Managing Intrusion-Detection Alerts Based on Fuzzy Comprehensive Evaluation

C.P. Mu H.K. Huang S.F. Tian

School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China

Abstract

An algorithm based on fuzzy comprehensive evaluation for correlating the alerts produced by intrusion detection systems is presented. The paper also gives an approach to learn the confidence metric for each type of alerts, which can be used to filter alerts further. The false positive alerts and duplicate alerts can be reduced significantly by using both the correlation algorithm and the confidence learning method

Key Words: intrusion detection, alert correlation, alert processing, fuzzy comprehensive evaluation

1.Introduction

Although intrusion detection systems have been studied and developed over 20 years, they are still on the stage of development. Many parts of the system are not perfect and need improvement. The criticism of the weaknesses of present IDSes focuses on the following points.

Firstly, both the false positive rate and false negative rate of IDSes are often unacceptable. Some IDSes could send thousands of false alerts each day, which make administrators of systems know nothing to deal with them.

Secondly, all the current IDSes focus on low-level attacks or anomalies; none of them can capture the logical steps or strategies behind these attacks [1].

The last but not least point is that the response capabilities of the current IDSes are weak. After detecting attacks, most IDSes do nothing except sending alerts.

The correlation algorithm and the confidence learning method proposed in this paper can significantly reduce false positive alerts and duplicate alerts produced by IDSes, which have been proved by our initial experiments. Our approach can be potentially extended to uncover high-level attack strategies and provide a base to choose appropriate response measures.

2.Related Work

In the past, the most of attentions in IDS research area were put on the research of Fore-Intrusion-Detection techniques that include collecting and preprocessing data form network and hosts for detection, finding new detection algorithms and improving old classification algorithms in order to promote the detection accuracy and reduce false alert rate. Recently in the IDS research area, there is a trend of integrating different security tools and network manage tools together to reduce false positive rate and false negative rate, and greatly increases the defence ability of networks. In the integration of different security systems, one of its key approaches is the correlation and fusion of alerts generated by different IDSes, which belongs to Post-Intrusion-Detection technique. Although very little research has been done on correlation or fusion of alerts, several alerts correlation techniques and models have been proposed since the beginning of 21 century.

The principles of correlating alerts in [1] and [2] are similar. Their proposed approach identifies the prerequisite (pre-condition) and consequence (post-condition) of each type of attacks, and correlates the corresponding alerts by matching the consequence of some previous alerts and the prerequisite of some later ones. To do this, we have to understand all possible attacks very well in order to know their prerequisites and consequences. Therefore, this approach highly depends on the experience of human experts and is hard to process new types of attacks.

In [3], an approach has been proposed to learn alert correlation models by applying machine learning techniques to training data sets embedded with known intrusion scenario. This approach can automatically build models for alert correlation. However, it requires training in every deployment, and the resulting models may overfit the training data, thereby missing attack scenarios not seen in the training date sets.

The article [4] provides a brief view of Intrusion Detection concepts and terms, an overview of the art and science of multi-sensor data fusion technology, which is helpful to develop our model.

The algorithm of correlating alerts proposed in my paper is suitable for online alert processing for its fast running speed. Using fuzzy comprehensive evaluation to correlate alerts and supervised confidence learning to filter alerts, we developed the Intrusion Detection Alerts Manage System (IDAMS).

