2006

A Knowledge Framework for Information Security Modeling

Shuangyan Liu
City University of Hong Kong

Ching-hang Cheung
City University of Hong Kong

Lam-for Kwok
City University of Hong Kong

DOI: 10.4225/75/57b65e7234771
Originally published in the Proceedings of 4th Australian Information Security Management Conference, Edith Cowan University, Perth, Western Australia, 5th December, 2006

A Knowledge Framework for Information Security Modeling

Shuangyan Liu, Ching-hang Cheung and Lam-for Kwok
Department of Computer Science, City University of Hong Kong
jenny@cs.cityu.edu.hk, Franky.CH.Cheung@gmail.com, cslfkwok@cityu.edu.hk

Abstract

The data collection process for risk assessment highly depends on the security experience of security staffs of an organization. It is difficult to have the right information security staff, who understands both the security requirements and the current security state of an organization and at the same time possesses the skill to perform risk assessment. However, a well defined knowledge model could help to describe categories of knowledge required to guide the data collection process. In this paper, a knowledge framework is introduced, which includes a knowledge model to define the data skeleton of the risk environment of an organization and security patterns about relationships between threat, entity and countermeasures; and a data integration mechanism for integrating distributed security related data into a security data repository that is specific to an organization for information security modelling.

Keywords

Risk Analysis, Information Security Management, Knowledge Representation, Data Integration

INTRODUCTION

Most enterprises agree that knowledge is an essential asset for survival and success in an increasingly competitive market; this awareness is one of the main reasons for the exponential growth of knowledge management in the past decade [Benjamins, et. al, 1998]. Knowledge can be of any kind: tacit, documented, or procedural. Information security knowledge is of prime importance in maintaining good information security of an organization, especially in the field of information security risk assessment.

Risk Assessment activity involves an onerous data collection process, which includes collecting information about assets, threats, vulnerabilities and so on. However, the competence to collect data efficiently is based on security experience about threat and countermeasures in different areas [Landoll, 2006]. A methodology which could help to gather security data based on previous security experience is necessary.

In this paper, a knowledge framework is presented to guide the data collection process for information security modelling. This framework includes a knowledge model to provide detailed definitions of security knowledge and a data integration mechanism to apply the security knowledge maintained in the model to guide the data collection process.

INFORMATION SECURITY MODELING

Information Security Model ([Kwok et. al, 1997], [Kwok et. al, 1999], [Kwok et. al, 2004]), which accumulates operational data and security experience, is formulated to assist the data collection process for risk analysis studies. It aims to collect all currently available information security data, and to evolve over time by incorporating new data.

A prototype of ISM ([Fung et. al, 2003], [Kwok et. al, 2001]) provides a directory structure to store security documentation electronically. The electronic security documentation provides a common source of information to a wide range of staff with information security responsibilities, and minimizes the duplication of data collection effort.

However, this directory structure makes security data and security knowledge mixed together. It is inconvenient to store and update security related data within an organization, and security knowledge cannot be reused and shared in a long term for periodical risk analysis process.

KNOWLEDGE FRAMEWORK FOR INFORMATION SECURITY OF AN ORGANIZATION

For the information security modelling, as mentioned above, our knowledge framework aims to provide a methodology to gather updated security data that are specific to the data requirements of the data collection process under the guidance of security knowledge that are maintained in this framework.

Overview of the Knowledge Framework

The framework mainly consists of two parts (Fig.1):

- A knowledge model; and
- A data integration mechanism.

