International Journal of Research

International Journal of Research (IJR)

e-ISSN: 2348-6848, p- ISSN: 2348-795X Volume 2, Issue 10, October 2015

Available at http://internationaljournalofresearch.org

Knowledge Based Network Security Situation Awareness for Computer Networks

Deepa T. Naik

M.E(Computer Sci. & Engg.) Pune, India

EMAIL: **deepamishrargh@gmail.com**

Abstract

Network security situation awareness is an evolving technique in the field of network security, which can effectively solve the problems, network systems are suffering from, such as worms, virus, network eavesdropping, sniffing etc. Traditional security controls cannot provide evaluation of threats as they work independently. The false positive and negative rates of these devices are too high. Therefore it is tedious to obtain the security state of the entire network. This paper proposes Network Security Situation Awareness (NSSA) based on Knowledge discovery. D-S evidence theory is applied as the fusion mechanism to analyze and fuse the security alert events gathered from various network security situation sensors. Network security situation is generated by extracting the frequent attack patterns based on knowledge discovery technique. The proposed system has improved results in terms of delay, energy consumption, packet delivery ratio and throughput as compared to without knowledge discovery technique.

Key Words: Security alert; False alarm; Situation awareness; Knowledge discovery

INTRODUCTION

Intrusion Detection Systems (IDS) generate large volume of alarms every day, maximum of which are false alarms that lead to blocking communication through incoming port connections [1]. These devices have pre-configured rules and they respond to network activities according to these rules. A mechanism to provide Network Security Situation Awareness (NSSA) is required to overcome the problems of the existing systems. Endsley defined Situation Awareness (SA) as “the perception of the elements in the environment within a volume of time and space; the comprehension of their meaning and the projection of their status in the near future” [2]. Four levels of SA have been proposed that includes perception, comprehension, projection and resolution. But SA in computer network security is in early stages. Bass was the first who introduced this concept into network security and proposed the network security perception frame based on multi-sensor data fusion [3-4].

The remainder of this paper is organized as follows: Section 2 presents related work, Section 3 introduces the proposed NSSA model, Section 4 provides simulation results and finally Section 5 provides the conclusion.

RELATED WORK

One major prerequisite to minimize the false alarm rates of IDS is fusing data from multiple heterogeneous sources in an efficient way. Data fusion methods can fulfill the requirement of NSSA and many fusion techniques have been proposed. Liu et al. [5] proposed a multi sensor data fusion method using multi-class support vector machines. Heterogeneous alert data fusion can also be done by using multi-layer feed-forward neural network [6] for generating network security situation.

Evolutionary neural network [7] reduces the parameters of neural network and extracts the network security situational factors that provide the quantification of network security situation. Juan et al. [8] proposed techniques for alert analysis and evaluation of threats that provide high level knowledge, such as different severity levels for different types of attacks to locate the most dangerous attack. This reduces the administrator’s time and efforts in processing large amount of security alert events. It generates attack graphs based on time and space dimension but no mechanism for alert fusion was considered.

Honeypot can be used to improve the false alarm rate of IDS [9] but it does not provide the security state of the whole network. Zhao-Yang et al. [10] presented a network security situation evaluation method based on D-S evidence theory as the information fusion technique to enhance performance. Fang Lan et al. [11] proposed a framework for network security situation awareness based on knowledge discovery method that supports for exact designing and automatic construction of network security situation graph

PROPOSED NSSA MODEL BASED ON KNOWLEDGE DISCOVERY

The NSSA model is presented in Figure 1. The network security situation sensors generate alert events either because of network attacks or whenever the monitored parameter exceeds the threshold value. Alert data are of different formats as they are gathered from heterogeneous sensors. Maintaining uniformity is a major issue when dealing with alert event data from different sources.

