

NETWORK SIMULATION MODELS

Major Ian A. McCulloh*
Cadet Joshua A. Lospinoso
Network Science Center
U.S. Military Academy
West Point, New York 10996
Email: ai6873@usma.edu

Professor Kathleen M. Carley
Center for Computational Analysis of Social and Organizational Systems
Institute for Software Research International, SCS
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

ABSTRACT

Accurate simulation models of social networks enable over time prediction, statistical testing of hypothesis about social groups, and enables researchers to develop more informed hypotheses for human experimentation, by evaluating them and reasoning about them using simulation. Several methods for simulating social network behavior are compared. The multi-agent simulation *Construct* is shown to be an excellent approach for simulating social behavior. The stochastic engine closely resembles an independently developed statistical framework for dynamic, temporal networks called the Link Probability Model (LPM). This paper illustrates that *Construct*, a multi-agent network model for the co-evolution of agents and socio-cultural environments, is a viable choice for most network simulation needs because it is based on the LPM concepts which perform well on empirical data. Further, with the ability to add additional network dependence, *Construct* is able to leverage flexibility to produce statistically greater conjectures on network structure and knowledge diffusion than alternatives.

1. INTRODUCTION

Network analysis of social groups allows investigators to study the effects of relational variables and relational dependence on group behavior. Traditional analysis focuses on the attributes of individuals. While this is still important, it overlooks important information about social dynamics. For example, if we look at the attributes of one of the London subway bombers, we could identify their income level, ethnic origin, educational level, age, and many other variables. We might also find similar individuals who possess all of these same attributes, yet they do not engage in terrorist activity. What is the difference? We submit that the terrorist maintains social connections with other terrorists, while the similar individual who does not engage in

terrorism maintains greater connections with the general population. Therefore, network analysis provides insight into an important dimension of group behavior.

Social network analysis (SNA) provides a mathematical framework to investigate relationships between social entities (i.e. people, groups, tasks, beliefs, knowledge, etc.). These entities are modeled with nodes and their connections or relationships are modeled with links. Not all nodes are connected to each other and some nodes may have multiple connections. This mathematical model is applicable in many content areas such as communications, information flow, and group or organizational affiliation (Tichy et al., 1979; Wasserman and Faust, 1994). SNA thus relies heavily on graph theory to make predictions about network structure and thus social behavior.

Recently, a great deal of literature has been focused on methods for simulating network structure. Simulation offers a number of advantages to the researcher. First, we can use simulation to emulate the behavior of individuals and predict behavior over time. For example, when analyzing over time data, real world data at time 1 can be used to initialize the simulation program. The simulation can then be used to predict data at time 2. Second, to the extent that such predictions are accurate, we can use the simulation to do hypothetical "what if" analyses. For example, we can use the simulation program to examine alternate hypothetical societies to see what differences in such societies might be necessary to get a different outcome than that perceived in the real data.

The value of such an exercise, is not that it proves why the group or society changed as it did, but that such an exercise provides a way of reasoning about the situation, and enables the researcher to create more informed hypotheses that can then be empirically tested. In sociology, as we move to dynamic models with feedback we will find that they capture more of the social

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the 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 this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE DEC 2008	2. REPORT TYPE N/A	3. DATES COVERED -			
4. TITLE AND SUBTITLE Network Simulation Models		5a. CONTRACT NUMBER			
		5b. GRANT NUMBER			
		5c. PROGRAM ELEMENT NUMBER			
6. AUTHOR(S)		5d. PROJECT NUMBER			
		5e. TASK NUMBER			
		5f. WORK UNIT NUMBER			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Network Science Center U.S. Military Academy West Point, New York 10996		8. PERFORMING ORGANIZATION REPORT NUMBER			
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		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 See also ADM002187. Proceedings of the Army Science Conference (26th) Held in Orlando, Florida on 1-4 December 2008					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified	UU	8	

situation, but that it is incredibly difficult for the researcher to think through, without mistakes, the implications of such models. Simulation becomes a tool for increasing the specificity of theory, thinking through the theoretical implications, and generating testable predictions.

