A}$ denote the connectivity matrix of a social network where element a_{ij} of $\mathbf{A}$ is binary valued as follows:

$$a_{ij} = \begin{cases} 1 & \text{if actor } i \text{ is connected to actor } j \\ 0 & \text{otherwise} \end{cases} \qquad \forall i = 1, 2, \ldots, n; j = 1, 2, \ldots, n; i \neq j$$

Let A'_j denote the set of actors in the network that precede and are connected to actor j, i.e. $A'_j = \{i : a_{ij} = 1\}$. Suppose that influence is a multi-state commodity that flows through the social network $G(V', E')$. It is assumed that this influence is flowing in a two terminal network such that $i = s$ denotes that influence is initiated from the source node and $j = t$ indicates influence arrival at the targeted node. Moreover, d is defined to be the required level of influence at the targeted node, t.

4.2.1 Conditional Probability

The definition of conditional probability throughout this work refers to the probability that actor i's strength of influence *onto* actor j equals e, given his *current* level of influence equals e_i. In the remainder of the chapter, conditional probability is denoted as $p_{ij}(e|e_i)$. The formal definition of $p_{ij}(e|e_i)$ is given by

$$P(s_{ij} = e|e_i) \tag{4.1}$$

where s_{ij} is the strength of influence between actor i and actor j. Note that for each actor, the current level of influence associated with actor j, e_j, is dependent on the influence levels associated with the actors directly preceding actor j. Therefore, e_j is formally defined to be a function of the influence levels associated with those actors in the set $A_j^{'}$,

$$e_j = f(s_{ij}) \qquad \forall i \in A_j^{'} \tag{4.2}$$

Note that this scenario implies that the values s_{ij} are elements of a stochastic matrix $\mathbf{S}$ which is defined such that the level of influence at the targeted node, e_t, maps $\mathbf{S}$ into a network of social influence behavior with reliability:

$$R(d) = P(e_t \geq d). \tag{4.3}$$

The definition of e_j in (4.2) assumes that an actor has a rule by which their level of influence is determined. As is shown in the remainder of this chapter, the specification of this rule has significant impact on the probability that a desired level of influence reaches the target of the network. The determination of an appropriate rule requires the consideration of how a specific actor might choose to weight differing levels of influence passed to him when determining the influence that he will have on subsequent actors in the network. In this chapter, three functions are considered as rules to determine current levels of influence for each actor: maximum influence, minimum influence, and median influence.

Actors that pass along the strongest (largest value) influence placed on him by a single connected actor is said to act according to the rule of maximum influence. In this case, the actor's influence is given by the following expression:

$$e_j = \max_{i \in A_j^{'}}(s_{ij}) \tag{4.4}$$

The maximum influence function reflects a scenario in which an actor behaves in an optimistic manner. That is, they assume the largest level of influence gathered from a single actor regardless of the remaining levels of influence communicated by all other connected actors in the network. Inversely, actors that are more conservative in how they process and communicate information are said to act in a pessimistic manner. In these cases, an actor's influence is given by the following:

$$e_j = \min_{i \in A'_j}(s_{ij}) \tag{4.5}$$

Actors adopting the minimum influence function accept a level of influence determined by the lowest level of influence amongst all actors connected to them. It is also interesting to consider a function that is reflective of a more analytical actor. Therefore, it is said that an actor who is considerate of the range of influence values observed from their connected actors has adopted the median influence function. The median influence function is given by the following:

$$e_j = \lfloor \text{median}_{i \in A'_j}(s_{ij}) \rfloor \tag{4.6}$$

This last influence function makes obvious use of the median operator used widely to assess the centrality of a range of values. Note that the floor of the median is found for ease and consistency in implementation.

4.2.2 Illustrative Example 1

To illustrate the modeling concepts introduced in this chapter, a simple social network example (Example 1) is presented. The purpose of the conditional influence model is to evaluate the network reliability of a particular social network in which the targets of influence must receive a predetermined level of influence in order to conform to the desires of the initiator of said influence. In this example, and all others considered in this chapter, actors within the network may be in one of four influence levels. An influence level of 3 indicates an actor perfectly receives a message and passes it along the communication path, and an influence level of 0 indicates the actor does not pass along the message. Influence levels of 1 and 2 account for the possibility of imperfect communication. It is assumed that the initially accessible actors receive perfect information from the initiator, and influence received by the targets flows directly into the sink. The remaining actor's conditional influence probabilities associated with each level of influence are assumed to be determined by experts familiar with the organization and are given below in Table 5.1.

e_i	$\Pr(x_j = s_{ij} \mid e_i)$			
	0	1	2	3
0	1	0	0	0
1	0.5	0.5	0	0
2	0.1	0.4	0.5	0
3	0.05	0.1	0.2	0.65

Table 4.1: Conditional Probability of Strength of Influence

Example 1 represents a social network with $n = 7$ actors and the social relationships depicted in Figure 4.1. In this example, actors 1 and 2 are assumed to receive perfect information from the source (an influence level of 3), while actors 6 and 7 are the targets of the influence. The total influence received from actors 6 and 7 is used to evaluate the reliability of the network. For this example, if actors 6 and 7 both receive perfect information, then the total influence observed at the sink will be $d = 6$ units. To ensure that both actors receive a non-zero level of influence, then the minimum total influence that can be observed at the sink is $d = 4$ units. Therefore, network reliability is evaluated given that the target must receive $d = 4$, $d = 5$, and $d = 6$ units of influence. Exhaustive enumeration (EE) is used to determine the system reliability for each influence rule, and the results are shown below in Table 4.2.

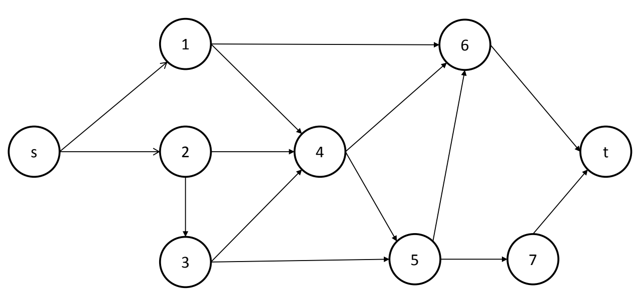


Figure 4.1: Notional Network for the Illustrative Example 1

Demand	**Maximum**	**Minimum**	**Median**
4	0.8845	0.0749	0.4104
5	0.7712	0.0342	0.2062
6	0.4742	0.0135	0.0921

Table 4.2: System Reliability for the Example 1 using Exhaustive Enumeration

These results indicate that the reliability of the network decreases as the required demand on the sink node increases. Importantly, these results illustrate the significant impact that influence function choice can have on the overall system reliability. When actors choose the maximum function to determine their current level of influence, the overall system reliability obtained is notably higher than the reliability found using the minimal influence function. Interestingly, in these results, the median function results in a level of influence midway between the two alternative influence functions when targeted demand is at its lowest level. However, as demand increases, the choice between the median and minimum influence function becomes less significant. Therefore, these results suggest that when the targeted demand is of moderate to low value, the manner in which the actor determines their influence level is an important factor. However, as targeted demand increases, the sensitivity of the influence functions becomes far less apparent.

In addition to evaluating the reliability of a fixed social network, this modeling paradigm is useful for organizations that wish to evaluate "what if" scenarios. For example, an advertising company may wish to determine how much they can influence a target group of people (which is inaccessible directly to them), by varying the number of people with which they have direct contact and can thus be influenced with the campaign. This would be done by varying the number of actors that are directly connected to the source node. In addition, the approach used in this work can be used to evaluate which actors should be targeted in order to improve the overall system reliability. To illustrate this application, reconsider the network given in Figure 4.1. Note that actors 1 and 2 are directly influenced by the source node. To evaluate the importance of actor 2, a perturbed version of the 9 node network is considered in which only nodes 1 and 3 are directly influenced by the source (see Figure 4.2). Note that in this perturbed network, actor 2 has been omitted and actor 3 is now directly connected to the influence source. This results in an increase in system reliability for all demand levels for all influence functions. Similarly, the importance of actor 1 is evaluated by considering a different permutation of the 9 node network in which only nodes 2 and 3 are directly influenced by the source (see Figure 4.3), and actor 1 has been omitted. The resulting system reliabilities under the set of influence functions for each demand level are shown in Tables 4.3 and 4.4.

These results may be used to quantify the impact of the removal of an actor from the initial influence campaign by investigating the resulting decrement in system reliability. For Example 1, omitting actor 1 results in the largest decrement to system reliability followed by actor 3 then actor 2. This ranking can be used by the initiators of the influence campaign to determine the actors with which to allocate limited resources to maximize the success of an influence campaign.

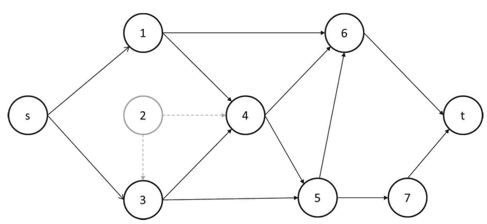


Figure 4.2: Notional Network with Actors 1 and 3 Directly Influenced

Demand	**Maximum**	**Median**	**Minimum**
4	0.9107	0.3902	0.1627
5	0.7602	0.1730	0.0811
6	0.5245	0.0833	0.0319

Table 4.3: System Reliability with Actors 1 and 3 Directly Influenced

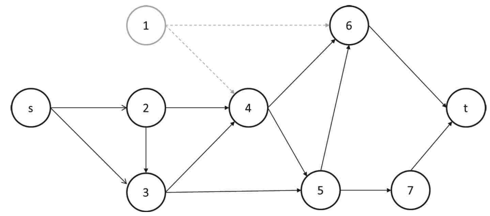


Figure 4.3: Notional Network with Actors 2 and 3 Directly Influenced

Demand	**Maximum**	**Median**	**Minimum**
4	0.7721	0.1678	0.1203
5	0.5996	0.0741	0.0642
6	0.3917	0.0319	0.0319

Table 4.4: System Reliability with Actors 2 and 3 Directly Influenced

4.3. Monte Carlo Simulation Model

Reliability computation is straightforward for small networks or for cases when minimum cut/path sets are easily obtained. However, as the size of the network (or the number of actor influence levels) increase, resorting to complete enumeration or minimum cut/path sets computation is computationally inefficient. Therefore, alternative reliability computation approaches are needed. In reliability engineering, Monte Carlo simulation has been effectively used for estimating two-terminal reliability for relatively large systems [36]. To analyze large social networks, a Monte Carlo (MC) simulation methodology is proposed for estimating equation (4.3). The methodology consists of using MC simulation to generate the strength of influence passed among network actors which are dependent on the influence rules.

Collaborators at Stevens Institute of Technology developed a MC simulation model to evaluate the system reliability under the current conditional influence model [44]. The influence algorithm developed has three main steps: 1) set up the initial influence value at the source node, 2) determine how the initial influence propagates throughout the network and 3) determine how much influence is received at the target node. It is on the propagation of the influence, where the MC simulation takes place. In order to determine the influence that each actor will give (see Table 5.1), the MC simulation serves as a means of constructing a stochastic network where the strength of each link is probabilistic. The network links need to be created as the influence flows through the network. Note that in order to determine the influence that a node will give, all its incoming influence needs to have been previously computed. Therefore, the simulation model only works with directed acyclic graphs and the order in which nodes need to be visited is given by a topological sorting of the graph [45]. The algorithms used for the simulation are explained in more detail below in Algorithms 1, 2, and 3. In Algorithm 3 the function getSynergy calculates the amount of influence that will be given for a particular node depending on the influence rule used (minimum, the floor of the median, or the maximum). The function getSampleLink, randomly generates a link depending on the influence received.

Input: $source, target, simulations, initialInfluence, targetInfluence, G = (V, E)$

Output: Reliability of the network

for $i = 0$**to**$simulations$ **do**

    visit($source, initialInfluence, G = (V, E)$);

    $received$ ←influenceReceivedAt($targetNode$);

    if $received \geq target$ **then**

        $success \leftarrow success + 1$;

    end

end

$reliability \leftarrow success/simulations$;

Algorithm 1: Main simulation

Input: $source, initialInfluence, G = (V, E)$

Output: All the links of the network are generated

$Influence_{source} \leftarrow initialInfluence$;

$nodeOrder \leftarrow$ topologicalSorting(G) ;

for $node \in nodeOrder$ **do**

  for *successor* ∈*successorsOf(node)* **do**

    processEdge(*node*, *successor*) ;

  end

end

Algorithm 2: Visit

Input: $source, destiny, G = (E, V)$

Output: Generate the stochastic link from *source* to *destiny*

$received \leftarrow$ getSynergy(*source*) ;

$pr =$ getRandom(0,1) ;

$transmit =$ getSampleLink($pr, received$) ;

$E \leftarrow E \cup transmit$;

Algorithm 3: Process Edge

4.3.1 Monte Carlo Simulation Experimentation

The use of the MC simulation allows for the consideration of larger social networks. However, the simulation-based approach does not guarantee that exact system reliability values will be obtained. To assess the quality of the MC simulation methodology, Example 1 is reconsidered and solved again using the MC procedure described above. A comparison of the MC and EE results are provided in Table 4.5.

Note that these results suggest that MC approximates the exact system reliability for variants of Example 1 very well. Each of the reliabilities obtained are accurate to nearest

Demand	Maximum		Median		Minimum	
	EE	**MC**	**EE**	**MC**	**EE**	**MC**
4	0.8845	0.8841	0.4104	0.4116	0.0749	0.0747
5	0.7112	0.7111	0.2062	0.2073	0.0342	0.0338
6	0.4742	0.4738	0.0921	0.0925	0.0135	0.0134

Table 4.5: System Reliability for Example 1: MC Versus EE

100^{th} of a decimal place. In fact, the maximum absolute deviation from the exact reliabilities for the MC results was 1.29%. Given the quality of the solutions obtain by MC, a larger social network example is considered.

4.3.2 Illustrative Example 2

The next example considered, referred to as Example 2, is almost double the size of Example 1 and cannot be solved efficiently using EE. Therefore, MC is used to consider a larger social network with $n = 13$ actors and the social relationships depicted in Figure 4.4.[1] In this example, actors 0, 1, 2, and 3 are assumed to receive perfect information from the source (an influence level of 3), while actors 11, 12, and 13 are the targets of the influence. The total influence received from actors 11, 12, and 13 is used to evaluate the reliability of the network. For this example, if actors 11, 12, and 13 all receive perfect information, then the total influence observed at the sink will be $d = 9$ units. To ensure that all actors receive a non-zero level of influence, the minimum total influence that can be observed at the sink is $d = 7$ units. Therefore, the network reliability is evaluated given that the target must receive $d = 7$, $d = 8$, and $d = 9$ units of influence. The results are shown below in Table 4.6.

Demand	Maximum	Median	Minimum
7	0.9998	0.3748	0.0123
8	0.9988	0.1736	0.0040
9	0.9934	0.0625	0.0012

Table 4.6: System Reliability for Example 2

[1] Network was taken from [46].

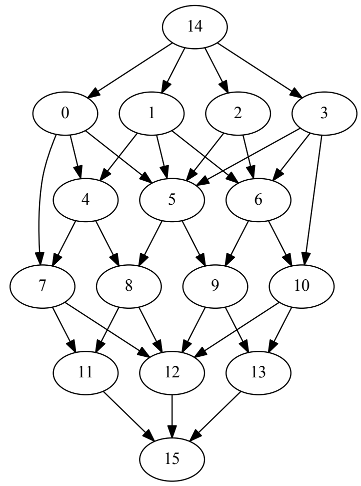


Figure 4.4: Notional Network for Example 2

Now, as was done in Example 1, the "what if" scenarios that allow for the quantification of the impact of actor removal from the initial influence campaign are investigated. Each scenario in Figure 4.5 represents a different set of actors that are targeted for direct influence. The links that appear in grey represent influence connections that have been removed from the original network. For example, figure (a) shows the original network without the source (node 14) influencing node 0. For each of the scenarios depicted in Figure 4.5, the reliability of the network is evaluated for demand values of 7, 8 and 9. The results are presented below in Table 4.7.

Ranking the set of uninfluenced actors as shown in Table 4.7 offers insights into the significance of each actor. The most important set of actors in ascending order are: $\{0,1\},\{0,3\},\{1,3\},\{2,3\},\{0,2\},\{0\},\{1,2\},\{3\},\{1\},\{2\}$. It can be concluded that removing actor 0 from the initial influence campaign results in the largest overall decrement to network reliability, given that it ranks first in the single node removal and it also appears in the first two configurations that include the removal of a pair of nodes. If a decision-maker had to choose one node to influence, they should choose this node. Note that in this setting influence is considered as a limited resource available for allocation.

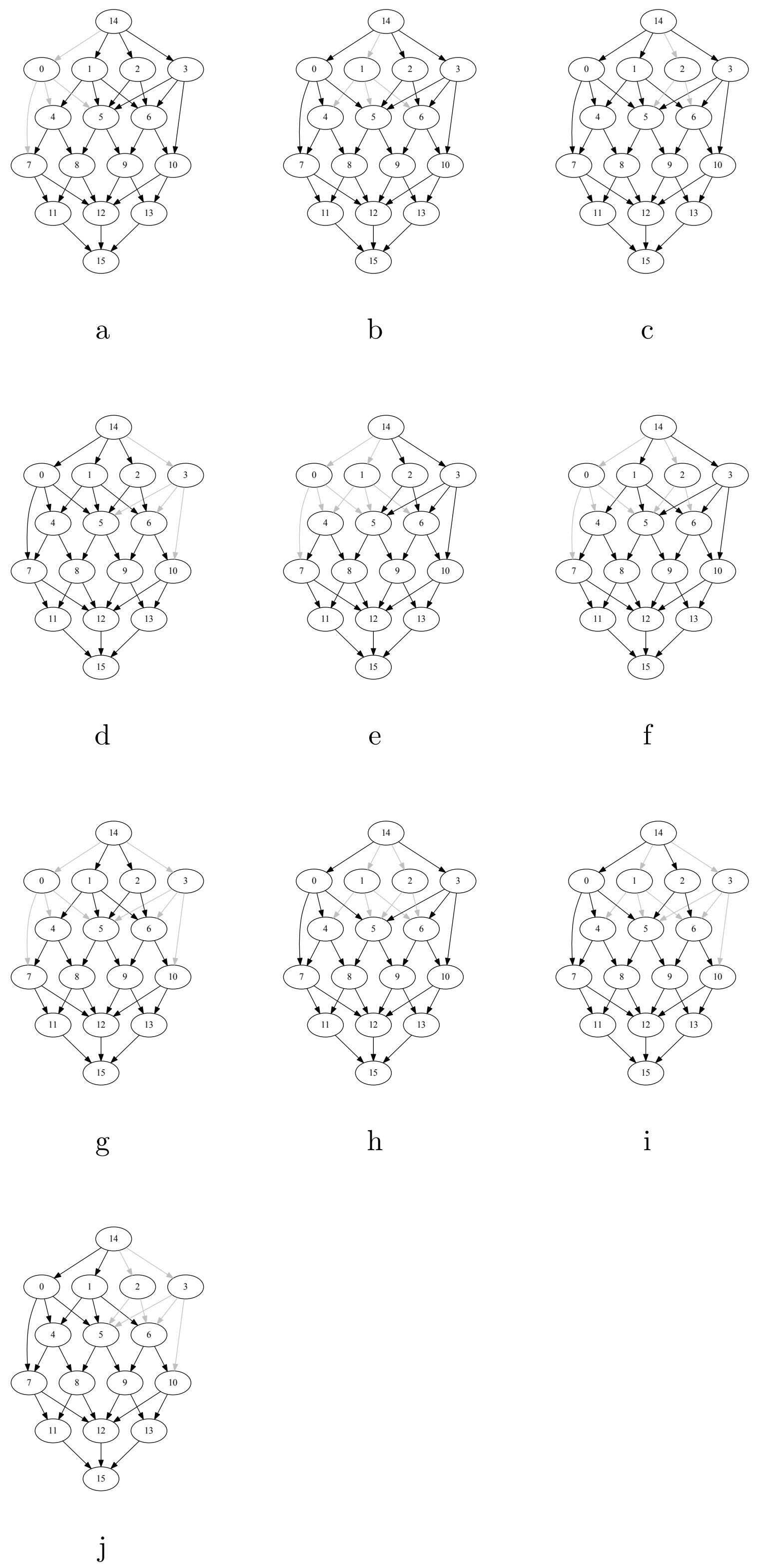


Figure 4.5: Variation on the group of actors initially influenced

network	Uninfluenced Actors	Demand	Maximum	Median	Minimum
a	**0**	**7**	0.9104	0.038	0.0012
b	**1**	**7**	0.9564	0.0196	0.0009
c	**2**	**7**	0.9646	0.0116	0.001
d	**3**	**7**	0.9377	0.0147	0.0013
e	**0,1**	**7**	0.7113	0	0
f	**0,2**	**7**	0.8976	0.0153	0.0022
g	**0,3**	**7**	0.8529	0.0185	0.0049
h	**1,2**	**7**	0.9373	0.0144	0.003
i	**1,3**	**7**	0.8837	0.0204	0.0029
j	**2,3**	**7**	0.8946	0.014	0.0038
a	**0**	**8**	0.7557	0.01	0.0001
b	**1**	**8**	0.8334	0.0046	0.0002
c	**2**	**8**	0.8628	0.003	0.0003
d	**3**	**8**	0.8081	0.0046	0.0002
e	**0,1**	**8**	0.4982	0	0
f	**0,2**	**8**	0.7277	0.0042	0.0008
g	**0,3**	**8**	0.6768	0.0058	0.0015
h	**1,2**	**8**	0.8028	0.0039	0.001
i	**1,3**	**8**	0.7029	0.0065	0.0017
j	**2,3**	**8**	0.7333	0.0038	0.0016
a	**0**	**9**	0.4787	0.0021	0
b	**1**	**9**	0.5802	0.0009	0.0002
c	**2**	**9**	0.6048	0	0.0002
d	**3**	**9**	0.5415	0.0008	0.0002
e	**0,1**	**9**	0.279	0	0
f	**0,2**	**9**	0.4629	0.0009	0.0002
g	**0,3**	**9**	0.4165	0.0011	0
h	**1,2**	**9**	0.5314	0.0004	0.0004
i	**1,3**	**9**	0.4359	0.0009	0.0002
j	**2,3**	**9**	0.4604	0.0006	0.0004

Table 4.7: Reliability Results of different scenarios

4.4. Additional Performance Metrics

The results from the MC simulation model described in section 4.3 can be further investigated to provide additional insights with regards to the propagation of influence through the organization [38]. In this section, performance metrics are developed to quantify the level at which individual actors within the organization are influenced under this model. During step 2 of the influence algorithm within the MC simulation model, a system state vector is created that details the propagation of influence throughout the network. Let $\mathbf{x}$ denote a system state vector, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ where x_i denotes the influence level of actor i. These state vectors can be analyzed to determine the expected influence level for each actor in the network under this model. The first performance metric is given by

$$q(a_i|\mathbf{x}) = \frac{E[a_i|\mathbf{x}]}{n} \tag{4.7}$$

where $E[a_i|\mathbf{x}]$ defines the expected number of actors influenced at level i, a_i, given influence vector $\mathbf{x}$. Recall that n is the total number of actors in the network. Note that a_i is stochastic since the influence is governed by the conditional influence at the actor level. Influence vector $\mathbf{x}$ describes the actor influence levels for each simulation run; allowing consideration of different influencing strategies. This metric provides the expected level at which a particular actor is influenced. This approach is equivalent to an all-terminal reliability evaluation [10], [39]. Relevant examples of this case may include cases on political and marketing operations where the interest is influencing portions of the organization.

