Social Network Analysis and Simulation of the Development of Adversarial Networks

Razvan Orendovici, Computer Science and Engineering
Frank E. Ritter, College of IST
Pennsylvania State University
University Park PA
razvan.orendovici@gmail.com, frank.ritter@psu.edu

Keywords:
network visualization, social network, simulation, adversarial networks, development of networks

ABSTRACT: *We present a novel way to monitor and analyze the time course of adversarial networks through a simulation tool and supporting mathematical analysis. Recent work on social networks has been used in the analysis of adversarial networks and their underlying structure with hopes of detecting and preventing future activity. In this paper we consider an adversarial network to be a network subgroup that works against the interests of the group studying it. ANA, the software package presented here, can portray the structure of such networks and allow analysts to find patterns and key players in the network while watching the network evolve. Finally, this work uses output from ANA to study how a simulated adversarial scenario grows in structure and compares it to more traditional social networks on standard measures and also analyze the changes over time of this network on those same dimensions.*

1. Introduction

We present a novel tool to monitor, analyze, and examineadversarial social networks through a visualization tool and supporting mathematical analysis. This tool is demonstrated by presenting a simulated adversarial network that was used to guide the development of the visualization tool. Finally, we analyze how this network evolved and present the implications for education, network science, social network analysis and measurement.

1.1 Motivation

Social network analysis and visualization of adversarial networks is a very powerful method of understanding networks and keeping track of relevant information about the network as it evolves and becomes more defined. Recent events and political pressures have led to a lot of research into adversarial networks and their identification. This, combined with recent popularity of social network analysis, has made a network centric analysis of these networks interesting and useful. Throughout this work we look at allowing users to take mock intelligence gatherings on the communications and individuals involved in a possible adversarial network and keep track of this information while visualizing it and performing statistical analysis.

After describing the background to this work, we present a visualization tool in Section 2, tailored specifically at the style of data that intelligence agencies capture that allows the user to store the information and visualize the graph as it changes. Once this information is entered and visualized, the tool provides additional features to allow this data to be analyzed over its evolution and new patterns detected. Section 3 looks at this analysis for the adversarial network created by intelligence data with analysis found on more traditional networks seen in the literature.

The adversarial network viewed in this paper is a series of intercepted communications in a mock terrorist plot (Shemanski, 2011). The communications are made to appear as intelligence reports from various national agencies and present links between the various actors in the plot. This tech report is used in security and risk analysis classes for students to analyze in the College of Information Sciences and Technology at The Pennsylvania state University.

1.2 Background

Social network analysis has been a growing field since the work done by Moreno (1978) and others (Anthonisse, 1971; Beauchamp, 1965; Freeman, 1979; Holland & Leinhardt, 1971, 1972; Sabidussi, 1966) to establish measures and calculations of the field in the 1950's to the 1970's. With the adoption of the Internet and growth of social network platforms this analysis has become even more widespread and been applied to a variety of fields.

Communities of authorship (Albert & Barabasi, 2002; Barabasi, Jeong, Nelda, Ravasz, Schubert, & Vicsek, 2002; Newman, 2003; Qiu, Ivanova, Yen, Liu, & Ritter, 2011), Economic communities and online games (Bakshy, Simmons, Huffaker, Teng, & Adamic,

2010), and many other fields have benefited from the application of network science and graph theories. In the past 20 years there has been some focus at looking at how adversaries and competition inside social networks (e.g., Baldwin, Bedell, & Johnson, 1997) changes a network's performance and influences it.

This competition and rivalry was studied in classroom settings (Yang & Tang, 2003), and also became a crime prevention and understanding topic in the early 2000s with the September 11th terrorist attacks. Krebs (2002) was the first to look at the terrorist plot as a social network and to diagram it out and present it to the world as a social network between the terrorists. At the same time Klerks (2001) looked at organized crime and mapped out the social network and money flow networks of criminal organizations. While it is difficult to understand these networks, the ability to target particular nodes in the network and disrupt the entire network is a very powerful feature of social network analysis (Carley, Lee, & Krackhardt, 2002). These specific applications to crime and terrorism are just a small portion of work done in adversarial networks.

We use this foundation to analyze incoming intelligence reports about a mock terrorist plot and view the entire organization as a social network with leaders and subgroups.