In this paper, we provide an overview of several competing methods of network simulation. Differences and similarities are identified. The link probability model (LPM) is briefly illustrated and we identify why it is in many cases favorable to the exponential random graph (ERG) model. We then move on to summarize *Construct* and its roots in structuralist sociological theory. We discover that the (LPM) provides a mathematical bridge between empirically observed data and the multi-agent simulation, *Construct*, which is based on structuralist theory. *Construct*, in turn, introduces additional relational dependence into the LPM correcting for its naïve assumption of independence. Finally, we depict how this sociological theory translates into the LPM, how *Construct* leverages the LPM, and relate the results of empirical studies conducted by others on the effectiveness of *Construct* vice other alternatives.

2. EXPONENTIAL RANDOM GRAPH MODELS

ERG models are used in social network analysis as a statistical model that enables an analyst to conduct inference on dependent relational data (Goodreau, 2007; Robins, et. al., 2007). The ERG model is therefore less restrictive than earlier models for social networks that assumed dyadic independence (Holland and Leinhardt, 1981). In many social network applications the relationship between two individuals depends on relationships between the individual and others in the network; cognitive limits on the number of relationships that can be maintained; similarity between individuals; and more. The ERG model framework for relaxing the dyadic independence assumption is thus essential for accurate inference in many data sets.

Estimating ERG model terms and parameters can be computationally challenging in large networks (Snijders, 2002; Pattison and Robins, 2002). Markov chain Monte Carlo estimation of ERG models has been used to fit these models to data (Goodreau, 2007; Robins, et. al., 2007; Handcock, 2003, 2002; Snijders, 2002; Pattison and Robins, 2002). The Markov dependence in these models leads to problems of degeneracy, which is discussed in detail by Handcock (Handcock, 2003, 2002). Essentially, model degeneracy occurs when the observed data is almost impossible under the specified model. This often occurs when explanatory terms are highly correlated and there is insufficient data to construct an appropriate model. Several advances in ERG models have been proposed to include curved exponential family models

(Hunter and Handcock, 2006) and neighborhood models (Robins, et. al., 2005). It is not clear that these advances have completely removed issues of model degeneracy, however.

3. LINK PROBABILITY MODEL

The LPM (McCulloh et al., 2007) has been proposed as an alternative model to the ERG model. The LPM framework for viewing the probability space of a social network avoids issues of model degeneracy, while preserving flexibility for modeling dyadic relationships. It provides researchers with an improved means to understand the probability space of the network, under certain conditions. The LPM is a square matrix where the rows and columns correspond to the nodes in a social network. The entries are the link probabilities of the directed link from the row node to the column node. This is not to be confused with an adjacency matrix, where the entries are either zero or some number representing the strength of a relationship between nodes. The link probability is a number between 0 and 1, and determines the likelihood of a link being present in an observed adjacency matrix.

The link probabilities can be derived from empirical data in several ways. Given network data collected over multiple time periods on a group of subjects, the link probabilities can be estimated by the proportion of link occurrences, $e(i,j)$, for each cell in the adjacency matrix, $a(i,j)$. In the case of communication networks, statistical distributions can be fit to the time between messages for each potential link in the network. For a specified period of time, t , the link probability p for each set of entities i and j can be found by integrating over the probability density function from 0 to t .

Relational dependence in link probability are accounted for in the LPM by the historic presence of links. Relational dependence in links can occur for many reasons. One example is if a boss sends an email to two employees telling them to work on a project; that will affect the probability of communication between the two employees. The LPM does not modify the link probability based on these perceived factors that may adjust the probability of two nodes having a relationship. The LPM accounts for the relational dependence, by assuming that it will be inferred by the historic presence or absence of links between nodes. If a boss often gives a task to two employees, then the presence of a link between the employees is likely to be more common when observing past networks. This does not account for all of the relational dependence in the network. In order to introduce a realistic degree of dependence, the LPM would need to be modified at each time step based on social theory established in the literature.

4. CONSTRUCT: MULTI-AGENT SIMULATION

Construct is a multi-agent simulation grounded in constructivist theory (Carley, 1990, 1995). The LPM provides the stochastic engine for the multi-agent simulation. At each time step the link probabilities are determined by the nodes' perceived homophily, socio-demographics, and proximity. These social factors re-introduce the additional relational dependence missing in the raw LPM.

