1-1-2009

The Applicability of Simulated Network Attack Results on Small, Terrorist Networks

Stephanie Young
Washington University in St. Louis

Washington University

Department of Political Economy and Public Policy

The Applicability of Simulated Network Attack Results on Small, Terrorist Networks

by

Stephanie Young

A thesis presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the
degree of Master of Arts

August 2009

Saint Louis, Missouri

Contents

List of Tables

List of Figures

1 Introduction

The terrorist attacks of September 11, 2001 spawned a greater interest in utilizing network science to create policies and defense strategies that would effectively deter future threats to homeland security. The listserv associated with the leading social network organization International Network for Social Network Analysis was bombarded with inquiries about the role of social network analysis in the war against terrorism.

The controversial Total Information Awareness Program (later renamed the Terrorism Information Awareness Program) was established in 2002 under Adm. John Poindexter in the Defense Advanced Research Projects Agency in an early attempt to mine large volumes of data to conduct subject-oriented link analysis and identify terrorists and terrorist groups.

When the National Security Agency began its warrant-less eavesdropping program and intercepted some 650 million communications worldwide every day, it was hoped that network analysis would help to focus its efforts and surveillances.[1] Researchers such as Kathleen Carley (2003; Carley, Lee, and Krackhardt 2002) looked into dynamic network analysis and multi-agent modeling to destabilize terrorist groups and economics researchers such as Lindelauf, Borm, and Hamers (2008) sought to understand the incentives behind the structuring of covert groups that would aid in predicting terrorist behavior.

The nature of terrorist, covert, and clandestine organizations lends itself to network analysis because its architecture is not dependent on predictable hierarchies intrinsic in formal states and institutional adversaries that the United States have traditionally faced, yet it must also exist with an underlying order and adhere to robust organizing principles of long-standing graph theory established by mathematicians and physicists. Theoretically, by mapping connections between individuals and analyzing the structure of their social networks, law enforcement agents should be more able to identify key figures of terrorist organizations and disconnect the network in an efficient manner. Eliminating important individuals in any network, in addition to disrupting coordinated activities, means depriving the network of resources such as expertise and funding. Studying network configurations also informs as to why terrorist groups and clandestine cells can be so difficult to infiltrate.

2 Brief Background on Network Science

Despite the current hoopla about network analysis today, it has been present and studied in some form for centuries and recognized as an important tool in fighting terrorism even before the attacks of September 11. Network analysis as a means of analyzing complex relational data traces as far back as Leonhard Euler's paper "Seven Bridges of Königsberg" in 1735 which presaged the idea of topology and laid the foundations of graph theory, a branch of mathematics that studies the properties of pair-wise relations in a network structure.

Forerunners of modern concepts of social network analysis, an important offshoot of network analysis that concretizes theoretical network graphs as sociographs, include the late 1800s works of Ferdinand Tönnies and Émile Durkheim. Tönnies conceptualizes the types of links that could exist between persons in a social group as either gemeinschaft, which are links derived from shared values and beliefs, or geselleschaft, which are derived from impersonal, formal, and instrumental relationships. Like modern analyses, Durkheim explains social phenomena independently of the properties of individual actors.

Contemporary social network analysis is the offspring of Stanley Milgram's 1967 small-world experiment, which mapped chains of acquaintances and established the concept of six degrees of separation. In the 1973 article "The Strength of Weak Ties," Mark Granovetter propels subsequent investigations into network links when he argues that more novel information flows between individuals through weak ties than through strong ties and as a result, more disperse, non-redundant networks have greater access to information and power than smaller, denser, and more interconnected networks because they supply more diversity of knowledge and information.

More recently, Duncan J. Watts and Steven Strogatz expands upon the discipline by presenting a mathematical theory of Milgram's small-world phenomenon (1998) while Kathleen Carley combines social network analysis, link analysis, and multi-agent systems to create dynamic network analysis, where the nodes and links in a network exist only probabilistically. Carley heads Computational Analysis of Social and Organizational Systems at Carnegie Mellon University, one of the largest computational model organizations that models terrorist networks. Using a computational tool called DyNet and various network models, she and her associates study how to destabilize terrorist networks, they study chemical and biological warfare effects on the populace, and the effectiveness of wiretapping programs in mapping the networks of rapidly evolving covert organizations.

Other computer scientists, mathematicians, and physicists have also come up with various models for simulating networks so that they may be systematically studied. All these simulations are used to study subjects ranging from a network's structural weaknesses to the evolution of cell formations and defensive mechanisms that can be employed by the network against various attacks. Despite the differing methodologies employed by various researchers, from social analysis to dynamic network analysis, these methods can and have been used to find the effects of attacks on the network structure, the weaknesses of different networks and how to exploit or negate those weaknesses, which are of particular interest to the military.

The approach of these recent studies complements much of the more traditional social scientific studies. In comparison to more conventional disciplines such as sociology and political science, which tends to focus more on the traits of individual actors, network theory and its derivatives focuses on the value of network structures in explaining social phenomena. Social network analysis propagates a view that characteristics of individuals are less important than their relationships and ties with other actors within the network; values, norms, and culture are devalued. Despite nearly eliminating the ability for individuals to influence their fate, this method has yielded useful explanations for real-world phenomena.

Governments have long understood that modern warfare and crises essentially necessitate superior knowledge of networks. The United States government began following communications traffic during World War II, not for its content, but for its pattern and continues to do so to this day (Meter 2001). In 1960s the CIA in Thailand mapped out the clandestine structure of local and regional Communist organizations and associated sympathetic groups. In 2004 the Department of Defense and the U.S. Army at the United States Military Academy established the Network Science Center to fund various research projects in the area of Network Science. Table 1 provides a partial list of major challenges identified by the United States military and provided by the Committee on Network Science for Future Army Applications in 2005, consisting of network research areas and objectives applicable to military operations.

Understanding the structural basis and its inherent connectivity behind natural occurrences informs on a multitude of things, from the spread of contagion, to various social processes including the evolution of organizations, many of which now span across continents and between classes. Network science has strategic applications for various anti-terrorism U.S. policies. It has the potential to inform us as to where we should direct our resources in a multitude of issues and provide us with increased situational awareness.

Research Area	Key Objective
Modeling, simulating, testing, and prototyping very large networks	Practical deployment tool sets
Command and control of joint/combined networked forces	Networked properties of connected heterogeneous systems
Impact of network structure on organizational behavior	Dynamics of networked organizational behavior
Security and information assurance of networks	Properties of networks that enhance survival
Relationship of network structure to scalability and reliability	Characteristics of robust or dominant networks
Managing network complexity	Properties of networks that promote simplicity and connectivity
Improving shared situational awareness of networked elements	Self-synchronization of networks
Enhanced network-centric mission effectiveness	Individual and organizational training designs
Advanced network-based sensor fusion	Impact of control systems theory
Hunter-prey relationships	Algorithms and models for adversary behaviors
Swarming behavior	Self-organizing unmanned aerial vehicles/unmanned ground vehicle; self-healing
Metabolic and gene expression networks	Soldier performance enhancement

Table 1: Network Research Areas of Importance to the Military

3 Previous Research on Networks and Attack Simulations

The focus of this paper will be on the applicability of research based on network and attack simulations on small terrorist networks. Many researchers and commentators take the results of most network and attack simulations in research papers to implicitly apply to small terrorist cells.[2] But without testing attack strategies on empirical, small terrorist networks, we cannot be sure that this is true.