Most work in the analysis of social networks takes a snapshot of an existing network and analyzes this network based on multiple centrality measures, distance measures, existence of power law patterns transitivity and clustering values. In this paper, however, we find that looking at a static view of the network is not as informative and useful as viewing how the network evolves over time. The evolution of a network is hard to analyze in most work due to difficulty in obtaining the data at different times or difficulty in observing the social network through its growth.

By using a social network that evolves from the communication of terrorists we can view the network from the beginning as it is seen through the intelligence reports about it. This allows us to look at standard social network measures mentioned earlier but with respect to how they change over time. This paper proposes these measures as being more interesting to analyze with respect to time, and also presents some inherent difficulties that exist in most social network analyses but that are not apparent until you view them with respect to time.

2. Adversarial Network Analyzer (ANA)

To best see these networks evolve and meet the requirements of the data set, this paper presents a graph visualization tool tailored to handling the simulated adversarial network. The Adversarial Network Analyzer (ANA) is a Java applet that allows users to input new connections about the graph and visualizes the state of the graph at all-time intervals. To provide powerful graph visualizations ANA is written on top of the Prefuse visualization toolkit (Heer, Card, & Landay, 2005). This feature-rich library provides a visualization library useful for displaying many aspects of datasets. The feature most commonly used in ANA is the use of graph visualization through Nodes and Edges. Prefuse handles all the calculations and work to layout and render the graph.

The application is a Java UI applet so that it could best be used by developers and analysts interested in network evolution or adversarial networks.

ANA 1.0 is planned to be used in the spring of 2012 by a security and risk analysis course taught in the College of Information Science and Technology at the Pennsylvania State University. The students who will be using this software will be studying simulations like the Shemanski (2011) dataset to understand an adversarial plot and identify it.

The UI shown in Figure 1 is broken down into four distinct sections, each providing a useful interface to managing the network graph. The various panels map to the supported tasks of ANA, described in Table 1, with some functionality moved into the menu.

Table 1 ANA 1.0 Supported Tasks

(a) Visualize graph
(b) Add nodes to existing graph
(c) Add edges to existing graph
(d) Modify existing edges and add more details
(e) Modify existing nodes and add more details
(f) Playback of graph expanding from the first node
(g) Save and reload graph
(h) Export state of graph to standard XML for mathematical analysis by ORA

An important part of the visualized data set is the requirement to handle unknown and incomplete data. As communications are intercepted and found by various intelligence organizations nobody knows the full layout of the cell network or the identities of everybody involved. The network has to be able to accept incomplete data and allow us to later fill in the blanks as more information becomes available. This is the reason behind allowing all information to be edited, and why nodes are added to the graph one at a time.

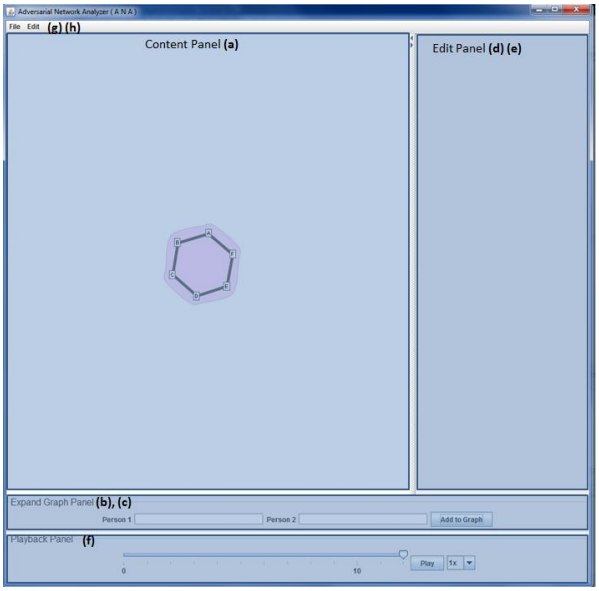

Figure 1 ANA 1.0 User Interface

Social networks do not just appear with hundreds of nodes and interconnections; they evolve slowly from nodes connecting to each other and new communications between actors showing up. We monitor this slowly from intelligence gatherings and intercepted communications.

ANA records all changes made in the graph, as edges and nodes are added or modified. This change detection and recording gives ANA the ability to play through the evolution of a graph from the beginning and animate it for the benefit of the user.

To run the mathematical analysis on any network created in ANA we built an exporting feature in ANA that allows the network to be output any time slice to a standard file format that tools such as the Organization Risk Analyzer, ORA (Carley & Reminga, 2004) can read. From here the calculations and analysis provided in Section 3 can be quickly computed.

