3-9-2006

A Graph Theoretic Analysis of the Effects of Organizational Structure on Employee Social Networks

John R. Hutzel

A GRAPH THEORETIC ANALYSIS OF THE EFFECTS OF ORGANIZATIONAL STRUCTURE ON EMPLOYEE SOCIAL NETWORKS

THESIS

John R. Hutzel, Captain, USAF

AFIT/GRD/ENV/06M-07

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

AFIT/GRD/ENV/06M-07

A GRAPH THEORETIC ANALYSIS OF THE EFFECTS OF ORGANIZATIONAL STRUCTURE ON EMPLOYEE SOCIAL NETWORKS

THESIS

Presented to the Faculty

Department of Systems and Engineering Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Research and Development Management

And

Degree of Master of Science in Systems Engineering

John R. Hutzel, B.S.

Captain, USAF

March 2006

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

AFIT/GRD/ENV/06M-07

A GRAPH THEORETIC ANALYSIS OF THE EFFECTS OF ORGANIZATIONAL STRUCTURE ON EMPLOYEE SOCIAL NETWORKS

John R. Hutzel, B.S.

Captain, USAF

Approved:

/signed/	9 Mar 06
Dennis D. Strouble, PhD, JD (Chairman)	Date
/signed/	9 Mar 06
John M. Colombi, Lt Col, USAF (Member)	Date
/signed/	9 Mar 06
Laura R. C. Suzuki, Maj, USAF (Member)	Date

AFIT/GRD/ENV/06M-07

Abstract

A simulation technique was used to investigate the impacts of organizational structure on an organization's social network. By simulating personnel in an organization as vertices in a graph and the aging of the corporation as the aging of the same graph, the maturation of an organization was realized. The characteristic path length of the graph was measured after each year returning an optimistic average organizational distance. Results include the finding that, per this model, an organization's characteristic path length can drop over 50% in a 20 year period with consideration of edges of all strengths. Next a series of random searches were performed to measure the ability of an individual to search for information. A concern of this research is the difference between a partitioned and a non-partitioned organization's social network. A partitioned organization is defined as one in which personnel are restricted to smaller communities during a career. Partitioning the organization decreases the travel distance, in this model, when searching within a smaller community by 6% and it increases the travel distance when searching within the larger community by 12%. This must be considered along with the organization's information flow needs when structuring an organization.

AFIT/GRD/ENV/06M-07

For my loving wife,

Thank You

Acknowledgments

I would like to express my sincere appreciation to my faculty advisor, Dr. Dennis Strouble, for his guidance and support throughout the course of this thesis effort. The insight and experience was certainly appreciated. Further, I would like to thank Lt Col John Colombi and Maj Laura Suzuki for the many hours spent contemplating our next steps on this journey. Thank you all.

John R. Hutzel

Table of Contents

Page

List of Figures

Page

A GRAPH THEORETIC ANALYSIS OF THE EFFECTS OF ORGANIZATIONAL STRUCTURE ON EMPLOYEE SOCIAL NETWORKS

I. Introduction

Background

In June 2004 the Aeronautical Systems Center (ASC) at Wright-Patterson AFB, OH restructured into an organization more representative of traditional Air Force units (ASC/PA, 2004). Under this new structure, personnel are assigned to Wings, Groups, and Squadrons reporting to the ASC Commander, a Lieutenant General. Whereas personnel, particularly civilians, were previously assigned to functional home-offices and assigned to programs in a matrixed fashion, under the new construct, the Wings now have direct authority over personnel placement. As a result, a civilian may have a much greater probability of staying within one organization throughout a career than under the previous construct.

Problem Statement

Under the new ASC construct, assuming employees stay predominantly within one Wing or Group throughout a career, logically, personnel will have a much lower probability of forming long-range relationships crossing the corporate ASC. The hypothetical reason for this decrease in long-range relationships is the result of not working together at various times throughout a career and then moving to different positions in different communities. Psychologists have shown us that personnel tend to form relationships with those they interact with on a regular basis (Festinger, 1950). Therefore, removing interactions by disallowing personnel from working with others throughout a career (due to residing in different Wings) would decrease the probability of forming those relationships. As the organization ages, personnel would have a greater amount of interactions with members of the same Wing, thus increasing the strength of these relationships, but at the consequence of relationships with members of other Wings. Particularly under a crisis situation, this lack of knowledge of others outside an individual's community could have a serious impact on an individual's ability to find solutions to a problem. This consequence is quite poignant in a problem solving organization where no particular sub-community will have solutions to all of the corporate entities problems.

Research Objectives/Hypotheses

The hypothesis for this research is that an organization that undergoes this type of partitioning will increase the social connections within each smaller community at the penalty of connections within the larger community. Further, this lack of strong connections to members of different communities within the corporate organization will hinder the ability of an individual to search for solutions within different parts of the organization. The objective of this research is to take an unbiased approach to characterizing the changes in an organization's ability to communicate and search within itself for information. Further, this research will consider the strength of relationships formed and some ways a manager could work to overcome potential setbacks due to a partitioned structure.

Methodology

This research will be based on the simulation of an organization as a graph. After initial creation, this graph will be matured by removing personnel (simulating firings, retirements, separations, etc.) and promoting or moving other personnel into the vacancies created with these removals. Further, personnel will participate in long-range collaborations throughout each time step in order to consider the effects of cross-organizational tiger teams and integrated product teams on the organization as a whole. These cross-organizational collaborations will allow the consideration of a way to overcome the effects of partitioning. In addition to moving personnel and commencing cross-organizational collaborations, each relationship will be aged to reflect the strengthening of a relationship when working together and the weakening of a relationship when not working together. This will allow for the consideration of the strengths of a relationship throughout the relationship's life.

After aging the organization, the graph representing that organization will be analyzed. The initial analysis will consider the characteristic path length of the organization. The characteristic path length of an organization is basically the average shortest length between any two personnel within the organization (similar to the "Kevin Bacon number" from the popular party game averaged over all pairs of people in an organization). Since this measure returns the average shortest path, it represents an optimistic distance between personnel. The next form of analysis will include a search of the organization. While several search algorithms already exist, this research will take advantage of one tuned for a military construct in order to more accurately reflect the military structure and ethos of ASC. The purpose of this search research is to consider the situation where an employee does not know the location of a solution and must ask multiple fellow employees, representing a more realistic distance than the characteristic path length measure. The characteristic path length measure in combination with the search algorithm will show how many personal links each person must follow to get to any other person in the organization. The number of links has a direct relationship to the ability of a person in an organization to search out information or knowledge from another part of the organization.

These analyses will be run to consider three organizational variables. The first variable considered will be the consideration of attrition rate. The second consideration will be the average number of collaborations members of the organization will participate in over a year. And the third consideration will be the effects of changing the organizational structure on the characteristic path length and the edge search length. The organizational changes considered will be a change from a non-partitioned organization to a partitioned one, following the choice of ASC. The second consideration will be an organizational structure change from a partitioned organization to a non-partitioned construct to consider the opposite case and provide insight to an organization that is partitioned and whishes to consider ways to potentially enhance cross-organizational communication.

Assumptions/Limitations

The primary assumption of this research is that each person in an organization is replaceable with any other on the same level or the level immediately below from the same community. A consideration of personnel ability to fill positions is outside the scope of this research and a potential extension for further research. These considerations would include those based on experience and tenure, subjects this simulation did not address.

A secondary assumption of this research is that each partitioned community has the same probability to communicate or search for information with any other community. The reason for this assumption was to ensure the model did not prefer any organization over any other in promotions or long-distance collaborations. Further, this assumption was made to clarify the model and the results. This assumption would not be true for all Wings within ASC, though, since there would be a natural tendency for certain Wings to communicate with the same Wings regularly.

Lastly, the model must estimate ASC's seven Systems Wings and Groups. ASC's seven Wings or Groups include:

1. Fighter Attack Systems Wing
2. Long-Range Strike Systems Wing
3. Reconnaissance Systems Wing
4. Mobility Systems Wing
5. Agile Combat Support Systems Wing
6. Special Operations Forces Systems Group
7. Training Aircraft Systems Group

In order to estimate these organizations, six communities were chosen as computationally able estimate for ASC. Further, the assumed number of levels for the organization is five levels. This returns a total organizational size of 1555 employees. While this is significantly smaller than ASC as a whole (ASC has approximately 11,800 personnel), this is a large enough number of employees to glean trends from and is computationally capable (ASC/PA, 2004). Further, this assumption of a branching ratio of six creates a highly regular organizational structure. As a result each Wing and Group has exactly the same number of personnel and the same structure.

Preview

This research will not only consider the effects of changes in an organizations structure on its social network, thus implying effects on that organizations ability to share information and search within itself, but also several factors affecting contemporary organizations. The first input would be a consideration of the speed with which an average relationship is formed between any two personnel. Next, consideration will be given to long-range collaborations. A superior might choose to encourage cross-organizational interactions by forming long-range tiger teams or integrated product teams. This measure is intended to gain an understanding of the effects of this decision on the organizations social network. The next consideration will be the effects of attrition on the organization. Any organization will go through times of varying attrition rates. Further, ASC is expecting a larger percentage of civilian retirements than the organization has traditionally experienced (Russo, 2003). The goal of this measure is to provide insight into the effects of these changing attrition rates on an organizations social construct regardless of the organizational structure. As a result of this simulation, a decision maker will gain a greater understanding of the effects of an organization's structure on an organization's social network and, by extension, the organization's ability to share information and search for solutions within an organization. This capability could have significant impacts not only on daily activities, but also on an organizations ability to respond to crises.

II. Literature Review

When conversing with friends, many people have commented on their observation of having mutual friends, often stating, “isn’t it a small world?” The concept of a world in which there are actually shorter distances between any two people than would be expected due to the large number of people has been studied in academia since the mid-to-late 1960s. Further, this concept has arisen in party games such as the “Kevin Bacon Game” or, for mathematicians, the “Paul Erdös Number” game. The first game relates the number of friends and friends of friends one must have before reaching the subject of the game, the actor, Kevin Bacon. The second game relates the number of links represented by coauthored research papers a researcher must cross prior to reaching a paper coauthored with the great mathematician, Paul Erdös (labeled an Erdös number, an Erdös number of one represents coauthorship with Paul Erdös and is quite coveted).

Small-World

In 1967, Stanley Milgram, sponsored by Harvard University, published “The Small World Problem.” In this document, he studied this occurrence in the real world by using the postal system. Milgram formulated his experiment as follows; he initially chose several people in one city and a single person in another city a great distance away. After sending a message to the initial people he gave these initial recipients the instructions that they could only send the message to a person they new on a first name basis that they believed would get the message closer to the intended final recipient. Through this process, he would measure the average number of people needed to get a message from point *a* to point *b*. His hypothesis was that the number of links needed to get from this *a* to *b* was actually much smaller than would be imagined by the average person due to the large population in the United States. In fact, he did run the experiment twice, with the starting points at Omaha, Nebraska and Wichita, Kansas and terminal points at Cambridge and Boston, Massachusetts, respectively. After receiving several messages and studying the results, Milgram found that the median number of intermediaries was five (Milgram, 1967). This leads to the common phrase, six degrees of separation (even becoming the title of a play by John Guare), representing the concept that we are only six links away from any other person (Guare, 1999).

Since this document and the coining of the label, "Small-World", researchers have continued to consider the phenomenon of large populations with small average degrees of separation. Of particular interest are the studies of Duncan Watts who, while studying at Cornell University, published "Collective Dynamics of 'Small-World' Networks" (Watts, 1998). In this document, Watts considered this small-world phenomenon from a network standpoint. This consideration allowed him to apply many tools developed in graph theory to mathematically define a small-world network. Further, he applied his findings to several real-world networks, including film actors, the western United States power grid, and the neural network of the worm *Caenorhabditis elegans*, thus demonstrating the large range of networks where this form of analysis is applicable (Watts, 1998).

Graph Theory

At this point, it is appropriate to introduce the concept and define several of the terms intrinsic to graph theory. The field of graph theory was originated in 1736 with the famous Königsberg bridges problem (Wilson, 1986). In this problem, the people of

Königsberg, Prussia wondered if there was any way to cross all seven bridges across the Pregel River within their city without crossing any bridge more than once. Although nobody had been able to find a way to do so, there was no proof that it was impossible rather than just a very challenging problem. Finally, in a paper written in 1736 Leonhard Euler proved there was no way of crossing all seven bridges only once (Euler, 1736). An English translation of this originally Latin paper can be found in Biggs' book, *Graph Theory 1736-1936* (Biggs, 1976). A drawing of the city, river, and island system can be seen below.

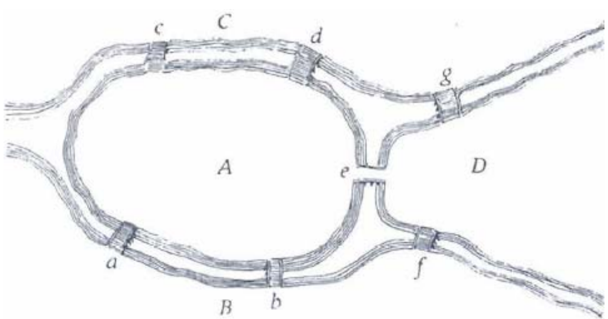

Figure 1 Bridge Network of Königsberg, Prussia, Now Kaliningrad, Russia. Euler used this interpretation of the bridge, island and Pregel River network through Königsberg, Prussia in his creation of graph theory (Biggs, 1976).

In his paper, Euler showed there was no way for a person to touch every landmass and cross each bridge only once. His argument was based on labeling the land masses as A,B,C, and D and the bridges as a,b,c,d,e, and f, in short, the masses became the vertices and the bridges became the edges of a graph. Euler then went on to state three rules for whether a journey could be made by only crossing each bridge once. An excerpt from the English translation of the 20th chapter of Euler's paper is as follows:

If there are more than two areas to which an odd number of bridges lead, then such a journey is impossible.

If, however, the number of bridges is odd for exactly two areas, then the journey is possible if it starts in either of these areas.

If, finally, there are no areas to which an odd number of bridges lead, then the required journey can be accomplished starting from any area (Euler, 1736).

Euler then continues to state that any arrangement of bridges and land areas can be solved with these considerations. Of note, though, is that, in this paper, Euler never provided a proof of the second statement. Rather, this was proven, autonomously to Euler's paper, in 1873 by Carl Hierholzer (Hierholzer, 1873). An English translation of the original German text can also be found in Biggs 1976 book, *Graph Theory 1736-1936* (Biggs, 1976). These aged papers have been built upon to provide the robust verbiage and techniques now in use, under the name graph theory to consider many contemporary problems.

For the purpose of this paper, the verbiage of graph theory will be relied on heavily as, while much of the terminology of networks is different, the concepts are generally similar or the same. In this case, a social network can be used as a metaphor for a graph where the people in the network are the vertices and the social links between these individuals would constitute the edges. Keeping this consideration in mind, a graph is formally defined as:

> A simple graph G with n vertices and m edges consists of a vertex set $V(G)=\{v_1,...,v_n\}$ and an edge set $E(G)=\{e_1,...,e_n\}$, where each edge is an unordered pair of vertices. We write *uv* for the edge $\{u,v\}$. If $uv \in E(G)$, then *u* and *v* are adjacent. We write $u \leftrightarrow v$ to mean "*u* is adjacent to *v*" (West, 1996).

In addition to the concept of a graph, an edge, and a vertex, the concept of adjacency is addressed within this definition. The concept of adjacency is the idea that there is an edge between any two vertices. Considering a social network as the network created where people are node and relationships of any strength are considered links between those nodes, adjacency in this social network construct would be when two people agree that they have a social connection. This concept of adjacency is represented by an adjacency matrix which is a square matrix (a column and a row for each vertex) in which adjacency is represented as a non-zero value and non-adjacency is represented as a zero. Again, borrowing the formal definition from West:

> Given a graph or digraph *G* with vertices indexed as $V(G)=\{v_1,\ldots,v_n\}$, the adjacency matrix of *G*, written *A(G)*, is the matrix in which entry a_{ij} is the number of copies of the edge v_iv_j in *G*. The degree of a vertex is the number of non-loop edges containing it plus twice the number of loops containing it (West, 1996).

This definition also gives us a definition for the degree of a vertex. Furthering the social network construct, the degree of any vertex within a graph describing the network would be the number of vertices that initial vertex was adjacent to, or the number of people each person knows. The concept of adjacency, as well as an adjacency network is best described with an example. A simple example of a graph can be seen in Figure 2. The vertices are numbered in the same way as they will be in the simulation used to prove our hypotheses.

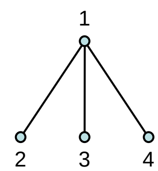

Figure 2 A Simple Tree Graph With 2 Levels and a Branching Ratio of 3. Note the vertex numbering convention used for this work where the top vertex is always one and the numbers increase as they go down and from left to right.

$$A(G) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

Figure 3 The Adjacency Matrix Corresponding to the Graph Displayed in Figure 2. This matrix assumes all relationships are of equal strength. Differences in strength will be considered by using non-zero values in the adjacency matrix other than one.

Finally, some of the ways to describe the relationship between the vertices in addition to the already presented concept of adjacency matrix must be given. The first concept is the distance between any two vertices. This concept relates to the number of edges that must be crossed in going from one point to another. An example from Figure 2 would be that the shortest distance, or geodesic, between vertices three and four is two since the path between them is from three to one and then from one to four. This is different from the geodesic from one to four, which is one since these two vertices are adjacent to one another. The characteristic path length of this graph relates to this concept since it is the average of all the geodesics within the graph. Watts provides an exceptional definition for the characteristic length of a graph:

> The characteristic path length (L) of a graph is the median value of the means of the shortest path lengths connecting each vertex $v \in V(G)$ to all other vertices. That is, calculate $d(v, j)\ \forall\ j \in V(G)$ and find $\bar{d}_v$ for each *v*. Then define L as the median of $\{\bar{d}_v\}$ (Watts, 1999).

The next descriptive concept involves the concept of a neighborhood and the clustering of the graph. The neighborhood for any vertex in the graph includes all vertices with which that vertex is adjacent. From our example in Figure 2, the neighborhood of vertex one is the set of vertices, {2,3,4} whereas the neighborhood of vertex 2 is only {1}. The definition of this concept is as follows:

> The neighborhood of *v*, Γ_v, is $\{x \in V(G) : x \leftrightarrow v\}$; *x* is a neighbor of v if $x \in \Gamma_v$ (West, 1996).

Watts builds upon this concept to define a clustering coefficient for a graph. To understand the clustering coefficient for a graph, first the concept of a clustering coefficient for a vertex must be understood. Per his usage, the clustering coefficient of a vertex is a ratio between the number of edges in a vertices neighborhood compared to the total number of possible edges. Watts formal definition is as follows:

> The clustering coefficient γ_v of Γ_v characterizes the extent to which vertices adjacent to any vertex v are adjacent to each other. More precisely,

$$\gamma_v = \frac{\# E(\Gamma_v)}{\binom{k_v}{2}} \tag{1}$$

where $\#E(\Gamma_v)$ is the number of edges in the neighborhood of ν and $\binom{k_v}{2}$ is the total number of possible edges in the neighborhood of ν, Γ_v when there are k_v vertices in that neighborhood (Watts, 1999).

