

Detecting Change in Human Social Behavior Simulation

Ian McCulloh and Kathleen M. Carley

August 19, 2008
CMU-ISR-08-135

Institute for Software Research
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213



Center for the Computational Analysis of Social and Organizational Systems
CASOS technical report.

This research is part of the Dynamics Networks project in CASOS (Center for Computational Analysis of Social and Organizational Systems, <http://www.casos.cs.cmu.edu>) at Carnegie Mellon University.

This work was supported in part by:

- The Army Research Institute for the Behavioral and Social Sciences, Army Project No. 611102B74F
- The Office of Naval Research (ONR), United States Navy Grant No. N00017-02-10973 on Dynamic Network Analysis
- The Army Research Labs Grant No. DAAD19-01-2-0009
- The Air Force Office of Sponsored Research (MURI: Cultural Modeling of the Adversary Org. 600322)
- Additional support on measures was provided by the DOD and the NSF IGERT 9972762 in CASOS.
- The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the National Science Foundation or the U.S. government.

We are grateful to Brian Hirschman from CMU for discussing the multi-agent simulation with us.

Keywords: Social network change detection, statistical process control, multi-agent simulation, military, organizational behavior, networks, network statistics

Abstract

The performance of social network change detection (SNCD) is evaluated using a multi-agent simulation of company level U.S Army Infantry organizations. Agent interaction is probabilistic, with increased likelihood of communication based on similarity in skills, role, sub-unit of assignment, military rank, and general personality homophily. Various social network measures are monitored for change over time with a Cumulative Sum (CUSUM) control chart, an Exponentially Weighted Moving Average (EWMA), a scan statistic, and a Hamming Distance. Findings show that the average betweenness, the average closeness, and the standard deviation of eigenvector centrality are social network measures that are well-suited for SNCD. This research further supports the efficacy of SNCD using statistical process control charts.

Table of Contents

1	Introduction	Error! Bookmark not defined.
2	Method.....	2
2.1	Docking	3
2.2	Verification and Validation	5
2.3	Virtual Experiment.....	6
2.4	The Hamming Distance.....	8
2.5	The Cumulative Sum Control Chart	8
2.6	The Exponentially Weighted Moving Average Control Chart	9
2.7	Scan Statistics.....	10
3	Results	10
3.1	Isolation of Headquarters	10
3.2	Loss of Subordinate Unit.....	14
3.3	Addition of New Subordinate Element	16
3.4	Sporadic Communication.....	17
4	Conclusion.....	18
5	References	20

1 Introduction

Social network analysis (SNA) has become an important analytic tool for friendly command and control structures. SNA quantifies relationships between individuals and represents them in a network. The mathematical, statistical, and qualitative analysis focuses on the relationships between individuals instead of traditional analysis that focuses on the attributes of individuals. Important factors such as homophily, geographic distance, collective belief formation, access to knowledge, resources, and events are all considered in a more tractable form through SNA (Wasserman and Faust, 1994).

This type of analysis continues to increase in importance to the U.S. military as it evolves into a more highly connected organization (McCulloh, et. al., 2007). Subordinate leaders and soldiers at the lowest level, now have access to cellular phones, email, the internet, among other sources of information and communication. This technology allows enhanced situational awareness and enables soldiers to solve problems at a much lower level in the chain of command. Unfortunately, handling these problems at a lower level isolates senior leaders from much of the decision process. The experience, judgment, and responsibility of these senior leaders are effectively divorced from many important problems. SNA offers senior leaders a valuable tool for identifying potential problems and for monitoring communication within their organization.

An important SNA tool for monitoring an organization is Social Network Change Detection (SNCD) (McCulloh, et. al. 2008). SNCD applies methods of statistical process control used in quality engineering to detecting changes in social networks. This approach was demonstrated to be effective in detecting change in a group of 24 graduate students as they prepared for their comprehensive exam as well as on a data set of Al-Qaeda members (McCulloh, Carley, and Webb, 2007). While SNCD has been effective in detecting change in a limited number of real-world data sets, the limitations and performance of this approach applied to different network measures has not been explored.

Simulation offers a controlled environment to explore the performance and limitations of SNCD. It is often very subjective to determine the true causes of change in real-world data. In the real-world, potential factors have high covariance, and there are multiple and competing causes for different types of change. For example, a major shift in Al-Qaeda was detected in 1997 (McCulloh, Carley, and Webb, 2007). This corresponds to the time that they were expelled from the Sudan, re-established bases in Afghanistan, and began to form the Islamic Front. If the change point had been identified as 1998, other causes could be identified. In 1998, Al-Qaeda bombed embassies in Tanzania and Kenya, and formed the Islamic Front (Marquand, 2001). In 1993, they were successful against the US in Somalia. Which of these events is a real change? With simulation, a virtual organization can experience a deliberate controlled change, and then the performance of an SNCD method can be objectively assessed. Replications can provide investigators with an estimate of the variance of performance metrics. When the

performance of SNCD is assessed and quantified in simulation, its usefulness and reliability for the real-world can be estimated.

This paper has six remaining parts. In the method section the logic of the simulation model is explained. The simulation is then aligned or docked with comparable models in the literature. Next, verification and validation are discussed, followed by several virtual experiments to simulate network change. A brief overview of SNCD methods is presented. The results section highlights SNCD performance before the conclusion section.

2 Method

A virtual experiment is set in the context of a military organization. Military units have a formal hierarchical chain of command as well as informal leaders and social relationships that extend beyond formally designated units. Both the formal and informal networks are used to share situational awareness, experience, skill development, and resources. Isolation of certain individuals or subordinate elements within a unit, due to radio failure, enemy attack, or poor coordination, can cause serious impacts to the unit's performance. Reconnaissance units, especially Special Forces, will also sporadically communicate with military units as they drop off the network to conduct their mission and enter the network to report information or request resources. The network can also experience change when new personnel or groups join the unit either permanently or for a particular phase of a military operation. The ability of SNCD to detect these types of change will be explored in a virtual experiment.

Using the social simulation program, *Construct* (Carley, 1990; Carley 1995; Schriber and Carley, 2004), military units of varying size are simulated. A variety of changes will be introduced to the network at a known point. The Cumulative Sum (CUSUM) (Page, 1961), Exponentially Weighted Moving Average (EWMA) (Roberts, 1959), and Scan Statistic (Fisher and Mackenzie, 1922), statistical process control charts will be applied to several social network graph level measures taken on the network at each time step. The number of time steps between the actual change and the time that an SNCD method signals a change will be recorded as the Detection Length. The Average Detection Length (ADL) over multiple independently seeded runs is then a measure of the SNCD method's performance. The ADL will be compared for different changes and different SNCD parameters.

