Algorithm of reducing the false positives in IDS based on correlation Analysis

To cite this article: Jianyi Liu et al 2018 IOP Conf. Ser.: Mater. Sci. Eng. 322 062016

You may also like

- Aiding Alarm Rationalization by Automatic Identification of various sequential patterns in large volume of Alarm and Event log data
  Kaushik Ghosh and Gokula Krishnan Sivaprakasam

- Real-time arrhythmia detection with supplementary ECG quality and pulse wave monitoring for the reduction of false alarms in ICUs
  Vessela Krasteva, Irena Jekova, Remo Leber et al.

- Pattern discovery in critical alarms originating from neonates under intensive care
  Rohan Joshi, Carola van Pul, Louis Atallah et al.
Algorithm of reducing the false positives in IDS based on correlation Analysis

Jianyi Liu¹, Sida Li¹, Ru Zhang²,*

¹School of Beijing University of Posts and Telecommunications, Beijing, China
²School of Beijing University of Posts and Telecommunications, Beijing, China

*Corresponding author e-mail: liujy@bupt.edu.cn

Abstract. This paper proposes an algorithm of reducing the false positives in IDS based on correlation Analysis. Firstly, the algorithm analyzes the distinguishing characteristics of false positives and real alarms, and preliminary screen the false positives; then use the method of attribute similarity clustering to the alarms and further reduces the amount of alarms; finally, according to the characteristics of multi-step attack, associated it by the causal relationship. The paper also proposed a reverse causation algorithm based on the attack association method proposed by the predecessors, turning alarm information into a complete attack path. Experiments show that the algorithm simplifies the number of alarms, improve the efficiency of alarm processing, and contribute to attack purposes identification and alarm accuracy improvement.

1. Introduction

Intrusion detection system is an important measure to protect network security. But the IDS triggered thousands of false alarms, which make network security analyst tired of dealing with these alarms. In order to reduce redundant alerts and false positives, discover real single step or multi-step attacks in the network, and carry out early warning, there is an urgent need to use algorithms to analyze and process alerts.

Nowadays, there are many solutions on how to reduce false positives. One of them is improving the IDS algorithm which can obtain more accurate matching rules, the other aims at analysing and filtering the alarm data to separate the correct and wrong alarm. In this paper, an improved algorithm based on correlation analysis is proposed to achieve the goal.

The input data of our system is the IDS alert logs and the output are the attack rules after the correlation analysis.[1-3]

2. Algorithm Introduction

The flow of the whole algorithm is:

1. Reducing the false positives: We find out the characteristics through the analysis of IDS alerts, and remove the false positives by the difference between them.
2. Alert clustering: After the first step, we aggregate similar alerts by calculating similarity function of each alarm by the source IP and destination IP, source port and destination port, and then further reduce the number of alerts.
3. Correlation analysis: After the two steps, we start to makes correlations of each alerts, turn the messy and meaningless alerts into meaningful attack rules.
2.1. Reducing the false positives

Before the correlation analysis, we analyzed a large number of intrusion detection system alarm log, through the observation of these log, we have the following three characteristics:

1. The number of neighboring related alerts is different
   According to the study, in a certain time interval, the number of the neighboring related alerts of the false positives is much less than half of the true positives. So the number of neighboring related alerts can be used as an indication of whether an alert is a false or a true positive.

2. The density of alerts is different
   The statistics show that the true positive often bring in large-scale alerts within 0-1 second and false positives in the range of 1-5 seconds, far less than the true positives. Through, the different density of alerts, it can also distinguish true and false positives.

3. Periodic false positives
   Due to the host configuration or other reasons, IDS may produce massive alerts regularly, so we just need find a certain time interval without any attack, then all the warning during this time is false positives. For this periodic false positives, it is necessary to find the period of alert information in the original. The periodic false positives can be deleted in the frequency domain by Fourier transform.

Using the characteristics of the true and false positives, we can screen the alerts generated by intrusion detection systems, thus reducing the useless alarms and improving the efficiency of correlation analysis.

