ASSURED INFORMATION SHARING
VOLUME 1: OVERVIEW

Dr. Bhavani Thuraisingham

Professor of Computer Science and
Director of the Cyber Security Research Center
Erik Jonsson School of Engineering and Computer Science
The University of Texas at Dallas
bhavani.thuraisingham@utdallas.edu

ABSTRACT

This paper describes issues, technologies, challenges, and directions for Assured Information Sharing (AIS). AIS is about organizations sharing information but at the same time enforcing policies and procedures so that the data is integrated and mined to extract nuggets. This is the first in a series of papers we are writing on AIS. It provides an overview including architectures, functions and policies for AIS. We assume that the partners of a coalition may be trustworthy, semi-trustworthy or untrustworthy and investigate solutions for AIS to handle the different scenarios.

1. INTRODUCTION

Data from the various data sources at multiple security levels as well as from different services and agencies including the Air Force, Navy, Army, Local, State and Federal agencies have to be integrated so that the data can be mined, patterns and information extracted**,** relationships identified, and decisions made. The databases would include for example, military databases that contain information about military strategies, intelligence databases that contain information about potential terrorists and their patterns of attack, and medical databases that contain information about infectious diseases and stock piles. Data could be structured or unstructured including geospatial/multimedia data. Data also needs to be shared between healthcare organizations such as doctors' offices, hospitals and pharmacies. Unless the data is integrated and the big picture is formed, it will be difficult to inform all the parties concerned about the incidences that have occurred. While the different agencies have to share data and information, they also need to enforce appropriate security and integrity policies so that the data does not get into the hands of unauthorized individuals. Essentially the agencies have to share information but at the same time maintain the security and integrity requirements.

This is the first in a series of reports we are writing on Assured Information Sharing. The reports that follow will include applying game theoretical techniques for AIS among semi-trustworthy partners, defending against malicious attacks while data sharing, applying RBAC (role-based access control) with UCON (Usage Control) extensions for AIS and carrying out offensive operations against untrustworthy partners. We are also investigating risk-based access control, data origin and provenance issues as well as geospatial data management for AIS.

In this paper we describe Assured Information Sharing that will ensure that the appropriate policies for confidentiality, privacy, trust, release, dissemination, data quality and provenance are enforced. We discuss technologies for AIS as well as novel approaches based on game theoretical concepts. In section 2 we will provide an overview of an AIS architecture. Data integration and analysis technologies for AIS will be discussed in section 3. Security policy aspects including confidentiality, privacy and trust policies will be discussed in section 4. Integrity and dependability issues such as data provenance and quality and real-time processing will be discussed in section 5. Balancing conflicting requirements including security vs. real-time processing will be discussed in section 6. Some novel approaches will be discussed in section 7. In particular applications of game theoretical techniques for handling semi-trustworthy partners will be discussed. Approaches for handling untrustworthy partners will be dis-

cussed in section 8. Discussion of the series of reports we will be writing on AIS is mentioned in section 9. The paper is concluded in section 10.

2. ORGANIZATIONAL DATA SHARING

A coalition consists of a set of organizations, which may be agencies, universities and corporations that work together in a peer-to-peer environment to solve problems such as intelligence and military operations as well as healthcare operations. Figure 1 illustrates an architecture for a coalition where three agencies have to share data and information. Coalitions are usually dynamic in nature. That is, members may join and leave the coalitions in accordance with the policies and procedures. A challenge is to ensure the secure operation of a coalition. We assume that the members of a coalition, which are also called its partners, may be trustworthy, untrustworthy or partially (semi) trustworthy.