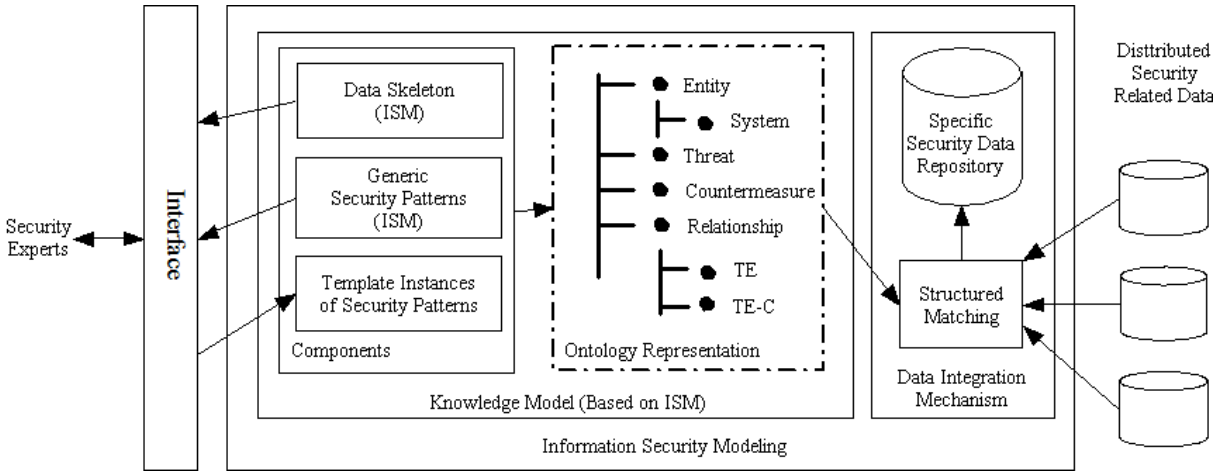

Fig. 1. The Knowledge Framework for Information Security Modelling

Three components of the knowledge model are presented: *Data Skeleton*, *Generic Security Patterns* and *Template Instances of Security Patterns*. The former two components are based on previous ISM, in which *Data Skeleton* represents the data structure about the risk environment of an organization and *Generic Security Patterns* depict relationships between threat, entity and countermeasures. In our study, Interface is provided to security experts to create *Template Instances of Security Patterns* based on their security experiences and the former two components in the model. The *Ontology Representation* in the knowledge model elaborates definitions of the components. The benefits of using ontology representation will be discussed later.

The data integration mechanism in the framework helps to integrate distributed security related data into *Specific Security Data Repository* under the guidance of security knowledge that are represented in the ontology (details described later). *Structured Matching* between data requirement schema (the global schema) and source schemas is proposed.

COMPONENTS OF THE KNOWLEDGE MODEL

Data Skeleton

The proposed data skeleton represents structure of the risk context of an organization, but will not contain any security related data specific to an organization, which is the capacity of Specific Security Data Repository. The skeleton includes [Anderson et. al, 1994]:

- *Environment* group including buildings, sites, and services, which could directly affect the operation of an organizational information processing system and indirectly affect the data resided in it;
- *System* group comprising information processing systems and essential personnel which may affect the secure operation of the systems; and

- *Assets* group referring to the information assets in the form of electronic data. This area aims to provide security officers the information about the business impact of a security event. The terminate influence of a threat can be stated as a business impact.

Generic Security Patterns

Similar to the data skeleton, generic security patterns are also based on the previous ISM, which depict patterns of threat propagating in the above groups and patterns of measures countering the threat.

The concept for threat propagation pattern is that a threat acting on an entity may cause another threat acting on another entity. Generic Threat-Entity relationship (TE) and TETE relationship presented in ISM [Kwok et. al, 2004] are used to express the generic security pattern of threat propagation in the model.

The definition of countermeasure pattern has two considerations: (1) an appropriate proposition of a countermeasure to each Threat-Entity pair; (2) deployment of additional countermeasures to assure the effectiveness of the original countermeasure. Generic ThreatEntity-Countermeasure relationship (TE-C) and TE-C-TE relationship presented in ISM [Kwok et. al, 2004] are applied to define the generic security pattern of countermeasure in the model.

Template Instances of Security Patterns

In fact, template instances of generic security patterns are template instances of the generic relationships mentioned above. For example, the threat scenario “user smart card is duplicated” can be represented as an instance of TE relationship. This instance can be assigned an identified name called “TE0001”. It has two properties: incidentThreat and targetEntity, in which incidentThreat has the value of “duplication of user smart card” and targetEntity has the value of “user smart card”.

