



A 501(C)(3)
NOT-FOR-PROFIT
CORPORATION

The National Security Agency's Data Mining Effort

John Brantley Halstead

Public Policy Analysis (IEUC-2006-01, IEUC-PPA-01)

June 9, 2006

Abstract: Thanks to the personal sacrifice of our men and women in uniform, much of the current world terrorist activity occurs outside of the United States. Thanks to intelligence and law enforcement, we are currently able to prevent terrorism from within our borders. Unfortunately the tremendous result of terrorism occurring elsewhere has removed the reality of terrorism from some of the public's thoughts, beliefs, and attitudes. Because of intelligence, law enforcement, and the armed forces' effectiveness, we are often insulated from a simple reality. The nation is currently conducting operations, while guarding most of our personal privacy, against an enemy who knows no national boundaries and is fanatically committed to the destruction of our nation.

The use of Data Mining by the National Security Agency to process communications logs and identify patterns of activity associated with the activities of terrorist networks may well be a key element in this ongoing success. However, it has proven highly controversial in the media and raises legitimate questions from privacy advocates. Fundamental to the NSA debate is an understanding of what data mining is and isn't.

Given the timely nature of this topic and the lack of any appreciable substantive media coverage of the true nature of the technology behind this important public policy debate, we have invited a data mining expert, John Brantley Halstead of The United States Military Academy at West Point to prepare this Public Policy Analysis. In it you will find an approachable nonpartisan introduction to this key End User Computing Application along with an exploration of what questions the public should be asking as it evaluates the appropriateness of the use of this powerful tool in the Global War on Terror.

**“... I fear for the future of the free world.”
- Carmen Bin Laden,¹2004**

The National Security Agency’s Data Mining Effort

John Brantley Halstead

Within the past few weeks, the media has exposed the general public to the National Security Agency’s (NSA) data mining initiative designed to thwart terrorism. The NSA is reported to be conducting a data mining application known as social network analysis to discover and establish terrorism patterns within vast amounts of data extracted from phone records.¹⁸ The NSA data collection effort does not involve listening, eavesdropping, or recording private conversations, but does collect the telephone call history of many citizens within the United States.¹⁹

On May 17, 2006 President Bush apprised the legislative branch’s intelligence committees of its scope and breadth. Senator Roberts correctly remarked that the president’s action was appropriate in assuring the government’s collective constitutional responsibilities.²⁰ Representative Hoekstra stated the agency’s data mining is “a critical tool in keeping the country safe.”²¹ Professor Charles Fried, a Harvard Law School professor, questions whether there are any violations in civil liberties and concludes that some are using mindless labeling.²²

The NSA is not the first government agency to use data mining and it is very likely not to be the last agency either. The Department of Homeland Security is researching and developing a data mining system entitled Analysis, Dissemination, Visualization, Insight, and Semantic Enhancement (ADVISE); the Defense Advanced Research Projects Agency (DARPA) created Total Information Awareness (TIA); and the National Visualization Analytics Center in Richland, WA uses Starlight, data visualization software.²³ Some of these systems have contributed to thwarting terrorism.²⁴ The government’s measures take advantage of recent technology in fighting terror and, consequently, securing our freedoms.

What the public may not know is that corporate entities use data mining and use it frequently and extensively. As an example, have you ever wondered how Amazon is able to recommend a book, movie, or music based on your current purchase? They use data mining techniques on personal purchasing information to pattern your consumer behavior. These techniques often advance a mutual relationship between the vendor and the consumer. The vendor can sell more products by intelligently determining the consumer’s purchasing pattern. Consumers feel as if they are receiving customized attention and that the vendor cares for their needs. As consumers, many appreciate Amazon’s use of their personal information for data mining. These techniques are used in many ways and provide many consumer conveniences. Data mining is not limited to vendors.

Data mining techniques are used throughout private industry. “Data mining has helped identify new chemical compounds for prescription drugs, detect fraudulent claims and purchases, create and maintain personal customer relationships, design better engines... and develop effective credit scoring rules.”²⁵ Their uses are vast and powerful. In many cases, they are imbedded in our lives. Data mining helped develop the cars we drive, purchase the homes we live in, make shopping easier, and enhance the drugs we use.