The basic military structure that will be simulated is an infantry training model. This is the most basic US military unit and is used for training soldiers and officers across the US Army Training and Doctrine Command (HQ, Dept of the Army, 1992). An organizational diagram is shown in Figure 1. Within this model, soldiers are organized into four man teams. Two teams and a squad leader form a 9 man squad. Three squads and a three person headquarters form a 30 man platoon. Three platoons and a 10 person command post form a company. Each soldier is trained in various skills that are

distributed throughout the organization. Each team for example will have an automatic gunner, a grenadier and two riflemen. One member on a team will also be trained as a medic, another in demolitions, and two will be able to search enemy prisoners of war. Each soldier possesses individual skill in stealth, situational awareness, physical fitness, intelligence, military rank, and motivation. Homophily in these individual skills create stronger bonds between members of a unit which will increase their probability of communication. Organizational proximity will also affect communication, with individuals in the same sub-unit being more likely to communicate. The objective of the simulation will be to model communication within the military unit.

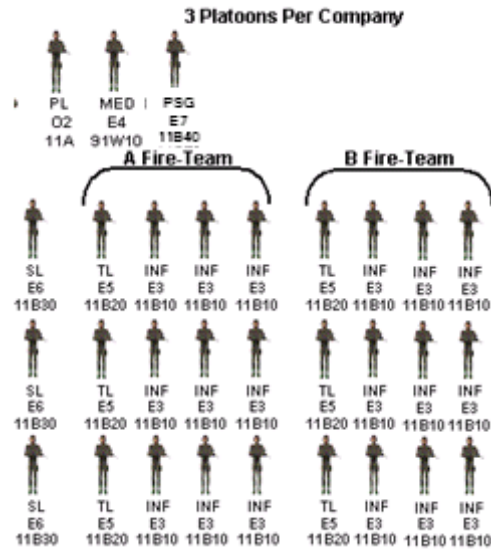


Figure 1. Platoon Organizational Diagram.

2.1 Docking

Other simulation studies have been performed on communication within military organizations (Kilduff, et.al., 2006; Rahimi, S., et.al.,). The Future Combat System (FCS) initiative has made extensive use of simulation to evaluate many system design considerations to include communication flow. The U.S.Army Research Laboratory's Human Research and Engineering Directorate (ARL-HRED) uses a simulation tool called C3TRACE (command, control, and communication: techniques for reliable assessment of concept execution). C3TRACE is a simulation environment that models organizations of varying sizes as they complete simulated tasks under various levels of workload. One study modeled communication within an infantry company very similar to the object of this paper's study. Some differences were introduced as part of the FCS revised personnel manning concept. The model equipped soldiers with differing communications devices to evaluate the impact on communication flow. The ARL-HRED focused on performance measures such as utilization, dropped messages, and decision quality. Their study did not look at the communication network however. The relationships between

the modeled agents were not investigated. Change detection was not applied to the communication patterns in the simulated organization.

Another study investigated wireless platoon communication alternatives through simulation (Rahami, Mohamed, and Paredes, 2007). The focus of the study was on designing robust communication network topologies. While this study considered the relationships between simulated agents, they were essentially static. The model was a discrete event simulation created in Arena, which is better suited for process flow than it is for modeling relational network data. More importantly, the simulation did not explore dynamic changes over time.

The C3TRACE and Arena simulation models provide some insight into modeling a military organization, but they are not well suited to modeling dynamic network change over time. *Construct* was specifically designed to model relational network data and evolve it over time. In addition, *Construct* has the ability to vary the interactions of simulated agents based on their relative homophily, knowledge, and expertise. Dynamic social networks are easily exported to social network analysis software, where change detection can be evaluated. *Construct's* ability to realistically simulate the social dynamics of an organization over time and provide network representations of communication makes this model uniquely well suited for SNCD exploration. An extensive search of the literature did not reveal any other relevant, similar models. Table 1 shows a comparison or docking of the three simulations.

While the three simulations are similar in terms of the organization being modeled, their objectives are different. The C3TRACE and Arena models are focused on measuring and improving unit performance. The purpose of this study is to realistically model the evolving social/communication network of an Army unit over time, so that methods of SNCD can be objectively compared to each other. In addition, *Construct* is much more sophisticated than the other methods at modeling how humans actually interact in an organization. This level of detail in modeling provides a much more rigorous test of the success of SNCD.

Table 1. Docking: *Construct*, C3TRACE, Arena

	<i>Construct</i>	C3TRACE	Arena
Simulated Organizations	Squad, Platoon, Company HQ	Squad, Platoon, part of Company HQ	Squad, Platoon
Size of Squad	9 men	9 men	9 men
Size of Platoon	30 men	49 men	undefined
Agent Details	Knowledge, expertise, beliefs, resources	Information quality, expertise	Knowledge
How agents interact	Uses an interaction sphere. Probability based on homophily or expertise.	Fixed in advance, based on doctrine	Interaction Strength is proportional to distance between agents
Output	Social network measures for each time step (Table 3).	Soldier utilization, soldier performance, and decision quality	Successful communication
Type of Simulation	Multi-Agent	Multi-Agent	Discrete Event
Virtual Experiment	Inject a change at a specified time point, to measure ARL	Change the agent interactions that will occur	No virtual experiment

2.2 Verification and Validation

Verification and validation will focus on the communication dynamics within the organization. The input to the model is abstract communication. The focus of the simulation is information sharing within a military organization. Issues of varying social capital for differing information is not considered in this model. Future research could investigate the effect of the importance of the information on communication dynamics. That type of investigation is beyond the scope of this study. Therefore, the model inputs are very simple.

There are several data sources that can be used to validate the process component of the simulation model. The author has served as an instructor at the US Military

Academy at West Point. During summer cadet training, he collected social network data on communication between soldiers ranging from team member to the company level. Data was collected for over 20 training missions from seven different company units. The communication within the simulation should be within the range of the real-world data. Simulated communication was also presented to four Army subject matter experts (Johnson, 2008; Gauthier, 2008; Smith, 2008; Trent, 2008) with recent combat experience to provide qualitative validation of model accuracy. The expert input was qualitative, where the soldiers were asked if a random sample of baseline (no change imposed) social networks appeared to reasonably describe communication patterns in an Army Infantry unit. Adjustments were made to weights placed on socio-demographic variables such as rank and job title to accurately reflect military communication. All four experts validated the final model.

The simulation output is validated by calculating several graph level social network measures (see Table 2) for the baseline simulation and comparing those results to the data collected on cadet summer training. There was no statistically significant difference in the average betweenness, average closeness, or density. Unfortunately, more detailed information was not available. The output validation coupled with subject matter expert review provides reasonable evidence of the model's accuracy.

2.3 Virtual Experiment

A virtual experiment is conducted using the *Construct* Infantry Model to provide a realistic data set for evaluating SNCD methods. Three different size infantry units (squad, platoon, and company) are simulated for 500 time periods. In these units, four changes are introduced. This creates 9 independent data sets that can be used to evaluate SNCD performance. Three of the changes are not feasible for the squad size element. The four network changes correspond to common military communication problems that might affect an infantry unit.

