Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Environmental Sciences 10 (2011) 1029 – 1034

Procedia Environmental Sciences

2011 3rd International Conference on Environmental Science and Information Application Technology (ESIAT 2011)

Multi-sensor Data Fusion for Cyber Security Situation Awareness

Yan Zhang[a,b], Shuguang Huang[a], Shize Guo[b], Junmao Zhu[b,a]*

[a]*Electronic Engineering Institute, Hefei 230037, China*
[b]*Institute of North Electronic Equipment, Beijing 100083, China*

Abstract

To analyze the influence of security incidents on a networked system and accurately evaluate system security, this paper proposes a novel cyber security situation assessment model, based on multi-heterogeneous sensors. By using D-S evidence theory, we fuse security data submitted from multi-sensors, according to the network topology and the importance of services and hosts. Moreover, we adopt the evaluation policy that from bottom to top and from local to global in this model. The evaluation of a simulated network indicates that the proposed approach is suitable for network environment, and the evaluation results are precise and efficient.

Keywords: Sensor; Security Data; Assessment; Hierarchical Method; Security Situation

1. Introduction

With the development of computer network, information security deteriorates rapidly. To cope with different types of network attacks, people often use different types of security devices (sensor). However, the management of these devices encounters many problems, including Alert Overload, Alert Conflict and high False Positive etc.

To solve these problems, many studies have tried to apply situation awareness to information security, e.g.: Bass [1] proposed multi-sensor data fusion architecture, Wang et al. [2] proposed the use of neural network in multi-heterogeneous sensor fusion. Wei [3] and Zhang [4] improved the framework of

* Corresponding author. Tel.: +0-86-0551-5767838; fax: +0-86-0551-5766654.
E-mail address: eeijack@163.com.

1878-0296 © 2011 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of Conference ESIAT2011 Organization Committee.
doi:10.1016/j.proenv.2011.09.165

situation awareness by joining in some environment factors (host number, host services, service attacks, etc.), and Lai [5] used a simple weighted and grey theory to implement security situation awareness.

In this paper, we use DS Evidence Theory to fuse alert that submitted from heterogeneous network sensors. We perform an experiment upon a simulated network environment. The results show that the proposed method not only provides the security situation in the macro system, but also provides three different levels of assessment of the security situation.

2. Related Work

2.1. Network security situation awareness

Briefly speaking, situation awareness is to know what is happening and how to respond. Endsley [6] viewed situation awareness as three levels: Perception, Comprehension and Projection. The first level perception is to collect data from different sources. The second level comprehension is to integrate and understand these data. And the third level projection is to predict what will occur within a period of time.

In order to create efficient network security situation awareness, Lai [5] proposed a Network Security Situation Awareness (referred to as NSSA) model. Inspired by Bass [1], Liu et al. [7] proposed an information fusion model for network security situation awareness.

According to these studies, current network security situation awareness only provides macro information, such as: What kind of network is being attacked (Probe, R2L, U2R, DoS ...). This can not help policy-makers to take prompt and effective response. To solve this problem, we introduced the concept of risk assessment, which can identify the most weakness point.

2.2. DS Evidence Theory

DS evidence theory is proposed by Shafer in 1976. It is used to describe different levels of accuracy and often applied to medical diagnostics, risk analysis and decision analysis [8].

Before using the DS fusion rule, the first step is to define the target framework. And then use BPA (Basic Probability Assignment) formula to allocate confidence to different sensors.

Suppose there are two IDS: O_1 and O_2, and an attack incident: H.

In O_1, confidence level $m_1(H)$ represents the probability that O_1 support the occurrence of H. In O_2, confidence level $m_2(H)$ represents the probability that O_2 support the occurrence of H. Through DS rules, the fusion result of O_1 and O_2 evidence is as follows:

$$m_{12}(H) = \frac{\sum_{B \cap C = H} m_1(B) m_2(C)}{\sum_{B \cap C \neq \phi} m_1(B) m_2(C)} \tag{1}$$

After the fusion of the two evidences, $m_{12}(H)$ is the final probability that the alert may occur. A number of studies have used this method to lower the False Positive Rate [7] [9].

3. Proposed Method

In this section, we proposed a Hierarchical Network Security Situation Assessment Model (referred HNSSAM) (see Figure 1). This model joins the DS evidence theory fusion rules with hierarchical quantitative risk assessment method, and makes use of confidence level, service importance and host importance. The advantages of this model are: a) to solve the problem of mass data processing; b) to provide three levels of intuitive security threat; b) to quickly find weaknesses in the system or the security situation.

