Information Sharing and Privacy Protection of Terrorist or Criminal Social Networks

Christopher C. Yang
College of Information Science and Technology
Drexel University

***Abstract*— Terrorist or criminal social network analysis is helpful for intelligence and law enforcement force in investigation. However, individual agency usually has part of the complete terrorist or criminal social network and therefore some crucial knowledge is not able to be extracted. Sharing information between different agencies will make such social network analysis more effective; unfortunately, it may violate the privacy of some sensitive information. There is always a tradeoff between the degree of privacy and the degree of utility in information sharing. Several approaches have been proposed to resolve such dilemma in sharing data from different relational tables. There is not any work on sharing social networks from different sources and yet try to minimize the reduction on the degree of privacy. In this paper, we propose a subgraph generalization approach for information sharing and privacy protection of terrorist or criminal social networks. Our experiment shows that such approach is promising.**

***Index Terms*— social network, information sharing, privacy protection, intelligence and security informatics**

I. Introduction

Terrorist or criminal social networks are useful for investigation, identifying suspects and gateways, and extracting communication patterns of terrorist or criminal organizations. An illustration of a terrorist social network as presented in [16] is shown in Figure 1. However, some important patterns or knowledge cannot be extracted if one is given only a partial terrorist or criminal social network. In reality, no agency has a complete social network but each agency only has information about some of the terrorists and their relationships. As a result, it is difficult to extract the timely and crucial knowledge for developing strategies in combating terrorism or crimes effectively unless information can be shared. On the other hand, sensitive information is usually confidential due to the privacy issue. That causes the limitation on the degree of information sharing. Ideally, we like to share the crucial information in order to be able to mine as much knowledge as if we have the complete information without violating the privacy. If we increase the degree of information sharing, the utility of the shared information increases and more knowledge can be extracted but the degree of privacy decreases. On the contrary, if we decrease the degree of information sharing, the degree of privacy increases but the utility of the shared information decreases. The challenge is how to develop a mechanism so that the information with high utility can be identified and shared with the minimum impact on privacy.

Information comes from data sources at different security level and different agencies, for examples, Federal Bureau of Investigation (FBI), Central Intelligence Agency (CIA), air force, army, navy, local, state and federal law enforcement agencies, and other private agencies. If information sharing is not available among these agencies, a complete picture of the terrorist community or crime organization cannot be obtained. Investigation will be delayed and the consequence can be severe. Indeed homeland security policy contend that, on the local level, information gathering, coordination with state and federal agencies, infrastructure protection, and enhanced development of police-community relationships will facilitate prevention, and aid response to potential terrorist attacks [5],[6],[11].

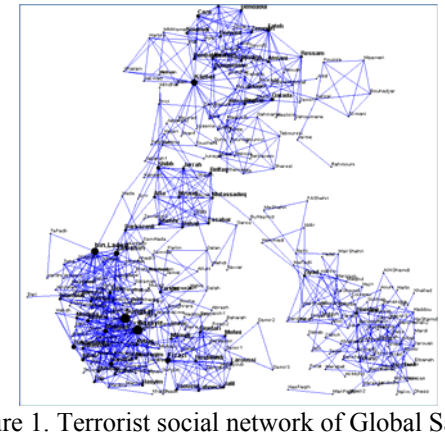

Figure 1. Terrorist social network of Global Salafi Jihad [16].

Thuraisingham [13] defined assured information sharing as information sharing between organizations but at the same time enforcing security and integrity policies so that the data is integrated and mined to extract nuggets. Thuraisingham described a coalition, in which members may join and leave the coalitions in accordance with the policies and procedures [13]. Members or partners conduct coalition data sharing in a dynamic environment. Baird et al. [2] first discussed several

ISI 2008, June 17-20, 2008, Taipei, Taiwan
1-4244-2415-3/08/$20.00 ©2008 IEEE

aspects of coalition data sharing in the Markle report. Thuraisingham [13] has further discussed these aspects including confidentiality, privacy, trust, integrity, dissemination and others.