3. Network Analysis

For the work in this section we entered the entire contents of the Shemansky (2011) simulation into ANA in order of how the intelligence reports appear. This simulation of 73 incident reports and 15 background reports were used to build the social network of the terrorist plot, as any analyst looking at the plot would see. We then exported the simulation to an ORA file format to compute the network statistics. The network created contained 30 nodes connected by 46 edges over a total of 117 time frames. We exported the state of the graph at intervals of 10 frames to visualize the time course of networks inside ORA.

Figure 2 presents this final network with labels and added color coding of the subgroups in the network.

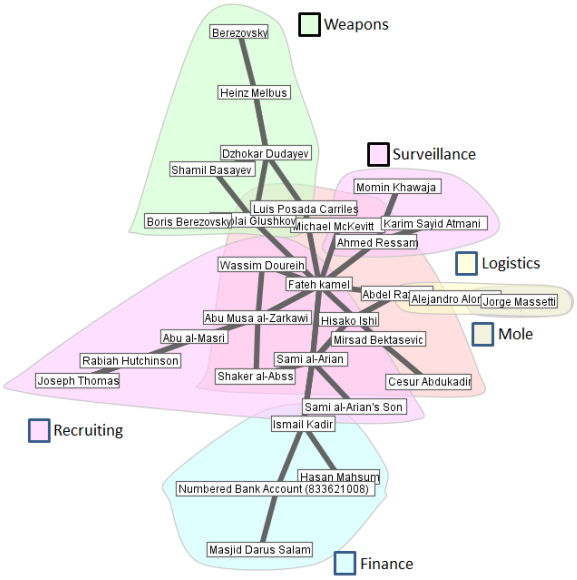

Figure 2 Final network of the Shemanski (2011) adversarial network simulation

3.1 Static Measures, Global and Local

In Table 2 we present the results of this analysis both on a global scale of the entire network, but also on a local scale where we look at particular agents and see how their position in the graph changes over time. The two players are Fateh Kamel, and Shaker al-Abssi who are the respective playmaker and leader of the plot.

Table 2 Static measures of the graph and key players

Measure	Graph Average	Fateh Kamel	Shaker al-Abssi
Degree Centrality	.053	.207	.034
Distance Centrality	.146	.305	.225
Betweenness Centrality	.048	.430	.012
Clustering Coefficient	.080	.042	.000
Distance	3.440		
Transitivity	.047		

The playmaker, Fateh Kamel, remains a more central node in the network than the average of the network and, more importantly, than the leader of the network.

This network is highly decentralized and spread out to protect the identities of its nodes. Compared to more traditional networks studied in the literature (e.g., co-authorship in various fields, World Wide Web, and movie actors) it has node degrees that are an order of magnitude smaller than some networks, and also clustering coefficients that are an order of magnitude smaller than previously studied social networks.

Distance in the network remains rather small both due to the size of the network and the connectivity of the playmaker. This network is much more similar to networks of things like the World Wide Web and power grids that are *non-social networks*.

Table 3 Comparison of Shemanski adversarial network with more standard social networks (Albert & Barabasi, 2002).

Network	Size	Average Degree	Average Distance	Clustering Coefficient
Shemanski Network	***30***	***1.53***	***3.44***	***.080***
WWW	153,127	35.21	3.10	.1078
Movie Actors	225,226	61	3.65	.79
LANL co-authorship	52,909	9.7	5.90	.43
MEDLINE co-authorship	1,520,251	18.1	4.60	.066
SPIRES co-authorship	56,627	173	4.00	.726
NCSTRL co-authorship	11,994	3.59	9.70	.496
Math. Co-authorship	70,975	3.9	9.50	.59
Neurosci. Co-authorship	209,293	11.5	6.00	.76
Power Grid	4,941	2.67	18.70	.08

3.2 Global Dynamic Measures

Rather than just comparing snapshots of the network, we analyze how it changes over time for various global measures of centrality and clustering of the network.

Figure 3 a, b, c presents the results for the change in centrality values over the time course of the network evolution. These values are already small, less than 50%, for all of the measures but present an interesting pattern of decreasing over the time of the simulation. Degree centrality and distance centrality show how both values decrease once the graph achieves a size of greater than 5 nodes, and then stabilize at a very small value. The deception forces inside this network keep it from becoming too centralized so that it may maintain its cell like structure remains hard to detect or infiltrate.