Albert, Jeong, and Barabási (2000) find that complex systems display a surprising degree of tolerance for errors; local failures of key components rarely lead to the loss of the global information-carrying ability of the network. This robustness is not due simply to redundant wiring, but is also a function of the network topology. In particular, scale-free networks are more robust against random network attacks are more robust against attacks targeted using node degrees.

Holme *et. al* (2002) expands Albert, Jeong, and Barabási's findings to include attacks on edges. They also focus on differences between local and global strategies of attacks, where local ones concentrate on eliminating the total number of edges and the global ones concentrate on eliminating as many geodesics as possible. They find that the changes in network topology are substantial under attacks and that the most effective strategies take these changes into account. They also reaffirm Albert, Jeong, and Barabási's findings and show that once a certain number of edges or vertices are removed, the network is disconnected.

Cohen *et. al* (2000) develop the first analytical approach to calculating the critical threshold for fragmentation of network experiencing random node failures while Callaway *et. al* (2000) independently propose an alternative approach. The results of their research support Albert, Jeong, and Barabasi's finding that scale-free networks become fragmented after a small fraction of high-degree nodes are removed.

Nagaraja and Anderson in 2006 at the Fifth Workshop on the Economics of Information Security demonstrate that network structure can influence the evolution of incentives of attackers and network defenders by simulating a network of 400 nodes with the ability to replenish itself and evolve, essentially defending itself from attacks. In contrast to many previous researchers, the network they study is not static and they explicitly focus on the anti-terrorist applications of network science. Previous researchers had focused on theoretical results and left others to discuss much of the implied applications of their work.

Dall'Asta, Barrat, Barthélemy, and Vespignani's 2006 study of weighted networks, airports, and air traffic alludes to the hijackings of 9/11. Their research directs attention to where defensive policies should be concentrated by addressing which types of airports required greater protections and were at greater risk for attack. They also assess which measures are better suited for quantifying network damage in the context of air-traffic to characterize the most effective attack and protection strategies.

Research of network attacks on other applied levels include the Albert, Jeong, and Barabasi's (1999) study on the attack tolerance of the World Wide Web; Jeong *et. al* (2000) study on cellular networks, specifically the responses of metabolic networks of various organisms to random and preferential node removal; and the Solé and Montoya (2001) study of ecological networks, specifically the response of food webs to the removal of species.

4 Differences Between Networks in Existing Research and Real-life, Small Terrorist Networks

Despite all the research on network attacks, there are fundamental differences between the cellular structure of a terrorist group and a computer-generated social network. The simulations used in research papers usually run with hundreds and thousands of nodes. The results with more nodes are more generalizable and the nuances of the effects of attacks are easier to see. But while any results are applicable to its paper's respective area of study—Dall'Asta, Barrat, Barthélemy, and Vespignani's 2006 study of weighted networks on airport and air traffic; Cohen, Erez, ben-Avraham, and Havlin's 2000 study of random breakdowns on a power grid and Zhao, Park, and Lai's 2004 study of cascading breakdowns on the internet—they may not hold for significantly smaller networks. Researchers such as Arquilla and Ronfeldt (2001) write about the effectiveness of vertex order attacks against scale-free networks by prematurely generalizing the results of studies on large, simulated networks without testing that assumption on smaller simulated networks first.

But while the number of terrorist sympathizers and terrorists in the world may be thousands, even millions, and may all somehow be linked with one another through internet forums, websites, and training camps, the active groups of people who commit crimes against formal states must operate in much smaller cliques and for the most part, those cliques function independently of one another. These small social groups that come together to operate and execute attacks on the United States in the near to immediate future are the ones I refer to and focus on in this paper, not the massive groups of people who might simply be guilty of terrorism by association.

Dunbar (1992; 1993) suggests that there is a theoretical cognitive limit to the number of people with whom one can maintain stable social relationships. He predicts that 147.8 is roughly the mean group size for humans with a very high incentive to stay together, which is consistent with census data on various village and tribe sizes of many cultures that he studied. (Today, the most commonly cited approximation of Dunbar's number is 150.) He suggests that groups that are considerably larger than this must expend more effort to socially groom to maintain social cohesion; when grooming fails to stem dissatisfaction and dissension, the group may partition into subgroups that may or may not be affiliated with one another.

This predicted size dynamic is evident in various criminal organizations. The largest mafia family in New York, the Genoveses totaled 152 members in 2002. The Gambinos, after losing 33 members in 2000-2001, totaled 130 in 2002, and the Luccheses followed closed behind with 113 members.[3] Islamic terrorist cells in Baghdad were estimated to be about 100 members organized into six cells. In the Anbar province, Maj. Gen. Charles H. Swannack Jr. of the 82nd Airborne Division reported about 50 to 80 foreign fighters operating in eight to 10 cells.[4]

If the sizes of terrorist cells that the government chooses to focus on truly are much smaller than the sizes of simulations being run in most research papers, then the results of large network simulations might not be wholly applicable. For example, technically different attack strategies might not be practically different; an attack strategy based on the vertex order might be extremely similar to an attack strategy based on betweenness centrality since both measures, though not equivalent, are related. Previous research shows that attack strategies should vary according to the network structure since the speed at which a network becomes fragmented is a function of the interaction of those two factors. But the heterogeneity of smaller networks is limited and the effects of attack strategies may converge as a result. Using an attack strategy that has been deemed more efficient in large network simulations might not confer any real payoffs to the attacker when it is used on a much smaller network.

Additionally, despite much recent research describing how we should strike at various network configurations, the difficulty in employing the tactics described by the research on revolutionary groups is that we usually lack complete information about the terrorist, covert and clandestine cells we wish to apply the attacks. For example, most of the structure of al-Qaeda these days is unknown. What little information we have comes from Jamal al-Fadl, a former associate of Osama bin Laden, through his cooperation with American authorities and even still the veracity of his testimony is disputed.[5] After all, the point of both covert and clandestine cell structures is that its details are completely hidden from the opposition.

Krebs (2002) highlights information problems when he maps the social networks of the hijackers involved with the 9/11 attacks. He links the individuals involved through shared addrsses, telephone numers, frequent-flier numers, and finds that while all 19 hijackers were only sparsely linked, a number of links converged on the ringleader Mohamed Atta. But in doing so, he discovers that such mapping is difficult because the media is plagued with inaccurate information. He acknowledges that he inevitably misses links and nodes and that it is difficult to determine which individuals associated with the hijackers should and should not be included in his mapping. He concludes in his paper that analyzing networks after an event if fairly straightforward for prosecutorial purposes, but mapping covert networks to prevent criminal activity is significantly more difficult.