In this work, we focus on the social network information sharing. We develop a sub-graph generalization method to share the only necessary information such that privacy can be protected. As a result, mining can be conducted to discover further knowledge to support social network analysis.

A social network is a graph, $G = (V, E)$, where V is a set of nodes representing persons and E is a set of edges ($V \times V$) representing the relationships between the corresponding persons.

We define our research problem as follow: Given two or more social networks from different sources, there is a desire to share the information between these social networks in order to conduct more accurate social network analysis for intelligence investigation. Without information sharing, we only have partial information of the complete social network that can be merged from these social networks. However, due to the privacy issue, releasing the information of the nodes and edges of one social network from one source to another is not permissible.

Example 1:
Given a social network $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, we can merge G_1 and G_2 to $\boldsymbol{G}$ if information sharing is possible.

$V_1 = (A,B,C,D,E,F,G,H)$
$V_2 = (A,B,C,D,E,F,G,H)$
$E_1 = ((A,B), (B,C), (C,D), (D,E), (D,F), (E,F), (F,G), (G,H))$
$E_2 = ((A,B), (B,C), (B,D), (B,F), (C,F), (D,E), (F,G), (F,H))$
Figure 2 shows G_1, G_2, and $\boldsymbol{G}$. Table 1 shows the adjacency matrix of $\boldsymbol{G}$.

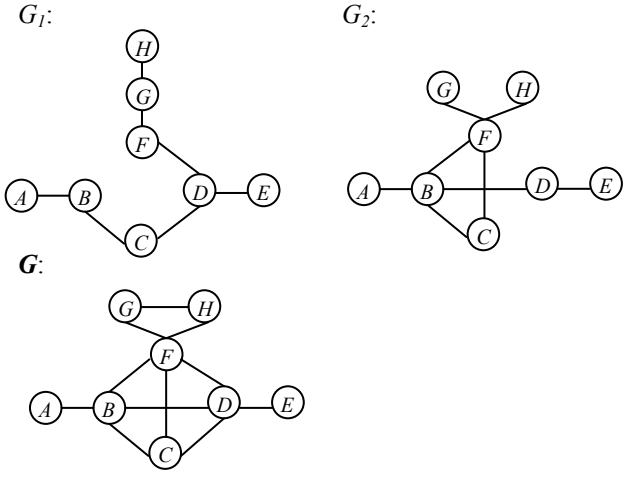

Figure 2. G_1, G_2, and the integrated graph $\boldsymbol{G}$.

Some common social network analysis techniques are:

$$degree\ centrality(u) = {degree\ of\ u} \Big/ {n-1}$$

$$closeness\ \ centrality(u) = \frac{n-1}{\sum_{v=1}^{n} d(u,v)}$$

where $d(u,v)$ is the shortest distance between u and v

$$betweennes\ s\ \ centrality\ (w) = \frac{\sum_{u<v} {p_{uv}(w)} \Big/ {p_{uv}}}{(n-1)(n-2)}$$

where $p_{uv}(w)$ is the number of shortest paths between u and v that pass through w

Table 1. Adjacency matrix of $\boldsymbol{G}$:

	A	B	C	D	E	F	G	H
A	0	1	0	0	0	0	0	0
B	1	0	1	1	0	1	0	0
C	0	1	0	1	0	1	0	0
D	0	1	1	0	1	1	0	0
E	0	0	0	1	0	0	0	0
F	0	1	1	1	0	0	0	0
G	0	0	0	0	0	1	0	1
H	0	0	0	0	0	1	1	0

The objective is to conduct the social network analysis tasks from the information given in two social networks without violating the privacy but still obtain the result as accurate as if it is conducted on the integrated social network. In this work, we propose the sub-graph generalization approach for information sharing and privacy protection and we focus on computing closeness centrality.