3. The Architecture of IDMAS and Components' Functions

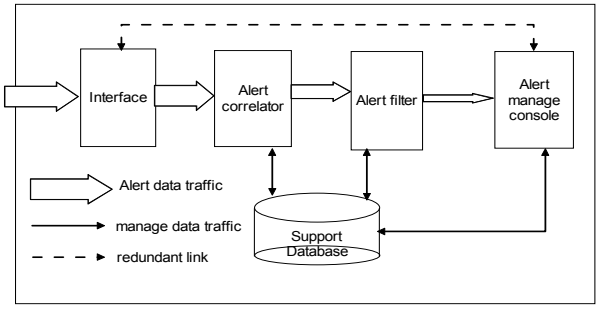

Fig. 1: The architecture of IDAMS

The functions of each component in Fig.1 are as follows:

The interface module is responsible for receiving and reading alerts. It can not only receive alerts online from multiple IDSes but also read alerts from a central database off-line. The main task of the alert correlator is to correlate current alert with each event (alert) stored in the historic template table of support database. In the alert filter, alerts are filtered further according to their corresponding confidences. Only an alert with a higher confidence value than the alert threshold can eventually reach the alert manage console. The console sends message to an administrator by displaying an alert window on the screen of the computer, sending Email, or sending mobile phone message etc as long as it receives a new alert. Through the alert manage console, an administrator could browse and manage alerts, maintain IDAMS, and configure its parameters. Also raw alerts could be read directly on the console through the redundant line.

There are 3 tables in the support database: the historic template table, the correlated alert table, and the intrusion event parameter table. The historic template table stores the alerts that recently happen. These alerts server as the templates and are compared with the newest alert. The correlated alert table stores correlated alert lists. Each list stores a series of alerts that correlated each other. The event severity, the total number of the corresponding alert, the number of the corresponding false alert and the confidence for each known attack are stored in the intrusion event parameter table. Section 5 shows how to determine these parameters in detail.

4. Correlating Alerts based on Fuzzy Comprehensive Evaluation (FCE)

4.1. Fuzzy Comprehensive Evaluation (FCE)

The mathematical model based on FCE could resolve the problem of uncertainty among multiple factors and give an objective evaluation result by comprehensively considering multiple related factors. Suppose $U=\{u_1,u_2,...,u_n\}$ and $V=\{v_1,v_2,...,v_m\}$ where U is the set of evaluation factors of FCE, and V is the set of evaluation remarks of FCE. From [5], we get the following model:

$$B = K \circ R \qquad (1)$$

where B is a subset on V, $K=(k_1,k_2,...,k_n)$ is a subset on U, R is a fuzzy matrix.

Different mathematical model can be obtained by choosing different operator $\circ$ (e.g. $M(\wedge,\vee)$, $M(\bullet,+)$, $M(\bullet,\vee)$, $M(\wedge,\oplus)$ in the formula (1). $\wedge$, $\vee$ and $\bullet$ are min, max and multiply operations respectively. $\oplus$ is the operation that $a \oplus b = \min(1, a+b)$. Here, we choose the model $M(\bullet,+)$ and get equation (2) from (1), which proportionally takes every factor into account according to its corresponding weight coefficient.

$$b_j = \sum_{i=1}^{n} k_i \bullet r_{ij} \qquad \square 2 \square$$

where r_{ij} (i=1,2,...,n; j=1,2,...,m; $r_{ij} \in [0,1]$) represents the membership degree of the subject to remark v_j from the viewpoint of factor u_i.

4.2. Correlating Alerts

Definition 1 correlated degree B. It represents how much two events relate each other or probability that two events belong to the same attack process.

The set of evaluation factors in our model is $U=\{u_1,u_2,u_3\}$, where u_1 is the time interval between two alerts, u_2 is the source IP address similarity between two alerts, u_3 is the signature similarity between two alerts. We will introduce more factors to improve our model in future. The $V=\{v_1,v_2,...,v_m\}$ is the set of evaluation remarks, where v_i (i=1,2,...,m) represents a template event stored in the historic template table.

The parameter r_{ij} in the fuzzy matrix is given by its corresponding membership function. The

determination of a membership function for each factor is based on the following knowledge and experience: the shorter the time interval between two alerts, the more similar the signatures between two alerts, and the more similar the source IP addresses between two alerts, then the more possible two alerts come from the same attack process [6].