The clustering coefficient for the graph is then the average clustering coefficient for all vertices in the graph.

Returning to Milgram's small-world theory, Watts, a formally trained mathematician considered whether the small-world phenomenon could be explained using graph theory. He hypothesized the small-world network topology could be found in a graph somewhere between total regularity and total randomness. In order to show this, he first constructed a highly regular graph with a ring lattice as its substrate. He then adjusted the endpoints of the edges within this graph at random with a probability that an one of an edge's endpoints had been changed of *p*. In this process, he was able to change the graph from being totally regular (p = 0) to fully random (p = 1) and investigate the dynamics in the middle area where $0 < p < 1$ (Watts, 1998). Watts' drawing of the process can be found in Figure 4.

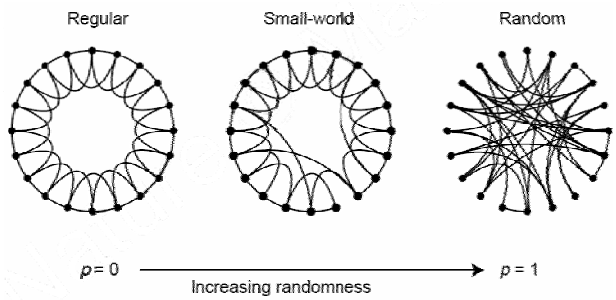

Figure 4 Watts' Small-World Interpretation Using a Ring-Lattice Substrate
(Watts, 1998).

After his simulation, he found global characteristics of a small-world network emerge within his ring-lattice graph during these intermediate steps. Of primary interest, though are the changes that occur within the graph during the period of change. Similar to the findings of Milgram, that, even for a large network (the population of the United States in the 1960s), the average number of links between any two is actually quite small, Watts found that the number of vertices within the graph does not have a linear affect on the characteristic path length of the network. His findings emphasized a handful of long-distance edges that had a disproportionate impact on the characteristic length in comparison to the normal edges of vertices that are next to one another in the ring-lattice substrate. Labeling these long-distance edges as shortcuts, he found these shortcuts had a non-linear effect on the characteristic path length of the graph since, not only did they create a much shorter link to the now connected vertices, but also to these vertices' immediate neighbors. An additional result of this phenomenon is the clustering coefficient is initially barely affected by this edge change. He actually found the graph clustering coefficient is nearly linear for the early parts of the transition from regular to random (Watts, 1998). A plot of the change in characteristic length and clustering coefficient can be seen in Figure 5.

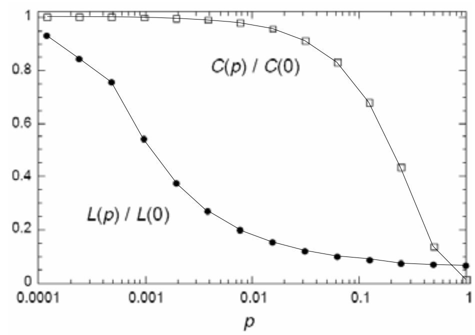

Figure 5 Effects of Increased Randomness on a Small World.
Note highly non-linear initial drop in characteristic path length (L) while clustering coefficient (C) initially stays stable with increasing randomness (p) (Watts, 1998)

After mathematically describing this phenomenon, Watts went on to consider several other readily accessible networks and found these attributes are not peculiar to sociology and the study of social networks. Watts found networks with characteristic path lengths not far removed from a totally random network, although much more clustered in a graph of film actors, power grids, and the neural network of the *Caenorhabditis elegans* worm (Watts, 1998). From this point, the study of this range of networks between regularity and randomness has grown to include several additional networks throughout science.

Power Law Networks

Of particular interest is the power law, or 'scale free' network described by Albert-László Barabási, a physicist from the University of Notre Dame. Barabási considered the evolution of actual real-world networks. He theorized there was a flaw in the writings of

Erdös and Rényi and their description of the evolution of random graphs (Erdös, 1960). Erdös and Rényi started their description of a network with a set number of vertices to which edges were added randomly to create the network (hence the name, random graphs). As a result of the assumption that edges were added at random, the vertex degree distribution within the graph follows a Poisson distribution described by Newman as:

$$p_k = \binom{N}{k} p^k (1-p)^{N-k} \simeq \frac{z^k e^{-z}}{k!} \tag{2}$$

where k is the vertex degree, N is the number of vertices and p_k is the probability of nodes having the associated degree (Newman, 2001). Barabási believed this initial assumption to be incorrect; rather, he believed real networks would not exhibit a truly random distribution. This hypothesis originated from comparing the degree distributions of actual networks to the Poisson distribution. These comparisons did not resemble the random distribution; rather, they resembled a power law distribution. Barabási then used simulation to create these networks with degree distributions resembling a power law. He hypothesized this power law distribution was the consequence of two functions:

i. true networks will continue to grow throughout their life and
ii. there is a preference for these new vertices to form edges with other well connected vertices (resulting in connections to predominantly older vertices) (Barabási, 1999).

By running several simulations with and without different combinations of these two considerations permitted, he was able to show that, in order to form this power law distribution, these two assumptions must be followed. At the same time as this publication, he coauthored a paper with Albert and Jeong in which the world-wide web was found to follow this power law distribution (Albert, 1999). In this paper, the authors considered web pages to be the vertices and hyperlinks on these web pages to be the edges within a large graph. First considering the Notre Dame intranet, then using a webcrawler to explore the entire world-wide web, they found this distribution. A further study has extended this description to the topology (on the router and inter-domain level) of the internet (Faloustos, 1999). Since the publication of these papers, research has gone into describing the dynamics of these networks with the consideration that, if a network can be identified as a power law, then a researcher could understand its characteristics without gathering all the information about the network. This enables additional studies of such large networks as the world-wide web, where gathering the names and hyperlinks of the entire network is currently not possible. Of particular import to this paper is research into the ability to search these networks.

The importance of a quick search routine for a power law network can be easily understood by anyone who has tried searching the world-wide web. The shear amount of information on the web seems insurmountable without some knowledge of where to go first. Knowledge that the web follows a power law, and by extension exhibits the traits of a small-world, means the network has short characteristic path lengths even with large numbers of vertices. As a result, the searcher knows he or she can get to the information relatively quickly, he or she just does not know where it is, so that person must search. A random search of the entire would not be optimal in this situation.

Social Networks

This situation is the basis for the 2000 research of Jon Kleinberg. Kleinberg investigated the ability of members of a small world to find the short connections between themselves within the graph without prior knowledge of the endpoint (Kleinberg, 2000). In this way, he considered the case of a person searching a social network with knowledge of who they need to reach, but without the knowledge of who could introduce them to one another. Unlike Watts with his ring-lattice substrate, Kleinberg used a *n* vertex by *n* vertex square matrix. He then allowed edges between immediate neighbors in the matrix as well as long-range edges. While his particular research was interested in the variance of search capability depending on the clustering coefficient of the graph, of particular interest to this research is his search heuristic. In Kleinberg's model, the location within the lattice of the target vertex was known to the searching vertex (ie, in the process of going from *a* to *b*, *a* knew where in the graph *b* was). He then adopted a greedy search heuristic where each successive node on the path from *a* to *b* was always chosen to move the path closer to the end state (Kleinberg, 2000). While this construct was appropriate in the situation where a searcher knows where the information they seek is, just not how to get there, Adamic was interested in the case where a searcher was aware of neither the way to get to the information nor the location of the information (ie, *a* does not know about *b* or the shortest path to *b* (Adamic, 2001)). Since, in Adamic's study, there was no global location knowledge, there is no way for a searcher to know whether the next step is any closer to the intended destination. She approached the search problem by considering two search heuristics intended for a power law network. She labeled the first heuristic the random walk (Adamic, 2001). In this construct, the search routine would simply choose the next vertex randomly from those that are adjacent to the current vertex. Further consideration for this search routine must be given to the ability to visit the same vertex twice. The second heuristic was based on a power law network's tendency to have vertices with large degrees. In this second search heuristic, upon each step, the next vertex chosen is the vertex adjacent to the current vertex with the greatest degree (Adamic, 2001). She noted this routine also presents problems since, once the vertex with the greatest degree is found, the routine must start looking toward vertices with lesser degrees. Further, the consideration of whether repeated visits to the same vertex must be considered when choosing this heuristic. Also, of particular interest to the study of a social construct is the ability or inability of this routine operating in a business environment. Restrictions might be in place due to a chain of command preventing direct communication between several layers of management. As will be discussed in the methodology, neither of Adamic's routines will be used for the search of a business construct, rather, a search routine designed for use within a graph of stable size and social regulations will be developed.

Search Routines

The next consideration is the application of these search routines to the social network found within a corporation. The first assumption to be made is that there is a greater propensity for acquaintances to be made between members of the same sub-unit. Following Krackhardt's argument in favor of this assumption, the first part of the validity of this argument is based on the number of interactions people within the same sub-unit will have over the course of any work period (Krackhardt, 1988). Initially, Krackhardt points out that psychologists have found that a high amount of interaction between people will lead to friendships (Festinger, 1950). Also, he points out findings that students who sit near one another will tend to form friendships (Byrne, 1952). Since these studies form a base for the assumption that there is a greater probability for friendships between those that interact, the lesser assumption, that personnel in the same sub-unit (in this case, assumed to be those who share a common superior and workplace) will form acquaintances is firm. Krackhardt builds upon this assumption, as well as psychological findings on the effects of experimental results to come to the proposition that organizations that are effective in times of crises are those that have friendship links between subunits (Krackhardt, 1981). In short, he argues that an organization that does not cultivate relationships between sub-units will not fair as well as those that do in times of crises. In addition to this consideration, critical to this research is his finding, based on past research that acquaintances will form within the sub-unit.

Conclusion

The considerations that lead to the saying, 'what a small-world it is,' are quantifiable. Through the use of graph theory, scientists and engineers have been able to take advantage of the tendency of a small-world to have a nonlinear relationship between the characteristic length of a network and the number of vertices within that network. Critical to this research is the consideration that these long-range edges, crossing an organization, can lead to very short distances needed to cross an organization efficiently. Particularly if the searcher knows the global location of the information desired and with no social construct inhibiting communication, a network such as this can be very quickly searched using a greedy algorithm. Even if the location of the knowledge one is searching for is not known, quick search routines already exist if the network has a power law distribution like the world-wide web. Further, an organization with social links across sub-units allows for a more efficient organization in times of crises such as quick taskers or wartime material support. These findings indicate that an organization that does not allow social contacts to grow over time between sub-units will not operate as efficiently as one that does allow these links to form. Lastly, even when not in crises, progression through an organizational community will create a small-world network within that community, but any searches that would lead outside the community would not benefit from the decreased characteristic length characteristic of a small-world.

III. Methodology

The primary intention of this research is to provide additional understanding of the effects of organizational structure on an organization's social network. While it is assumed an organization's social network enables quick, trusted communications on a regular basis, particularly important to this research is communications between personnel in times of crises. Further, the ability of an individual person to rapidly search for information within the organization during crises is considered paramount. The approach used for this research is to model an organization as a large hierarchical network of personnel. Primary analysis of the network is accomplished by considering the network as one connected graph where individual personnel are represented as vertices and the social (and organizational) ties between them are represented as the edges.

Our methodology for modeling and analyzing the organizational and social relationships within an organization consists of three main parts: (*a*) an algorithm for constructing a graph representative of an organization, (*b*) an algorithm for maturing that graph, and (*c*) several techniques for measuring the results of the preceding algorithms. The initial algorithm, as detailed below, follows the initial network creation algorithm detailed by Dodds, et al. The network created by this algorithm then provides the basic framework, or substrate, for the growth of organizational social connections (Dodds, 2003). After the organizational network is created, the network maturing algorithm is run. This algorithm models the maturation of an organization over a pre-selected number of years by removing personnel (represented by nodes or vertices in the organizational network) and either bringing another node on the same level into the position, promoting a node into the vacated position, or bringing in a new node if the vacancy is on the bottom level. These removals model firings and/or separations using pre-selected annual attrition rates. In addition to promotions, the algorithm selects a series of internodal collaborations which increase the strength of each link within the network. These collaborations model typical interdepartmental or cross-organizational efforts such as tiger teams or cross-organizational integrated product teams. Lastly, the measurement portion runs several routines that analyze the graph created representing the organizational network.

Organizational Construction Algorithm

The following algorithm takes advantage of the first part of the network construction algorithm laid out by Dodds, et al (Dodds, 2003). The initial inputs required to fulfill this portion of the organizational construct are the number of levels desired (*Lvl*) and the branching ratio for the organization (*b*). The branching ratio for the organization is defined as the number of subordinates directly reporting to each superior. These inputs combine to create a graph with $V(b,Lvl)=(b^{Lvl}-1)/(b-1)$ vertices and $E(b,Lvl)=V(b,Lvl)-1$ edges. After these inputs are made, the algorithm starts by creating a tree graph representing a strict hierarchical organization. An image of this graph can be seen in Figure 6. From this figure, the strict hierarchical structure the entire organization is built from is displayed. Further, the level count and the branching ratio can be observed from this figure.

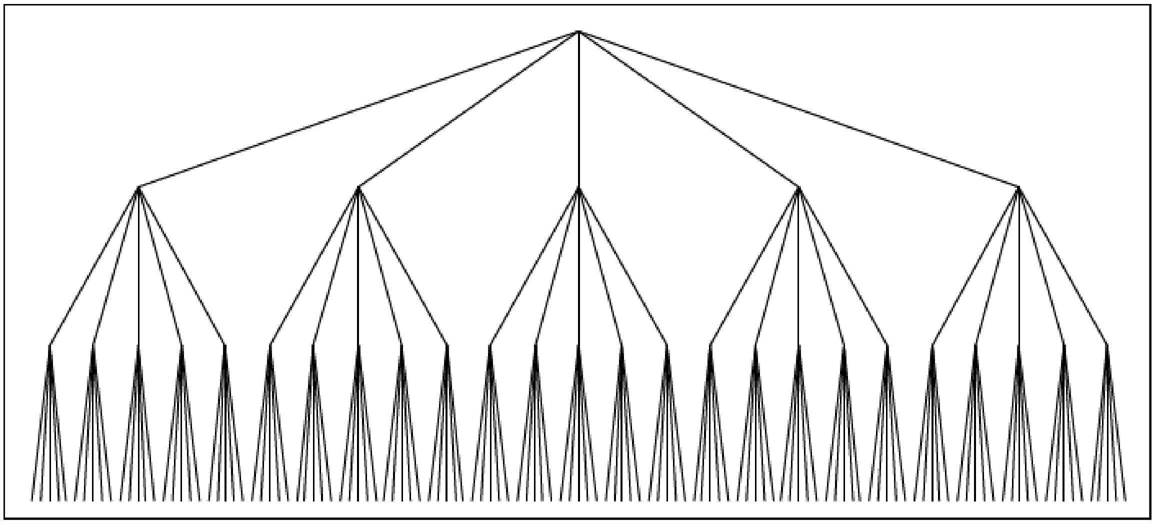

Figure 6 A Visual Display of an Organizational Network Graph (*Lvl=4, b=5*)

After the initial tree network is created, the initial social connections are added. Following the research of Krackhardt, et al, these initial social connections are drawn between members of a subunit (Krackhardt, 1988). Following this research, members of the same subunit are defined as having the same immediate superior. After adding the initial social connections, the graph will now still have the same number of vertices, but now the number of edges will increase to $E(b, Lvl) = (b^{Lvl+1} + b^{L} - b^{2} - b) / 2(b-1)$. Figure 7 displays the final organizational network after these social connections are made.

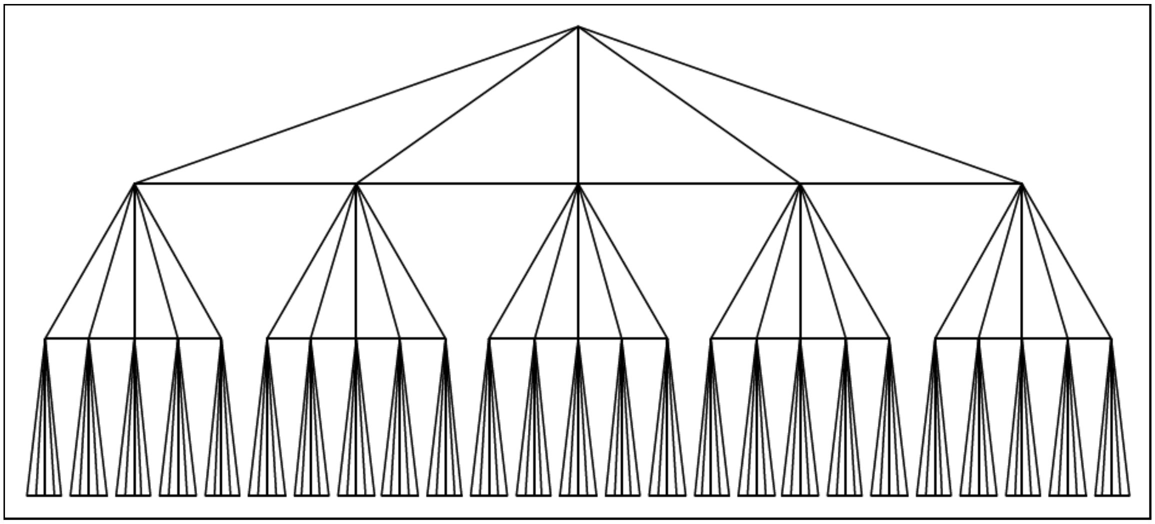

Figure 7 An Organizational Network Before Maturation (*Lvl=4, b=5*).
The organizational hierarchy is the base tree graph and the subunit connections are the horizontal lines.

Organizational Maturation Algorithm

Now that the organizational construction algorithm has built the organizational network, the maturing algorithm will begin. This algorithm has three main components. The first removes nodes and replaces those vacancies by either moving another node on the same level laterally, promoting a node from a lower level into that position, or bringing in a new node if the vacancy is on the lowest level. The second part of the algorithm sends a series of collaborations through the network. Finally, the third component actually ages the links between any two nodes within the network. This is accomplished by changing the weight of the edge representing the relationship within the graph representing the organizational and social ties of the organization.

Attrition Modelling.

