UNITED STATES MILITARY ACADEMY

WEST POINT, NEW YORK

HONORS THESIS

DYNAMIC ANALYSIS OF NODAL DEPENDENCIES TO DETERMINE INTRINSIC AND TIME-DEPENDENT NODAL VALUES WITHIN A SOCIAL NETWORK

by

Kegan Reilly

May 2008

Thesis Advisor:	MAJ Anthony N. Johnson
Second Reader:	MAJ Ian McCulloh
Second Reader:	LTC Donald A. Outing

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.

1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** May 2008	**3. REPORT TYPE AND DATES COVERED**

4. TITLE AND SUBTITLE Dynamic Analysis of Nodal Dependencies

5. FUNDING NUMBERS

6. AUTHOR(S) Kegan Reilly

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
United States Military Academy
West Point, NY 10996

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
NA

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)
Communications within networks are interdependent connections to specific events at a set time. These connections are based on previous interrelations established between the different entities or nodes of an organization. Social network analysis (SNA) has become an important tool for analyzing a wide variety of networks, and can be used to determine which agents in a particular network are the most influential within a closed group. The influence, however, that an individual or group has on a specific event in time is in constant flux. For example, the coach of a basketball team might have the most influence within the team early in the game, but when the game is in its final seconds, the player chosen to execute the game winning three-point play becomes the most influential individual for that specific event. We propose a way to show that through dynamic event analysis, we can uncover indicators which point to the player executing that shot as being the influential node with the team network!

Analysis of current nodal or agent valuation methods, such as centrality, hubs and connectors, capability, prestige, closeness, and betweeness will be explored to find isomorphic ties to the Black-Scholes financial model used to stochastically depict the randomness of stock-option values. Data collected from a group of 24 officers in a one year graduate program at Columbia University will serve as the basis for a model which will be used to answer the question: who is the most influential individual or group with in a social network given prior knowledge of a specific event occurring in time? The mathematical concepts involved include stochastic differential equations, graph theory, and network science.

14. SUBJECT TERMS	**15. NUMBER OF PAGES** 30
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT OF**
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

DYNAMIC ANALYSIS OF NODAL DEPENDENCIES TO DETERMINE INTRINSIC AND TIME-DEPENDENT NODAL VALUES WITHIN A SOCIAL NETWORK

Kegan Reilly
CDT LT, Class 08
B.S., United States Military Academy, 2008

Submitted in partial fulfillment of the
requirements for the degree of
BACHELOR OF SCIENCE
in **MATHEMATICAL SCIENCES**
with Honors
from the
UNITED STATES MILITARY ACADEMY
May 2008

Author: Kegan Reilly

Advisory Team: MAJ Anthony N. Johnson
Thesis Advisor

MAJ Ian McCulloh
Second Reader

LTC Donald A. Outing
Second Reader

COL Michael Phillips
Chairman, Department of Mathematical Sciences

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Communications within networks are interdependent connections to specific events at a set time. These connections are based on previous interrelations established between the different entities or nodes of an organization. Social network analysis (SNA) has become an important tool for analyzing a wide variety of networks, and can be used to determine which agents in a particular network are the most influential within a closed group. The influence, however, that an individual or group has on a specific event in time is in constant flux. For example, the coach of a basketball team might have the most influence within the team early in the game, but when the game is in its final seconds, the player chosen to execute the game winning three-point play becomes the most influential individual for that specific event. We propose a way to show that through dynamic event analysis, we can uncover indicators which point to the player executing that shot as being the influential node with the team network!

Analysis of current nodal or agent valuation methods, such as centrality, hubs and connectors, capability, prestige, closeness, and betweeness will be explored to find isomorphic ties to the Black-Scholes financial model used to stochastically depict the randomness of stock-option values. Data collected from a group of 24 officers in a one year graduate program at Columbia University will serve as the basis for a model which will be used to answer the question: who is the most influential individual or group with in a social network given prior knowledge of a specific event occurring in time? The mathematical concepts involved include stochastic differential equations, graph theory, and network science.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank Dr. Kathleen Carley and the Center for Computational Analysis of Social and Organizational Systems (CASOS) at Carnegie Mellon University for their support on this project. Their efforts and custom modifications to the Organizational Risk Analyzer (ORA) software were essential to the successful completion of this research.

Further, I would like to thank my advisor MAJ Anthony Johnson as well as the following for helping me through the research process either by providing me with research information, giving helpful advice, or by being a reader:

Dr. Brian Winkel
MAJ Ian McCulloh
LTC Donald Outing
Dr. Amy Erickson
COL Michael Phillips
Dr. Elisha Peterson

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

E-mail has significantly changed how people communicate and interact [1]. In many ways communication is easier and more reliable with e-mail, however, there many new communication challenges introduced. In order to integrate the use of this new communication tool into Army protocol, leaders need to be aware of the strengths and weaknesses of e-mail. As a result, Army leaders will be able to use this resource to their advantage to distribute information to subordinates.

E-mail provides several advantages over classic communication methods. Peers that were once unreachable due to distance, conflicting time schedules, or other duties, can now be reached with ease using e-mail. If the intended recipient is not available, the message is stored in an in-box until it can be read. Leaders are able to issue guidance and direction efficiently to a much lower level or rank than before. While it was once infeasible for a Brigade Commander to assemble the brigade and issue guidance, except in rare circumstances, he/she can now e-mail everyone in the organization. Those Soldiers, who wouldn't have been present due to some other duty, are able to read the guidance or directive later when they have time. Soldiers are able to have greater access to headquarters elements such as personnel, finance, and supply, through the use of e-mail and Web based services [2]. All of these advantages can greatly improve the efficiency of military organizations. However, there are new challenges that must be understood before the Army can efficiently modify doctrine to exploit the advantages e-mail offers.

With each advantage e-mail brings, there are challenges that must be understood as well. The increase in peer to peer collaboration allows problems to be solved at a much lower level or rank than before [3]. Unfortunately, military commanders may now be unaware of issues and problems in their organization. While e-mail may be good for subordinates to solve problems, it does not allow commanders to provide guidance or take advantage of their experience [4]. It is great for commanders to reach all of their subordinates at once; however, if that is the case, what role do intermediate commanders now play? How does a commander handle feedback, questions, and recommendations from a large volume of subordinates?

Social Network Analysis (SNA) offers a valuable methodology to understand e-mail communication. A social network is a collection of individuals and their relationships

[5]. In SNA, we analyze groups of individuals and their relationships mathematically to gain insight about their behavior [6]. Some examples of networks that can be studied include those of friendship, respect, migration, biological relation, or, in this project, e-mail they received. Before e-mail networks are described in detail, a background in social networks is presented.

We describe a social network in terms of its members and their connections [7]. Members of the network can be connected through these relations as groups or individuals. In the following figures, the shaded circles (or nodes) represent members of a social network. The lines (or edges) between these nodes represent their relations.

The connections between members can be either directional or non-directional relationships [3]. An example of a directional social network is very similar to a military chain of command (see Figure 1). This type of social network has a directed flow of information. The commander gives orders or directions to his subordinate units, so the arrows point down to the lowest level. This diagram is also representative of a connected graph where everyone is in the network. Every node in the diagram receives information from a superior.