Some do not agree with using personal information in business applications. A professional engineering organization, IEEE, has published concerns involving corporate and government data mining. In October 2003, the IEEE had posited that today’s technology permits extraction of various patterns and relationships from data warehouses, “putting consumers’ privacy in jeopardy.”²⁶ The IEEE further contends that consumers provide information for billing purposes and implicitly agree that the data may be used for billing. However, they do not implicitly agree that the data may be used for data mining, so any such data mining applications would exceed the intent of the data collection.²⁷ The implication of the IEEE opinion is that limits should exist for the ethical use of personal information.

The current NSA data mining issue is centered on the use of private information in the global war on terror (GWT). The issue is important, involving both our constitutional freedoms and our collective security. The public should be correctly informed of these intelligence initiatives, as much as national security permits. Through intelligent dialog and a clear understanding of the facts, we can achieve our collective security and protect our individual freedoms. Facilitating this dialog, we should seek to understand the underlying assumptions related to the NSA intelligence collection and analysis along with a basic understanding of data mining and queries.

What Is At Stake with the NSA Effort

Rather than focus on the politics of this issue, we should be questioning fundamental assumptions concerning the government’s use of personal information. Depending on our personal view of government, these assumptions may and should include: 1) government will/will not stay within the bounds of terrorism pattern recognition; 2) government is/isn’t eavesdropping, listening, or recording private conversations with/without court consent; 3) government will/will not use information protected by the Foreign Intelligence Surveillance Act (FISA);²⁸ and 4) government will/will not use knowledge gained to immediately prosecute private citizens and mistakenly make false positive arrests. A false positive is incorrectly identifying or classifying an innocent as a terrorist. Coupled with these assumptions is a balance with the GWT on the other scale.

We should be concerned that the government uses the information strictly for determining terror pattern recognition. Given the vast amounts of data,²⁹ the government is naturally forced into using data mining techniques. The essential difference between data mining techniques and a database query is a presumption of a fact or trend. In data mining, methods are employed to

discover a trend or pattern across the full spectrum of a database. A query, conversely, requires some known information or hypothesis regarding a single entity within a database. Queries, as a consequence, delve into individual records and, by chance, discover information. A query has the potential to risk our privacy. Queries are extremely time consuming and, given the size of the data and applying common sense, may be impossible for the government to conduct. The use of data mining techniques does provide most citizens with protection from the government discovering personal information not related to terrorism. Regardless of the data mining technique, terrorism defines the function and method that searches for patterns and trends within the data. We can protect our personal freedoms and continue the search for terrorists through legislation that provides assurances that data mining techniques are used to discover trends rather than conducting individual queries on records.

Without legal consent, any government attempt to eavesdrop, listen, and record a conversation of a private citizen is wrong. Quite honestly, it's also a waste of time, resources, and manpower. Government potentially risks our security by investing time and resources into such endeavors. With the limited time available to prevent possible terrorist attacks, the government is better served searching for patterns within the full spectrum of the available data rather than focus on a few (relative to the data's size) entities. This concern is a genesis of conspiracy theory; we shouldn't overly concern ourselves with it. Again, legislation mandating the strict use of data mining reduces the temptation of the activity.

To date, the government has claimed to stay within the legal parameters of FISA.³⁰ The FISA statute protects citizens by restricting government's intelligence collection. Government is not able to use personal identifiers in data collection without court consent. The government is restricted from using personal information such as names, addresses, and social security numbers within a search.³¹ Data mining techniques typically don't require personal information to be effective. If a terrorist pattern exists and is discovered with data mining, personal information may be contained within non-personal information used to discover the trend. The non-personal information determines the classification, trend, or pattern rule rather than the personal information. Enforcement of FISA protects our individual freedoms while facilitating the GWT.

The public should be concerned about misclassification (i.e. a false positive classification). The damage to an individual falsely accused of terror activities could be extensive. There are many examples of people fired from professions as a result of incorrect information gained in a query.³² Such examples highlight two concerns: one of accuracy and another of procedures. Procedurally, this concern should drive an operational requirement to confirm or deny the information gained from data mining's discovered patterns and trends. The procedures necessary to perform these operations currently exist within our laws under FISA which authorizes a special court to evaluate government requests for authorization to conduct further searches and direct surveillance. Our current laws provide the court with the framework to guide its decision.

From a technical perspective, accuracy issues diminish with the size of the data. A direct relationship exists between data size and accuracy. The more records contained within the data,

the more accurate a data mining method. Ironically, the size of the NSA data collection significantly reduces a misclassification error.