For the non-centrality based measures, clustering coefficient, transitivity, and average distance, presented in Figure 3 we see very similar patterns of the graph aiming to be more spread out and less tightly connected as more actors are brought into the network.

The clustering coefficient does increase between frame 40 and 50 due to a few connections developing. These connections bring the finance and weapons subgroups closer together so they can more effectively work inside of their subgroup. Over time this clustering coefficient does not continue to grow and the groups grow farther apart. These changes may be indicative of normal network growth or may be an artifact of this simulated network. In any case, the changes suggest that studying the time course of network growth can be interesting.

Distance across the network continues to increase over the entire time course of the network as more nodes are added to the far ends without connecting them to the center of the graph for quick communication paths. This keeps the two far ends of the network far apart and allows one end of the network to remain safe if anything were to compromise the other end.

Finally, transitivity only increases when weapons and finance subgroups become connected but decreases afterwards similar to the clustering coefficient. This shows that very few ties are created between triads of actors and instead the network chooses to communicate through the longer pre-existing chains of command and communication.

3.3 Local Dynamic Measures

Many of the values calculated in the previous sections can be calculated for particular actors. In Figure 4 we show a comparison between centrality and clustering of the two key players of the network, Fateh Kamel and Shaker al-Abssi. The important pattern seen in all of these graphs is that while both actors have low values for all of these network measures, the leader tries to remains less central and less visible compared to the playmaker. The leader, Shaker, is always looking to be more obscured by the surrounded network, while the playmaker, Fateh Kamel, is at times looking to grow more connections so that he can more quickly work with the various subgroups and leaders of those subgroups. Surprisingly, the leader's *distance centrality* in the graph becomes lower over time.

The betweenness centrality, degree centrality, and clustering coefficient for the playmaker actually increase in the graph as he becomes more tightly coupled to some of the people he is directing and organizing. This allows him to be effective, and yet leave the actual leader less detectable.