In essence, this approach provides a description of how the influence in the network is exchanged and propagated by the social network actors as a function of the influence vector, the various rules of influence, and their conditional influence probabilities.

The second performance metric is given by

$$Median(e_i|\mathbf{x}) \tag{4.8}$$

which provides a description of the median influence level attained by actor i within the network given influence vector $\mathbf{x}$.

To evaluate the model under these performance metrics, the network from Example 2 in section 4.3.2 is considered (see Figure 4.4). The MC simulation model is used to generate 1,000 state vectors which are used to determine the median and average influence levels for each actor in the network. Finally, the expected value of the percentage of actors in each influence level is determined for the various influence rules. The results are shown in Table 4.8. Figure 4.6 provides a graphical representation of the overall median actor influence level under each influence rule. Figures 4.7, 4.8, and 4.9 illustrate the results of eight individual state vectors obtained from a simulation run considering the three influence rules. To facilitate understanding of the results, the influence levels are color-coded as follows: 0 (red), 1 (pink), 2 (yellow), 3 (green).

Function	**Influence**			
	0	**1**	**2**	**3**
Minimum	35.83	14.51	10.84	38.83
Median	7.33	21.89	23.76	47.03
Maximum	0.17	1.26	6.23	92.34

Table 4.8: Expected Percentage of Actors in Each Influence State

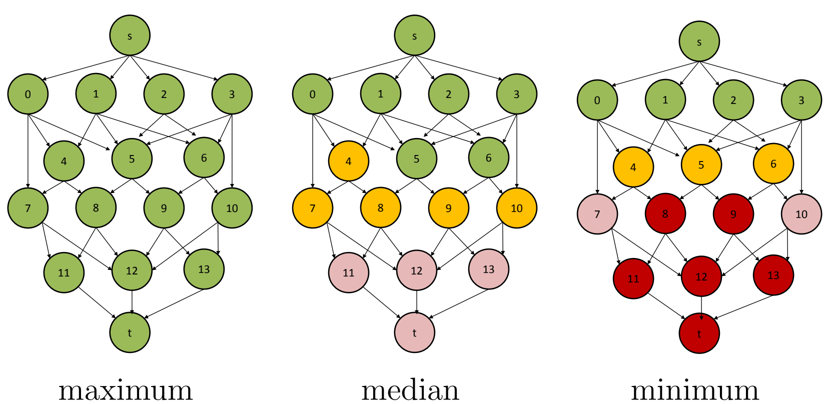


Figure 4.6: Overall Median Actor Influence Level Under Various Influence Rules

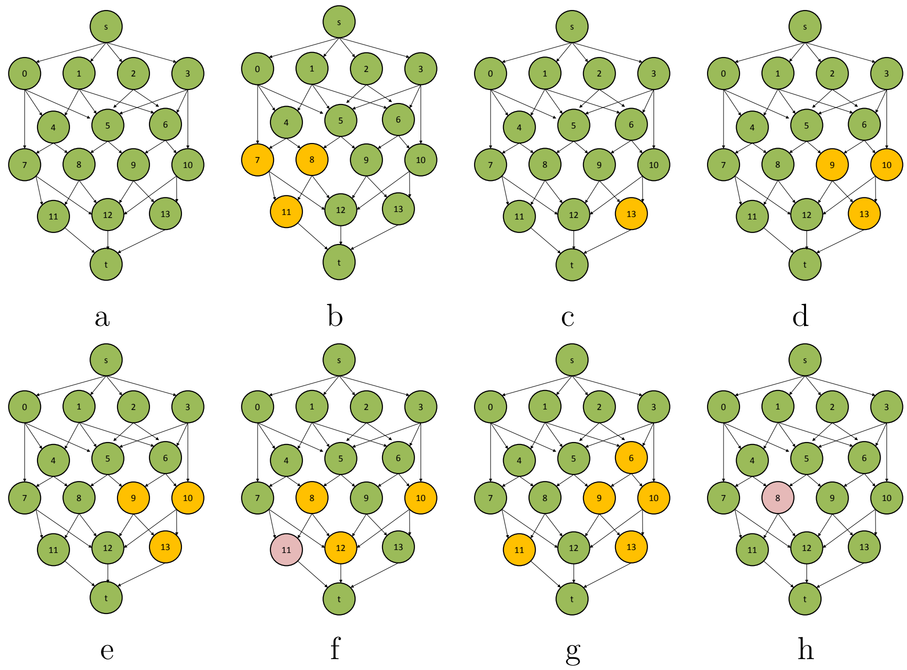


Figure 4.7: Select Instances of Actor Influence Level Under the Rule of Maximum Influence

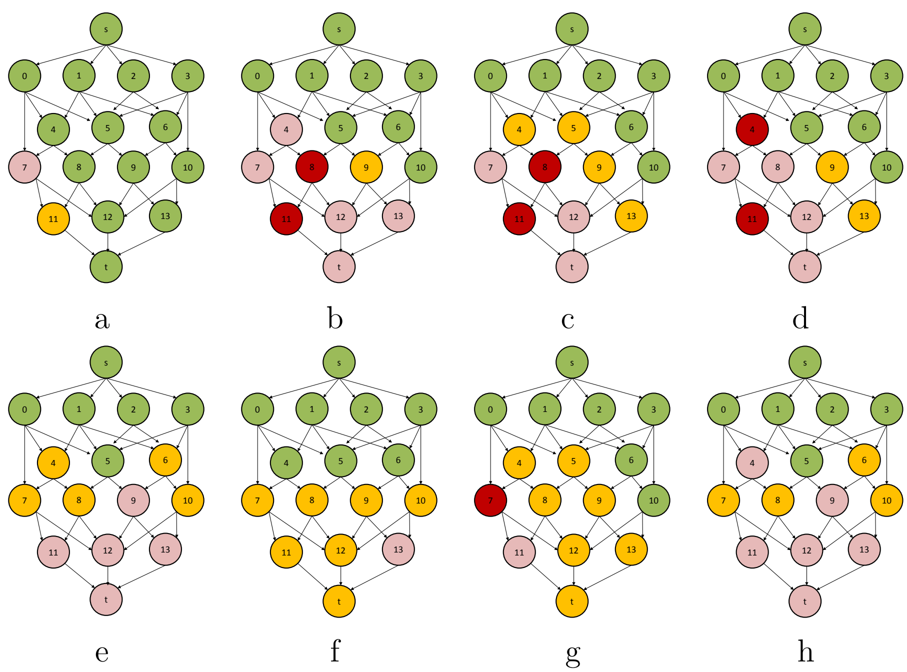


Figure 4.8: Select Instances of Actor Influence Level Under the Rule of Median Influence

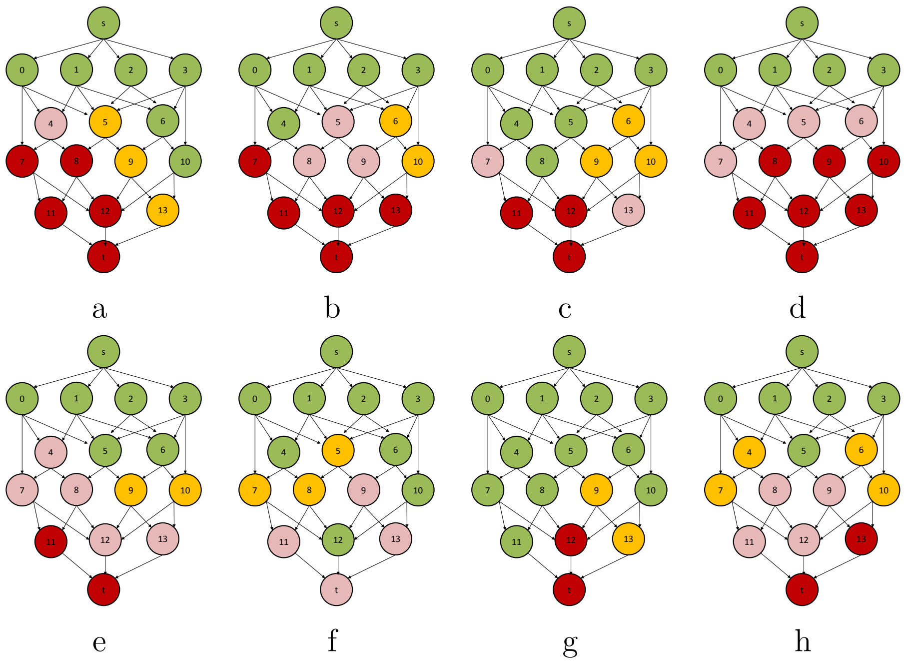


Figure 4.9: Select Instances of Actor Influence Level Under the Rule of Minimum Influence

4.5. Conclusions and Future Work

In this chapter, the necessary model paradigms to evaluate the exchange and propagation of influence within a social network have been presented, and a Monte Carlo simulation model has been developed to evaluate the probability that the flow of influence that reaches the target actor(s) is greater than or equal to some threshold value, i.e. network reliability. The model developed in this chapter differs from traditional social network and network flow models in that actor influence is considered multi-valued, stochastic, and a function of an actor's interactions with others within the network. Although in this chapter only homogeneous influence levels for all actors are considered, it is important to note that in real life, the influence level of each actor could be drastically varied with respect to various reasons (i.e. personal characteristics or the characteristics of a situation at hand). The models presented allow for addressing heterogeneous actors with distinct levels of influence and influence functions. However, this requires the compilation of data to supports such model.

Exhaustive enumeration was used to evaluate the reliability of a small, illustrative example, and a Monte Carlo simulation model was developed to efficiently approximate the value of social network reliability for a larger, more realistic problem. Two additional performance metrics were defined to further evaluate the propagation of influence through the network under study. These approaches give decision-makers a model that can to be used as a tool to identify the network actors that most significantly impact how the targeted nodes are influenced. In this case, influence is considered a limited resource available for allocation.

Future research in this area will focus on identifying additional performance metrics that can be used to assist decision makers with allocating resources related to influence campaigns. The emphasis will be on developing metrics with closed-form expressions that can be evaluated without the use of simulation modeling. Initial studies will focus the effects of alternating communication paths or eliminating the uncertainty associated with an actors level of influence.

Chapter 5

Optimizing Social Networks under Conditional Influence

5.1. Introduction

Recently, researchers have shown interest in the optimization of social networks. Key issues under investigation include identifying the important actors within a network [5, 22], minimizing the time to influence all actors in a network [34], and maximizing the diffusion of influence throughout a network subject to limited resources [7, 16, 33]. Borgatti [5] defines two Key Player Problems (KPP-1 and KPP-2) that allow a decision maker to identify a set of k individuals that achieves a specified goal. KPP-1 identifies a set of actors whose removal results in a residual network that exhibits the least cohesion possible while KPP-2 identifies the set of actors that are maximally connected to all others. Hamill [22] extends the work of Borgatti [5] in several ways. He defines the input parameters of the problem such that directed graphs may be considered, formulates a set of mathematical programs in which the main decision variables are defined such that if an actor is chosen as a key player, it must be able to reach its assigned members within m steps, and incorporates constraints to account for characteristics (other than network location and connectivity) that may be associated with certain actors. Ni et al. [34] formulate models to minimize the expected value of the amount of time required to influence all actors in a network while Evan-Dar and Shapira [16] develop models to maximize the spread of influence throughout a network. Cao et al. [7] present a resource allocation model to maximize influence diffusion through a densely connected social networks. In these papers, the authors use the term influence to refer to a person's willingness to adopt a new technology or purchase a new product.

Recently, an optimization model was developed to determine the best way to make connections among actors within a disconnected multi-layer social network where the layers of the network represented different affiliations among the actors [41]. A graphical representation of the multi-layer social network is shown below in Figure 5.1. The model was used to determine the manner in which actors should communicate within and across layers to ensure that the communication not only reliably reaches the intended recipient but is also communicated to a minimum number of actors. In this model, the reliability of an actor was defined as the probability that an actor accurately conveys a received communication to another actor, and we assumed actor reliabilities were known. This model utilized concepts from both reliability and network optimization to provide decision makers with a tool for measuring the tradeoff between system reliability and the degree of information sharing within the network. The primary weakness of this model is that each actor in the network communicates with no more than one other actor. This results in a single, serial communication among actors in the network such that network reliability is calculated as a series system. An enumeration scheme was developed to identify the optimal solutions to the models developed.

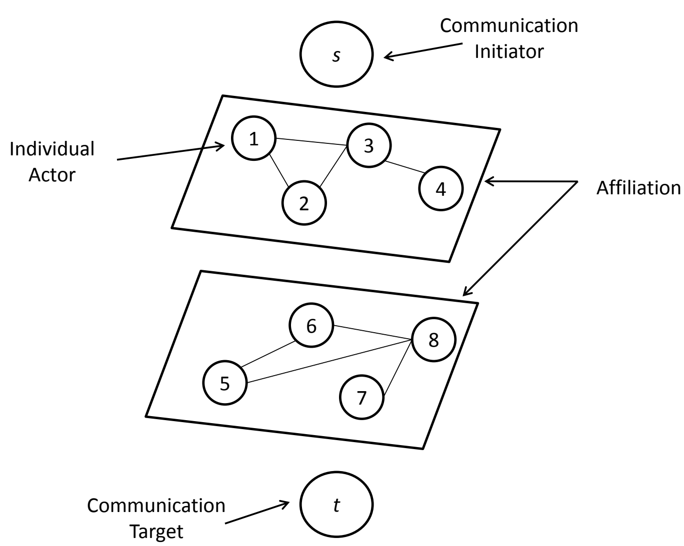


Figure 5.1: Multi-Layer Social Network Representation

In this chapter, we continue to investigate models in which actor influence levels are multi-valued, stochastic, and dependent upon the influence they receive from their predecessors.

Also, we modify the layered network representation presented above in that we consider hierarchical communication that occurs between actors in different levels on the network. The specifics of this model are discussed next.

5.2. Model Development

Consider the model presented in Section 4.2 where a social network is modeled as a graph, $G(V, E)$. The size of the network is defined as the number of actors within the network, n. Let $V = (v_1, v_2, \ldots, v_n)$ denote the set of actors in the network, and let $E = (e_1, e_2, \ldots, e_m)$ denote the number of relationships (or edges) between the actors. Let $\mathbf{A}$ denote the connectivity matrix of a social network where element a_{ij} of $\mathbf{A}$ is binary valued as follows:

$$a_{ij} = \begin{cases} 1 & \text{if actor } i \text{ is connected to actor } j \\ 0 & \text{otherwise} \end{cases} \qquad \forall i = 1, 2, \ldots, n; j = 1, 2, \ldots, n; i \neq j$$

Let A_j' denote the set of actors in the network that precede and are connected to actor j, i.e. $A_j' = \{i : a_{ij} = 1\}$, and let N denote the cardinality of the set, $\left|A_j'\right|$. Each actor within the organization has some level of influence that quantifies their willingness/ability to pass along information. Influence is modeled as a multi-state commodity that flows through the social network. It is assumed that this influence is flowing in a two terminal network such that $i = s$ denotes that influence is initiated from the source node. In this model, we assume actors within the network communicate in a hierarchical fashion. That is, actors are arranged in levels, and actors in a given level only communicate with actors in the level immediately below them (see Figure 5.2) . Influence arrival at the targeted node is denoted by $j = t$, and d is defined to be the required level of influence at the targeted node, t.

In this model, an actor's influence state is probabilistic and dependent on the influence states associated with the actors directly preceding actor j. Let e denote an actor influence state where $e = 0, 1, \ldots, \omega$. The probability distribution on actor influence level is assumed

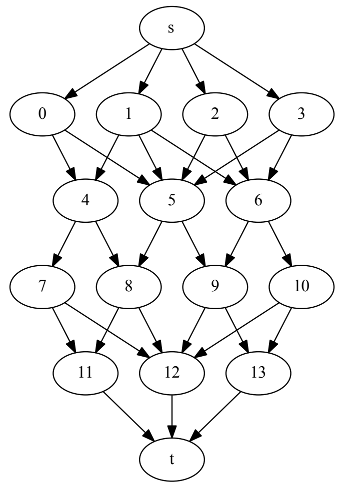


Figure 5.2: Hierarchical Network

to be the same for all actors in the network. Based on this assumption, the definition of conditional probability section 4.2.1, $p_{ij}(e|e_j)$, can be further simplified to $p(e|e_j)$. Recall that e_j is a defined as a function of the influence states associated with those actors in the set $A_j^{'}$.

Based on this scenario, the following research questions are presented:

- If resources exist to add or remove connections within the network, where should those changes be made?

- If it is possible to eliminate the uncertainty in influence state associated with actors in the network, how should those actors be chosen?

5.3. Illustrative Example

Consider a modified version of the network from Illustrative Example 2 [46] in section 4.3.2. In this case, the social network contains $n = 14$ actors with the social relationships depicted in Figure 5.2 and influence probabilities depicted in Table 5.1. In this example, we again consider four influence states, i.e. $\omega = 3$. An influence state of 3 indicates an actor perfectly receives a message and passes it along the communication path, and an influence state of 0 indicates the actor does not pass along the message. Influence state of 1 and 2 account for the possibility of imperfect information.

e_i	$p(e\|e_i)$			
	0	1	2	3
0	1	0	0	0
1	0.5	0.5	0	0
2	0.1	0.4	0.5	0
3	0.05	0.1	0.2	0.65

Table 5.1: Conditional Probability of Strength of Influence

In this modified example, actors within the network communicate in a hierarchical fashion. That is, actors in a given level within the network only communicate with actors in the level immediately below them. In this example, actors 0, 1, 2, and 3 are considered to be in "Level 1" of the organization, and they communicate with actors 4, 5, and 6 in "Level 2" of the organization. These actors, in turn, communicate with actors in "Level 3" of the organization (actors 7, 8, 9, 10), who finally communicate with the targets of the influence campaign (actors 11, 12, and 13).