II. Related Work

The concerns on privacy of publishing data have been increasing in the recent years. Nowadays, many organizations are publishing data containing sensitive personal information for analysis. This data is usually represented as table, which include medical, census, customer transactions, voter registrations. This unaggregated information about individuals is important for trend analysis, medical research, and allocation of resources. However, individual privacy is always an important issue in the public. A solution is publishing distorted version of tables so that identifications cannot be easily detected. The degree of distortion affects the degree of privacy and utility of data. Utility refers to the usefulness of the distorted table for analyzing the characteristics of data. If we increase the degree of distortion, it increases the degree of privacy but decreases the degree of utility. On the contrary, if we decrease the degree of distortion, it decreases the degree of privacy but increases that of utility.

A simple approach is removing attributes that uniquely identifying a person, such as names and identification numbers. However, trivial linking attack proves that this simple approach does not work. It is very likely to identify a person with his sensitive data by using the apparently innocuous sets of attributes such as ages, gender and zip code from other tables. These sets of attributes support linking attack are known as *quasi-identifiers*. An earlier study estimates that 87% of US

population can be uniquely identified by linking attack [9]. According to Sweeney [10], it can identify the medical records of Massachusetts governor by linking several publishing tables. As illustrated, if one knows that Charles is a registered voter, he/she looks up Charles' registration record and uses the quasi-identifiers to cross-check with the hospital that he has visited. It is simple to identify that Charles has HIV which is his sensitive information.

Table 2. Medical records and voter registration records in information sharing using attribute removing and *k*-anonymity.

Medical Records – Removing Names

Name	Age	Sex	Location	Disease
Peter	8	M	00330	Viral Infection
Paul	14	M	01540	Viral Infection
Andrew	18	M	18533	Viral Infection
Stephen	20	M	14666	Viral Infection
Charles	29	M	35667	HIV
Gordon	30	M	43986	Cancer
Linda	35	F	31147	Cancer
Mary	39	F	45246	Cancer
Stella	45	F	85103	Heart Disease
Angel	51	F	96214	HIV

quasi – identifiers

Voter Registration Records

Name	Age	Sex	Location
Charles	29	M	35667
Paul	14	M	01540
David	25	M	00338

…

Medical Records – *k*-anonmity

Age	Sex	Location	Disease	
[5,20]	M	[00300,02000]	Viral Infection	quasi group *A*
	M		Viral Infection	
	M		Viral Infection	
	M		Viral Infection	
[20,40]	M	[20001,50000]	HIV	quasi group *B*
	M		Cancer	
	F		Cancer	
	F		Cancer	
[41,60]	F	[80000,99999]	Heart Disease	quasi group *C*
	F		HIV	

Voter Registration Records

Name	Age	Sex	Location
Peter	29	M	35667
Paul	14	M	00332
David	25	M	00338

…

k-anonymity [8,10] is the first approach to protect privacy in publishing data rather than only removing identification attributes by generalizing the dataset. If every record in a table is indistinguishable from at least *k*-1 other records with respect to every set of quasi-identifier attributes, this table satisfies the property of *k*-anonymity. However, *k*-anonymity fails when there is a lack of diversity in the sensitive attributes or the one who attacks has other background knowledge. As illustrated, there is a lack of diversity on the attribute values of disease in the quasi group with age = [5,20] and location = [00300, 02000]. When one identify Paul's attributes values of the quasi identifiers in the voter registration records matches with the values of the quasi group in the medical records, it is easy to identify Paul's disease as viral infection.

l-diversity[2] ensure that there are at least *l* well-represented values of the sensitive attribute for every set of quasi-identifier attributes. However, one can still estimate the probability of a particular sensitive value. As illustrated, quasi group *A* is 1-diverse and therefore the attribute value of Disease of any records within this group can be identified as viral infection. However, quasi group *B* and *C* are 2-diverse. As a result, we cannot identify the disease attribute value for any records within these two groups. Unfortunately, we can still estimate that the probability of the disease attribute values as cancer for any records in quasi group *B* is 75%. If one has the background knowledge about Mary that she is HIV-negative in addition to matching the quasi identifiers, Mary's disease must be cancer.