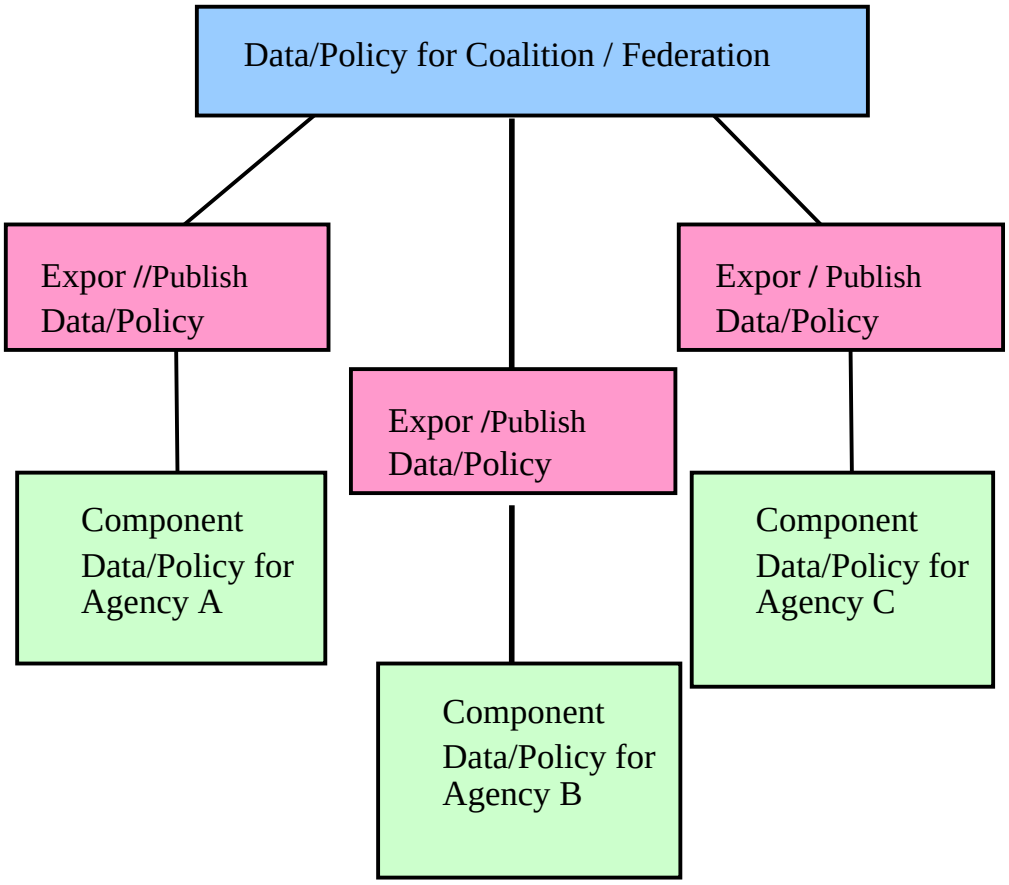

Figure 1. Architecture for Organizational Data Sharing

Various aspects of coalition data sharing are discussed in the Markle report [MARK03]. However, security including confidentiality, privacy, trust, integrity, release and dissemination has been given little consideration. Much of the prior work on security in a coalition environment has focused on secure federated data sharing. Thuraisingham was one of the first to propose multilevel security for federated database systems [THUR94]. Discretionary security was proposed in [OLIV95]. None of the previous work has focused on determining the amount of information that is lost for conducting military operations by enforcing security. Furthermore, developing flexible policies in a coalition environment are yet

to be examined. Enforcing security while meeting timing constraints remains a largely unexplored topic. A discussion of information survivability issues and the need for flexible policies for enforcing security and meeting timing constraints are given in [THUR99] and [SON95]. However, to our knowledge, no research has been reported on secure (including confidentiality, privacy, trust and integrity) and timely data sharing for a coalition environment. Some of the challenges include the following:

Data Sharing: One of the main goals of coalition data sharing is for organizations to share the data but at the same time maintain autonomy. For example, one database could be used for travel data while another database could be used to manage data pertaining to airplanes. For counter-terrorism applications and military operations, the key is to make links and associations as rapidly as possible. We need policies and procedures to determine what data to share under what conditions.

Data Mining: Data mining techniques extract patterns and trends often previously unknown from large quantities of data [THUR98]. However data mining tools could give out false positives and false negatives. This is especially critical for applications such as counter-terrorism and military operations as it could result in catastrophic consequences [THUR03]. Therefore, we need human analysts to examine the patterns and determine which ones are useful and which ones are spurious. The challenge is to develop automated tools to sift through the data and produce only the useful links and associations.

Security: Confidentiality, privacy, integrity, trust, real-time processing, fault tolerance, authorization and administration policies enforced by the component organizations via the local agencies have to be integrated at the coalition level. As illustrated in Figure 1, each organization may export security policies and data to the coalition. The component systems may have more stringent access control requirements for foreign organizations. The challenge is to ensure that there is no security violation at the coalition level.

In sections 3 through 6 we discuss various aspects on AIS assuming that the partners are trustworthy. Semi-trustworthy partners will be discussed in section 7. Untrustworthy partners will be discussed in section 8.

3. DATA INTEGRATION AND ANALYSIS TECHNOLOGIES

Data Integration: As illustrated in Figure 2, data from the various data sources at multiple levels such as local, state and federal levels have to be integrated so that the data can be mined, patterns extracted and decisions made. Data integration has been attempted for about 20 years. Until recently brute force integration techniques consisting of translators and gateways were used between the multiple data management systems. Standards such as RDA (Remote Database Access) were developed initially for client-server interoperability. Later object-base wrappers were used to encapsulate the multiple systems including the legacy systems. For example, distributed object management standards were used to encapsulate systems and applications into objects. However, common representation of the data remained a challenge. It is only recently that we have a good handle on syntactic integration through standards such as XML (eXtensible Markup Language). The idea is as follows: each data system publishes its schema (also called metadata) in XML. Since all the systems now represent their schema in XML, the systems can talk to each other in a seamless fashion.

A major challenge for data integration is semantic heterogeneity. While much progress has been made on syntactic integration, not much work has been reported on semantic integration. For example, multiple systems may use different terms for the same data; the procedure EKG (Electro Cardiogram) is called ECG in the United Kingdom. Even within the same state, different hospitals may use different terms to mean the same entity. For example, one hospital may use the term influenza while another hospital may use the term flu. In some cases, the same term may be used to represent different entities. While repositories and dictionaries have been built, a satisfactory solution for semantic heterogeneity is

still not available. The development of semantic web technologies including the Resource Description Framework (RDF) language standard shows promise to handle semantic heterogeneity.