In Chapter 4, a Monte Carlo simulation model was developed to estimate network reliability. Here we evaluate the *expected influence state of actors at the target node*. A closed-form expression for this performance measure is presented, and a simulation model is not required to evaluate this performance metric. Empirical results associated with the illustrative example described above are detailed below.

5.3.1 Expected Influence Values for Level 2 Actors

In this example, actors in Level 1 are assumed to receive perfect information from the source (an influence state of 3). Therefore, the conditional probabilities associated with actors in Level 2 are simplified from $p(e|3)$ to p_e. Recall from section 4.2, that three rules are considered to determine current levels of influence for each actor: maximum influence, minimum influence, and median influence. This work focuses on the rules of maximum and minimum influence since closed-form solutions for these order statistics exist. For each influence rule, we can build the probability mass function (PMF) and compute the expected influence level for each actor in Level 2. The PMF is constructed using a set of multinomial series, and the expected influence state for each actor is the weighted sum of these series. Note that k_e denotes the number of actors in influence state e and that $N = k_0 + k_1 + \ldots + k_e$.

The PMF under the rule of maximum influence is shown below in equations 5.1 - 5.4. In order for an actor in level 2 of the network to receive an influence state of zero under the rule of maximum influence, all of its predecessors (actors in the set $A_j^{'}$) must have passed to it an influence state of zero. The probability of this occurring is computed using equation 5.1. For an actor to receive an influence state of one, at least one actor in the set $A_j^{'}$ must have passed an influence state of one. All other actors must have passed an influence state of one or less. The first term after the summation in equation 5.2 captures the number of combinations associated with having k_0-out-of-N actors in influence state zero. Note the remaining $N - k_0$ actors are in state one. Similarly, for an actor to receive an influence state of two, at least one of the preceding actors must have passed an influence state of two, and all other actors must have passed an influence state of two or less. Again, the term after the summation signs in equation 5.3 captures the number of combinations associated with having k_0 actors in state zero and k_1 actors in state one. Note that the remaining $N - k_0 - k_1$ actors are in state two. Finally, for an actor to receive an influence state of three, at least one of the actors in the set $A_j^{'}$ must have passed an influence state of three, and the remaining actors can be in any one of the other states. As with the previous equations, the term after the summation in equation 5.4 captures the number of combinations associated with having k_0 actors in state zero, k_1 actors in state one, k_2 actors in state two, and the remaining $N - k_0 - k_1 - k_2$ actors in state three.

$$f_{i,max}(0) = p_0^N \tag{5.1}$$

$$f_{i,max}(1) = \sum_{k_0=0}^{N-1} \binom{N}{k_0} p_0^{k_0} p_1^{N-k_0} \tag{5.2}$$

$$f_{i,max}(2) = \sum_{k_0=0}^{N-1} \sum_{k_1=0}^{N-k_0-1} \frac{N!}{k_0!k_1!(N-k_0-k_1)!} p_0^{k_0} p_1^{k_1} p_2^{N-k_0-k_1} \tag{5.3}$$

$$f_{i,max}(3) = \sum_{k_0=0}^{N-1} \sum_{k_1=0}^{N-k_0-1} \sum_{k_2=0}^{N-k_0-k_1-1} \frac{N!}{k_0!k_1!k_2!(N-k_0-k_1-k_2)!} p_0^{k_0} p_1^{k_1} p_2^{k_2} p_3^{N-k_0-k_1-k_2} \tag{5.4}$$

The PMF under the rule of minimum influence is shown below in equations 5.5 - 5.8. It is, in essence, the opposite of the PMF under the rule of maximum influence. For an actor to receive an influence state of zero, only one of the preceding actors must have passed an influence state of zero, and the remaining actors may be in any one of the other state. For an actor to receive an influence state of one, at least one actor from the set $A_j^{'}$ must have passed an influence state of one, and the remaining actors must have passed an influence state one or more. Similar logic follows for influence states two and three. As with the PMF under the rule of maximum influence, the terms after the summation signs capture the number of combinations associated with the N preceding actors being in a particular set of states.

$$f_{i,min}(0) = \sum_{k_0=0}^{N-1} \sum_{k_1=0}^{N-k_0-1} \sum_{k_2=0}^{N-k_0-k_1-1} \frac{N!}{k_0!k_1!k_2!(N-k_0-k_1-k_2)!} p_0^{k_0} p_1^{k_1} p_2^{k_2} p_3^{N-k_0-k_1-k_2} \tag{5.5}$$

$$f_{i,min}(1) = \sum_{k_1=0}^{N-1} \sum_{k_2=0}^{N-k_0-1} \frac{N!}{k_1!k_2!(N-k_1-k_2)!} p_1^{k_1} p_2^{k_2} p_3^{N-k_1-k_2} \tag{5.6}$$

$$f_{i,min}(2) = \sum_{k_2=0}^{N-1} \binom{N}{k_2} p_2^{k_2} p_3^{N-k_2} \tag{5.7}$$

$$f_{i,min}(3) = p_3^N \quad (5.8)$$

The resulting PMF and expected influence states for actors in Level 2 under the rules of maximum and minimum influence are shown Table 5.2 and 5.3, respectively.

State	$i = 4$	$i = 5$	$i = 6$
0	0.0025	6.25E-06	1.25E-04
1	0.0200	0.0005	3.25E-03
2	0.1000	0.0145	0.0395
3	0.8775	0.9850	0.9571
$E(e)$	2.8525	2.9844	2.9536

Table 5.2: Level 2 PMFs and Expected Influence Values under Maximum Influence

State	$i = 4$	$i = 5$	$i = 6$
0	0.0975	0.1855	0.1426
1	0.1800	0.2925	0.2433
2	0.3000	0.3435	0.3395
3	0.4225	0.1785	0.2746
$E(e)$	2.0475	1.5150	1.7461

Table 5.3: Level 2 PMFs and Expected Influence Values under Minimum Influence

To further investigate the structure of the expected influence values for Level 2 actors, additional values of N were considered, and the resulting expected influence values by influence rule are shown in Figure 5.3. Additionally, a different probability distribution (Distribution 2) on strength of influence was investigated. Specifically, the values of p_e considered with Distribution 2 were $p_0 = 0.40, p_1 = 0.25, p_2 = 0.20$, and $p_3 = 0.15$. These results are shown for each influence rule in Figures 5.4 and 5.5. Based on empirical evidence from initial experimentation, it can be conjectured that the expected influence states are concave under the rule of maximum influence and convex under the rule of minimum influence.

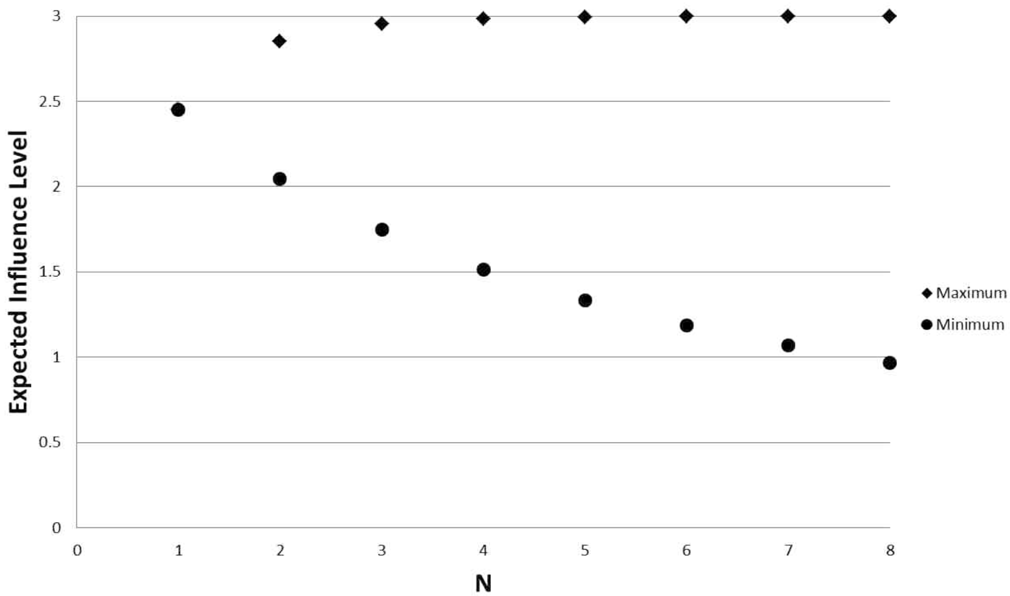


Figure 5.3: Expected Influence Values by Influence Rule

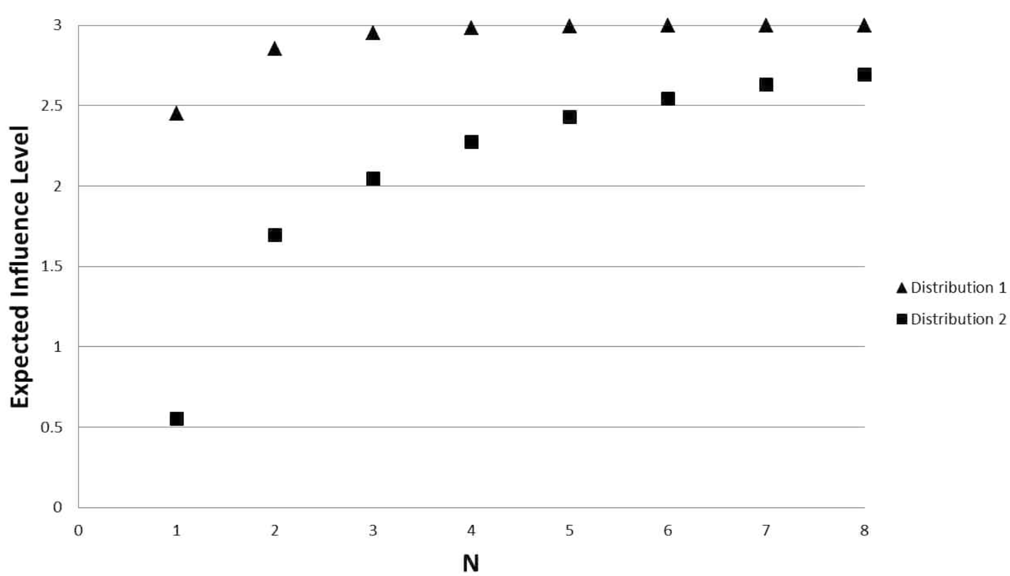


Figure 5.4: Expected Influence Values under Rule of Maximum Influence

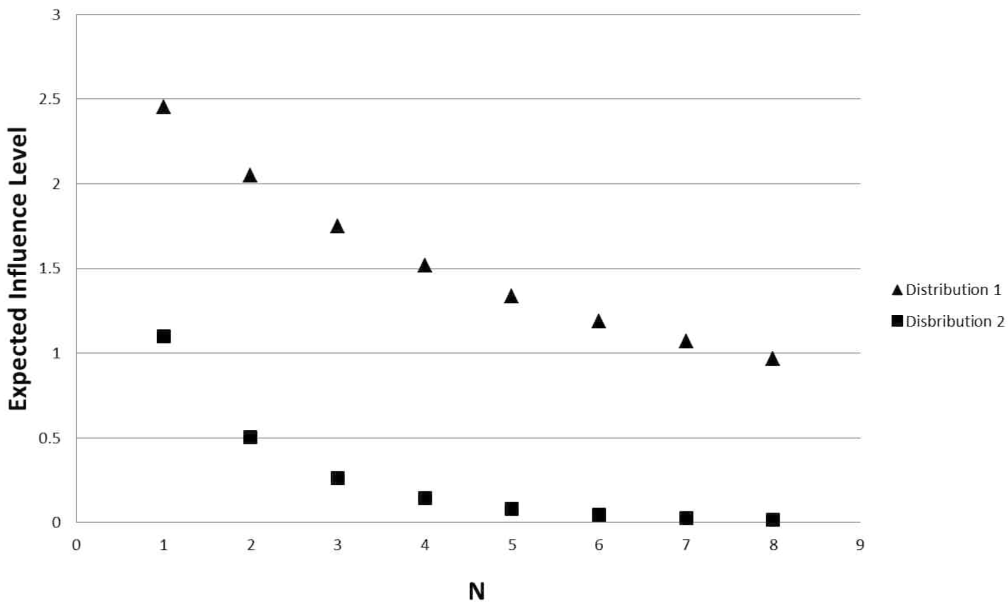


Figure 5.5: Expected Influence Values under Rule of Minimum Influence

5.3.2 Expected Influence Values under Conditional Influence

For the remaining levels in the network, the actor-specific PMF values from the previous levels are necessary to obtain the PMF values for the current level. Recall that $A_j^{'}$ is defined as the set of actors that precede and are connected to a particular actor. Since an actor's influence state is dependent on the influence received from preceding actors, simply specifying the number of combinations associated with actors being in the various states does not provide enough information to compute the PMF for actors in Level 3 and beyond. Let $S(N, e)$ denote a set of combinations of actor influence states from the set $A_j^{'}$ that could result in a given influence state, e. Let this combination under the rule of maximum influence be denoted by $S_{max}(N, e)$ and be denoted by $S_{min}(N, e)$ for the rule of minimum influence

As an example, consider actor 8 from Figure 5.2. Note that $A_j^{'} = \{4, 5\}$ and $N = 2$. For actor 8 to receive an influence level of zero ($e = 0$), both actors must be in state zero. Therefore,

$$S_{max}(2, 0) = \{0, 0\}.$$

To receive an influence state of one ($e = 1$), at least one actor must be in state one, and the other actor must be in or below state one, i.e.

$$S_{max}(2, 1) = \{0, 1\} \cup \{1, 0\} \cup \{1, 1\}.$$

Similarly, for actor 8 to receive an influence state of two ($e = 2$), at least one actor must be in state two, and the other actor must be state two or below, i.e.

$$S_{max}(2, 2) = \{0, 2\} \cup \{1, 2\} \cup \{2, 0\} \cup \{2, 1\} \cup \{2, 2\}.$$

The same sentiment holds for actor 8 to receive an influence state of three ($e = 3$) where

$$S_{max}(2, 3) = \{0, 3\} \cup \{1, 3\} \cup \{2, 3\} \cup \{3, 0\} \cup \{3, 1\} \cup \{3, 2\} \cup \{3, 3\}.$$

This can be expressed using set notion to identify all possible combinations of actor states excluding those that do not contain at least one actor in influence state e where

$$S_{max}(N, e) = \{0, 1, \dots, e\}^N \setminus \{0, 1, e - 1\}^N \tag{5.9}$$

Similarly, the set can be defined when actors follow the rule of minimum influence. Under this rule to receive an influence state of zero ($e = 0$), only one actor must pass a state of zero. Again we consider actor 8 as an example which gives us

$$S_{min}(2, 0) = \{0, 0\} \cup \{0, 1\} \cup \{0, 2\} \cup \{0, 3\} \cup \{1, 0\} \cup \{2, 0\}, \cup\{3, 0\}.$$

To receive an influence level of one ($e = 1$), at least one actor must be in state one and all other actor must be in a higher state, i.e.

$$S_{min}(2, 1) = \{1, 1\} \cup \{1, 2\} \cup \{1, 3\} \cup \{2, 1\} \cup \{3, 1\}.$$

Similarly, for actor 8 to receive an influence state of two ($e = 2$), at least one actor must be in state two, and the other actor must be state two or higher, i.e.

$$S_{min}(2, 2) = \{2, 2\} \cup \{2, 3\} \cup \{3, 2\}.$$

Finally, for actor 8 to receive an influence state of three ($e = 3$), both actors must be in state three where

$$S_{min}(2,3) = \{3,3\}$$

This can be expressed using set notion to identify all possible combinations of actor states excluding those that contain at least one actor above influence state e where

$$S_{min}(N,e) = \{0,1,\ldots,e\}^N \setminus \{e+1, e+2, \ldots, \omega\}^N \tag{5.10}$$

Once the sets of actor influence states have been identified, they can be used to compute the PMF values for actors in the network beyond Level 2. Let x denote the cardinality of $S(N,e)$, and let s_i^x denote the i^{th} element of the x^{th} combination of the set. In general, the PMF for actors in Level 2 and beyond can be expressed in terms of the set S as follows:

$$f_j(e) = \sum_{\substack{k=1 \\ s_i^k \in S(N,e)}}^{x} \prod_{i \in A_j'} \sum_{e=0}^{\omega} p(s_i^k|e) f_i(e) \tag{5.11}$$

where $f_i(e)$ denotes the PMF for actor i in the set A_j'. The computations associated with actor 8 under the rule of maximum influence are shown for illustrative purposes. As mentioned above, for actor 8 to receive a state of zero, both actors 4 and 5 must pass a state of zero. The probability of actor 8 being in state zero under the rule of maximum influence is

$$\begin{aligned} f_{max,8}(0) = {} & [p(0|0)f_4(0) + p(0|1)f_4(1) + p(0|2)f_4(2) + p(0|3)f_4(3)] \times \\ & [p(0|0)f_5(0) + p(0|1)f_5(1) + p(0|2)f_5(2) + p(0|3)f_5(3)] \end{aligned} \tag{5.12}$$

For this example, there are three ways in which actor 8 can receive an influence level of one, and that probability is computed by

$$
\begin{aligned}
f_{max,8}(1) = {} & [p(0|0)f_4(0) + p(0|1)f_4(1) + p(0|2)f_4(2) + p(0|3)f_4(3)]\times \\
& [p(1|0)f_5(0) + p(1|1)f_5(1) + p(1|2)f_5(2) + p(1|3)f_5(3)] \\
& +[p(1|0)f_4(0) + p(1|1)f_4(1) + p(1|2)f_4(2) + p(1|3)f_4(3)]\times \\
& [p(0|0)f_5(0) + p(0|1)f_5(1) + p(0|2)f_5(2) + p(0|3)f_5(3)] \\
& +[p(1|0)f_4(0) + p(1|1)f_4(1) + p(1|2)f_4(2) + p(1|3)f_4(3)]\times \\
& [p(1|0)f_5(0) + p(1|1)f_5(1) + p(1|2)f_5(2) + p(1|3)f_5(3)]
\end{aligned}
\tag{5.13}
$$

There are five ways in which actor 8 can receive an influence level of two, and that probability is given by

$$
\begin{aligned}
f_{max,8}(2) = {} & [p(0|0)f_4(0) + p(0|1)f_4(1) + p(0|2)f_4(2) + p(0|3)f_4(3)]\times \\
& [p(2|0)f_5(0) + p(2|1)f_5(1) + p(2|2)f_5(2) + p(2|3)f_5(3)] \\
& +[p(1|0)f_4(0) + p(1|1)f_4(1) + p(1|2)f_4(2) + p(1|3)f_4(3)]\times \\
& [p(2|0)f_5(0) + p(2|1)f_5(1) + p(2|2)f_5(2) + p(2|3)f_5(3)] \\
& +[p(2|0)f_4(0) + p(2|1)f_4(1) + p(2|2)f_4(2) + p(2|3)f_4(3)]\times \\
& [p(0|0)f_5(0) + p(0|1)f_5(1) + p(0|2)f_5(2) + p(0|3)f_5(3)] \\
& +[p(2|0)f_4(0) + p(2|1)f_4(1) + p(2|2)f_4(2) + p(2|3)f_4(3)]\times \\
& [p(1|0)f_5(0) + p(1|1)f_5(1) + p(1|2)f_5(2) + p(1|3)f_5(3)] \\
& +[p(2|0)f_4(0) + p(2|1)f_4(1) + p(2|2)f_4(2) + p(2|3)f_4(3)]\times \\
& [p(2|0)f_5(0) + p(2|1)f_5(1) + p(2|2)f_5(2) + p(2|3)f_5(3)]
\end{aligned}
\tag{5.14}
$$

Finally, there are seven ways in which actor 8 can receive an influence level of three, and the probability computation is given by

$$\begin{aligned}
f_{max,8}(3) = {} & [p(0|0)f_4(0) + p(0|1)f_4(1) + p(0|2)f_4(2) + p(0|3)f_4(3)]\times \\
& [p(3|0)f_5(0) + p(3|1)f_5(1) + p(3|2)f_5(2) + p(3|3)f_5(3)] \\
& +[p(1|0)f_4(0) + p(1|1)f_4(1) + p(1|2)f_4(2) + p(1|3)f_4(3)]\times \\
& [p(3|0)f_5(0) + p(3|1)f_5(1) + p(3|2)f_5(2) + p(3|3)f_5(3)] \\
& +[p(2|0)f_4(0) + p(2|1)f_4(1) + p(2|2)f_4(2) + p(2|3)f_4(3)]\times \\
& [p(3|0)f_5(0) + p(3|1)f_5(1) + p(3|2)f_5(2) + p(3|3)f_5(3)] \\
& +[p(3|0)f_4(0) + p(3|1)f_4(1) + p(3|2)f_4(2) + p(3|3)f_4(3)]\times \\
& [p(0|0)f_5(0) + p(0|1)f_5(1) + p(0|2)f_5(2) + p(0|3)f_5(3)] \\
& +[p(3|0)f_4(0) + p(3|1)f_4(1) + p(3|2)f_4(2) + p(3|3)f_4(3)]\times \\
& [p(1|0)f_5(0) + p(1|1)f_5(1) + p(1|2)f_5(2) + p(1|3)f_5(3)] \\
& +[p(3|0)f_4(0) + p(3|1)f_4(1) + p(3|2)f_4(2) + p(3|3)f_4(3)]\times \\
& [p(2|0)f_5(0) + p(2|1)f_5(1) + p(2|2)f_5(2) + p(2|3)f_5(3)] \\
& +[p(3|0)f_4(0) + p(3|1)f_4(1) + p(3|2)f_4(2) + p(3|3)f_4(3)]\times \\
& [p(3|0)f_5(0) + p(3|1)f_5(1) + p(3|2)f_5(2) + p(3|3)f_5(3)]
\end{aligned} \tag{5.15}$$

The resulting PMF and expected influence states for actors in Level 3 under the rules of maximum and minimum influence are shown Table 5.4 and 5.5, respectively, and the resulting PMF and expected influence states for the target actors under the rules of maximum and minimum influence are shown Table 5.6 and 5.7, respectively.