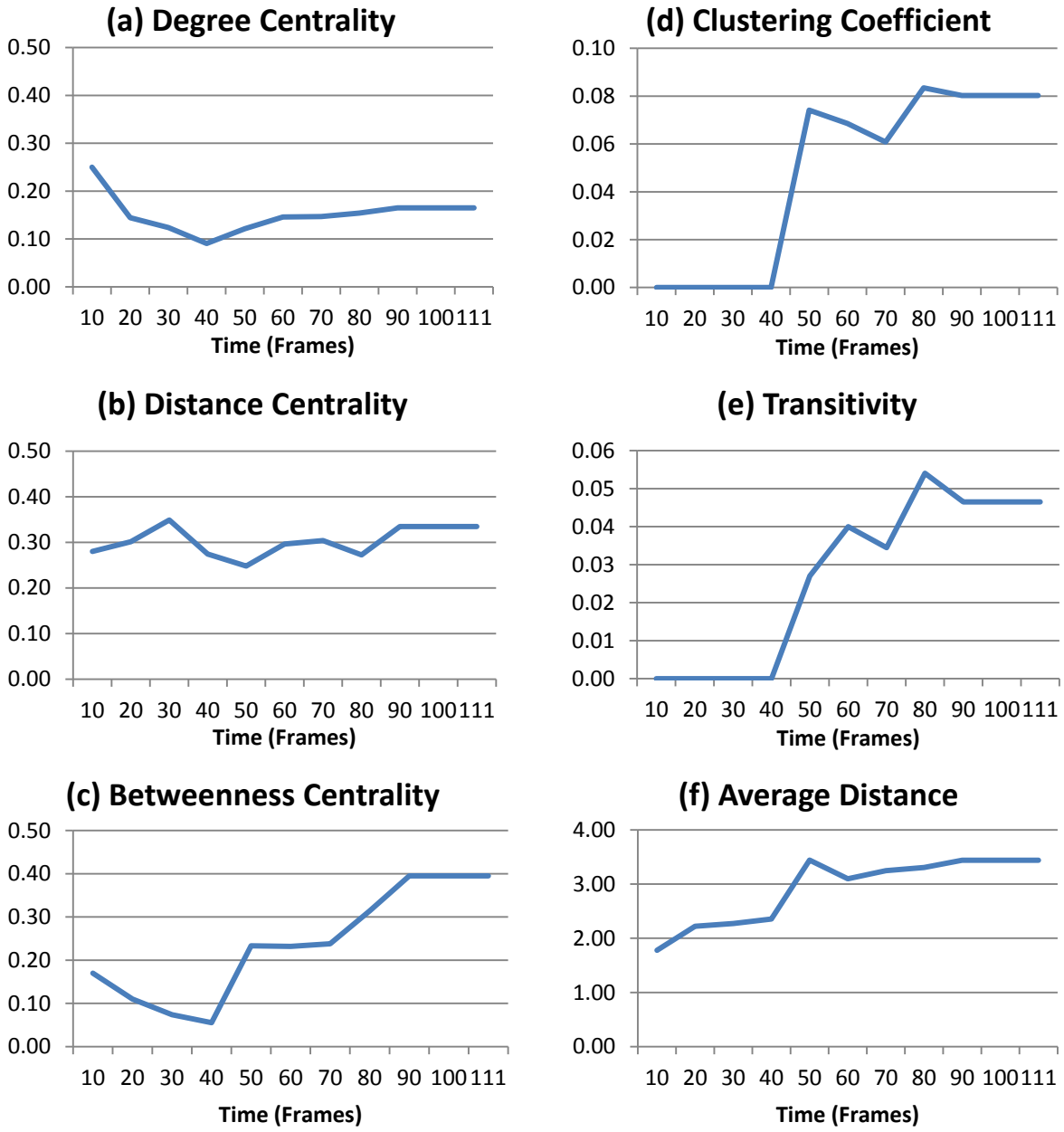

Figure 3 Global Dynamic Measures a) Degree Centrality b) Distance Centrality c) Betweenness Centrality d) Clustering Coefficient, e) Transitivity f) Distance of the entire network measures over the development period of the network.

3.4 Summary

Adversarial networks maintain a much smaller degree centrality than other networks. They are not interested in having each person be connected to as many other people in the network as possible. Each person is only connected to one, perhaps two other people that are strictly necessary to accomplish tasks. Additionally the clustering coefficient is small to minimize triangles of connections between actors.

In terms of dynamic measures, these networks do not follow a simple pattern of increasing connectivity, centrality, and clustering. The values actually decrease over the time of the network, and the average distance between members of the network increases. The networks push to be more spread out and increase the distance between recruits with high risk of being compromised and the leaders and playmaker of the network that carry out key actions.

The leader of such a network remains hidden from almost all members of the network. While this dataset does not specifically say how many people inside the network could identify the leader, our information and analysis of this shows that almost nobody would be able to identify the leader. Shaker al-Abssi remains nearly invisible from all but one or two members of the

network. These members are the playmaker who actually carries out all of his orders and missions, and an additional buffer person. These covert strategies allow the playmaker to be the most visible person in the network. If he was to be compromised, he could be easily replaced by another member to carry out the orders of the leaders and bring the various functions together for a successful plot. The leader knows operational details of the group but does not actively use the existing links in the network.