Attributes in alerts are denoted as follows: v_j.time — the time when the event j (j=1,2,…,m) happens; v_j.Sip — the source address of the event j; v_i.signature — the signature of the event j; v_0 — the current event (or the newest alert).

The membership function of the time interval is as follow:

$$r_{1j} = \begin{cases} 1 & T_j < 10\,\text{min} \\ \dfrac{60 - T_j}{50} & 10\,\text{min} \le T_j \le 60\,\text{min}) \\ 0 & T_j > 60\,\text{min} \end{cases} \quad (3)$$

where T_j=v_0.time-v_j.time.

The membership function of the source IP address similarity is as follow:

$$r_{2j} = \frac{x}{32} \quad \square 4 \square$$

where x is the number of the same bits while comparing v_0.Sip with v_j.Sip from left to right. Here, the mask address of source IP addresses of two alerts should be taken into account as well.

The membership function of the signature similarity is as follow:

$$r_{3j} = \begin{cases} 0 & v_0.signature \neq v_j.signature \\ 1 & v_0.signature = v_j.signature \end{cases} \quad (5)$$

The process of correlating the current alert with template events is as follows:

1) The fuzzy matrix $\underset{\sim}{R}$ is obtained from equations (3) (4) (5)
2) The weight vector K=(k_1,k_2,k_3) is obtained according to the experience of experts in the area.
3) Calculating the correlated degree B from equation (2).
4) According to the principle of the maximum membership degree, getting b_L=max{b_1,b_2,...,b_m}□1≤L≤m, which means that the current event v_0 (alert) is most likely to be in the same attack process with the event v_L stored in the historic template table.
5) If $b_L \ge b_{threshold}$ (the threshold of the correlated degree), then v_0 and v_L are correlated. In other words, both of v_0 and v_L are in the same attack process. v_0 is written into the correlated alert table and replace v_L in the historic template table as a new template.
6) If $b_L < b_{threshold}$, then v_0 and v_L are not correlated. In other words, v_0 is either a beginning of a new attack process or a positive false alert. Under the condition, v_0 is transferred to the alert filter.

5. Supervised Confidence Learning and Alerts Filtering

In most cases when we receive an alert from IDS, we could not conclude that this definitely is an attack. In other words, we could not exclude the existence of false alerts. It is more reasonable to send alerts with corresponding confidences than those alerts without confidence parameter. However, there are no such confidence parameters in many IDSes (esp. Misuse IDSes), which causes inconvenience to process alerts further.

We propose a supervised approach to learn a confidence to improve the IDS without such parameter. The confidence parameter could be used to filter alerts with low confidence or false alerts and provide an important metric to choose response measures against attacks. The approach is expressed as follow equation:

$$C_i = \frac{S_i + S_0 - E_i}{S_i + S_0} \quad (6)$$

where i represents the type of an alert (For convenience, an alert whose type is i is referred as alert i in the following depiction); C_i is the confidence of an alert whose type is i; S_i is the total number of alerts i that ever received, it will be added 1 whenever an alert i is received; E_i is the total number of false alerts i, it will be added 1 whenever an alert is received and confirmed to be false by an administrator; S_0 is an initial value set by an administrator, which determines the speed of learning and the tolerance to false alerts. We have to balance these two aspects. We usually set S_0□[10,100].

An administrator could set different alerting threshold values according to different severities of alerts because different severities cause different threats and possible damages to the protected system. In this way, those alerts with lower confidence values than the alerting threshold would be filtered.

6. Experiments and Analysis

In the experiment, Snort 2.0 IDS and IDAMS were deployed on a subnet (211.71.75.0/24) of our lab, which connects Internet. BlackICE PC Protection and Norton Internet Security 7.0 IDSes were installed on each host in the subnet respectively as well. We collected alert data from 9:30 am to 5:30 pm and carried simulated attacks (such as Vertical Scan, CGI Vulnerability Scan etc.) on the subnet from other different subnets in our campus network during the 8 hours data collection. The experiment results are

shown in Table 1 and Table 2.