Note that influence is modeled as flowing directly from the target actors into the sink node, t. Therefore, the expected influence value at the target node, $E(d)$ is the sum of the expected influence values for the target actors. In this example, under the rule of maximum influence, $E(d) = 8.1844$, and under the rule of minimum influence, $E(d) = 0.4036$.

State	$i = 7$	$i = 8$	$i = 9$	$i = 10$
0	0.0664	0.0034	0.0027	0.0536
1	0.1378	0.0284	0.0232	0.1131
2	0.2255	0.1228	0.1100	0.2112
3	0.5704	0.8454	0.8641	0.6221
$E(e)$	2.2999	2.8103	2.8354	2.4019

Table 5.4: Level 3 PMFs and Expected Influence States under Maximum Influence

State	$i = 7$	$i = 8$	$i = 9$	$i = 10$
0	0.2386	0.5242	0.5700	0.3119
1	0.2523	0.3112	0.2996	0.2849
2	0.2345	0.1328	0.1097	0.2247
3	0.2746	0.0319	0.0207	0.1785
$E(e)$	1.5451	0.6724	0.5812	1.2698

Table 5.5: Level 3 PMFs and Expected Influence States under Minimum Influence

State	$i = 11$	$i = 12$	$i = 13$
0	0.0134	0.0001	0.0111
1	0.0751	0.0067	0.0658
2	0.1949	0.0672	0.1841
3	0.7165	0.9260	0.7389
$E(e)$	2.6146	2.9190	2.6508

Table 5.6: Target PMFs and Expected Influence States under Maximum Influence

State	$i = 11$	$i = 12$	$i = 13$
0	0.8174	0.9748	0.8621
1	0.1499	0.0246	0.1188
2	0.0291	0.0006	0.0176
3	0.0037	5.78E-06	0.0016
$E(e)$	0.2191	0.0258	0.1586

Table 5.7: Target PMFs and Expected Influence States under Minimum Influence

We can evaluate the original network shown in section 4.3.2 by creating a set of dummy nodes in Level 2 of the hierarchical network as shown in Figure 5.6. In this case, the influence levels passed from actors 0 and 3, flow through nodes 0.1 and 3.1, respectively. There is no uncertainty associated with the passing of influence through the dummy node. Under the rule of maximum influence, the expected influence value at the target node increases by 3.4% to $E(d) = 8.4648$, and under the rule of minimum influence, the expected influence value at the target node decreases by 11.2% to $E(d) = 0.3582$. The updated PMFs for actors 7, 10, 11, 12, and 13 are shown in Tables 5.8 and 5.9.

We can also evaluate the difference between the two networks using network reliability Under the rule of maximum influence and demand level $d = 9$, the original model has a network reliability of $R(d) = 0.6254$, and the hierarchical model has a network reliability of $R(d) = 0.5013$. Removing the arc between 0-7 and 3-10 results in a 19.8% reduction in network reliability. Under the rule of minimum influence and demand level $d = 1$, the original model has a network reliability of $R(d) = 0.3257$, and the hierarchical model has a network reliability of $R(d) = 0.3541$. In this case, removing the arc between 0-7 and 3-10 results in an increase of 8.7% in network reliability.

State	$i = 7$	$i = 10$	$i = 11$	$i = 12$	$i = 13$
0	0.0033	0.0027	0.0052	2.40 E-05	0.0047
1	0.0273	0.0223	0.0428	0.0021	0.0393
2	0.1198	0.1073	0.1538	0.0364	0.1472
3	0.8496	0.8677	0.7983	0.9615	0.8089
$E(e)$	2.8157	2.8401	2.7452	2.9593	2.7603

Table 5.8: Updated PMFs for Original Network Under Maximium Influence

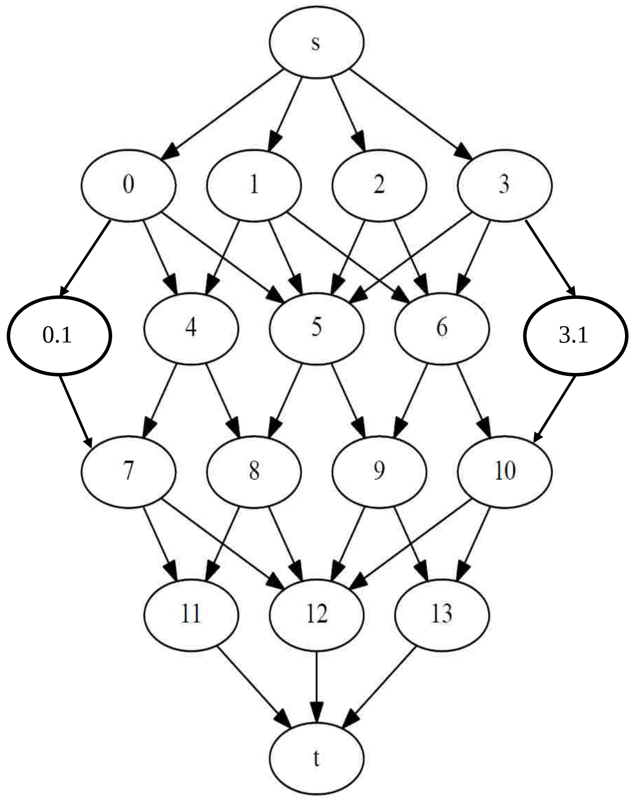


Figure 5.6: Original Network Modeled as Hierarchical Network

State	$i = 7$	$i = 10$	$i = 11$	$i = 12$	$i = 13$
0	0.2767	0.3463	0.8340	0.9791	0.8740
1	0.2906	0.3110	0.1400	0.0205	0.1106
2	0.2543	0.2267	0.0237	0.0004	0.0143
3	0.1785	0.1160	0.0028	2.44 E-06	0.0010
$E(e)$	1.3346	1.1124	0.1945	0.0213	0.1424

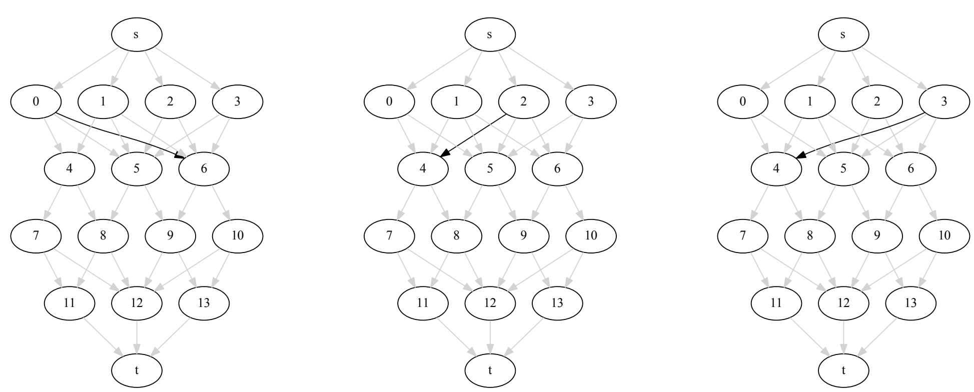
Table 5.9: Updated PMFs for Original Network Under Minimum Influence

5.3.3 Altering Communication Paths

In this section, we begin to address the research question posed in section 5.2. That is, if resources exist to add or remove connections within the network, where should those changes be made? As demonstrated in Figure 5.4, under the rule of maximum influence, adding arcs increases the expected influence of an actor, and it is demonstrated in Figure 5.5 that the converse is true under the rule of minimum influence. Suppose decision makers associated with an influence campaign have the resources to alter communication paths within the network. That is, if resources exist to add (or remove) connections within the network, where should those connections be added (or removed)? We begin by investigating the addition (or removal) of a single arc between existing layers. To assess the impact of altering the communication paths we consider the expected influence value at the target node, $E(d)$.

Adding a Single Arc between Level 1 and Level 2

Here, the focus is on adding a single arc between actors in Level 1 and Level 2. For the network shown in Figure 5.2, there are three possible scenarios in which a single connection can be added. These scenarios are presented in Figure 5.7.



1 2 3

Figure 5.7: Scenarios Adding Communication Paths Between Levels 1 and 2

Under the rule of maximum influence, the expected value of influence at the target is

8.184. Adding an arc from actor 0 to actor 6 (scenario 1) results in an expected value of influence at the target of 8.199 which is an increase of 0.18% from the baseline scenario. However, adding an arc from actor 2 to actor 4 (scenario 2) or actor 3 to actor 4 (scenario 3) results in an expected value of influence at the target is 8.227 which in an increase of 0.53% from the baseline scenario. These results are summarized in Table 5.10. Recall from Figure 5.4 that adding arcs under the rule of maximum influence increases expected influence at a decreasing rate. Therefore, we present the following rule:

Rule 1: In the case where actors in Level 1 receive perfect information from the source and communicate under the rule of maximum influence, adding a single arc to the actor with the lowest value of N maximizes expected influence at the target node.

Actor	**Baseline**	**Add 0-6**	**Add 2-4**	**Add 3-4**
4	2.853	2.853	2.954	2.954
5	2.984	2.984	2.984	2.984
6	2.954	2.984	2.954	2.954
7	2.300	2.300	2.402	2.402
8	2.810	2.810	2.835	2.835
9	2.835	2.844	2.835	2.835
10	2.402	2.434	2.402	2.434
11	2.615	2.615	2.651	2.651
12	2.919	2.922	2.926	2.926
13	2.651	2.663	2.651	2.651
$E(d)$	8.184	8.199	**8.227**	**8.227**
% increase	n/a	0.18%	**0.53%**	**0.53%**

Table 5.10: Expected Influence States when adding a single connection from Level 1 to 2

In the previous chapter, the performance metric under study was network reliability. Therefore, it may also be interesting to consider the effects on network reliability when adding arcs. Recall that network reliability is defined as the probability that at least d units of influence are received at the target node. Network reliability is evaluated using Monte Carlo simulation. Refer to section 4.3 for details on the simulation model. Because reliability is a proportion, the number of replications on the MC simulation model can be selected by specifying the worst case half-width of the associated confidence interval (based on the s-normal approximation of the binomial distribution). To have a worst-case half-width of 0.001 on a 95% confidence interval on network reliability, 960,400 replications were performed.

Under the rule of maximum influence, the network reliability at $d = 9$ for the baseline scenario is estimated to be 0.4997. Adding an arc under scenario 1 results in an estimated network reliability of 0.5068. This is an increase of 1.42% from the baseline scenario. However, adding an arc under scenario 2 or scenario 3 results in an estimated network reliability of 0.5208 and 0.5198, respectively. Note the difference between these two scenarios is not statistically significant and results in approximately a 4% increase in network reliability. For this example, the rule that maximizes the expected value of influence at the target node also maximizes network reliability.

Baseline	**Add 0-6**	**Add 2-4**	**Add 3-4**
0.4997	0.5068	**0.5208**	**0.5198**

Table 5.11: Network Reliability when adding a single connection from Level 1 to 2

Adding a Single Arc between Level 2 and Level 3

In this section, the focus is on adding a single arc between actors in Level 2 and Level 3. For the network shown in Figure 5.2, there are six possible scenarios which are presented in Figure 5.8. The expected value of influence at the target node is computed for each of the six scenarios. The expected value of influence state for each actor, the total expected influence at the target, and the percent increase from the baseline scenario is presented in Table 5.12. For this example, the expected value of influence at the target node is maximized when a single arc is added from actor 5 to actor 7. Adding this arc results in an increase of 1.88%.

When making connections between Levels 2 and 3, the cardinality of the predecessors for the actors involved in the new connection must be taken into account. Emperical evidence suggests that the following rule should be employed maximize the expected value of influence at the target node:

> ***Rule 2:*** Identify the actor in Level 3 with the lowest value of N, and add an arc from the most connected actor in the previous level.

Again, it may be interesting to consider the effects on network reliability under this scenario. In this case, adding an arc from actor 5 to actor 7 also maximizes network reliability. Adding this arc results in an increase of 13.11% over the baseline scenario.

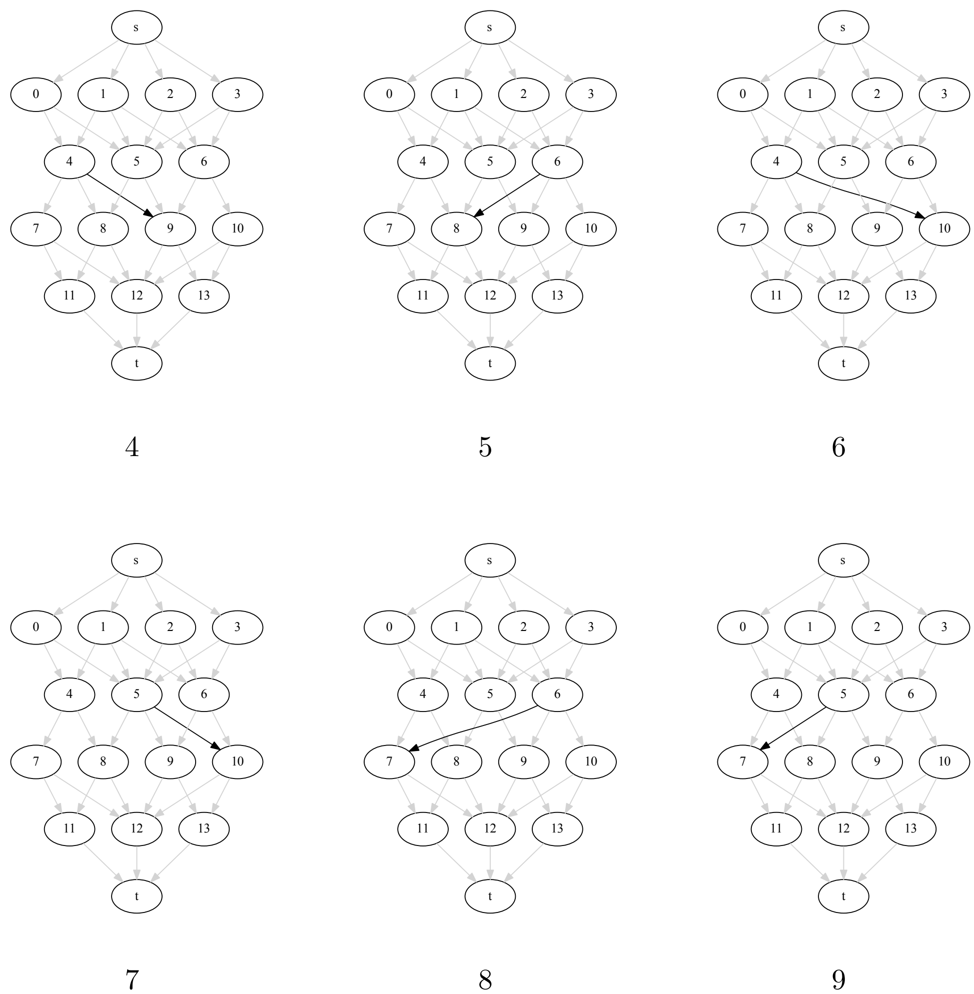


Figure 5.8: Scenarios Adding a Single Communication Path Between Levels 2 and 3

Actor	Add 4-9	Add 6-8	Add 4-10	Add 5-10	Add 6-7	Add 5-7
4	2.853	2.853	2.853	2.853	2.853	2.853
5	2.984	2.984	2.984	2.984	2.984	2.984
6	2.954	2.954	2.954	2.954	2.954	2.954
7	2.300	2.300	2.300	2.300	2.800	2.810
8	2.810	2.936	2.810	2.810	2.810	2.810
9	2.936	2.835	2.835	2.835	2.835	2.835
10	2.402	2.402	2.800	2.835	2.402	2.402
11	2.615	2.676	2.615	2.615	2.740	2.743
12	2.929	2.931	2.939	2.942	2.943	2.943
13	2.697	2.651	2.748	2.759	2.651	2.651
$E(d)$	8.240	8.258	8.302	8.315	8.334	**8.338**
% increase	0.50%	0.72%	1.26%	1.41%	1.65%	**1.70%**

Table 5.12: Expected Influence States when adding a single connection from Level 2 to 3

Add 4-9	Add 6-8	Add 4-10	Add 5-10	Add 6-7	Add 5-7
0.5240	0.5329	0.5502	0.5521	0.5639	**0.5652**

Table 5.13: Network Reliability when adding a single connection from Level 2 to 3

Adding a Single Arc Between Level 3 and Target Actors

One of the primary assumptions of the models developed in this dissertation is that the initiators of an influence campaign do not have direct access to the target actors. Therefore, the case of adding a single arc between level 3 and the target actors is not addressed for this example.

Removing a Single Arc between Level 1 and Level 2

When actors communicate under the rule of minimum influence, it may be beneficial to reduce the number of communications an actor has with the previous level. Here, the focus is on removing a single arc between actors in Level 1 and Level 2. For the network shown in Figure 5.2, we will consider three scenarios, remove a single arc from actor 4 (by removing either 0-4 or 1-4), remove a single arc from actor 5 (by removing either 0-5, 1-5, 2-5, or 3-5) and remove a single arc from actor 6 (by removing either 1-6, 2-6, or 3-6). To investigate the effects associated with the removal of arcs, we again consider the expected value of influence at the target. Under the baseline scenario, $E(d) = 0.404$. Removing a single arc from actor 4 results in an expected value of influence of the target of $E(d) = 0.509$ which is a 26% increase over the baseline. Removing a single arc from actors 5 and 6, results in an expected value of influence at the target node of 0.471 and 0.473, respectively. These represent increases of 16.6% and 17%. These results are summarized in Table 5.16. Recall from Figure 5.5 that removing arcs under the rule of minimum influence increases expected influence at an increasing rate. Therefore, we present the following rule:

Rule 3: In the case where actors in Level 1 receive perfect information from the source and communicate under the rule of minimum influence, removing a single arc to the actor with the lowest value of N maximizes expected influence at the target node.

Actor	**Baseline**	**Actor 4**	**Actor 5**	**Actor 6**
4	2.048	2.450	2.408	2.048
5	1.515	1.515	1.746	1.515
6	1.746	1.746	1.746	2.048
7	1.545	1.923	1.545	1.545
8	0.672	0.790	0.778	0.672
9	0.581	0.581	0.668	0.672
10	1.270	1.270	1.270	1.545
11	0.219	0.315	0.255	0.219
12	0.026	0.035	0.033	0.034
13	0.159	0.159	0.183	0.219
$E(d)$	0.404	**0.509**	0.471	0.473
% increase	n/a	**25.99%**	16.58%	17.08%

Table 5.14: Expected Influence States when removing a single connection from Level 1 to 2

Again, it may be interesting to consider the effects on network reliability when removing arcs. To evaluate this performance metric under the rule of minimum influence, the level for the network reliability evaluation is set to $d = 1$. This value is chosen because at larger values of d, the reliability values are very small and not statistically different. Under the rule of minimum influence, the network reliability at $d = 1$ for the baseline scenario (scenario 0) is estimated to be 0.354. Removing an arc from actor 4 results in an estimated network reliability of approximately 0.394 which is an increase of 11% over the baseline scenario. However, removing an arc from actor 5 results in an estimated network reliability of approximately 0.398 which is an increase of 12%, and finally, removing an arc from actor 6 results in an estimate network reliability of approximately 0.382 which is an increase of 8%. These results are summarized in Tables 5.15 and 5.16. For this example, removing a connection to actor 5 maximizes network reliability.