One goal in this paper will be to see whether there are, as in large networks, significant differences in the speed of the decay of a small network given its architecture and given the nature of attacks. The two types of networks I use are Barabási and Albert's (1999) model and Erdõs and Rényi's (1959) model. The two types of attacks I implement are a random attack and a degree-based attack. These are chosen because despite the ubiquitous literature on the varieties of attacks and structures and the minutiae of attack effects on structures, there is a general consensus amongst the literature that for unweighted, undirected graphs, the type of graph generated by Barabási and Albert's is considerably less resilient against degree-based attacks but much more resilient against random attacks in comparison to Erdõs and Rényi's model. The point is not to find the most effective attack or to dispute the results of previous research, but merely to question whether those results generally apply to much smaller networks. And as mentioned earlier, I will also address whether, given that we have a small network, varying levels of information affect the policy we should take in our network attack efforts and how much information levels affect our ability to predict attack outcomes. Small network size and limited information availability are both features that are incongruous with the simulations used in most research papers. Hopefully this paper will reveal how applicable the results of those simulations are on small terrorist cells.

5 Definitions of Quantities and Terms

In general, the complex networks studied here and in much of the modern literature can be represented in an undirected and unweighted graph, $G = (V, E)$, where V is the set of vertices (or nodes), E is the set of edges (or links), and $N = |V|$ denotes the number of vertices in the network. Every element of the set E is mapped to an ordered pair of points from the set $V \times V$ so that each edge connects exactly one pair of vertices and a vertex-pair can be directly connected by maximally one edge. The concept of vertices and nodes is abstract, but for a social network, V is the set of actors and E is the set of interpersonal ties defined as information-carrying connections between people.

For an undirected graph, the degree of a node $v \in V$ is the number of edges incident on that node. The degree is also synonymous with the vertex order. Paths in an undirected graph are alternating sequences of vertices and edges in which an edge between v_i and v_j is incident on both. The betweenness centrality, denoted $C_B(v)$ is a weighted measure of the number of shortest paths node v sits on and calculated by:

$$C_B(v) = \sum_{\substack{s \neq v \neq t \in V \\ s \neq t}} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

where σ_{st} is the number of shortest paths from node s to node t and $\sigma_{st}(v)$ is the number of shortest paths from s to t that pass through vertex v.

The data we will use to compare the effects of the attack are the size of the giant component and average inverse geodesic length. These are two measures of network functionality. The giant component size measures the number of nodes in the largest connected subgraph. The largest subgraph poses the greatest threat because it likely retains the most resources of its nodes. Henceforth, the size of largest connected component will be denoted S'.

To understand the average inverse geodesic length, it is necessary to first explain the average geodesic length. The average geodesic length is defined by:

$$\ell \equiv \langle d(v,w) \rangle \equiv \frac{1}{N(N-1)/2} \sum_{v \in V} \sum_{w \neq v \in V} d(v,w),$$

where $d(v,w)$ is the geodesic length (the number of edges in the shortest path) between v and w for $v, w \in V$, and $N(N-1)/2$ is the number of pairs of vertices. If there is no path between v and w, the geodesic length is taken to be infinite. The larger ℓ is, the slower the dynamics and information flow in the network.

The average inverse geodesic length, not the reciprocal of the average geodesic length, is defined by:

$$\ell^{-1} \equiv \left\langle \frac{1}{d(v,w)} \right\rangle \equiv \frac{1}{N(N-1)/2} \sum_{v \in V} \sum_{w \neq v \in V} \frac{1}{d(v,w)}, \tag{1}$$

If nodes v and w are disjoint, then $d(v,w) = \infty$ and $1/d(v,w) = 0$. The larger ℓ^{-1} is, the better the network functions. The average inverse geodesic length provides insight to the internal mechanics of the network; falls in its value indicate increasing difficulty in internode communication. Low network efficiency could be deliberately created by terrorists in exchange for greater network resilience so that captured members would be have less information to provide the enemy. But that is not always the case and at some point, less network efficiency benefits the enemies of the terrorists.

6 Networks

The types of networks that studied in this paper are the random network and the scale-free network. These two types of graphs have very divergent characteristics, but both have been consistently and thoroughly studied in the past and they will be used both to assess how arrangement and incomplete information changes attack effects for small networks.

The random graphs are generated by some random process and have been studied largely because they have interesting mathematical properties and because many of their properties can be determine using probabilistic arguments. Random-graph theory is also regularly used in the study of complex networks because since networks with a complex topology and unknown organizing principles often appear random. Random graphs can also be manipulated to exhibit various characteristics.

The scale-free networks are so-called because their degree-distributions asymptotically follow a power-law $P(k) \sim k^{-\gamma}$, where $P(k)$ is the fraction of nodes in the network with degree k and γ is a constant. This fundamental characteristic is what allows the features of scale-free networks to parallel those of many real-world networks, such as the World Wide Web, Internet, cellular networks, film-actor collaboration networks, science collaboration networks, citation networks, and other social networks. A large scale-free network exhibits high-degree nodes called "hubs," with major hubs followed by smaller hubs. Low-degree nodes in the network connect to rather dense sub-graphs and those sub-graphs, in turn, connect to each other through hubs.

While it is technically possible to build random graphs with power-law distributions, the core difference between mechanisms that have been used to generate random graphs and those that have been used to generate scale-free networks is the approach-random graph generators model network topology while scale-free graph generators models network assembly and evolution. The goal of the former models is to construct a graph with correct topological features; the goal of the latter models attempt to capture network dynamics. The underlying assumption between the evolving or dynamic network models, such as those that generate scale-free networks, is that if the processes that generated the networks are captured, then the correct topology will be obtained as well. Dynamics take the driving role and topology is only a byproduct of the modeling philosophy. This is evident later when the simulation models are set up.

7 Network Setups

I create the scale-free network using Barabási and Albert's (1999) generative algorithm in the following steps:

1. Randomly connect 5 initial nodes.

2. For 145 rounds, a new node with 2 edges is added each round. The probability that a new node connects with node i is $\prod(k_i) = \frac{k_i}{\sum_j k_j}$ where k_i is the degree of node i.

Our network size is small, but the Barabási-Albert (BA) model is one of several models meant to generate scale-free networks and is likely to be consistent with a social group whose organization develops organically. It incorporates two important general concepts that are believed to contribute to the clustering and organization seen in real-world networks: growth and preferential attachment. Growth means that the size of the network increases over time. Preferential attachment means that the more connections a node already has determines how likely it will receive links with new, added nodes. This is analogous to the social phenomenon in which popular people, by being more visible, are more likely to make new contacts and gain more friends while more isolated people would have limited opportunities to meet newcomers to a community. Likewise, new websites link preferentially to traffic hubs.

I create a random network using the $G(n, p)$ variant of the classical Erdõs-Rényi (1959) algorithm, where n is the number of nodes in the graph and p is a value between zero and one. In this model, a graph is constructed by randomly connecting nodes; each edge in the graph is exists with a fixed probability p and completely independent of any other edges. The parameters chosen for the ER model are such that the average degree per node is approximately the average degree per node for the BA model.