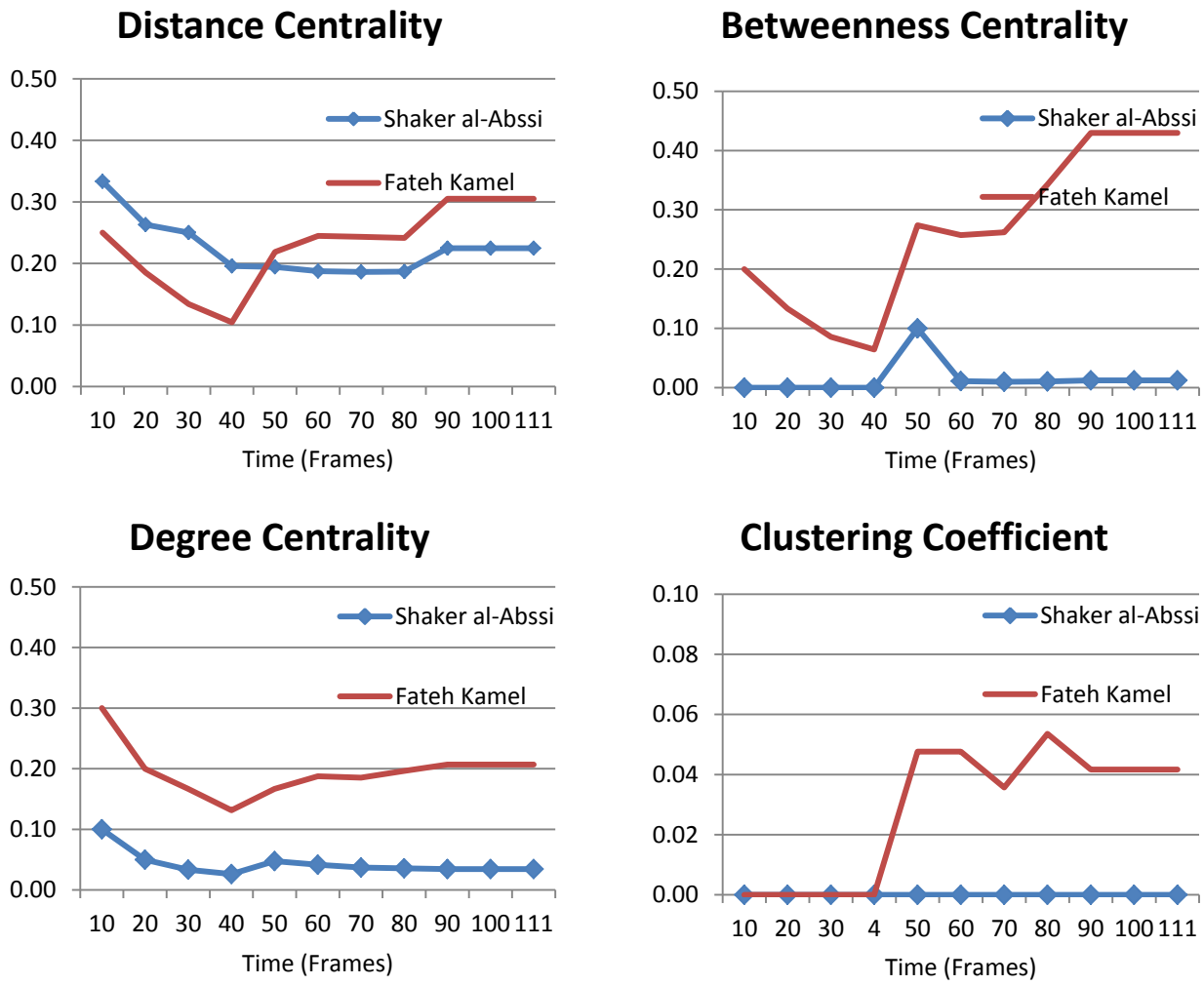

Figure 4 Local Dynamic Measures a) distance centrality b) betweenness centrality c) degree centrality d) clustering coefficient of the key players in the network measured over the development period of the network.

4 Conclusions

This paper introduced a new way to look at social networks, in particular adversarial networks and to analyze them for new patterns. These networks are seen through the eyes of a new animated visualization tool, ANA, that can build the network as information about its actors and connections emerge. One such network, the Shemanski (2011) simulation of an adversarial plot, is visualized through ANA, and using ANA's interface to standard file formats it is mathematically analyzed through social network measures.

In a time-wise analysis of the social network for two important actors, this analysis of these two actors shows additional differences between them. The leader, as expected, avoids building more connections and rather allows himself to be less and less central to the network as the network evolves. He is always less central and connected than the playmaker. The playmaker while not being heavily connected must build a number of connections between actors so the functional subgroups can work together.

Time-based analyses of adversarial social networks show concretely how these differ fundamentally from normal social networks. The network does not evolve to be more connected, nor does it evolve to be more central. The actors remain spread far apart so that each

subgroup is separated and protected from problems that may occur in other parts of the network.

4.1 Limitations

Social Network analysis of these types of networks is a challenging task due to the nature of the networks. The network itself is adversarial and remains covert or tries to hide its underlying structure to improve its own performance. This causes error in the data that is obtained about the network and can complicate the analysis of such a network.

The observation methods used to record and look at social networks suffer from an inherent lag that can cause error in the analysis. All analysis is done on the network evolution of when we observe connections to be created. This is an analysis of our understanding of the social network rather than an analysis of the underlying evolution of the social network. Not all of the connections that are appearing through communications between actors are the first interaction between them. Many of these connections could have been formed days, months, or even years earlier but only been called into action when we observed it.

Inherent differences between our view of the network and the underlying structure of the network presents a source of error and remains as something to be looked at in future time-based analyses of social networks. Even work that does not study adversarial networks suffers from such a lag. Connections on popular social networks (Facebook, Twitter, LinkedIn) are not formed in a vacuum and are usually representative of an earlier interaction between actors. The same limitation applies to studies of publication networks that are frequent in network science. These publication databases suffer from a lag between when those researchers met each other and began sharing ideas and working together and the time when a collaborative paper is published.