Multimedia and Geospatial Data: Data will include structured data as well as unstructured data such as text, voice, video and audio. Data emanating from multiple data sources including sensor and surveillance data have to be integrated and shared. Managing, integrating and mining multimedia data remains a challenge. We need efficient indexing techniques as well as XML and RDF based representation schemes. Furthermore, the data has to be mined so that patterns and trends are extracted. Video data could be data emanating from surveillance cameras or news feeds such as CNN (Cable News Network) video data. Emergency response systems have to integrate geospatial data such as maps together with structured data, make sense out of the data and rapidly produce summaries so that the emergency response teams can read and understand the data [ASHR06].

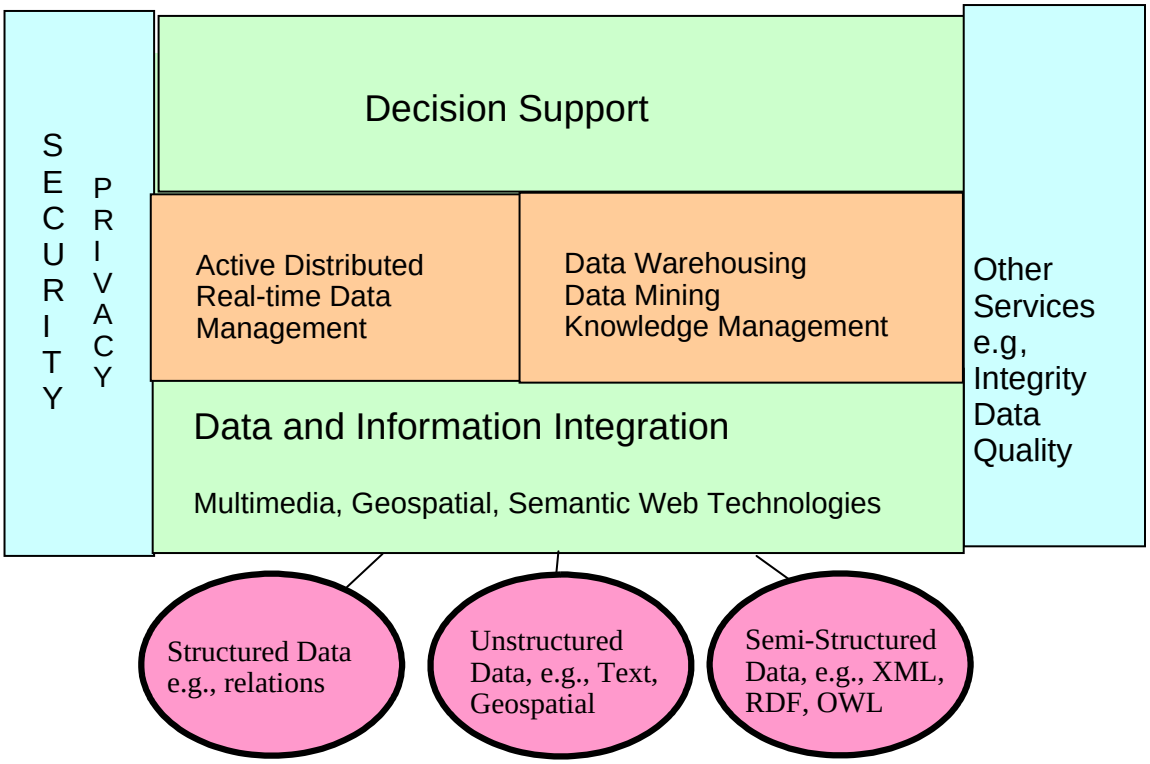

Figure 2. Data Integration and Analysis

Data Mining: Integrated data may be mined to extract patterns for suspicious and unusual behavior. Much of the work in data mining has focused on mining relational and structured databases. While some work has been reported on text, image, audio and video data mining, much remains to be done. For example, how can one mine integrated geospatial and multimedia data? How can false positives and false negatives be eliminated or at least reduced? What are the training models used for multimedia data? What are the appropriate outcomes for multimedia data mining? Does it make sense to extract metadata and then mine the metadata? Much remains to be done before operational tools for multimedia and geospatial data mining are developed.

Web Services: The Department of Defense (DoD) as well as other agencies is migrating toward service oriented architectures (SOA). For example, the Network Centric Operations Architecture is based on

SOA and the services are called Network Centric Enterprise Services (NCES). Furthermore, The Global Information Grid (GIG) is based on SOA. In a coalition environment, the agencies will publish their policies and schema as illustrated in Figure 1, and communicate with each other using web services technology.

Semantic Web: Semantic web is the vision of Tim Berners Lee and is utilized by web services and other applications including e-business [LEE01]. Due to the extensive investments by the DoD (Department of Defense) and other agencies, many semantic web technologies such as XML, RDF and Ontologies have been developed for applications such as interoperability. Furthermore, semantic web technologies are being developed for different communities. These technologies are critical for AIS. For example, we need ontologies specified in languages such as OWL (web ontology language) to specify objects so that multiple systems can work with the ontologies to handle semantic heterogeneity. A member organization of a coalition can publish its schema in languages such as XML or RDF to facilitate interoperability and information extraction.

While semantic webs are being developed for different communities, there is little work on enforcing security, privacy and trust for these semantic webs. XML, RDF and Ontologies have to be secure. Furthermore, there is a need to incorporate trust negotiation for the semantic web. We are developing secure semantic web technologies for AIS [BERT04], [THUR05a].