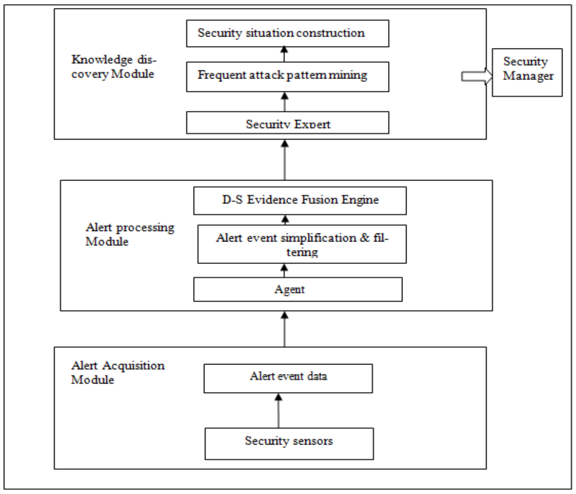


Fig. 1. The NSSA model

Events are processed and converted to an identical format that is represented as a multi-tuple:

e_i= {detectTime $_i$,eventType $_i$, attack $_i$,srcIP$_i$, desIP$_i$, srcPort$_i$, desPort$_i$, protocol $_i$, sensorID $_i$ }.

Data processing Module

The data processing module performs different operations on the alert events collected by the agents in order to reduce the number of effective alert events. It includes following operations:

Event simplification [e $_1$, e $_2$, • •, e $_n$] ---> e $_m$ the repeated alert events are simplified, which have occurred number of times, to reduce the total number of effective events.

Event filtering [e$_i$, P (e$_i$)/€ H] $\longrightarrow \varnothing$ if the property P (e$_i$) does not belong to a specific valid set of H the event is filtered out. Such events occur due to some network conditions that are not related with any type of attack.

Knowledge discovery Module

Knowledge Discovery is a method of finding new attack patterns from the set of security alert events gathered from sensors, which are meaningful for security situation generation. The frequent pattern (FP) mining algorithm [12-13] is chosen to extract the security context knowledge from the set of security alert events. The FP tree algorithm is used to generate the frequent patterns of attack. It abbreviates ample databases to compressed tree structure and efficiently mines the set of frequent patterns. The FP tree algorithm executes faster as it avoids repeated database scanning and does not use candidate item generation. It includes tree generation algorithm and FP_growth frequent pattern mining algorithm.

FP_growth adopts a divide and conquer strategy as follows. First, it compresses the database representing frequent items into a frequent –pattern tree, or FP tree,

Table 1. Set of events

TID	Event ID
1	E1,E2,E5
2	E2,E4
3	E2,E3
4	E1,E2,E4
5	E1,E3
6	E2,E3
7	E1,E3
8	E1,E2,E3,E5
9	E1,E2,E3

Which retains the item- set association information. It then divides the compressed database into a set of conditional databases; each associated with one frequent item or “pattern fragment”, and mines each such database separately.

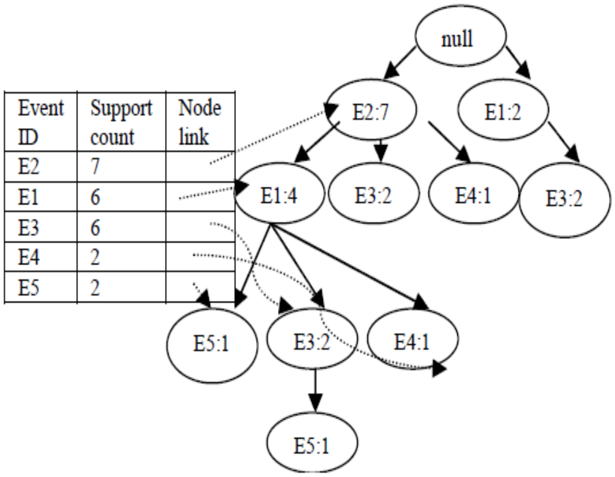
Table 1 shows set of events occurring during different time intervals. The database is mined as follows. The first scan of the database derives the set of frequent items (1-itemsets) and their support counts (frequencies). Let the minimum support count be 2. The set of frequent items is sorted in the order of descending support count. This resulting set or list is denoted L. Thus L= {{E2:7}, {E1:6}, {E3:6}, {E4:2}, {E5:2}}

An FP tree is then constructed as follows. First create the root of the tree, labelled with “null”. Scan the database second time.