Other enhanced techniques of *k*-anonymity and *l*-diversity have also been proposed recently. Personalized anonymity [15] allows a person to specify the degree of privacy protection for his sensitive values. Instead of publishing the exact sensitive value, it publishes a value in a higher level of the taxonomy of attribute that is acceptable to the user. (α,k)-Anonymity [14] allows a person to specify a threshold α on the relative frequency of the sensitive data in every equivalence class.

The approaches of privacy protection on records presented in table formats are mainly focused on *domain generalization*. Domain generalization partitions the values of a domain *D* into a number of partitions P_1, P_2, … such that the union of these partitions equal to the range of *D*. In general, the partitions do not have any overlapping. That means $\cup P_i = D$ and $P_i \cap P_j = \emptyset$. By partitioning the domains of several quasi-identifiers, it increases the difficulties for attackers to uniquely identify a person and his sensitive information.

III. Sub-graph Generalization

A social network is not represented as a table but can be represented as an adjacency matrix. Domain generalization is not applicable because we don't have a domain of any particular attribute. We propose an approach in privacy protection of social networks by generalizing sub-graphs of a social network. By generalizing a sub-graph, we want to protect the identification of nodes that are not known in public, the structure of the sub-graph, and the types of relationships among the nodes within the sub-graph. However, generalizing a sub-graph will lose the important information for social network analysis when a social network is integrated with other social networks.

In order to maintain a certain degree of utility of a generalized sub-graph, we provide other general information about sub-graphs for the purpose of analysis. Such general

information should not reduce the degree of privacy; on the other hand it should provide additional utility for social network analysis. This general information are the general properties of graphs such as number of nodes in a sub-graph, length of shortest paths, possible types of relationships, etc.

Given a social network, $G = \{V, E\}$, the sub-graph generalization approach partitions G into several connected sub-graphs $G_1, G_2, \ldots$ Each sub-graph has a set of nodes and a set of edges. $G_i = \{V_i, E_i\}$. There are edges that connecting G_i together as the original graph G. E_{ij} denotes the edge connecting G_i and G_j. Since these partitioned sub-graphs do not have any overlapping (i.e. $V_i \cap V_j$), the union of the nodes from all the sub-graphs is the original set of nodes (i.e. $\cup V_i = V$). The union of the edges from all the sub-graphs and other connecting edges of the sub-graphs is the original set of edges (i.e. $\cup E_i + \cup E_{ij} = E$).

The partitioned sub-graph (or known as generalized node) is represented as a node in the generalized graph. The total number of generalized nodes in the generalized graph equals to the number of partitioned sub-graphs. The total number of edges in the generalized graph equals to the number of connecting edges of sub-graphs. $G = \{\{\ldots, G_i, \ldots\}, \{\ldots, E_{ij}, \ldots\}\}$.

The adjacency matrix of the generalized graph is transformed by merging the cells that representing the relationships of the nodes within a generalized sub-graph or between generalized sub-graphs. The adjacency matrix of a generalized graph with K generalized sub-graphs has $K \times K$ merged cells. The value in a cell between two generalized sub-graphs represents the type of relationship between the representing nodes of the sub-graphs (i.e. the node of the known identification in public). The values in the cell between the same sub-graphs represent the generalized information of the generalized sub-graph. E_{ij} exists between G_i and G_j if there is at least one edge between the nodes of G_i and the nodes of G_j; otherwise, E_{ij} does not exist.

Policy of Sharing

Identity of a person and his relationship with another person can be disclosed if his name and relationship is available in a public source. The relationship between two persons can be disclosed only if the identities of both persons are available in a public source. The public source can be made available by agencies in the higher hierarchy or the public information providers such as news providers.

Text mining techniques have been developed to extract entities such as persons from text. The latest techniques are going further to extract the relationships between these entities. As a result, general terrorist social networks can be automatically extracted from news articles on the Web. For example, the Joint-Research Center of the European Commission [4] is developing the Europe Media Monitor system to derive quantitative data from unstructured text. However, information available on the Web is limited. Much confidential information is still relying on the security agencies to provide.