Arc Removed	**Reliability**
0-4	0.3943
1-4	0.3944
0-5	0.3983
1-5	0.3984
2-5	0.3984
3-5	0.3988
1-6	0.3823
2-6	0.3812
3-6	0.3835

Table 5.15: Network Reliability when removing a single connection from Level 1 to 3

Removing a Single Arc between Level 2 and Level 3

For this example, removing the connection between actor 4 and 7 or actor 6 and 10 would, in essence, remove an entire actor from the network. Therefore, in this section, we only consider removing arcs that do not result in the removal of an actor. Specifically, we investigate the effects for removing arcs 4-8, 5-8, 5-9 or 6-9 on expected value of influence at the target.

For this example, removing arc 4-8 leaves actor 8 connected to actor 5. Note that the cardinality of the set A_5' is $N = 4$. This results in an expected value of influence at the target node of $E(d) = 0.549$ which is an increase of 36% over the baseline scenario. However removing arc 5-8 leaves actor 8 connected to actor 4. Note that the cardinality of A_4' is

Actor	Baseline	Remove 4-8	Remove 5-8	Remove 5-9	Remove 6-9
4	2.048	2.048	2.048	2.048	2.048
5	1.515	1.515	1.515	1.515	1.515
6	1.746	1.746	1.746	1.746	1.746
7	1.545	1.545	1.545	1.545	1.545
8	0.672	1.604	1.545	0.672	0.672
9	0.581	0.581	0.581	1.270	1.064
10	1.270	1.270	1.270	1.270	1.270
11	0.219	0.352	0.513	0.219	0.219
12	0.026	0.038	0.052	0.051	0.044
13	0.159	0.159	0.159	0.348	0.292
$E(d)$	0.404	0.549	**0.724**	0.618	0.555
% increase	n/a	35.89%	**79.21%**	52.97%	37.38%

Table 5.16: Expected Influence States when removing a single connection from Level 2 to 3

$N = 2$. This results in an expected value of influence at the target node of $E(d) = 0.724$ which is an increase of 79% over the baseline. Removing arc 5-9 leaves actor 9 connected to actor 6. Note that the cardinality of the set A'_6 is $N = 3$. This results in an expected value of influence at the target node of $E(d) = 0.618$ which is an increase of 53% over the baseline scenario. Removing arc 6-9 leaves actor 9 connected to actor 5. This results in an expected value of influence at the target node of $E(d) = 0.555$ which is an increase of 37% over the baseline scenario. When removing connections between Levels 2 and 3, the cardinality of the predecessors for the actors involved in the removal must be taken into account. Empirical evidence suggests that the following rule should be employed to maximize the expected value of influence at the target node:

> ***Rule 4:*** Remove an arc such that the cardinality of the set A'_j for an actor in Level 3 is as small as possible.

As with the previous examples, it may be interesting to investigate the effects of removing an arc on the network reliability. We again consider $d = 1$ so that the reliability values obtained from the simulation model are statistically different, and the results are shown in Table 5.17. Recall for the baseline scenario, network reliability at $d = 1$, is estimated to be 0.3540. Removing arcs 4-8 and 6-9 resulted in an estimated network reliability of approximately 0.844 which was an increase of approximately 9%. Removing arc 5-9 resulted in a network reliability of approximately 0.5071 which is an increase of 43%, and removing arc 5-8 resulted in a 55% increase in network reliability. For this example, the removing arc 5-8 maximizes both the expected value of influence at the target node and network reliability.

Arc 4-8	**Arc 5-8**	**Arc 5-9**	**Arc 6-9**
0.3884	0.5503	0.5071	0.3883

Table 5.17: Network Reliability when removing a single connection from Level 2 to 3

Removing a Single Arc Between Level 3 and Target Actors

One of the primary assumptions of the models developed in this dissertation is that the initiators of an influence campaign do not have direct access to the target actors. Therefore, the case of removing a single arc between level 3 and the target actors is not addressed for this example.

5.4. Reducing Uncertainty in Influence Levels

The focus of this section is to investigate the effects on the expected value of influence at the target node, $E(d)$, when the uncertainty associated with an actor passing influence through the network is eliminated. The hierarchical network shown in Figure 5.2, the conditional probabilities associated with strength of influence shown in Table 5.1, and the rules of maximum and minimum influence are still considered. A one-at-a-time analysis is performed for each actor in the network in which the uncertainty associated with passing information is eliminated as shown in Table 5.18

For example, consider the actors in Level 1. Each actor is assumed to receive perfect information from the initiator of the influence campaign (influence state 3) and passes an

e_i	$p(e\|e_i)$			
	0	1	2	3
0	1	0	0	0
1	0	1	0	0
2	0	0	1	0
3	0	0	0	1

Table 5.18: Eliminate Uncertainty for a given actor

influence state based on the probability distribution from Figure 5.1. Suppose, however, that resources exist to persuade an actor from Level 1 to pass along influence in the same state it was received. The effects on the expected value of influence at the target node under the rule of maximum influence and minimum influence are shown below in Tables 5.19 and 5.20, respectively. In both cases, eliminating the uncertainty associated with actor 1, the actor with the most connections to Level 2 results in the largest increase in the expected value of influence at the target node.

	Baseline	Actor 0	Actor 1	Actor 2	Actor 3
0	2.450	3.000	2.450	2.450	2.450
1	2.450	2.450	3.000	2.450	2.450
2	2.450	2.450	2.450	3.000	2.450
3	2.450	2.450	2.450	2.450	3.000
4	2.853	3.000	3.000	2.853	2.853
5	2.984	3.000	3.000	3.000	3.000
6	2.954	2.954	3.000	3.000	3.000
7	2.300	2.450	2.450	2.300	2.300
8	2.810	2.853	2.853	2.816	2.816
9	2.835	2.840	2.853	2.853	2.853
10	2.402	2.402	2.450	2.450	2.450
11	2.615	2.671	2.671	2.617	2.617
12	2.919	2.930	2.933	2.924	2.924
13	2.651	2.653	2.671	2.671	2.671
$E(d)$	8.184	8.254	8.275	8.212	8.212
% increase	n/a	0.85%	1.11%	0.33%	0.33%

Table 5.19: Expected Influence States Under the Rule of Maximum Influence when Removing Uncertainty from Level 1 actors

Now consider the actors in Level 2 of the network. In general, each actor passes an

	Baseline	Actor 0	Actor 1	Actor 2	Actor 3
0	2.450	3.000	2.450	2.450	2.450
1	2.450	2.450	3.000	2.450	2.450
2	2.450	2.450	2.450	3.000	2.450
3	2.450	2.450	2.450	2.450	3.000
4	2.048	2.450	2.450	2.048	2.048
5	1.515	1.746	1.746	1.746	1.746
6	1.746	1.746	2.048	2.048	2.048
7	1.545	1.923	1.923	1.545	1.545
8	0.672	0.922	0.922	0.778	0.778
9	0.581	0.668	0.778	0.778	0.778
10	1.270	1.270	1.545	1.545	1.545
11	0.219	0.371	0.371	0.255	0.255
12	0.026	0.046	0.062	0.045	0.045
13	0.159	0.183	0.255	0.255	0.255
$E(d)$	0.404	0.600	0.688	0.555	0.555
% increase	n/a	48.6%	70.4%	37.4%	37.4%

Table 5.20: Expected Influence States Under the Rule of Minimum Influence when Removing Uncertainty from Level 1 actors

influence state based on the probability distribution from Figure 5.1. Suppose, however, that resources exist to persuade an actor from Level 2 to pass along influence in the same state it was received. The effects on the expected value of influence at the target node under the rule of maximum influence and minimum influence are shown below in Tables 5.21 and 5.22, respectively. Under the rule of maximum influence, the expected value of influence at the target nodes is maximized when the uncertainty associated with actor 6 is eliminated. This is followed by actor 4 then actor 5. Under the rule of minimum influence, the order is reversed. Reducing the uncertainty of actor 5 maximizes expected value of influence at the target node followed by actor 4 then actor 6.

Finally, suppose that resources exist to persuade an actor from Level 3 to pass along influence in the same state it was received. The effects on the expected value of influence at the target node under the rule of maximum influence and minimum influence are shown below in Tables 5.23 and 5.24, respectively. Under the rule of maximum influence, the expected value of influence at the target nodes is maximized when the uncertainty associated with

	Actor 4	Actor 5	Actor 6
0	2.450	2.450	2.450
1	2.450	2.450	2.450
2	2.450	2.450	2.450
3	2.450	2.450	2.450
4	2.853	2.853	2.853
5	2.984	2.984	2.984
6	2.954	2.954	2.954
7	2.853	2.300	2.300
8	2.952	2.993	2.810
9	2.835	2.994	2.984
10	2.402	2.402	2.954
11	2.800	2.706	2.615
12	2.955	2.950	2.959
13	2.651	2.726	2.835
$E(d)$	8.405	8.382	8.409
% increase	2.70%	2.41%	2.74%

Table 5.21: Expected Influence States Under the Rule of Maximum Influence when Removing Uncertainty from Level 2 actors

	Actor 4	Actor 5	Actor 6
0	2.450	2.450	2.450
1	2.450	2.450	2.450
2	2.450	2.450	2.450
3	2.450	2.450	2.450
4	2.048	2.048	2.048
5	1.515	1.515	1.515
6	1.746	1.746	1.746
7	2.048	1.545	1.545
8	0.847	0.935	0.672
9	0.581	0.803	0.766
10	1.270	1.270	1.746
11	0.365	0.309	0.219
12	0.041	0.049	0.045
13	0.159	0.221	0.288
$E(d)$	0.564	0.579	0.553
% increase	39.8%	43.4%	37.0%

Table 5.22: Expected Influence States Under the Rule of Minimum Influence when Removing Uncertainty from Level 2 actors

actor 8 is eliminated. This is followed by actor 9 then actor 7 then actor 10. Under the rule of minimum influence, the order is reversed for the last two actors. Reducing the uncertainty of actor 8 maximizes expected value of influence at the target node followed by actor 9 then actor 10 then actor 7.

	Actor 7	Actor 8	Actor 9	Actor 10
0	2.450	2.450	2.450	2.450
1	2.450	2.450	2.450	2.450
2	2.450	2.450	2.450	2.450
3	2.450	2.450	2.450	2.450
4	2.853	2.853	2.853	2.853
5	2.984	2.984	2.984	2.984
6	2.954	2.954	2.954	2.954
7	2.300	2.300	2.300	2.300
8	2.810	2.810	2.810	2.810
9	2.835	2.835	2.835	2.835
10	2.402	2.402	2.402	2.402
11	2.757	2.889	2.615	2.615
12	2.946	2.974	2.976	2.950
13	2.651	2.651	2.909	2.796
$E(d)$	8.354	8.514	8.500	8.360
% increase	2.07%	4.02%	3.85%	2.15%

Table 5.23: Expected Influence States Under the Rule of Maximum Influence when Removing Uncertainty from Level 3 actors

5.5. Conclusions and Future Work

In this chapter, a hierarchical network in which actor influence levels are multi-valued, stochastic, and dependent upon the influence they receive from their predecessors was investigated. To address the two research questions presented in section 5.2 a new performance metric, the expected value of influence at the target node, $E(d)$ was defined. A closed-form expression for the PMF function for actor influence states was developed, and the expected actor influence state was computed for each actor in the network. Since influence flow is modeled as flowing directly from the target actors into the sink node, the expected influence

	Actor 7	Actor 8	Actor 9	Actor 10
0	2.450	2.450	2.450	2.450
1	2.450	2.450	2.450	2.450
2	2.450	2.450	2.450	2.450
3	2.450	2.450	2.450	2.450
4	2.048	2.048	2.048	2.048
5	1.515	1.515	1.515	1.515
6	1.746	1.746	1.746	1.746
7	1.545	1.545	1.545	1.545
8	0.672	0.672	0.672	0.672
9	0.581	0.581	0.581	0.581
10	1.270	1.270	1.270	1.270
11	0.286	0.348	0.219	0.219
12	0.033	0.040	0.042	0.035
13	0.159	0.159	0.258	0.216
$E(d)$	0.477	0.547	0.519	0.470
% increase	18.3%	35.5%	28.5%	16.5%

Table 5.24: Expected Influence States Under the Rule of Minimum Influence when Removing Uncertainty from Level 3 actors

value at the target node was defined as the sum of the expected influence states for the target actors.

Under the rule of maximum influence, adding arcs between Levels 1 and 2 resulted in an increase in expected influence value for each actor, while adding arcs resulted in an decrease in expected influence value for each actor under the rule of minimum influence. Based on empirical evidence from initial experimentation, it was conjectured that the expected influence states are concave under the rule of maximum influence and convex under the rule of minimum influence. A series of simple rules were presented to assist decision-makers when resources existed to either add or remove a single communication path.

Finally, the expected value of influence at the target node was investigated for the case in which the uncertainty associated with passing influence for a single actor was eliminated. For actors in Level 1, expected influence at the target node was maximized when the uncertainty was eliminated for the actor with the most connections to Level 2 under both influence rules. For actors in Levels 2 and 3 a simple rule did not seem to apply, and the results were dependent up on the influence rule selected.

This chapter presents many opportunities for future work both with modeling efforts and solution techniques. In this chapter, actors within a level did not communicate with one another, and communication did not "skip" a level. In reality, such a strict hierarchical model probably does not exist. More realistic models will include communication between actors in a given level, the ability of an actor to communicate with actors across levels and perhaps include feedback mechanisms. Now that performance metric for which a closed-form expression has been developed, more sophisticated solutions techniques can be investigated for the decision-making process that do not require for simulation.

Chapter 6

Eliciting Influence Probabilities for a Social Network

6.1. Introduction

In the earlier chapters of this dissertation, influence within a social network is defined as an actor's ability "to induce a change in behavior of another that conforms to the influencing actor's desires" [21, 22]. In each chapter, levels of influence are defined, and the probabilities associated with an actor being in each influence level is given. It is assumed that both the influence levels and the probability associated with an actor having a particular level of influence have been determined by experts familiar with the organization under study. However, no formal method for identifying influence levels and/or their associated probability distributions is currently available in the literature. In this chapter, a method for eliciting influence probabilities within a college classroom environment is developed. It is worthwhile to note that the methodology developed in this chapter would not apply directly to a clandestine or terrorist organization since those organizations, by definition, attempt to hide their social structure. However, this methodology does provide a baseline for quantifying perceived influence levels and the probabilities associated with actors being in each level. It also provides a mechanism to quantify the difference between an actor's perceived influence and actual influence through a set of experiments.

When developing a social network for use in an educational setting, the primary factors to consider are choosing an appropriate level of analysis, ensuring the quality of the data collected, and addressing missing, incomplete, or inaccurate data. The level of analysis of a social network depends on the type of data to be collected. Egocentric analysis focuses on the set of connections of a particular actor. Cognitive network analysis includes not only the connections of a particular actor but also that actor's perceptions regarding the connections between other actors in the network [11]. Whole network analysis involves the collection of data on the relationships between all actors in a bounded network where the network bounds (or actors included in the study) are defined by the analyst [11]. The advantages of this type of analysis include:

> "(1) the identification of the naturally existing peer networks within a given setting or context; (2) the simultaneous collection of data on the units of the system under analysis and on the structures generated by the relations developed among those units; and (3) the identification of indirect ties between and among actors." [11]

To ensure high quality data is collected on the social network of interest, it is important for the researcher to adequately define the network boundaries, identify all relevant actors, ensure the highest response rate possible for surveys relating to the structure of the social network, and address inaccuracies and inconsistencies within the survey responses [11].

6.2. Background Information

The Freshman Engineering Program (FEP) is the first-year experience for students in the College of Engineering at the University of Arkansas. The FEP is divided into two sub programs - an academic program and a student services program. The academic program consists of a two-semester, thirty-credit hour course load including a two-semester sequence of Introduction to Engineering (Intro), two math classes, science classes, and university core requirements. The math and science classes are based on student's K-12 preparation. For the 2011-2012 school year seven teaching assistants (TA) supported the FEP. Each section of Intro was assigned a primary TA who served as the main point of contact for students in their section. The primary TA was responsible for attending class each day and grading assignments. All FEP TA hold office hours and assist students as necessary.

The student services program compliments the academic program by providing academic assistance, professional development, academic advising, and peer mentoring. The peer mentoring component of FEP is required for all students in the Intro course sequence. Students are matched with an upper-class engineering student and meet for 30 minutes each week during the Fall semester and for the first eight weeks during the Spring semester.

6.3. Research Methodology

In this chapter, the relationships between students in a subset of the Spring 2012 offerings of GNEG 1121 - Introduction to Engineering II are evaluated using the whole network analysis approach briefly described above. Since the focus of this study involved particular sections of the course, the natural boundary of "is enrolled in this section" was used to define the students included in the network. The students that participated in the study were allowed class time to complete the surveys required to build the social network, quantify existing relationships, and identify potential initiators of influence. The following 4 step methodology was used to gather the data required for this study:

1. Establish consent from participants;
2. Determine existing relationships and quantify social closeness;
3. Determine initiators of influence;
4. Determine influence levels and probabilities.

To conform with standard protocols associated with research involving human subjects, a protocol form and survey script were submitted to the IRB Program Manager, and approval was granted for this study. The approval letter for this study is available in the Appendix.

6.3.1 Establish Consent from Participants

The first step in completing this study was to establish participation consent from students in the sections of GNEG 1121 under study. Participation in this study was completely voluntary, and students had the right to withdraw from the study at any point. A copy of the consent form distributed to the students is also available in the Appendix.

6.3.2 Determine Existing Relationships and Quantify Social Closeness

The first step in this analysis was to determine the existing relationships of the actors (students) within the network (GNEG 1121 section) under study. After establishing the consent of the participants, an initial survey was distributed to the students to establish existing relationships. The *roster technique* was employed during the administration of the survey [11]. That is, each student was provided with a complete list of all the students enrolled in the course and then asked to introduce him\herself so students could match a name with a face. Each student was then asked to indicate their affiliation with all of the other students. These affiliations are used to quantify the 'social closeness' of the individuals in the network. For the purposes of this dissertation, social closeness is quantified as follows:

M - This is me.

0 - I do not know this person.

1 - I recognize this person from class.

2 - I have more than one class with this person.

3 - We live in the same residence hall.

4 - We participate in similar extracurricular activities and/or study together.

5 - We associate socially outside of class.

Students were also asked to provide the name of their peer mentor.

6.3.3 Determine the Initiators of Influence

The next step in this analysis was to define the term 'influence' for the purposes of this study and to have the students determine the actors within the network they consider influential. The students were presented with Hamill's definition "to induce a change in behavior of another that conforms to the influencing actor's desires" [21, 22]. It was assumed that the course instructor and primary teaching assistant were influential, and the students were asked to identify which (if any) of the other teaching assistants they considered influential. They were also given the opportunity to identify up to five students in the class they considered influential. A copy of the initial survey is shown in Figure 6.1

6.3.4 Determine Influence Levels and Probabilities

After the initial survey was administered and evaluated, a follow-up survey was created to assess levels of influence. Students were reminded that their participation in the study was completely voluntary. For each initiator of influence identified in the initial survey, an attempt was made to elicit the probability of that initiator being in each state of influence for various methods of communication (i.e. announcement in class, e-mail announcement, direct contact, etc.). In the previous chapters of this dissertation, four levels of influence have been considered. This chapter will continue to use four influence levels defined as follows:

3. Highly Influential

2. Somewhat Influential

1. Barely Influential

0. Not Influential

The questions for the follow-up survey are presented in Figures 6.2 and 6.3.

Initial Survey – Developing the Social Network

For each student in the list below, please indicate your affiliation. Check all that apply.

ID #	Student	This is me.	I do not know this person.	I recognize this person from class.	I have more than one class with this person.	We live in the same residence hall.	We live on the same floor in our residence hall.	We participate in one or more of the same extracurricular activities.	We are study partners.	We associate socially outside of class.
1										
2										
3										
4										

Who is your peer mentor?

Which teaching assistant(s) do you consider influential? Check all that apply.

- Stephanie Clark
- Coby Durham
- Chris Farnell
- John Judkins
- Michael May
- Hayley Moore
- Lora Strother

Identify up to five (5) students from the list above that you consider influential.

Figure 6.1: Initial Survey

Follow-Up Survey – Assessing Levels of Influence

Participation in this survey is voluntary.

- I would like to participate in this survey.
- I do not want to participate in this survey.

Consider the following definition of influence. Influence is the ability "to induce a change in behavior of another that conforms to the influencing actor's desires" (Hamill, 2007). Also, consider that influence may come in levels. Specifically, consider the following influence levels.

1. Highly Influential
2. Somewhat Influential
3. Barely Influential
4. Not Influential

Consider each statement below. For each statement, allocate a total of 10 points. These points should be used to indicate the likelihood that you would perform a specified task given the manner in which it is presented.

Example:

Statement	Highly Influential	Somewhat Influential	Barely Influential	Not Influential	Sum
Instructor makes an announcement in class.	3	3	3	1	10
Instructor sends an e-mail.	4	3	3	0	10

Consider the statements below for the Course Instructor, Ms. Schneider.