The events in each time interval are processed in L order and a branch is created for each time interval. To facilitate tree traversal, an event header table is built so that each item points to its occurrences in the tree via a chain of node links. The tree obtained after scanning all of the time intervals is shown in Figure 2 with the associated node-links.

The FP Tree is mined as follows. Start from each frequent length-1 pattern, construct its conditional pattern base, then construct its conditional FP tree and perform mining recursively on such a tree. The pattern growth is achieved by the concatenation of the suffix pattern with the frequent patterns generated from a conditional FP tree

Mining of the FP tree is summarized in Table 2 and detailed as follows. First E5 is considered. E5 occurs in two branches of FP tree of figure. The paths formed by these branches are {E2, E1, E5:1} and {E2, E1, E3, E5:1}. Therefore considering E5 as a suffix, its corresponding two prefix paths are {E2, E1:1} and {E2, E1, E3:1}, which forms its conditional pattern base. Its conditional FP tree contains only a single path, {E2:2, E1:2}, E3 is not included because its support count of 1 is less than the minimum support count.



Event ID	Support count	Node link
E2	7	
E1	6	
E3	6	
E4	2	
E5	2	

Fig. 2. Constructing a FP Tree

The single path generates all the combinations of frequent patterns:{E2,E5:2},{E1,E5:2},{E2,E1,E5:2}.Similarly frequent patterns are generated for E4, E3, and E1.

Table 2. FP Tree mining by creating conditional pattern bases

Event	Conditional Pattern base	Conditional FP Tree	Frequent patterns generated
E5	{{E2,E1:1} ,{E2,E1,E3:1}}	<E2:2,E1:2>	{E2,E5:2}, {E1,E5:2}, {E2,E1,E5 :2}
E4	{{E2,E1:1}, {E2:1}}	<E2:2>	{E2,E4:2}
E3	{{E2,E1:2},{E2:2}, {E1:2}}	<E2:4,E1:2> ,<E1:2>	{E2,E3:4}, {E1,E3:4}, {E2,E1,E3 :2}
E1	{{E2:4}}	<E2:4>	{E2,E1:4}

Algorithm: FP_growth.

Input:

A database

Min_sup, the minimum support count threshold

Output: the complete set of frequent patterns.

Method:

The FP tree is constructed.

The FP tree is mined by calling FP_growth (FP_tree, null), which is implemented as follows:

```
FP_growth (Tree, $\alpha$)
For each (ai in the header of Tree) do
$\beta := ai \cup \alpha$
Generate ($\beta$ with support = ai.support)
Construct $\beta$'s conditional base pattern
And $\beta$'s conditional FP-Tree Tree$\beta$
If Tree$\beta \neq O$
Then call FP-growth (Tree$\beta$, $\beta$)
Initially call: FP-growth (Tree, null)
```

Where α & β are combination of the nodes in a path of the tree.

Event Fusion Module

Event fusion $e_i \ldots e_j$-----fuse-----→e_k It is used for merging two different events so as to track attack behavior by combining multiple alert events by using the information fusion techniques such as the D-S evidence theory so as to minimize the false positive rate. The most elementary concept of D-S evidence theory is the basic probability assignment function m defined as:

$$(P\ \{C, W\}) \longrightarrow [0, 1],$$

$$m\ (\text{¢}) = 0,$$

$$m(C) + m(W) + m(\{C, W\}) = 1\ ,$$

Where, C refers to correct alert, W refers to wrong alert m(C) refers to the confidence level of a security alert detected by a sensor.