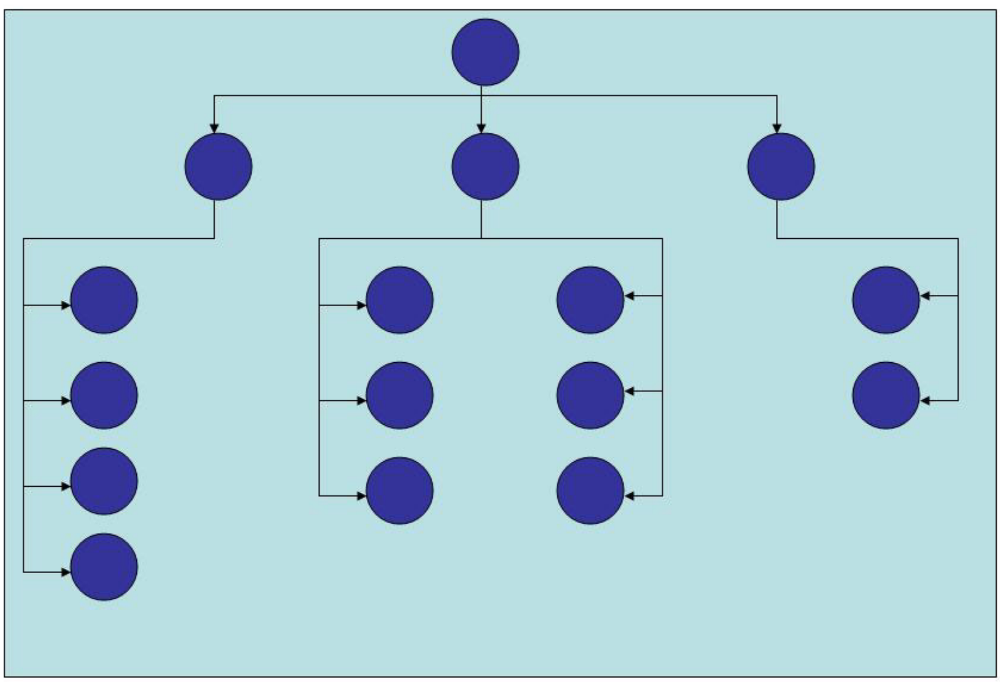


Figure 1. Chain of Command Network

The family and friendship network shown in Figure 2 is, by contrast, an example of a non-directional, non-connected graph. In this figure, we can see that there is no predictable structure to the social network because there are no restrictions when talking to family members. For example, the youngest child may talk to anyone in the family where as in the military a private cannot simply talk to a general. This graph is disconnected because a particular family member may not talk to anyone, such as a disgruntled teenager, yet he/she is still part of the family.

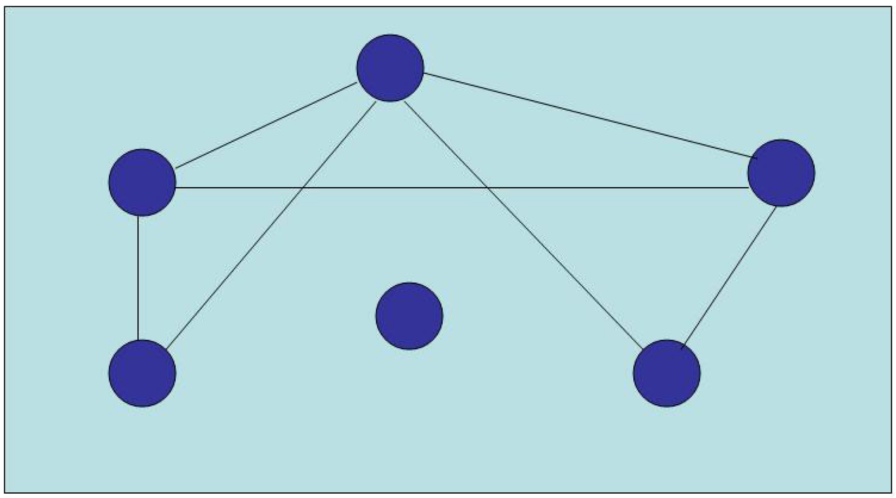


Figure 2. Family and Friendship Social Network

These are only two examples of the many social networks that exist. Applications of social network analysis are beneficial not only in the military, but also in other organizations. For example, a mayor may want to conduct an analysis of a local neighborhood to improve cohesiveness, or a CEO may want to conduct some analysis to improve a department in his company.

The present research focuses on the application of SNA techniques to understand a group of US Army officers. This effort focuses on the use of e-mail communication patterns as the basis for these analysis. This background is in three parts: first, we will discuss the implementation of the data collection effort; second, we will explore the characteristics of the collected networks; and third, we will explore the relationship between e-mail networks and networks generated through self-report questionnaires.

This first section focuses on the practical implementation of the data collection and the organizational and technical challenges that must be overcome for such data collection to occur. It concludes with suggestions for efficient, ethical data collection, and serves not only as the basis for the remainder of this report, but also for the continued collection of data for the remainder of the Ike NEt project, discussed in the next section.

B. PARTICIPANTS

This project monitors the e-mail traffic of 24 mid-grade Army officers, selected to serve as company tactical officers at the United States Military Academy, along with cadets, faculty, and researchers. Before serving as tactical officers, ELDP students go through a one-year Master's Degree in Leadership program at Columbia University. Students in the program take some classes on the Columbia main campus and some classes in the Department of Behavioral Sciences and Leadership at the U.S. Military Academy. The program is called the Eisenhower Leadership Development Program (ELDP), thus our project is called IkeNet.

The overall demographics consisted of the mid-grade officers, cadets, faculty, and researchers. Participants were treated in accordance with ethical standards established by the American Psychological Association. The research methods used in this experiment were approved by the U.S. Military Academy Human Subjects Use Committee.

The subjects willing to participate in this investigation permitted us to place a patch onto their Microsoft Outlook e-mail accounts. This patch allowed us to collect information from e-mails in their sent items folder, found on participants' personal computers. The information included: the time each e-mail we sent, the FROM/TO/BC/CC e-mail addresses/names, and the subject line. We were not able to see the content of any e-mail itself. The patch was triggered by two conditions; 1,000 minutes of elapsed time since the previous trigger and sending an e-mail. Upon triggering, the patch would search the subject's sent mail folder and compile information into a comma-separated values (CSV) spreadsheet from all sent e-mails since the last search. The patch then initiated an e-mail to send the CSV spreadsheet to the principal investigator. The principle investigator would then compile the spreadsheets from all 24 subjects into one master spreadsheet and ensure anonymity of the names. The anonymous names are coded as P1 through P24.

In addition to the e-mail data, the ELDP coordinator provided other sources of important data. The planning calendar for the entire year in which the data were collected was provided. This calendar included graded events and important dates that would potentially have an impact in the e-mail communication of the subjects, such as Army football games, Thanksgiving, and Christmas. Army home football games are big events at USMA in which the entire corps of cadets and tactical officers are required to attend. The ELDP coordinator also administered several surveys to the subjects to compare e-mail communication with self-reported social network data. Finally, several subjects were interviewed to gain greater insight into the activities of the subjects and to identify workload demands on the subjects.

C. COMPUTER PATCH

The first research task in the project was to create a method for collecting e-mail data for SNA. A Visual Basic for Applications (VBA) program was written that could be installed on a personal computer (PC) in the session window of Microsoft Outlook. It is installed in Outlook by opening the program, selecting Alt-F11, and copying the code into the session window. Once the code is installed in this manner, that PC will send regular CSV spreadsheets of e-mail communication to the principal investigator defined in the code.

There were several advantages to the VBA approach to the problem. Microsoft Outlook is proprietary software that causes difficulty in allowing a program such as Organizational Risk Analyzer (ORA) to directly pull information from a subject's outbox. The VBA patch is an easier software implementation. Another advantage is that the patch is installed on each individual subject's computer; therefore, the people who run the e-mail exchange server do not have to be involved in the research project. It often can be difficult, in research, to be dependent upon another group of people for your data; they do not necessarily care about your research. Researchers at the Center for Computational Analysis of Social and Organizational Systems (CASOS) have written software that allows the same e-mail information to be pulled directly from the exchange server as an extension of this project. The final advantage offered by the VBA patch is the control the subjects feels in the project. The subject actually installs the patch himself or herself. The subject then feels they have control over the information they send. Most of the subjects knew how to remove the patch when their participation in the project ended. Several subjects said they felt more comfortable knowing the software sending the principal investigator information was on their computer, and ”Big Brother” wasn't pulling their information from somewhere else.