Erdõs and Rényi's model (ER model) produces random graphs with interesting mathematical properties, but have very limited real-world characteristics. A network like this is more likely to occur in cases where members of a group deliberately construct the group as a random graph as a defensive mechanism. Deliberate structuring of a group is not unheard of; hierarchies are examples of organizational structuring employed nearly universally and according to the *al-*

Qaeda Training Manual al-Qaeda's minimal core group is prearranged in as a ring or chain network.

8 Attack Simulations

Technically, there are multitudes of ways the government can attack; it can strategically remove nodes from the network or focus on specific edges and connections between the nodes. The government also has various strategies it can employ and those strategies can vary by round. For the purposes of this paper, the two forms of attacks we will employ will be a random attack and an attack based on each node's vertex order.

In the random attack, nodes that can be attacked are randomly selected with equal probability and disconnected from the system. Once a node has been attacked, it cannot be selected again. In the vertex order attack, the attacker selects the node with the highest degree and disconnects it from the system. In general, random attacks are considered easier for the attacker to initiate in the sense that a real-world attacker would not have to exert much effort to decide which individual to target. While degrees are not difficult to measure, they require more time and resources to be used in order to implement a vertex order attack.

Both attack strategies will be employed on each type of network and all attacks in the simulations are performed in rounds. In each round of attacks, the attacker will remove a fixed number of nodes from the network. An attack strategy is considered more efficient if fewer rounds are necessary to disconnect the terrorists.

The data from the resulting networks after each round of attacks will be recorded and plotted so that the effect of attacks on each type of network can be compared.

Initially, the simulations will be run assuming an attacker has complete information about the network. In subsequent simulations, the attacker will have less and less information about the network. Of the nodes that are available, the attacker will only be able to see a percent of that network and those are the nodes that can be attacked; the attacker will only be able to attack nodes it sees and it cannot see edges to nodes it does not see. The nodes that are masked are randomly chosen and all have an equal chance of being hidden; whether a node is hidden is independent of whether any other nodes are hidden. In the real world, even if the government can see all the individuals suspected terrorists interact with, not all acquaintances of a suspected terrorist must be suspected of terrorism, too.

9 BA Graph vs. ER Graph

9.1 Random Attacks

The results of the random attack simulations on the BA and ER models by an attacker with complete information are superimposed in Figure 1.

Surprisingly, under random attacks, the network structured in accordance with the BA model deteriorates just as fast as, if not faster than, the random network generated by the ER model. Differences in the slope of both S' and ℓ^{-1} are nearly indistinguishable for a majority of rounds. The decay in S' is practically linear for both models. The trend of ℓ^{-1} is similar for both as well, with the changes in

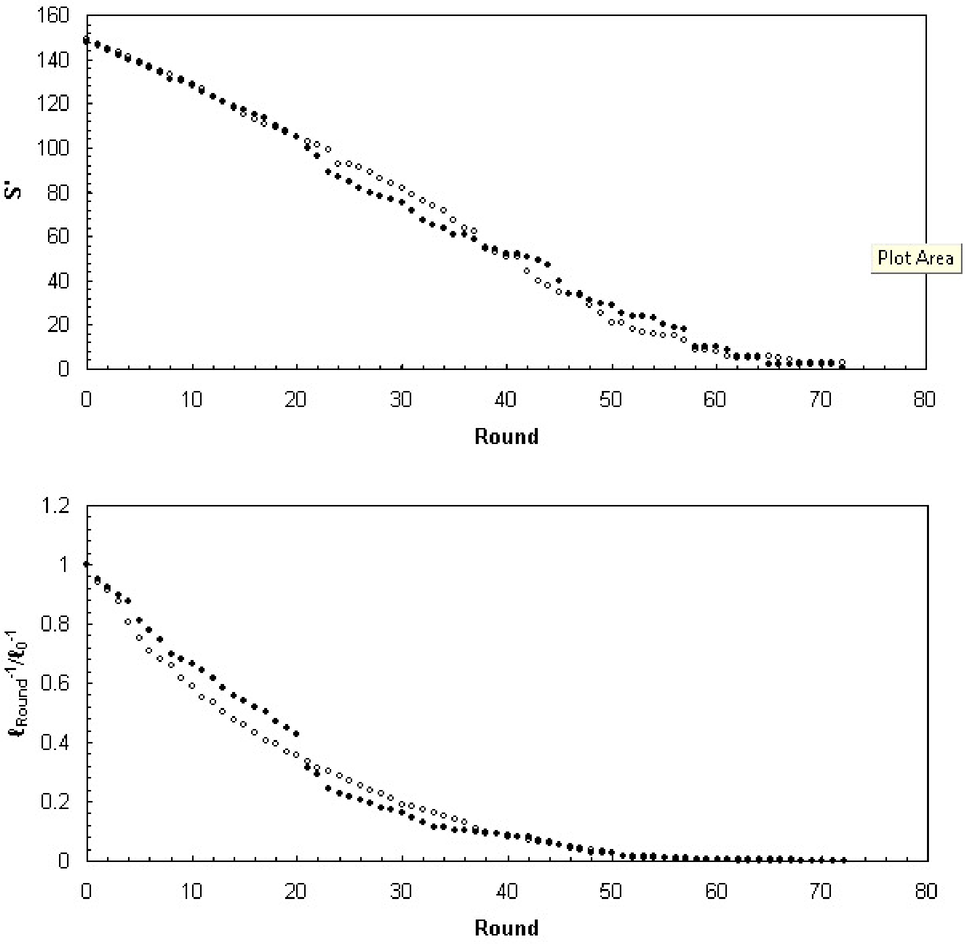


Figure 1: The results of the simulation on a network structured by the BA model are designated by solid dots. The results of the simulation on a network structured by the ER model are designated by hollow circles.

ℓ^{-1} decreasing over time.

Typically, a larger network constructed under the BA model would have maintained greater resilience against random attacks because a vast majority of the nodes are of small degree; the chances that a random attack strikes a well-connected node are much lower than the chances that a random attack strikes a low-degree node. And even if we remove a hub from the network, the network would not lose its connectedness because remaining hubs would maintain the connection.

However, since the BA network in our simulation is so small, the most connected nodes are more vulnerable and the probability that a high-degree node is selected randomly is much greater. There are also much fewer clusters that can form make up for the loss of a relatively connected node. This increased vulnerability, coupled with the lack of fewer hubs to make up for lost links makes the network more vulnerable than would have been predicted given the results of research done with thousands of nodes.

There is a lot of hope that network science would help us strategically target important individuals driving revolutionary movements, but this particular simulation demonstrates that the initial enthusiasm might have been premature. Network simulations and tests are done in generic and theoretical terms that might have limited applicability on real, small terrorist networks.

Even if we model the networks with excruciating attention to detail and add different node characteristics to proxy human traits and roles in the network, it might be that such information, assuming it is available, does not yield results proportionate to the effort that is required. Gathering information on private citizens without sufficient cause might not always be as easy as it was immediately following 9/11. Furthermore, if we knew all the specific traits about each actor in the network, we might not need network science to help us to figure out who the key players in the network are—we'd already know who they are. Part of network science's appeal is that it allows us to understand a network without looking at the actors. The goal should be to understand when we can use it to our advantage and when its power is limited.