4. SECURITY POLICY ENFORCEMENT

Security policies include policies for confidentiality, privacy, trust, release, dissemination and integrity. A broader term is dependable systems or trustworthy systems that also include real-time processing and fault tolerance. We will discuss dependability in the next section. By confidentiality we mean that data is only released to individuals who are authorized to get the data. Privacy in general deals with the situation where an individual determines what information should be released about him/her. (Note that different definitions of privacy have been proposed.) Trust policies may add further restriction to privacy and confidentiality policies. For example, a user may be authorized to get the data according to the confidentiality policies, but the system may not trust the individual in which case the data is not released. Similarly a person may give permission to release certain private information about him or her but that person may not trust a particular web site in which case the private information is not released to the web site. Alternatively one could argue that one needs to establish trust first before establishing the confidentiality and privacy policies. For example, a user's (or web site's) trust is established before determining that the user (or web site) can received confidential (or private) information. Release policies specify rules for releasing data while dissemination policies specify rules for disseminating the data. Integrity within the context of security ensures that only authorized individuals can modify the data so that the data is not maliciously corrupted [TSYB06]. We are conducting extensive investigation on privacy preserving data mining [LIU05]. We are also investigating the use of these techniques for AIS [LIU06].

Security for relational databases has been studied extensively and standards such as secure SQL (Structured Query Language) have been developed. In addition several secure data management system products have been developed. There has been research on incorporating security into next generation data management systems. There is also work on data quality as well as trust management. Security has also been investigated for secure object request brokers as well as for secure e-commerce systems. Finally W3C (World Wide Web Consortium) is specifying standards for privacy such as the P3P (Platform for Privacy Preferences). While there is research on incorporating security for semantic webs and heterogeneous data systems, this research is in the early stages. There is an urgent need to develop operational systems that enforce security. Furthermore, security has conflicting requirements with real-time processing. We need to enforce flexible policies and subsequently standards for specifying these policies.

Security is critical for many of the information technologies we have discussed here. For a discussion of secure data sharing and related standards we refer to [THUR05b].

Security Policy Integration: There is a critical need for organizations to share data as well process the data in a timely manner, but at the same time enforce various security policies. Figure 3 illustrates security policy integration in a coalition environment. In this example, A and B form a coalition while B and C form a second coalition. A could be California, B could be Texas and C could be Oklahoma. California and Texas could form a coalition as part of the larger states in the US and Texas and Oklahoma could form a coalition as part of the neighboring states in the South of US for emergency management. There is also an urgent need for multiple organizations to share data and at the same time enforce security policies. These policies include policies for confidentiality, privacy, and trust. For example, patient data may be shared by multiple organizations including hospitals, levels of government and agencies. It is important to maintain the privacy of patient data. However it is also important that there are no unnecessary access controls so that information sharing is prohibited. One needs flexible policies so that during emergency situations it is critical that all of the data is shared so that effective decisions can be made. During normal operations, it is important to maintain confidentiality and privacy. In addition, trust policies ensure that data is shared between trusted individuals. The standards efforts in this area include Role-based access control (RBAC) [SAND96] as well as P3P (Platform for Privacy Preferences). Our partners at George mason University are examining the use of models such as RBAC and UCON for AIS [SAND06].