D. AUTOMATION OF DATA COLLECTION

Data compiling and collection needed to be automated. There were 24 subjects sending daily e-mail to the principal investigator. This meant saving 24 files from e-mail to a hard drive every day. The principal investigator would have to wait at least a week before compiling the data because there would be an occasional subject who would not log on to the e-mail for a day or two. This would cause missing data, until the subject eventually did log in and sent a bigger data file. Data were typically compiled once per month throughout the project. Once the files were saved on a hard drive, each file would be opened and the contents copied and pasted into a master file. This process merged all of the data from all 24 subjects for a one-month period into a single data file. With 24 subjects and 25 average files per month, this meant processing 600 files, which would have taken the principal investigator several hours to complete. Instead, significant time-saving software was developed by Freytag Industries LLC to deal with this complication.

SNA was performed on the data using the proprietary software ORA owned by the CASOS and free for government use. The CASOS created several custom modifications to their software to facilitate this research. The software now has a feature that loads e-mail social network data directly into ORA using the format of the CSV file from the VBA Outlook patch. Once data files were compiled, they were easily loaded into ORA for analysis. Investigators on this project sorted data for different time and date periods and looked at certain key subjects in the group using Microsoft Excel and were then able to easily load these files into ORA. With only a couple of hours of familiarization with ORA, the novice investigator could conduct analysis on social network data thanks to these custom modifications.

This research effort was highly successful at collecting a unique sensitive data set for 31 weeks, running from 15 October 2006 through 11 May 2007. Collecting the data in the format of a CSV spreadsheet allowed for the manipulation of the data in ORA to conduct the necessary research analysis. The data was broken down into 20 time blocks within the 31 weeks of collection in order to analyze the different nodal values assigned to the subjects as the network changed through time. By creating visualizations and data files through ORA, the IkeNet data was able to be analyzed to allow the formulation of a stochastic model.

II. METHODOLOGY

A. SOCIAL NETWORK ANALYSIS

Social Network Analysis (SNA) is a quantitative approach to measuring the connectivity of individuals and groups based on a specific metric or related event [8]. SNA can serve as an important tool to linking the connections of individuals and groups to predestined events giving insight as to who was the main cause of the event. There are three strategies used to quantitatively assign values to specific individuals and groups in a social network. They are nodal, dyadic, and triadic strategy [9].

The nodal strategy breaks down the social network into single nodes that are linked together based on a shared connection. The measurement of nodal value within the nodal strategy is based on visibility [9]. A nodes visibility is the number of links it shares with other nodes within the network relating to the properties of the individual or the situation he or she encounters [10]. The nodal strategy focuses on the occupant of the node in a social network.

The dyadic strategy breaks down the network into nodal pairs. The nodal value of a nodal pair in the dyadic strategy is based on three strategies, network proximity, spatial representation, and structural indices [9]. The proximities are derived through many techniques such as, social circles, chain measures, and row and column correlations of a relational matrix [9]. A spatial representation of a network can be in the form of a sociogram, which are scattered points connected by lines. Structural indices are used to measure the cohesiveness and equivalence of a networks structure. The dyadic strategy focuses on the pairs relationship within the network.

The triadic strategy inventories three node groupings or triads in a network [11]. The triadic strategy is primarily used for theoretical testing of social relations [9]. The focus of the triadic strategy is on the relationships between the three nodes of a triad within a network [12]. This strategy analyzes the network as a whole for sufficient connections among the different triad groupings.

Among these different typologies of networks, there are many ways to generate nodal values. Nodal value generation will be further researched in order to efficiently provide a means to assist the generation of a stochastic differential equation that can be used to expose who was the most influential individual or group that caused a specific event to occur. Mapping out interactions among a large number of agents within a communications network to generate nodal values is not a trivial task. The practical difficulty being that the network is transient and changes at every time step leaving little to no information as to what it was or what it will be. As more tools become available to record the interactions within a communications network (i.e. email traffic and phone records), time dependent, dynamic analysis of this data will become vital to capturing key characteristics and properties of the network. Most notably, dynamic analysis can uncover information about nodes and edges that are not exposed in a static model. Recently, the United States Military Academy collected email header, and cellular phone traffic from cadets, staff, and faculty under a project called the Eisenhower Leadership Network (IkeNet), to attempt to create a smaller scale social network that can be analyzed. Currently, they have found that even with a smaller scale network, it is still computationally intensive to model the network as an entity that changes with time.

B. CENTRALITY, PRESTIGE, AND CAPABILITY

One important application of graph theory in social network analysis is in identifying the most important or prominent actors within a network at a given time. There are two derived methods of assigning prominence to actors within a social network. They are centrality and prestige. Centrality is made up of the sub-categories degree, closeness, and betweenness that approximate the involvement of actors within a graph [13]. Prestige is made up of the sub-categories degree and proximity that are methods of approximating the amount of ties sent and received by the actors within a graph [13].

Degree Centrality stems from the concept that the most prominent actors are those that have the most ties to other actors in the network, which is termed activity [13]. Activity of an actor within a directional graph is the measure of the outgoing ties of a node also known as the outdegree [13]. Degree Centrality of a node in a directional graph is the proportion of the nodes in the graph that are connected to it [13]. The proportion is the outdegree of a node over the group size or total entities in the network, which is annotated in the equation below.

$$C'_D(n_i) = d(n_i)/g - 1$$

According to Degree Centrality, the node with the highest proportion is the most prominent within the graph.

Closeness Centrality is based on how close, according to distance, an actor is from all the other actors within the social network [13]. According to closeness centrality, a centrally located actor is one who can interact with the others the fastest. The fastest actor, in theory, is the actor that has the closes distance between other actors [14]. In a graph, the shortest paths linking the nodal entities measure the minimum distance. By equating closeness to minimum distance, closeness centrality is inversely proportional to distance [13]. Thus, the farther a node is relative to other nodes in a network, means that the centrality of that node will be smaller then the nodes whose distance is shorter. For a directional graph, the relation between nodes n_j and n_i is different then the relation between n_i and n_j. The equation for calculating closeness centrality is achieved by dividing the minimum possible total distance by the sum of the i^{th} row of the distance matrix to obtain the total distance for n_i is from other nodes [13].

$$C'_C(n_i) = (g-1)/[\sum_{j=1}^{g} d(n_i, n_j)]$$

One drawback with closeness centrality is that within the digraph used to represent the network every node has to have a connection, or else the equation will be undefined.

Betweeness Centrality is based on the influence that the "actors in the middle" have on two nonadjacent actors within a network [13]. The shortest path between two nonadjacent nodes is said to be influenced by the nodes in between the path that connects the nonadjacent nodes [15]. According to betweeness centrality, the actors that are more central are those that lie between the most paths of other actors, because they have more control over connections of other actors [16]. To calculate the betweeness centrality of an actor the probability of communication or having a path needs to be examined. It is assumed that the there is the same likelihood for an actor to choose one path over another [13]. Let g_{jk} be the number of paths linking two actors j and k and let $g_{jk}(n_i)$ be the number of paths linking the two actors that contain actor i [13]. The actor betweeness index for n_i is the sum of all the probabilities that actor i is between the pair of actors j and k [13].

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i)/g_{jk}$$

The equation for the betweeness centrality for a directional graph of an actor is the betweeness index divided by the maximum number of pairs not including n_i within the graph [13].

$$C'_B(n_i) = C_B(n_i)/[(g-1)(g-2)]$$

The advantage to betweeness centrality is that the graph does not have to be completely connecting in order for the equation to be defined.