Statement	Highly Influential	Somewhat Influential	Barely Influential	Not Influential	Sum
Instructor makes an announcement in class.					10
Instructor sends an e-mail.					10
Instructor makes an announcement in class and sends a follow up e-mail.					10
Instructor provides direct, one-on-one communication.					10

Consider the statements below for the Primary Teaching Assistant.

Statement	Highly Influential	Somewhat Influential	Barely Influential	Not Influential	Sum
Primary TA makes an announcement in class.					10
Primary TA sends an e-mail.					10
Primary TA makes an announcement in class and sends a follow up e-mail.					10
Primary TA provides direct, one-on-one communication.					10

Figure 6.2: Follow Up Survey to Elicit Influence Probabilities - Page 1

Consider the statements below for your peer mentor.

Statement	Highly Influential	Somewhat Influential	Barely Influential	Not Influential	Sum
TA 1 sends an e-mail.					10
TA 1 provides direct, one-on-one communication.					10

Consider the statements below for the following TA (identified as influential from the initial survey).

Statement	Highly Influential	Somewhat Influential	Barely Influential	Not Influential	Sum
TA 1 sends an e-mail.					10
TA 1 provides direct, one-on-one communication.					10

Statement	Highly Influential	Somewhat Influential	Barely Influential	Not Influential	Sum
TA 2 sends an e-mail.					10
TA 2 provides direct, one-on-one communication.					10

Statement	Highly Influential	Somewhat Influential	Barely Influential	Not Influential	Sum
TA 3 sends an e-mail.					10
TA 3 provides direct, one-on-one communication.					10

Consider the statements below for the following students (identified as influential from the initial survey).

Statement	Highly Influential	Somewhat Influential	Barely Influential	Not Influential	Sum
Student 1 sends an e-mail.					10
Student 1 provides direct, one-on-one communication.					10

Statement	Highly Influential	Somewhat Influential	Barely Influential	Not Influential	Sum
Student 2 sends an e-mail.					10
Student 2 provides direct, one-on-one communication.					10

Statement	Highly Influential	Somewhat Influential	Barely Influential	Not Influential	Sum
Student 3 sends an e-mail.					10
Student 3 provides direct, one-on-one communication.					10

Figure 6.3: Follow Up Survey to Elicit Influence Probabilities - Page 2

Section	Students Enrolled	Not Participating	Absent	Total Responses
006	13	0	0	13
H004	28	0	2	26

Table 6.1: Class Sizes and Number of Responses for Initial Survey

6.4. Discussion of Results

A detailed discussion of the results from the methodology and surveys described above is provided in this section. The software UCINET [6] was used to create the sociograms and evaluate the centrality metrics used to determine the existing relationships and potentially identify influential actors within the network. Finally, a few experiments were performed to begin to assess the validity of the elicited influence probabilities.

6.4.1 Establishing Consent and Determining Existing Relationships

Two sections of GNEG 1121 - Introduction to Engineering II were identified for inclusion in this study. The first section selected for study (section 006) had a total enrollment of 14 students and was taught by Dr. Asya Galbraith, and the second (section H004) had a total enrollment of 28 students and was taught by me. These sections were chosen because of their relatively low enrollment. The small class size is more conducive to the administration of surveys and the assessment of data collected. (Note sections of GNEG 1121 had an enrollment capacity of 55). On January 26, 2012, students in the two sections were provided information on the study and given a copy of the consent form. The class sizes and number of survey responses for each class is shown in Table 6.1.

The students that were present and agreed to participate in the study completed the initial survey form, and a social closeness matrix was created for each course section in the study. The matrix for section 006 is shown in Table 6.2, and the matrix for section H004 is shown in Table 6.3. The rows in the table are the respondents, and the columns are their classmates.

	1	3	4	5	6	7	8	9	10	11	12	13	14
1	M	0	0	1	0	0	0	2	0	0	0	0	3
3	1	M	0	1	0	5	0	0	1	0	0	0	5
4	0	1	M	0	0	0	1	0	0	0	0	0	1
5	0	0	0	M	0	0	1	0	0	0	1	1	0
6	0	0	0	0	M	0	0	0	0	0	0	0	1
7	0	5	0	0	0	M	0	0	3	0	0	0	3
8	0	0	0	0	0	0	M	0	0	0	0	0	0
9	1	1	1	2	1	1	3	M	1	2	1	1	2
10	1	1	0	1	0	3	2	2	M	1	1		5
11	0	0	1	0	0	0	0	0	1	M	0	0	0
12	0	0	0	0	0	0	0	0	0	0	M	1	0
13	0	0	0	0	0	0	0	0	0	0	0	M	0
14	3	5	0	0	0	0	0	0	5	0	0	0	M

Table 6.2: Social Closeness Matrix for GNEG 1121-006

There are several interesting things to note from Table 6.2. Actor 2 withdrew from the course, so he has been omitted from the matrix. Actor 9 indicated that he recognized each person in the class, and both actors 8 and 13 indicated that they did not know anyone in the class. The next thing to note is that not all reported relationships are reciprocal. For instance, actor 1 reports that he recognizes actor 5 from class, but actor 5 reports that he does not know actor 1, and actor 7 reports that he lives in the same residence hall as actor, but actor 14 reports that he does not know actor 7. Similar discrepancies exist in the data for section H004.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
1	M	1	0	3	0	0	0	2	0	0	1	0	0	0	0	2		0	0	5	0	0	0	0	0	0	0	1
2		M	0	0	1	0	0	0	0	1	0	1	0	0	0	1		5	0	0	0	0	0	0	0	0	0	0
3	0	0	M	1	2	3	0	2	2	2	1	2	0	1	0			2	0	0	1	1	0	0	0	0	1	1
4	1	0	0	M	0	1	0	0	0	0	0	0	5	1	0	0	0	0	0	5	0	0	1	1	1	1	5	0
5	0	0	0	0	M	1	0	0	0	5	5	0	0	0	1	0	0	3	0	0	0	0	0	1	0	0	2	5
6	0	0	1	0	2	M	0	0	0	2	2	1	0	0	0	0	0	2	2	0	0	0	0	0	0	0	2	2
7	0	1	0	1	1	1	M	0	0	0	0	0	0	0	1	0	0	1	0	1	0	1	0	0	0	0	0	0
8	0	0	0	0	0	0	0	M	0	0	5	0	0	3	0	0	0	1	0	0	0	0	0	0	0	0	0	0
9	1	1	1	1	1	1	1	1	M	2	1	1	1	1	1	1	1	3	1	1	1	1	1	1	1	5	1	1
10	0	1	0	0	5	1	0	0	2	M	5	1	0	0	0	2	0	1	0	0	0	0	0	0	0	4	1	4
11	0	0	0	0	5	0	0	5	0	3	M	0	0	0	0	0	0	1	5	5	0	1	0	0	0	0	1	1
12	0	1	0	0	3	1	0	0	0	3	0	M	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
13	0	0	0	5	0	0	0	0	0	1	0	0	M	0	0	1	0	0	0	5	0	0	0	0	0	1	1	0
14	0	1	0	1	0	0	0	3	0	0	0	0	0	M	0	0	0	3	0	0	0	1	1	0	0	0	0	0
15	0	0	0	0	0	0	1	0	0	0	1	0	0	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0
16																												
17																												
18	0	5	1	0	2	2	0		1	0	3	0	0	2	3	0	0	M	0	0	0	0	0	0	0	0	0	3
19			0	1	1	2	1	1	1	0	5	0	0	0	0	0		0	M	0	1	1	1	1	1	0	0	0
20	5	0	0	5	1	0	0	0	0	0	2	0	1	0	0	0	0	0	0	M	0	0	1	0	0	0	2	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	M	0	0	0	0	0	0
22	0	1	0	0	0	0	2	0	0	0	2	0	0	1	2	0	0	0	0	0	0	M	1	0	1	0	2	2
23	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	M	0	0	0	4	0
24	0	0	0	0	0	2	0	0	0	0	0	0	2	0	0	0	0	0	2	0	0	0	0	M	4	0	0	0
25	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0	0	0	0	2	M	0	0	1
26	0	0	0	0	4	0	0	0	5	5	0	0	1	0	0	2	0	0	0	0	1	0	0	0	0	M	2	0
27	1	0	0	5	1	1	0	1	0	1	1	0	1	1	0	0	0	0	0	1	0	1	4	0	0	1	M	0
28	0	0	0	0	5	0	0	0	1	5	1	0	1	2	3	0	0	1	1	1	0	2	0	0	3	1	0	M

Table 6.3: Social Closeness Matrix for GNEG 1121-H004

To begin the analysis of relationships within each section, the magnitude of associated with social closeness was disregarded, and a sociogram built based simply on a 0/1 relationship, and the resulting sociograms [6] are shown in Figures 6.4 and 6.5. Next, the sociogram was updated to include values associated with social closeness as shown in Figure 6.6 and 6.7.

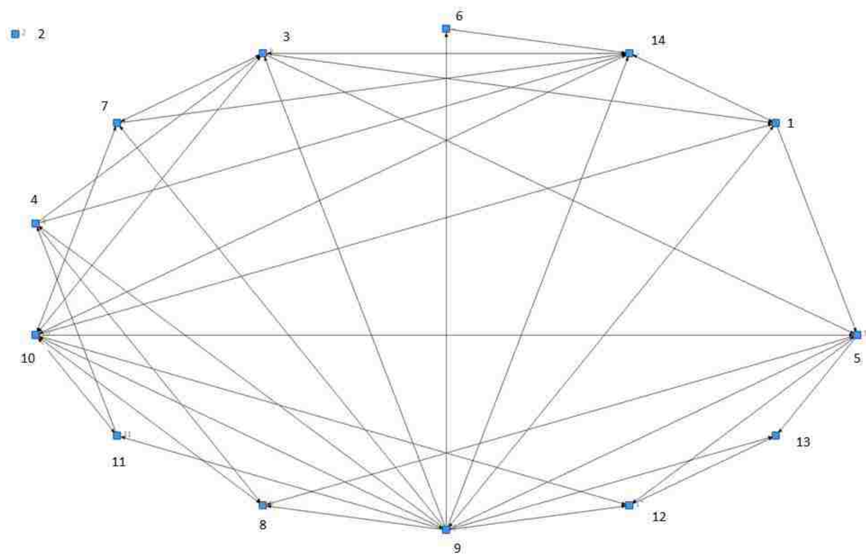


Figure 6.4: Section 006 Sociogram

6.4.2 Determining the Initiators of Influence

Traditional social network analysis can be used to evaluate the relationships obtained from the initial survey, and centrality metrics can be used to identify and rank important actors. The out-degree, in-degree, and betweenness centrality for actors in each section are shown in Tables 6.4 and 6.5. The top three or four ranked students for each performance metric are identified in bold type. Tables 6.6 and 6.7 include the influential actors in the network as determined by both centrality metrics and their peers. The number in the last column indicates the number of students identifying a particular actor as influential.

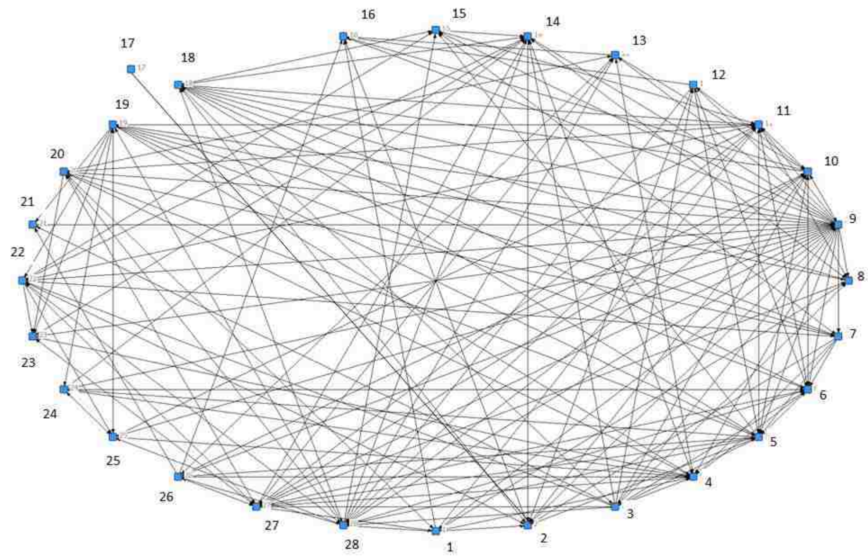


Figure 6.5: Section H004 Sociogram

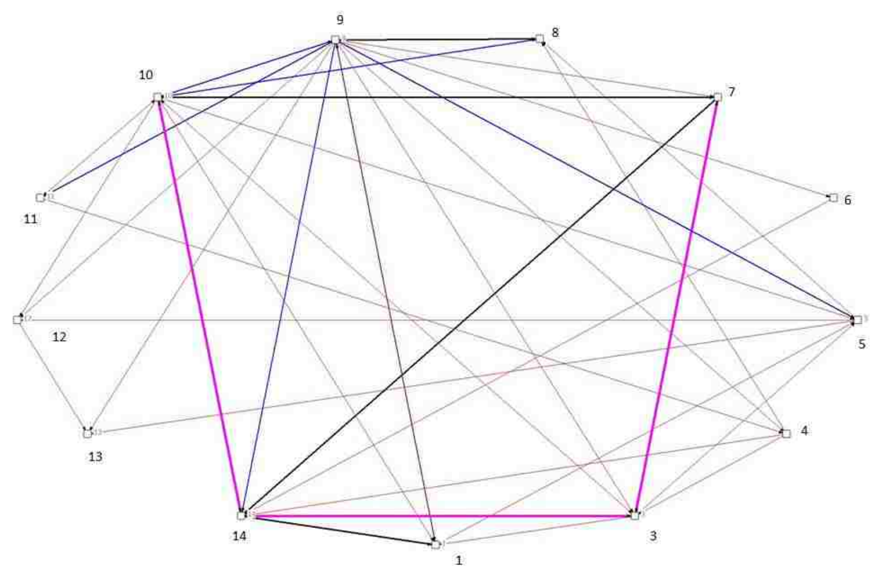


Figure 6.6: Section 006 Sociogram with Social Closeness

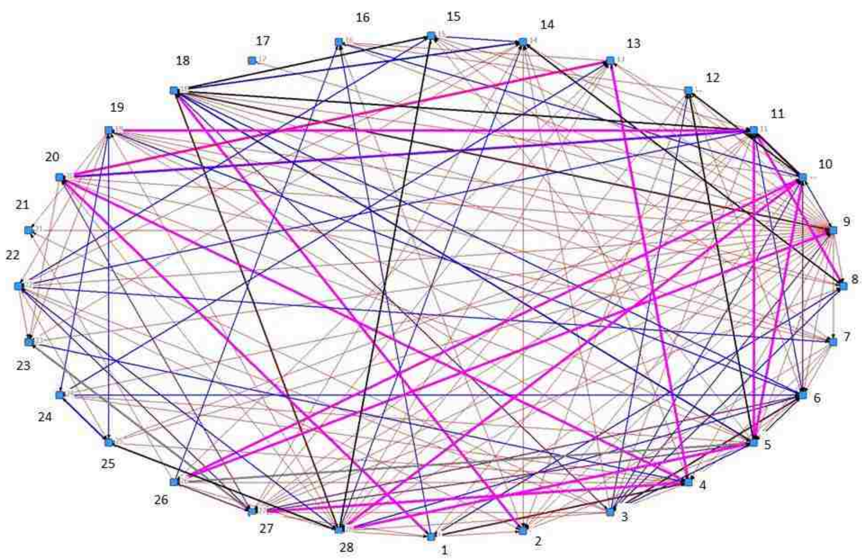


Figure 6.7: Section H004 Sociogram with Social Closeness

Actor	Out-Degree	In-Degree	Betweenness
1	0.2308	0.3077	0.1720
3	**0.3846**	**0.3846**	0.2727
4	0.2308	0.1538	0.0442
5	0.2308	0.3077	0.2064
6	0.0769	0.0769	0
7	0.2308	0.2308	0
8	0	0.3077	0
9	**0.9231**	0.1538	**0.5037**
10	**0.6923**	**0.3846**	**1.000**
11	0.1538	0.1538	0.0590
12	0.0769	0.2308	0.0369
13	0	0.2308	0
14	0.2308	**0.5385**	**0.4472**

Table 6.4: Centrality Metrics for GNEG 1121-006

Actor	Out-Degree	In-Degree	Betweenness
1	0.2593	0.1481	0.1065
2	0.1852	0.2963	0.2468
3	**0.4815**	0.1111	0.1199
4	0.3704	0.3704	0.7495
5	0.2963	**0.5185**	0.4694
6	0.3333	0.4074	0.6107
7	0.2963	0.1481	0.1306
8	0.1111	0.2593	0.0397
9	**0.9630**	0.2222	**0.9242**
10	0.4074	0.4074	0.6432
11	0.3333	**0.5185**	0.7429
12	0.2222	0.1852	0.049
13	0.2222	0.2222	0.0783
14	0.2222	0.3704	0.3848
15	0.1481	0.2222	0.0767
16	0	0.2593	0
17	0	0	0
18	0.3333	**0.4444**	0.7786
19	0.4444	0.2222	0.5567
20	0.2593	0.2963	0.1735
21	0	0.1481	0
22	0.3333	0.2963	0.3355
23	0.0741	0.2593	0.0134
24	0.1852	0.2222	0.0693
25	0.1481	0.2222	0.0936
26	0.2593	0.2222	0.2628
27	**0.4815**	**0.4444**	**1.000**
28	**0.4815**	0.4074	**0.8937**

Table 6.5: Centrality Metrics for GNEG 1121-H004

Actor	Out-Degree	In-Degree	Betweenness	Selected by Peers
3	x	x		3
7				1
9	x		x	
10	x	x	x	1
14		x	x	4

Table 6.6: Influential Students in GNEG 1121-006

Actor	Out-Degree	In-Degree	Betweenness	Selected by Peers
2				1
3	x			
4				1
5		x		2
8				1
9	x		x	
10				2
11		x		3
14				1
18		x		1
20				3
22				1
26				1
27	x	x	x	1
28	x		x	2

Table 6.7: Influential Students in GNEG 1121-H004

The students selected most often as "influential" by their peers were included in the follow-up survey aimed at eliciting influence probabilities. Therefore, actors 3 and 14 were included for section 006, and actors 11 and 20 were included for section H004. In addition to identifying classmates they found influential, students were asked to identify teaching assistants (other than their primary teaching assistant) within the program they found to be influential. In section 006, students identified Coby, Michael, and Lora while the students in section H004 identified only Coby and Michael.

6.4.3 Determining Influence Probabilities

Students were presented with various scenarios involving both an initiator of influence and a method of communication (see Figures 6.2 and 6.3). Students were asked to distribute a total of 10 points across the influence levels to indicate the likelihood they would perform a specified task. To elicit integer responses, the students were asked questions such as "If the instructor sent an announcement via e-mail asking you to do something on 10 different occasions, how many times would you do exactly as instructed (highly influential)? How many times would you completely ignore the request (not influential)? How many times would you do something in between (somewhat or barely influential)?" This process was repeated for the various initiators of influence and communication methods. The composite results are shown in Tables 6.8 and 6.9.

It is interesting to note that the probabilities associated with the instructor influence levels elicited from the students are very similar to the probability distribution used in earlier chapters (see Table 6.10). Students indicated that overall, on average, they would do what their instructor asked exactly as instructed approximately 60% of the time. Another interesting thing to note is the difference in perceived influence of the primary TA between the two sections. Students in section 006 indicated that overall, on average, they would do what their primary TA asked exactly as instructed approximately 62.5% of the time, while students in section H004 responded with 39.3%. Finally, the perceived influence of the peer mentors is noteworthy. In fact, in section H004 the students indicated their peer mentors were more influential than the primary TA.

6.4.4 Investigation of Elicited Influence Probabilities

A series of four scenarios were used to investigate the influence probabilities elicited in section 6.4.3. For the first scenario, student attendance is evaluated. Attendance in Intro in mandatory and recorded as part of the final grade. The next two scenarios involved the completion of two surveys (one bonus, one required), and the final scenario involved an e-mail request from the primary TA for students to retrieve a set of papers from their designated mailfolder.