Templates of generic security patterns such as TE, TE-TE, TE-C and TE-C-TE, indicating the required attributes of these relationship classes (as the middle part of Fig.2), should be provided to security experts to create template instances of security patterns (as the right part of Fig.2). The value of the required attributes should be selected from the concepts represented in the data skeleton, which stores basic concepts of threat, entity and countermeasures.

Template instances of generic security patterns can help generate data requirements for data integration mechanism. Instances of concepts of threat, entity and countermeasures, could be created through data integration mechanism that will be discussed later. They are known as common instances of security patterns in our study.

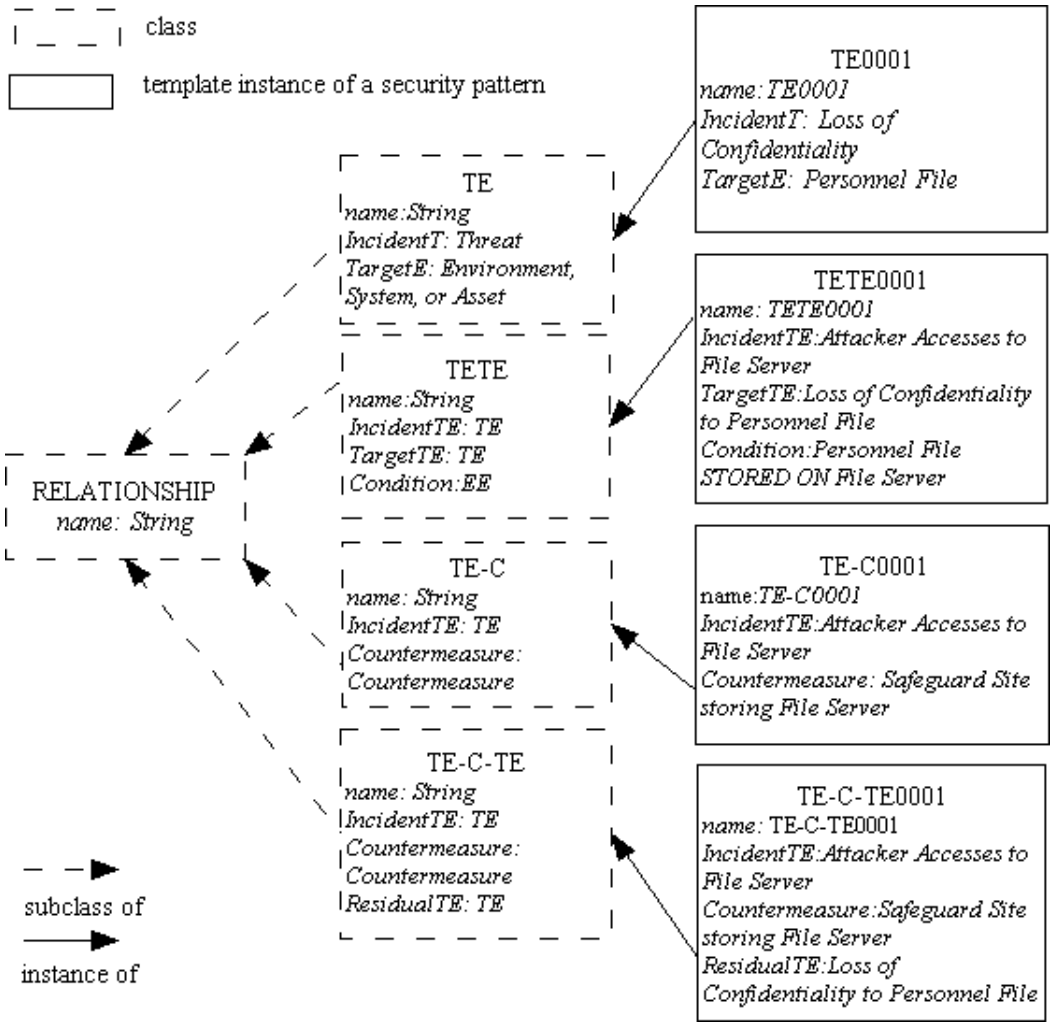

Fig. 2. Template Instances of Security Patterns

Ontology Representation of the Components

A frame-based ontology is constructed to represent the components in the knowledge model. The frames refer to basic concepts in the ontology, e.g. classes, slots or instances. This ontology consists of a set of classes organized in a taxonomic hierarchy. We also define slots for these classes. Instances of classes are defined by providing specific value and additional restrictions in these slots. For example, a template instance of security pattern "TE" could be represented as an instance of TE classes in the ontology.