Attrition modeling is accomplished through a node removal and replacement algorithm, representing employee removal and replacement, which begins after the initial graph has been created by the organizational construction algorithm. The inputs for this algorithm are, in addition to the initial graph, the desired annual attrition rate, the desired partitioning level, and the number of years the simulation should run for. The purpose of the partitioning level is to specify the various communities organizational personnel will be restricted to throughout a career. Further, it helps to define a sub-community, which, for this model, are members of the same community within a partitioned organization that report to the same supervisor. After choosing the partitioning level, each node that level becomes a community leader. Due to the symmetric structure of the tree, the number of communities formed below this level will be $C=b^{L-1}$ where C is the number of communities, b is the branching ratio, and L is the partitioning level. The level numbering convention places l_1 at the level for the top node and flowing down to l_n at the bottom of the tree with $l_1, l_2, ..., l_n \in Lvl(G)$. No personnel movement is permitted between communities. Due to this inability to move between communities, these communities become highly connected over the course of the simulation. This simulates an organization that would not permit movement between major subdivisions and the resultant social connections. In this graph, the nodes on the partitioning level would represent vice presidents or ASC Wing Commanders with responsibility for the progression of personnel within their subdivision or Wing. The next input is the annual attrition rate, which is simulated in the graph as the node removal rate for the organization for each time step. Finally, the simulation run defines the total years of simulation maturation.

Once the organizational network has been created and the inputs have been made, the maturing algorithm will begin. This algorithm initially chooses which vertices to remove from the graph. The number of vertices to remove with each time step (*NR(G)*) is found by multiplying the number of vertices (*N(G)*) in the graph by the attrition rate $NR(G)=N(G)*AttRate$. Then a list of the full set of vertices is randomly permuted and the number of vertices desired to remove is chosen from this list. In this removal process, the assumption is made that very few vertices will be removed twice in the same step. As a result of this assumption, the removals are made without replacement and the probability of any node being removed is $P_{removal}=NR(G)/N(G)$. Once the procedure has chosen the number of vertices to remove from the graph in the time step, it begins looping through this list, simulating removal of personnel.

Once the initial vertex to remove is chosen, the vertex is found in the adjacency matrix. The algorithm does not remove the vertex in the traditional sense (i.e. remove that column and row from the adjacency matrix) as this would ruin the cross-linked tree substrate, potentially creating a disconnected graph. Rather, a vertex is removed from this graph by setting all entries in the vertices row and column in the adjacency matrix to zero. As in a traditional organization, when one person is removed, either another is laterally assigned over from the same level, there is a promotion from within, or somebody is hired from outside. An assumption made for this model is that personnel would only be brought into an organization at the lowest level since the model is simulating a government employee construct where personnel primarily enter at the lowest levels. Within the algorithm, the probability of a lateral is $P_{lateral}=0.5$. Since the only other option to fill the position, when that position is not on the bottom level, is to promote a person from below, the probability of a promotion is also $P_{promotion}=0.5$. In either case, the partitioning of the graph becomes critical in the consideration of which node to either move laterally or promote since a promotion or lateral can only come from within the same community as the employee that was removed. By requiring the promoted or the laterally moved employee to come from the same community, the eligible pool to fill the position is reduced. Further, a lateral is only permitted by a vertex on the same level and a promotion is only permitted from one level below the removed vertex.

Once a node, representing an employee, is chosen to move, the movement is accomplished within the algorithm by replacing the zero row and column by the mover's row and column within the adjacency matrix. Additionally, the subunit social links for the new position are added to the adjacency matrix. Lastly, the row and column in the adjacency matrix for the moved vertex, representing the employee's prior position, are set to zero. Once a vertex is moved, a new vacancy in the graph, representing a new open position in the organization, is created and the process repeats itself until the node removed is from the bottom row. In this case consideration is given to a lateral move, and, if a lateral move is not chosen, only non-subunit edges are removed and the loop moves to the next vertex in the removal list. This simulates a new hire that will have an initial relationship with his/her coworkers and superior after the first time step. This process continues to repeat until all nodes in the list of those to be removed in the time step have been removed and those positions have been filled. Then the algorithm goes on to the next time step, where a new set of removals and promotions are made. Node removal and replacement is accomplished 12 times per a simulated year, representing the 12 months in a year.

Two examples of the output of these algorithms can be seen in Figure 8 and Figure 9. The plots display the results on cross corporation connections due to partitioning the graph at separate levels, creating different organizational, and social, structures when no long-distance collaborations are considered.

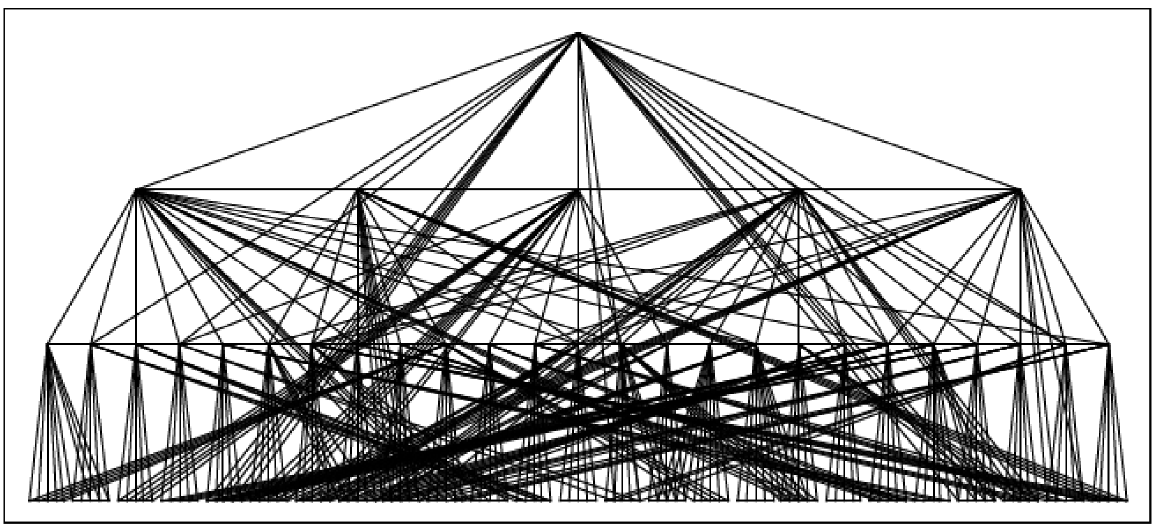

Figure 8 A Display of the Effects of Maturation on a Graph with no Partitions.
(Lvl=4, b=5, PartLevel=1, AttRate=0.1, RunYrs=50)

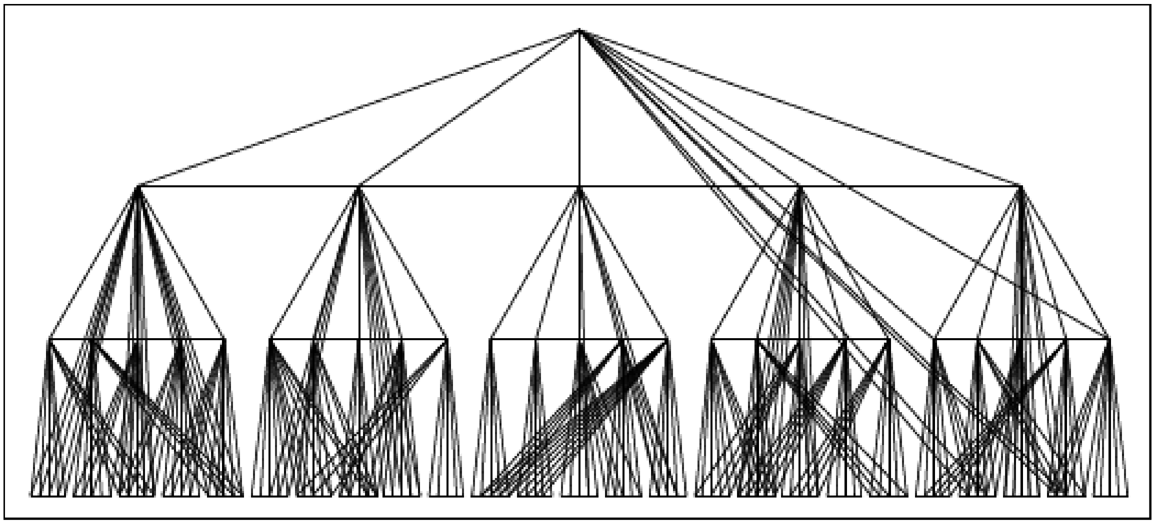

Figure 9 A Display of the Effects of Maturation on a Graph with Partitions.
The partition is drawn at level 2, creating 5 communities.
(Lvl=4, b=5, PartLevel=2, AttRate=0.1, RunYrs=50)

Maturation of Social Ties

While the removal and replacement algorithm, simulating normal annual attrition and movement to fill vacancies in an organization, focuses on moving nodes, representing personnel within the graph, another algorithm is running to change the weighting, or strength, of each edge, or social tie, in the graph. Changing the weight of each edge models the strength of any relationship which is assumed to have a linear relationship with the probability of the use of any edge within the graph for communication. The possible measures for any edge weight in this model vary from 0.0, representing no link is formed, to 1.0, representing a strong link. Further, the assumption is made that the strength of any social connection within our model is affected primarily by the time since two individuals served in the same subunit as well as the amount of repetitive collaborations incurred in a time step between the two individuals.

The first assumption, which is that the interpersonal connections will strengthen with time when personnel are in the same sub-unit (and weaken otherwise), is modeled using a logarithmic function. Similar to a learning curve, the basic formula for the strength of each link is $y(t)=1-\eta e^{-\gamma t}$ where y represents the strength, t is time and α and γ are constants. The initial condition for any edge in the graph is $y(0)=0$, so $\eta=1$. After this initial constant is found, the model solves for γ based on an initial input. This input is the desired strength of the connection after six months for the aging algorithm. For the purpose of this research, the desired relational strength will be assumed to be 0.8 after six months of growing in strength. This input, along with the assumption of 50 work weeks per year, is used to solve for γ. For this model, since y(25weeks)=0.8, the γ value would be $\gamma=-ln(0.2)/25=-0.0644$. The strength of each relationship is changed by adjusting the edge value within the adjacency matrix with every time step. With each increment, the strength is changed by *50/#Increments* (assuming the 50 work week year). If v_i and v_j share the same sub-community, the change is positive, reflecting a strengthening of the connection. If they are not in the same sub-community, the change is negative, reflecting a decrease in the strength of the connection. The importance of this consideration is that, while the personnel are working together on a regular basis, the connection is strengthening, but, once they are no longer working together as often, the connection is weakened. A derivation of the aging calculations can be found in Appendix A.

Collaborations

The other way the edge weight is changed in the algorithm is by repetitive long-distance collaborations. Similar to the assumption that those who work together on a regular basis will have a stronger connection, this one is based of the assumption that personnel will have an increased connection to other personnel they collaborate with on temporary large projects. Although these collaborations can vary in intensity from working together on tiger teams to planning a complex social event, the primary consideration is that personnel interact heavily on these efforts. This consideration is modeled by initially choosing several pairs of vertices that collaborate throughout the year. This set of pairs represents interpersonal collaborations throughout the time step. These vertex pair choices are not made at random; rather, there is a higher probability of a link based on how close one vertex is to another within the organizational graph. Following the assumptions of Dodds, the probability of communication between any two nodes is:

$$P(i, j) \ \alpha \ e^{-D_{ij}/\lambda} e^{-x_{ij}/\delta} \qquad (3)$$

where D_{ij} is the depth of the lowest common superior between the vertices, x_{ij} is the organizational distance between the vertices, and λ, δ, and ζ are normalizing constants. For this research, Equation 3 will be modified for application to a military construct.

Dodds makes the assumption that the amount of communication between two vertices is relative to the depth of these vertices (Dodds, 2003). This assumption is not necessarily relevant to a military construct, where it is just as likely for two colonels to communicate as it is for two lieutenants. Due to this consideration, D_{ij}/λ will be set to zero to prevent this from scaling the probability function. Additionally, in order to use this equation as a function instead of a relativity, another variable, labeled ζ, will be added to scale the function. Lastly, in order to prevent a probability of communication greater than 100%, the minimum of one and the value found will be used. After making these changes, the new probability function, representing the probability of communication between two employees if they are chosen to communicate will be:

$$P(i, j) = \min(1, \zeta e^{-x_{ij}/\delta}) \qquad (4)$$

where x is the organizational distance and δ and ζ are normalizing constants. The tuning of the organizational constants, δ and ζ, will be accomplished in the analysis and results section using a discrete probability function. The organizational distance is based on the relative levels of the two vertices. The equation for the organizational distance is:

$$x(i, j) = \sqrt{d_i^{\,2} + d_j^{\,2} - 2} \qquad (5)$$

where the d_i and d_j are the differences of the depths of these vertices below the depth of their lowest common superior (Dodds, 2003). Once the probability of communications is found, the model will assign a preselected number of collaborations to the edges based on this probability. Then, for each time step, the weight will be incremented by a weight equivalent to three weeks of working together for each collaboration throughout the time step. This simulates the strengthening of each connection based on the number of temporary projects throughout the year. Further, this takes into consideration the assumption that those that work more closely will collaborate more often throughout any time-step. Lastly, the increment for strengthening is always positive. This is due to the assumption that any collaboration can only increase the connection strength between any two people. As a result of these assumptions, the model could randomly choose or even add an edge to the graph that would cross the graph, but if this connection is not used on a regular basis, the age of the edge will increase, weakening the connection until the edge no longer exists. At this point, the graph will be fully aged and the analysis of the model can begin.

Organizational Network Analysis

After the organizational graph has been created and aged by the previous algorithms, the model begins analysis of the resulting social and organizational network. Central to this research is the hypothesis that the number of edges that must be crossed, or search distance between two vertices, v_i and v_j, increases for a search between any two vertices in separate communities. Conversely, the search distance between two vertices, v_i and v_j, from the same community will be shortened in a partitioned organization in comparison to one with no partitions, as there are fewer vertices to search through in the smaller community. Further, we hypothesize that the increase in efficiency for searches within the smaller community is smaller than the decrease in efficiency in searches between communities. Finally, the analysis must demonstrate if there is a way the corporate entity can show if there would be advantages to partitioning an organization based on the amount of inter- vice intra-community communication.

These hypotheses flow from the assumption that the organization will grow from its highly hierarchical network to a small-world network after the aging process. This assumption is based on the consideration that, as can be noted in Figure 8 and Figure 9, when vertices are permitted to traverse the entire network, long distance edges are drawn, whereas, when partitioned, these edges are not drawn based on working in the same community. These long distance edges, creating a weaker small world network throughout the organization, are provided by long-distance collaborations instead of working together in a sub-unit in the case of a partitioned organization. As Watts noted, these long distance edges have a non-linear effect on the characteristic path length of the graph (Watts, 1998). Due to this phenomenon, the intra-community distances will drop due to partitioning, and the inter-community distances will not drop as quickly or as far as they would in a non-partitioned organization due to the lack of the cross organization edges, assuming this is a small-world network, especially when the strength of the relationship is considered. Therefore, the analysis must initially show that the organization, after the aging process, follows the small world tendency of a decreasing characteristic path length (considering various edge weights representing relational strengths) with aging.

In order to show this decrease in characteristic path length with age, representing an optimum average distance between any two vertices in the graph, the characteristic path length must be measured after each simulated year. The first measure is the characteristic path length of the graph prior to aging to form a baseline for additional measurements. Once this value has been found for the non-aged graph, the model will record the characteristic path length for each time-step. This procedure will record the change in the graph as edges are added and removed through the aging process.

Once it has been shown that the characteristic path lengths decrease with age, due to the small-world phenomenon, a series of searches will be made. The purpose of these searches is to find the length between any two vertices, v_i and v_j, when the starting vertex does not know the location in the graph of the ending vertex. This models the situation where a person needs information, but does not know who has that information.

A new search heuristic for an organization with a strong reporting chain, such as would be seen in many traditional corporate structures as well as the military is used for this analysis. This situation is different than the random walk search and the degree distribution searches proposed by Adamic et al, which where more appropriate to a network such as the internet where there is neither a social strata nor organizational boundaries to prevent vertical communication through the network (Adamic, 2001). Critical to this type of organization is that a person on the bottom level of the organization is not going to go directly to the Wing Commander with a question. Further, the Wing Commander is not going to go directly to somebody on the bottom level, rather, he or she would go through Squadron Commanders, then mid-level managers until finally arriving at the person with the information. As a result of this construct, the search would begin at a chosen initial vertex and check the adjacency matrix for a connection to the terminal vertex. If the terminal matrix is listed as a connection with appropriate edge strength, then the search completes with one edge crossed (since there is still a crossing between two individuals). If there is no connection or the edge is too weak, then, following the reporting chain construct, the individual would ask his or her superior and the superior would ask his or her colleagues. This is modeled by changing the initial node to the superior, incrementing an edge counter, and repeating the search through the adjacency matrix. This progression up through the network continues until either there is a direct connection to the terminal node or there is a connection to somebody in the reporting chain of the terminal node. At this point, the search would go down and result in incrementing the edge counter by the difference in the depth of the terminal node and the person on the reporting chain contacted (since, as part of the assumption, this superior will not go directly to the subordinate, rather, that person would go down through the reporting chain).

In order to prove these hypotheses, the choice of starting and finishing nodes as well as the partitioning construct of the graph is critical. To show the first hypothesis, searches must be accomplished between partitioned communities and within the same community. Further, this set of searches must be compared against the same set of searches within a non-partitioned organization. In order to accomplish this, an organization will be grown according to the previous organizational graph aging procedure. At this point, two sets of several search points, v_i and v_j, will be chosen. The first of these two sets will include random pairs that, even after partitioning, would be in the same community. The second set will include random pairs that will always be of different communities regardless of organizational structure. Once these two sets are chosen, a series of searches will be accomplished on the two different sets of pairs. These searches will represent the two extremes (exclusively within the same community and exclusively within separate communities) for searches prior to partitioning the graph into separate communities. Now the aging algorithm will be run again, but this time with the aged organization as an input. Further, the new aging routine will require at least one partition of the organization. While the aging routine is running, the search routine will run at the end of each year increment on the same two sets of pairs. In accomplishing this, the model will reflect the changes in the search distance resultant from the partitioning of a mature organization. Next, the previous procedure will be run in reverse, by starting with an aged partitioned organization as the initial input and flowing to a non-partitioned organization. This simulation will show the result of allowing free flow through an organization when this was not previously allowed.

Following these procedures, an understanding of the social connections of a maturing organization can be gained. These social connections become very useful to an organization, particularly in times of crises, when an employee must search quickly throughout the organization to find information. When this happens, the social connections formed over a career enable that employee searching for information to quickly share knowledge. By partitioning an organization into several smaller communities, the corporation enhances short ranged connections, often at the expense of long ranged ones. This combination of algorithms measures these changes in the organization as the organization matures under the several different partitioning constructs.

IV. Analysis and Results

Introduction

The primary consideration of this analysis is to consider the effects of an organization's structure on that organization's social network. Based on sociological and psychological research, it has been found that this network plays an important role in an organization's ability to respond to crises. Further, this social network influences an employee's ability to search for and find information within the organization. In addition to the consideration of the organizational structure, consideration will be given to the effects of the attrition rate and the number of collaborations, discussed in the methodology, under both organizational constructs.