Shared Information in a Generalized Node

Information available for sharing in a generalized node in a generalized graph determines the utility of the generalization. In order to conform to the policy of sharing, identities of non-public persons and their relationships cannot be released. However, general information such as the number of nodes in a generalized node and the lengths of shortest paths do not violate the privacy information of individuals.

We propose publishing a combination of the following information in a generalized node depending on the intended degree of privacy.

(1) The number of nodes: N
(2) The connecting node on a link between two generalized nodes if this node is known in a public source:
(3) The maximum length of the shortest paths between any two nodes in a generalized node, W: $L\_SP(W)$
(4) The average length of the shortest paths between any two nodes in a generalized node, W: $AVG\_SP(W)$
(5) The length of the shortest path between two known nodes, a and b, in a generalized node, W: $SP(a,b,W)$
(6) The length of the shortest path between a known node, a, and any node in a generalized node, W: $SP(a,W)$
(7) The set of relationship types for all nodes in a generalized node, W: $R(W)$
(8) The range of degrees for all nodes in a generalized node, W: $D(W)$
(9) The number of shortest path going through a known node a between any two nodes in a generalized node, W: $P(a, W)$ if w is a node known in a public source
(10) The total number of shortest path for all pairs of nodes in a generalized node, W: $P(W)$

When a policy is developed to determine the information to be shared in a generalized node, there are a number of criteria to be considered [13]. We give a few examples that are considered to be important.

Need to Know: Some information is essential for conducting a social network analysis but some other information is not as important depending on the types of analysis. As a result, the information needs to be shared depends on what the information requester needs to do. For example, if the information requester is computing closeness centrality of a criminal social network, information such as $P(w)$ and R are not necessary.

Need to Share: Some information is crucial for discovering knowledge. Without such, there is very limited knowledge one can discover.

Privacy: The private information of individuals that is under a high risk of violating laws or endangering the person involved should not be disclosed.

Trust: Information providers may have different degree of trust on different information requesters. The amount information to be shared can be adjusted by the degree of trust. The complete information can be disclosed to an information requester who has a hundred percent trust from the information provider regardless of the need. On the other hand, the

information provider may decide not disclose any information if the information requester has zero percent trust from itself.

Computation with shared information

If we don't know the connecting nodes between two generalized nodes, we can compute the maximum, minimum and average distances between any pairs of nodes in any two generalized nodes as follows:

$$d^{\max}(u,v)=\begin{cases} L\_SP(U)+L\_SP(V)+ \\ \quad \sum_{\forall W} L\_SP(W)+|W|+1 & if\ U\neq V \\ L\_SP(U) & else \end{cases}$$

$$d^{\min}(u,v)=\begin{cases} |W|+1 & if\ U\neq V \\ 1 & else \end{cases}$$

$$d^{\text{avg}}(u,v)=\begin{cases} AVG\_SP(U)+AVG\_SP(V)+ \\ \quad \sum_{\forall W} AVG\_SP(W)+|W|+1 & if\ U\neq V \\ AVG\_SP(U) & else \end{cases}$$

where U and V are subgraphs (generalized nodes) in a generalized graph,

$u \in U, v \in V$

W is a subgraph (generalized node) in the shortest path between u and v in a generalized graph

If we know the connecting nodes of a subgraph W which is in the shortest path between u and v in a generalized graph and we know the length of the shortest path between these two nodes SP, we can replace the maximum length of the shortest path between any pairs of nodes in the subgraph $L\_SP(W)$ and the average length by the shortest path between the two know connecting nodes.

$L\_SP(W)$ is substituted by $SP(a,b)$

where a is a known connecting node between W and the subgraph before W in the shortest path between u and v in a generalized graph and

b is a known connecting node between W and the subgraph after W in the shortest path between u and v in a generalized graph.

If the known connecting nodes of a subgraph which is in the shortest path between u and v in a generalized graph and these connecting nodes in a subgraph is indeed the same node, then $L\_SP(W)$ can be replaced by 0.