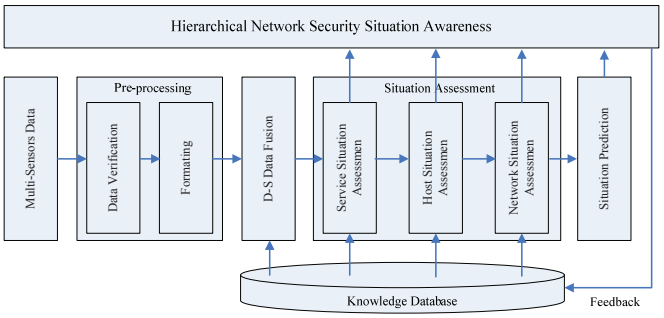

Fig. 1. Hierarchical framework for network security situation awareness

3.1. Pre-processing

Data preprocessing module is designed to collect security data from different sensors. Data verification mechanism is adopted to determine whether there is a successful attack. By comparing the conditions and the configuration information (e.g.: OS version, services running, etc.) necessary for a successful attack, we could simply remove non-impact attack alert. For example: IDS detected a large number of serv-u directory traversal attacks which aim at serv-u software running on Windows system. However, the target host is running on Linux system, so attack can not be succeeded. Therefore, these invalid alerts should be removed to reduce the number of alerts. Finally, the security data will be converted into a uniform format so as to meet the HNSSAM architecture.

3.2. D-S Data Fusion

According to the basic definition of DS, we set the target framework $\Theta = \{True\_Positive, False\_Positive\}$. Because the alerts generated by security equipment clearly have two possibilities: (1) True Positive; (2) False Positive, we define the confidence values of an alert as True Positive Rate (TPR): m (correct alerts) = TPR. We obtain TPR by supervised training of security devices in various attacks. Then the confidence values will be stored in Knowledge Base for further use. The process of fusion is illustrated in Figure 2.

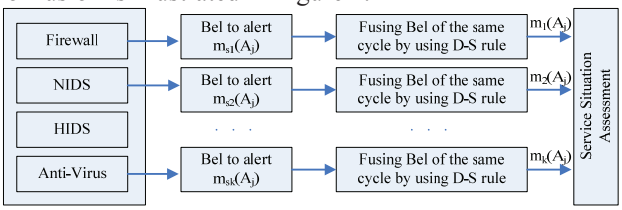

Fig. 2. D-S rule-based alert fusion model

3.3. Hierarchical Quantitative Situation Assessment

For narrative convenience, we define the following concepts:

Definition 1. A: Attack: The activities that trigger FW, IDS, Anti-Virus and other security devices to generate an alert. After pre-processing and DS data fusion, attack can be expressed as $A = \{ATK\_TYPE, ATK\_BEL, SEVERITY, TIME, DIP\}$, where ATK_TYPE is representatives of attack

type, ATK_BEL is representatives of the confidence level, $SEVERITY$ is representatives of threat index obtained from the Knowledge Base, $TIME$ is representatives of the time attack taken place, and DIP is representatives of the target host.

Definition 2. Service Security Situation: The insecurity degree of a service after been attacked.

Definition 3. Host Security Situation: The impact of a number of insecurity services on a host.

Definition 4. Network Security Situation: The impact of a number of insecurity hosts on a network.

We first evaluate how serious the services provided by the host are under attack. Note that, the impact of attack on services is not only related to the level of threat but also related to the network traffic of user activities. Beside, the impact of attack varies with different period of time [10]. So we proposed the service security situation assessment formula as follows:

$$SS_{KS_i}(t) = W_T(t)\left(\sum A_j 10^{B_j}\right) \tag{2}$$

In formula (2), S_i is representative of service that is under attack. W_T is representative of the weight of time. We divide one day into three continuous sections: $\{t_1, t_2, t_3\}$. Based on the statistical results, network administrators can set each section a separated traffic description: low, medium, and high. The corresponding quantitative values are 1, 2, and 3. Followed by normalization, we obtain the weight of each time period W_{T_i}.

$$W_{T_i} = T_i \Big/ \sum T_i \tag{3}$$

A_j is representative of the value of confidence after alert fusion, $j = \{FTP\_ATK\_1, FTP\_ATK\_2, HTTP\_ATK\_1, ...\}$. B_j is representative of the severity of the attack. The higher the value of $SS_{KS_i}(t)$ is, the higher the level of service threats at time t is.