Construct is a dynamic-network multi-agent simulation model that can be used to examine the evolution of social, knowledge and activity networks in response to external interventions and the normal course of human interaction (Carley, 1990, 1991)¹. Network evolution and the diffusion of information and beliefs through social networks can be examined using *Construct* (Carley, 1995; Hirshman & Carley, 2007b, Hirshman, Martin & Carley, 2008). *Construct* captures group dynamic dynamics under diverse cultural and technological configurations (Schreiber & Carley, 2004). Consequently organizational change (Carley & Hill, 2001), socio-cognitive inconsistencies (Carley & Krackhardt, 1996), the impact of communication technologies (Carley, 1995; Carley 2002). To use *Construct* the researcher specifies both the agents replete with information processing capabilities (Hirshman, Carley & Kowalchuk, 2007a) and the networks in which they are embedded (Hirshman, Carley & Kowalchuk, 2007b).

Before, we explore the ability for network simulation to represent reality, we must first lay the foundational theory behind constructivism as it applies to the multi-agent simulation *Construct*. Advances in both cognitive science and network theory have engendered the belief that it should be possible to develop analytical models of the relationships between individuals that would enable quantitative predictions of changes in interaction and that take into account both the self and the society, the individual and the group, the cognitive and the social. These advances have rekindled the dream, originally seen in social comparison theory (Festinger, 1954), cognitive dissonance theory (Festinger, 1957), and balance theory (Heider, 1958), that it is possible to build a mathematics of group change as a function of individual change, yet there is still a gap between the more cognitive and individual perspective in which changes in relationships between individuals result from independent dyadic encounters and the more social and structural perspective which changes in relationships between individuals result from gross changes to the group. Currently a great deal of research is directed at bridging

this gap. On the individual side the linking of symbolic interactionism and role theory can be viewed as a move to incorporate social or group factors into an otherwise predominantly cognitive.

Similarly, affect control theory is a move to incorporate the social, in terms of task constraints and social knowledge, into a cognitive and affective model of the individual's evaluation of; and hence determination of future action (Heise 1971, 1979, 1987; Smith-Lovin 1987). The focus on the change in the individual or his or her relationships to an actual or a generalized other, treats the group or social world as present, but relatively fixed. This implicitly assumes that social or group behavior is somehow an aggregate of the results of independent encounters between pairs of individual. This last assumption is not exclusive to those who propose more cognitively rich models of behavior.

For example, we also see it in the work on status and dominance where hierarchies are viewed to result from independent dyadic encounters (Berger, Conner, and Fisek 1974; Rosa and Mazur 1979; Lamb 1986). On the up side, evidence is being amassed that group behavior cannot be accounted for by aggregating independent dyadic encounters (Chase 1974, 1980; Ridgeway and Diekema 1989) but is rather an emergent property of the simultaneous actions of all group members (Bales 1950; Homans 1950; Chase 1974, 1980; Fararo and Skvoretz, 1986). The mechanism by which such group behavior emerges remains elusive. As a step toward locating this mechanism, research in the structural and network traditions has been moving toward providing explanations, and hence predictions, of individual cognitive change in terms of the individual's social position.

This can be seen in Burt's model of action (1982) where perceived similarity and hence norms, attitudes, likelihood of adopting innovations, and so on is a function of social position. This is further supported by Krackhardt's notion (1985, 1986, 1987) that the individual's social cognition (which he defines as the individual's perception of who interacts with whom) is a function of social position. These works reveal a more cognitive actor than that revealed by classic structuralist whose behavior is nonetheless socially situated. Yet, like the more cognitive individual models, these social models of individual change, still focus on the change in the individual while maintaining a relatively fixed social world. Thus, both the individual and the social perspectives treat the social world as fundamentally stable. Consequently, neither perspective provides a mechanism by which such individual changes can produce social change. Neither approach is sufficient to explain, let alone quantitatively predict, changes in the interaction patterns for all members of the society at once.