Two primary measurements of the effectiveness of an organization's social network will be considered. The first measure will be the characteristic path length. As stated in the literature review, the characteristic path length is a measure of the average shortest path length averaged over all pairs of vertices within the graph. This represents an optimistic view of the average shortest distance between any two employees in an organization. Additionally, following the small-world construct, it is assumed that the characteristic path length will decrease non-linearly with maturation of the graph representing an organization's organizational and social networks. The next measure which will be used in the analysis of the organizational and social network is an edge search length based on a reporting chain heuristic. This measure will return a realistic distance between pairs of employees when one employees needs to search for information throughout the organization. As a result of this realism, it is expected that this measure will tend to be longer than the characteristic path length.

Consideration of the Distribution of Collaborations

As stated in the methodology, throughout the simulation, a series of connections will be made across the graph simulating collaborations. These connections can be assumed to simulate any effort where two people interact regularly. The mechanics of implementing these collaborations will begin with a random number generator choosing a pair of vertices from within the graph. The probability of communication for these two vertices, based on Equation 4 from the methodology, and restated below

$$P(i, j) = \min(1, \zeta e^{-x_{ij}/\delta})$$

will then be compared to another random number in order to decide whether the collaboration based edge will be created

In order to compute the probability though, the parameters, δ and ζ must be considered first. Initially, while ζ is assumed to be equal to one, a choice will be made for δ reflecting observations of collaborations within ASC. After this value is decided upon, ζ will be varied to address scaling the probability. By scaling these values, the proportion of long-distance to short-distance collaborations will be adjusted. Following observations, it is assumed that the majority of projects one will work on will be with those individuals he or she is relatively close to within the organization. In order to consider the effects of varying the value of δ on the graph, it is necessary to compare the probability of a collaboration happening at a certain distance in the graph and the probability of a collaboration pair being chosen based on that organizational distance. This consideration can be made on a level-by-level basis using a discrete probability distribution function based on the following equation:

$$q_i = \frac{\# d_i}{\sum d} P(d_i) \tag{6}$$

where q_i represents the expected number of pairs at distance d_i that would have communicated if given the opportunity, $\#d_i$ is the number of pairs of organizational distance d_i, Σd is the total number of pairs of vertices at all distances, and $P(d_i)$ is the result computed in Equation 4. A q-value will be found for each distinct distance d_i and then normalized by the sum of the q's returning the equation used to create a discrete probability density function

$$PDF(d_i) = \frac{q_i}{\Sigma q} \tag{7}$$

where, as before, the d_i represents the distinct organizational distance, q_i represents the expected number of pairs at distance d_i that would have communicated if given the opportunity and Σq represents the sum of the q_i's. This combined probability takes into account the relative geometric placement of the vertices as well the number of vertices within the graph. A δ value can then be chosen to preclude choosing a pair of low probability vertices (because they are geometrically distant) due to the large number of low probability vertices (because many of the vertices are geometrically distant). A choice of $\delta = 2/3$ has been made for this study while ζ is chosen as one since it will be considered later. These choices result in the discrete probability distribution function plot found in Figure 10 for the distribution of those members of the bottom two levels of the organization.

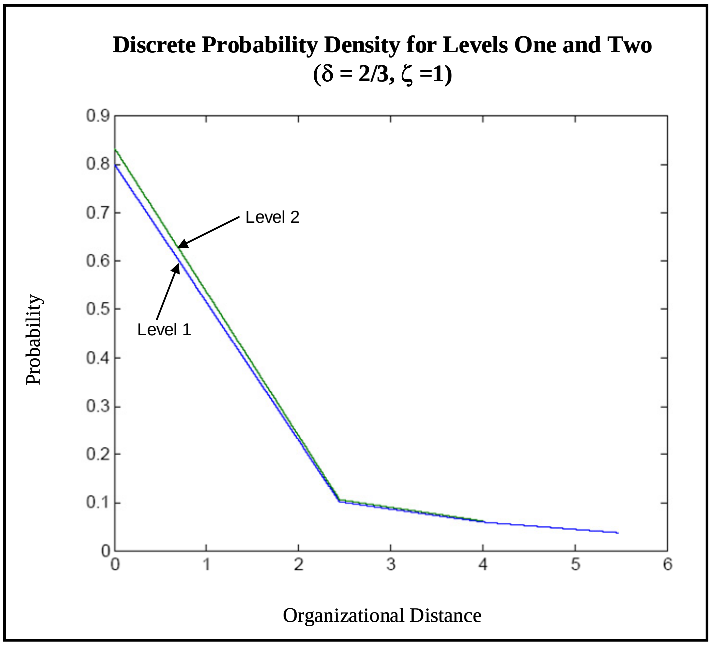

Figure 10 Probability Distribution of Collaborations.

Figure 10 reveals that, when the δ value is placed at 2/3, approximately 80% of collaborations will occur within the same sub-community when personnel are on the lowest level whereas slightly more than 80% of the collaborations for members on level two will be within the same sub-community. This proportion of collaborations drops to only a 10% with members at the same level just outside the sub-community. It is assumed these values are representative for a military construct, therefore we chose $\delta \equiv 2/3$.

Since the value of δ has been chosen, the next consideration is the effect of ζ on the distribution of collaborations. In addition to analyzing this value using a probability distribution function, as in Figure 10, the effect of the distribution of collaborations on the characteristic path length will be evaluated. Due to the normalization of the *q*-values during the calculation of the discrete probability function in the previous analysis, varying the value of ζ between just greater than zero and one will have no effect on the resultant distribution of collaborations. Since these small values will have no effect on the distribution of collaborations, they will also have no effect on the characteristic path length. Larger values of ζ will affect the distribution of collaborations, though. This effect can be seen in Figure 11, displaying the increasing preference for longer distance edges as the value for ζ increases logarithmically.

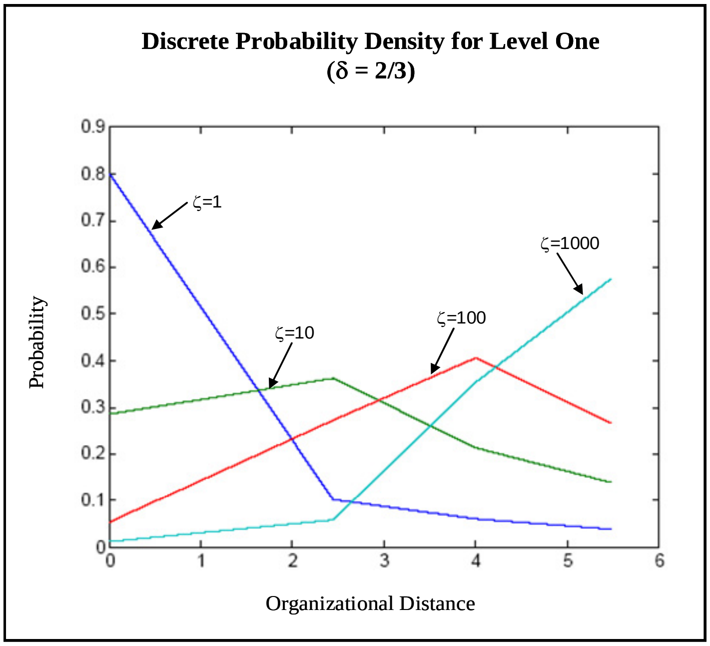

Figure 11 Probability Distribution of Edge Creation Due to Organizational Distance.

Figure 11 reveals that large ζ's will make a difference to the proportion of long-distance versus short-distance edges, representing differences in an organizations preference for long-distance versus short-distance collaborations. Further, an increased proportion of long-distance edges can be expected to decrease the characteristic path length of the graph due to the small-world effect.

As predicted previously and seen in Figure 12 for a one community and a six community construct, due to the small-world effects, there is a quick drop in the graph's characteristic path length with very large values of ζ as well as disparity in the characteristic path lengths resultant from different ζ's. In addition to the variance of the characteristic path length due to the ζ-value, the steep decrease and leveling off of the characteristic path length with aging, due to long-distance edges formed from collaborations, follows the definition of a small-world. This resemblance to a small-world is the reason for such a small difference between the characteristic path length of the one large community graph and that of several smaller communities. Although there is a difference in the results between the two organizational constructs toward the end of the simulation for the case of small values for ζ, these results are very similar, reflecting the non-linear effect of a handful of long-distance edges within the graph on the graph's characteristic path length. Further evidence of this consideration can be seen in the characteristic path lengths resulting from larger ζ values, reflecting a larger probability for the simulation to choose long-distance collaborations.

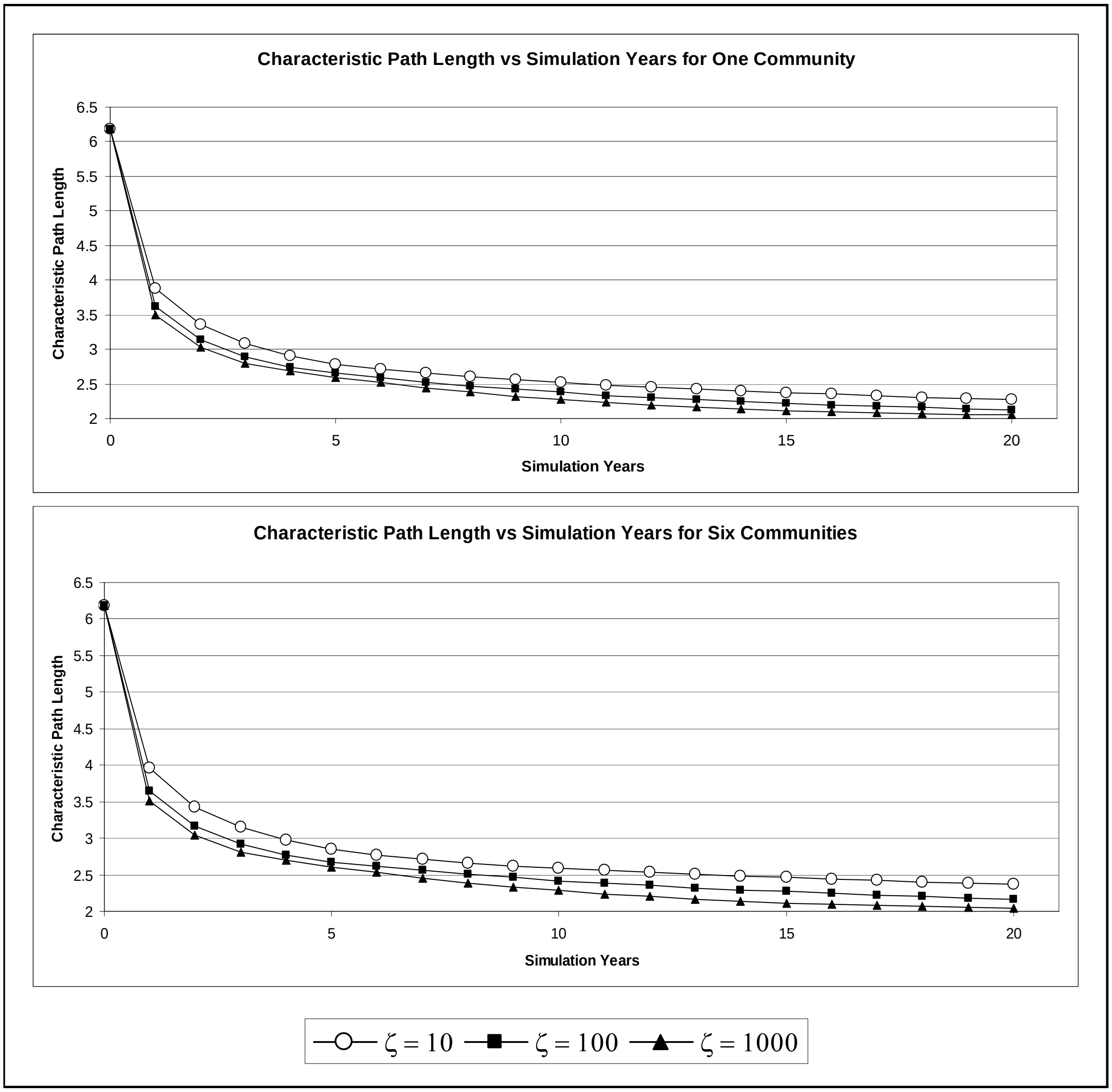

Figure 12 Change in Characteristic Path Length Due to Different ζ Values.

As shown above, the ratio of short-distance versus long-distance relationships resultant from collaborations will affect the characteristic path length of the graph representing the combined organizational and social network for an organization. Therefore the scaling parameter, ζ, should be chosen such that the distribution of ties is representative of the organization being modeled. For purposes of the remainder of this analysis, a value of $\delta \equiv 2/3$ was chosen, which as discussed above and shown in shown in Figure 10, results in approximately 80% of social ties being formed between personnel within the same sub-community. Further, a value of $\zeta \equiv 1$ was chosen in the remainder of this research to prevent changing the proportion of long-distance versus short-distance collaborations.

Characteristic Path Length Considerations

The characteristic path length of a graph is a fairly typical measure and is used in the analysis of many models. Typically, in such models, all edges, which in this case represent relationships within an organization, are assumed to be of equal value. However, as discussed in the methodology, the model developed here incorporates the strengthening of relationships when personnel are in the same sub-unit and the weakening of relationships otherwise. As a result of this consideration a variant of the classical definition of the characteristic path length was developed where weaker edges will be removed from the graph prior to finding the characteristic path length.

Of particular interest to a decision maker would be the actual strength of the edges within the graph. Due to the strengthening and weakening previously discussed, these varying edge strengths would reflect the strength of bonds between coworkers where zero would represent no knowledge of one another and one would represent a very strong relationship. Assuming a linear relationship between the strength of a relationship and the probability of communicating for information, a stronger relationship would have a greater probability of being used than a weaker one. In order to test for the weights of these graphs, three values, labeled α, were chosen as cutoff values for the graph. The set of α's chosen were $\alpha=\{0.0,0.1,0.2\}$. Any values within the final adjacency matrix below the α value chosen was then removed prior to finding the characteristic path length. This simulates removal of weaker links from the social network prior to analyzing the network. A physical interpretation of these strengths is as follows:

i. $\alpha = 0.0$ Strength of an unformed relationship

ii. $\alpha = 0.1$ Strength of a 1.6 week old relationship

iii. $\alpha = 0.2$ Strength of a 3.5 week old relationship

The values of 1.6 weeks and 3.5 weeks represent the time a relationship has developed due to working together prior to the time that relationship is removed from the social network. Therefore, the α-values corresponding to these values represent the strength of a relationship that has developed for that amount of time prior to removal from the network. The removal of these lower-strength edges, representing short-term relationships or relationships that formed long ago and have since aged, limited the characteristic path length calculation only to those edges with increased probabilities of use. The value of α value of 0.0 represents an unformed relationship. Since no relationship is weaker than an unformed relationship, when $\alpha=0.0$, no edges are removed from the graph. The resulting plots of these varying α values can be seen in Figure 13 for both one large community and six smaller communities.

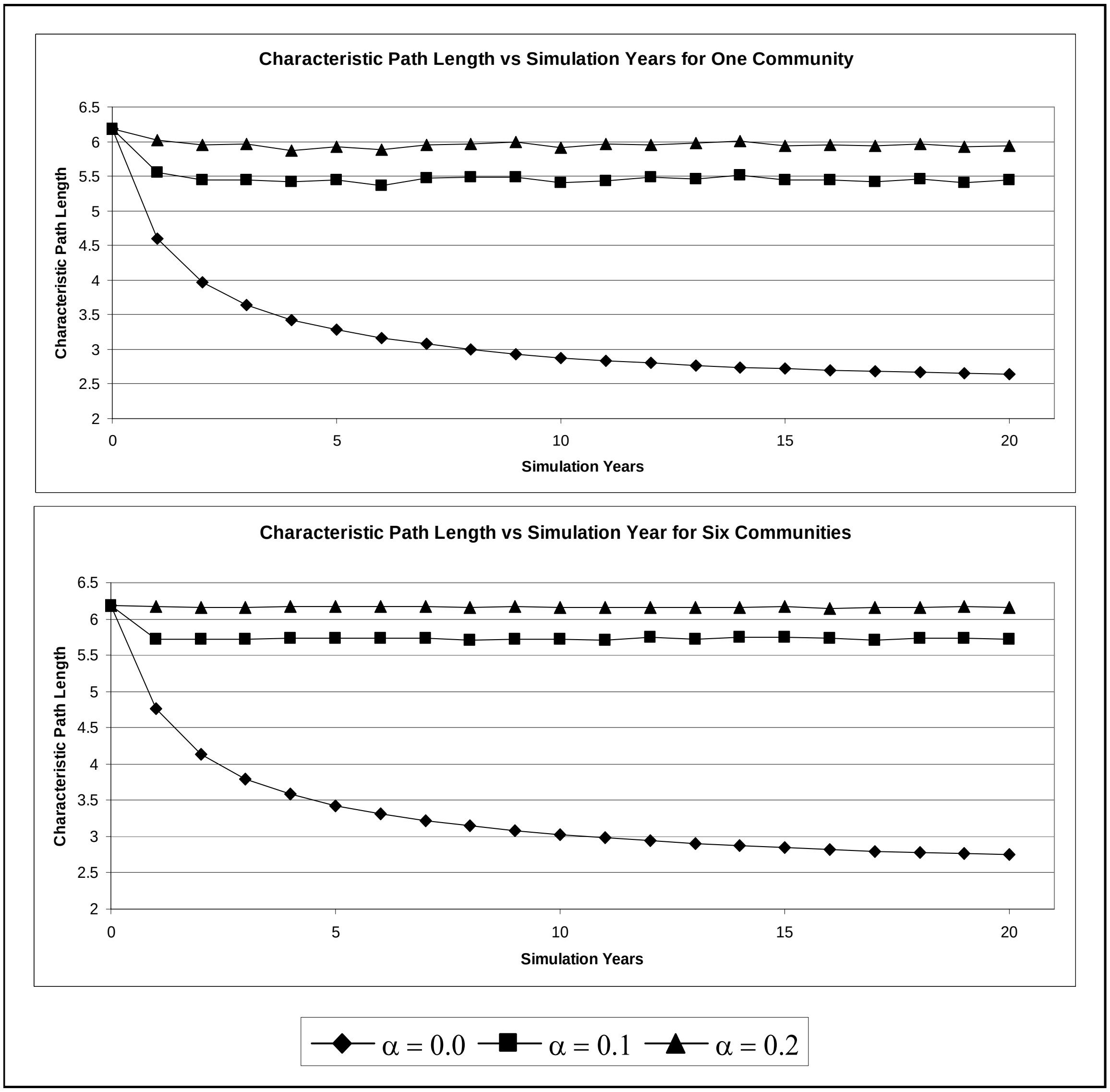

Figure 13 Change in Characteristic Path Length Due to Different α Values. Decrease in the characteristic path length for an organization with six communities. Note that α = 0.1 corresponds to 1.6 weeks of working together and α = 0.2 corresponds to 3.5 weeks of working together, reflecting the weakness of the edges in the graph.