After a top-down development process, a classification scheme of the ontology is formed (Table 1). The ontology currently contains 89 classes, including classes from the top level to the bottom level.

Table 1. The Classification Schema of the Ontology

Top Level Class	Middel Level Class	Bottom Level Class

System	Hardware		PC, Server, Workstation, Hub, Router, Switchboard, Gateway	
	Software	System Software	Loader, Operating System, Device Driver, Programming tools	
		Application Software	...	
	Platform		Physical Platform	
	Network		Physical Network, Virtual Network	
	User			
Asset	Data			
Environment	Location, Site, Building, Floor, Passage		Room	
	Service		Power Supply, Communication Link, ...	
Security	Threat	Threat Type	Environmental Threat	Fire, Flood, ...
			Personnel Threat	Insider Attack, ...
			Network Threat	Intrusion Attack, ...
			Physical Threat	Equipment Failure
				Equipement/Data Theft
		Threat Tree		
	Defence	Countermeasure		
		Threat Countermeasure		
Relationship	Threat-Entity Relationship (TE)			
	ThreatEntity-ThreatEntity Relationship (TETE)			
	ThreatEntity-Countermeasure Relationship (TE-C)			
	IncidentTETE-Countermeasure-ResidualTETE Relationship (TETE-C-TETE)			
	Entity-Entity Relationship (EE)		Asset-Platform Relationship	
			Asset-Application Relationship	
			...	

The use of ontology to represent the components in the knowledge model aims to help maintain a global schema for the data integration mechanism which will be discussed later.

DATA INTEGRATION MECHANISM

Distributed Security Related Data

Security related data are entities regarding the physical and logical aspects of security. Entities of Units, Physical Networks, and Physical Platforms are correlated to the physical aspects of security; and entities of Information Assets, Application Systems, and Virtual Networks are related to the logical aspects of security.

An internal structure of security entities is shown in Figure 3. It describes the inter-relationships between the entities described above. Information Assets are resided in the top end of the structure, which are processed by Application Systems. Virtual Networks host Application Systems and they themselves are hosted by Physical

Networks. Interconnected Units form the Physical Networks, which are located in Physical Platforms. Physical Platforms are located in a Physical Environment. Details of the inter-relationships are discussed in [Kwok et. al, 2001].

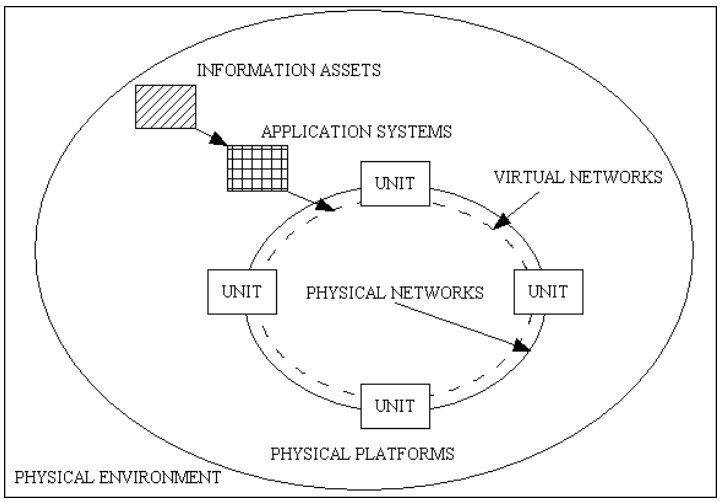

Fig. 3. Security Related Data

Data about these entities can be recorded and are supposed to distribute in different departments of an organization. These data provide security officers the origins of security related data that are required for our data integration mechanism.