Attendance

Since attendance is an integral component of Intro, the scale in Table 6.11 was used to match attendance grades with influence levels. Note that to receive attendance points for a given day, students must arrive to class on time and "click-in" using a Response Card. The results of this scenario are somewhat surprising. In general, one would assume students

Instructor Influence				
Method	3	2	1	0
Announcement	6.00	2.45	1.30	0.40
E-Mail	4.91	2.73	1.73	0.64
Both	6.55	2.36	1.00	0.09
Direct	8.09	1.73	0.18	0.00
Overall	6.39	2.32	1.05	0.28
Primary TA Influence				
Method	3	2	1	0
Announcement	5.45	2.55	1.45	0.55
E-Mail	5.36	1.91	1.82	0.91
Both	6.27	1.82	1.27	0.64
Direct	7.91	1.73	0.27	0.09
Overall	6.25	2.00	1.20	0.55
Peer Mentor Influence				
Method	3	2	1	0
E-Mail	4.64	2.91	1.27	1.18
Direct Contact	6.09	2.27	0.64	1.00
Overall	5.36	2.59	0.95	1.09
Other FEP TA Influence				
Method	3	2	1	0
Coby E-Mail	3.36	2.09	1.82	2.73
Coby Direct	4.50	2.40	1.10	2.00
Michael E-Mail	2.91	2.18	1.91	3.00
Michael Direct	4.50	2.50	1.00	2.00
Lora E-Mail	3.27	2.45	1.91	2.36
Lora Direct	4.80	2.60	1.20	1.40
Overall	3.89	2.37	1.49	2.25
Other Student Influence				
Method	3	2	1	0
Student 3 E-Mail	1.00	2.09	3.00	3.91
Student 3 Direct	1.30	3.80	2.00	2.90
Student 14 E-Mail	2.60	2.30	2.20	2.90
Student 14 Direct	3.89	3.67	1.11	1.33
Overall	2.20	2.96	2.08	2.76

Table 6.8: Elicited Influence Levels for GNEG 1121-006

Instructor Influence				
Method	3	2	1	0
Announcement	4.48	3.26	1.56	0.70
E-Mail	4.15	3.52	1.67	0.67
Both	6.15	2.42	0.96	0.46
Direct	6.75	1.92	0.92	0.42
Overall	5.38	2.78	1.28	0.56
Primary TA Influence				
Method	3	2	1	0
Announcement	3.28	3.28	2.36	1.08
E-Mail	2.92	3.23	2.77	1.08
Both	4.36	3.08	1.72	0.84
Direct	5.17	2.67	1.38	0.79
Overall	3.93	3.06	2.06	0.95
Peer Mentor Influence				
Method	3	2	1	0
E-Mail	4.58	2.69	1.42	1.31
Direct Contact	5.50	2.38	1.29	0.83
Overall	5.04	2.53	1.36	1.07
Other FEP TA Influence				
Method	3	2	1	0
Coby Email	2.35	3.31	2.81	1.54
Coby Direct	3.31	3.15	2.42	1.12
Michael Email	2.50	3.23	2.77	1.50
Michael Direct	3.36	3.00	2.28	1.36
Overall	2.88	3.17	2.57	1.38
Other Student Influence				
Method	3	2	1	0
Student 11 Email	1.58	1.73	2.81	3.88
Student 11 Direct	2.08	2.04	3.08	2.80
Student 18 Email	1.16	2.04	3.48	3.32
Student 18 Direct	1.84	2.64	3.28	2.24
Overall	1.66	2.11	3.16	3.06

Table 6.9: Elicited Influence Levels for GNEG 1121-H004

Scenario	**3**	**2**	**1**	**0**
Chapter 4 & 5	0.650	0.200	0.100	0.050
Section 006	0.639	0.232	0.105	0.028
Section H004	0.538	0.278	0.128	0.056

Table 6.10: Comparison of Instructor Influence Levels

in an Honors section of a class would have a better overall attendance rate than students in a 'regular' section. However, the attendance grades for Sec 006 more closely match the instructor influence levels from Table 6.10 than those for Sec H004.

Attendance Grade	"Influence Level"	Sec 006	Sec H004
$\geq$ 90%	3	61.5%	33.3%
$\geq$ 80%	2	23.1%	33.3%
$\geq$ 70%	1	0%	18.5%
$<$ 70%	0	15.4%	14.8%

Table 6.11: Attendance Grades Matched to Influence Levels

Survey Response

Two surveys were administered during the class. Both surveys were announced in class and a reminder sent via e-mail. In Sec 006, 8/13 students (61.5%) completed the Decision Day survey and 9/13 (69.2%) completed the Peer Mentor Survey while in Sec H004, 20/27 students (74.1%) completed the Decision Day survey and 19/27 (70.4%) completed the Peer Mentor Survey.

Retrieving Items

Each student in Intro was assigned a mail folder in the FEP study lounge where his or her assignments were returned. The primary TA sent an e-mail to the students requesting they retrieve all their items from their mail folder by a given time. In Sec 006, 7/13 students (53.8%) removed all their items, 2/13 students (15.4%) removed none of their items, and 4/13 (30.1%) removed a subset of the items from their mail folder. In Sec H004, 16/27 students (59.3%) removed all their items, 6/27 students (22.2%) removed none of their items, and 5/27 students (18.5%) removed a subset of the items from their mail folder.

There are several interesting things to note from the results of this survey. The first is with the inconsistencies in social closeness. There were several instances in which relationships were not reported as equally reciprocal. That is, a student may report that he or she shares a residence hall with a particular actor while that actor indicates that they do not know the other person. Although it was not the focus of this study, it may be interesting to explore the implications of these inconsistencies with respect to influence or information flow within the network.

The second noteworthy observation is with the identification of influential students. Students that responded to the survey indicating that, basically, they knew everyone in the class had the highest rankings with respect to out-degree centrality. However, these students did not necessarily have a high incidence of being identified as influential by their peers. In section 006, actors 3, 9, and 10 ranked highest by out-degree centrality, but only actor 3 was chosen for inclusion for the set of influential students (actors 3 and 14) for that class. Similarly, in section H004, actors 3, 9, 27, and 28 ranked highest in out-degree centrality, but none were chosen for inclusion in the set of influential students (actors 11 and 18) for the class. However, the in-degree centrality metric captures the extent to which an actor is known by others in the network and more closely aligns with the actors included in the set of influential students. In section 006, actors 3, 10, and 14 ranked highest by in-degree centrality while in section H004, actors 5, 11, 18, and 28 received the highest rankings.

The next observation to note is with the elicited influence levels for the instructor (see Table 6.10). These align almost perfectly with the conditional probability distribution under perfect influence (see Table 5.1). This result did not happen by design. The development of the probability distributions for the illustrative examples in Chapters 4 and 5 was completely detached from the elicited influence levels in this study. It will be interesting to see if these values hold across different student groups in future research.

Finally, the student response to the activities used to investigate the elicited probabilities is interesting. Although attendance was a required component of the course and assigned a grade, the honors students attended class at a much lower rate than the regular students but responded to requests for surveys at a higher rate. There are many factors that may have contributed to this observation. First may be the "points" associated with the various activities. Another may have been the perceived value of the activities. However, when the students were asked to retrieve their items from their mail folders (an activity which did not have a point value associated with them), a significant portion of the students did exactly what was requested of them.

6.5. Conclusions and Future Work

In this chapter, an initial study aimed at identifying social relationships, influential actors, and influence level probabilities inside a required course was completed. An initial set of experiments aimed at investigated the elicited influence probabilities was performed. However, more work needs to be done in this area of quantifying student perceptions of influence. More extensive experimentation involving all of the combinations of influential actors and communication methods needs to be performed, and student perception on the "importance" of tasks within the course should also be considered as that could have a tremendous impact on the outcome of the experiments.

In future work, it would be worthwhile to model the relationships as students progress through their degree. In this work, students were in their second semester of their freshman year. I would be interesting to follow the students to see how and when their relationships form and/or change. Another avenue for future work it to continue to explore the idea of an "influential" student. Anecdotally, the students identified by their peers as the most "influential" did not necessarily have the highest GPA or make the best grades in the class.

Chapter 7

Appendix

7.1. Chapter 3 Sensitivity Analysis

	Level of Influence			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 1 \| x_i = b_{ij})$			
1	0.1426	0.8681	0.8680	0.8678
2	0.7918	0.7957	0.7937	0.7946
3	0.7948	0.7962	0.7959	0.7961
4	0.7636	0.9675	0.9670	0.9663
6	0.1166	0.9153	0.9144	0.9156
7	0.7789	0.7980	0.8008	0.7975
	$\Pr(\phi(\mathbf{x}) \geq 1) = 0.7952$			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 2 \| x_i = b_{ij})$			
1	0.0876	0.1336	0.7728	0.7728
2	0.6317	0.6401	0.6425	0.6404
3	0.6368	0.6404	0.6428	0.6424
4	0.5989	0.7455	0.9304	0.9306
6	0.0703	0.1144	0.8250	0.8234
7	0.6198	0.6349	0.6448	0.6472
	$\Pr(\phi(\mathbf{x}) \geq 2) = 0.6379$			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 3 \| x_i = b_{ij})$			
1	0.0376	0.0765	0.1141	0.5400
2	0.2362	0.2497	0.2531	0.2536
3	0.2400	0.2519	0.2555	0.2548
4	0.1773	0.5511	0.6857	0.8201
6	0.0236	0.0681	0.1044	0.4307
7	0.2148	0.2460	0.2558	0.2615
	$\Pr(\phi(\mathbf{x}) \geq 3) = 0.2518$			

Table 7.1: System Reliability under Various Levels of Demand with Actor 4 at Low Influence

	$d = 1$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	2	0.2903	3	1.0915	2	0.2736	2	1.3766
2	5	0.0020	6	1.0002	5	0.0023	5	1.0023
3	6	0.0010	5	1.0011	6	0.0002	6	1.0002
4	3	0.1823	1	1.2160	3	0.0132	3	1.0134
6	1	0.3461	2	1.1508	1	0.2845	1	1.3975
7	4	0.0090	4	1.0045	4	0.0068	4	1.0069
	$d = 2$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	2	0.4415	3	1.1410	2	0.5511	2	2.2276
2	5	0.0052	6	1.0049	5	0.0032	5	1.0033
3	6	0.0043	5	1.0062	6	0.0006	6	1.0006
4	3	0.2439	1	1.3620	3	0.0204	3	1.0208
6	1	0.4897	2	1.1947	1	0.5702	1	2.3264
7	4	0.0124	4	1.0085	4	0.0110	4	1.0111
	$d = 3$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	2	0.2718	2	1.3815	2	0.6979	2	3.3103
2	5	0.0069	6	1.0041	5	0.0234	5	1.0240
3	6	0.0062	5	1.0090	6	0.0156	6	1.0159
4	1	0.4587	1	2.7229	3	0.0986	3	1.1094
6	3	0.2461	3	1.2368	1	0.7404	1	3.8521
7	4	0.0188	4	1.0181	4	0.0567	4	1.0601

Table 7.2: General CIM Results with Actor 4 at Low Influence

	$d = 1$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	2	0.1307	2	1.0823	3	0.0821
2	5	0.0012	6	1.0001	5	0.0015
3	6	0.0008	5	1.0009	6	0.0001
4	3	0.0526	3	1.0324	2	0.0338
6	1	0.2038	1	1.1283	1	0.1280
7	4	0.0058	4	1.0032	4	0.0041
	$d = 2$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	2	0.2134	2	1.1692	2	0.1653
2	5	0.0037	6	1.0048	5	0.0010
3	6	0.0037	5	1.0055	6	0.0003
4	3	0.0678	3	1.0543	3	0.0520
6	1	0.2770	1	1.2187	1	0.2155
7	4	0.0103	4	1.0102	4	0.0059
	$d = 3$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	1	0.2018	1	1.4006	1	0.4008
2	6	0.0030	6	1.0049	5	0.0070
3	5	0.0039	5	1.0084	6	0.0070
4	3	0.1284	3	1.2584	3	0.2515
6	2	0.1789	2	1.3552	2	0.3552
7	4	0.0138	4	1.0244	4	0.0305

Table 7.3: Alternative CIM Results with Actor 4 at Low Influence

	Level of Influence			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 1 \mid x_i = b_{ij})$			
1	0.9081	0.9617	0.9619	0.9629
2	0.9312	0.9588	0.9598	0.9600
3	0.9390	0.9595	0.9589	0.9597
4	0.7641	0.9667	0.9672	0.9672
6	0.7488	0.9932	0.9936	0.9935
7	0.8444	0.9854	0.9841	0.9852
	$\Pr(\phi(\mathbf{x}) \geq 1) = 0.9577$			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 2 \mid x_i = b_{ij})$			
1	0.8035	0.8720	0.9204	0.9214
2	0.8313	0.8899	0.9157	0.9151
3	0.8648	0.8976	0.9155	0.9155
4	0.6006	0.7445	0.9295	0.9306
6	0.6440	0.7429	0.9793	0.9800
7	0.7338	0.8253	0.9551	0.9553
	$\Pr(\phi(\mathbf{x}) \geq 2) = 0.9054$			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 3 \mid x_i = b_{ij})$			
1	0.6214	0.7240	0.7761	0.8116
2	0.6303	0.7347	0.7778	0.8013
3	0.6573	0.7602	0.7872	0.8018
4	0.1762	0.5544	0.6868	0.8202
6	0.3995	0.6333	0.7273	0.9279
7	0.4633	0.6900	0.7758	0.8838
	$\Pr(\phi(\mathbf{x}) \geq 3) = 0.7695$			

Table 7.4: System Reliability under Various Levels of Demand with Actor 4 at High Influence

	$d = 1$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	5	0.0210	4	1.0047	4	0.0173	4	1.0176
2	5	0.0107	5	1.0019	5	0.0092	5	1.0093
3	6	0.0079	6	1.0017	6	0.0065	6	1.0066
4	2	0.0739	3	1.0097	2	0.0674	2	1.0723
6	1	0.1054	1	1.0373	1	0.0727	1	1.0784
7	3	0.0650	2	1.0284	3	0.0394	3	1.0411
	$d = 2$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	4	0.0554	4	1.0114	4	0.0498	4	1.0524
2	5	0.0365	6	1.0074	5	0.0330	5	1.0341
3	6	0.0229	5	1.0074	6	0.0178	6	1.0181
4	2	0.1717	3	1.0182	1	0.1715	1	1.2069
6	1	0.1908	1	1.0547	2	0.1561	2	1.1849
7	3	0.1171	2	1.0367	3	0.0927	3	1.1021
	$d = 3$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	4	0.0808	4	1.0211	4	0.0839	4	1.0915
2	5	0.0714	6	1.0174	5	0.0754	5	1.0815
3	6	0.0572	5	1.0217	6	0.0526	6	1.0556
4	1	0.3139	3	1.0220	1	0.3860	1	1.6287
6	2	0.2356	1	1.0686	2	0.2376	2	1.3116
7	3	0.1688	2	1.0522	3	0.1671	3	1.2006

Table 7.5: General CIM Results with Actor 4 at High Influence

	$d = 1$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	4	0.0091	4	1.0043	4	0.0052
2	5	0.0045	5	1.0020	5	0.0028
3	6	0.0043	6	1.0016	6	0.0029
4	3	0.0187	3	1.0094	3	0.0101
6	1	0.0618	1	1.0318	1	0.0327
7	2	0.0445	2	1.0228	2	0.0237
	$d = 2$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	4	0.0259	4	1.0136	4	0.0149
2	5	0.0170	5	1.0088	5	0.0099
3	6	0.0142	6	1.0073	6	0.0084
4	3	0.0459	3	1.0250	3	0.0257
6	1	0.1112	1	1.0616	1	0.0613
7	2	0.0757	2	1.0413	2	0.0423
	$d = 3$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	4	0.0371	5	1.0230	4	0.0252
2	6	0.0334	6	1.0208	6	0.0226
3	5	0.0368	4	1.0235	5	0.0243
4	3	0.0877	3	1.0560	3	0.0579
6	1	0.1589	1	1.1029	1	0.1035
7	2	0.1293	2	1.0833	2	0.0847

Table 7.6: Alternative CIM Results with Actor 4 at High Influence

	Level of Influence			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 1 \mid x_i = b_{ij})$			
1	0.6522	0.6878	0.6869	0.6854
2	0.6087	0.6896	0.6898	0.6912
3	0.6346	0.6901	0.6934	0.6923
4	0.1354	0.8189	0.8191	0.8202
6	0.6281	0.9790	0.9790	0.9798
7	0.1467	0.8149	0.8170	0.8178
	$\Pr(\phi(\mathbf{x}) \geq 1) = 0.6843$			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 2 \mid x_i = b_{ij})$			
1	0.5539	0.5626	0.5889	0.5846
2	0.4268	0.5329	0.6037	0.6079
3	0.5037	0.5566	0.6067	0.6045
4	0.0808	0.1241	0.7463	0.7421
6	0.5334	0.6238	0.9506	0.9502
7	0.0946	0.1327	0.7375	0.7397
	$\Pr(\phi(\mathbf{x}) \geq 2) = 0.5773$			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 3 \mid x_i = b_{ij})$			
1	0.2781	0.3073	0.3154	0.3254
2	0.1945	0.2605	0.3175	0.3560
3	0.1866	0.2981	0.3338	0.3557
4	0.0175	0.0701	0.1104	0.5203
6	0.2560	0.5167	0.5976	0.8214
7	0.0422	0.0784	0.1116	0.5107
	$\Pr(\phi(\mathbf{x}) \geq 3) = 0.3137$			

Table 7.7: System Reliability under Various Levels of Demand with Actor 6 at Low Influence

	$d = 1$							
	SAD		*MRAW*		*MFV*		*MRRW*	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	6	0.0131	6	1.0035	6	0.0156	6	1.0159
2	4	0.0311	5	1.0086	3	0.0368	3	1.0382
3	5	0.0242	4	1.0112	5	0.0242	5	1.0248
4	1	0.3181	2	1.1974	1	0.2674	1	1.3650
6	2	0.3137	1	1.4310	4	0.0274	4	1.0281
7	3	0.3115	3	1.1933	2	0.2619	2	1.3548
	$d = 2$							
	SAD		*MRAW*		*MFV*		*MRRW*	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	6	0.0190	6	1.0109	6	0.0220	6	1.0225
2	4	0.0840	4	1.0329	3	0.1125	3	1.1268
3	5	0.0503	5	1.0327	4	0.0544	4	1.0576
4	1	0.4278	2	1.1927	1	0.5484	1	2.2141
6	3	0.2789	1	1.4577	5	0.0253	5	1.0260
7	2	0.4166	3	1.1863	2	0.5354	2	2.1525
	$d = 3$							
	SAD		*MRAW*		*MFV*		*MRRW*	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	6	0.0185	6	1.0142	6	0.0446	6	1.0467
2	4	0.0728	5	1.0490	3	0.1832	3	1.2243
3	5	0.0683	4	1.0660	4	0.1516	4	1.1787
4	2	0.3166	2	1.2195	1	0.7896	1	4.7530
6	1	0.3508	1	2.0568	5	0.0613	5	1.0653
7	3	0.3020	3	1.2093	2	0.7533	2	4.0530

Table 7.8: General CIM Results with Actor 6 at Low Influence

	$d = 1$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	6	0.0051	6	1.0028	6	0.0047
2	5	0.0131	5	1.0080	4	0.0110
3	4	0.0140	4	1.0096	5	0.0109
4	1	0.2182	3	1.1585	1	0.1604
6	3	0.0920	1	1.0647	3	0.0698
7	2	0.2140	2	1.1556	2	0.1571
	$d = 2$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	6	0.0116	6	1.0135	6	0.0066
2	4	0.0423	4	1.0395	4	0.0338
3	5	0.0333	5	1.0314	5	0.0263
4	1	0.2464	1	1.2156	1	0.2113
6	3	0.0770	3	1.0687	3	0.0646
7	2	0.2401	2	1.2102	2	0.2057
	$d = 3$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	6	0.0091	6	1.0155	6	0.0134
2	5	0.0357	5	1.0588	5	0.0550
3	4	0.0451	4	1.0731	4	0.0707
4	1	0.2257	1	1.3622	1	0.3573
6	3	0.0988	3	1.1585	3	0.1563
7	2	0.2148	2	1.3454	2	0.3394

Table 7.9: Alternative CIM Results with Actor 6 at Low Influence

	Level of Influence			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 1 \mid x_i = b_{ij})$			
1	0.7810	0.9813	0.9811	0.9817
2	0.9481	0.9619	0.9634	0.9616
3	0.9524	0.9634	0.9642	0.9631
4	0.8561	0.9879	0.9877	0.9882
6	0.6294	0.9795	0.9779	0.9781
7	0.9296	0.9685	0.9686	0.9707
	$\Pr(\phi(\mathbf{x}) \geq 1) = 0.9600$			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 2 \mid x_i = b_{ij})$			
1	0.6990	0.7673	0.9577	0.9581
2	0.8711	0.9037	0.9197	0.9188
3	0.8902	0.9091	0.9191	0.9177
4	0.7186	0.8431	0.9692	0.9687
6	0.5338	0.6256	0.9495	0.9505
7	0.8553	0.8943	0.9291	0.9286
	$\Pr(\phi(\mathbf{x}) \geq 2) = 0.9144$			
	0	1	2	3
Actor	$\Pr(\phi(\mathbf{x}) \geq 3 \mid x_i = b_{ij})$			
1	0.4475	0.6720	0.7418	0.9226
2	0.6681	0.7532	0.7761	0.7874
3	0.6944	0.7642	0.7802	0.7884
4	0.2971	0.7045	0.8216	0.9236
6	0.2579	0.5150	0.5985	0.8229
7	0.7001	0.7352	0.7706	0.7950
	$\Pr(\phi(\mathbf{x}) \geq 3) = 0.7691$			