Degree Prestige is the measure of the indegree or amount of ties received by an actor, which means that actors that are more prestigious tend to receive more ties or nominations then others [13]. The equation used to calculate the degree prestige of an actor n_i is the number of arcs terminating at n_i divided by the group size of the graph [13].

$$P'_D(n_i) = x_i/g - 1$$

This produces the proportion of actors within a network who choose to interact or share a connection with actor i.

Proximity prestige considers the distance that an actor n_i is to other actors within the network [13]. An actor who is closer to all other actors within a network is considered more prestigious according to proximity prestige. The proximity prestige of an actor calculated according the ratio of the proportion of actors who can reach i to the average distance these actors are from i [13].

$$P_P(n_i) = (I_i/(g-1))/\sum d(n_j, n_i)/I_i$$

According to this equation, as actors who can reach the actor i become closer, then the ratio becomes larger, so the closer an actor's proximity ratio is to 1, the more prestigious that actor is within the network [13].

Nodal Capability assigns value to nodes within a network based on how capable or available a node is to their associated network[17]. The capability measurement ensures that every node with a network receives a value greater then zero. The purpose of the capability measurement is to show that every node within a network is able to be reached at some point in time. Even if a node does not share any ties with the network at a certain time, at a time in the future that node still is capable of being reached. The following is the equation for capability [17]:

$$C_i(T) = 1/(1 + e^{(-10(x^*-1/2))})$$

where

x^* =Total number of messages sent or received for node$_i$ at time T / Maximum possible nodal links at time T

The capability equation was developed by ORA and it is assigned to each node$_i$ at time T throughout the time interval based on a normalization of the messages sent/received for the node as compared to the maximum possible messages.

C. VALUE DIRECTED GRAPHS

Social Network Analysis quantifies the relationship among various groups and individuals within the network. One method that is commonly used to best represent the strength and intensity of each linked relation within a social network is a valued directed graph. A value directed graph is a graph that consists of valued lines or arcs that connect each node based on a directional valued relation. A directional valued relation means that the value associated with the connection from node n to node m is different from the value associated with the connection from node m to node n. A real-life example of this type of relationship would be a superior-subordinated relationship, where there would be a higher valued connection between the superior and the subordinate then the subordinate to the superior. In a directed value graph, the nodes are assigned a nodal degree that is equal to the number of arcs incident with the node and the values associated with the arcs [13]. The arcs that connect two or more nodes within the directed value graph are called valued paths. The value assigned to each path is equal to the smallest value arc within the path [18].

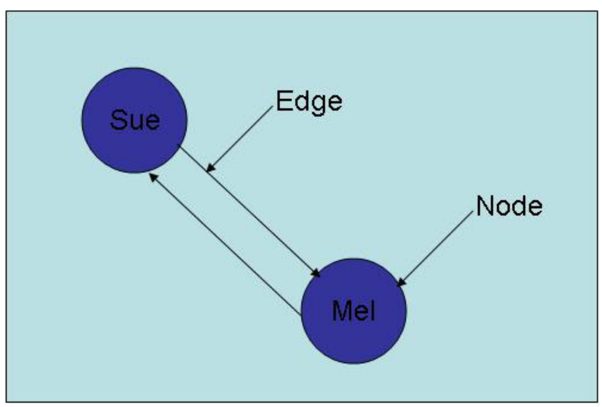


Figure 3. Directed Network Graph

The path value is considered the most restricted amount of communication or interaction within the path. Nodes that share higher valued arcs within a directed graph are considered more reachable, because they possess a stronger interaction [19]. Within a value directed graph, the values assigned to nodes are based on the reachability of the arcs incident with the nodes.

D. HUBS AND CONNECTORS

Over the years the idea that nature can be modeled according to some sort of intricate network of various entities that share common connections has emerged. Erdős and Rényi were the first to introduce the idea that the universe consisted of properties that could be modeled using randomized graphs, but they concluded that the connections shared among the entities in the graph were completely random [20]. Malcolm Gladwell, discovered that there are certain individuals in a social network that tend to have a more influential role within the network then others based on the amount of connections that they share. These individuals are labeled as connectors. Connectors are very important to a social network. According to Albert-László Barabási, author of Linked, "they are the thread of society..." [20]. The idea of connectors was the final discovery that forced researchers to abandon the idea of a random universe. Connectors within the graph of a social network are the nodes that have an unusually large number of links, known as hubs [20]. Social networks such as Hollywood actors and actresses and the World Wide Web have clusters of nodes that are relatively close to a centrally located node, which is a hub of the network. Hubs are responsible for the majority of the connections that occur within its specified network [20]. With the World Wide Web, a hub is a site that is most highly visible and it is commonly used to connect web surfers from web-site to web-site, such as the sites Google and Yahoo. If the Web was a randomized network then all the web-sites would have the same probability to be seen and heard, which is not true [20]. According to Barabási, Web-sites that are connected to hubs like Google have a much higher probability of being noticed then web-sites that are not linked to Google. Connectors or hubs "dominate the structure of all networks in which they are present..." [20]. Hubs distinct and unusual amounts of links to other nodes, allows them to create the shortest path between two entities in the same network. Hubs can be found in society, the Web, and even cells. Connectors and hubs have opened a new door to the science of social networks and they have given rise to the idea that the universe may not be a random entity that it was once conceived to be.

E. BLACK AND SCHOLES

The Black-Scholes model is a stochastic differential equation that is used to value a call and put option for common stock [21]. An option is "a security giving the right to buy or sell as asset, subject to certain conditions, within a specified period of time." [22]

There are two different types of options, a European and American option. The European option can only be exercised on a specific future date, while an American option can be exercised anytime prior to the future date [23]. The price that is paid after the option is exercised is known as the strike price and the last day that the option can be exercised is known as the expiration date or maturity date [24]. A call is an option to buy a stock at the strike price prior to the expiration date and a put is an option to sell a stock at the strike price [22]. If the market value of the option is higher then the strike price at the expiration date, then holder will receive the difference. The focus of the Black-Scholes Model is for the European call option, because given the option can only be exercised on a specific expiration date at time t, the model will predict the value of the option in terms of price at that time under ideal conditions. The Black-Scholes model consists of five key values that relate to the stock and its behavior in the market. The five critical variables are what drives the formula. The assumptions, variables, and derivation of the Black-Scholes model is applied to the derivation of the Stochastic Valuation Transform, which will be presented in this paper, in order make the connection between stock behavior in the stock market and nodal behavior in a social network.