The first type of network change is the isolation of the Headquarters section. For a squad, this is simply the squad leader. For a platoon, this consists of the Platoon Leader, Platoon Sergeant, and the Radio Telephone Operator (RTO). For a Company, this includes the 10 person command post, also known as the headquarters element. A military headquarters is most often isolated from the rest of the unit as a result of radio failure or a deliberate attack from enemy forces. This is perhaps one of the most significant changes that commonly happen in a military situation, as it requires a rapid and efficient transfer of command and control, as the formal hierarchy is significantly adjusted. In the simulation, this is modeled by isolating the Headquarters section beginning at time period 20. These individuals remain isolated for the remainder of the simulation. Network measures are calculated on the organization for all time periods.

Another significant change in a military organization is the loss of a subordinate element. A subordinate element might be lost as a result of a task organization change, radio failure, or enemy attack. This change is not modeled for the infantry squad, since

this would mean losing half of the organization. For the Platoon, this change is modeled by isolating a squad at time period 20 for the remainder of the simulation. For the Company, this is also modeled by isolating a squad at time period 20 for the remainder of the simulation. While it is conceivable to isolate any number of individuals in the simulation, these changes are used to demonstrate the performance of the SNCD methods. Perhaps SNCD methods that have similar performance could be evaluated under greater conditions of change in a future paper. For now, it is beyond the scope of this paper to exhaustively address all conceivable types of network change.

A similar change is the addition of a new subordinate element. This is usually a result of a task organization change. This is modeled by adding a squad in both the Company and Platoon level models. It is not modeled for a squad, because squad organizations are not usually capable of managing an additional subordinate element. Again, this simple change is used to evaluate SNCD and not meant to be an exhaustive comparison of different types of organizational change.

The final type of change simulated, is sporadic communication. Sporadic communication can be either deliberate, or unplanned. An example of deliberate sporadic communication is a reconnaissance operation, where radio power must be conserved and noise discipline is important. An example of unplanned sporadic communication is radio failure. This is modeled in the simulation by introducing a squad from time period 20 to time period 30. Network measures will be recorded throughout the simulation. This change is only modeled for the Platoon and Company level simulations.

The social network measures listed in Table 2 are measured for every simulated network. Table 3 illustrates the combinations of the virtual experiment. The outputs of the simulation are the graph level measures recorded for each simulated time step. Different SNCD methods are then used to identify possible changes in the network over time.

Table 2. Social Network Measures

Average Betweenness	Standard Deviation of Closeness
Maximum Betweenness	Average Eigenvector Centrality
Standard Deviation of Betweenness	Maximum Eigenvector Centrality
Average Closeness	Minimum Eigenvector Centrality
Maximum Closeness	Standard Deviation of Eigenvector

Table 3. Virtual Experiment

Variable	Number / Nature of Values	Values
Network Size	3	9, 30, 100
Network interaction sphere is fully connected. Edge probability determined by simulation.		
Type of Change in Network		
Isolation of leadership	2	Isolated headquarters after 20 time periods
Sporadic communication (Reconnaissance)	2	Initially absent, present for 10 time periods, then absent for remainder of simulation (omitted for squad)
Loss of subordinate unit	2	Removal of the immediate subordinate unit after 20 time periods (omitted for squad)
Gain an attached unit	2	Addition of a squad after 20 time periods. (omitted for squad)
Cells	18	3 Network sizes x 4 Changes x 2 Levels – Squad omissions
Replications	25	
Independent Runs	450	

2.4 The Hamming Distance

The Hamming distance is a measure of change between two networks (Hamming, 1950). For an unweighted network, where edges are either present or absent, the Hamming distance can be used to measure the similarity of two networks. The Hamming distance is calculated by counting the percentage of possible edges that are different between two networks. The Hamming distance is recorded between sequential time steps of simulated networks. It is possible that a significant change in Hamming distance may indicate network change over time.

2.5 The Cumulative Sum Control Chart

Statistical process control charts (Montgomery, 1996; Page, 1954, 1961) are used to detect changes in temporal data. The cumulative sum (CUSUM) control chart statistic has been shown to be effective in detecting changes in real-world social networks (McCulloh et. al, 2007; McCulloh et. al, 2008). With control charts helping to distinguish process abnormality, graph level social network measurements from the simulation are recorded and used to compute a test statistic. When the test statistic exceeds the limits of the control chart, the process is deemed abnormal. This indicates that a change in the process

may have occurred. The process (in this case organizational behavior) can then be investigated to identify the potential cause of the change. The CUSUM statistic is given by,

$$C_t^+ = \max \left\{ 0, Z_t - \frac{\delta}{2} + C_{t-1}^+ \right\}$$

where Z_t is the normalized social network measure, and δ is the magnitude of change that the CUSUM is optimized to detect. For this study, $\delta = 1$ for all calculations. A control limit of 4.0 is used, which corresponds to a type I error of approximately 0.01. Therefore, this SNCD scheme would signal a false alarm every 100 observations on average, when used on an organization that does not experience change. The type I error rate was estimated using Monte Carlo simulation of a baseline *Construct* Infantry Model (no change) with 2,500 independently seeded runs. The standard deviation of the error estimate is 0.0002.

2.6 The Exponentially Weighted Moving Average

The exponentially weighted moving average (EWMA) was introduced by Roberts (1959) for monitoring changes in the mean of a process over time. The EWMA associated with subgroup t is $w_t = \lambda \bar{x}_t + (1 - \lambda)w_{t-1}$, where $0 < \lambda \leq 1$ is the weight assigned to the current subgroup average and $w_0 = \mu_0$. Common values of λ are $0.1 \leq \lambda \leq 0.3$. Having observed a total of T subgroups, the statistic w_T is plotted against the control limits

$$\mu_0 \pm L \sigma_{\bar{x}} \left(\frac{\lambda}{2 - \lambda} \left[1 - \lambda^{2T} \right] \right)^{1/2}$$

where L is a constant that scales the width of the control limits.

Lucas and Saccucci (1990) investigated the impact of different combinations of L and λ on the ADL performance of the EWMA control chart. The combinations that were investigated were chosen such that the each chart was calibrated for the same sensitivity to false alarm. They found that EWMA charts with small values of λ perform well at detecting small changes in a process mean. Conversely, EWMA charts with large values of λ perform well at detecting large changes in a process mean. Hunter (1986) and Montgomery (1996) investigated the ADL performance of the EWMA chart and concluded that it is similar to the ADL performance of the CUSUM chart. For this study, parameter values of $\lambda = 0.1, 0.2,$ and 0.3 are used. The control limit parameter is set to $L = 2.5, 2.6,$ and 2.7 respectively. These combinations of L and λ are calibrated to have the same sensitivity to false alarm as the CUSUM settings indicated above based on Monte Carlo simulation as described above.