Balancing these concerns, we might ask ourselves who we fear or mistrust more, the terrorist or the United States government. Thanks to the personal sacrifice of our men and women in uniform, much of the current world terrorist activity occurs outside of the United States. Thanks to intelligence and law enforcement, we are currently able to prevent terrorism from within our borders. Unfortunately the tremendous result of terrorism occurring elsewhere has removed the reality of terrorism from some of the public's thoughts, beliefs, and attitudes. Because of intelligence, law enforcement, and the armed forces' effectiveness, we are often insulated from a simple reality. The nation is currently conducting operations, while guarding most of our personal privacy, against an enemy who knows no national boundaries and is fanatically committed to the destruction of our nation.³³ The threat is real and our national sacrifice hasn't been extensive beyond those combating terrorism and the victims of terrorism. Until the GWT is over, we should balance sacrificing some of our personal freedoms with our collective security. The majority of the American public still values this tradeoff. Many public opinion polls concur with sacrificing some freedoms for security.³⁴ As Carmen Bin Laden, who has unique insight concerning the relationships between Saudi Arabia, the Bin Laden family, and terrorism, warns "... if we, in the Western World, are not vigilant enough, there will be no end to their terrorism. They will use our tolerance to infiltrate our society with their intolerance."³⁵

Data Mining versus Database Queries

Fundamental to the NSA debate is an understanding of what data mining is and isn't. First, data mining is not drilling or querying a database for information. Both drilling and queries assume known information. They use a form of standard query language that searches a database for records containing the predefined information. Data mining distinguishes itself by discovering unknown trends and patterns in data. To discover unknown trends, data mining employs methods such as statistical learning, machine learning, mathematics, and artificial intelligence. A popular graduate level textbook defines data mining as "the process associated with discovering patterns and relationships in extremely large data sets."³⁶ A University of California, Los Angeles professor defines data mining as "the process of analyzing data from different perspectives and summarizing in into useful information."³⁷ Wikipedia has defined data mining as "the process of automatically searching large volumes of data for patterns."³⁸ In summation, data mining distinguishes itself from queries by its exploratory nature of discovering patterns rather than developing a hypothesis and searching for entities that match the hypotheses. Our personal freedoms are protected and our security is enhanced within the subtle difference. Queries delve into our personal privacy, while data mining observes the entire data and determines a trend or pattern. Data mining certainly is not analogous to an Orwellian practice; queries may not survive similar scrutiny.

Data mining is an extensive and expanding body of knowledge. Significant contributions and advances are rapidly developed as mathematicians, engineers, economists, and scientists

continue creating in their respective fields. The information contained in this article is not sufficient for a full understanding of data mining. However, it is intended to begin thoughtful discussion on data mining and its relation to personal freedom. For an extensive discussion on data mining and pattern recognition, you may consult the two popular cited works within the body of knowledge. Hastie, et al authored a thorough textbook containing a nearly full catalog of statistical learning algorithms.³⁹ Keinosuke Fukunaga authored the popular cited textbook concerning statistical pattern recognition.⁴⁰ The IEEE remains the primary and most reliable source for recent and significant advances in data mining.⁴¹

When defining a pattern and trend, data mining methods perform classification, regression, clustering, associations, or determine sequential patterns. Classification sorts entities into groups by using supervised learning techniques. The unsupervised learning twin to classification is clustering, which typically uses statistical distances to cluster entities within similar groups. The distinguishing factor between supervised and unsupervised learning is whether the data contains a response. An example of a response is indicating that the entity is a known terrorist or not a terrorist. Supervised learning requires a response. Unsupervised learning doesn't require a response. Regression involves establishing a relationship between data and a response and uses supervised learning methods to obtain the relationship. Associations can be inferred through supervised and unsupervised learning methods. Sequential patterns involve determining relationships in time series data, where the next pattern is dependent and related to the previous pattern.

Supervised learning requires a response to discover patterns and trends. To overcome a response requirement, unsupervised learning uses a distance measure, often statistical distance, to establish patterns by grouping entities together. The only response required in a terrorism data mining application is a field within an entity indicating whether the entity is a known terrorist or not. Typically, supervised learning uses a known data set, a data set containing a known response for every entity within the data. Patterns are determined using a method capable of relating the data to the response by training an algorithm on a data set randomly selected from the known data. Data miners often call this a training data set. The validity and accuracy of the data mining method is tested on separate data randomly selected from the known data. The data mining method predicts the response. Accuracy is determined by comparing the predicted response with the true response. These data are often labeled validation data or test data. If the errors are very low, the data mining method is able to generalize. Generalizing speaks of the method's ability to accurately classify or predict new and unseen entities.