In the special case that the connecting nodes are the same in all subgraphs along the shortest path between u and v, $d^{max}(u,v)$ equals to $d^{min}(u,v)$.

Similarly, we replace $AVG\_SP(W)$ by $SP(a,b)$ or 0 in computing $d^{avg}(u,v)$ when the shortest path between the known connecting nodes in W is available or the connecting nodes in W are the same.

Example 2:

G_2 in Example 1 can be generalized to a graph with four generalized nodes based on what we want to further analyze from what we know about the structure of G_1. Table 3 shows the adjacency matrix of G_2. Table 4 presents the adjacency matrix of the generalized graph of G_2 with four generalized nodes. The first node generalizes the subgraph of two nodes, A and B. The second node is C. The third node generalizes the subgraph of D and E. The forth node generalizes F, G, and H.

Table 3. Adjacency matrix of G_2:

	A	B	C	D	E	F	G	H
A	0	1	0	0	0	0	0	0
B	1	0	1	1	0	1	0	0
C	0	1	0	1	0	1	0	0
D	0	1	1	0	1	1	0	0
E	0	0	0	1	0	0	0	0
F	0	1	1	1	0	0	0	0
G	0	0	0	0	0	1	0	1
H	0	0	0	0	0	1	1	0

Table 4. Adjacency matrix of the generalized graph of G_2:

	A	B	C	D	E	F	G	H
A	$N=2$ $L\_SP=1$		1	1		1		
B								
C	1		$N=1$ $L\_SP=0$	0		1		
D	1		0	$N=2$ $L\_SP=1$		0		
E								
F	1		1	0		$N=3$ $L\_SP=1$		
G								
H								

The generalized graph of G_2 is presented in Figure 3. The generalized node is labeled by one of the known node in the subgraph.

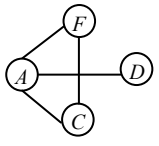

Figure 3. The generalized graph of G_2

The second and third column of Table 5 show the distance computed from G_1 and $\boldsymbol{G}$. $\boldsymbol{G}$ is the integrated graph of G_1 and G_2. The closeness centrality of A computed from G_1 has an error of 42.9% taking the closeness centrality of A computed from $\boldsymbol{G}$. However, if we use the shared information from the generalized graph of G_2, the error can be reduced to 11% using d^{avg} for computation.

Table 5. Distance between A and other nodes and closeness centrality of A as computed from G_1, $\boldsymbol{G}$, and the information shared from the generalized graph of G_2

	G_1	$\boldsymbol{G}$	information sharing from the generalized graph of G_2		
			d^{max}	d^{avg}	d^{min}
$d(A,B)$	1	1	1	1	1
$d(A,C)$	2	2	2	2	2
$d(A,D)$	3	2	3	3	2
$d(A,E)$	4	3	3	3	2
$d(A,F)$	5	2	3	3	2
$d(A,G)$	6	3	3	3	2
$d(A,H)$	7	3	3	3	2
Closeness centrality(A)	7/8 = 0.25	7/16 = 0.44	7/18 = 0.39	7/18 = 0.39	7/13 = 0.54

IV. Experiment

We have conducted an experiment to evaluate the effectiveness of using the subgraph generalization approach for information sharing between social networks to compute the closeness centrality. In this experiment, we randomly generated ten pairs of graphs. For each pair of graphs, we created an integrated graph. We computed the closeness centralities of the nodes in the integrated graph as the benchmark. By using the first graph in each pair of graphs, we computed the closeness centralities of the nodes in this graph and the computed the errors. We then generalized the second graph and used the shared information to compute the closeness centralities of the same set of nodes and the errors again. For the ten pairs of graphs, the average error of closeness centrality computed from the first graph without information sharing is 35%. The average error of closeness centrality computed with shared information of the second using d^{max}, d^{avg}, and d^{min}, are 23%, 17%, and 28%, respectively. It shows that the subgraph generalization approach of information sharing between social networks improve the social network analysis such as closeness centrality computation substantially.

