A risk recommendation approach for information security risk assessment

Ya-Chi Chu, Yu-Chih Wei, and Wen-Hsuan Chang,
Telecommunication Laboratories, Chunghwa Telecom Co., Ltd, Taoyuan, Taiwan, R.O.C
Email: { gyh2211, vickrey, eifie } @cht.com.tw

***Abstract*—Nowadays, information security becomes a critical issue on protecting the benefits of business operation. Many organizations introduce security risk management to ensure the security of business processes. However, in the processes of risk assessment, it is difficult and time-consuming to identify the threats and vulnerabilities for each asset. Furthermore, if the identified results diverged from the real situation, the organization may implement unnecessary controls to prevent the non-existing risk. In order to resolve these problems, we adopt data mining approach to find the relationship between asset and threat-vulnerability. And then, we propose a recommendation scheme for assisting user identifying threat and vulnerability. The experiment result shows that our recommendation mechanism can improve the efficiency and accuracy of the risk assessment.**

***Index Terms*—Threat, Vulnerability, Recommendation, Data mining.**

I. INTRODUCTION

In recent years, companies much more rely on information technology to achieve their business goals. By the rapid development of information technology, making them vulnerable to IT security incidents. According to ISO 27005 [4], giving a definition of information security risk: potential that a given threat will exploit vulnerabilities of an asset or the group of assets and thereby impact the organization. So if the company expects the success of adopting risk management, finding the appropriate threat-vulnerability pair of each asset is the key point. However, in the process of identify threat-vulnerability; it is difficult to recognize the feasible combination, especially for the people who lacks of security concept. Without recommendation mechanism, asset owner may face as least three challenges: first, in spite of the threat and vulnerability lists are provided as a candidate list for owners, it is still time-consuming for them to choose the suitable one from more than hundreds items. Second, the threat-vulnerability pairs do not matched frequently. For example, application server equipments may have a vulnerability which may lack of physical protection. Theoretically, environmental damage and physical breakage are matched threats, however, user may choose another unmatched threat such as power outage. Third, not all the users have the ability to find the security issue for the asset, and finally choose non-exist threat-vulnerability. Non-exist threat-vulnerability pairs make organization spend unnecessary cost on preventing the risk happened. These vulnerability identification errors may mislead the manager to neglect the real weaknesses, or invest in improper security measures.

Due to the deficiencies mentioned above, we propose a mechanism to resolve this problem. After user identifies assets, in this research, different asset categories have been classified. By the use of Predictive Apriori to mining the common threat-vulnerability pairs for each category, the recommendation list is prepared. User can choose the pair correspondence to the real circumstance from the recommendation list. The main contribution of our proposed approach is to improve the efficiency and accuracy on identifying the threat-vulnerability pairs and the recommended items are suitable for all kind of industries.

The remainder of the paper is organized as follows: Section 2 describes relevant background on data mining approach and risk identification problems in the past. Section 3 presents our research model which recommends vulnerabilities-threat pair for different categories of asset. Section 4 contains experimental design and results. Conclusions and future directions are given in Section 5.

II. RELATED WORK

ISO/IEC 27001, ISO/IEC 27005 and other standards not only form the basis of any domain-specific information security (IS) risk assessment standard, but also developed risk assessment approaches. In these standards, it may not explicit suggestion the potential treat and vulnerability for each asset. And in risk assessment phase, treat and vulnerability must be identified by brainstorming, questionnaire or other technical tools. This may take too much time and not intuitive for users.

Other Existing risk management mechanism such as CORAS[5] and OCTAVE[2], also propose their own methods in risk assessment based on standards. In CORAS, a Platform for Risk Analysis of Security Critical Systems is proposed, it uses threat and vulnerability modeling alongside with threat diagrams and structured brainstorming to identify risks [6]. These approaches suggest some common security principles or security best practices. However, they do not determine and evaluate specific security needs of assets to identify their risks. The other approaches, such as OCTAVE, determine the criticality and impact of vulnerabilities from the review of security requirement. It considers the possible condition or situations that can threaten an organization's information assets by existing security checklists, standards or brainstorming [1]. Although OCTAVE uses security requirements, they determine

only the impact of vulnerabilities by security requirements or only the requirements of a product; or using asset unspecific security standards.

In addition to the above we mentioned, there are some models use different ways to identify risks. AURUM is an ontology-based method [3]. It supports the decision maker to answer the following questions: Which threats threaten critical assets? Which threat is a multiplier? Which vulnerabilities have to be exploited by a threat to become effective? And it shows the potential threat of selected asset by threat tree. For each threat highly granular vulnerabilities, which a threat could exploit, have been modeled in the ontology. All the threats AURUM provided are recommended from standard, and may lack of flexibility to adapt the new information technology. Furthermore, it is complicated for users to build new ontology depend on themselves.