In terms of discovering terrorist patterns, the results of supervised and unsupervised learning methods are similar. With supervised learning, if an entity is classified as a suspected terrorist, the entity shared the same pattern and features as the terrorist. In the unsupervised case, entities would share the same cluster as a known terrorist. By sharing the same cluster, the entity also shares the same pattern and features as the terrorist. Once clustered or classified as a potential terrorist, under FISA, intelligence agencies are able seek court consent to continue surveillance with the intent to confirm or deny the cluster or classification. The court would have to weigh

the classification accuracy in determining search consent. Most citizens have little to fear. These methods will not discover personal information. They rather relate a person to terrorist or, more frequently, don't relate a person to terrorists. They are designed to seek and discover terrorist patterns.

Data mining methods and algorithms are extensive. They may involve support vector machines and regression, random forest, genetic algorithms, artificial neural networks, decision trees, logistic or linear regression, nearest neighbor methods, rule induction, and data visualization and reduction methods. Improvements within these methods are discovered frequently as scientist, engineers, mathematicians, and economists optimize the methods to deliver the best generalization possible.

Social network analysis is a promising method used to cluster groups of people together. Recently, some have included social network analysis as a data mining application instead of a data mining method.²⁶ For example, other applications may be drug discovery, marketing, finance, or earth sciences.

Social Network Analysis

Social networks are interconnected structures made of nodes and ties. The nodes represent people or organizations and the ties are linkages established by communication between nodes. For an example of a terrorist social network, one can view the network discovered by Valdis Krebs in the Christian Science Monitor.²⁷

Many network analysts believe a person's social network has tremendous impact on how they live their lives.²⁸ They further believe the success and failure of the organization can be measured from the structures' pattern. The implication of social network analysis and terrorism is that if a person belongs to the network, their lives are influenced by the network and they are in communications with members of the network.

From the beginning, social network analysis has been committed to two governing principles.²⁹ The analysis is guided by formal theory explained by mathematical terms. Further, it is grounded in the systematic analysis of empirical data. With these governing principles, social network analysis is a formal and scientific method for determining social networks.

Analysis of the network involves discovering the nodes' location.³⁰ Locations are measured by numerous metrics that provide information regarding the node's importance and role within the network. These metrics include: "Degree Centrality," "Between-ness Centrality," "Close-ness Centrality," "Network Centralization," "Network Reach," "Boundary Spanners," and "Peripheral Players."³¹ The method applied to terrorism and anti-terrorism implies that at least one member of a potential terrorist network is a known terrorist. The metrics help clarify the members of the network and their roles.

Many of the current experts in social network analysis, including Krebs, question the government's extensive data search.³² They claim social network analysis doesn't require the vast amounts of data collected by NSA. Extending the claim, they view the data collection as wasted effort. However, some counter the NSA data collection as necessary and propose the NSA methods and applications may be more than social network analysis.³³

Conclusion

We may not know if the NSA data mining strictly involves social network analysis as an application, or the employed NSA data mining methods. We can conclude the size of the NSA data collection helps reduce classification and clustering errors. We can also conclude that the data size, coupled with data mining methods, provides technical safeguards to our private information. With vast amounts of data, the NSA is naturally restricted to employing terrorism data mining, focusing critical resources of time, resources, and manpower to the effort. Doing otherwise jeopardizes our collective security and is a waste of resources. We can also be assured that if a person is classified or clustered as a potential terrorist, they share many of the same features and/or are communicating with know terrorists. For these reasons, in the interim, our personal information is safe from government exploitation.

During the Global War on Terror, government could provide better assurance to the public that it stays within the bounds of FISA and the Patriot Act. Working within the current framework of our laws, we can maintain security and provide for personal freedom, more so than in any other previous military action. Once the GWT is over, the people may want to explore the possibility of legislation that reduces the current NSA data mining scope but provides a capability to monitor and protect us from potential threats. Data mining has the potential to increase our security, while not risking our personal freedom provided we remain within a data-mining framework. These methods should be retained and leveraged to safeguard against future threats. The public may consider legislation that provides government with the means to provide us security while protecting our personal information. Part of this legislation would permit data mining for potential threats and restrict the use of query without court consent.