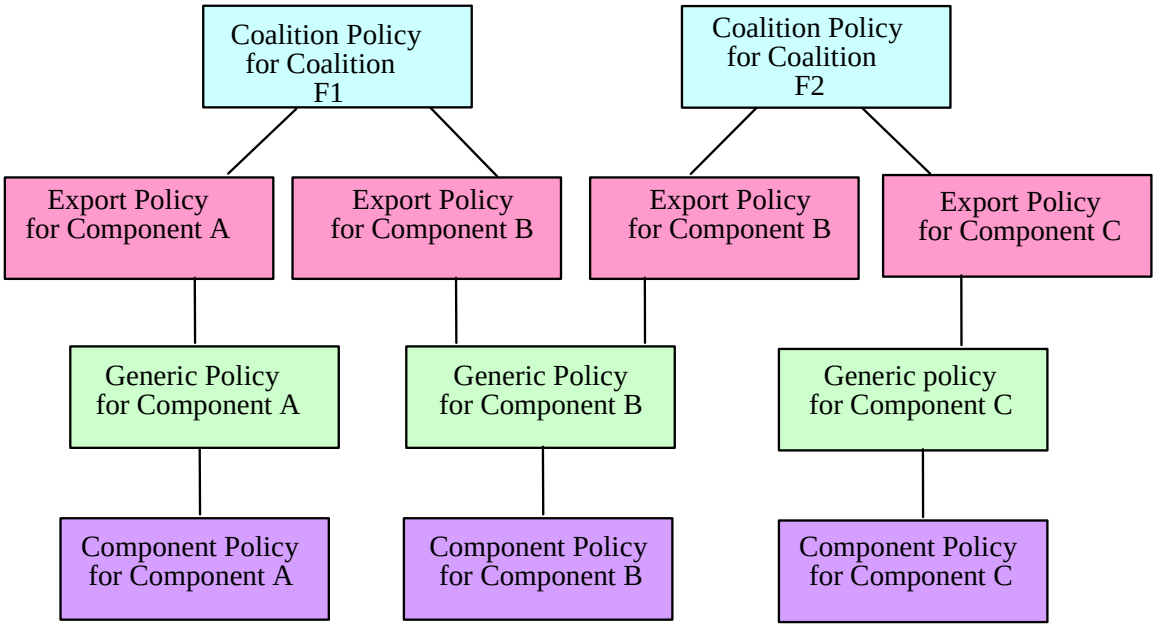

Figure 3. Security Policy Integration and Transformation for Coalitions

5.DEPENDABILITY ASPECTS

By dependable systems we mean systems that are fault tolerant and meet timing constraints. The time-critical, information-sensitive goals of managing a crisis include actions such as the early confirmation of cases and correct identification of exposed populations over a relevant time period. Early confirmation means that triggers have to be activated when certain situations (such as anomalies) occur. Suppose a hospital is flooded with 30 patients within 15 minutes who are all reporting a temperature of 105 de-

grees. There has to be a rule such as “If more than 30 patients register at a hospital within 20 minutes with temperature greater than 102 degrees then alert the emergency response system”. To effectively process a large number of rules, we need active data management. Furthermore, the various parties involved such as federal, state and local governments have to be informed within a certain time. That is, if the authorities are notified after say 2 hours then it will be difficult to contain the spread of the disease. This means we need real-time data management capabilities. Some initial research on dependable and secure systems is discussed in [KIM06a].

While there are techniques for active real-time data management, the challenge is to develop an integrated system for end-to-end data management. For example, the data manager will ensure that the data is current and the transactions meet the timing constraints. However in an emergency situation there are numerous dependencies between different data sources. For example when rule A gets triggered, that would result in rules C, D, and E getting triggered in multiple data management systems. Such chain rule processing remains a challenge. We also need end-to-end real-time processing. That is, in addition to the data manager, the infrastructure, the network and the operating system have to meet timing constraints. This remains a challenge. Incorporating security into real-time processing techniques remains largely unexplored. For example, in an emergency situation, real-time processing and activating triggers may be more critical than enforcing access control techniques. Furthermore, the system must ensure that the deadlines are not missed due to malicious code and attacks (e.g., denial of service).

While integrity within the context of security implies that the data is not maliciously corrupted, integrity also includes policies for data quality and data provenance management. Data quality determines the accuracy of the data. This would depend on who updated the data, who owns the data and what is the accuracy of the source of the data. That is, as data moves from organization to organization, its quality may vary. Some measure to compute the quality of the data is needed. Data provenance is about maintaining the history of the data. That is, information as to who accessed the data from start to finish is needed to determine whether data is misused [KIM06b].

6. BALANCING CONFLICTING REQUIREMENTS

There are two types of conflicting requirements: one is security vs. data sharing. The goal of data sharing is for organizations to share as much data as possible so that the data is mined and nuggets obtained. However when security policies are enforced then not all of the data is shared. The other type of conflict is between real-time processing and security. The war fighter will need information at the right time. If it is even say 5 minutes late the information may not be useful. This means that if various security checks are to be performed then the information may not get to the war fighter on time.

We are conducting research in both areas. For example, we are integrating the data in the coalition databases without any access control restrictions and apply the data mining tools to obtain interesting patterns and trends. In particular, we are developing associations between different data entities such as “A and B are likely to be in a location 50 miles from Baghdad”. Next we are using the same tool on the integrated data after enforcing the policies. We can then determine the patterns that might be lost due to enforcing the policies (note that there is some relationship between this work and the research on privacy preserving data mining). Our research is described in [AWAD06].

In addition, we are conducting research on examining the extent to which security affects timing constraints. For example, we enforce timing constraints on the query algorithms. That is, we first process the query using the enforcement algorithms without enforcing any of the policies. Then we enforce the security policies and determine whether the timing constraints can be met. This will determine the extent to which security impacts timely information processing.

Our goal is to develop flexible approaches and balance conflicting requirements. That is, if timely processing of data is critical then security has to be relaxed. Similarly say during non combat operations, security will have to be given full consideration. The same applies for data sharing vs. security. If during an emergency operation such as say the operation just before, during or soon after Hurricane Katrina, then several agencies will need the data without any restrictions. However during non emergency operations, security policies need to be enforced. Our research is reported in [KIM06c]. In particular, we are examining the application of RBAC and UCON models for timely data sharing.

Another aspect of our research on AIS is risk analysis. For example, if the security risks are high and the cost to implement security features are low, then security should be given high consideration. If the risks are low and the cost is high, one needs to evaluate whether it is worth the effort and cost to incorporate security. Our research on risk based access control is reported in [CELI06].

7. GAME THEORY APPLICATIONS AND SEMI-TRUSTWORTHY PARTNERS

In the previous sections we assumed that the organizations were trustworthy and would enforce the policies while data sharing. However in many cases the organization may be semi-honest or completely dishonest. In the case of semi-honest partners, organizations may have to play games to extract data. In the case of dishonest and untrustworthy partners, one may not only have to defend against malicious code, but also have to figure out what the partner is up to by monitoring his machine. In this section we will address semi-trustworthy partners and in the next we will discuss untrustworthy partners.