9.2 Vertex Order Attacks

The results of the vertex order attack simulations on the BA and ER model by an attacker with complete information are superimposed in Figure 2.

In contrast to the previous simulation on random attacks, the vertex order simulations appear to confirm previous findings—the ER model is more resilient against vertex order attacks than the BA model under complete information. It only takes 8 rounds to break the network in half under the BA model, but takes more than twice that many rounds in order to halve the ER model. The S' decreases at an almost monotonically increasing rate until the attacks halve the network under the BA model while the S' decreases at a generally increasing, but very inconsistent, rate before the attacks halve the network under the ER model. After the network is split, the S' decreases at a decelerating then inconsistent rate under the BA model and decreases at a variable and inconsistent rate under the ER model. Throughout the simulation, $S'_{BA} \geq S'_{ER}$.

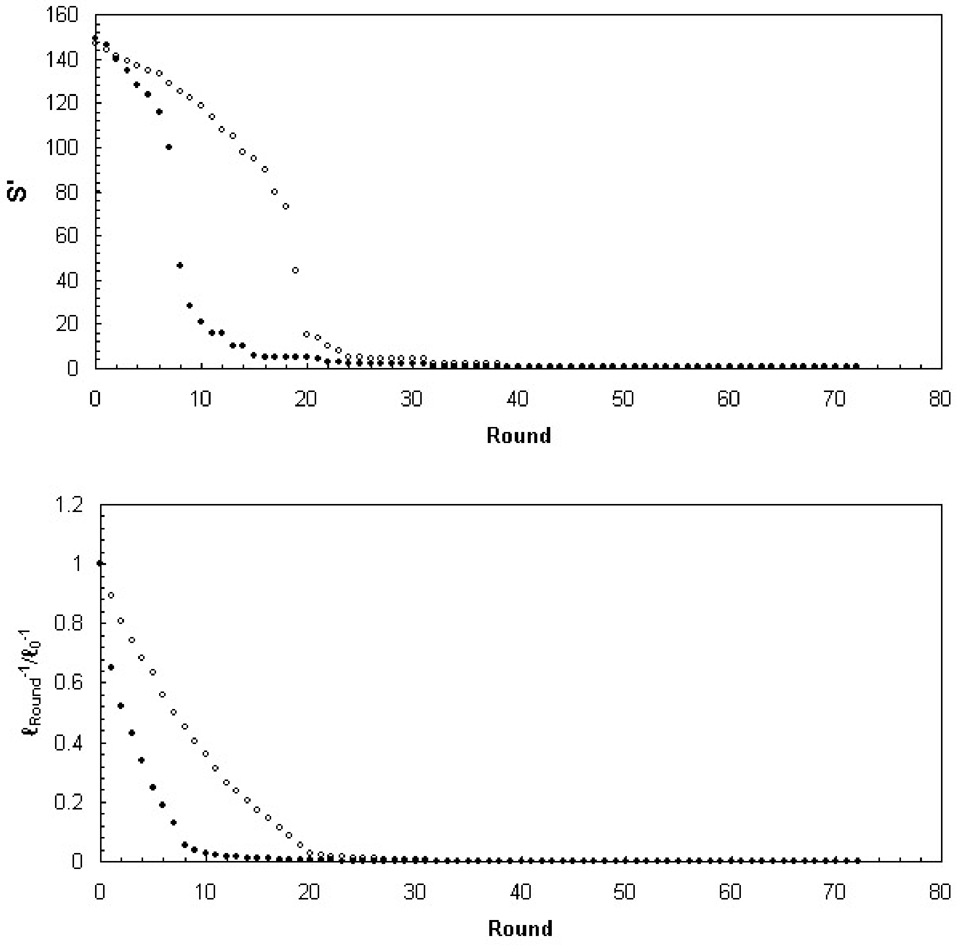


Figure 2: The results of the simulation on a network structured by the BA model are designated by solid dots. The results of the simulation on a network structured by the ER model are designated by hollow circles.

As for the progression of ℓ^{-1}, under the BA model, the inverse average geodesic length becomes zero after 32 rounds; the same is accomplished under the ER model only after 38 rounds. Much of the difference can be attributed to the rate at which shortest path lengths increase since many of the shortest paths are likely to cross through the highest vertex ordered nodes in the BA model while the shortest paths in the ER model are relatively more evenly distributed on different nodes. Thus, given that $S'_{BA} \leq S'_{ER}$ and $\ell^{-1}_{BA} \leq \ell^{-1}_{ER}$, it appears that the ER model is superior to the BA model in fending off non-random failures.

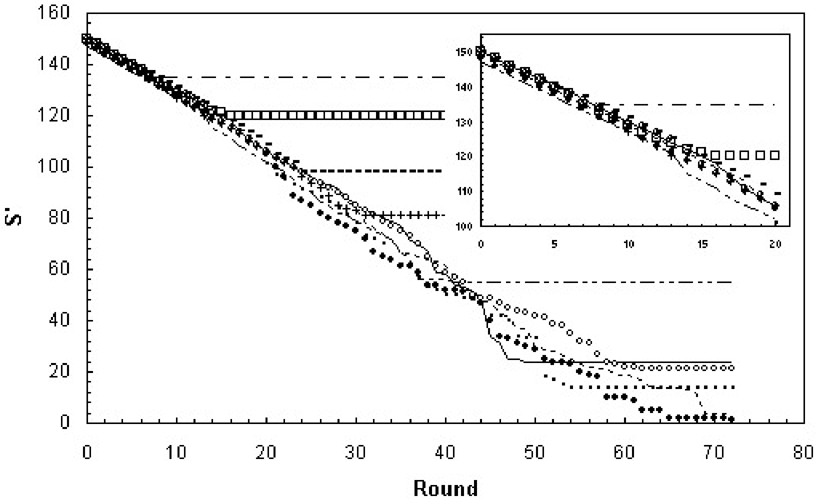
10 Incomplete Information

Additionally, all the previous data on the effectiveness of attacks were gathered assuming attackers know the entire network topology of interest. In reality, this assumption can be problematic. In light of Krebs's (2002) conclusions, it is necessary to assess how well simulated network attacks fare when the attacker has limited information about the network.

The results of network attacks on the BA and ER model are shown in Figures 3 through Figure 6.