¹ The Construct system itself is freely downloadable from the CASOS website, <http://www.casos.cs.cmu.edu/projects/construct>

Rather, the explanations of social change are highly contextual relying on situation specific factors, forces, and constraints such as goals, coercion, bureaucratization, change in group size, and membership rituals.

Every group has a population consisting of some number of individuals. In every group there is a set of information or facts that is potentially learnable by the members of the group. This set of information contains each piece of information that is known by at least one group member. The number of such facts will be denoted by K . At a particular point in time, say time period t . The individual, for any piece of information, such as k , either knows that fact or does not. This is denoted by $F(t) = 1$ if the fact is known by individual at time period t and 0 otherwise.

Every society has a culture, which can be thought of as the distribution of information across the population. At a particular point in time, say time period t , an individual i has a certain probability to interact with another other member of the society, j . This is exactly where the LMP comes into consideration. Every society has a social structure, which can be thought of as the distribution of interaction probabilities across the population. The initial make-up of these probabilities and the transition of these probabilities at different time points are thus determined by several factors.

The first assumption of the *Construct* model posits that interaction leads to shared knowledge. It is generally demonstrable that individuals acquire information (and hence will come to share knowledge) during interactions. In order to represent this process a variety of simplifying assumptions are made. All pieces of information are entirely unstructured and undifferentiated. Thus, the individual may know conflicting information such as the sky is blue and the sky is green. Consequently, the overlap in what two individuals' know is just the sum of the pieces of information that they both know. When two individuals interact each communicates one fact to the other. Individuals always learn the piece of information that is communicated to them. Consequently, if individual i knows that the sky is blue and individual j knows that the sky is green and individual j communicates to individual i that the sky is green, the overlap in their knowledge increases. Hence they have more shared knowledge. All facts known by the individual are equally likely to be communicated.

According to constructivism, both the individual cognitive world and the socio-cultural world are continuously constructed and reconstructed as individuals concurrently go through a cycle of action, adaptation, and motivation. During this process not only does the socio-cultural environment change, but social structure and culture co-evolve in synchrony. Carley (1991a) defined

the following primary assumptions in describing constructivism:

1. Individuals are continuously engaged in acquiring and communicating information
2. What individuals know influences their choices of interaction partners
3. An individual's behavior is a function of his or her current knowledge

In addition to these primary assumptions there were a series of implicit assumptions that upon explication serve to clarify and expand the primary assumptions. Following is an expanded list of assumptions, numbered to clarify their relation to the primary assumptions:

- 1a. Individuals, when interacting with other individuals, can communicate information
- 1b. Individuals, when interacting with other individuals, can acquire information
- 1c. Individuals can learn the newly acquired information thus augmenting their store of knowledge
- 2a. Individuals select interaction partners on the basis of relative similarity and availability
- 2b. individuals engage in interaction concurrently thus an individual's first choice of interaction partner may not be available.
- 3a. individuals have both an information processing capability and knowledge which jointly determine the individual's behavior
- 3b. individuals have the same information processing capabilities
- 3c. individuals differ in knowledge as each individual's knowledge depends on the individual's particular socio-cultural-historical background
- 3d. individuals can be divided into types or classes on the basis of extant knowledge differences.

These assumptions lead to a simulation template, which features a dynamic LPM as the stochastic engine. We briefly present *Construct* in this fashion, and go on to show that it performs well in simulated empirically obtained networks.

5. DATA

The LPM and ERG models are both used to model the Sampson (1969) Monk data and the Newcomb (1961) Fraternity data. The fit of each of these models is compared to the data.

Sampson recorded social network data on the strength of "liking" between monks in a monastery at three different points in time. Between surveys, four of the monks were actually expelled from the monastery.

The social network of these individuals was therefore changed over time.

Newcomb provided 17 college transfer students with fraternity style housing in exchange for their participation in a study on friendship formation. Every week they were required to rate on a scale of 1 to 16 their preference for others in the house. Since ERG models require binary data, we use the dichotomous version of the Newcomb data proposed by Krackhardt (1998), which records a directed link between node i and node j if node i rated node j as one of their top 8 closest relationships in the network. There are 15 time periods in the Newcomb data.