Table1 the experiments results of correlating alerts

	$b_{threshold}$ (k_1,k_2,k_3)	Detection rate	False alert rate	Incorrect correlated alert rate
Snort	N/A	65.6%	93.8%	N/A
IDAMS	1.5 (1.0,0.5,0.5)	64.6%	30.9%	21.3%
	1.5 (0.5,1.0,0.5)	64.6%	23.2	15.4%
	1.5 (0.5,0.5,1.0)	65.6%	25.1	26.4%

Table2 the experiments results of single attack

Attack Tool	Attack item	Number of Alerts/ Number of alert classes			
		□	□	□	□
DDoS uperKill	DoS attack	290/1	177/2	0	1/1
x-scan 3.0	Vertical scan	26/4	68/8	4/4	4/4
x-scan 3.0	CGI vulnerability scan	108/8	3/3	4/1	8/8
Swan	Serv-U Ftp MDTM overflow	1/1	1/1	0	1/1

Note□ □Snort;□BlackICE;□Norton;□IDAMS;
DoS (Denial of Service) attack lasted 1 minute.

From Table 1, when k_3 (the weight coefficient to the signature similarity) is greater than k_1 (the weight coefficient to time interval similarity) and k_2 (the weight coefficient to IP address similarity), the corresponding incorrect correlated alert rate is high. The false positive alert rate is high while k_1 greater than other 2 weight coefficients. Other experiments also show the above result. Therefore, let $k_2 \geq k_1 \geq k_3$ when these weight coefficients are decided.

The experiment in Table 2 was carried out without background traffic, which shows that IDAMS can greatly reduce duplicate alerts and balance the number of alert and number of alert classes very well.

In the experiment of alert confidence learning, we found students in our campus network often visit a hot website XXXX.net whose service port is not normal port 80 but 8080. The IDS would generate alerts if a computer visits an abnormal port (not 80 port) of a web server in a network. Obviously the alert is false. After a period of learning according to equation (6), IDAMS could filer this kind of alert.

7.Conclusions

1) The proposed approach solves the problem of duplicate alerts by merging these duplicate alerts into a statistic report that can reflect attacking condition better than a single alert.
2)IDAMS could effectively filter false positive alerts through alert correlation and alert confidence learning.
3)Both of correlated alerts and alert confidences produced by IDAMS are very helpful to catch intrusion logical steps and choose appropriate intrusion responses.

In order to improve IDAMS, we are studying the dynamic weight vector K, which means different alerts' correlation uses different weight vectors. More evaluation factors will be introduced into our model. These approaches will make correlation more exact and objective.

8.Reference

[1] Peng Ning, Y.Cui., "An intrusion alert correlator based on prerequisites of intrusion," *Available as Technical Report TR-2002-01*, Department of computer science, North Carolina State University, January 2002.

[2] Frédéric Cuppens, Alexandre Miége. Alert, "Correlation in a Cooperative Intrusion Detction Framework," *Proc. of the 2002 IEEE Symposium on Security and Privacy (S&P'02).*

[3] O.Dain, R.K.Cunningham, " Fusing a heterogeneous alert stream into scenarios," *Proc. of the 2001 ACM Workshop on Data Mining for Security Application*, November 2001.

[4] Tim Bass, " Intrusion Detection Systems and Multisensor Data Fusion," *Communications of the ACM*, Apr. 2000, 43(4).

[5] XIE Jijan, LIU Chengping, "The methodology and application of fuzzy mathematics (in Chinese)," *China Hua Zhong University of Science and Technology Press*, May 2000, pp.197-230.

[6] Curistis A. Carver, John M.D.Hill and Udo W. Pooch, "Limiting uncertainty in intrusion response," *Proc. of the 2nd IEEE information Assurance and Security Workshop*, West point, NY, USA,June 2001.