John Brantley Halstead, Ph.D., Lieutenant Colonel, United States Army, Assistant Professor, United States Military Academy, Department of Systems Engineering.

The views expressed in this article are singularly the author's. The views do not represent the official policy of the United States Government, the Department of Defense, the United States Army, the United States Military Academy, and the Department of Systems Engineering. Nor do they represent official technical findings or public policy positions of The Institute for End User Computing, Inc. The author acknowledges the review comments and editorial feedback of The IEUC's Peter J. Wasilko, who commissioned this Public Policy Analysis. This material may not be published, broadcast, rewritten, redistributed, translated or otherwise be used as the basis for a derivative work without the consent of the author. The author has granted a non-exclusive use right to The Institute for End User Computing, Inc. under which this material has been posted on The Institute's website and made available for your download without cost as a public service. While you may link to this article, index it for search purposes, and exercise any other fair use rights provided by law, all other rights are reserved by the author. The author may be contacted at: john.halstead@ieuc.org.

Endnotes

¹ Carmen Bin Laden, *Inside the Kingdom, My Life in Saudi Arabia* (New York, NY: Warner Books, 2004), 199.

² Leslie Cauley, “NSA has a massive database of Americans’ phone calls,” 11 May 2006; available from http://www.usatoday.com/news/washington/2006-05-10-nsa_x.htm; Internet; accessed 16 May 2006. The article outlines the NSA data mining effort to pattern terrorist activities. It reports that the NSA is not listening or recording private conversations; rather, it uses data to analyze patterns in an effort to detect terrorist activity. The data mining method used is social network analysis, which links people via their phone calls. The methodology uses known terrorists and establishes links from the terrorists to others, strength of link can be determined by frequency and length. Furthering the idea, the article implies that the government has gained a secret window into the communication habits of millions of Americans. Furthering the idea, the article implies the government will reach beyond the scope of the anti-terrorism goal. A key question is whether the data collection is bounded by the Foreign Intelligence Surveillance Act (FISA) or even the Patriot Act. FISA doesn’t prohibit data mining provided personal identifiers are not included.

³ Ibid.

⁴ Pam Benson and Cheryl Bronson, “White House to brief House, Senate panels on NSA wiretaps,” 17 May 2006; available from <http://www.cnn.com/2006/POLITICS/05/17/nsa/index.html>; Internet; accessed 17 May 2006. The article reports on the executive branch providing more information to the house and senate so that all committee members know the full width and breadth of the intelligence program. Senator Pat Roberts is quoted “... there is no way we could fulfill our collective constitutional responsibilities without that knowledge.” Representative Peter Hoekstra, House Intelligence Committee Chairman, states the program is “a critical tool in keeping the country safe.” The article also contained a key element to understanding data mining. It differentiated the difference between listening to private calls and pattern recognition, “the agency uses the data, which includes numbers, times and locations, to look for patterns that might suggest terrorist activity.”

⁵ Ibid.

⁶ Mark Clayton, “Mining Data to Nab Terrorists; Fair?,” 15 May 2006, available from <http://www.csmonitor.com/2006/0515/p01s04-usju.html>; Internet; accessed 16 May 2006. The article presents cases for and against social networking without providing an opinion. Also it contains a nice example of social networking performed by Valdis Krebs that associated bombers of the USS Cole with 9/11 hijackers. Also contains commentary from Charles Fried, Harvard Law Professor. He states that no civil liberties are violated. The article closes with polling results that demonstrate the public’s equal divide on NSA data mining.

⁷ Mark Clayton, “US Plans Massive Data Sweep,” 9 February 2006, available from <http://www.csmonitor.com/2006/0209/p01s02-uspo.html>; Internet; accessed 16 May 2006. The article provides limited information on other data mining applications. The government applications are Analysis, Dissemination, Visualization, Insight, and Semantic Enhancement (ADVISE) developed by the Department of Homeland Security; Starlight maintained by the National Visualization Analytics Center in Richland, WA; and the Total Information Awareness (TIA) developed by the Department of Defense.

⁸ Ibid.

⁹ Richard A. Johnson and Dean W. Wichern, *Applied Multivariate Statistical Analysis*, 5th Edition (Upper Saddle Hill, NJ: Prentice Hall, 2002), 731.