F. SIMULATION STRATEGY AND INVESTIGATION

The simulation strategy was focused on the generation of a social network through the IkeNet research in order to develop an accurate stochastic transform model. The generation of a network through the ORA creates a visual observation and data collection of the IkeNet data. A computer script created in Perl converted the networking data from IkeNet to data files that could be used in the ORA and Excel programs to allow for research and simulation. The use of ORA and Excel were crucial to conducting a social network analysis that can connect to the Black-Scholes stock analysis. Figure 1 is an example of how nodal valuation can change over time while Figure 2 is an example of how stock value changes over time. Both behave similarly to the Wiener processes of stochastic modeling, which means the exhibit a random projection through time or a random "walk." From investigating the IkeNet network behavior in ORA, we uncovered, through observation, a connection between the randomness of a stock value to the randomness of the value of a node as it changes through time. This is an important breakthrough to the research because if the behavior of nodal values can be connected to the behavior of stock prices, then it is valid to make the assumption that the Black-Scholes closed-form equations can be applied to social

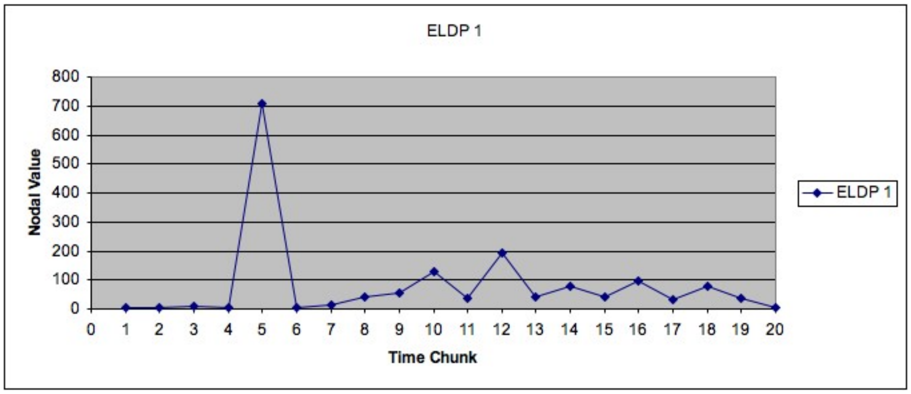


Figure 4. Nodal values vs. Time

network valuation. From this assumption, discrete nodal values were extracted at different time increments in the study. Then, those extracted values were used with a transformed version of the Black-Scholes model, which will be discussed further in the next section, to develop new values that that predict their involvement with a given event that occurred within the network.

Figure 5. Stock values vs. Time

III. STOCHASTIC VALUATION TRANSFORM

A. DEFINING VARIABLES AND THE CLOSED-FORM SOLUTION

To stochastically model the changing values of the entities within a social network, the assumption that the percentage change in the nodal values follow a Weiner process with a known nodal drift, or nodal value change, and standard deviation of that drift, so that on the date of the significant event within the network, the nodal value has a standard normal distribution with a known mean and variance [23]. The same assumption was made in the formulation of the Black-Scholes model. In concurrence with the Black-Scholes model, the Stochastic Valuation Transform (SVT) will consist of five different variables that will drive the formula based on the assumption of ideal conditions within the social network. Assuming ideal conditions within a social network means:

- There is a designated set of entities that define a social network.
- There is consistent communication within the specific entities defined in the social network.
- The network status rate of a node is known and constant through time.
- The value of a node follows a random walk with continuous time with a constant variance rate of change, thus making the finite interval of a node lognormal.
- The changing of nodal capability is continuous.
- Increasing nodal communication increases nodal influence in a social network

Following the variable characteristics of the Black-Scholes model, the variables for the SVT were created in order to be applied to social network analysis and valuation. The variables are used to assign values to the entities within the social network, which can be applied directly to the SVT formula:

V = current value of the node$_i$
σ = nodal volatility
T = time until the event date
r = Network Status Rate
P = predicted nodal value at event date
s = slope of nodal value change for node$_i$ from current value to final value

The variable V is assigned to every node within a network using the following formula:

$$V = L_i(T) + C_i^*(T) * M_i$$

where $L_i(T)$ = total number of messages sent or received for node$_i$ at time T [17]

$$C_i^*(T) = 1/(1 + exp(-10(x^* - 1/2)))$$

x^* = $L_i(T)$/(maximum possible links for nodes in time T), and M_i = cumulative total of messages sent or received for node$_i$ during time interval

The variable $L_i(T)$ is used to account for two factors:

1. The amount of connections a node has within the social network.
2. The frequency of the connections that the node has with other nodes.

The variable $C_i^*(T)$ is a value acquired from the ORA Program that measures how capable or available a node is within the network [17]. Even though a node n may not have sent or received a message from node j during a time period in the network, the fact that he could have interacted with node j is what generates the capability value for node n. The variable capability variable is used in the nodal value formula because it prevents the calculation from becoming zero, in order to make the value behave similar to a stock price.

The variable M_i is used in the formula to provide the intrinsic value that a node has within each time n. When M_i is multiplied by $C_i^*(T)$ it creates an amount of messages that the node could have sent or received base on how capable it was during the specific time period. The variable σ is assigned using the formula:

$$\sigma = \sqrt{((\sum(\Delta - \bar{X})^2)/T_{Total})}$$

where

$$\Delta = (V_T - V_{T+1})/V_T,$$

$\bar{X}$ is the Average Nodal Value (V) of node$_i$ for the time interval (T_{Total}).

The formula for σ is derived from the equation for Standard Deviation of a Population [25]. The one alteration to the equation is the Δ which is used to calculate the nodal value change for node$_i$ between each time n. The σ formula coincides to the volatility variable used in the Black-Scholes model, which was the standard deviation of the annual rate of return [23]. The SVT uses the standard deviation of the nodal value change to produce the nodal volatility.

The variable r is assigned using the formula:

$$r = \mu(\sup(V))/\sup(\sup(V))$$

where $\sup(V)$ is the Maximum Nodal Value for node$_i$, $\mu(\sup(V))$ is the Mean Maximum Nodal Value for all the nodes in the network, and $\sup(\sup(V))$ is the Maximum of all the Maximum Nodal Values.

The formula for r averages the maximum nodal value for each node$_i$ and divides that by the supremum of the Maximum Nodal Values of all the nodes in the network. This creates the Network Status Rate, which approximates how much status each node within the social network must at least grow through the time interval, just by being involved in the network.

The variable P may vary based on the analyst's prediction for each node within the network with respect to their contribution to an event. This value allows for simulation and analysis.

The variable s is the slope between the beginning value for node$_i$ for the time interval and the value of the node$_i$ at the time of the event. The variable s is assigned using the formula:

$$s_i = (V(T_n) - V(T_1)/(T_n - T_1)$$

The purpose of the Stochastic Valuation Transform is to produce an "RJ-Value" for each node within the network at every different time n, which can be used to predict how much each node contributed at each time step leading up to the time of the event. Applying the social network variables defined above to the closed-form equations of the Black-Scholes model produces the SVT [23]:

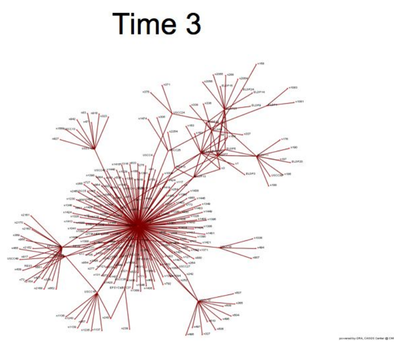
$$RJ\,Value = \begin{cases} VN(d_1) - Pe^{-rT}N(d_2) & \text{if } s \geq 0 \\ Pe^{-rT}N(-d_2) - VN(-d_1) & \text{if } s < 0 \end{cases}$$

where

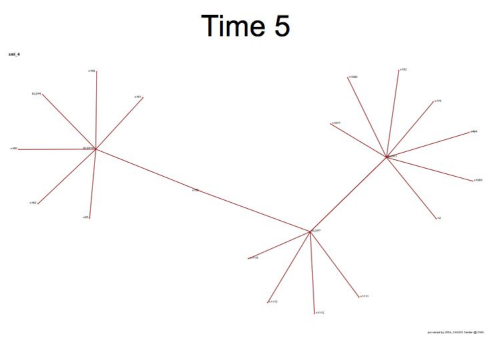
$$d_1 = \ln(V/P) + (r + \sigma^2/2)T$$

$$d_2 = d_1 - |\sigma|\sqrt{T}$$

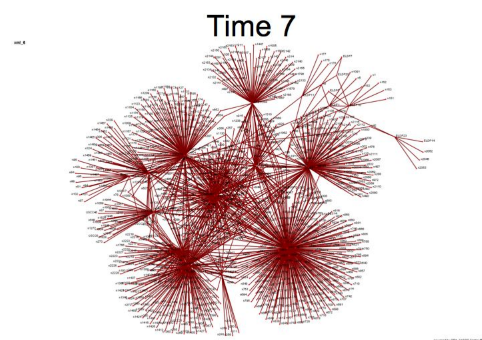
N stands for the standard normal cumulative distribution function with a mean of zero and standard deviation of one. N is a probabilistic distribution that contributes to the predictive nature of the transform. With the SVT, an analyst can search through a social network from a date where an event occurred to determine the key nodes that contributed to that event's occurrence.