Then the host security situation is evaluated. The security situation of host is affected by the services and the security mechanisms [4]. The assessment formula is designed as follows:

$$SS_{H_k}(t) = \sum W_{kS_i} SS_{kS_i}(t) \Big/ \sum W_{SA_p}\left(\sum SP_{kS_{pq}}\right) \tag{4}$$

$SS_{H_k}(t)$ is representatives of the security situation of host H_k at the time t, $H_k = \{Host\_A, Host\_B, Host\_C, ...\}$. W_{kS_i} is representative of the importance weight of the service S_i on host H_k. SS_{kS_i} is representatives of the security index of the service S_i on host H_k. W_{SA_p} is representatives of the importance weight of the security standards on host H_k, $SA_p = \{Confidentiality, Availability, Integrity, Authentication, Non\_Repudiation\}$. $SP_{kS_{pq}}$ is representatives of the level of influence of security mechanism q on security standard p, $q = \{encrypting, digital\_signaturing, access\_controlling...\}$, $p = SA_p$. The greater the value of $SS_{H_k}(t)$ is, the higher the level of the threat to the host H_k at time t is. Therefore, the network security situation assessment formula is designed as follows:

$$SS_N(t) = \sum W_{H_k} SS_{H_k}(t) \tag{5}$$

W_{H_k} is representatives of the importance weight of host k in the network. The larger the value of $SS_N(t)$ is, the higher threat level of network at time t is.

4. Experimental Results and Analysis

To test the effect of this HNSSAM model, we simulated a multi-sensor network environment (see Figure 3). In this simulated network environment, we deployed four different sensors, the firewall at Internet entry, network intrusion detection system, the host intrusion detection system and anti-virus software installed on the hosts.

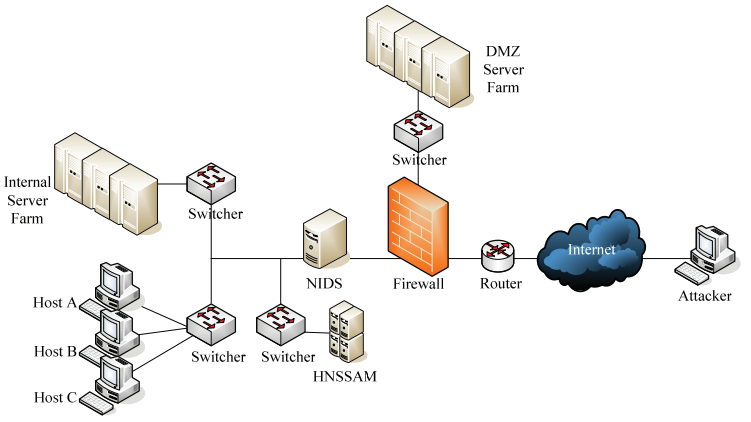

Fig. 3. Experimental network topology

Section 1. We divided one day into three periods: t_1 = 00:00 ~ 08:00, t_2 = 08:00 ~ 18:00, t_3 = 18:00 ~ 24:00. Each period is assigned with a different importance level. Observed time T1 and T6 fall in period t_3, T2 ~ T3 fall in period t_1, T4 ~ T5 fall in period t_2. From T1 to T6, collection the information that FW, NIDS, HIDS, Anti-Virus detect attack on Host A, B, C.

Section 2. According to the security data collected from section 1, we look up the confidence level corresponding to the security data in the Knowledge Base. By DS fusion rule, we fuse the confidence value. Then the results are multiplied by the severity value of attacks in the knowledge base. Finally, the service security situation value is obtained by equ. 2.

Section 3. In accordance with the analysis of section 2, we draw the results in Figure 4 a). We could clearly see in Figure 4 a) that the RPC services on the Host A suffer higher level of threats, which should be deal with firstly. According to equ. 4, we obtain the host security situation, as shown in Figure 4 b). It can be clearly seen that the attacks are active during time T1 to T4.