In fact, security related data are commonly embedded in documentation intended for another purpose other than documentation intended primarily for security purposes. In our study, this documentation refers to the records in rational database.

Specific Security Data Repository

The Specific Security Data Repository in the framework (Fig. 1) stores the results of data integration mechanism, i.e. instances of concepts as mentioned above. The concepts are the values of attributes in the template instance of security patterns. In this paper, instances of concepts are named as common instances of security patterns.

Security related data change quickly due to the rapidly changing risk environment. If we can dynamically and continuously integrate data satisfying data requirement from the dispersed data sources in an organization into the Specific Security Data Repository, accurate and updated data can be supplied for risk assessment activities.

Overview of Data Integration Mechanism

Previous studies in using ontologies in data integration suggest that it is an efficient solution to solve the problem [Isabel et. al, 2005]. We, thus, propose a similar data integration mechanism for our framework (Fig.4), which aims to integrate the distributed security related data into the Specific Security Data Repository, in which:

- a global schema is established as a global ontology for the knowledge model. Local ontologies of operational data are also constructed to represent the local data sources; and
- a structured matching between the global schema and source schemas is addressed.

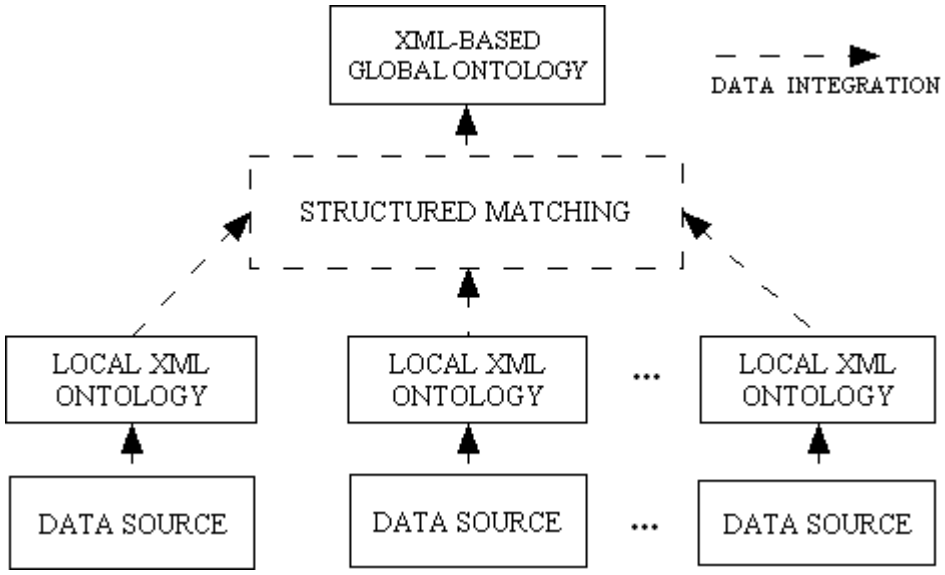

Fig. 4. The Architecture for Data Integration Mechanism

The Global Schema

The global schema in the mechanism is an XML-based representation for the ontology. This global XML format serves our data integration solution a central normalized data structure. For example, the following XML schema defines the frames in the ontology:

```
<class>
    <superclass>System</ superclass>
    <name>Hardware</name>
    <slot>Description</slot>
    <slot>Vulnerability</slot>
    <slot>Date</slot>
    <slot>Replacement Cost</slot>
    <slot>Hardware Maintenance Manual</slot>
    <slot>Environmental Impact</slot>
</class>
<templateinstance>
    <ofclass>EE</ofclass> (EE stands for Entity-Entity Relationship)
    <name>Personnel File Stored On File Server</name>
    <incident>Personnel File</incident>
    <target>File Server</target>
    <link>Stored On</link>
</templateinstance>
…
```

Structured Matching between Global Schema and Source Schema

A structured mapping process between the global schema and the source schemas is proposed. Local ontology is needed to be built before the mapping. XML schema is used to represent both the global ontology and the local ontology in our approach (Fig.5).