10.1 Random Attacks Under Incomplete Information

Regardless of the information available to the attacker or the structure of the network, the effect of random attacks on a small network's giant component size appears to be linear. The pattern of effects on S' is similar; this is particularly



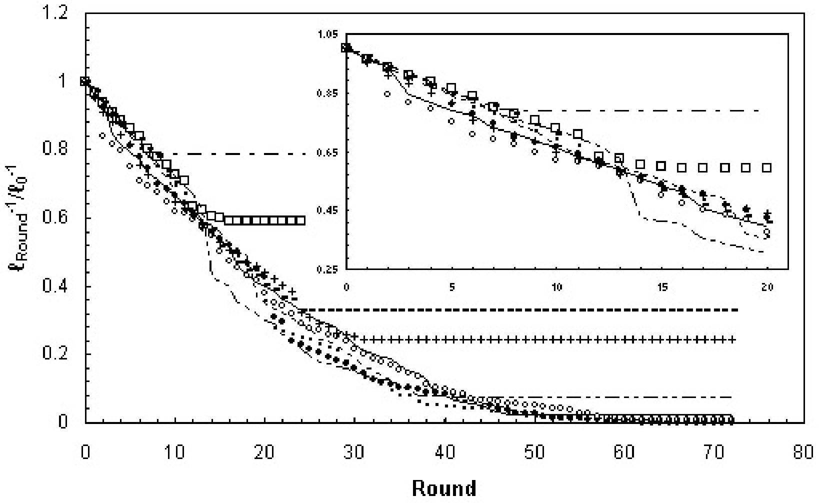


Figure 3: The results of the random attacks on the BA model are shown above. Network functionality is measured by the average inverse geodesic length ℓ^{-1} and the size of the giant component S'. Solid circles mark the progression of the attacks under perfect information. Dotted lines mark when 10% of the network is unseen. Hollow circles mark when 20% of the network is unseen. A square dot marks when 30% of the network is unseen. A solid line marks when 40% of the network is unseen. A line in a dot-dot-dash pattern marks when 50% of the network is unseen. Plus signs mark when 60% of the network is unseen. The small, solid rectangle marks when 70% of the network is unseen. A hollow square marks when 80% of the network is unseen. And a line in a dot-dash pattern marks when 90% of the network is unseen.

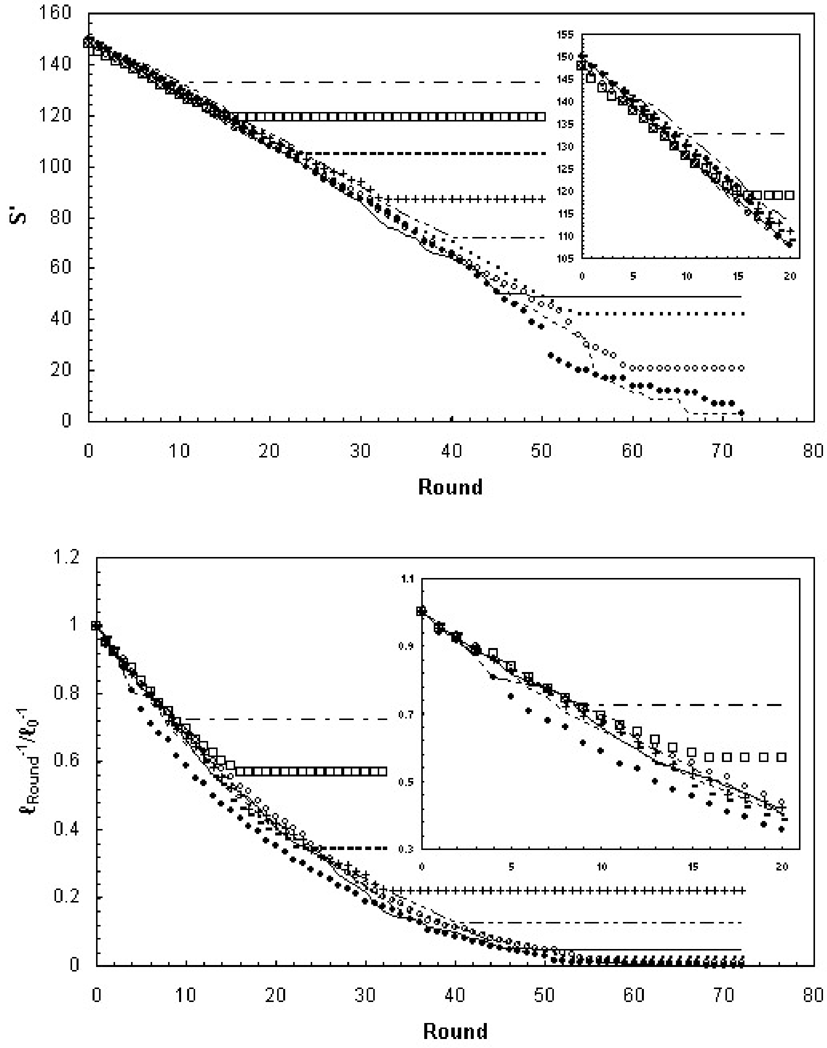


Figure 4: The results of the random attacks on the ER model are shown above. Network functionality is measured by the average inverse geodesic length ℓ^{-1} and the size of the giant component S'. Solid circles mark the progression of the attacks under perfect information. Dotted lines mark when 10% of the network is unseen. Hollow circles mark when 20% of the network is unseen. A square dot marks when 30% of the network is unseen. A solid line marks when 40% of the network is unseen. A line in a dot-dot-dash pattern marks when 50% of the network is unseen. Plus signs mark when 60% of the network is unseen. The small, solid rectangle marks when 70% of the network is unseen. A hollow square marks when 80% of the network is unseen. And a line in a dot-dash pattern marks when 90% of the network is unseen.

true for the earlier rounds. The evolution of S' is very uniform when fewer than 30 attacks (15 rounds) have taken place and remain relatively unvarying up around 60 attacks (30 rounds). This suggests it is not entirely necessary to completely determine the network structure if the government's goal is to simply fractionalize a small terrorist network; it also implies that the government's efforts can be more cost-effective than recognized by previous research.

The S' also both models also plateau at similar levels, though neither model has a distinct advantage in terms of final giant component size. Under some circumstances, the effects of random attacks plateau at much higher levels for the ER model that the BA model so that much more of the network stays functional after the attacker runs out of information and nodes to attack and under others, the reverse is true; there is no general rule that dictates which model plateaus higher in terms of S'.

The effect of the attacks on the average inverse geodesic length on both the BA and ER model are non-linear; on average the decreases in the percent from the initial ℓ^{-1} are steeper and relatively constant for each information level available to the attacker for the initial rounds, though after around 15 rounds, it the effects of the attack become less homogeneous and more subtle and slowly flatten the percent changes of ℓ^{-1}.

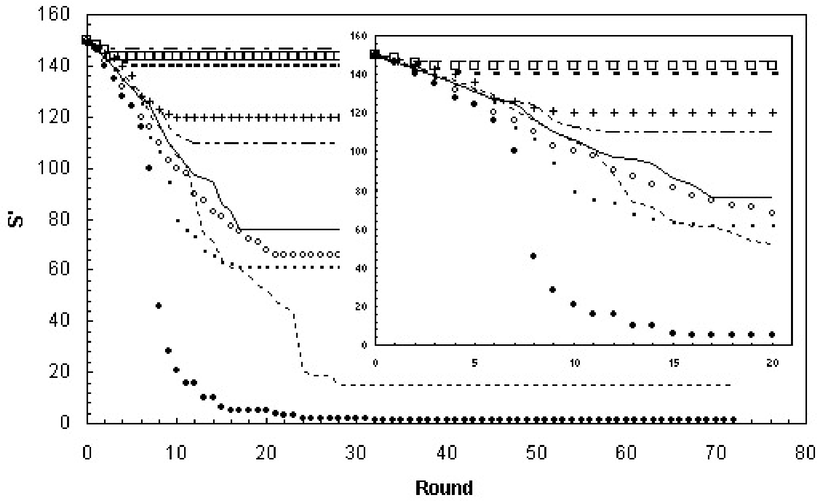
The ER model does not appear to be any more resilient than the BA model. The effect of attacks on the percent from the initial ℓ^{-1} converge when 10% to 40% of information is unavailable to the attacker. But when less of the network is seen by the attacker, it appears that the BA model eclipses the ER model in terms of maintained functionality as measured by ℓ^{-1}. The greater discrepancies in the final levels of ℓ^{-1} between the BA and ER model occur when fewer nodes are visible, whereas no such rule applies to the development of S'.