6. COMPARING MODELS

The ERG model and LPM are investigated for their strengths and weakness in modeling longitudinal data in McCulloh (2008). We re-present the results here. For the Sampson (1969) monk data, an ERG model fit by Hunter, et. al. (2008) is used. An ERG model is also fit to the Newcomb (1961) fraternity data. An LPM is also fit to both the Sampson and Newcomb data sets. Monte Carlo simulation is used to generate instances of the Sampson Monk social network and the Newcomb Fraternity social network under the ERG model and the LPM.

A distance measure is required to compare the similarity between the dichotomous networks generated using the ERG model, the LPM, and the empirical data. Hamming distance (1950) is a logical choice, since it evaluates a distance between dichotomous networks. If the data were weighted networks and the models generated weighted networks as well, then a Euclidean

distance would be appropriate. The quadratic assignment procedure (QAP) (Krackhardt, 1987) could be used to compare the correlation between networks; however, the correlation coefficient does not change linearly with network distance. The average Hamming distances from each empirical data set to every other empirical data set and from each simulated network to each empirical data set were calculated. These average Hamming distances were then compared using a t-test. The results of this test indicate whether the LPM or the ERG model, models the empirical networks with more or less error. Table 1 shows the distance between the Sampson Monk data to both the ERG and LPM. Table 2 shows the distance between the Newcomb Fraternity data to both the ERG and LPM. It can be seen in both tables that the p-values are significant at the 0.05 level. This means that there is a significant difference between how well the ERG and LPM model empirical data. The positive values for the test statistic indicate that the LPM does a significantly better job of modeling empirical data than the ERG.

Table 1. ERG and LPM Distance to Empirical Data for the Sampson Monk Data

Time period	1	2	3
Mean Hamming Distance for ERG model	98.7	99.1	103.7
ERG Standard Deviation	5.697	6.2263	6.2902
Mean Hamming Distance for LPM	27.67	24.99	24.66
LPM Standard Deviation	3.5922	3.5935	3.5945
T-Test Statistic	39.43	37.64	39.74
P-value	0.0006	0.0007	0.0006

Table 2. ERG and LPM Distance to Empirical Data for the Newcomb Fraternity Data

Time Period	Mean Hamming Distance for ERGM	ERG Standard Deviation	Mean Hamming Distance for LPM	LPM Standard Deviation	t-test	p-value
1	139.7	8.3938	91.9	5.1913	18.0147	0.0353
2	138.9	8.1847	75.1	5.2128	24.6573	0.0258
3	137.3	8.2872	48.3	5.2226	33.9732	0.0187
4	135.5	9.3363	49.7	5.2340	29.0460	0.0219
5	134.1	8.9870	50.1	5.2319	29.5558	0.0215
6	136.3	8.5251	45.5	5.2440	33.6983	0.0189
7	133.9	9.0609	47.3	5.2397	30.2202	0.0211
8	134.1	7.2946	51.9	5.2591	35.6377	0.0179
10	133.7	5.1865	64.2	5.2223	42.3990	0.0000
11	132.7	6.0562	53.4	5.2074	41.4119	0.0006
12	136.3	8.4466	51.1	5.2147	31.8930	0.0200
13	134.9	9.0117	46.6	5.2311	30.9989	0.0205
14	133.9	5.4457	46.1	5.2230	50.9574	0.0000
15	133.1	5.7242	47.2	5.2378	47.4518	0.0004

A similar test was done to compare the Hamming distance between the empirical data at each time point, with the empirical data at all other time points. The results are omitted due to space restrictions. The LPM was found to have no more error than that present between different time points in the empirical data. This provides evidence to validate the LPM as an effective method for simulating data.

The LPM has additional advantages. The LPM avoids the issues of model degeneracy inherent in the ERG model. The probability of link occurrence is based on the historic presence of links and does not use a Markov assumption or over specify a statistical model. For these reasons, the LPM provides an alternative method for modeling and conducting longitudinal social network analysis. For our purpose in this paper, the LPM's success makes it a reasonable stochastic engine for the *Construct* multi-agent simulation model. The multi-agent simulation simply adds additional relational dependence into a model that already performs well to make it more realistic and capable of evolution over time.