Figure 13 reveals that the primary cross graph links are relatively weak for both the partitioned and non-partitioned organization. As can be seen, when the α is set to 0.2, equivalent to 3.5 weeks of working together, there is little effect on the characteristic path length of the graph due to long distance edges. Of note is the consideration that all collaborations increase the weight of any chosen edge an additional 3 weeks. As a result, when α=0.2 it will remove any long-distance edge resultant from only one collaboration in the previous time step. Further, the difference between the one community scheme and the six communities' scheme is still marginal. The combination of these results follows the small-world construct since the construct relies heavily on cross-graph links. The marginal difference in edge strengths, favoring the one community construct, are due to these cross-organizational movements, which are permitted when there are no partitions throughout the 20 year timeframe. Therefore, as a result of the small-world reality, the long-distance links decrease the potential characteristic path length considerably.

Effects of Varying the Number of Collaborations.

In order to further consider the effects of long-distance edges on the graph, the number of collaborative efforts was varied. In the preceding comparisons, the number of long-distance collaborations each pair of vertices would experience throughout the time step was approximately four per a year. By varying this number of collaborations, the number of the long-distance edges was varied accordingly. Figure 14 reveals the results of varying the number of collaborations in a time step for both the consideration of one large community as well as six smaller communities.

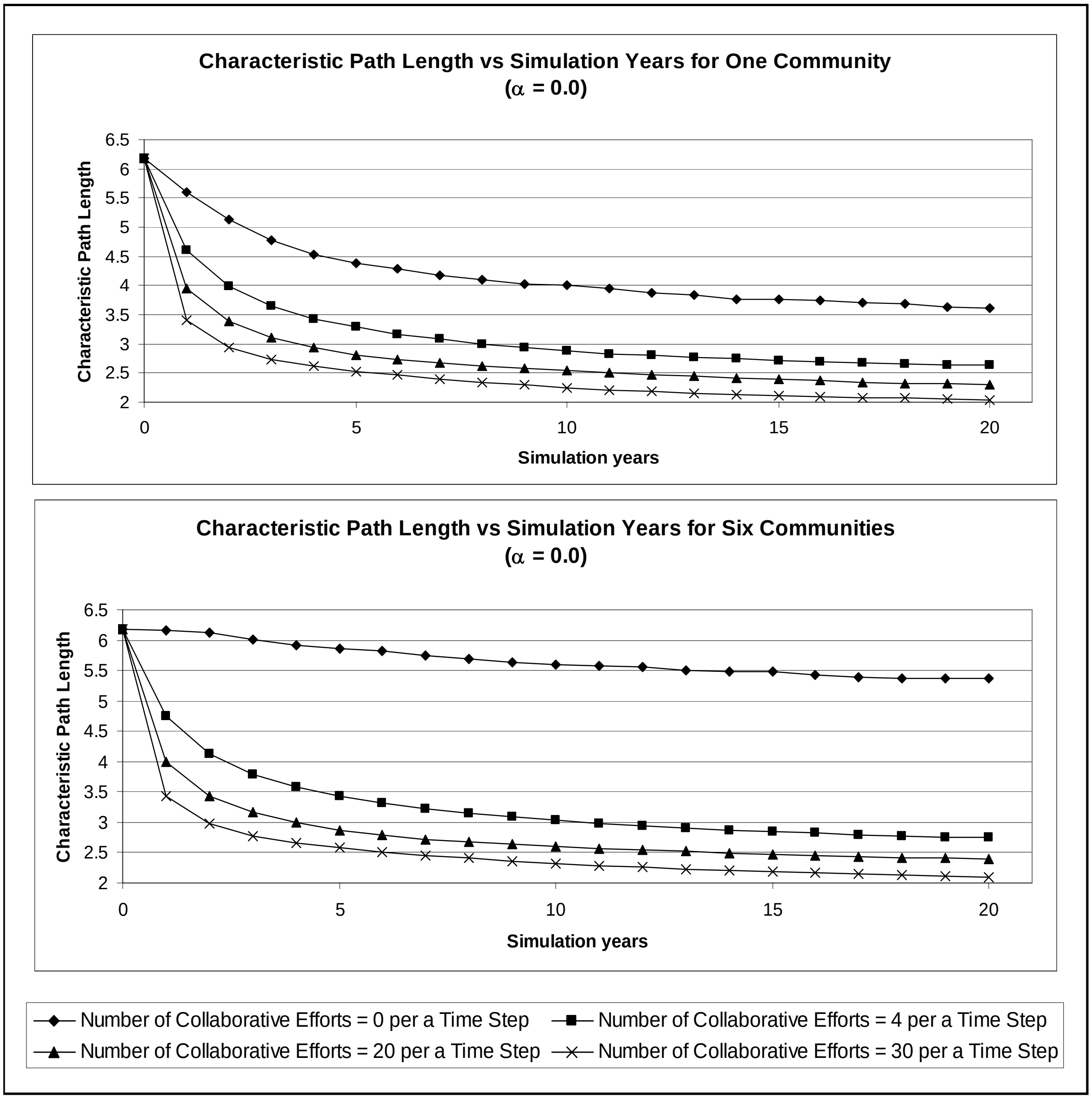

Figure 14 Characteristic Path Length Changes with Long-Range Collaborations.

These plots reveal the importance of weak links that form across the graph, to the organizations characteristic path length. The most obvious display of the importance of the importance of the characteristic path length can be seen in the top two lines in the two plots in Figure 14. Due to the ability for personnel to move throughout the organization, the characteristic path length drops to below 60% of the characteristic path length of a non-aged graph when no long-distance collaborations occur when considering the one community construct. A great disparity exists between this consideration and the results of an organization with six communities. In the case of six communities, the organization is more reliant on these collaborations in order to decrease the characteristic path length. As a result of this reliance, the characteristic path length only drops minimally with time when there are no long-distance collaborations as seen in the situation of zero collaborative efforts in the six community construct.

Of interest is the impact of removing some of the weaker edges, reflecting the removal of weaker relationships from the graph. In Figure 15, edges with a value less than 0.1, representing the strength of relationships that are 1.6 weeks old, were removed from the graph. The result on the characteristic path length of the graph, representing the most optimistic shortest path across an organization differed only marginally depending on the type of organization. There is a slight difference, though, between the one and the six community construct. This difference reflects stronger edges in the one community organization due to a reliance on edges formed from co-working in subunits where these edges are strengthened repeatedly and then moving throughout the organization instead of collaborative efforts.

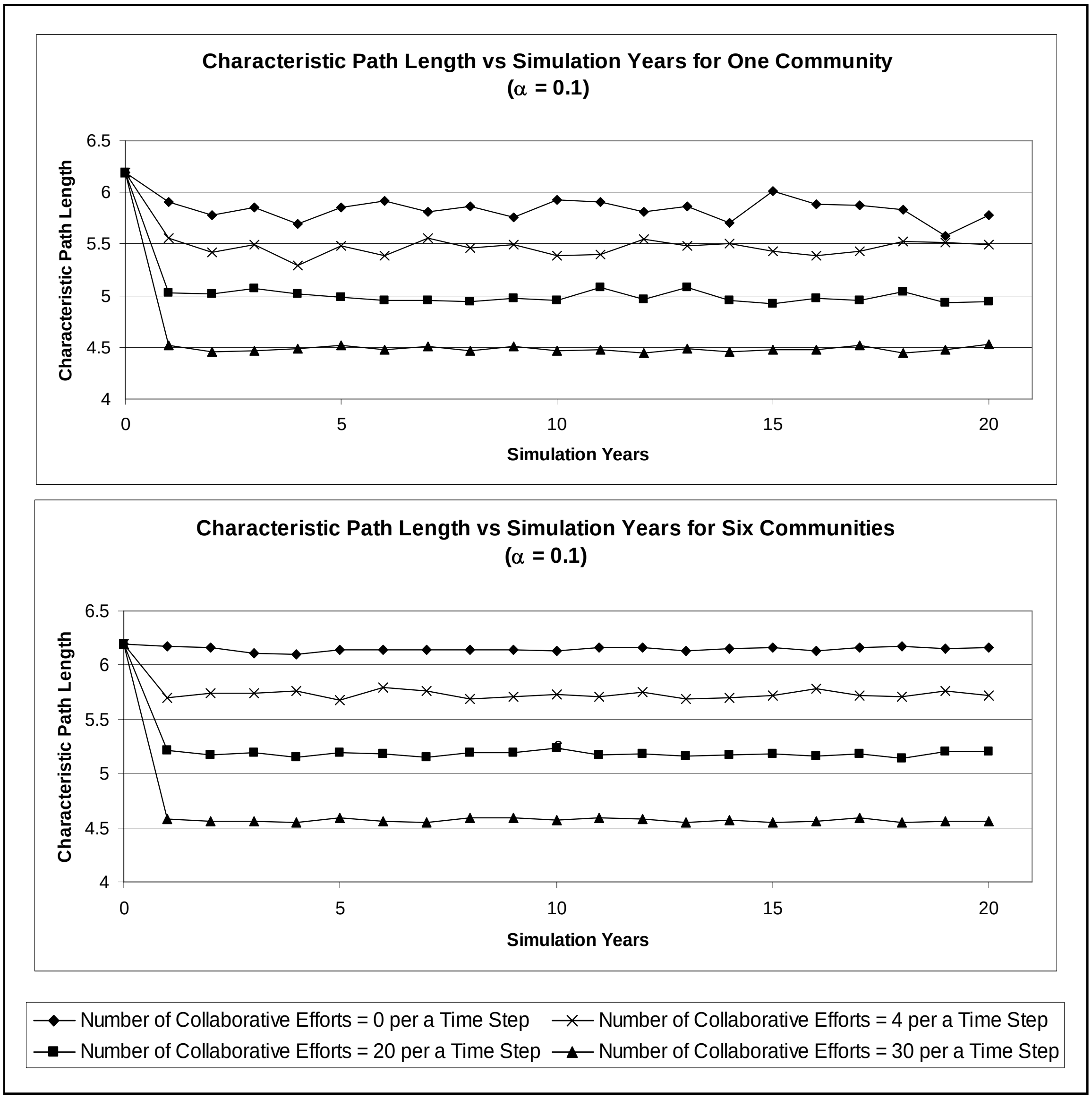

Figure 15 Characteristic Path Length Changes with Long-Range Collaborations.

From these results, it can be concluded that the characteristic path length does decrease as a result of an increase in the number of collaborative efforts members of an organization participate in each year. While the interactions used in this model were representative of short-term projects, one can easily extrapolate that any interactions that allow the development social ties across long distances in the organization would benefit the organization. Further, an organization that is partitioned will require these collaborations in order to come close to the short path lengths experienced in the one community construct even without the collaborations.

Effects of Varying the Attrition Rate.

The next consideration of interest to this study is the result of attrition rate to the small-world phenomenon within the organizational graph. While an organization will generally have an annual attrition rate of approximately 5-10% (6% has been used throughout this research to this point based on Aeronautical Systems Center historical data (Russo, 2003)) organizations will go through times of rapid as well as slow personnel turnover. As a result, an organization must have the ability to anticipate the effect of these increased turn-over rates on the ability of information to cross an organization. Figure 16 displays plots of the changes in characteristic path length due to varying attrition rates.

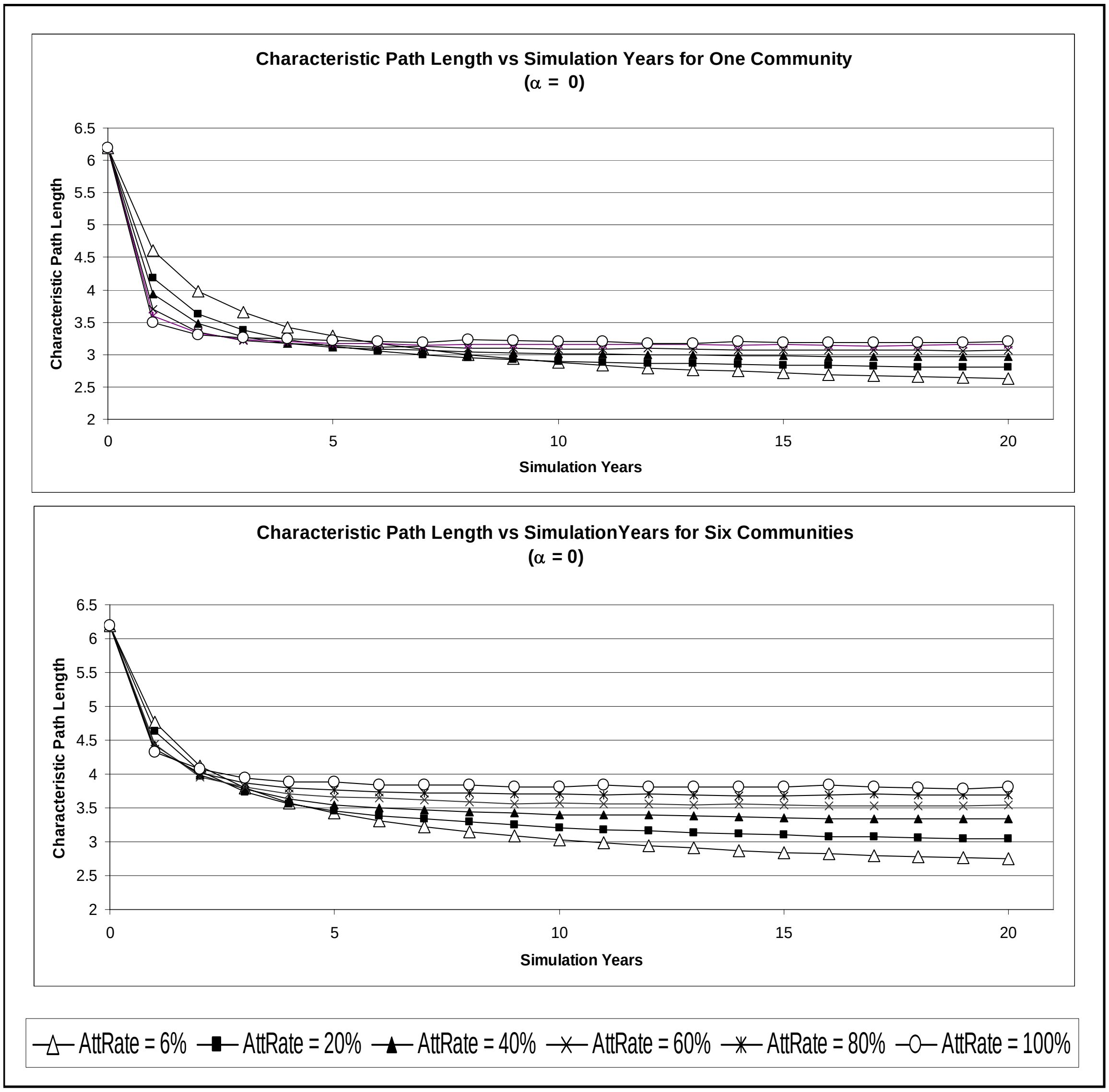

Figure 16 Characteristic Path Length Changes with Varying Attrition Rates.

As can be seen in these plots, when the attrition rate is decreased, the organization will experience a decrease in the characteristic path length. The reason for this phenomenon can be explained logically. The model used for this research requires positions to be filled, forcing new employees to enter a position either from a lower level or by a lateral move. While these promotions and laterals are commencing, long-term collaborations and shared sub-units are increasing the strength of edges throughout the organization, reflecting stronger interpersonal relationships. When the attrition rate is set closer to zero, for each time step fewer personnel are being removed from the organization, a situation realized by a decrease in the number of vertices removed from the graph although the edges within the graph are still forming and gaining strength. When the attrition rate is set to zero percent, the characteristic path length would, eventually reach one, reflecting all personnel know every other person in the organization. On the other extreme, when the attrition rate is set to one hundred percent, all positions are vacated each year. This consideration results in a characteristic path length that drops quickly due to a large amount of movement throughout the organization to fill vacancies and stays level throughout the twenty year timeframe.

Since communications based on long-distance collaborations tend to be weaker, as can be seen in Figure 17, these edges along with those built from working together long ago in the simulation, tend to be removed more readily when considered against the already described α value of 0.2. Of particular interest is the disparity between the spread of values of the six community and the one community characteristic path lengths. This difference reflects the reliance on edges formed by collaborations or working together long ago rather than working in the same sub-units recently for the six community construct since these edges are largely removed with an α equal to 0.2. This is evidenced by the decreasing characteristic path length for increasing attrition rates in the one community construct, where personnel are permitted to move throughout the organization.

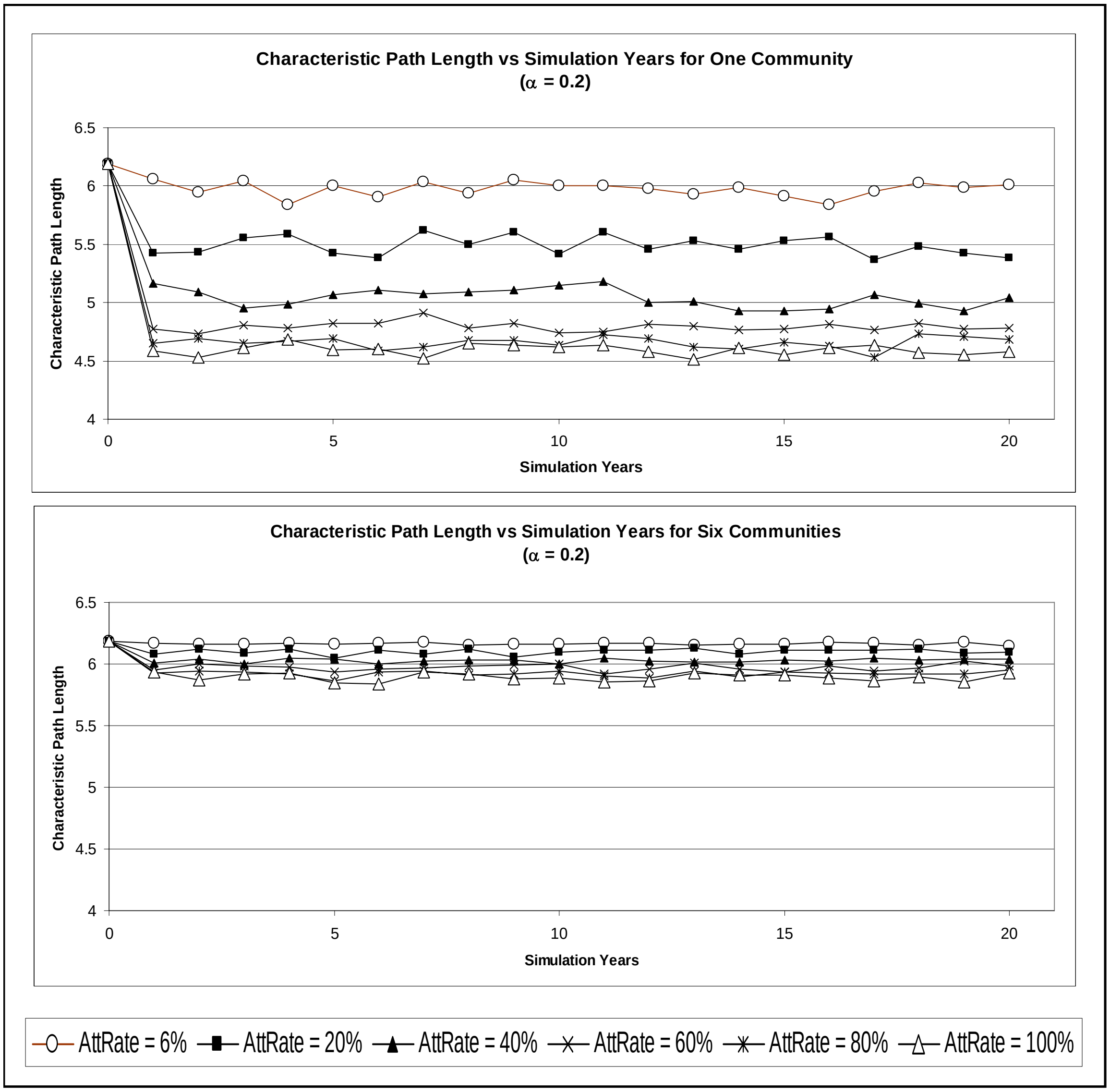

Figure 17 Characteristic Path Length Changes with Varying Attrition Rates.