2.7 Scan Statistics

Scan statistics (Fisher and Mackenzie, 1922; Naus, 1965; Priebe, et. al., 2005), also known as moving window analysis, investigates a random field for the presence of a local signal. A small window of observations is used to calculate a local statistic. In this paper a window size of 7 observations proceeding the current time period is used, and the window mean is used for the local statistic. If the statistic exceeds a control limit, then inference can be made that a change in the random field may have occurred. A control limit of 1.0 is used to calibrate the scan statistic to have the same sensitivity to false alarm as the CUSUM and the three EWMA approaches. The control limit was approximated using Monte Carlo simulation as described above.

3 Results

3.1 Isolation of Headquarters

Investigating the isolation of the headquarters element in three different organizations will provide insight into how the network size affects the performance of change detection measures. In each organization, 30 man platoon, 100 man company, and 9 man squad; 10% of the network will be removed. In a sense, the magnitude of change is the same, however, the network size is different.

The isolation of the platoon headquarters is modeled by removing the three headquarters members at time period 20 for the duration of the simulation. Social network measures are recorded for all time periods. There was no significant change in the Hamming distance. Table 4 displays the ADL performance of the SNCD methods. It can be seen that the average of the betweenness is a better measure to use for SNCD than either the maximum of the standard deviation of betweenness. This is generally true for all magnitudes of change and sizes of organization investigated. For the closeness measure, both the maximum closeness and average closeness generally outperform the standard deviation of closeness. However, for a EWMA with $r = 0.3$, the maximum closeness measure has relatively poor performance. This might suggest that the average closeness measure is a more robust measure of change detection. In a single variant application of the EWMA, the parameter, r , makes the control chart more or less sensitive to a particular magnitude of change (Lucas and Saccucci, 1990; McCulloh, 2004). It is reasonable to consider that for the isolation of a platoon headquarters, the maximum closeness EWMA with $r \leq 0.2$ is sensitive to detecting the change, yet the maximum closeness EWMA with $r \geq 0.3$ is less sensitive. This will be explored with other magnitudes and types of changes throughout the paper. For eigenvector centrality, the maximum eigenvector centrality and the standard deviation of eigenvector centrality appear to be more sensitive measures of change detection than the average or minimum of the eigenvector centrality. It also appears that the eigenvector centrality measures dominate all other measures for performance in this case.

Table 4. ADL Performance of SNCD on Isolation of Platoon Headquarters

	CUSUM $k = 0.5$	EWMA $r = 0.1$	EWMA $r = 0.2$	EWMA $r = 0.3$	Scan Statistic
Average Betweenness	9.32	8.24	10.16	11.52	6.76
Maximum Betweenness	14.36	14.72	15.72	17.08	13.24
Std Deviation Betweenness	16.44	16.24	16.92	18.52	15.24
Average Closeness	10.68	9.08	13.60	17.52	10.48
Maximum Closeness	8.76	6.00	10.60	37.96	8.64
Std Deviation Closeness	34.48	34.72	34.52	35.68	27.08
Average Eigenvector Cent	31.28	31.28	31.28	31.28	24.00
Minimum Eigenvector Cent	14.36	14.36	14.28	15.56	14.88
Maximum Eigenvector Cent	5.24	5.40	5.80	7.52	4.00
Std. Dev Eigenvector Cent	5.92	4.88	6.40	6.96	3.64

Statistical process control is a powerful statistical method for detecting the change. Figures 2 and 3 show the average betweenness score for a baseline simulation run (no change) and one of the simulation runs with the headquarters isolated. It can be seen that there is no dramatic difference between the figures. Figures 4 and 5 show the CUSUM statistic value for the baseline simulation run and the simulation run with the headquarters isolated. The dramatic difference in the plots can clearly be seen, paying attention to the values on the y-axis of the plots.