Template instances of generic security patterns are defined in the global XML-based schema. Converters of the global schema can generate a *Data Requirement Table*. Each column in this table may map to the actual data item in the local data source through the local XML-based schema.

For instance, a Data Requirement Table with three columns including a threat ("Attacker"), its direct target entity ("File Server"), and its indirect target entity ("Personnel File") can be derived from the template instances in the XML-based global schema as shown in Figure 5.

After the Data Requirement Table is established, a semantic mapping between the elements in local XML-based schemas and those in global schema is needed for filling the actual data item in local data source into each column of the Data Requirement Table. The matched elements should have the same semantic meaning to the Data Requirement Table. For example, an element "attacker" in source schema of Network center can be matched to the node "incidentT" in global schema that has a value of "Attacker" (Fig.5).

An automatic semantic mapping can borrow the concept of Learning Source Descriptions system (LSD), which uses machine learning to (semi-)automatically compute the semantic mappings between the local data source schema and the global schema. Details of this mapping process can be found in [AnHai et. al, 2000]. However, manually mapping process is also applicable for filling the Data Requirement Table.

After the integration, common instances of security patterns can be created using the data items in the Data Requirement Table.

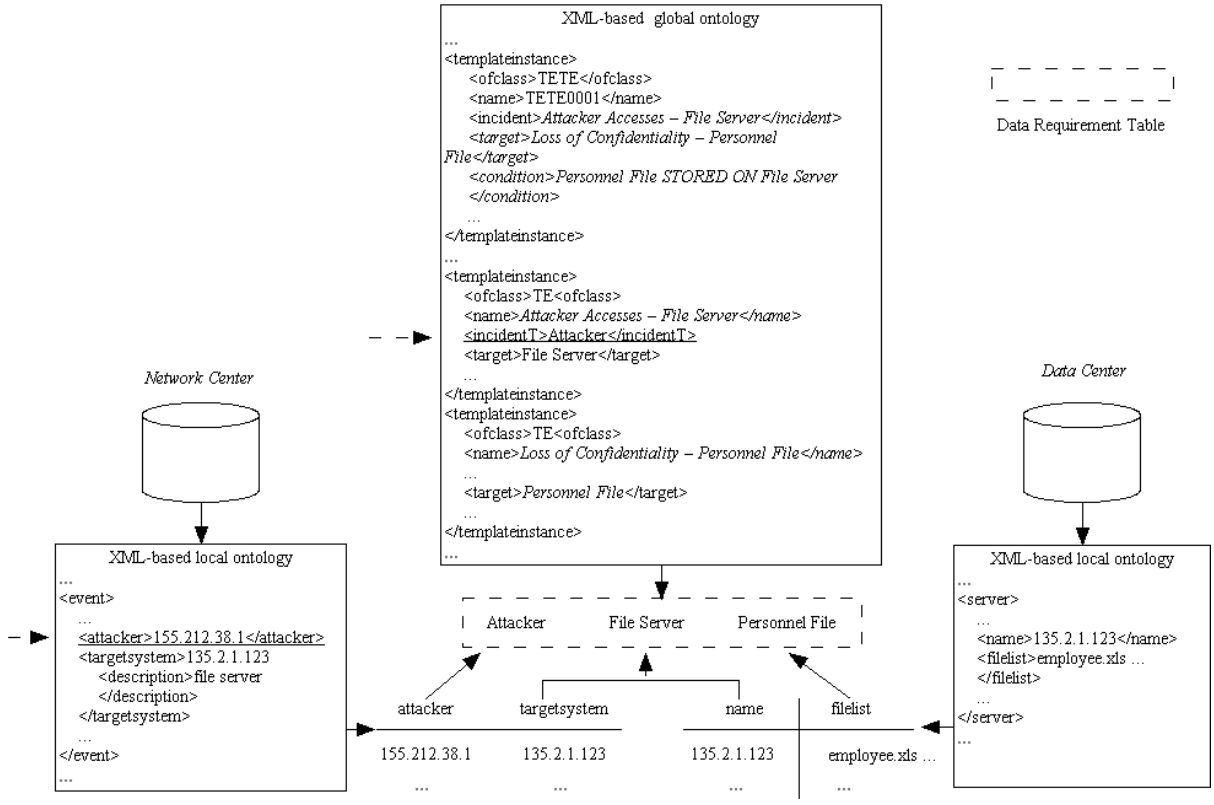

Fig. 5. Source Network Center and Data Center returns data that is in their local schemas, which then need to be mapped to the global schema during data integration