These results show that the attrition rate greatly affects the organizations characteristic path length, reflecting the social connections of an organization and, therefore, the ability for information to cross an organization. A typical attrition rate (on the order of 5-10%) will result in comparable characteristic path lengths for both a one and a six community organization due to long-range collaborations. Of particular interest, though is the result when removing some of the lower strength edges. When this occurs, edges formed due to sharing a sub-unit greatly outweigh those from collaborations. As a result of the reliance on these stronger relationships, the one community organization retains a relatively (as compared to the six community organization) short characteristic path length for increasingly larger attrition rates. This result is due to the large amount of movement across the corporate entity within the single community organization, increasing time within sub-units, which is precluded when personnel are restricted to one smaller community throughout a career. This realization reflects well on the robustness of the one community construct's social network in relationship to the social network of a partitioned organization. Additionally, this model indicates that the social network, thus the ability to communicate and handle crises, for an organization facing a large attrition rate in the near future would fare better under a non-partitioned organizational construct than a partitioned one.

Effects of Changing the Organizational Structure.

The final primary considerations for characterizing the most optimistic traits of these organizations are to consider the result of using a graph that has already been aged for 20 years as an input to an aging routine. The first graph, as seen in Figure 18, displays the exact opposite, transitioning from six communities to one large community. The second graph, as seen in Figure 18, displays the effect on the characteristic path length of transitioning an organization from one large community to six smaller communities.

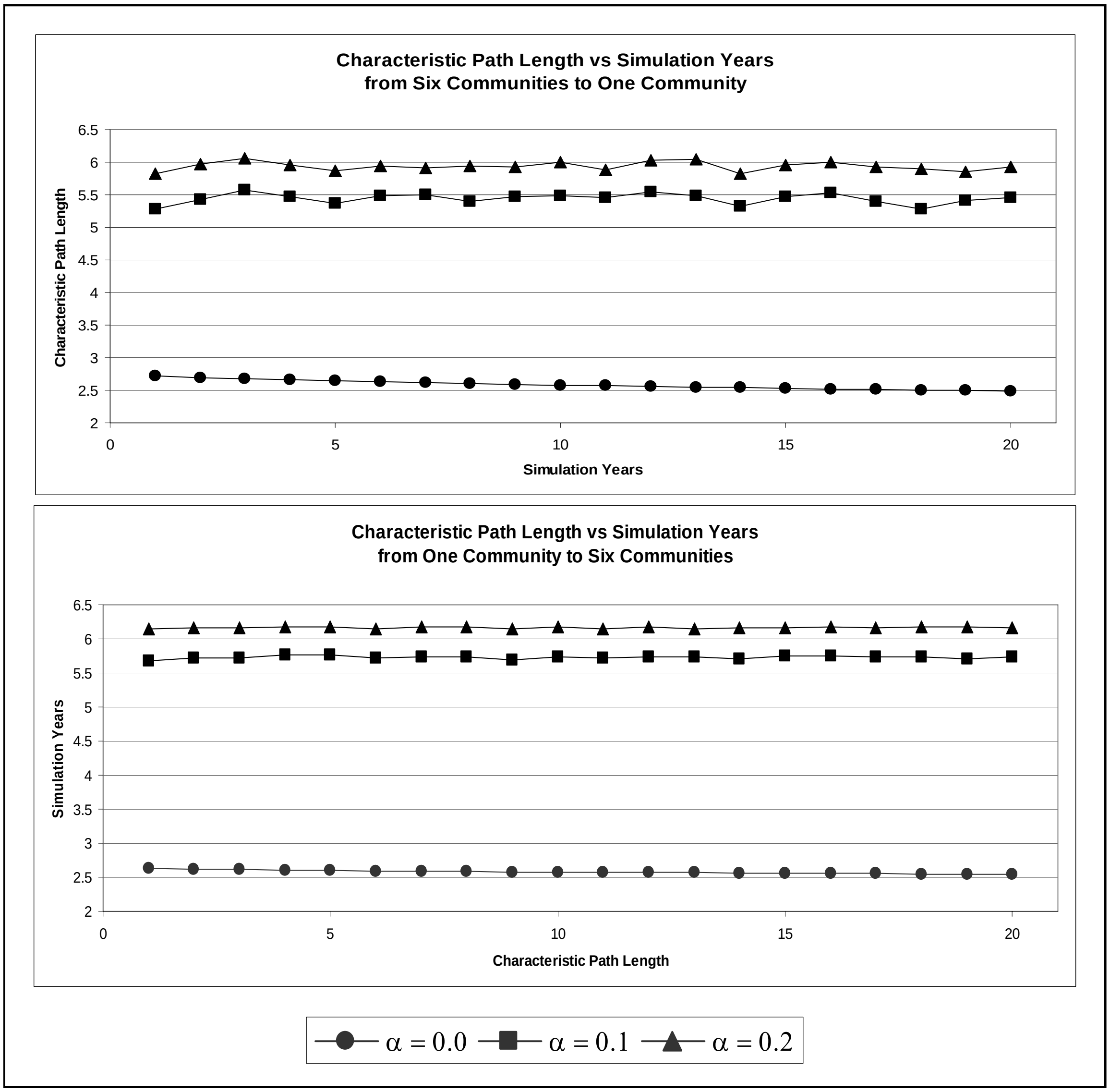

Figure 18 Characteristic Path Length Changes with Changing Community Structure.

As can be from these two plots, when all other constants are held steady, there are only marginal affects on the characteristic path length due strictly to changing the structure of the organization. Due to the scale of Figure 18, there appears to be little difference between the two considerations. By looking at the percent difference between the two, as shown in Figure 19, a difference favoring a one community construct can be seen, though. The dots shown on Figure 19 are plots of the percent difference between the results of a change from one community to six and the change from six communities to one, shown in **Figure 18**, with a linear curve fit for each value of α. By following the linear trend lines in Figure 19, one can see that the difference between transformation from six communities to one community and from one community to six communities grows as the respective organizations transform.

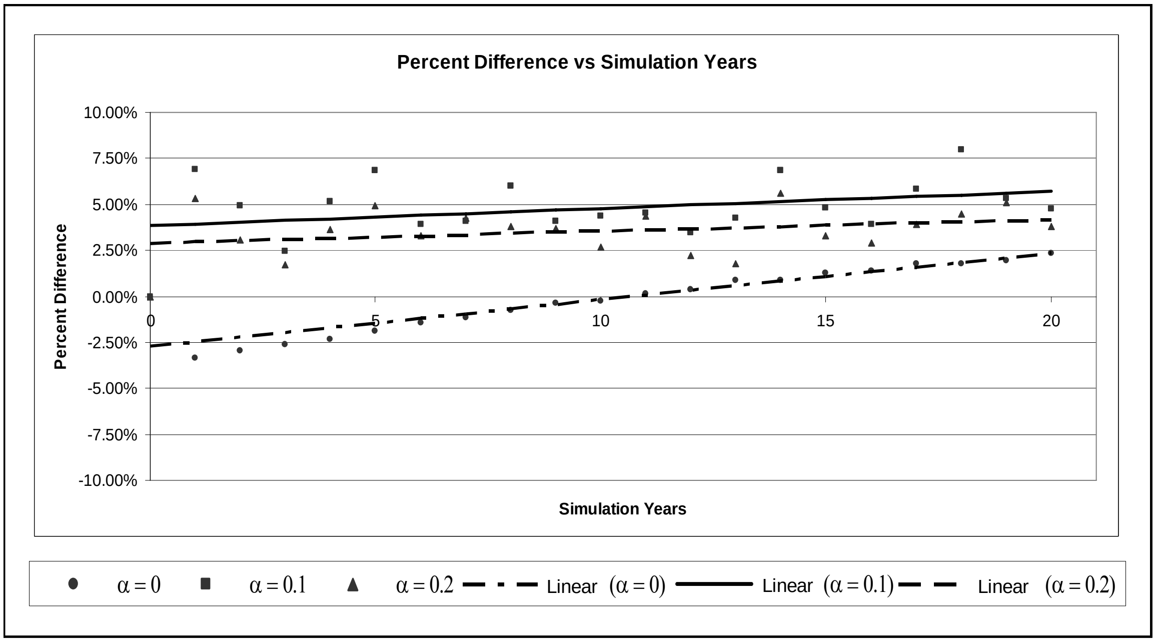

Figure 19 Percent Difference in Path Length Due to Community Structure.

These differences vary up to over a five percent difference, reflecting an advantage for an organization transforming from six communities to one community based on this model.

Characteristic Path Length Conclusion.

The characteristic path length of an organizational graph reflects the most optimistic average shortest path for between any two people within an organization. This value can be used as a gauge for the ability of an individual to share information with any person within the organization. From the preceding considerations, it can be seen that encouraging collaborative efforts from disparate parts of an organization can rapidly decrease the absolute characteristic path length, although, for these assumptions, the edges formed from these connections tend to be weak in comparison to those formed by sharing the same sub-unit. Further, an increase in cross-organizational collaborative efforts decreases the characteristic path length regardless of any formal partitioning schema. An increase in these efforts is a viable choice for an organization employing a partitioned structure that wishes to keep the characteristic path length, and, hence the ability to share information across an organization, short. Next, an increase in the attrition rate favors a one community organization over a six community organization when weaker edges are removed. This consideration reflects the reliance of collaborations on a partitioned organization, reaffirming the necessity of a partitioned organization to encourage these cross-organizational efforts. Finally, there is a difference in the effects on the characteristic path length when transitioning from a non-partitioned organization to a partitioned one and vice versa. The end result of removing the partitions on an organization returns a shorter characteristic path length than when installing these partitions. From this model's results, an organization should endeavor to remove partitions whenever possible to reduce the characteristic path length, and, as a result, enhance the ability to share information across an organization. When removal of partitions cannot occur, an organization should ensure personnel are working on many efforts shared throughout the organization.

Edge Search Considerations

The preceding considerations reflected the effects of varying several inputs on the characteristic path length of a graph representing an organization. This characteristic path length then could be used as a most optimistic measure reflecting the ability of information to cross an organization. The measure does not necessarily address the reason many organizations accomplish a partitioning for, though. The reason many organizations will partition their organization is to enhance the ability to accomplish short-range searches for information, due to a smaller organization to search, realizing that there may be penalties to the ability to search throughout the organization. The next portion of this analysis addresses the ability to search an organization using the search heuristic previously described in the methodology.

Discussion of Random Number Generator Mean.

One critical aspect of the search algorithm is a random number generator. Throughout this search algorithm, the random number generator within MATLAB was used to produce search choices as well as to decide whether to pursue a link. As a result of this reliance on the number generator, the mean of the number generator was adjusted to more accurately estimate the probability of an edge being used by estimating the mean of the edge strengths. If this is not accomplished, the organizational search will always default to the organizational structure. This consideration realizes that the distribution of the random number generator does not follow that of the distribution of edge weights in the adjacency matrix, rather, it is intended to allow the random number generator to produce numbers on the same order as those in the adjacency matrix. After considering several of the adjacency matrices, the mean edge strength was approximately 0.0051, which corresponds to a full length edge degrading for 82 weeks. Considering the assumption of a 50 work week year, two sets of runs, one where the mean of the random number generator is set at 75 weeks (1.5 years since having a strong relationship) and one at 100 weeks (2 years since having a strong relationship) will be considered. This mean will be accomplished by raising any random number (always from 0 to 1) by 6.966 and 9.288 for 75 and 100 weeks, respectively. By placing the mean at 75 weeks, the search will, on average, reject most edges and by placing the mean at 100 weeks, the search will, on average, accept most edges.

Figure 20 displays the result of leaving the mean of the random number generator at 0.5. As predicted, the search algorithm defaults to the strict organizational links in this case since there are very few edges within the graph greater than any number created by the random number generator. Since the bare organization for a six community and a one community organizational structure are the same, the edge search lengths are the same for both structures. These results reinforce the importance to the interpretation the model of using the different values for the random number generator mean. A further consideration in the search algorithm is the choice of the pairs to run a search between. These pairs, chosen at random within a community, represent an *a* and a *b* in a *a* to *b* search. For these considerations, two search types were considered, based on a variable labeled ϕ. The value of ϕ represented the amount of the search pairs for any run that were within the same community. Under this construct, when $\phi = 0\%$ all of the searches were between separate communities. Conversely, when ϕ = *100%* all of the searches were within the same community.

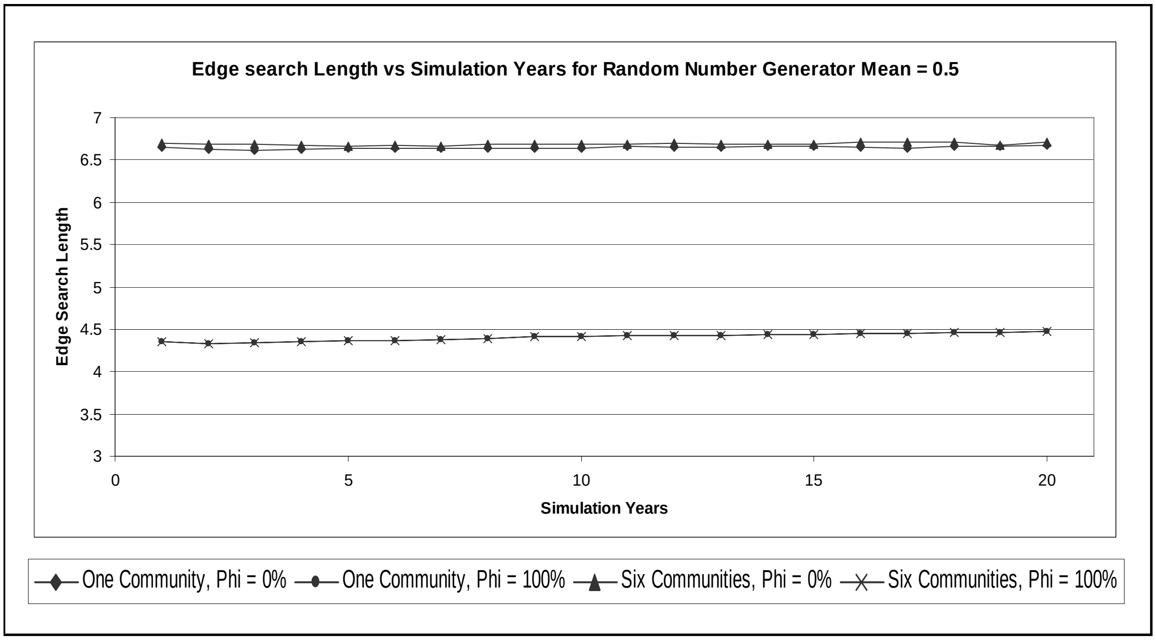

Figure 20 Edge Search Length with No Scaling of Random Number Generator.

Effects of Varying the Number of Collaborations.

Following the pattern of the considerations used when looking at the characteristic path length, the preliminary consideration will be the effects of long-distance collaborations on the ability to search for information. In accordance with the previous findings from the characteristic path length studies, the edge search length for the organization should drop when the number of long-distance collaborations is increased. Further, these collaborations will tend to result in weak relationships that will not always be used when searching the organization. Lastly, due to the cross-organizational nature of these long-range collaborations, they should have little to no impact on searches within the same smaller community.

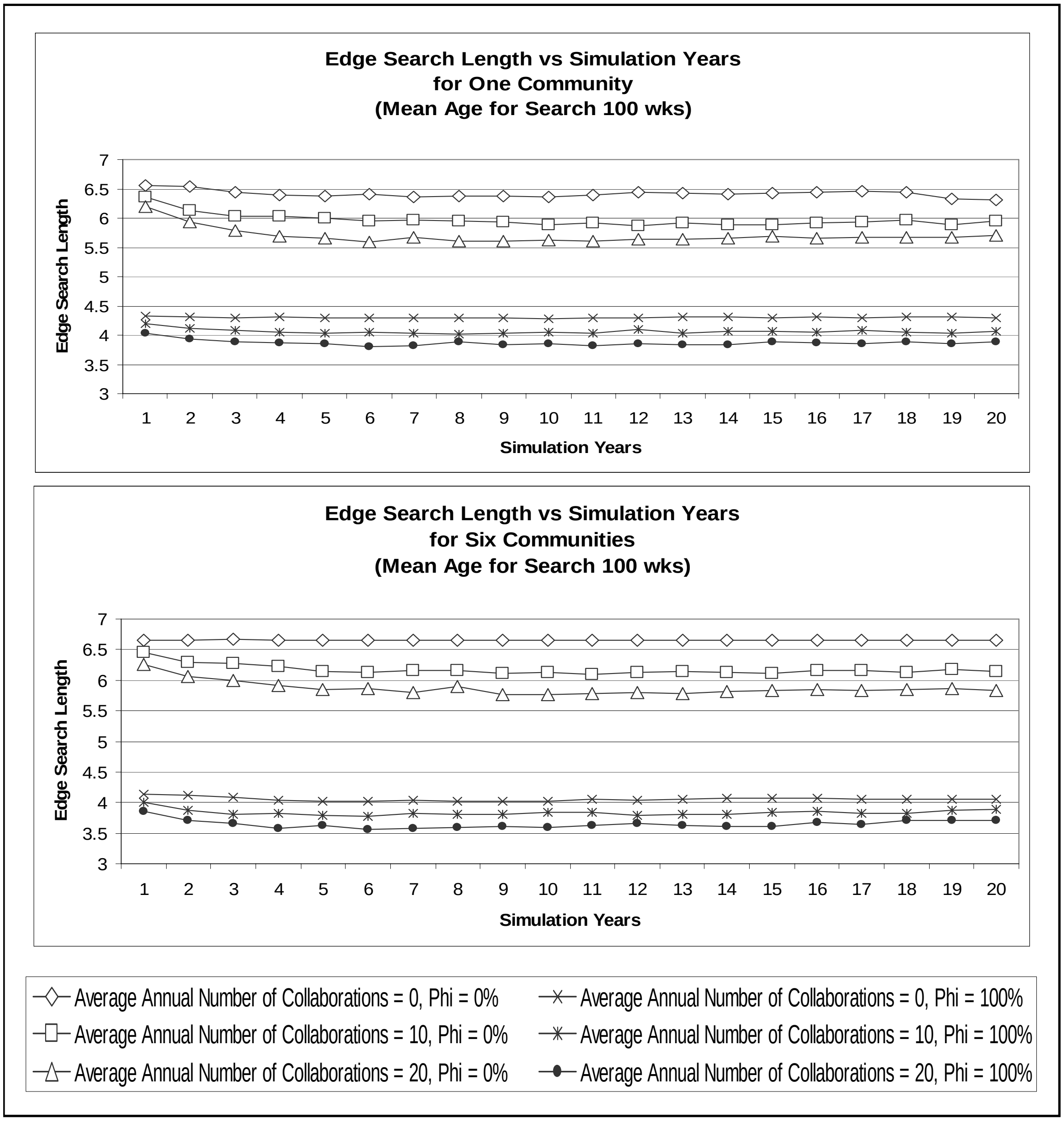

Figure 21 Edge Search Length Changes Due to Long-Range Collaborations.