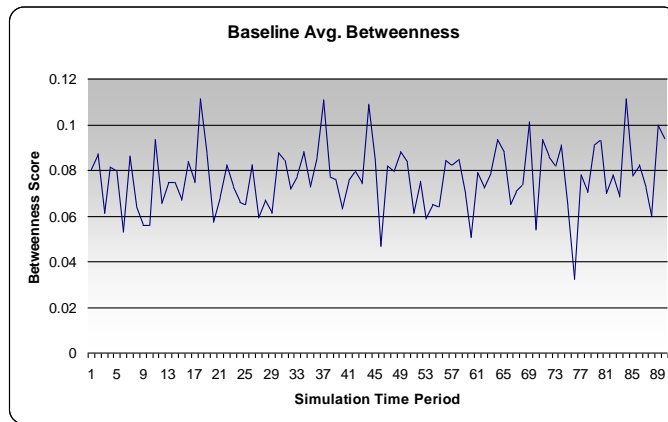


Figure 2. Baseline Betweenness Score

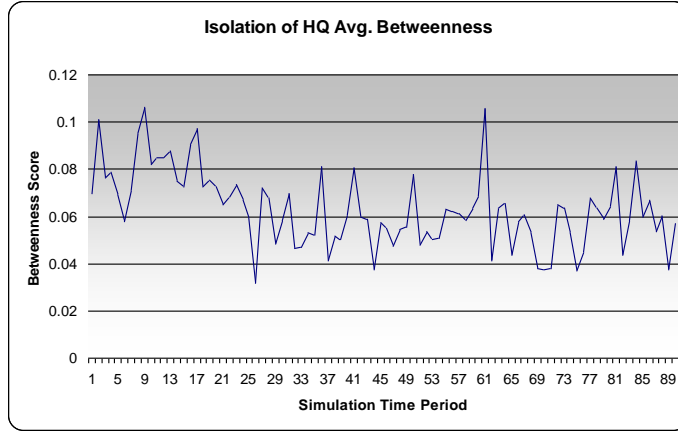


Figure 3. Isolation of HQ Betweenness Score

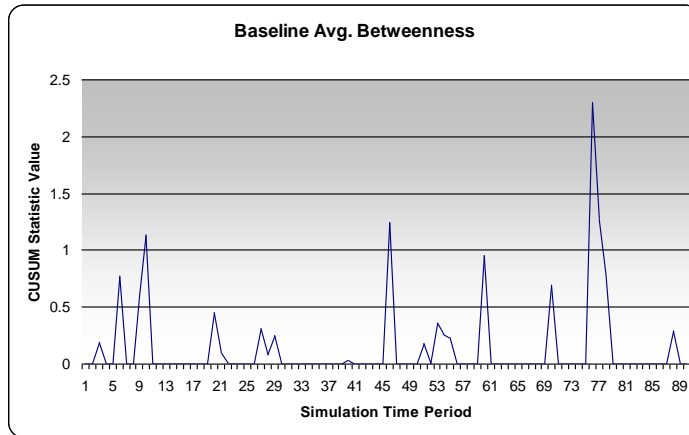


Figure 4. Baseline CUSUM Statistic Value

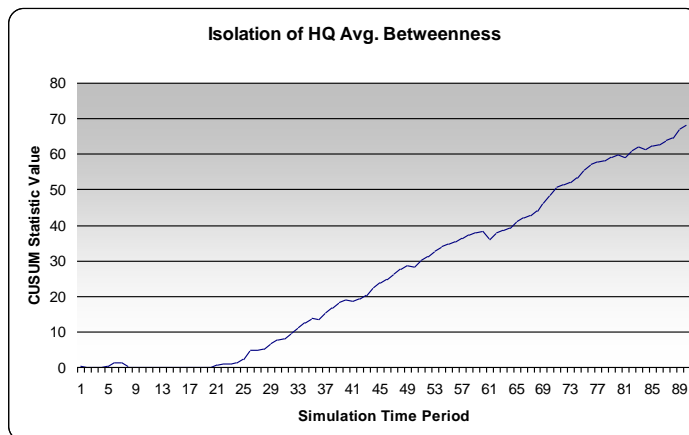


Figure 5. Isolation of HQ CUSUM Statistic Value

The plot in Figure 5 clearly shows a sharp and sudden increase beginning at time period 20, which is when the isolation of the HQ element occurs. There is a similar

performance for other types of change imposed on the network, and other SNCD schemes that are used. The CUSUM is simply used to illustrate the power of the general change detection approach. Other magnitudes and types of change will be compared by simply reporting the ADL from when a change occurs until the SNCD scheme signals.

The isolation of the Company Headquarters was modeled by removing the 10 soldier headquarters section at time 20 for the remainder of the simulation. This is very similar to the platoon example, in that 10% of the organization is removed. Social network measures are again recorded for all time periods. The Hamming distance was not effective at signaling a change in the network. Table 5 displays the ADL performance of each of the SNCD methods applied to the 100 node network. Again, it can be seen that the average of the betweenness is a more effective measure of change detection than the maximum or the standard deviation of betweenness. The performance of the closeness measures behave as they did in the case of platoon headquarters isolation. In this case, the maximum eigenvector centrality does not appear to be as effective of a measure for detecting change as does other measures. However, the standard deviation of eigenvector centrality still dominates all other measures for change detection performance.

Table 5. ADL Performance of SNCD on Isolation of Company Headquarters

	CUSUM $k = 0.5$	EWMA $r = 0.1$	EWMA $r = 0.2$	EWMA $r = 0.3$	Scan Statistic
Average Betweenness	11.16	11.08	10.20	13.48	6.96
Maximum Betweenness	17.32	17.76	18.20	20.12	13.72
Std Deviation Betweenness	18.08	19.40	20.88	22.52	17.36
Average Closeness	11.16	9.44	12.52	15.64	9.40
Maximum Closeness	10.44	9.72	12.64	51.76	9.60
Std Deviation Closeness	41.88	39.48	42.20	43.44	40.76
Average Eigenvector Cent	35.84	36.72	34.84	34.84	29.24
Minimum Eigenvector Cent	16.00	17.96	17.88	16.76	13.60
Maximum Eigenvector Cent	26.40	30.76	29.64	29.24	25.44
Std. Dev Eigenvector Cent	10.40	10.72	9.36	9.48	6.44

The isolation of squad leadership was modeled by removing the squad leader at time 20 for the remainder of the simulation. This is also similar in that 11% of the organization is isolated. Table 6 shows the SNCD performance at the squad level, 9 node network. It is not clear that certain measures perform better than others for change detection in the 9 node network. It appears that the measures of average betweenness, average closeness, and the standard deviation of eigenvector centrality become better measures of network change as the size of the network increases. However, they do not necessarily perform worse on a small network. While an extensive study of the

sensitivity of each measure to the network size is beyond the scope of this paper, it holds the promise of fruitful future research.

Table 6. ADL Performance of SNCD on Isolation of Squad Leader

	CUSUM $k = 0.5$	EWMA $r = 0.1$	EWMA $r = 0.2$	EWMA $r = 0.3$	Scan Statistic
Average Betweenness	16.12	15.76	16.32	17.92	12.32
Maximum Betweenness	16.64	17.40	19.52	18.56	11.56
Std Deviation Betweenness	17.68	17.76	18.20	18.72	12.08
Average Closeness	15.16	15.84	16.48	15.60	11.72
Maximum Closeness	18.72	19.60	18.68	23.80	14.32
Std Deviation Closeness	16.20	16.08	15.52	16.24	12.88
Average Eigenvector Cent	24.12	24.12	24.12	24.12	15.12
Minimum Eigenvector Cent	17.84	18.48	17.04	18.08	12.36
Maximum Eigenvector Cent	19.36	21.56	20.56	20.56	13.84
Std. Dev Eigenvector Cent	17.08	18.72	18.36	17.44	12.36

3.2 Loss of Subordinate Element

The loss of a subordinate element provides insight into how the magnitude of change affects change detection performance. For the 30 man platoon and the 100 man company, a nine man squad is isolated. This represents 30% of the platoon and 9% of the company. This change is obviously not feasible for the nine man squad, since it would involve removal of the entire organization.

The infantry platoon had one squad removed from the simulation at time period 20, for the remainder of the simulation. Social network measures were recorded for each time period. There were no significant or marginally significant changes in Hamming distance between sequential networks. The ADL for each measure is reported in Table 7. Again, it can be seen that the average of the betweenness outperforms other betweenness measures. The closeness measures perform as in previously investigated cases. The minimum eigenvector centrality outperforms the maximum eigenvector centrality for most of the SNCD schemes for this particular type and magnitude of change. The standard deviation of eigenvector centrality still outperforms other eigenvector centrality measures, however, it is no longer dominates all other measures.

Table 7. ADL Performance for Loss of Subordinate Element in a Platoon

	CUSUM $k = 0.5$	EWMA $r = 0.1$	EWMA $r = 0.2$	EWMA $r = 0.3$	Scan Statistic
Average Betweenness	6.96	6.00	8.68	12.16	8.12
Maximum Betweenness	9.52	7.44	11.12	13.24	7.80
Std Deviation Betweenness	9.16	7.40	9.48	12.72	6.84
Average Closeness	9.64	8.36	12.72	19.28	11.40
Maximum Closeness	9.32	9.16	12.36	31.56	9.52
Std Deviation Closeness	18.96	16.44	19.40	26.24	17.04
Average Eigenvector Cent	29.36	29.36	29.36	29.36	20.60
Minimum Eigenvector Cent	10.08	9.64	12.24	12.60	10.28
Maximum Eigenvector Cent	11.72	12.04	11.88	20.60	10.84
Std. Dev Eigenvector Cent	8.48	6.28	9.80	10.44	6.88

The Infantry Company also had one squad removed at time 20 for the remainder of the simulation. The results for the Company network are shown in Table 8. It generally takes longer to detect the changes in the Company network. This was also observed in the isolation of the headquarters. This implies that the size of the network could impact the speed of change detection. The average betweenness, average closeness, and standard deviation of eigenvector centrality appear to outperform other measures for change detection performance. The maximum closeness measure dominates other measures in all cases except for the EWMA with $r = 0.3$.