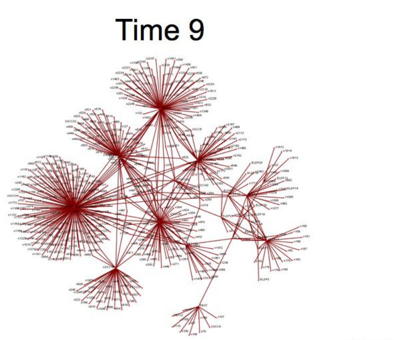
(a) Moderate Activity



(b) Deescalation of Activity



(c) Robustness



(d) Modest Lull

Figure 6. Snapshots of the Network over time

B. DERIVATION

Following the Black-Scholes derivation, a derivation of the partial differential equation (PDE) that is used to create the SVT closed-form piecewise defined function will be shown.

The nodal value V is not a deterministic variable. It is assumed that the movement of the nodal value follows the geometric Brownian motion given by

$$dV = \eta V dt + \sigma S dW \tag{III-1}$$

where W is a Wiener process or Brownian motion and η (the percentage drift) and σ (the percentage volatility) are constants [23].

Now the derivation will use Itō's Lemma to formulate the SVT Equation. Let $x(t)$ be an Itō (or generalized Wiener) process. That is let

$$dx(t) = a(x,t)dt + b(x,t)dW \tag{III-2}$$

where W is a Wiener process and let $f(x,t)$ be a function with continuous second derivatives [23].

Then $f(x(t),t)$ is also an Itō process, and

$$df(x(t),t) = (a(x,t)\frac{\partial f}{\partial x} + \frac{\partial f}{\partial t} + \frac{1}{2}b(x,t)^2\frac{\partial^2 f}{\partial x^2})dt + b(x,t)\frac{\partial f}{\partial x}dW \tag{III-3}$$

Itō's Lemma is crucial for the derivation of the SVT equation [23].

Now let RJ be some sort of value on V - mathematically RJ is a function of V and t. $RJ(V,t)$ is the value of the node at time t if the value of the underlying node at time t is V. The value of the node at the time that the event occurs is known. To determine its value at an earlier time we must know how the value evolves as we go backward in time.

By Itō's Lemma, we have

$$dRJ = (\eta V\frac{\partial RJ}{\partial V} + \frac{\partial RJ}{\partial t} + \frac{1}{2}\sigma^2 V^2\frac{\partial^2 RJ}{\partial V^2})dt + \sigma V\frac{\partial RJ}{\partial V}dW \tag{III-4}$$

Now lets construct a fictitious node of value Π consisting of one large RJ-Value and a number $-\Lambda$ (to be determined later) of the underlying node, which gives

$$\Pi = RJ - \Lambda V \tag{III-5}$$

Differentiating this gives the change in the node's value as

$$d\Pi = dRJ - \Lambda dV \tag{III-6}$$

where Λ is assumed fixed for the time step dt [23]. Substituting Equation (4) into (6) yields

$$d\Pi = (\eta V \frac{\partial RJ}{\partial V} + \frac{\partial RJ}{\partial t} + \frac{1}{2}\sigma^2 V^2 \frac{\partial^2 RJ}{\partial V^2})dt + \sigma V \frac{\partial RJ}{\partial V} dW - \Lambda dV \tag{III-7}$$

Using Equation (1) we get

$$\begin{aligned} d\Pi &= (\eta V \frac{\partial RJ}{\partial V} + \frac{\partial RJ}{\partial t} + \frac{1}{2}\sigma^2 V^2 \frac{\partial^2 RJ}{\partial V^2})dt + \sigma V \frac{\partial RJ}{\partial V} dW - \Lambda(\eta V dt + \sigma V dW) \\ &= (\eta V \frac{\partial RJ}{\partial V} + \frac{\partial RJ}{\partial t} - \Lambda \eta V + \frac{1}{2}\sigma^2 V^2 \frac{\partial^2 RJ}{\partial V^2})dt + (\frac{\partial RJ}{\partial V} - \Lambda)\sigma V dW \end{aligned} \tag{III-8}$$

which is a random change in the value of the node in time dt [23].

Next, carry out the no diminishing status rule as follows: take the node with value Π and place it into a social network with a network status rate r to get a gain of

$$d\Pi' = r\Pi dt$$

in time dt [23]. Since there is no diminishing status possibilities, we must expect that

$$d\Pi = d\Pi'$$

for $d\Pi > d\Pi'$, then add the node into the network at a rate r to get a status rate for the node. If $d\Pi < d\Pi'$, then construct the reversed or negative impacting node, and lend the value of the node at rate r to get a status rate for that node. Therefore we must have

$$(\eta V \frac{\partial RJ}{\partial V} + \frac{\partial RJ}{\partial t} - \Lambda \eta V + \frac{1}{2}\sigma^2 V^2 \frac{\partial^2 RJ}{\partial V^2})dt + (\frac{\partial RJ}{\partial V} - \Lambda)\sigma V dW = r\Pi dt \tag{III-9}$$

By equating the change in the node's involvement with an event to the status rate of Π, we have eliminated the randomness associated with the node [23]. This is only true if the coefficient of dW is equal to zero, therefore

$$\Lambda = \frac{\partial RJ}{\partial V} \tag{III-10}$$

and Equation (8) becomes completely deterministic [23]. From Equations (5) and (9) we get

$$\eta V \frac{\partial RJ}{\partial V} + \frac{\partial RJ}{\partial t} - \eta V \frac{\partial RJ}{\partial V} + \frac{1}{2}\sigma^2 V^2 \frac{\partial^2 RJ}{\partial V^2} - r\Pi = 0$$
$$\frac{\partial RJ}{\partial t} + \frac{1}{2}\sigma^2 V^2 \frac{\partial^2 RJ}{\partial V^2} - r(RJ - V\frac{\partial RJ}{\partial V}) = 0$$

Rearranging gives

$$\frac{\partial RJ}{\partial t} = -\frac{1}{2}\sigma^2 V^2 \frac{\partial^2 RJ}{\partial V^2} + r(RJ - V\frac{\partial RJ}{\partial V}) \tag{III-11}$$

which is the **Stochastic Valuation Transform equation** for assigning RJ-Values to nodes within a network. The solution to Equation (11) for $RJ = max(V - P, 0)$ on the event date gives the formula for the nodal RJ-Value if $s \geq 0$ [23]. The solution for Equation (11) for $RJ = max(P - V, 0)$ on the event date gives the formula for the nodal RJ-Value if $s < 0$ [23].