7. DISSCUSSION

The theoretical underpinnings of constructivism as manifested in *Construct* lead us to a multi-agent simulation which utilizes a dynamic LPM as a stochastic engine for the development of knowledge diffusion and relationship building. What does this simulation provide the user?

The simulation provides an accurate, realistic simulation of social dynamics. We envision several ways in which this will be important to the military in particular and the wider academic audience in general.

Construct can be used as a valuable decision support tool for military commanders. The social dynamics of terrorist organizations, local culture, or friendly military forces can all be modeled with the simulation. A commander can war-game potential courses of action, and evaluate alternatives using *Construct*. It can be very difficult to reason through the many potential interactions, factors, and competing theories. This simulation provides a framework that is grounded in social theory, and validated against empirical evidence, that can be used to evaluate potential courses of action.

For example, a commander might consider detaining one or more suspected terrorists. By modeling the course of action in *Construct*, he can observe the impacts of removing the individual, on the organizations performance, situational awareness, and overall effectiveness. Given limited resources, the commander

could even use the simulation to optimize the individuals to remove from the social group. The simulation provides the military analyst the ability to predict the future social dynamics of an organization. This is a powerful combat multiplier for today's non-kinetic asymmetric war fighter.

The Army could also use *Construct* to evaluate the organizational structure of newly formed doctrinal units, such as the Future Combat System (FCS) operational units. The simulation can evaluate which personnel communicate more or less frequently. This can help inform efficient organization of soldiers from staff organizations to vehicle crews.

Focused research on social groups can follow better experimental design, and yield greater knowledge, if an array of research questions is first evaluated in simulation. Social dynamics are complex and it can be difficult to correctly reason through different scenarios. Simulation can provide insight that may shape the research questions to be more effective.

Finally, the normal behavior of an organization can be simulated many times. From the simulations, statistical distributions can be fit to various measures of group behavior. These statistical distributions can be used to evaluate statistical hypotheses or to detect statistically significant differences between observations of the group and normal behavior. This statistical framework, therefore, increases the relevant findings one can discover in socially dynamic organizations.

8. CONCLUSION

We have presented two models for describing the behavior of social networks: the ERG model and the LPM. Both models were fit to two famous data sets in the literature, the Sampson Monk data, and the Newcomb Fraternity data. The LPM modeled the data with a statistically significant better fit than the ERG model. The benefit of the LPM was further demonstrated by finding that the difference between the LPM fit and the empirical data, was no larger than the average difference between any two samples of the empirical data.

The key limitation of the LPM is that it does not account for all of the relational dependence that is known to exist in socially connected groups. The multi-agent simulation *Construct* conveniently overcomes this limitation. *Construct* essentially uses the LPM as its stochastic engine. The link probabilities at each time step are affected by constructivist theory established in the literature. Factors such as perceived homophily, shared knowledge, proximity, and socio-demographic variables all affect the link probabilities at each time

period. These factors introduce relational dependence into the LPM. The relative weighting that these factors have can be adjusted by the user. This creates a flexible simulation tool, grounded in empirical evidence and sociological theory.

While *Construct* may be a powerful simulation tool, the current user interface limits its' capability. The Organizational Risk Analyzer (ORA) is a software package maintained by the Center for Computational Analysis of Social and Organizational Systems (CASOS) at Carnegie Mellon University. ORA has an interface for near-term impact, which allows the user to isolate certain agents in a socially networked group and evaluate the impact of the isolations through simulation using *Construct*. Other than this interface, simulation runs must be conducted using an xml script. Future research will hopefully provide funding to better develop the user interface for the simulation. An improved user interface might make *Construct* available to a brigade staff to better evaluate various courses of action. This improved ability to war-game various scenarios may enhance the effectiveness of those military units.