Table 8. ADL Performance for Loss of Subordinate Element in a Company

	CUSUM $k = 0.5$	EWMA $r = 0.1$	EWMA $r = 0.2$	EWMA $r = 0.3$	Scan Statistic
Average Betweenness	13.64	11.72	13.80	20.60	12.68
Maximum Betweenness	23.80	19.64	23.80	30.72	25.44
Std Deviation Betweenness	24.84	18.12	24.96	25.52	22.04
Average Closeness	9.72	7.4	13.44	14.96	9.80
Maximum Closeness	6.92	4.92	7.48	53.16	6.32
Std Deviation Closeness	45.44	47.92	47.96	50.88	43.68
Average Eigenvector Cent	34.72	36.60	34.72	34.72	30.64
Minimum Eigenvector Cent	18.68	19.96	19.64	23.88	18.32
Maximum Eigenvector Cent	18.28	25.80	25.00	27.20	25.88
Std. Dev Eigenvector Cent	9.52	9.92	11.88	15.32	8.72

3.3 Addition of New Subordinate Element

Another type of change is the addition of a new subordinate element. A squad is added to both the 30 man platoon and the 100 man company.

The infantry platoon had one squad that was not present initially, and added at time period 20. Social network measures were calculated for each time period. SNCD methods were applied to the data. Results are shown in Table 9. Although the speed of change detection is much faster for this type of change, the same performance trends are seen as before. For betweenness measures, the average outperforms the maximum or the standard deviation. The average closeness and maximum closeness measure perform well, however, the maximum closeness does not perform well with an EWMA $r = 0.3$ scheme. The standard deviation of eigenvector centrality almost completely dominates other measures.

Table 9. ADL Performance for Addition of Subordinate Element in a Platoon

	CUSUM $k = 0.5$	EWMA $r = 0.1$	EWMA $r = 0.2$	EWMA $r = 0.3$	Scan Statistic
Average Betweenness	1.60	1.52	1.68	1.72	1.00
Maximum Betweenness	2.32	2.16	2.20	2.00	1.00
Std Deviation Betweenness	2.36	2.36	2.40	2.24	1.00
Average Closeness	1.48	1.52	1.56	1.52	1.00
Maximum Closeness	1.24	1.28	1.20	5.00	1.00
Std Deviation Closeness	3.44	4.60	4.20	3.48	2.64
Average Eigenvector Cent	31.76	31.76	31.76	31.76	25.56
Minimum Eigenvector Cent	6.24	5.6	6.16	6.80	4.20
Maximum Eigenvector Cent	4.52	4.88	4.80	4.80	3.56
Std. Dev Eigenvector Cent	1.16	1.60	1.24	1.24	1.00

The company model had a squad added at time period 20 for the remainder of the simulation. Again the platoon level performance is better than the company level performance, shown in Table 10. The average betweenness, average closeness, and maximum closeness all perform well at detecting the change. Surprisingly, the standard deviation of eigenvector centrality is not an effective measure for this type and magnitude of change.

Table 10. ADL Performance for Addition of Subordinate Element in a Company

	CUSUM $k = 0.5$	EWMA $r = 0.1$	EWMA $r = 0.2$	EWMA $r = 0.3$	Scan Statistic
Average Betweenness	9.64	9.52	9.84	10.28	5.04
Maximum Betweenness	14.52	16.96	15.80	17.44	12.16
Std Deviation Betweenness	12.88	13.16	13.32	14.56	8.92
Average Closeness	5.32	5.8	5.36	5.24	1.44
Maximum Closeness	4.24	5.12	4.48	6.04	1.04
Std Deviation Closeness	10.40	18.52	12.96	12.32	10.00
Average Eigenvector Cent	35.56	37.04	38.64	37.60	30.24
Minimum Eigenvector Cent	38.16	39.32	38.04	40.84	36.40
Maximum Eigenvector Cent	30.20	33.48	34.44	29.52	30.92
Std. Dev Eigenvector Cent	33.88	33.72	37.80	44.48	33.96

3.4 Sporadic Communication

Sporadic communication was modeled with a squad communicating from time period 20 to time period 30 only. It can be seen in Table 11 that the performance of different measures is much more similar than in previous types of change. It is also interesting that all of the ADL values are greater than 10, which means that the change was detected after the organization returned to its original state. This might be a result of the SNCD statistic being moved closer to the decision interval from time period 20 to time period 30. When the organization returned to its original state, the statistic is much closer to the decision interval than it was before the change occurred. Therefore, the statistic is much more likely to signal a false alarm after the sporadic change than it was before the sporadic change. This increased sensitivity can therefore provide an alert that a sporadic change may have occurred. Another interesting affect of the sporadic change is that the Hamming distance was more effective at detecting this change than it was for detecting other types of change. For 73% of the simulation runs the Hamming distance ADL was 8.00. For the other 27% of the simulation runs the Hamming distance did not signal a change. The increased performance of the Hamming distance for sporadic changes might be due to the fact that there are essentially two changes of equal magnitude. At time period 20, a squad was added, and at time period 30 it was removed. The Hamming distance only looks at the difference between the networks of two consecutive time periods. With two changes, the Hamming distance is twice as likely to detect the change. Unfortunately, the Hamming distance method failed to detect change in 27% of the simulation runs. This failure rate renders the Hamming distance an unreliable change detection method by itself.

Table 11. ADL Performance for Sporadic Communication

	CUSUM $k = 0.5$	EWMA $r = 0.1$	EWMA $r = 0.2$	EWMA $r = 0.3$	Scan Statistic
Average Betweenness	15.08	14.20	16.12	17.56	17.76
Maximum Betweenness	15.24	16.52	16.88	18.24	17.84
Std Deviation Betweenness	14.28	14.80	16.04	17.40	17.48
Average Closeness	13.72	13.68	16.84	16.80	17.52
Maximum Closeness	12.44	12.16	15.32	18.32	17.20
Std Deviation Closeness	23.16	19.96	21.76	21.36	17.24
Average Eigenvector Cent	24.32	24.32	24.32	24.32	18.84
Minimum Eigenvector Cent	12.76	14.32	11.92	12.80	14.56
Maximum Eigenvector Cent	12.96	12.68	14.36	14.36	18.84
Std. Dev Eigenvector Cent	12.88	14.20	16.80	16.48	21.28

All methods of SNCD were ineffective for the Company network. The sporadic change did not persist long enough to signal a possible change in most of the runs. The squad level network was not investigated for this type of change, due to a lack of context.