III. THREAT-VULNERABILITITY PAIRS RECOMMENDATION MECHANISM

In this paper, we will explain the processes to find the threat-vulnerability pairs for each asset in the remainder parts.

A. Data source

In this paper, the original data is collected from six different kinds of business units (network management, billing and information management) in the same telecommunications organization. Each of the business unit has been certified compliant with ISO/IEC 27001:2005. In order to collect enough and real data which selected by users, we develop an online risk management system without any recommend list. When users identify the risk, they can choose the threat-vulnerability from the hundred items. All the threats and vulnerabilities of the system provide are extended from ISO/IEC 27005:2011. After this step finished, system will collect many assets, and each asset has several threat-vulnerability pairs.

B. Classify asset

In ISO/IEC 27005:2011, assets can be classified into number of categories, such as hardware, software, people, etc. Therefore, we classify the asset into five categories base on the standard: hardware, software, people, information and service. When getting original data from the online risk management system, we cluster the assets into five categories. Then, for each category we classify several groups which function is similar. For example, windows XP and Microsoft office are both belong to software category. But, they play different role in system operation. Windows XP is a kind of operation system software, and Microsoft office is a kind of package software. Because they have different function, we establish operation system group and package software group. Windows XP and Microsoft office belong to two groups respectively. In this paper, we create many groups in each category, as shown in Table 1.

C. Mining the data by Weka

After we classify each asset into different groups, each group will contain many assets and threat-vulnerability pairs which chosen by users. Then we use Weka to mining each group to find the associate relationship between threats and vulnerabilities. Weka provides several data mining algorisms, such as Apriori, Predict Aprori and Tertius. In Weka, it only supports ARFF file format, so before mining association rule, we must translate the original data from csv into arff.

First, we arrange the data of each group, and translate the file to ARFF form. Second, users can set the parameters provide by Weka GUI to find the results they want. Because Predictive Apriori algorithm aims to discovery of *n* most predictive association rules, we adopt the default settings. Finally, it may output the top *n* relevant rule in Weka.

TABLE 1 EXAPLE OF ASSET CATEGORY

Hardware	**Software**	**Information**	**People**	**Service**
Web server	Operation system	SOP	System manager	Electric service
Application server	Application program	Installation Manual	Network manager	Air conditioning
Database server	Application system	User manual	Information security manager	Network management
Log server	Development tool	Operation manual	Help desk	MIS
Personal computer	Package	Planning	Operate person	Email service
Notebook	Network management tool	System design document	Database manager	VPN service
Network equipment	Compression tool	Test report	System analyst	LDAP service
End point equipment for system operation	Audit tool	Contract	System quality manager	Telephone communications
Printer	Analysis tool	Confidential consent	End point user	System maintenance service
Scanner	Statistics tool	Audit report	Auditor	VPN
Storage equipment	Execution file	System log file	File manager	Equipment maintenance Service
Server room	Utility program	Parameter profile file	Supplier	LDAP
Office	Execution files	Database	Safeguard	Authentication service
Control area	Self-developed software	Source code	Administrative staff	System operation

IV. EXPERIMENT

In order to evaluate the experiment result of the threat-vulnerability pairs, in this paper, we invite two risk assessment experts to determine the recommendation list whether actually improves the efficiency and accuracy in identifying threat-vulnerability. One of the two experts is major in risk management, and has many experiences in consulting the organization staff about risk assessment approach and get ISO/IEC 27001 certification. Another one is a professional auditor in ISO/IEC 27001, who with a great deal of maturity. From the results of the experiment, we take two examples to proof the proposed mechanism is effective and explained in rest parts.

Log file is an import clue in system operation. It records

significant information when users execute some commands and even they access files or objects. So keep the integrity and accuracy of log file is a crucial planning for the organization. Table 2 list top 5 threat-vulnerability pairs of log file.

Vul and threat are two attributes used in the experiment. They represent vulnerability and threat respectively. WEKA will count each vulnerability and threat appear in our data source, and shows in each rule. Let each association rule express vul ==> threat. Every rule has an accuracy value which represent like acc:(0.98954). The accuracy value calculates between 0 to 1 and closer 1 represents the better result.