ACKNOWLEDGEMENTS

This work was supported in part by the National Science Foundation IGERT in CASOS (DGE 997276) . Additional support for the fundamental development of Construct was provided in part by the Office of Naval Research (N00014-06-1-0104), the Internal Revenue Service, the National Science Foundation (SES-0452487), the Air Force Office of Sponsored Research (MURI: Cultural Modeling of the Adversary, 600322) for research in the area of dynamic network analysis., and the Air Force (FA865007C6769) for examining the unintended consequence of action using multi-agent simulation. Support for LPM was provided by the Army Research Institute for the Behavioral and Social Sciences Army Project No. 611102B74F Additional support was provided by CASOS - the center for Computational Analysis of Social and Organizational Systems at Carnegie Mellon University. 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 Office of Naval Research, the Internal Revenue Service, the Air Force Office of Sponsored Research, the US Air Force, the National Science Foundation , the Army Research Lab or the U.S. government.

REFERENCES

Bales, R.F. (1950). *Interaction Process Analysis: A Method for the Study of Small Groups*. Cambridge, MA: Addison-Wesley.

Berger, J., Conner, T.L., Fisek, M.H. (1974). *Expectation States Theory: A Theoretical Research Program*. Cambridge MA: Winthrop.

Burt, R.S. (1982). *Toward a Structural Theory of Action*. New York: Academic Press.

Carley, K.M. (1991). A theory of group stability. *American Sociology Review*, 56(3):331–354.

Carley, K.M. (1990). Group stability: A socio-cognitive approach. In Lawler E., Markovsky B., Ridgeway C., and Walker H. (Eds.) *Advances in group*

processes: Theory & research. VII., 1-44. Greenwich, CN:JAI Press.

Carley, K.M. (1995). Communication Technologies and Their Effect on Cultural Homogeneity, Consensus, and the Diffusion of New Ideas. *Sociological Perspectives*, 38(4): 547-571.

Carley, K.M. (2002). Smart agents and organizations of the future. In L. Lievrouw & S. Livingstone (eds), *The Handbook of New Media*, 206-220. Thousand Oaks, CA: Sage.

Carley, K. M., Hill. V. (2001). Structural change and learning within organizations. In A. Lomi and E. R. Larsen (eds), *Dynamics of Organizations: Computational Modeling and Organization Theories*. Menlo Park, CA: MIT Press/AAAI.

Carley, K.M., Krackhardt. D. (1996). Cognitive inconsistencies and non-symmetric friendship, *Social Networks*, 18: 1-27.

Chase, I.D. (1974). Models of Hierarchy Formation in Animal Societies. *Behavioral Science*, 19, 374-82.

Chase, I.D. (1980). Social Process and Hierarchy Formation in Small Groups. *American Sociological Review*, 45, 905-24.

Fararo, T.J., Skvoretz, J. (1986). E-State Structuralism. *American Sociological Review*, 51, 591-602.

Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, 7:117-140.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson

Goodreau, S.M., (2007). Advances in exponential random graph (p*) models applied to a large social network. *Social Networks*, 29, 231-248.

Hamming, R.W. (1950). Error Detecting and Error Correcting Codes, *Bell System Technical Journal* 26(2):147-160

Handcock, M.S., (2003). Statistical models for social networks: degeneracy and inference. In: Breiger, R., Carley, K., Pattison, P. (Eds.), *Dynamic Social Network Modeling and Analysis*. National Academies Press, Washington, DC, 229–240.

Handcock, M.S., (2002). Statistical models for social networks: degeneracy and inference. In: Breiger, R., Carley, K., Pattison, P. (Eds.), *Dynamic Social Network Modeling and Analysis*. National Academies Press, Washington, DC, 229-240.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. New York: Wiley.

Heise, D. (1977). Social Action as the Control of Affect. *Behavioral Science*, 22: 163-77.

Heise, D. (1979). *Understanding Events: Affect and the Construction of Social Action*. New York: Cambridge University Press.

Heise, D. (1987). Affect Control Theory: Concepts and Models. *Journal of Mathematical Sociology*, 13:1-34.