THIS PAGE INTENTIONALLY LEFT BLANK

IV. TEST AND ANALYSIS

Most Influencial Nodes for Event at Time 20		
Node	Integrated RJ Value	Normalized RJ Value
ELDP 1	-475.5854137	-0.017570833
ELDP 2	-2.390954248	-8.83355E-05
ELDP 6	187.1791033	0.006915462
ELDP 7	-145.7831345	-0.005386059
ELDP 10	-70.13551985	-0.002591205
ELDP 11	257.8341823	0.009525863
ELDP 13	105.1517542	0.003884905
ELDP 23	-17.75048466	-0.000655804
ELDP 24	-20.83364456	-0.000769713
EPSYCH 2	-20.93786959	-0.000773564
EPSYCH 4	1509.499252	0.055769498
RES 13	9852.760142	0.364017064
RES 7	-113.0164874	-0.004175473
USCC 10	25738.50805	0.95092705
USCC 12	11665.01715	0.430972158
USCC 13	1161.43599	0.042910059
USCC 15	27066.75348	1
USCC 16	11088.94771	0.409688872
USCC 18	-458.4746941	-0.016938666
USCC 20	-1653.757534	-0.06109922
USCC 23	-356.6607547	-0.013177079
USCC 24	-832.1648313	-0.030744907
USCC 25	-2139.234245	-0.079035494
USCC 26	166.3793055	0.006146999
1st	2nd	3rd

Figure 7. Most Influential Nodes

A. SIMILARITIES AND DIFFERENCES

The key difference between the RJ-values produced from the SVT and the call option value produced in the Black-Scholes model is that the RJ-values can be less then zero. In a call option, if the price of the stock goes below the strike price at expiration, then the buyer will not exercise the option thus assigning the option a value of zero. In a social network, a node that has a lower Nodal Value (V) than its Predicted Value (P) at the time of an event is assigned a negative RJ-value. Unlike the call option, a negative RJ-value has meaning within a social network. A node that is assigned a negative RJ-value at a certain time measures the lack of contribution or how much deterrence the node had with that event as compared to the other nodes in the network. A node i that is assigned a more negative RJ-value then a node j at time n means that node i is farther removed from involvement with a particular event at time n then node j. On the other hand, a positive RJ-value measures the amount of involvement or contribution a node had with an event that occurred at time n. A node with a large positive RJ-value relative to the other nodes within the network at the time of an event means that the node was a more prominent contributor to causing that particular event to occur then the nodes with smaller positive values. For example, if the RJ-values were calculated for the basketball situation explained in the abstract, then at the game ending buzzer, the game winning shooter would have the relatively largest positive RJ-value, the coach would have a low positive value, a player sitting the on bench would have a small negative value, and the team mascot would have the relatively largest negative value based on their contribution to the event of winning the game.

B. SIMULATION AND RESULTS

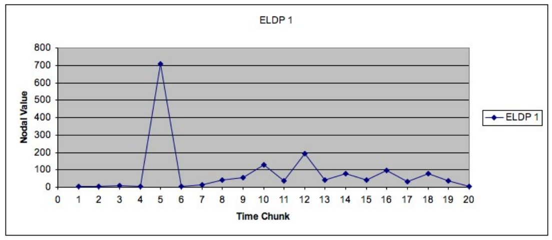
In order to analyze the formulation of the SVT, a simulation was run in Excel on the 24 critical subjects involved in the IkeNet social network research. In Excel, the RJ-values were calculated for the subjects (nodes) at each of the 20 time steps. The lined scatter plot graphs depicted below show how the RJ-values for each node fluctuated through the entire time interval of the study. By observing and comparing the graphs of the RJ-values with the graphs of the Nodal Values (V), they show that the SVT is working properly because there is a correlation between the RJ-values and the behavior of the Nodal Values (V). If a node had a low Nodal Value because it did not send or receive any messages and had a low capability at a time n, then it was assigned a negative RJ-value, because it mostly likely would not have any involvement in an event that occurred at that time n. It was the same result for positive values as well. Based on the assumptions made and the essence of stochastic value modeling, the SVT produces substantive data. In order to find the most influential nodes over the whole time interval T leading up to the event that occurred at time step 20, the Trapezoidal Integration Method was used to integrate the changing RJ-Values for each of the 24 critical nodes. The equation for the Trapezoidal Integration Method is

$$T_n = \frac{1}{2}\sum_{i=1}^{n}[\frac{f[x[i-1]]\Delta x + f[x[i]]\Delta x}{2}]$$

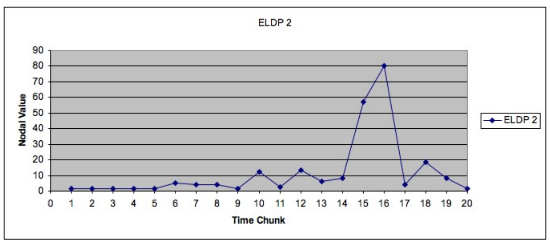
where

$$\Delta x = \frac{b-a}{n}$$

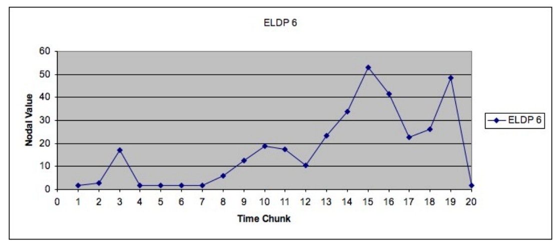
The integration gave the overall RJ-Values for each node, which was then normalized to divulged most influential nodes within the network that caused an event to occur at time 20 according to their RJ-values. Below is a chart that shows the resulting data:



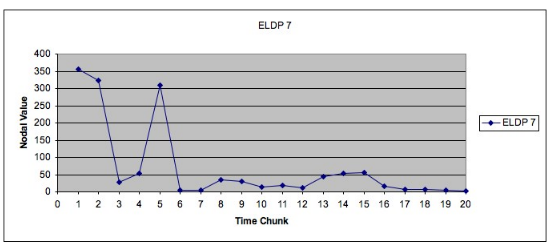
(a) Strong Concentrated



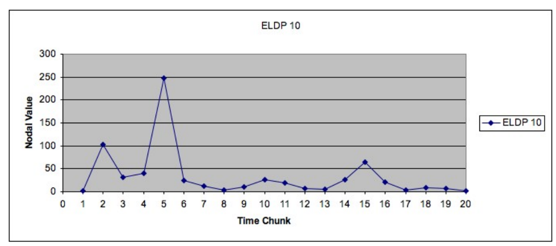
(b) Mild End Concentration



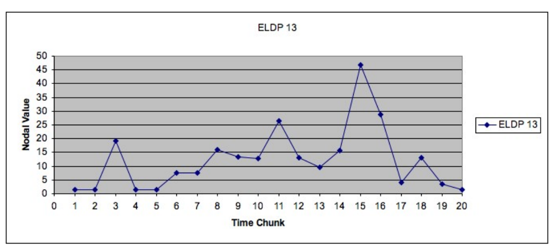
(c) Positive Slope



(d) Negative Slope

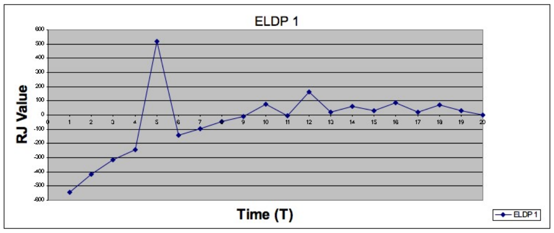


(e) Mild Volatility

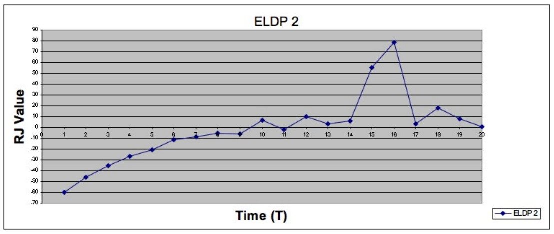


(f) Moderate Volatility

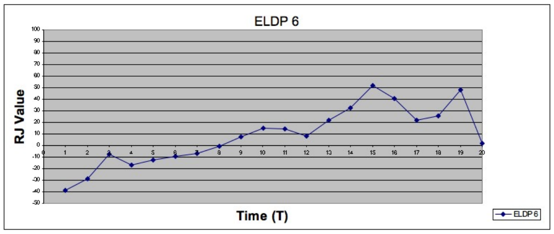
Figure 8. Network Nodal Values Over Time