4 Conclusion

Control charts are a critical quality-engineering tool that assists manufacturing firms in maintaining profitability. The ability for control charts to identify changes in social networks has been demonstrated in a limited number of real world examples. Introducing control charts to the field of network science allows one to allocate minimal resources to tracking the general patterns of a network and then shift to full resources when changes are determined.

This paper contributes to the growing body of SNCD literature in two important ways. Human social behavior simulation is used as a virtual experimentation platform. The simulation allows different SNCD schemes to be calibrated to have the same sensitivity to false alarms. Once calibrated, the performance of different schemes and measures can be readily compared. The simulation also allows a single and deliberate change to be introduced into the network. This eliminates speculation in real-world data about whether and when a change actually occurred. The virtual experiment removes collinear factors and allows the performance of SNCD to be studied more objectively.

This particular virtual experiment provides insight into certain social network measures that are effective for change detection. The average betweenness and average

closeness performed very well in the virtual experiment. These two measures were also found to be effective for detecting changes in real-world data (McCulloh, et. al., 2007; 2008; McCulloh and Carley, 2008). The virtual experiments also suggest the standard deviation of eigenvector centrality as a powerful SNCD measure. Perhaps future SNCD investigation will explore using this measure on real-world data. While the maximum eigenvector centrality was an effective SNCD measure for certain changes, it was ineffective for others. In all cases, the standard deviation of eigenvector centrality was either similar in performance to maximum eigenvector centrality or it performed better. Therefore, the recommended SNCD measures are average betweenness, average closeness, and the standard deviation of eigenvector centrality.

These results do not imply that the question of SNCD is settled. Instead, it provides early evidence that this type of change detection is effective. Some thought must be devoted to determining what types of changes and magnitudes of change are of interest for applications of SNCD. Further investigation is also required to determine the limitations of change detection in terms of network measures, SPC approaches, magnitude of change and type of change. Some networks may also contain periodic effects. Communication networks for example, may have different patterns of communication during daytime working hours than they have at night. There may be more communication during a Monday through Friday work week than on the weekends. There also may be seasonal or annual trends. Methods are required to control for this type of periodicity in network change detection.

Future research should also focus on optimizing the parameters within SNCD schemes. In the virtual experiment presented in this paper, it was seen that the efficacy of maximum closeness for the EWMA scheme, depended on the choice of the parameter, r . The ADL of a control chart scheme is also affected by the choice of parameter, as seen in the three different EWMA schemes in Tables 4-11. A more complete analysis and construction of receiver operating characteristic (ROC) curves for different types and magnitudes of change would certainly benefit from the simulation approach taken in this paper. It is likely that different parameterizations of SNCD will perform better or worse than others for certain types and magnitudes of change.

Further investigation into SNCD using multi-agent simulation will likely reveal detection schemes that may be better for detecting certain types of changes such as sudden large changes or slow creeping shifts; changes in informal leaders or the isolation of a subordinate organization. Usage of control charts on comparing models and observations should also be studied to see what specific conclusions can be obtained. Thus, the mathematical study of change detection in social networks is far from settled. These studies have a broad spectrum of possible applications in fields as diverse as monitoring terror networks to improving command and control of friendly military forces. We will likely see a great deal of research in this area over the next several years.

5 References

- Anderson, B.S., Butts, C. and Carley, K.M. (1999). The Interaction of Size and Density with Graph-Level Measures to Social Networks. *Social Networks*, 21, 239-267.
- Arquilla, J. and Ronfeldt, D. (2001) *Networks and Netwars: the future of terror, crime and militancy*. RAND, Santa Monica, CA.
- Banks, D. and Carley, K.M. (1994). Metric Inference for Social Networks. *Journal of Classification*, 11, 121-149.
- Banks, D. and Carley, K. M. (1996). Models of Social Network Evolution. *Journal of Mathematical Sociology*. 21(1-2), 173-196.
- Behrman, R. and Carley, K.M. (2004). Social Network Influences on Strategic Choices. *Proceedings of the NAACSOS 2004 Conference*. Pittsburgh, PA.
- Breiger, R. Carley, K. M. and Pattison, P. (2003). Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers. *Committee on Human Factors, Board on Behavioral, Cognitive, and Sensory Sciences*. Washington, DC: National Academy Press.
- Carley, Kathleen M. (1995). Communication Technologies and Their Effect on Cultural Homogeneity, Consensus, and the Diffusion of New Ideas. *Sociological Perspectives*, 38(4), 547-571.
- Carley, Kathleen. (2003). Destabilizing Terrorist Networks. *Proceedings of the 8th International Command and Control Research and Technology Symposium. Conference held at the National Defense War College, Washington DC. Evidence Based Research, Track 3, Electronic Publication*,
WebSite:http://www.dodccrp.org/events/2003/8th_ICCRTS/pdf/021.pdf
- Carley, Kathleen M. (2003). Dynamic network analysis. *Dynamic social network analysis: Workshop summary and papers*: 133-145. P. Pattison (Ed.), Washington D.C.: The National Academies Press.
- Carley, Kathleen M. (1990). Group Stability: A Socio-Cognitive Approach. *Advances in Group Processes*, 7, 1-44.
- Carley, Kathleen M. (1999). On the evolution of social and organizational networks. In S. B. Andrews & D. Knoke (Eds.), *Research on the Sociology of Organizations* 16, 3-30. Greenwich, CT: JAI Press.
- Carley, Kathleen & Reminga, Jeffrey & Borgatti, Steve. (2003). Destabilizing Dynamic Networks Under Conditions of Uncertainty. *International Conference on Integration of Knowledge Intensive Multi-Agent Systems*. Boston MA: IEEE KIMAS.

- Carley, Kathleen & Lee, Ju-Sung & Krackhardt, David. (2001). Destabilizing Networks. *Connections*, 24(3), 31-34.
- Fisher, R.A., and Mackenzie, W. (1922). The Accuracy of the Plating Method of Estimating the Density of Bacterial Populations, with Particular Reference to the Use of Thornton's Agar Medium with Soil Samples. *Annals of Applied Biology*, 9, 325-359.
- Frantz, Terrill & Carley, Kathleen. (2005). *A Formal Characterization of Cellular Networks*. Carnegie Mellon University, School of Computer Science, Institute for Software Research International, Technical Report CMU-ISRI-05-109.
- Cebrowski, A.K. and Garstka, J.J. (1998). Network-centric warfare: Its origin and future. *Proceedings of the U.S. Naval Institute*.
<http://www.usni.org/Proceedings/Articles98/PROcebwski.htm>.
- Gauthier, S. (2008) Interview conducted for model validation. Major Gauthier is an Assistant Professor of Systems Engineering at the U.S. Military Academy with 15 years of service as a commissioned officer. He also has extensive experience with simulations and modeling.
- Ghosh, B.K. (1970). *Sequential Tests of Statistical Hypotheses*, Addison-Wesley Publishing Company, Inc., Reading, Massachusetts.
- Hamming, R.W. (1950). Error Detecting and Error Correcting Codes. *Bell System Technical Journal*. 26(2), 147-160.
- Headquarters, Department of the Army (1992). *Field Manual 7-8, Infantry Rifle Platoon and Squad*. U.S. Army Infantry School, Ft. Benning, GA.
- Hunter, J.S. (1986). The Exponentially Weighted Moving Average. *Journal of Quality and Technology* 18, 203-210.
- Johnson, A. (2008) Interview conducted for model validation. Major Johnson is an Academy Professor of Mathematics at the U.S. Military Academy with 20 years of service as a signal soldier in positions ranging from private to commander.
- Kilduff, P.W., Swoboda, J.C., and Katz, J. (2006). *A Platoon-Level Model of Communication Flow and the Effects on Operator Performance*. Army Research Laboratory Modeling Report, MR 0656. Aberdeen, MD.
- Lorden, G. (1971). Procedures for Reacting to a Change in Distribution. *Annals of Mathematical Statistics*. 42, 1897-1908.