APPLICATION OF THE KNOWLEDGE FRAMEWORK

After security data are integrated into the Specific Security Data Repository through the data integration mechanism, they can be applied in several risk assessment activities. One of the activities is performing risk simulation.

Risk simulation could help security officers to obtain an entire set of impacted entities specific to an organization if a threat happens. This simulation includes the following tasks: (1) identify threats to an organization; (2) demonstrate the outcome of a threat acting on one entity or multiple entities. These tasks raise several data requirements for the data collection process, which includes preparation of threat profile, aggregation of impacted entities.

Common instances of TE relationship and TETE relationship stored in the Specific Security Data Repository provide the required security data for risk simulation which can be consumed by applications to visualize the simulation results.

CONCLUSION

The knowledge framework presented in this paper provides a methodology to guide the data collection process for risk assessment activity. It has several benefits, which mainly includes: (1) security knowledge are maintained in a formal way for periodical information security modelling process; (2) updated security data specific to an organization could be obtained under the guidance of security knowledge maintained in our framework.

Prototype of the framework is under development to test the concepts of applying ontology as a global schema for the data integration mechanism and structured matching between the global schema and source schemas. The report of such a prototype is due shortly.

REFERENCES

Anderson A, Kwok L F and Longley D. Security Modelling for Organizaitons. Proceedings of 2nd ACM Conf on Computer and Communication Security. Fairfax Virginia, USA, 2-4 November, 1994, pp.241-250.

AnHai Doan, Pedro Domingos, and Alon Y. Levy. Learning source description for data integration. In WebDB (Informal Proceedings), pages 81--86, 2000.

Benjamins, R., Fensel, D. and Gomez Perez A. Knowledge Management through Ontologies. In U. Reimer (editor), Proceedings of the Second International Conference on Practical Aspects of Knowledge Management. 29-30 October, 1998, Basel, Switzerland.

Caelli W, Longley D, and Tickle A.B. A Methodology for Describing Information and Physical Security Architectures. IT Security: The Need for International Cooperation, Proc. IFIP TC11 8th Int. Conf. on Information Security, IFIP Sec'92, Singapore, 27-29 May 1992, pp.277-296.

Fung P, Kwok L F, and Longley D. Electronic Information Security Documentation in (Eds. Johnson C, Montague P, Steketee C) ACSW Frontiers 2003, Australasian Information Security Workshop (AISW2003), February 2003, Adelaide, Australia, pp25-31.

Isabel F. Cruz and Huiyong Xiao. The Role of Ontologies in Data Integration. Journal of Engineering Intelligent Systems: 13(4), December, 2005.

Kwok L F. A Hypertext Information Security Model for Organizations, Information Management and Computer Security, Vol. 5, No. 4, 1997, pp.138-148.

Kwok L F and Longley D. Information Security Management and Modeling, Information Management and Computer Security, Vol. 7, No. 1, 1999, pp.30-39.

Kwok L F, Fung P P K and Longley D. Security Documentation in (Eds. Eloff JHP, Labuschagne L, von Solms R and Dhillon G) Advances in Information Security Management & Small Systems Security, IFIP TC11.1/WG11.2 8th Annual Working Conf. On Information Security Management & Small Systems Security, 27-28 September 2001, Las Vegas, USA, pp127-140.

Kwok LF and Longley D. Security Modelling for Risk Analysis. Proc. 18th IFIP World Computer Congress, IFIP 2004, 22-27 August 2004, Toulouse, France, pp29-45.

Landoll, Douglas J. The Security Risk Assessment Handbook: a Complete Guide for Performing Security Risk Assessments. Boca Raton, FL: Auerbach Publications, 2006.

COPYRIGHT