Semi-Honest Partners and Game Playing

To handle secure data sharing especially with semi-trustworthy partners, modeling the query processing scenario as a non cooperative game may be more appropriate especially between two partners. The players are the partners, which could be agencies or countries of a coalition. Lets assume we have Agency A and B as two partners. The objective of agency A is to extract as much information as possible from agency B. Essentially agency A wants to compromise information managed by Agency B. B's goal is to prevent this from occurring. Cooperative games on the other hand may have applications among friendly partners of a coalition. A mixture of cooperative and non-cooperative strategies may be applied for multi-party coalition.

Two-party information sharing: Information sharing between two agencies A and B may be modeled as a non-cooperative game. A has a specific objective; for example, it may know that B has some sensitive data and it wants to extract the value of that data from B. B knows A's objective. A move made by A is a query. A move made by B is the response. The game continues until A achieves its objectives or gets tired of playing the game. As stated in [JONES80], the game can be represented as a graph theoretic tree of vertices and edges. The tree has a distinguished vertex, which is the initial state. There is a payoff function, which assigns a pair of values say (X,Y) where X is the payoff for A and Y is the pay for B for each move. The payoff for A is high if it is close to obtaining the sensitive value. The payoff for B is high if the response does not reveal anything about the sensitive value. Note that if B does not give out any information or if it gives erroneous information then it cannot be regarded as a game, That is, the aim here is for B to participate in the game without giving away sensitive information.

Multi-party information sharing: The idea here is that certain parties play cooperative games while certain other parties play non-cooperative games. We illustrate with an example consisting of three parties. Let's consider an example. Suppose the year is 2006 and the UK has obtained some sensitive information on Operation Iraqi Freedom that the US needs. However, the UK is reluctant to share this information. The US in the meantime has formed an alliance with Argentina by giving some incentive either in the form of money or weapons. When the UK hears this, it is concerned thinking about the Falklands situation. However, in reality the US has no intention of doing anything about the Falklands but

does not want the UK to know the truth. So the UK may reason about the benefits it receives by sharing the data with the US and makes a determination.

Cooperative games have also been called Coalition games. In a true coalition the players are friendly and therefore share the information and determine a collective payoff. However in our environment, organizations form coalitions only to solve a particular problem. An agency that is a trustworthy party in a particular coalition may turn against its partner at a later time and divulge the information gathered during the coalition operation.

We have conducted some initial research on game theory applications for AIS. Our objective has been to consider the interaction of participants within a loose coalition. In particular, we are interested in a scenario in which those involved have made a reluctant but necessary decision to trade information to achieve some goal. A great deal of work has already been done in the areas of secret sharing and protocol enforcement. However, even if agreements to exchange are kept, there is no guarantee what is shared is legitimate. The ultimate goal of this research is to create a behavior which works optimally against lying agencies while taking advantage of implicit trust. Our results at this point in the research suggest our algorithm is effective against basic opponents, though more refinement is needed. We report which behaviors work for the players and why, with regards to the motivating factors for each strategy. Our research will be described in Volume 3 of these series [LAYF06].

8.HANDLING UNTRUSTWORTHY PARTNERS

Note that in fighting the global war on terrorism we have to work with our allies as well as with countries that we may not trust. If our partners our untrustworthy, then we have to not only defend against malicious code but also figure out what the partners are doing both with their computers as well as their activities. Essentially we need to conduct information operations [SPIT02]. We will first discuss our research on defensive operations and then discuss some aspects of offensive operations.

Defensive Operations: In the case where partners are untrustworthy we have to defend ourselves against malicious code such as viruses and worms planted by our partners. In order to accomplish this, we are applying data mining techniques to detect such malicious code. Some of our research in this area can be found in [MASU06] and will be published in Volume 4 of these series [KHAN06].

Offensive Operations: There is little work in the unclassified published literature on offensive operations. However recently we are seeing articles published in Signal magazine on the importance of monitoring the adversaries' computing activities [SIGN05a], [SIGN05b]. Three of the techniques for handling untrustworthy partners include the following:

Trojan Image Exploitation: Modern anti-virus and anti-spy ware detection packages rely on the presence of malicious code within an executable or script to prevent attacks. This is done by detection methods that are carried out when the program first loads. In theory, it is possible to circumvent this detection by designing a program without any explicit malicious code; instead, a memory leak in this program's security is purposefully created. This weakness is exploited by downloading a tailored file from the Internet, such as a picture, after the program is loaded. As a result, this program could be used as a staging area for a malicious attack.

Web Browser Customization: Web browsers have been enhanced dramatically in the past year to prevent attacks from malicious web pages. For the benefit of the user, these features are frequently made optional, allowing a great deal of customization. By compromising a user's customization features covertly, it becomes possible to execute potential attacks without the user detecting any warning signs normally visible in the user's browser such that the attacker's methods can be hidden from the user. The attacker could use browser customization, such as enabling JavaScript, to create a shadow copy of the web and gain classified information from the victim without certain warning signs, such as URLs being

correctly displayed. All user-entered information would be funneled through the attacker's spoofed world and thus the attacker could easily take advantage of the situation in order to retrieve any type of information.

Message Interception: Enron data set (publicly available) may be used to send emails to the partners of the coalition as well as to those outside of the coalition. Messaging may be simulated in such a way that they are sent at random intervals. We can then determine whether interception techniques can be used to extract some of the messages sent. This is a very challenging problem.