2.2. Alert clustering

This method grouping and merging those alerts by defining the similarity membership function to compare the degree of similarity in some important properties of the alarm, and reduce alarm. The algorithm flow is as follows:

- Alert X $(X_1, X_2, ..., X_n)$
- Alert Y $(Y_1, Y_2, ..., Y_n)$
- Specific Similarity Algorithm
- $(S_1, S_2, S_3, ..., S_n)$
- Minimum Similarity $(M_1, M_2, ..., M_n)$
- Feature Similarity $(Sim_1, Sim_2, ..., Sim_m)$
- Similarity Expectation $(E_1, E_2, ..., E_n)$
- Overall Similarity $SIM(X, Y)$

Fig.1 the flow of Alert clustering

As shown above, first of all, calculate the similarity of the alarm, involved source IP, destination IP, source port, destination port, and so on. Among them, each alarm attribute has a similarity function corresponding to it. The function eventually returns a similarity between 0 and 1, 1 means that the two alarms are exactly the same, and 0 means completely different. The membership function of similarity is expressed as a formula:

$$SIM(X,Y) = \frac{\sum E_i SIM(X_i,Y)}{\sum E_j}$$
Among them, X and Y represent two alerts, j is the index of the alarm attribute value, and \( \text{Ej} \) is the expected value of the attribute J, that is, the weight of the j attribute in the membership function. When \( \text{SIM}(X, Y) > \Theta \), it is considered that X and Y can be correlated, otherwise they cannot be linked. Among them, \( \Theta \) is a given threshold\(^4\).

The advantages of alert clustering method is it can be good on the alert classification, and it has good real-time performance. But it can not show the causal relationship between the alarms, and therefore it can not tell us the attack rule among the alert logs.

### 2.3. Alert correlation

A complete attack is usually composed of multiple attack steps. The alarm generated by traditional IDS is too singleness, can not find such a complex attack. Therefore, we have to find the attack rules through the alert correlation.

#### 2.3.1. Causal correlation algorithm

The causal correlation algorithm is to represent the preconditions of each attack and the consequences of the attack in a logical predicate. If the result of an attack is the precondition of another attack, it is considered a dependency exists between the two of them, and it is linked, so as to realize the reconstruction of the attack scene.

The algorithm defined the meaning of hyper-alert, which includes fact, prerequisite and consequence. The fact is a set of property names, prerequisite is a logical combination of predicates, and consequence is a set of predicates such that all the free variables in consequence are in fact\(^5\).

Algorithm steps:
1. Defines the preconditions and consequences of each known attack type in advance
2. Generates hyper-alert instances according to time constraints and hyper-alert definitions of type, and the process of generating hyper-alert instances is actually a aggregation process of raw alert information.
3. Correlate hyper-alert instances.

This method can find the causal relationship between the alarms, but the correlation results depends on the knowledge base, new attack emerge in an endless stream, this method do not have the ability to do a real-time correlation analysis.

#### 2.3.2. Reverse causal correlation algorithm

In the following scenario (such as botnets), the attacker can do IP Sweep attack with thousands of hosts, and then found several loopholes from the hosts, then use the host as a springboard to attack the others, may only one host can attack successfully in the last time. If we use the causal correlation algorithm, we will get thousands of correlation rules in the first step, and then only a few can go to the next step, and may be only one multi-step attack reached the last step, causing a great waste of resources.

Based on that, this paper proposes an reverse causal correlation algorithm. when we start doing correlation analysis, we consider the final attack firstly, then go step by step to the origin, get a complete attack path finnally.

Algorithm steps:
1. Same as causal correlation algorithm;
2. Same as causal correlation algorithm;
3. Correlate hyper-alert instances, starting from the final attack, until the first step of multi-step attacks.

Experiments show that the method overcomes the efficiency of the causal correlation algorithm, and greatly reduces the redundant correlation results, reduces the correlation complexity, and improves the efficiency of correlation algorithm.

### 3. Experiment

In order to verify the effectiveness of the system, this paper selects the industry standard data set DARPA\(^6\) as experimental data. We use DARPA2000 LL DOS1.0 which contains a multi-step attack scenario: first the attacker performs a scripted IP sweep on the network, then the attacker try to find
which hosts might be vulnerable to the exploit that he has. Then the attack script attempts the sadmind Remote-to-Root exploit several times against each host. Entering next phase, the script performs a telnet login, and finally, the attacker manually launches the DDoS.