Hirshman, B.R., Carley, K.M., Kowalchuck, M.J. (2007a). *Specifying Agents in Construct*. Carnegie

- Mellon University, School of Computer Science, Institute for Software Research, Technical Report, CMU-ISRI-07-107.
- Hirshman, B.R., Carley, K.M., Kowalchuck, M.J. (2007b.) *Loading Networks in Construct*. Carnegie Mellon University, School of Computer Science, Institute for Software Research, Technical Report, CMU-ISRI-07-116.
- Hirshman, B.R., Martin, M.K., Carley, K.M. (2008). *Modeling Information Access in Construct*. Carnegie Mellon University, School of Computer Science, Institute for Software Research, Technical Report, CMU-ISR-08-115.
- Holland, P.W., Leinhardt, S., (1981). An exponential family of probability distributions for directed graphs (with discussion). *Journal of the American Statistical Association* 76, 33–65.
- Homans, G.C. (1950). *The Human Group*. New York: Harcourt, Brace and Co.
- Hunter, D.R., Handcock, M.S., (2006). Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics* 15, 565–583.
- Krackhardt, D., Porter, L. (1985). When Friends Leave: A Structural Analysis of the Relationship between Turnover and Stayers' Attitudes. *Administrative Science Quarterly*, 30(2): 242-261.
- Krackhardt, D., Porter, L. (1986). The Snowball Effect: Turnover Embedded in Communication Networks. *Journal of Applied Psychology*, 71(1): 50-55.
- Krackhardt, D. (1987). QAP Partialling as a Test of Spuriousness. *Social Networks*, 9, 171-186.
- Krackhardt, D. (1998) Simmelian ties: Super strong and sticky. In *Power and Influence in Organizations* (eds R. Kramer, M. Neale), 21-38. Sage, Thousand Oaks, CA.
- Lamb, T.A. (1986). The Familiarity Effect in Small-group Hierarchy Research. *Journal of Social Psychology*, 126, 51-56.
- Lazarsfeld, P. and Merton, R., (1954). Friendship as social process: a substantive and methodological analysis. In *Freedom and Control in Modern Society* (Eds M. Berger, T. Abel, and C. Page), 18-66. New York: Van Nostrand.
- McCulloh, I., Lospinoso, J., and Carley, K.M., (2007). Social network probability mechanics. In *Proceedings, 12th International Conference on Applied Mathematics, World Scientific Engineering Academy and Society*, Cairo, Egypt, 319-325.
- McCulloh, I., and Carley, K.M., (n.d.) The Link Probability Model: An Alternative to the Exponential Random Graph Model for Longitudinal Data. Submitted to the *Journal for Social Networks*.
- Newcomb, T.N. (1961). *The Acquaintance Process*. New York: Holt, Rinehart and Winston,
- Pattison, P.E., Robins, G.L., (2002). Neighbourhood-based models for social networks. *Sociological Methodology* 32, 301–337.
- Ridgeway, C. Diekema, D. (1989). Dominance and Collective Hierarchy Formation in Male and Female Task Groups. *American Sociological Review*, 54, 79-93.
- Robins, G.L., Pattison, P.E., (2005). Interdependencies and social processes: generalized dependence structures. In: Carrington, P., Scott, J., Wasserman, S. (Eds.), *Models and Methods in Social Network Analysis*. New York: Cambridge University Press, 192–214.
- Robins, G.L., Pattison, P.E., Kalish, Y., Lusher, D., (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, 29, 173-191.
- Rosa, E. Mazur, A. (1979). Incipient Status in Groups. *Social Forces*, 58, 18-37.
- Sampson, S.F., (1969). *Crisis in a cloister*. Ph.D. Thesis. Ithaca: Cornell University,.
- Smith-Lovin, L. (1987). Impression Formation from Events. *Journal of Mathematical Sociology*, 13, 35-70.
- Snijders, T.A.B., (2002). Markov chain Monte Carlo estimation of exponential random graph models. *Journal of Social Structure* 3, 2.
- Schreiber, C., Carley, K.M. (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.
- Tichy, N.M., Tushman, M.L., Fombrun, C., (1979). Social Network Analysis for Organizations, *The Academy of Management Review*, 4(4): 507-519.
- Wasserman, S., and Faust, K., (1994). *Social Network Analysis: Methods and Applications*. New York: Cambridge University Press.