9.SERIES OF REPORTS

As we have stated in section 1, this paper is the first in a series of papers we will publish as part of the AIS Technical Report series at the University of Texas at Dallas. In this section we briefly discuss the contents of some of the other reports.

Experimental Analysis: In this report we will discuss the experiments we are conducting on how much information ism lost by enforcing security policies in a coalition environment.

Game Theory Applications: In this report we will discuss the application of game theoretic techniques for extracting information when partners are semi-trustworthy.

Defensive Operations: In this report we will discuss our approach to defending the systems from worms when the partners are untrustworthy.

RBAC for AIS: In this report our partners at GMU will discuss the application of Role-based access control for assured information sharing.

Offensive Operations: In this report we will discuss techniques for finding out the activities of untrustworthy partners.

We are conducting research in related topics that will support AIS. Some related reports that we will publish include the following:

Risk-based access control: In this report we will discuss data sharing when taking security risks into consideration.

Data provenance: In this report, we will use healthcare applications as an example and discuss data provenance issues for AIS.

Dependable Data Sharing: In this report we will describe our approach to systems meeting security as well as real-time requirements.

Standards: In this report we will discuss data integration standards for AIS.

Privacy Preserving Data Sharing: In this report we will discuss data sharing and at the same time ensuring privacy of the individuals using healthcare applications.

Geospatial data: In this report we will discuss assured information sharing for geospatial and unstructured data.

Semantic web: In this report we will explore the use of web services and semantic web technologies for AIS

Social network analysis: In this report we will examine how organizations form networks and discuss approaches for supporting AIS.

Infrastructure: In this report we will investigate how infrastructures such a data grids support AIS.

In addition to the above reports, we will also publish reports on the implementation of the designs of systems for AIS. For example, implementation of the systems we have designed for geospatial data sharing, risk-based access control and game theory applications will be described in future technical reports.

10. SUMMARY AND DIRECTIONS

In this paper we have defined Assured Information Sharing (AIS) and discussed issues, technologies, challenges and directions for this area. The goal of AIS is for organizations to share data but at the same time enforce security policies. Security includes confidentiality, privacy, trust, and integrity policies. We discussed approaches for AIS when the partners of a coalition are trustworthy, semi-trustworthy and untrustworthy. In particular, we discussed security policy enforcement, game theory applications and defending against worms and viruses. We also discussed AIS technologies including data integration, data mining, and the semantic web.

There are several areas that need further investigation. We need to develop policies for accountability. This is especially important in a coalition environment. In such an environment, there are numerous pieces of hardware and software that interact with each other. Therefore, the action of all the processes has to be recorded and analyzed. Furthermore, risk analysis studies are needed to determine the risks and developing appropriate solutions. For example, in a high risk low cost security environment, there will be no questions about implementing security solutions. However in a low risk high cost environment one needs to think twice before enforcing the security policies. Essentially we need some form of risk-based AIS. We also need to develop web services for AIS. Essentially we need to integrate AIS and semantic web technologies. Finally we need to investigate several additional technologies such as collaborative services, social network analysis, surveillance data sharing, digital identity management, metadata extraction and management as well as policies for identification and authentication for AIS. We also need to investigate the use of standards as well as infrastructures such ass data grids for AIS. Some of our preliminary research in some of these topics is reported in [THUR05b], [ZHU06], [LAVE05], [LAYF05].

We are conducting extensive investigation on AIS with our partners George Mason University and Purdue University. In addition to the technical aspects discussed in this paper, we are also investigating the connection between AIS and the Global Information Grid as well as Network centric Operations. While our primary application is counter-terrorism, we are also focusing on other applications such as Emergency preparedness and Healthcare. Future papers will focus on the design of our approaches as well as our experimental results for AIS.

ACKNOWLEDGEMENTS

I thank Dr. Robert Herklotz for funding our research on Information Operations Across Infospheres from which I got the ideas to write this paper. I thank my colleagues Profs. Latifur Khan, Murat Kantarcioglu, Ravi Sandhu and Elisa Bertino as well as Dr. Mamoun Awad and Dr. Ebru Celikel for discussions and inputs on AIS. I also thank my students Ryan Layfield, Nathalie Tsybulnik, Li Liu, Alam Ashraful, Ganesh Subbiah, Gal Lavee and Kim Jungin as well as many others for discussions on AIS, and especially Ryan Layfield, Nathalie Tsybulnik and Li Liu for writing the techniques for information operations in section 8.

REFERENCES

[AWAD06] M. Awad, B. Thuraisingham, and L. Khan, et al, Assured Information Sharing: Volume 2: Experimental Analysis of Data Integration, Mining and Security, Technical Report, The University of Texas at Dallas, 2006 (to appear)

[ASHR06] A. Ashraful, G. Subbiah, L. Khan, and B. Thuraisingham, Geospatial Semantic Web, Technical Report, The University of Texas at Dallas, 2006 (to appear).

[BERT04] E. Bertino, B. Carminati, E. Ferrari and B. Thuraisingham, *Secure Third Party Publication of XML Documents*, IEEE Transactions on Knowledge and Data Engineering, October 2004

[CELI06] E. Celikel, M. Kantarcioglu and B. Thuraisingham, Assured Information Sharing: Risk-based Data Sharing, Technical Report, The University of Texas at Dallas, 2006 (to appear)

[JONE80] A. Jones, Game Theory, Mathematical Models of Conflict, Halstead Press, 1980.