10.2 Vertex Order Attacks Under Incomplete Information

The effect of vertex order attacks on both models are more drastic than the effects of the random attacks by both measures of functionality, S' and ℓ^{-1}. In contrast to the random attacks, the effect of degree-based attacks on both the BA and ER model accelerates much more quickly.

It appears at first glance that the BA model is disproportionately hurt by the vertex order attacks. The overall effects of the attacks on S' always surface quicker for the BA model, but the more rounds there are and the more information the attacker has the more limited the BA model's disadvantage becomes. The S'_{BA} is actually capable of plateauing at equal or higher levels than S'_{ER}, regardless of information level. And once the attacker is unable to see 50% or more of the population, there is relatively little difference in final S' between the BA and ER model.

By measures of percent change in ℓ^{-1}, the attacks disproportionately disadvantages the BA model, especially in the earlier rounds and when the attacker has more information about the network. The BA model suffers significantly greater declines over fewer rounds in percent change in ℓ^{-1} when 60% or fewer of the nodes are unseen, but the final level of ℓ^{-1} tends to be lower for BA model. The fact that the ER model is better able to retain its initial level of ℓ^{-1} is due largely to the fact



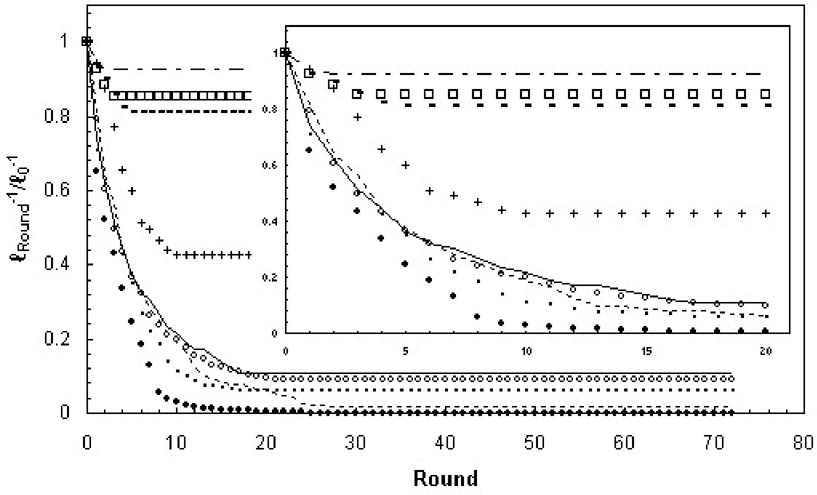


Figure 5: The results of the vertex order attacks on the BA model are shown above. Network functionality is measured by the average inverse geodesic length ℓ^{-1} and the size of the giant component S'. Solid circles mark the progression of the attacks under perfect information. Dotted lines mark when 10% of the network is unseen. Hollow circles mark when 20% of the network is unseen. A square dot marks when 30% of the network is unseen. A solid line marks when 40% of the network is unseen. A line in a dot-dot-dash pattern marks when 50% of the network is unseen. Plus signs mark when 60% of the network is unseen. The small, solid rectangle marks when 70% of the network is unseen. A hollow square marks when 80% of the network is unseen. And a line in a dot-dash pattern marks when 90% of the network is unseen.

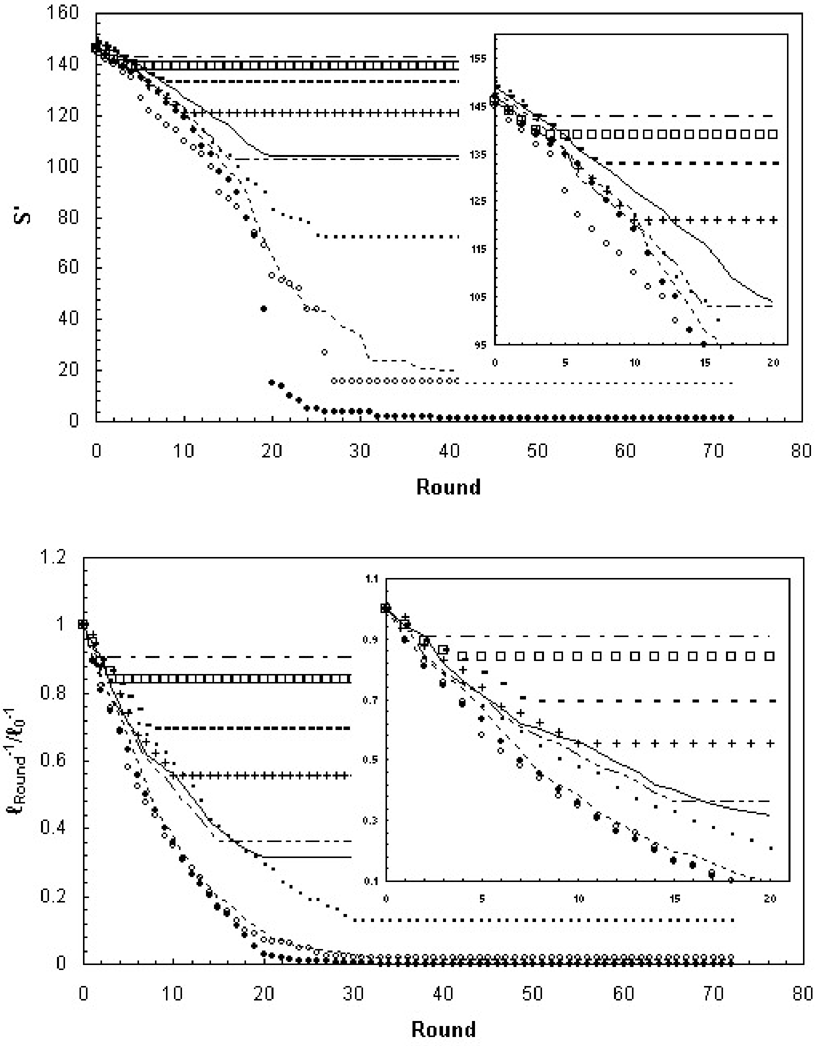


Figure 6: The results of the vertex order attacks on the ER model are shown above. Network functionality is measured by the average inverse geodesic length ℓ^{-1} and the size of the giant component S'. Solid circles mark the progression of the attacks under perfect information. Dotted lines mark when 10% of the network is unseen. Hollow circles mark when 20% of the network is unseen. A square dot marks when 30% of the network is unseen. A solid line marks when 40% of the network is unseen. A line in a dot-dot-dash pattern marks when 50% of the network is unseen. Plus signs mark when 60% of the network is unseen. The small, solid rectangle marks when 70% of the network is unseen. A hollow square marks when 80% of the network is unseen. And a line in a dot-dash pattern marks when 90% of the network is unseen.

that information levels have a much more systematic effect on the ℓ^{-1} of the ER model than the BA model. Decreasing levels of information strictly increases final ℓ^{-1} for both models, but the ℓ^{-1} under the BA model tends to collapse so long as at least 60% of the network is visible while the change in ℓ^{-1} for the ER model is more homogenous so that even modest protection from the attacker affords the network greater resilience. This result shows that the ER model's advantage over the BA model against vertex order attacks holds relatively well even when the attacker's sight is limited.