We use snort as the IDS, and use the forward and reverse correlation method to reconstruct the attack scene, as shown in Fig. 2, we can correlate IDS alarm log, forming one or several meaningful attack path.

| Source Address | Target Address | Alert Event      | Intention Alert |
|----------------|----------------|------------------|-----------------|
| 1 202.77.162.213 | 172.16.112.0/24 | Icmp_Ping_Sweep  | JP_Sweep        |
| 1 202.77.162.213 | 172.16.113.0/24 | Icmp_Ping_Sweep  | JP_Sweep        |
| 1 202.77.162.213 | 172.16.114.0/24 | Icmp_Ping_Sweep  | JP_Sweep        |
| 1 202.77.162.213 | 172.16.115.0/24 | Icmp_Ping_Sweep  | JP_Sweep        |
| 2 202.77.162.213 | 172.16.112.0/24 | Sadmind_Port_Request | Port_Scan |
| 2 202.77.162.213 | 172.16.113.0/24 | Sadmind_Port_Request | Port_Scan |
| 2 202.77.162.213 | 172.16.114.0/24 | Sadmind_Port_Request | Port_Scan |
| 3 202.77.162.213 | 172.16.112.20 | Sadmind_Overflow_Attemp | Remote_Overflow_Attemp |
| 3 202.77.162.213 | 172.16.112.50 | Sadmind_Overflow_Attemp | Remote_Overflow_Attemp |
| 4 202.77.162.213 | 172.16.112.20 | Rsh_Login         | Remote_Login    |
| 4 202.77.162.213 | 172.16.112.50 | Rsh_Login         | Remote_Login    |
| 5 172.16.112.20  | 172.16.112.50  | Attempted_Dos     | DDos Attack     |
| 5 spoofedIP      | 131.84.110.31  | Successful_Dos    | DDos Attack     |

**Fig. 2** the results of correlation analysis

From the Fig. 3, we can see that there are 3328 alerts produced by snort, after the Reducing of false positives, only remain 440, and then after alert clustering, there are only 7 categories of types of attacks, and finally we produce 3 rules by reverse causal correlation algorithm. Using this algorithm, the number of alerts is reduced from the 3 thousand to 3 complete attack rules, it also produce meaningful attack rules, instead of disordered alarm log.

| Reducing the false positives | Alert clustering | association analysis |
|-----------------------------|------------------|---------------------|
| input                       | 3328             | 440                 | 7                   |
| output                      | 440              | 7                   | 3                   |

**Fig. 3** the results of the algorithm

Comparing with causal correlation algorithm and reverse causal correlation algorithm, we can see the results from the Fig. 4, the reverse causal correlation algorithm eventually leaving only one complete attack path, and other attack paths as shown in left, are not our primary concern.

In this experiment, we have only a small number of hosts, while in reality, IDS produce ten of thousands of alerts. If using the traditional causal correlation algorithm, we will waste a lot of time and energy in dealing with unfinished attack. So the reverse correlation algorithm can greatly improve the efficiency of correlation analysis, and reduce the unfinished attack alerts, improve the analysis efficiency of warning log.
4. Conclusion
A new alert correlation model is proposed in this paper. The model takes advantage of the causal correlation algorithm. Not only the attack scene can be reconstructed accurately, but also it has good time complexity. It can extract attack patterns from the massive alarm logs which helps to reduce false positives. Alert correlation analysis is dependent on professional knowledge base, so the next step is to study how to generate a complete attack graph automatically [7].

Funding
This work was supported by the National key Research and Development Program of China (2016YFB0800903), the NSF of China (U1636212, U1636112).

References
[1] DHANAKOTI V, NEDUNCHEZHIAN R, et al. Correlated alerts and non-intrusive alerts [J]. Journal of Control Engineering and Applied Informatics.
[2] TIANZH, et al. Network intrusion detection model based on context verification [J]. Journal of Computer Research and Development
[3] ZHANG Y, ZHAO K. Alert correlation graph: a new method to evaluate the vulnerability of networks [J]. Journal of Defense Sci & Tech University, 2012, 34(3): 109-112
[4] Valdes A. Probabilistic alert correlation [A]. Proc of RAID 2001 [C]. Heidelberg, 2001. 54-68
[5] P-Ning, Y.Cui, D.S.Reeves. Constructing attack scenarios through concision of intrusion alerts [C]. In Proceedings of the 9th ACM Conference on Computer and Communications Security, Washington, D.C., November
[6] MIT Lincoln Lab. 2000 DARPA intrusion detection scenario specific datasets [EB]. https://ll.mit.edu/ideval/data/2000data.html
[7] Yu Fu, Cheng Hong, Survey of APT attack detection based on data analysis Journal of Communications