Table 7.10: System Reliability under Various Levels of Demand with Actor 6 at High Influence

	$d = 1$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	2	0.0810	2	1.0223	2	0.0622	2	1.0663
2	5	0.0063	6	1.0024	5	0.0041	5	1.0041
3	6	0.0061	5	1.0037	6	0.0026	6	1.0026
4	3	0.0626	1	1.0291	3	0.0361	3	1.0374
6	1	0.1287	3	1.0193	1	0.1148	1	1.1297
7	4	0.0194	4	1.0097	4	0.0106	4	1.0107
	$d = 2$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	2	0.1498	2	1.0317	2	0.1321	2	1.1523
2	5	0.0212	5	1.0035	5	0.0197	5	1.0201
3	6	0.0125	6	1.0029	6	0.0108	6	1.0109
4	3	0.1254	1	1.0398	3	0.0974	3	1.1079
6	1	0.2469	3	1.0260	1	0.2440	1	1.3228
7	4	0.0360	4	1.0105	4	0.0289	4	1.0297
	$d = 3$							
	SAD		$MRAW$		MFV		$MRRW$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean	Rank	Mean
1	3	0.1998	2	1.0665	3	0.1933	3	1.2396
2	4	0.0474	5	1.0110	4	0.0507	4	1.0534
3	6	0.0367	4	1.0132	6	0.0345	6	1.0357
4	2	0.2479	1	1.0897	2	0.2326	2	1.3030
6	1	0.3299	3	1.0233	1	0.4056	1	1.3824
7	5	0.0434	6	1.0119	5	0.0446	5	1.0467

Table 7.11: General CIM Results with Actor 6 at High Influence

	$d = 1$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	2	0.0371	2	1.0200	2	0.0186
2	6	0.0034	6	1.0023	5	0.0012
3	5	0.0041	5	1.0030	6	0.0012
4	1	0.0432	1	1.0234	1	0.0216
6	3	0.0338	3	1.0180	3	0.0172
7	4	0.0141	4	1.0084	4	0.0063
	$d = 2$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	2	0.0710	2	1.0380	2	0.0396
2	5	0.0093	5	1.0042	5	0.0059
3	6	0.0071	6	1.0027	6	0.0051
4	1	0.0836	1	1.0446	1	0.0467
6	3	0.0659	3	1.0355	3	0.0366
7	4	0.0236	4	1.0118	4	0.0140
	$d = 3$					
	MAD		$MMAW$		$MMFV$	
Actor	Rank	Mean	Rank	Mean	Rank	Mean
1	2	0.1079	2	1.0699	2	0.0704
2	6	0.0218	6	1.0132	6	0.0152
3	5	0.0231	5	1.0142	5	0.0158
4	1	0.1931	1	1.1241	1	0.1269
6	3	0.0925	3	1.0595	3	0.0608
7	4	0.0300	4	1.0189	4	0.0201

Table 7.12: Alternative CIM Results with Actor 6 at High Influence

7.2. Chapter 6 IRB Protocol, Script, and Approval

RSSP Project Number ______________________

INSTITUTIONAL REVIEW BOARD
UNIVERSITY OF ARKANSAS
PROTOCOL FORM

"The University Institutional Review Board is responsible for recommending policies and monitoring their implementation on the use of human beings as subjects for physical, mental, and social experimentation in and out of class... Protocols for the use of human subjects in research, and in class experiments, whether funded internally or externally, must be approved by the IRB prior to the implementation of the human subject protocol.... Violation of procedures and approved protocols can result in the loss of funding by the sponsoring agency or the University of Arkansas and may be interpreted as 'scientific misconduct.'" (Faculty Handbook, p. 3 – 25)

Supply the information requested in items 1-14 as appropriate. Type entries in the spaces provided using additional pages as needed. In accordance with the college/department policy, submit the original and one (1) copy of this completed protocol form and all attached materials to the appropriate Human Subjects Committee. In absence of an IRB-authorized Human Subjects Committee, submit the original and one (1) copy of this completed protocol form and all attached materials to the IRB Program Manager, Carol Rodlun, E214B ANSC.

1. Title of Project: Eliciting Influence Probabilities for a Social Network
2. (Students must have a faculty member supervise the research. The faculty member must sign this form and all researchers and the faculty advisor should provide a campus phone number.

	Name	Department	Campus Address	Campus Phone
Principal Researcher	Kellie Schneider	ENGD	ENGR 324	5-4271
Co-Researcher	Ed Pohl	INEG	BELL 4207	5-6042
Co-Researcher	Chase Rainwater	INEG	BELL 4207	5-2687

3. Researcher(s) Status. Check all that apply.

   X Faculty   Staff   X Graduate Student(s)   Undergraduate Student(s)

4. Project Type

   Faculty Research   X Thesis/dissertation   Class Project   Independent Study   Honors Project

5. Is the project receiving extramural funding?

   X No   Yes. Specify the source of funds.

6. Brief description of the purpose of the proposed research and all procedures involving people. Use additional pages if needed. (**Do not** send thesis or dissertation proposals. Proposals for extramural funding must be submitted in full.)

   Purpose of research: The purpose of this research is to quantify levels of influence and the probabilities associated with being in each influence level for members within a social network.

   Procedures involving people: Students in one of my GNEG 1121/1121H classes for Spring 2012 will be asked to complete an initial survey in class as well as a follow-up web based survey. The survey results will be used to create a social network. Students will be asked to complete a series of tasks throughout the semester. The tasks will be initiated by members of the social network identified as influential. The completion rate of these tasks will be used to and gauge actual vs perceived influence levels.

Figure 7.1: IRB Protocol Page 1

Survey Consent Form

Dear Students,

As part of my dissertation research, I will be creating a social network of this class. Participation in this study is completely voluntary and will not affect your grade in GNEG 1121 – Introduction to Engineering II. Students that agree to participate in the study will be asked to complete an initial survey, become part of the social network along with the other participants from the class, and asked to complete a final survey. Students that initially agree to participate in the study may stop participation and withdraw from the study at any time. To ensure the confidentiality of data collected, each student will be assigned a code number, and all responses and activities associated with the study will be recorded with the code number. At the end of the study, the code listing will be destroyed rendering all responses anonymous.

If you have any questions about this study, please let me know, or you may contact one of my advisors, Dr. Edward A Pohl (epohl@uark.edu) or Dr. Chase Rainwater (cer@uark.edu). You may also contact Ro Windwalker, Compliance Coordinator of Research Support and Sponsored Programs by phone at 479-575-3845 or by e-mail at irb@uark.edu, if you have any questions regarding survey administration.

Sincerely,

Kellie Schneider
Instructor
Freshman Engineering

Please indicate below whether or not you would like to participate in this study.

I would like to participate in this study.

____
Printed Name

____
Signature

I would not like to participate in this study.

____
Printed Name

____
Signature

Figure 7.2: Consent Form for Survey Participants

7. Estimated number of participants (complete all that apply)

Children under 14 | Children 14-17 | 60 UA Students (18 yrs and older) | Adult non-students

8. Anticipated dates for contact with participants:

First Contact: January 23, 2012 | Last Contact: May 5, 2012

9. Informed Consent Procedures: The following information must be included in any procedure: identification of researcher, institutional affiliation and contract information; identification of Compliance Officer and contact information; purpose of the research, expected duration of the subject's participation; description of the procedures; risks and/or benefits; how confidentiality will be ensured; that participation is voluntary and that refusal to participate will involve no penalty or loss of benefits to which the subject is otherwise entitled. See *Policies and Procedures Governing Research with Human Subjects*, section 5.0 Requirements for Consent.

Signed informed consent will be obtained. Attach copy of form.
Modified informed consent will be obtained. Attach copy of form.
X Other Method (e.g., implied consent). Please explain on attached sheet.
Not applicable to this project. Please explain on attached sheet

10. Confidentiality of Data: All data collected that can be associated with a subject/respondent must remain confidential. Describe the methods to be used to ensure the confidentiality of data obtained.

All responses by individual will remain confidential. Once the survey and data collection is closed, students will be assigned a random identification number. The responses will be analyzed and reported in the dissertation using the random identification number.

11. Risks and/or Benefits:

Risks: Will participants in the research e exposed to more than minimal risk? Yes X No
Minimal risk is defined as risks of harm not greater, considering probability and magnitude, than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests. Describe any such risks or discomforts associated with the study and precautions that will be taken to minimize them.

Benefits: Other than the contribution of new knowledge, describe the benefits of this research.
This research will benefit educators' ability to disseminate important information in an effective manner.

12. Check all of the following that apply to the proposed research. Supply the requested information below or on attached sheets:

A. Deception of or withholding information from participants. Justify the use of deception or the withholding of information. Describe the debriefing procedure: how and when will the subject be informed of the deception and/or the information withheld?
B. Medical clearance necessary prior to participation. Describe the procedures and note the safety precautions to be taken.
C. Samples (blood, tissue, etc.) from participants. Describe the procedures and note the safety precautions to be taken.
D. Administration of substances (food, drugs, etc.) to participants. Describe the procedures and note the safety precautions to be taken.
E. Physical exercise or conditioning for subjects. Describe the procedures and note the safety precautions to be taken.

Figure 7.3: IRB Protocol Page 2

F. Research involving children. How ill informed consent from parents or legally authorized representatives as well as from subjects be obtained?

G. Research involving pregnant women or fetuses. How ill informed consent to obtained from both parents of the fetus?

H. Research involving participants in institutions (cognitive impairments, prisoners, etc.) Specify agencies or institutions involve. Attach letter of approval. Letters must be on letterhead with original signature; electronic transmission is acceptable.

I. Research approved by an IRB at another institution. Specify agencies or institutions involved. Attach letters of approval. Letters must be on letterhead with original signature: electronic transmission is acceptable.

J. Research that must be approved by another institution or agency. Specify agencies or institutions involved. Attach letters of approval. Letters must be on letterhead with original signature; electronic transmission is acceptable.

13. Checklist for Attachments

The following are attached:
Consent form (if applicable) or
X Letter to participants, written instructions, and/or script of oral protocols indicating clearly the information in item #9.
Letter(s) of approval from cooperating institution(s) and/or other IRB approvals (if applicable)
X Data Collection Instruments

14. Signatures

I/we agree to provide proper surveillance of this project to insure that the rights and welfare of the human subjects/respondents are protected. I/we will report any adverse reactions to the committee. Additions to or changes in research procedures after the project has been approved will be submitted to the committee for review. I/we agree to request renewal of approval for any project when subject/respondent contact continues more than one year.

Principal Researcher Kellie Schneider Date: 11/11/11

Co- Researcher [illegible] Date: 11/11/11

Co- Researcher [illegible] Date: 11/11/11

Figure 7.4: IRB Protocol Page 3

UNIVERSITY OF ARKANSAS

Office of Research Compliance
Institutional Review Board

November 30, 2011

MEMORANDUM

TO:	Kellie Schneider Ed Pohl Chase Rainwater
FROM:	Ro Windwalker IRB Coordinator
RE:	New Protocol Approval
IRB Protocol #:	11-11-290
Protocol Title:	*Eliciting Influence Probabilities for a Social Network*
Review Type:	☒ EXEMPT ☐ EXPEDITED ☐ FULL IRB
Approved Project Period:	Start Date: 11/30/2011 Expiration Date: 11/29/2012

Your protocol has been approved by the IRB. Protocols are approved for a maximum period of one year. If you wish to continue the project past the approved project period (see above), you must submit a request, using the form *Continuing Review for IRB Approved Projects*, prior to the expiration date. This form is available from the IRB Coordinator or on the Research Compliance website (http://vpred.uark.edu/210.php). As a courtesy, you will be sent a reminder two months in advance of that date. However, failure to receive a reminder does not negate your obligation to make the request in sufficient time for review and approval. Federal regulations prohibit retroactive approval of continuation. Failure to receive approval to continue the project prior to the expiration date will result in Termination of the protocol approval. The IRB Coordinator can give you guidance on submission times.

This protocol has been approved for 60 participants. If you wish to make *any* modifications in the approved protocol, including enrolling more than this number, you must seek approval *prior to* implementing those changes. All modifications should be requested in writing (email is acceptable) and must provide sufficient detail to assess the impact of the change.

If you have questions or need any assistance from the IRB, please contact me at 210 Administration Building, 5-2208, or irb@uark.edu.

210 Administration Building • 1 University of Arkansas • Fayetteville, AR 72701
Voice (479) 575-2208 • Fax (479) 575-3846 • Email irb@uark.edu

The University of Arkansas is an equal opportunity/affirmative action institution.

Figure 7.5: IRB Approval Letter

References

[1] R. Ahuja, T. Magnangti, and J. Orlin, *Network flows: Theory, algorithms, and applications*, PH, Upper Saddle River, 1993.

[2] T. Aven, *Reliability evaluation of multistate systems with multistate components*, IEEE Transactions on Reliability **34** (1985), no. 5, 473–479.

[3] R.E. Barlow and A.S. We, *Coherent systems with multi-state components*, Mathematical Operations Research **3** (1978), no. 4, 275–281.

[4] R. Boedigheimer and K. Kapur, *Customer driven reliability models for multistate coherent systems*, IEEE Transactions on Reliability **43** (1994), no. 1, 46–50.

[5] S.P. Borgatti, *Identifying sets of key players in a social network*, Computational and Mathematical Organization Theory **12**, 21–34.

[6] S.P. Borgatti, M.G. Everett, and L.C. Freeman, *Ucinet for windows: Software for social network analysis*, Harvard, MA: Analytic Technologies.

[7] T. Cao, Z. Wu, S. Wang, and X. Hu, *Maximizing influence spread in modular social networks by optimal resource allocation*, Expert Systems with Applications **38**, no. 10, 13128–13135.

[8] P.J. Carrington, J. Scott, and S. wasserman, *Models and methods in social networks*, CAMB, New York, 2005.

[9] C.R. Clark, *Modeling and analysis of clandestine networks*, Master's thesis, Air Force Institute of Technology, 2005.

[10] J.L. Cook and J.E Ramirez-Marquez, *A reliability analysis of cluster-based ad-hoc networks*, Reliability Engineering & System Safety **93** (2008), no. 10, 1512–1522.

[11] A.J. Daly (ed.), *Social network theory and educational change*, Harvard Education Press, Cambridge, 2010.

[12] P.K. Davis and K. Cragin, *Social science for counterterrorism: Putting the pieces together*, RAND Corporation, Santa Monica, CA, 2009.

[13] A. Degenne and M. Forse, *Introducing social networks*, SAGE, Boston, 1999.

[14] C.R. Ebeling, *An introduction to reliability and maintainability engineering*, McGraw-Hill, Boston, 1997.

[15] E. El-Neweihi, F. Proschan, and J. Sethuraman, *Multistate coherent systems*, Journal of Applied Probability **15** (1978), 675–668.

[16] E. Evan-Dar and A. Shapira, *A note on maximizing the spread of influence in social networks*, Information Processing Letters **111** (2011), no. 4, 184–187.

[17] G.S. Fishman, *The distribution of maximum flow with applications to multistate reliability systems*, Operations Research **35** (1987), no. 4, 607–618.

[18] L.C. Freeman, *Centrality in social networks: Conceptual clarification*, Social Networks **1** (1979), 215–239.

[19] J.L. Geffre, R.F. Deckro, and S.A. Knighton, *Determining critical members of layered operational terrorist networks*, Journal of Defense Modeling and Simulation **6** (2009), no. 2, 97–109.

[20] W. Griffith, *Multistate reliability models*, Journal of Applied Probability **17** (1980), 735–744.

[21] J. T. Hamill, R.F. Deckro, V.D. Wiley, and R.S. Deckro, *Gains, losses, and thresholds of influence in social networks*, International Journal of Operational Research **2** (2007), no. 4, 357–379.

[22] J.T. Hamill, *Analysis of layered social networks*, Ph.D. thesis, Air Force Institute of Technology, 2006.

[23] R.A. Hanneman and R. Riddle, *Introduction to social network methods*, University of California, Riverside, Riverside, CA, 2005.

[24] J. Hudson and K. Kapur, *Reliability analysis for multistate systems with multistate components*, IIE Transactions **2** (1983), no. 4, 127–134.

[25] ———, *Reliabliity bounds for multistate systems with multistate components*, Operations Research **33** (1985), no. 1, 153–160.

[26] C.C. Jane and Y.W. Laih, *A practical algorithm for computing multi-state two-terminal reliability*, IIE Transactions on Reliability **57** (2008), no. 2, 295–302.

[27] ———, *Computing multi-state two-terminal reliability through critical arc states that interrupt demand*, IEEE Transactions on Reliability **59** (2010), no. 2, 338–345.

[28] ———, *A dynamic bounding algorithm for approximating multi-state two-terminal reliability*, European Journal of Operational Research **205** (2010), 625–637.

[29] J.S. Lin, C.C. Jane, and J. Yuan, *On reliability evaluation of a capacitated-flow network in terms of minimal patsets*, Networks **25** (1995), 131–138.

[30] G.H. McCormick, *Terrorist decision making*, Annual Review of Political Science **6** (2003), 473–507.

[31] B. Natvig, *Two suggestions of how to define a multistate coherent system*, Advances in Applied Probability **14** (1982), 434–455.

[32] B. Natvig, S. Kormo, A.T. Holen, and G.Hogasen, *Multistate reliability theory - a case study*, Advances in Applied Probability **18** (1986), 921–932.

[33] Y. Ni and Z.Q. Liu, *Heuristic search for optimizing diffusion on finfluence in a social network under the resource constraint*, Soft Computing **15** (2011), no. 2, 335–344.

[34] Y. Ni, L. Xie, and Z.Q. Liu, *Minimizing the expected complete influence time of a social network*, Information Sciences **180** (2010), no. 13, 2514–2527.

[35] J.E. Ramirez-Marques and D.W. Coit, *Composite importance measures for multi-state systems with multi-state components*, IEEE Transactions on Reliability **54** (2005), no. 3, 517–529.

[36] J.E. Ramirez-Marquez and D.W. Coit, *A monte-carlo simulation approach for approximating multi-state two-terminal reliability*, Reliability Engineering & System Safety **87** (2004), 253–264.

[37] J.E. Ramirez-Marquez, D.W. Coit, and M. Tortorella, *A generalized multistate-based path vector approach to multistate two-terminal reliability*, IIE Transactions **38** (2006), no. 6, 477–488.

[38] J.E. Ramirez-Marquez, I. Hernandez, K. Schneider, C. Rainwater, and E.A. Pohl, *Reliability model for influencing individuals in the social network setting*, Proceedings of the PSAM 11 & ESREL 12 Conference.

[39] J.E. Ramirez-Marquez and C. Rocco, *All-terminal network reliability optimization via probabilistic solution discovery*, Reliability Engineering & System Safety **93** (2008), no. 11, 1689–1697.

[40] M. Rausand and A. Hyland, *System reliability theory: Models and statistical methods*, John Wiley & Sons, Hokoken, NJ, 2004.

[41] Reliability and Maintainability Symposium (RAMS), 2011 Proceedings, *Assessing multi-layered social networks using reliability models*, 2011.

[42] R. S. Renfro and R.F. Deckro, *A flow model social network analysis of the iranian government*, Military Operations Research **8** (2003), no. 1, 5–16.

[43] S. Satitsatian and K.C. Kapur, *An algorithm for lower reliability bounds of multistate two-terminal networks*, IEEE Transactions on Reliability **55** (2006), no. 2, 199–206.

[44] K.S. Schneider, J.E. Ramirez-Marquez, I. Hernandez, C. Rainwater, and E.A. Pohl, *Social network analysis via multi-state reliability and conditional influence models*, Reliability Engineering & System Safety **109** (2013), 99–109.

[45] S. Skiena, *The algorithm design manual*, Springer-Verlag, New York, 1998.

[46] R. Kevin Wood, *Deterministic network interdiction*, Mathematical and computer modeling **17** (1993), no. 2, 1–18.

[47] S Yang and D. Knoke, *Optimal connections: strength and distance in valued graphs*, Social Networks **23** (2001), no. 4, 285–295.

[48] M.J. Zuo, Z. Tian, and H.Z. Huang, *An efficient method for reliability evaluation of multistate networks given all minimal path vectors*, IIE Transactions **39** (2007), 811–817.