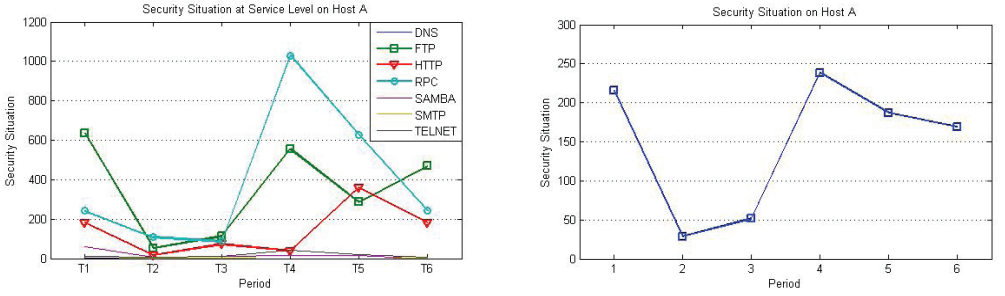

Fig. 4. (a) Security situation of all services on host A; (b) Security situation of host A

Section 4. In accordance with the analysis of the results generated in section 3, it is easy to draw the security situation of all hosts in Figure 5 a). From this figure, administrators could find out the threat level on each host. According to equ.. 5, we obtain the network security situation. The results are shown in Figure 5 b). We could find out through this figure that this network is suffering more attacks from afternoon to midnight.

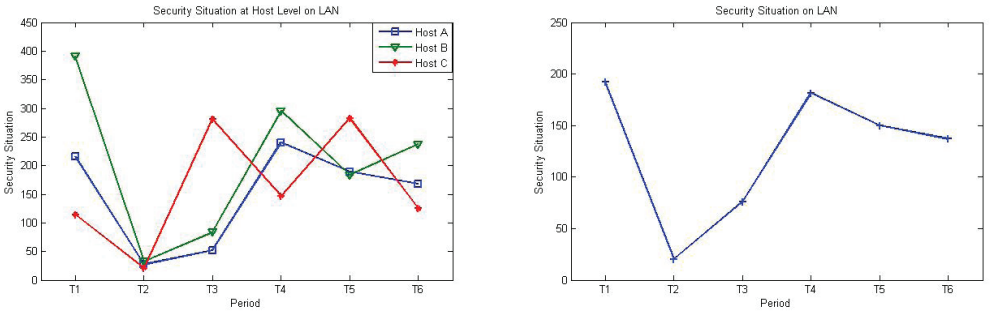

Fig. 5. (a) Security situation of all hosts on LAN; (b) Security situation of LAN

According to the analysis above, it can be seen that HNSSAM framework provides three levels of cyber security situation. This method overcomes the shortcomings of current hierarchical situation awareness systems. It can also assist decision-makers to adjust policy to enhance security.

5. Conclusions

In this paper, we analyzed the existing situation assessment algorithms, proposed a novel cyber security situation assessment model based on multi-heterogeneous sensors. According to the proposed model, we implemented a situation awareness system. The evaluation of a simulated network indicates that the approach is suitable for network environment, and the evaluation results are precise and efficient.

References

[1] Bass, T., "Intrusion Detection Systems and Multisensor Data Fusion," Communications of the ACM, Vol. 43, No. 4, April 2000.

[2] Wang Huiqiang, Lai Jibao, and Ying Liang, "Network Security Situation Awareness Based on Heterogeneous Multi-Sensor Data Fusion and Neural Network," Second International Multisymposium on Computer and Computational Sciences, 2007.

[3] Wei Yong, Lian Yi-Feng, A Network Security Situational Awareness Model Based on Log Audit and Performance Correction[J], Chinese Journal of Computers, 2009,(04)

[4] Zhang Yong; Tan Xiao-bin; Cui Xiao-lin; Xi Hong-sheng, Network Security Situation Awareness Approach Based on Markov Game Model[J], Journal of Software, 2011,(03)

[5] Lai Ji-bao; Wang Ying; Wang Hui-qiang Zheng Feng-bing Zhou Bing, Research on Network Security Situation Awareness System Architecture Based on Multi-source Heterogeneous Sensors[J], Computer Science, 2011,(03)

[6] Endsley, M., "Design and evaluation for situation awareness enhancement," In Proceedings of the Human Factors Society 32nd Annual Meeting, Human Factors Society, pp. 97-101, 1988.

[7] Liu Mixi, Yu Dongmei and Zhang Qiuyu et al., "Network Security Situation Assessment Based on Data Fusion," 2008 Workshop on Knowledge Discovery and Data Mining, 2008.

[8] Sentz, K. and Ferson, S., "Combination of Evidence in Dempster-Shafer Theory," SAND 2002-0835, Unlimited Release, April 2002.

[9] Mei Haibin and Gong Jian, "Intrusion Alert Correlation Based On D-S Evidence Theory," Communications and Networking in China, Second International Conference on IEEE, 2007.

[10] Chen XZ, Zheng QH and Guan XH et al., "Quantitative hierarchical threat evaluation model for network security," Journal of Software, 2006(04)