11 Conclusion

It is interesting to note that while size of a network impacts the research outcomes, and though, in general, the more information that the attacker is missing about the network, the less effective his overall attack strategies are, attack strategies remain effective even if large portions of the network are unseen. These counterintuitive results are interesting, but not completely surprising. While many commentators in the media were enthusiastic about network science, academics have consistently cautioned that the field is still in its infancy and that its conclusions are very context-specific, especially since any of the measures of centrality people are using to find key nodes in any network are very sensitive to the addition of new nodes or edges.

Network science is not without its strengths. The results of the attack simulations bodes well for the application of network science to terrorist networks since size, which does impact results, is easily corrected for in future research and since attack strategies can remain effective in the face of incomplete information. For random attacks, the differences in information levels do change the maximum amount of damage the attacks can do, but the trends of network deterioration in terms of inverse geodesic lengths and giant component sizes, in general, do not change. For vertex order attacks, information levels are of greater importance as it affects both the maximum amount of damage the attacks can do, but the speed of network deterioration.

The results are largely preliminary, but are important to keep in mind whenever people hope to extrapolate the results of network science. Future work can build on this by taking into account the fact that how much an attacker knows about a network is not random and that not all nodes are seen or unseen with equal probability. An intelligent attacker would be more likely to see individuals that are commonly linked to known terrorists and less likely to see those that are attached to unseen nodes. Additionally, it remains to be seen precisely at what size the results of large network simulation no longer apply. Network science is still a growing discipline and much of its contribution remains to be seen. But its interdisciplinary nature makes it particularly versatile and powerful. In order for it to be useful, it is our duty to determine when, where, how, and if we should utilize it. This paper, hopefully, contributes to that endeavor.

Notes

[1] Patrick R. Keefe, “Can Network Theory Thwart Terrorists?” New York Times, 12 March, 2006.

[2] Joel Garreau, “Disconnecting the Dots” Washington Post, 17 September 2001.

[3] “NYC Mafia Famlies Hold Recruitment Drive” BBC News, 20 May, 2002.

[4] Rajiv Chandrasekaran, “Iraq Attacks Blamed On Islamic Extremists,” Washington Post, 19 March, 2004.

[5] Johanna McGeary, Massimo Calabresi, and Elaine Shannon, “A Traitor’s Tale,” Time, 19 February, 2001.

References

[1] Albert, Réka, Hawoong Jeong, and Albert-Lásló Barabási. 1999. "Internet: Diameter of the World-Wide Web." *Nature* 401 (September): 130-31.

[2] Albert, Réka, Hawoong Jeong, and Albert-Lásló Barabási. 2000. "Error and Attack Tolerance of Complex Networks." *Nature* 406 (July): 378-82.

[3] Barabási, Albert-Lásló and Réka Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286 (October): 509-12.

[4] Arquilla, J. and D. Ronfeldt. 2001. /emphNetworks and Netwars: The Future of Terror, Crime, and Militancy. Santa Monica, CA: RAND.

[5] Callaway, Duncan S., M.E.J. Newman, Steven H. Strogatz, and Duncan J. Watts. 2000. "Network Robustness and Fragility: Percolation on Random Graphs." *Physical Review Letters* 85 (December): 5468 - 471.

[6] Carley, Kathleen M. 2003. "Destabilizing Terrorist Networks." Proceedings of the 8th International Command and Control Research and Technology Symposium, Washington, D.C.

[7] Carley, Kathleen M., Ju-Sung Lee, and David Krackhardt. 2002. "Destabilizing Networks." *Connections* 24 (3): 79-92.

[8] Cohen, Reuven, Keren Erez, Daniel ben-Avraham, and Shlomo Havlin. 2000. "Resilience of the Internet to Random Breakdowns." *Physical Review Letters* 85 (November): 4626-628.

[9] Committee on Network Science for Future Army Applications and the National Research Council. 2005. *Network Science.* Washington, D.C.: National Academies Press.

[10] Dall'Asta, Luca, Barrat Alain, Marc Barthélemy, and Alessandro Vespignani. 2006. "Vulnerability of Weighted Networks." *Journal of Statistical Mechanics: Theory and Experiment* 2006 (April): P04006.

[11] Dunbar, R.I.M. 1992. "Neocortex Size as a Constraint on Group Size in Primates." *Journal of Human Evolution* 22: 469-93.

[12] Dunbar, R.I.M. 1993. "Coevolution of Neocortical Size, Group Size, and Language in Humans." *Behavioral and Brain Sciences* 16 (4): 681-735.

[13] Erdõs, P. and A. Rényi. 1959. "On Random Graphs, I." *Publicationes Mathematicae* 6 (Debrecen): 290-97.

[14] Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78 (May): 1360-1380.

[15] Holme, Peter, Beom Jun Kim, Chang No Yoon, and Seung Kee Han. 2002. “Attack Vulnerability of Complex Networks.” *Physical Review E* 65 (May): 056109.

[16] Jeong, H., B. Tombor, R. Albert, Z.N. Oltvai, and A.-L. Barabási. 2000. “The Large-Sale Organization of Metabolic Networks.” *Nature* 407 (October): 651-54.

[17] Krebs, Valdis E. “Mapping Networks of Terrorist Cells.” *Connections* 24 (3): 43-52.

[18] Lindelauf, Roy, Peter Borm, and Herbert Hamers. 2009. “The Influence of Secrecy on the Communication Structure of Covert Networks.” *Social Networks* 31 (May): 126-137.

[19] Nagaraja, Shishir, and Ross Anderson. 2006. “The Topology of Covert Conflict.” Presented at the Fifth Workshop on the Economics of Information Security, Cambridge.

[20] Solé, Ricard V., and José M. Montoya. 2001. “Complexity and Fragility in Ecological Networks.” *Proceedings of the Royal Society of London B* (October): 2039-45.

[21] Travers, Jeffery and Stanley Milgram. 1969. “An Experimental Study of the Small World Problem.” *Sociometry* 32 (December): 425-43.

[22] Watts, Duncan J., and Steven H. Strogatz. 1998. “Collective dynamics of ’small-world’ networks.” *Nature* 393 (June): 440-42.

[23] Zhao, Liang, Kwangho Park, and Ying-Cheng Lai. 2004. “Attack Vulnerability of Scale-Free Networks Due to Cascading Breakdown.” *Physical Review E* 70 (September): 035101.