4.2 Future Work

ANA in its current state does a good job of fulfilling a use case for students and intelligence analysts with a very simple interface and easy to use features. Future improvements can add more strength to ANA by providing support for more types of networks. The three additional types of networks that should be supported by future versions of ANA include networks with positive and negative relationships, directional relationships, and multi-mode networks (Qiu, Ivanova, Yen, Liu, & Ritter, 2011) to support the inclusion of events, and multi-person meetings. These features would allow more complete modeling of the interactions of an adversarial network but would have to be carefully implemented as to not overly complicate the interface and visualization of the network.

5. Acknowledgements

This work was supported by a grant from the Defense Threat Reduction Agency (HDTRA1-09-1-0054). We would like to thank Don Shemanski, Chris Dancy, and Jonathan Morgan for their expertise and help with this work.

6. References

Albert, R., & Barabasi, A.-L. (2002). Statistical mechanics of complex networks. *Review of Modern Physics, 74*, 47-97.

Anthonisse, J. M. (1971). *The rush in a graph.* Amsterdam: Mathematische Centrum.

Bakshy, E., Simmons, M. P., Huffaker, D. A., Teng, C.-Y., & Adamic, L. A. (2010). The social dynamics of economic activity in a virtual world. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media.* Washington DC, 1-10.

Baldwin, T. T., Bedell, M. D., & Johnson, J. L. (1997). The social fabric of a team based M.B.A. program: Network effects on student satisfaction and performance. *Academy of Management Journal, 40*(6), 1369-1397.

Barabasi, A.-L., Jeong, H., Nelda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A 311*, 590-614.

Beauchamp, M. A. (1965). An improved index of centrality. *Behavioral Science, 10*, 161-163.

Carley, K. M., & Reminga, J. (2004). *ORA: Organization Risk Analyzer*. Carnegie Mellon University, School of Computer Science, Institute for Software Research International, Technical Report.

Carley, K. M., Lee, J.-S., & Krackhardt, D. (2002). Destabilizing Networks. *Connections 24(3)*, 79-92.

Freeman, L. C. (1979). Centrality in Social Networks: I. Conceptual Clarification. *Social Networks, 1*, 215-239.

Heer, J., Card, S. K., & Landay, J. A. (2005). Prefuse: A toolkit for interactive information visualization. *SIGCHI Conference on Human factors in Computing Systems* (pp. 421-430). ACM.

Holland, P. W., & Leinhardt, S. (1971). Transitivity in structural models of small groups. *Comparative Group Studies, 2*, 107-124.

Holland, P. W., & Leinhardt, S. (1972). Some evidence on the transitivity of positive interpersonal sentiment. *American Journal of Sociology.* 72, 1205-1209.

Klerks, P. (2001). The Network paradigm Applied to criminal Organizations: Theoretical nitpickign or a

relevant doctrine for investigators? Recent developments in the Netherlands. *Connections 24(3)*, 53-65.

Krebs, V. E. (2002). Mapping networks of terrorists. *Connections 24(3)*, 43-52.

Moreno, J. L. (1978). *Who shall survive?: Foundations of sociometry, group psychotherapy, and sociodrama.* Beacon, NY: Beacon House.

Newman, M. E. (2003, June). The structure and function of complex networks. *SIAM Review, 45*, 167-256.

Qiu, B., Ivanova, K., Yen, J., Liu, P., & Ritter, F. E. (2011). Event-driven modelling of evolving social networks. *Int. J. Social Computing and Cyber-Physical Systems, 1*(1), 13-32.

Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika, 31*, 581-603.

Shemanski, D. R. (2011). *Stop the Terrorists! Team-based Simulation of an International Terrorist Plot to Acquire and Use a Weapon of Mass Destruction.* ACS Tech Report 2011-1.

Yang, H.-L., & Tang, J.-H. (2003). Effects of social network on students' performance: A web-based forum study in Taiwan. *Journal of Asynchronous Learning Networks, 8*(3), 93-107.

Author Biographies

RAZVAN ORENDOVICI is a graduate of the Master's program in Computer Science and Engineering from the Pennsylvania State University and is now working at Microsoft.

FRANK E. RITTER is a professor in the College of Information Sciences and Technology at the Pennsylvania State University.