Figure 21 presents the results of searches for both a one community construct and a six community construct. As can be seen in these plots, additional collaborations did result in a shorter edge length, particularly when considering searches exclusively between different communities. In both the one community and the six communities case, the benefit of additional collaborations to this model is as high as 10% of the baseline edge strength. Further, as predicted, there was little benefit to short-range searches from long-distance collaborations (although a little in the one community construct).

With an appropriate appreciation of an organizations structure, attrition rate, and necessary levels of information flow, a decision-maker can take these measurements to consider how many tiger teams or long-reaching integrated product teams he or she should form throughout the organization. If the organization tends to need information to be found easily by all personnel anywhere in the organization, regardless of the organizational structure, the manager could increase the number of collaborative efforts. If the organization tends to not have a knowledge base that is separated amongst several specialists, the organization might need not take the time to encourage these efforts.

Effects of Varying the Attrition Rate.

The next consideration a decision maker would care about involves the effects of varying the attrition rate on an organization with different structures. Following the findings in the characteristic path length, changes in the attrition rate should have little effect on the search length other than to weaken the edges within the graph and to create marginal disparity in the results of the one community construct. After running the search algorithm for both a one community and a six communities organization, these expected results were found again. The results of this analysis can be seen in Figure 22 for both organizational structures after the mean of the random number generator was changed.

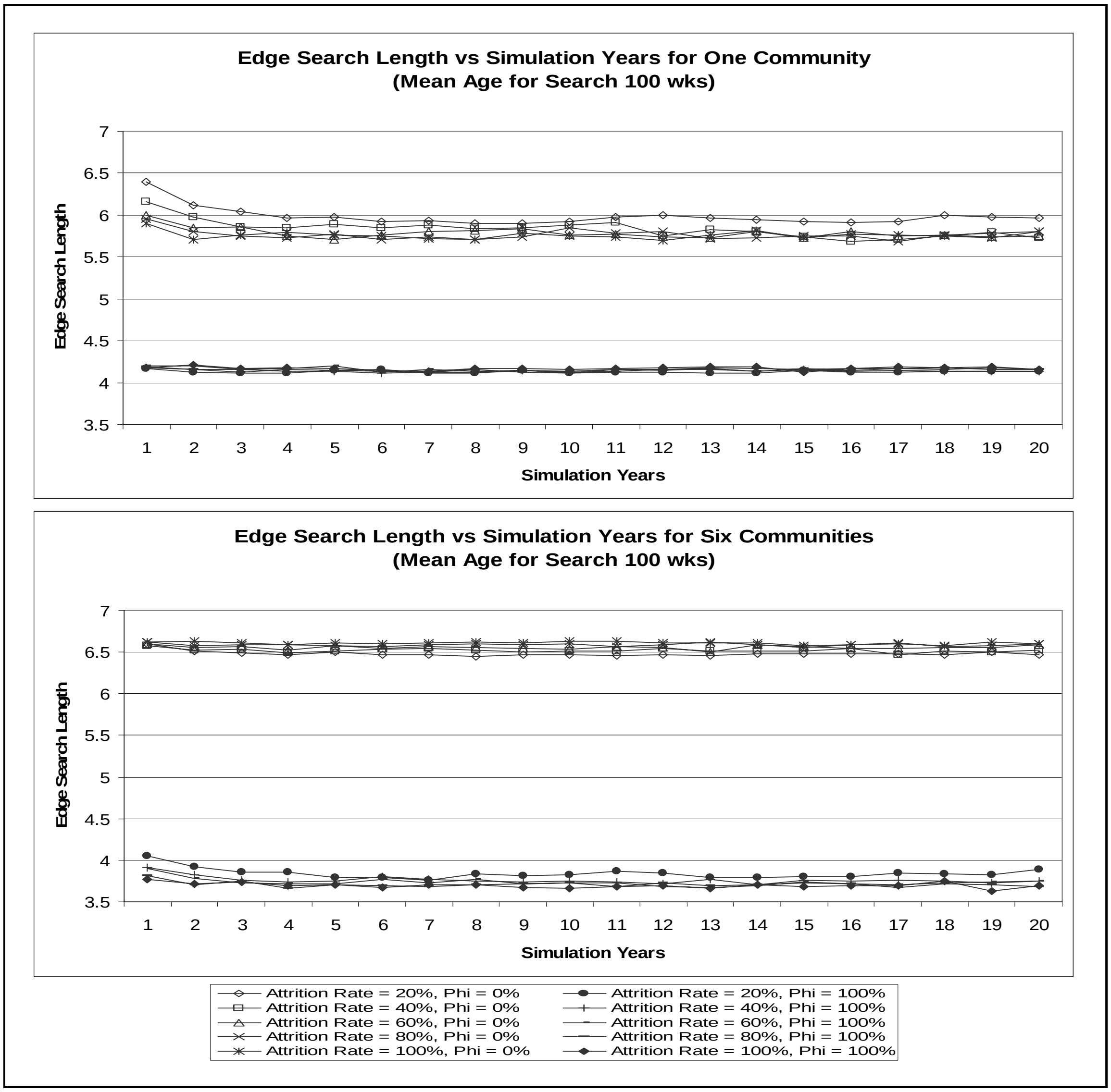

Figure 22 Edge Search Length Changes Due to Varying Attrition Rates.

As can be seen in these plots, the attrition rate did tend to create disparity in the results for the cross organizational (ϕ=0%) search for the one community construct. Like found in the characteristic path length analysis, this is due to an increased amount of personnel movement throughout the organization permitted by the one community construct. This spread is not seen in the long-range search results for the six community construct due to personnel's inability to move throughout the organization and form relationships in various sub-units. Additionally, the low strength of the edges is evidenced by the fact that the average of the random number generator had to be increased to a two-year mean to prevent the model from choosing the organizational substrate instead of the social network. When the random number generator was left unchanged (ie the mean of the random number generator was not changed from 0.5), there was no difference in the results from the two organizational structures.

The next point of interest that can be drawn from Figure 22 is a consideration of the average edge strength for the attrition rates for each organizational structure. The average of each grouping of lines from Figure 22 is plotted in Figure 23. These plots exhibit disparity in edge search lengths within both organizational types.

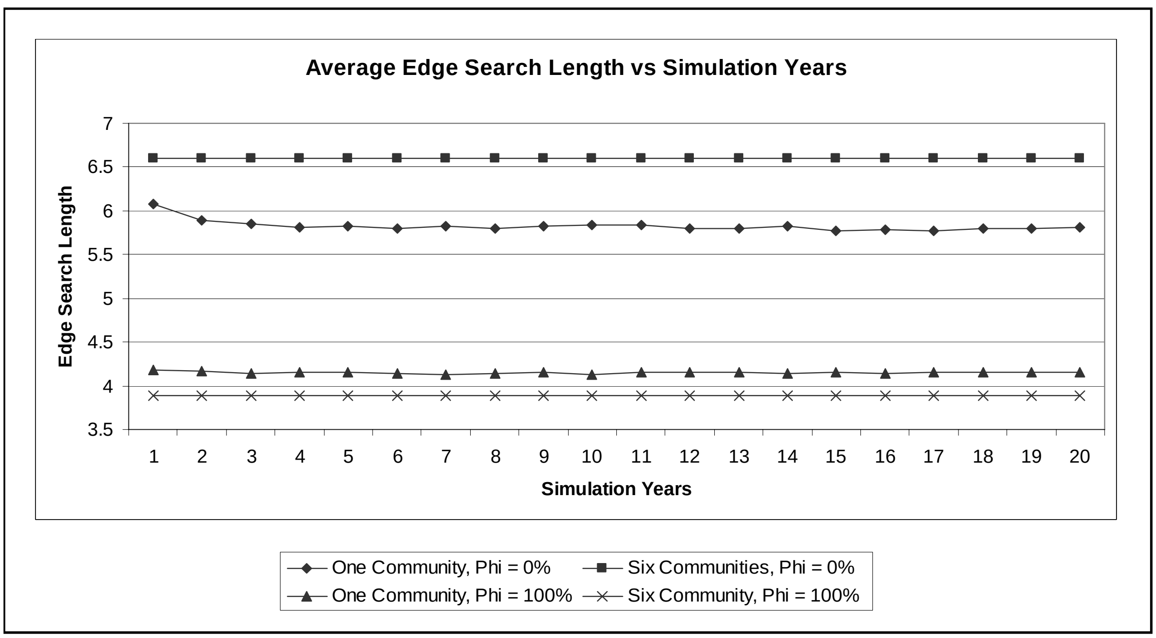

Figure 23 Average Edge Search Lengths for Different Search Considerations.

As can be seen, when accomplishing a cross-organizational search, there was a significant advantage to having a one community construct. To the contrary, when accomplishing a search within a smaller community, it is advantageous to have a partitioned organization. Both of these are logical outcomes due to the consideration that allowing cross-organizational connections to form (by allowing movement throughout an organization) enhances a one organizations ability to perform long-range searches whereas a decreased search pool with more local edges formed enhances the ability to perform short-range searches. Ultimately, in deciding upon an organizational structure, a decision maker must consider the mix of searches that he or she foresees in the future for his or her organization.

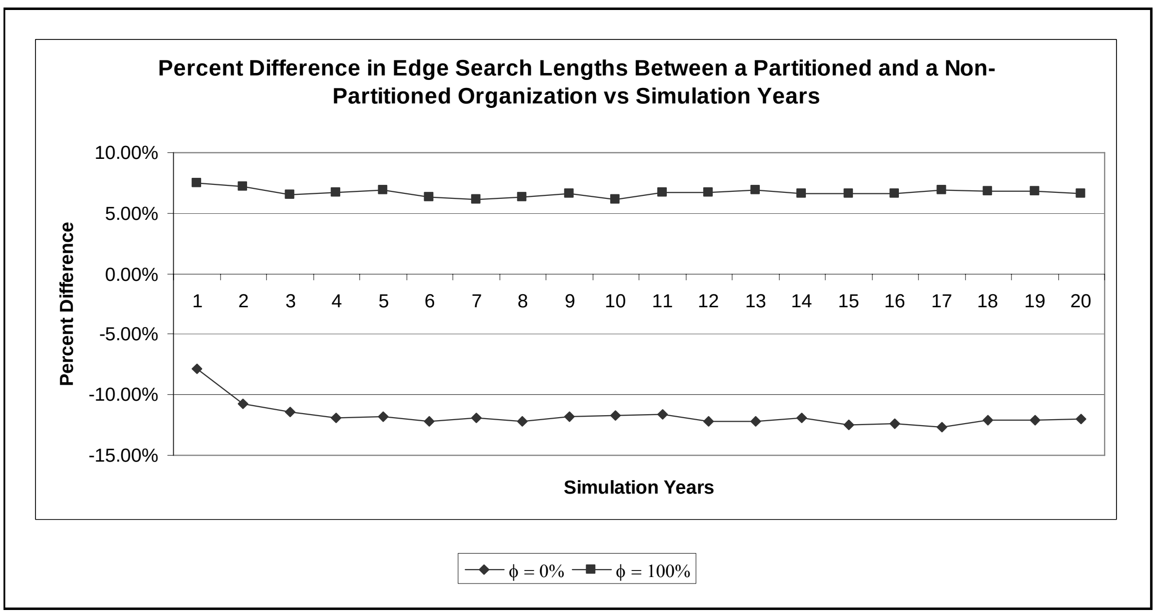

Figure 24 Percent Difference in Edge Search Length Due to Organizational Structure.

Based on Figure 24, a plot of the percent difference of the average edge search lengths presented in Figure 23, in this model, an organization choosing a partitioned organization will enjoy a roughly 6% benefit in short-term searches by choosing a partitioned organization at the penalty of a 12% decrease on potential long-distance searches. An organization that performs many short-range searches might choose to accept this long-distance search penalty, though if they accomplish very few long-range searches.

Effects of Changing the Organizational Structure.

The final consideration for searches will be the consideration of changing the organizational structure after a fully matured network has formed. Similar to the way this was accomplished in the characteristic path length measurements, for these sets of runs, after 20 years maturation, the organizational structure was changed from a one community to a six community structure (or the reverse) and them matured for an additional 20 years. Further, the various mean ages were used to consider whether to pursue a link or not. Lastly, the consideration of whether to look exclusively within one, smaller community, or look exclusively amongst different communities was considered.

Plots of an already aged one community organization transitioning to a six community construct as well as vice versa can be seen in Figure 25. From these plots, it can be seen there is little difference between the two constructs, as well as the different age means. Examination of these plots validates the arguments for both a one community and a six community organizational construct. The transition from six communities to one community does result in a shorter average edge search length than the transition from one community to six communities when considering searches exclusively between different partitioned communities. Further, the transition from one community to six communities results in either a shorter edge search length or minimal change in edge search length for searches exclusively within the same community.

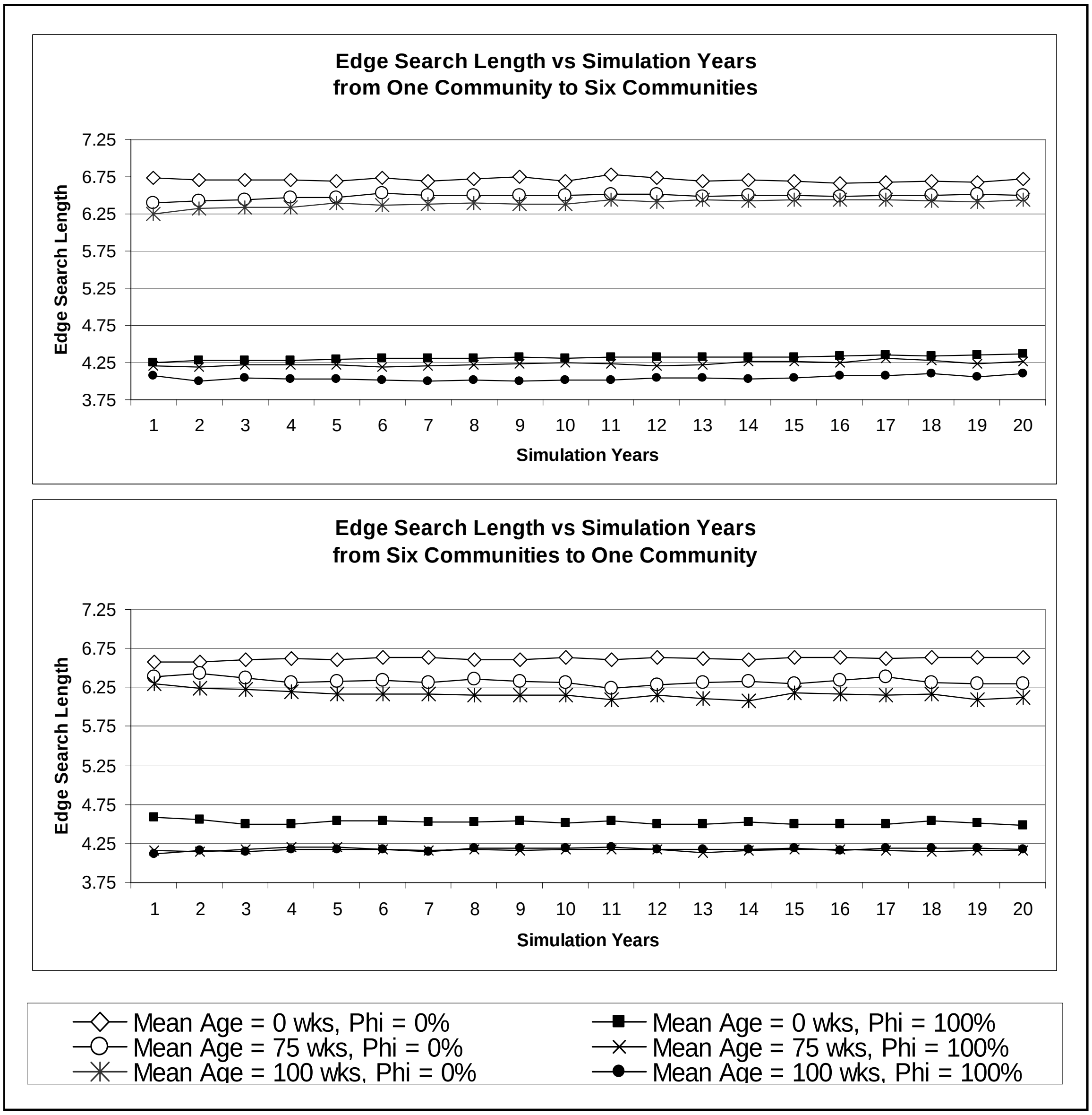

Figure 25 Edge Search Length Changes Due to Changing Organizational Structure.