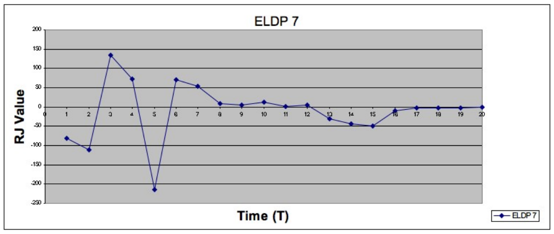
(a) Concentrated with Increasing Influence



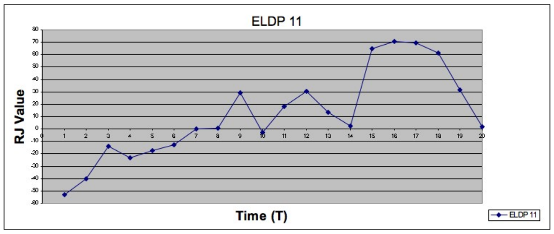
(b) Rising End Peak Influence



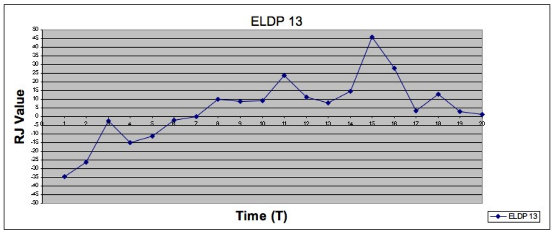
(c) Mild Stable Climb



(d) Volatile and Flat



(e) Strong Influence



(f) Slight Climb

Figure 9. Values Computed Using RJ Method Over Time

V. FUTURE WORK AND CONCLUSION

The motivation behind the formulation of the Stochastic Valuation Transform is the direct application to the Army. If the SVT proves to be an accurate means of nodal value generation after intense testing, simulation, and analysis, it could be applied in hostile environments such as Iraq, in order to provide a means to tracing the key insurgent cells that contribute to attacks on friendly forces. Since the SVT is the first use of stochastic modeling for Social Network Analysis, there is much room for future work in this field. Further research and analysis of social network behavior could expose more precise variables for nodal valuation that increase the predictive accuracy of the SVT or it could lead to a better stochastic model all together. The SVT uses the measurements of email traffic, capability, nodal volatility, network status rate, and time to compute the RJ-value. Increased amounts of model simulation on multiple types of social environments are needed to test the quality and accuracy of the variables and assumptions of the model.

Similar to predicting the option value of a call within the stock market, in a social network it is difficult to predict with 100 percent accuracy who the key individuals or groups are that cause events to occur within the network at a given time[26]. What makes a social network unpredictable is its fluid nature. There are critical factors that need to be accounted for and there are assumptions that need to be made in order to develop a sufficient predictive model, but since the social network is in constant change, the factors and assumptions may not always account for the real-world dynamics of a social network. The Stochastic Valuation Transformation offers a new stochastic technique to assigning values to nodes based on an event that occurred within a given network. The SVT is a transformation of the Nobel Prize Winning Black-Scholes closed-form equations for value pricing an option. The transform allows the closed-form Black-Scholes equations to be applied to Social Network Analysis based on the assumption that the behavior of nodes in a social network is similar to the behavior of stock in the stock market. The SVT allows social network analysts to compute a value called an RJ-value for each entity within the network that predicts their involvement or lack thereof in a specific event that occurred at a given time within the system. The SVT assigns the RJ-Value to the nodes within a network in order analyze and predict future behavior and interactions of the entities involved, which could lead to new discoveries in the field Social Network Analysis.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

[1] S. T. Fiske and S. E. Taylor. Social Cognition. McGraw-Hill, Inc., New York, 1991.

[2] D. Krackhard. Assessing the political landscape: Structure, cognition, and power in organizations. Administrative Science Quarterly, Vol. 35:342–369, 1990.

[3] C. A. Ellis, S. J. Gibbs, and G. L. Rein. Groupware: Some issues and experiences. Communications of the AMC, Vol. 34(1):39–58, 1991.

[4] C. Gutwin and S. Greenburg. The importance of awareness for team cognition in distribution collaboration. Team Cognition: Understanding the Factors that Drive Process and Performance., pages 177–201, 2004.

[5] H. R. Bernard, P. D. Killworth, and L. Sailer. Informant accuracy in social network data iv: A comparison of clique-level structure in behavioral and cognitive network data. Social Networks, Vol. 2:191–218, 1980.

[6] D. A. Kashy and D. A. Kenny. Do you know whom you were with a week ago friday? a re-analysis of the bernard, killworth, and sailer studies. Social Psychology Quarterly, Vol. 53:55–61, 1990.

[7] H. R. Bernard and P. D. Killworth. Informant accuracy in social network data ii. Human Communication Research, Vol. 4:3–18, 1977.

[8] M.E.J. Newman. The Mathematics of Networks. University of Michigan, Ann Arbor, MI, 2005.

[9] Charles J. Fombrun. Strategies for network research in organizations. The Academy of Management Review, Vol. 7(2):280–291, 1982.

[10] Tom A. B. Snijders. The statistical evaluation of social network dynamics. Sociological Methodology, Vol. 31(1):361–395, 2001.

[11] P. D. Killworth and H. R. Bernard. Informant accuracy in social network data iii: A comparision of triadic structure in behavioral and cognitive data. Social Networks, Vol. 2:19–46, 1979.

[12] P. D. Killworth and H. R. Bernard. Informant accuracy in social network data. Human Organization, Vol. 35:269–96, 1976.

[13] S. Wasserman and K. Faust. Social Network Analysis, Methods and Applications. Cambridge University Press, New York, 1994.

[14] Stanley S. Wasserman. Analyzing social network as stochastic processes. Journal of the American Statistical Association, Vol. 75(370):280–294, 1980.

[15] L. C. Freeman. Centrality in social networks: I. conceptual clarification. Social Networks, Vol. 1:215–239, 1991.

[16] Milton E. Harvey. The identification of development regions in developing countries. Economic Geography, Vol. 48(3):229–243, 1972.

[17] ORA. ORA: Organizational Risk Analyzer v. 1.7.8. Carnegie Mellon University, Pittsburgh, PA, 2007.

[18] Stanley S. Wasserman. Models for binary directed graphs and their applications. Advances in Applied Probability, Vol. 10(4):803–813, 1978.

[19] Han Dieperink and Peter Nijkamp. Multiple criteria location model for innovative firms in a communication network. Economic Geography, Vol. 63(1):66–73, 1987.

[20] A. Barabási. Linked. Penguin Group Inc., New York, 2003.

[21] D. J. Bartholomew. Recent developments in nonlinear stochastic modelling of social processes. The Canadian Journal of Statistics, Vol. 12(1):39–52, 1984.

[22] Fischer Black and Myrom Scholes. The pricing of options and corporate liabilities. The Journal of Political Economy, Vol. 81(3):637–654, 1973.

[23] Donald Outing. Mathematical Modeling Lecture Notes Black-Scholes Model. United States Military Academy Department of Mathematics, West Point, NY, 2006.

[24] Fischer Black and Myrom Scholes. The valuation of option contracts and a test of market efficiency. The Journal of Finance, Vol. 27(2):399–417, 1971.

[25] Mario F. Triola. Elementary Statistics Using Excel Third Edition. Pearson Education Inc., Boston, MA, 2007.

[26] H. R. Bernard, P. D. Killworth, and L. Sailer. Informant accuracy in social network data v: An experimental attempt to predict actual communication from recall data. Social Science Research, Vol. 11:30–66, 1982.