When we concern security issue of the sensitive system, it is crucial to detect unauthorized person accessing. Therefore, managers should set the parameter to log users' behavior and view the records regularly. Then, according to log files' importance, it is necessary to protect the log file so that it may not be tampered, loosed or damaged. Our recommendations shown in Table 2, which list 5 common threat-vulnerability pairs might happen with high probability. For example, log files may have vulnerability that administrator set insufficient parameters to log users' behavior or don't check the log files regular. Then, the threat like unauthorized access may utilize this vulnerability to access to system programs or data and the managers can't detect in time. In this way, it may bring out very huge loss to an organization. The other vulnerabilities list in Table 2 such as lacking of a movable device/media control or lacking of suitable physical protection should be pay close attention by organizes.

TABLE 2 RECOMMENDATION LIST FOR LOG FILE

Log file
Rule 1： vul= Lack of settings and inspection the system log 13 ==> threat= Unauthorized access to system programs or data 12 acc:(0.86677)
Rule 2： vul= Lack of appropriate data access control 2 ==> threat= Malicious damage physical devices or information 2 acc:(0.74021)
Rule 3： vul= Lack of a movable device and media control 2 ==> threat= Leakage of confidential data, or programs 2 acc:(0.74021)
Rule 4： vul= Lack of suitable physical protection 2 ==> threat= Leakage of confidential data, or programs 2 acc:(0.74021)
Rule 5： vul= lacking or insufficient printout information security management mechanism 2 ==> threat= Leakage of confidential data, or programs 2 acc:(0.74021)
Rule 6： threat= Natural damage 8 ==> vul= Lack of proper data backup mechanism 6 acc:(0.6994)
Rule 7： vul= Lack of or insufficient retain and inspect operation log file 20 ==> threat= Tampering with the data or program 14 acc:(0.68182)
Rule 8： threat= Operation fail 8 ==> vul= Lack of proper data backup mechanism 6 acc:(0.6994)
Rule 9： vul= Important information fails to properly stored offsite or implementing off-site backup 4 ==> threat= Tampering with the data or program 3 acc:(0.65862)
Rule 10： threat= Tampering with the data or program 21 ==> vul= Lacking or insufficient retain and inspect operation log file 14 acc:(0.65217)

As we mentioned above, the experiment result for log files recommend the most common threat-vulnerability pairs in real situation. In addition to this, our recommendation lists comprehensive consider all kind of risks which may happen. Not only physical issue, but also technical and artificial. Therefore, from the point of view two domain experts, our mechanism is really helpful for risk assessment and satisfies the requirements of the information security risk management system ISO/IEC 27005:2011.

V. CONCLUSION AND FUTURE WORK

In this paper, we propose a mechanism that can recommend users the suitable threat-vulnerability pairs when they identify risk for each information asset, especially the user who has insufficient security knowledge. From the result of experiment, our recommendation list can actually help user to filter the appropriate risk item no matter they are common or specific usage. It not only improves the efficiency, but also the accuracy in the process of risk assessment.

In the future, our recommendation mechanism can easily been expanded to any specific field when organization has other new architecture for security control. The more data we collected, this model will more complete. The algorithm of finding association rule used here can be refined and extended, so that to improve performance and scalability. Eventually, we can spread this recommendation mechanism being incorporated into our risk management system for supporting business in information security planning. Finally, much more researches in general needs to be done to assist organization for protecting their assets from harm within acceptable cost.

REFERENCES

[1] Alberts, C., Dorofee, A., Stevens, J. and Woody, C. (2003), 'Introduction to the OCTAVE approach', Software Engineering Institute (SEI), Carnegie Mellon University, Pittsburgh, USA.

[2] ASNZ (2009), 'Australian/New Zealand Standard Risk Management AS/NZS ISO 31000:2009 - Risk Management – Principles and Guideline'.

[3] Ekelhart, A., Fenz, S. and Neubauer, T. "Ontology-based Decision Support for Information Security Risk Management", 'International Conference on Systems, 2009. ICONS 2009. IEEE Computer Society, 2009, pp. 80-85.

[4] ISO (2011), 'ISO 27005:2011 Information technology - Security techniques - Information security risk management', International Organization of Standardization (ISO).

[5] Scheffer, T. (2001). Finding association rules that trade support optimally againstvconfidence. Proceedings of the European Conference on Principles and Practiceof Knowledge Discovery in Databases.

[6] Stølen, K., den Braber, F., Dimitrakos, T., Fredriksen, R., Gran, B. A., Houmb, S.-H., Lund, M. S., Stamatiou, Y. C. and Øyvind Aagedal, J. (2002), 'Model-based risk assessment – the CORAS approach', in 'NIK (2002) informatics conference, Kongsberg'.