The basic probability assignment function of combined evidence is

$$m\ (e) = K^{-1} \sum_{e_1 \cap e_2 = e} m_1(e_1) m_2(e_2)$$

$$= m_1\ (e_1) \oplus m_2\ (e_2)$$

Where, K refers to the normalization factor,

$$K = \sum_{e_1 \cap e_2 \neq \text{¢}} m_1\ (e_1)\ m_2\ (e_2)$$

SIMULATION AND EVALUATION OF RESULTS

The NSSA system developed is applied in the experiment. A wireless network of 100 nodes is created. In the simulation attacker nodes are created to attack the sink node continuously. The simulation parameters are presented in Table 3. The attacker uses ICMP ping sweep technique to perform DDoS attack. Constant bit rate traffic is generated to test the working of the system using

UDP agent. The agents gather the alert events generated by the intrusion detection sensors and pass them to the data processing module. The knowledge discovery module applies the FP tree algorithm to generate frequent attack patterns and then the fusion module fuses the alert events using D-S evidence theory. Finally the security situation data is passed to the security manager for decision making.

Table 3. Simulation Parameters

Parameters	Description
Number of nodes	100
Routing	Adhoc
CST	250m
Topology	Flat grid
X dimension of topology	100m
Y dimension of topology	100m
Mac type	802.11

The performance of the system is tested by varying the number of nodes from 100,120,140, 160, 180 and 200. From the Fig. 3 it is observed that the maximum value of average end to end delay with KD technique is considerably less as compared to without knowledge discovery technique. As the number of effective security alert events are greatly minimized by event filtering and fusion.

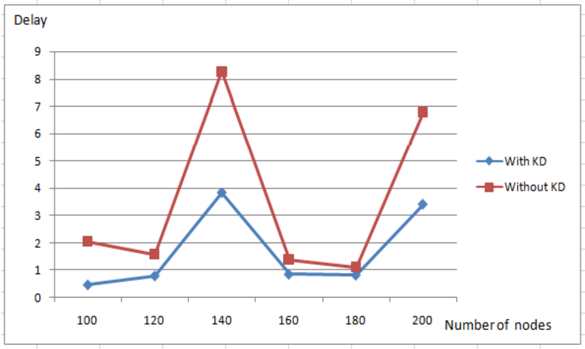


Fig. 3. Number of nodes vs Avg. end to end delay

Fig. 4 shows that the average energy consumption with KD is less as compared to without KD hence the performance of the system improves.

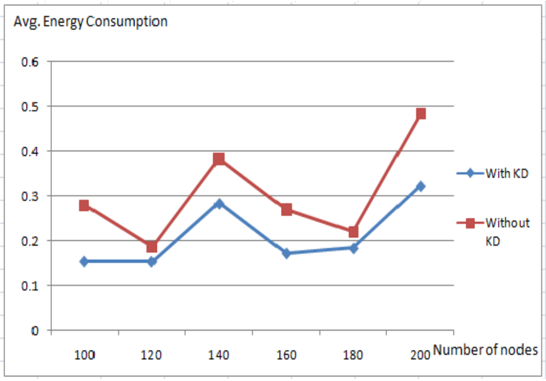


Fig. 4. Number of nodes vs Avg. energy consumption

The packet drop greatly reduces and packet delivery ratio shows a significant improvement with knowledge discovery technique as attacker is detected and prevented from sending packets. While performing the DDoS attack the attacker tries to send many packets to the same node in a short interval of time thus preventing the node from accepting legitimate traffic. Fig. 5 shows that with KD technique the packet delivery ratio remarkably increases as compared to without KD.

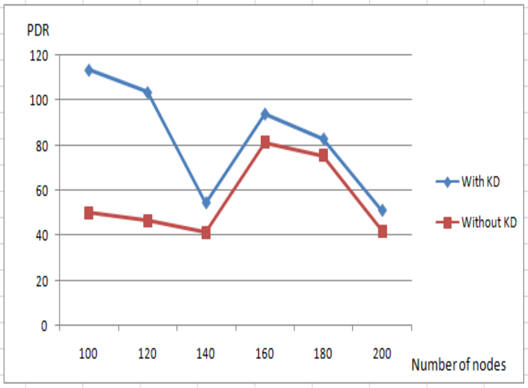


Fig. 5. Number of nodes vs Packet delivery ratio

The throughput of the system increases as the average rate of successful packet delivery over the communication channel increases due to the decrease in packet loss. Fig. 6 shows that the throughput with KD technique is significantly higher than without KD.