- Lucas, J.M. and Saccucci, M.S. (1990). Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements. *Technometrics* 32, 1-12.
- Marquand, R. (2001). The tenets of terror. *Christian Science Monitor*, 18, Oct 2001.
- McCulloh, I.A., Carley, K.M. (2008). *Social Network Change Detection*. Carnegie Mellon University, School of Computer Science, Technical Report, CMU-CS-08-116.
- McCulloh, I.A., Webb, M., Carley, K.M., Graham, J. (2008). *Detecting Changes in Social Networks Using Statistical Process Control*. Army Research Institute Technical Report 1265.
- McCulloh, I.A., Garcia, G., MacGibbon, J., Tardieu, K., Dye, H., Moores, K., Graham, J. (2007). *IkeNet: Social Network Analysis of e-mail Traffic in the Eisenhower Leadership Development Program*. Army Research Institute Technical Report 1218.
- McCulloh, I.A., Carley, K.M., Webb, M. (2007). *Social Network Monitoring of Al-Qaeda*. *Network Science*. 1, 25-30.
- Montgomery, D.C. (1991). *Introduction to Statistical Quality Control*, 2nd Edition, John Wiley and Sons, New York.
- Moon, I. and Carley, K.M. (2006). Estimating the near-term changes of an organization with simulations, *AAAI Fall Symposium*, Arlington, VA, Oct 12-15, 2006
- Moustakides, G.V. (1986). Optimal Stopping Rules for Detecting Changes in Distributions. *Annals of Mathematical Statistics* 14, 1379-1387.
- Moustakides, G.V. (1998). Quickest Detection of Abrupt Changes for a Class of Random Processes. *IEEE Transactions* 44, 1965-1968.
- Moustakides, G.V. (2004). Optimality of the CUSUM Procedure in Continuous Time. *Annals of Statistics*. 32, 302-315.
- National Research Council, (2006). *Network Science*, <http://www.nap.edu/catalog/11516.html>.
- Naus, J. (1965). Clustering of Random Points in Two Dimensions. *Biometrika*, 52, 263-267.
- Newman, M., Watts, D.J. and Barabasi, A. (2006). *The Structure and Dynamics of Networks*, Princeton University Press.

- Neyman, J. and Pearson, E.S. (1933). On the Problem of the Most Efficient Tests of Statistical Hypotheses. *Philosophical Transactions Royal Society Series A*. 231, 289-337.
- Page, E.S. (1954). Continuous Inspection Schemes. *Biometrika* 41, 100-115.
- Page, E.S. (1961). Cumulative Sum Control Charts. *Technometrics* 3, 1-9.
- Priebe, C.E., Conroy, J.M., Marchette, D.J., and Youngser, P. (2005) Scan Statistics on Enron Graphs. *Computational and Mathematical Organization Theory*, 11, 229-247.
- Rahami, S., Mohamed, H., and Paredes, N. (2007). Company Command Support System CS². *IEEE Systems and Information Engineering Design Symposium*. Charlottesville, VA.
- Roberts, S.V. (1959) Control chart tests based on geometric moving averages. *Technometrics* 1, 239-250.
- Ryan, T.P. (1999). *Statistical Methods for Quality Improvement*, 2nd Edition, John Wiley and Sons, New York.
- Samuel, T.R., Pignatiello, J.J., Jr., and Calvin, J.A. (1998). Identifying the Time of a Step Change with \bar{X} Control Charts. *Quality Engineering*. 10, 521-527.
- Schreiber, C. & Carley, K. (2004). *Construct - A Multi-agent Network Model for the Co-evolution of Agents and Socio-cultural Environments*. Carnegie Mellon University, School of Computer Science, Institute for Software Research International, Technical Report, CMU-ISRI-04-109.
- Scott, J. (2001). *Social Network Analysis*, Sage.
- Shiryayev, A.N. (1996). Minimax Optimality of the Method of Cumulative Sums (CUSUM) in the Case of Continuous Time. *Russian Mathematics Survey*. 51, 750-751.
- Siegmund, D. (1986). Boundary Crossing Probabilities and Statistical Applications. *Annals of Mathematical Statistics*. 14, 361-404.
- Smith, I. (2008). Interview conducted for model validation. Lieutenant Colonel Smith is an Academy Professor of Sociology at the U.S. Military Academy. His 20 years of commissioned service include service as a platoon leader, company commander, and battalion operations officer in Ranger and regular Army Infantry units.

- Trent, S. (2008). Interview conducted for model validation. Major Trent is an Assistant Professor of Engineering Psychology at the U.S. Military Academy. He is familiar with modeling and simulation as well as military leadership.
- Topper, Curtis & Carley, Kathleen. (1999). A Structural Perspective on the Emergence of network Organizations. *Journal of Mathematical Sociology*. 24(1), 67-96.
- Wald A. (1945). Sequential Tests of Statistical Hypotheses. *Annals of Mathematical Statistics*. 16, 117-186.
- Wald A. (1947). *Sequential Analysis*, Wiley, New York.
- Wald, A. and Wolfowitz, J. (1948). Optimum Character of the Sequential Probability Ratio Test. *Annals of Mathematical Statistics*. 19, 326-339.
- Wasserman, S. and Faust, K. (1994). *Social Network Analysis Methods and Applications*. Cambridge University Press.
- Weiss, L. (1953). Testing One Simple Hypothesis Against Another. *Annals of Mathematical Statistics*. 24, 273-281.
- Weiss, L. (1962). On Sequential Tests Which Minimize the Maximum Expected Sample Size. *Journal of the American Statistical Association*. 57, 551-557.
- Vance, L.C. (1986). Average Run Lengths of Cumulative Sum Control Charts for Controlling Normal Means. *Journal of Quality Technology*. 18, 189-193.