V. Conclusion

We have proposed a subgraph generalization approach for sharing information between terrorist or criminal social networks. The subgraph generalization tries to retain the degree of privacy in information sharing and yet share information with high utility so that an effective social network analysis can still be conducted. In our experiment, it shows that the proposed technique is promising to reduce the error substantially when we compare the computation of closeness centrality without and with information sharing by subgraph generalizations. In the future, we shall further investigate other generalization techniques and determine what information to be shared under different circumstances or analysis.

References

[1] G. Aggarwal, T. Feder, K. Kenthapadi, Samir Khuller, R. Panigrahy, D. Thomas, and A. Zhu, "Achieving Anonymity via Clustering," Proceedings of PODS, June 26-28, Chicago, Illinois, US, 2006.

[2] Z. Baird, J. Barksdale, and M. Vatis, Creating a Trusted Network for Homeland Security, Markle Foundation, 2003.

[3] Z. Baird and J. Barksdale, Mobilizing Information to Prevent Terrorism: Accelerating Development of a Trusted Information Sharing Environment, Markle Foundation, 2006.

[4] C. Best, J. Piskorski, B. Pouliquen, R. Steinberger and H. Tanev, Chapter 2. Automating Event Extraction for the Security Domain, Intelligence and Security Informatics: Applications and Technique, Editors: H. Chen and C. C. Yang, Springer-Verlag, to appear in 2008.

[5] K. Caruson, S. A. Macmanus, M. Khoen, and T. A. Watson, "Homeland Security Preparedness: The Rebirth of Regionalism," *Publius*, 35(1), 2005, pp.143-189.

[6] R. R. Friedmann and W. J. Cannon, "Homeland Security and Community Policing: Competing or Complementing Public Safety Policies," *Journal of Homeland Security and Emergency Management*, 4(4), 2005, pp.1-20.

[7] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam, "l-Diversity: Privacy Beyond k-Anonmity," Proceedings of the 22nd International Conference on Data Engineering, 2006.

[8] P. Samarati, "Protecting Respondents' Identities in Microdata Release," IEEE Transactions on Knowledge and Data Engineering, 2001.

[9] L. Sweeney, "Uniqueness of Simple Demographics in the US population," Technical Report, Carnegie Mellon University, 2000.

[10] L. Sweeney, "k-Anonymity: A Model for Protecting Privacy," International Journal on Uncertainty Fuzziness Knowledge-based Systems, 10(5), 2002, pp.557-570.

[11] D. Thacher, "The Local Role in Homeland Security," *Law & Society*, 39(3), 2005, pp.557-570.

[12] B. Thuraisingham, "Security Issues for Federated Databases Systems," Computers and Security, North Holland, December, 1994.

[13] B. Thuraisingham, Chapter 1. Assured Information Sharing: Technologies: Challenges and Directions, Intelligence and Security Informatics: Applications and Technique, Editors: H. Chen and C. C. Yang, Springer-Verlag, to appear in 2008.

[14] R. C. Wong, J. Li, A. Fu, and K. Wang, "(α,k)-Anonymity: An Enhanced k-Anonymity Model for Privacy-Preserving Data Publishing," Proceedings of SIGKDD, August 20-23, Philadelphia, Pennsylvania, US, 2006.

[15] X. Xiao and Y. Tao, "Personalized Privacy Preservation," Proceedings of SIGMOD, June 27-29, Chicago, Illinois, 2006.

[16] C. C. Yang, N. Liu, and M. Sageman, "Analyzing the Terrorist Social Networks with Visualization Tools," Proceedings of the IEEE International Conference on Intelligence and Security Informatics, San Diego, CA, US, May 23 – 24, 2006.

[17] C. C. Yang and T. D. Ng, "Terrorism and Crime Related Weblog Social Network: Link, Content Analysis and Information Visualization," Proceedings of the IEEE International Conference on Intelligence and Security Informatics, New Brunswick, New Jersey, May 23-24, 2007.