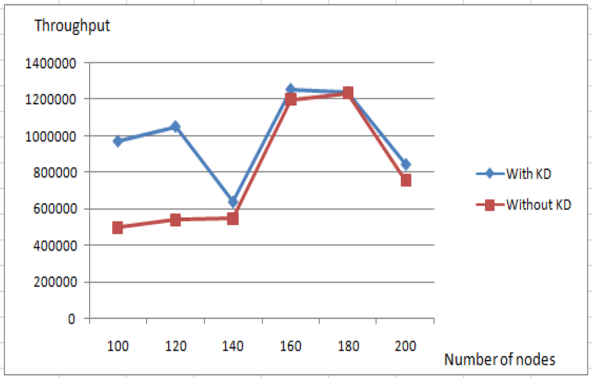


Fig. 6. Number of nodes vs Throughput

CONCLUSIONS & FUTURE WORK

NSSA based on Knowledge discovery technique efficiently overcomes the limitations of the existing security controls and provides effective construction of network security situation. The event fusion technique converts many trivial attacks into DDoS attack. Security analysts do not need to process thousands of security alerts that prevent them from managing and controlling their networks effectively.

The results show that the performance of the system remarkably improves with knowledge discovery technique in terms of average end to end delay, average energy consumption, packet delivery ratio and throughput.

As this work progresses, the difficulties of actual forecasting of emerging fatal malicious activities need to be explored.

REFERENCES

[1] J.R. Goodall, W.G. Lutters and K. Anita, "The work of intrusion detection: rethinking the role of security analysts," in Proc. of the Tenth Americas Conf. on Information System, New York, 2004, pp. 1421-1427.

[2] M. R. Endsley, "Design and evaluation for situation awareness enhancement", Proceeding of the human factors society 32nd annual meeting, Santa Monica, CA, pp.97-101, 1988.

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

[4] T.Bass, "Multi sensor Data Fusion for Next Generation Distributed Intrusion Detection Systems", Invited Paper1999 IRIS National Symposium on Sensor and Data Fusion, pp.24-27, May 1999.

[5] Liu Xiaowu, Wang Huiqiang, Lai Jibao, and Liang Ying, "Network Security Situation Awareness Model Based on Heterogeneous Multi-sensor Data Fusion", 2007 IEEE.

[6] Liu Xiaowu, Yu Jiguo, Wang MaoLi, "Network Security Situation Generation and Evaluation Based on Heterogeneous Sensor Fusion", 2009 IEEE

[7] Ying Liang, Hui-Qiang Wang, Ji-Bao Lai,"Quantification of Network Security Situational Awareness Based on Evolutionary Neural Network", Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 19- 22 August 2007 IEEE.

[8] Juan Wang, Feng-li Zhang, Jing Jin, Wei Chen, "Alert Analysis and Threat Evaluation in Network Situation Awareness", 2010 IEEE.

[9] Babak Khosravifar, Jamal Bentahar, "An Experience Improving Intrusion Detection Systems False Alarm Ratio by Using Honeypot", 22nd International Conference on Advanced Information Networking and Applications, 2008 IEEE.

[10] Zhao-Yang Qu, Ya-Ying Li, Peng-Li, "A Network Security Situation Evaluation Method Based on D-S Evidence Theory", 2nd Conference on Environmental Science and Information Application Technology, 2010 IEEE.

[11] Fang Lan, Wang Chunlei, and MaGuoqing, "A Framework for Network Security Situation Awareness Based on Knowledge Discovery" 2nd International Conference on Computer Engineering and Technology 2010 IEEE.

[12] Jia Han, Micheline Kamber., "Data Mining concepts and techniques", second edition 2006, Elsevier Inc.

[13] J Hall, J Pei, Y Yin, "Mining frequent patterns without candidate generation".2000 ACM, SIGMOD int'I Conf on Management of Data (SIGMOD'OO), DallaS, TX, 2000