These results boost the argument for both a partitioned and a non-partitioned organization. These findings must be considered against an organization's need for communications throughout the organization. The search choices were made based on a random choice of vertices within the graph, the majority of which were on the lower two levels, explaining the long edge search lengths seen in these outputs (since the search would predominantly be from a bottom vertex to a bottom vertex). This consideration, along with the weak edges within the graph meant that the search defaulted to the formal organizational tree quite often when a search across the organization was necessary. This observation should also enter into the considerations of an organization when deciding upon a formal structure. The organization must consider how easily information requests flow through management. If the search would meet a middle manager and then stop, the search time would become much longer for a search outside the community. In fact, if the manager was reluctant to allow personnel to look outside the host community, the search would stop and personnel could not follow formal chains to find information. In this case, these personnel would have to default to social connections or not get the information necessary, reinforcing the necessity for long-range collaborations.

V. Conclusions and Recommendations

Conclusions of Research

Through this research, a model was created reflecting the dynamics of an organization's social network under varying organizational structures. While the variables, such as the strength of collaborations, have not undergone benchmarking, the model can be used to extrapolate trends that an organization can take advantage of when undergoing or considering a structural change. The product of this model is a mathematical way of evaluating the effects of organizational structural change on an organization's social network, and, thus its ability to share information and react to crises based on two measures. The first measure, based on the traditional graph theory measure of characteristic path length, has been modified to consider the strength of an organization's relationships by removal of weaker edges from the adjacency matrix representing the organization. The second measure is based on a search heuristic that was developed to consider the way personnel search for solutions within a military organization. This search heuristic was then used to evaluate the graph to find a more realistic representation of the number of social links a person must cross when searching for information in the organization.

As a result of this research, a greater knowledge of the effects of community structure on the social network of an organization has been provided. Using this model, there is a 12% longer distance that must be traveled when searching for information outside of one community when the organization is partitioned. Conversely, when looking inside the smaller communities created by partitioning, the distance that must be traveled when searching is 6% shorter than it was prior to partitioning the organization. These effects benefit a partitioned organization during day-to-day activities where, it is assumed, an individual does not have to search far for information. This construct hinders an organization in a crisis where an individual must search outside that person's community in a timely manner for a solution, though. In order to counter this negative effect on the average distance between personnel, an organization can take proactive steps to encourage cross-organizational relationships amongst its personnel. The two ways investigated here were for organizations to encourage teams with personnel from throughout the organization or to vary attrition rates and personnel movement. Particularly long-distance collaborations, which could include tiger teams or planning complex social events such as the organizational bar-b-que, increase the number of cross-organizational relationships, resulting in a drop in the average distance between personnel to the same level as that of a non-partitioned organization. When these collaborations are not encouraged, there is as much as a 25% advantage, in this model, to a non-partitioned organization. Further, an organization's attrition rate impacts the optimal path length between any two people within the organization favoring the non-partitioned organization by as much as 10%, in this model, for very high attrition rates. Conversely, organizational structure has little effect on the average distance between employees with typical (5-10% annually) attrition rates. Further, the attrition rate has little effect on the ability to randomly search either a partitioned or a non-partitioned organization. These results validate a small-world assumption for an organization since these long-distance relationships tend to have a non-linear effect on the characteristic path length of an organization which is directly related to the size of an organizations social network.

Significance of Research

This research provides a way to quantify the effects on an organizations social network of converting an organization from one in which personnel were permitted free movement throughout the corporate entity to one in which those individuals are restricted to a particular community. An organization that enforces a partitioned construct can expect information to flow more quickly within the now smaller communities, but at the penalty to information flow between different communities. In addition to the enlarged social distance between personnel in different communities, the partitioning often results in information searches being forced through increasingly higher levels of management instead of staying with those who actually need the information. This results in an increased amount of paperwork staffing, ultimately decreasing an organization's ability to react to crises, in a timely manner, when different parts of the corporate entity must participate in the response.

Recommendations for Action

An organization should consider the effects of its organizational structure on its social network based on the amount of information flow it requires to operate on a daily basis as well as in emergency situations. An organization in which little communication at the lowest levels such as is the case with personnel working on different assembly lines for the same manufacturing company would benefit from a partitioned organization. This is due to the assumption that employees accomplishing extremely repetitive tasks benefit little with increased knowledge of their colleagues' capabilities or knowledge set. In contrast, and organization such as an aerospace engineering firm must encourage as large of a knowledge base as possible. As a result of this, assuming a construct restricting interpersonal contact and transitions due to partitioning should not be chosen. Personnel in these organizations must be able to rapidly search for solutions for varying problems very few will have all the answers for. As a result, partitioning an organization like this would hinder timely solution finding, especially during crises, by increasing the number of personnel a person must go through to find an answer.

Recommendations for Future Research

An assumption of this research was a uniform distribution of personnel throughout an organization. While mathematically elegant, a true organization will have a varying number of subordinates reporting to each superior. Further, when considering collaborations, a δ-value of 2/3 was used throughout the research. Additional study could consider the effects of varying this value on the ability to communicate and search for information within an organization. Additionally, this research assumed an organization with 1555 due to limits in computer capability. Further research should consider larger organizations with varying superior to subordinate structures. Further research should also consider disparity in employee types as well as different search algorithms such as those proposed by Adamic (Adamic, 2001). Finally additional research should be accomplished on the effects an organizations structure has on its corporate knowledge base.

Summary

When the decision was made to consolidate several of the disparate ASC functional offices into a handful of Wings or Groups, one of the primary hypotheses was that this would fix a lack of communication between similar offices. This was in part due to a weakened engineering functional which was originally responsible for long-term growth, by relocation, of its engineers. As has been found in this research, placing these offices into a handful of smaller communities (product focused Wings and Groups) within the corporate construct does decrease the average distance between personnel within these smaller communities. This type of partitioning also has the downside of increasing the distance between members of separate communities. The impact of this finding is reinforced when considering a search within the organization. When the organization was partitioned, there was a 12% increase in the number of people that had to be asked for an answer before the solution was found over the comparable finding when the organization was not partitioned. In comparison, though, searches exclusively within the smaller community were 6% shorter for the partitioned organization in comparison with the non-partitioned organization as seen in Figure 24.

The reason for these results has to do with the small-world effect discussed by Duncan Watts (Watts, 1998). He found that the characteristic path length of a graph was highly dependent on the number of long-distance edges within the graph. This finding impacts an organization when it is partitioned because, as the organization goes through normal cycles of attrition, hiring, and promotion, a partitioned organization disallows movement across the entire organization, decreasing the number of cross-organizational links created throughout a career. A consequence of this restriction is that personnel do not form the far-reaching relationships necessary to accomplish a quick search through the different communities in the organization, although they gain a stronger knowledge of the people within their same community.

Realizing the non-linear relationship between characteristic path length and the number of long-distance links within an organization, a manager can force these long-distance relationships by supporting participation in cross-organizational collaborations. An increasing number of these collaborations does increase the number of long-distance relationships, thus decreasing the characteristic path length of the organization although these links tend to be quite.

When considering an organizational structure, a decision maker must consider the amount of information sharing necessitated by the primary corporate activity accomplished. A manufacturing organization in which little interaction is required between personnel such as those on an assembly line might prefer a partitioned community in which solutions can be found quickly within a product cell. Members of a firm accomplishing advanced (aerospace) engineering design or scientific research where communication with members accomplishing often very different efforts would prefer a one community construct where personnel are encouraged to interact in many communities throughout a career.

Appendix A: Definition of Edge Aging Calculation

The initial consideration for the aging of an edge is that the edge will age logarithmically. The curve this strengthening will follow will be based on the following equation:

$$y(t) = 1 - e^{-\gamma t} \tag{8}$$

where y = edge strength, t = simulation time interval, and γ = inputted constant based on a desired edge strength.

When plotted, Equation 8 shows the logarithmic growth desired as seen in Figure 26 .

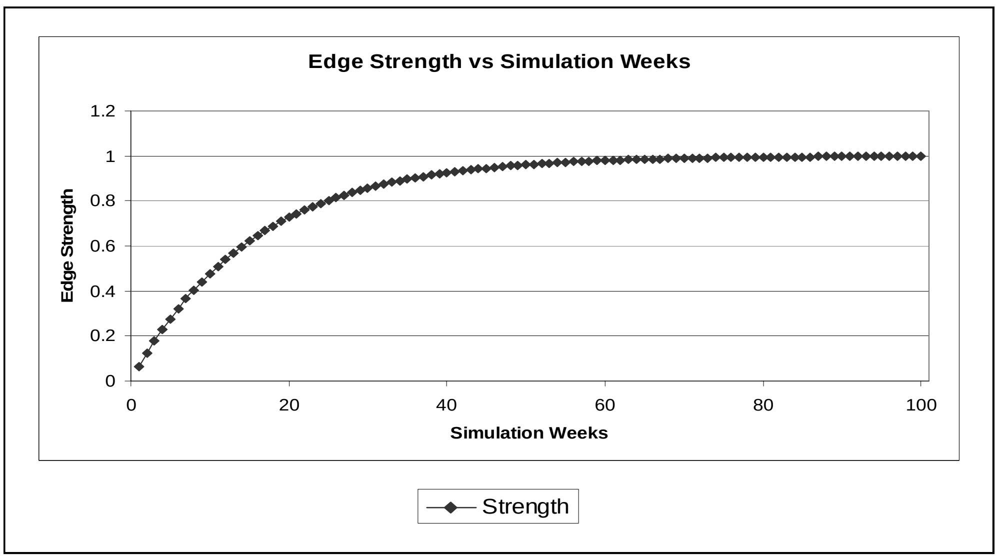

Figure 26 Edge Strengthening Profile.

while weakening of the edge will be accomplished through the following equation:

$$y(t) = e^{-\gamma t} \tag{9}$$

where y = edge strength, t = simulation time interval, and γ = inputted constant based on desired edge strength.

When plotted, Equation 9 shows the logarithmic growth desired as seen in Figure 27.

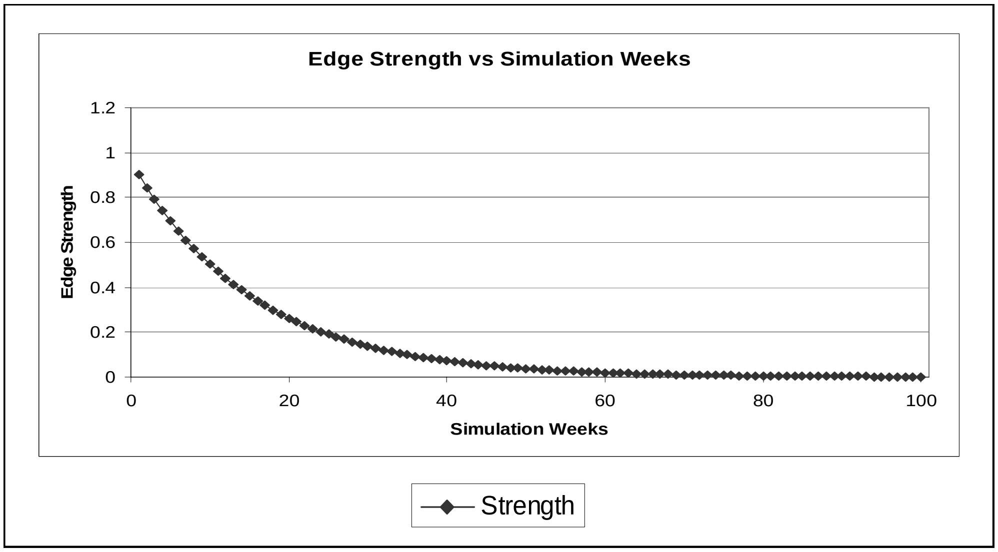

Figure 27 Edge Weakening Profile.

In order to solve for the γ, a desired edge strength (y_t) at a time (t) must be entered. After entering this data, γ is found using Equation 10.

$$\gamma = -\frac{\ln(1 - y_t)}{t} \tag{10}$$

Edge Strengthening

In order to find the change in y, Δy, due to some change in t, Δt, when strengthening the edge, begin by considering Equation 8 at two different times.

$$\Delta y = y(t + \Delta t) - y(t) \tag{11}$$

Substituting Equation 8 into Equation 11:

$$\Delta y = (1 - e^{-\gamma(t+\Delta t)}) - (1 - e^{-\gamma t}) \tag{12}$$

After collecting terms and simplifying, this returns:

$$\Delta y = e^{-\gamma t}(1 - e^{-\gamma\Delta t}) \tag{13}$$

Using $e^{-\gamma t} = 1 - y(t)$, we have:

$$\Delta y = (1 - y(t))(1 - e^{-\gamma\Delta t}) \tag{14}$$

Equation 14 represents the change in *y*, for a prescribed time change, from the original value when strengthening. To find the value of y at this new time find $y+\Delta y$.

$$\begin{aligned} y(t+\Delta t) &= y(t) + \Delta y \\ &= y(t) + (1 - y(t))(1 - e^{-\gamma\Delta t}) \\ &= 1 + e^{-\gamma\Delta t}(y(t) - 1) \end{aligned} \tag{15}$$

Equation 13 allows the calculation of a new edge strength when an initial strength and a change in time is known.

Edge Weakening

At this point, there is a necessity to be able to weaken the edge as well. The weakening will follow the curve found in Equation 9. When considering the weakening, the change in strength, Δy, will be based on the following:

$$\Delta y = y(t+\Delta t) - y(t) \tag{16}$$

Substituting Equation 9 into this:

$$\begin{aligned} \Delta y &= e^{-\gamma(t+\Delta t)} - e^{-\gamma t} \\ &= e^{-\gamma t}(e^{-\gamma\Delta t} - 1) \end{aligned} \tag{17}$$

Now that the change in strength for a change in time is found, add this to the original strength:

$$\begin{aligned} y(t+\Delta t) &= y(t) + \Delta y \\ &= y(t) + e^{-\gamma t}(e^{-\gamma \Delta t} - 1) \end{aligned} \tag{18}$$

Substituting Equation 9 into this and collecting terms:

$$y(t) + \Delta y = y(t)(e^{-\gamma \Delta t}) \tag{19}$$

Equation 19 returns the strength of an edge after weakening.

Bibliography

1. Adamic, L. A., Lukose R. M., Puniyani A. R., & Huberman, B. A. (2001). Search in power-law networks. *Physical Review E, 64*(046135).

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

3. ASC/PA (2004). *Aeronautical Systems Center Fact Sheet.* http://88abw.pa.public.wpafb.af.mil/Docs/Fact%20Sheets/ASC%20Factsheet.htm

4. Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science, 286*, 509-512.

5. Biggs, N. L., Lloyd, E. K., & Wilson, R. J. (1976). *Graph Theory 1736-1936.* London, UK, Oxford University Press.

6. Bryne, D., & Buehler, J. A. (1955). A note on the influence of propinquity upon acquaintenceships. *Journal of Abnormal and Social Psychology*, *51*, 147-148.

7. Dodds, P. S., Watts, D. J., & Sabel C. F. (2003). Information exchange and the robustness of organizational networks. *Proceedings of the National Academy of Sciences*, *100*(21), 12516-12521.

8. Erdös, P., & Rényi, A. (1960). On the Evolution of Random Graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Science*, *5*, 17-61.

9. Euler, L. (1736). Solutio Problematis ad Geometriam Situs Pertinentis. *Commentarii Academiae Scientarum Imperialis Petropolitanae, 8,* 128-140.

10. Faloutsos, M., Faloutsos, P., & Faloutsos, C. (1999). On Power-Law Relationships of the Internet Topology. *Computer Communications Review, 29,* 251-262.

11. Festinger, L., Schachter, S., & Back, K. (1950). *Social pressure in informal groups*. New York: Harper.

12. Guare, J. (1999). *Six Degrees of Separation*. New York, NY: Vintage Books.

13. Hierholzer, C. (1873). Über die Möglichkeit, Einen Linienzug ohne Wiederholung und ohne Unterbrechnung zu Umfahren. *Mathematische Annale, 6*, 30-32.

14. Kleinberg, J. M. (2000). Navigation in a small world. *Nature*, *406*, 846.

15. Krackhardt, D., & Stern, Robert N. (1988). Informal Networks and Organizational Crises: An Experimental Simulation. *Social Psychology Quarterly, 51*(2), 123-140.

16. Milgram, S. (1967). The small world problem. *Psychology Today*, *1*(2), 60-67.

17. Russo, V. J., Executive Director, Aeronautical Systems Center. "An Overlooked Asset: The Defense Civilian Workforce." Presentation to the Committee on Governmental Affairs Subcommittee on Oversight of Governmental Management. United States Senate. 12 May 2003.

18. Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University Press.

19. Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*, 440-442.

20. Watts, D. (1999). *Small world: the dynamics of networks between order and randomness*. Princeton, NJ: Princeton University Press.

21. West, D. B. (1996). *Introduction to Graph Theory*. Upper Saddle River, NJ: Simon & Schuster.

22. Wilson, R. J. (1986). An Eulerian Trail through Königsberg. *Journal of Graph Theory, 3*(10), 265-275.

Vita

Captain John R. Hutzel was born in Sacramento, California. He graduated from Edward S. Marcus High School in Flower Mound, Texas in the summer of 1995 and received a four-year ROTC scholarship to attend Texas A&M University in College Station, Texas. After graduating with a Bachelor of Science Degree in Mechanical Engineering, he was commissioned a Second Lieutenant in the United States Air Force and was assigned to Wright Patterson Air Force Base, Ohio in September 2000. While stationed at Wright Patterson, Captain Hutzel worked as a program manager on the B-2 program before entering the Air Force Institute of Technology School of Engineering and Management to pursue two Master of Science degrees, one in Research and Development Management and the other in Systems Engineering. His research has focused on complex systems and their application to the social networks of organizations. He has active memberships in Sigma Iota Epsilon and the Association of Old Crows.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 074-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, 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 the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to an penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
23-03-2006	Master's Thesis	Mar 2005 - Mar 2006

4. TITLE AND SUBTITLE

A Graph Theoretic Analysis of the Effects of Organizational Structure on Employee Social Networks

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Hutzel, John, R., Captain, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way, Building 640
WPAFB OH 45433-8865

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GRD/ENV/06M-07

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

N/A

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

A simulation technique was used to investigate the impacts of organizational structure on an organization's social network. By simulating personnel in an organization as vertices in a graph and the aging of the corporation as the aging of the same graph, the maturation of an organization was realized. The characteristic path length of the graph was measured after each year returning an optimistic average organizational distance. Results include the finding that, per this model, an organization's characteristic path length can drop over 50% in a 20 year period with consideration of edges of all strengths. Next a series of random searches were performed to measure the ability of an individual to search for information. A concern of this research is the difference between a partitioned and a non-partitioned organization's social network. A partitioned organization is defined as one in which personnel are restricted to smaller communities during a career. Partitioning the organization decreases the travel distance, in this model, when searching within a smaller community by 6% and it increases the travel distance when searching within the larger community by 12%. This must be considered along with the organization's information flow needs when structuring an organization.

15. SUBJECT TERMS

Graph Theory, Social Network, Small World, Network Analysis, Communications Networks, Knowledge Management, Social Communication

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dennis Strouble, AFIT/ENV
					19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	98	(937)255-3355x3323 (dennis.strouble@afit.edu)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18