[KHAN06] L. Khan, B. Thuraisingham et al, Assured Information Sharing: Volume 4: Data Mining Applications for Defensive Operations in a Coalition, Technical Report, The University of Texas at Dallas, (to appear).

[KIM06a] J. Kim and B. Thuraisingham, Dependable and Secure TMO Scheme, Proceedings of IEEE ISORC Conference, April 006.

[KIM06b] J. Kim, B. Thuraisingham, et al, Data Provenance in Healthcare Systems: Survey and Research Issues, UTD Technical Report, to appear.

[KIM06] J. Kim and B. Thuraisingham, Applying RBAC and UCON to TMO, Technical report, University of Texas at Dallas, to appear.

[LAVE05] G. Lavee et al, Suspicious Event Detection with Surveillance Data, Proceedings of the ACM SIGKDD Conference Workshop on Multimedia Data Mining, 2005.

[LAYF05] R. Layfield, et al, Design of a Social Network Analysis System, Proceedings of the ACM SIGKDD Conference Workshop on Multimedia Data Mining, 2005.

[LAYF06] R. Layfield, M. Kantarcioglu and B. Thuraisingham, Assured Information Sharing: Volume 3: Using Game Theory to Enforce Honesty Within a Competitive Coalition, Technical Report, The University of Texas at Dallas, 2006 (to appear)

[LEE01] Berners Lee, T., et al., The Semantic Web, Scientific American, May 2001.

[LIU05] L. Liu, M. Kantarcioglu, N. Thuraisingham, L. Khan, An Adaptable Perturbation Model of Privacy Preserving Data Mining, Proceedings of the IEEE ICDM Data Mining Conference Workshop on Privacy preserving Data Mining, 2005 (also published as technical report, UTDCS-03-06, January 2006).

[LIU06] L. Liu, et al, Privacy Preserving Data Sharing, Technical Report, The University of Texas at Dallas, 2006 (to appear)

[MARK03] Creating a Trusted Network for Homeland Security, Markle Report, 2003 (Editor: M. Vatis)

[MASU06] Masud, M, L. Khan, B. Thuraisingham and M. Awad, Detecting New malicious Executables Using Data Mining, UTDCS-27-06 Technical Report, The University of Texas at Dallas, June 2006, also submitted for publications. (version to be published as UTD AIS Technical Report series)

[NCW05] The Implementation of Network Centric Warfare, Office of Force Transformation, 2003.

[OLIV95] Martin S. Olivier: Self-protecting Objects in a Secure Federated Database, Proceedings of the IFIP Database Security Conference, NY, August 1995.

[SAND96] Ravi Sandhu, Edward Coyne, Hal Feinstein and Charles Youman, "Role-Based Access Control Models." *IEEE Computer*, Volume 29, Number 2, February 1996.

[SAND06] R. Sandhu et al, RBAC for AIS, to be published as AIS Technical Report Series, 2006.

[SIGN05a] Signal Magazine, AFCEA, May 2005

[SIGN05b] Signal Magazine, AFCEA, February 2005

[SPIT02] Lance Spitzner, Honeypots, Tracking Hackers, Addison Wesley, 2002.

[SON95] S. Son, R. David and B. Thuraisingham, *An Adaptive Policy for Improved Timeliness in Secure Database Systems*, Proceedings of the 9th IFIP Working Conference in Database Security, New York, August 1995.

[THUR90] B. Thuraisingham, *Novel Approaches to the Inference Problem*, June 1990, Proceedings of the 3rd RADC Database Security Workshop, New York.

[THUR94] B. Thuraisingham, *Security Issues for Federated Database Systems*, 1994, Computers and Security (North Holland), December 1994.

[THUR98] B. Thuraisingham, Data Mining: Technologies, Techniques, Tools and Trends, CRC Press, December 1998.

[THUR99] B. Thuraisingham and J. Maurer, *Information Survivability for Real-time Command and Control Systems*, IEEE Transactions on Knowledge and Data Engineering, January 1999

[THUR03] B. Thuraisingham, Web Data Mining and Applications in Business Intelligence and Counter-terrorism, CRC Press, Boca Raton, FL, 2003.

[THUR05a] B. Thuraisingham, Security Standards for the Semantic Web, Computer Standards and Interfaces Journal, 2005.

[THUR05b] B. Thuraisingham, Database and Applications Security: Integrating Information Security and Data Management, CRC Press, May 2005

[THUR06] B. Thuraisingham, D. Harris, L. Khan, R. Paul, "Standards for Secure Data Sharing across Organizations," Accepted in Computer Standards and Interfaces Journal, 2005. (version to be published as part of UTD AIS technical report series)

[TSYB06] N. Tsyblinik, B. Thuraisingham, A. Ashraful, CPT: Confidentiality, Privacy and Trust for the Semantic Web,UTDCS-06-06, Technical Report, the University of Texas at Dallas, March 2006, Also to appear in the Journal of Information Technology Management.

[ZHU06] J. Zhu, B. Thuraisingham, Grid Computing and Grid Security, Technical; Report, The University of Texas at Dallas, to appear.