¹⁰ George W. Zobrist, "Data Mining and Privacy Issues," IEEE – USA Today's Engineer, October 2003, available from <http://www.todaysengineer.org/2003/Oct/data-mining.asp>; Internet; accessed 16 May 2006. The article defines data mining and distinguishes data mining from query. The author provides a problem statement concerning privacy and states corporate entities and government may exceed intent of data collection. The article provides more information on TIA (see endnote 7). The article closes with legislation initiatives: Citizens Protection in Federal Databases Act and a Memorandum of Understanding with Defense Advanced Research Projects Agency (DARPA) and TIA.

¹¹ Ibid.

¹² Federation of American Scientist, "Foreign Intelligence Surveillance Act," available from <http://www.fas.org/irp/agency/doj/fisa/>; Internet; accessed 18 May 2006. Website provides reliable information of FISA.

¹³ Cauley.

¹⁴ Ibid.

¹⁵ Federation of American Scientists

¹⁶ Brian Bergstein, "In This Data Mining Society, Privacy Advocates Shudder," 2 January 2004; available from http://www.seattlepi.nwsource.com/business/154986_privacychallenge02.html; Internet; accessed 15 May 2006. Article provides concerns on privacy issues. The author incorrectly defines queries as data mining.

¹⁷ Vice Director of Strategic Plans and Policy, Joint Staff, J-5, "Fighting the Long War – Military Strategy for the War on Terrorism," unclassified briefing presentation with notes, Joint Staff, Department of Defense. The presentation is an excellent source containing the goals of Islamic Jihad and the nation's strategy towards peace. The presentation summarizes radical Islamic goals obtained from many reliable sources. The United States strategy is also explained in an unclassified manner.

¹⁸ Clayton, "Mining Data to Nab Terrorists: Fair?"

¹⁹ Bin Laden, 2003.

²⁰ Johnson, 731.

²¹ Jason Frand "Data Mining: What is Data Mining?," unknown date, available from <http://www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/datamining.html>; Internet; accessed 16 May 2006. The article provides basic data mining definitions and capabilities.

²² Wikipedia, "Data Mining," unknown date, available from http://www.en.wikipedia.org/wiki/Data_mining; Internet; accessed 15 May 2006. The site provides a public domain for defining data mining. The site is considered only moderately reliable since there are few standards governing the authors' entries.

²³ Trevor Hastie, Robert Tibshirani, and Jerome Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction (New York, NY: Springer, 2001). This is an excellent text that provides the vast catalog of powerful algorithms used in data mining.

²⁴ Keinosuka Fukunaga, Introduction to Statistical Pattern Recognition, 2nd Edition (New York, NY: Academic Press, 1990). The text is the most definitive source on pattern recognition. On a side note, Fukunaga is co-discoverer of branch and bound, an exhaustive technique to determine the optimal pairs of data.

²⁵ IEEE, “The Worlds Leading Professional Association for the Advancement of Technology,” date unknown; available from <http://www.ieee.org>; Internet; accessed on 18 May 2006. The IEEE homepage serves as a gate to discovering recent and relevant technological innovations.

²⁶ Society for Industrial and Applied Mathematics, “2006 SIAM Conference on Data Mining,” unknown date; available from <http://www.siam.org/meetings/sdm06/>; Internet; accessed 21 May 2006. The site provides conference proceedings and lists Social Network Analysis as an application of data mining rather than as a data mining method.

²⁷ Clayton, “Mining Data to Nab Terrorists: Fair.”

²⁸ Lin Freeman, “The Study of Social Networks,” unknown date; available from http://www.insna.org/INSNA/na_inf.html; Internet; accessed 21 May 2006.

²⁹ Ibid.

³⁰ Valdis Krebs, “An Introduction to Social Network Analysis,” 2006; available from <http://www.orgnet.com/sna.html>; Internet; accessed 18 May 2006. The site provides a summary of Social Network Analysis. Author was one of the first to construct a 9/11 social network of the terrorist involved in the 9/11 attacks and relate them to the bombers of the USS Cole from public information.

³¹ Ibid.

³² Clayton, “Mining Data to Nab Terrorist: Fair.”

³³ Noah Shachtman, “NSA Sweep ‘Waste of Time,’ Analyst Says,” 11 May 2006; available from <http://defensetech.org/archives/002399.html>; Internet; accessed 21 May 2006. The Defense Technology Organization’s web log contains commentary discussing the pros and cons of the NSA data collection effort. The site also implies that the NSA may be using other methods.