Computer Communications 38 (2014) 50–59

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Adaptive non-critical alarm reduction using hash-based contextual signatures in intrusion detection ☆

Yuxin Meng *, Lam-For Kwok

Department of Computer Science, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong Special Administrative Region

ARTICLE INFO

Article history:
Received 16 March 2013
Received in revised form 22 October 2013
Accepted 1 November 2013
Available online 9 November 2013

Keywords:
Network security and performance
Intrusion detection
Non-critical alarm reduction
Contextual signature
Hash function

ABSTRACT

Signature-based intrusion detection systems (IDSs) have been widely deployed in network environments aiming to defend against different kinds of attacks. However, a large number of alarms, especially non-critical alarms could be generated during the detection, which can greatly lower the effectiveness of detection and increase the difficulty in analyzing the generated IDS alarms. The main reason is that the detection capability of a signature-based IDS heavily depends on its signatures, whereas current IDS signatures are short of information related to actual deployment (i.e., lacking of contextual information). In addition, the traditional signature matching is a key limiting factor for IDSs in which the processing burden is at least linear to the size of an input string. To mitigate these issues, in this paper, we propose a novel scheme of hash-based contextual signatures that combines the original intrusion detection signatures with contextual information and hash functions. By using hash functions, our scheme can be used to construct an adaptive hash-based non-critical alarm filter which can further improve the performance of existing contextual signatures in filtering out non-critical alarms. Some examples of contextual information matching are also provided. In the evaluation, we discuss how to choose appropriate hash functions and investigate the performance upon implementation of the scheme with a real dataset and in a real network environment. The experimental results are positive and indicate that our scheme is encouraging and effective in filtering out non-critical alarms.

1. Introduction

Nowadays, network intrusions are being a big challenge to network security. To mitigate this problem, intrusion detection systems (IDSs) [3] have been widely deployed in various network environments. In general, intrusion detection is the process of monitoring computer system events or network events in order to detect the signs of possible *incidents* that violate the pre-defined security policies or standard security practices, such as malware (e.g., virus, Trojan, spyware), attackers gaining unauthorized access to systems and users' mis-behaviors (e.g., authorized users misuse their privileges). By automating the intrusion detection procedure, intrusion detection systems have the capability of defending against different kinds of attacks.

Traditionally, intrusion detection systems can be roughly classified into two categories: *signature-based IDS* and *anomaly-based IDS*. The signature-based IDS (or misuse-based/rule-based IDS) [4,9] identifies an attack by comparing current system events or network events with its signatures.[1] The detection ability of a signature-based IDS is heavily depending on the capability of its signatures (i.e., the number of signatures), therefore, this kind of detection systems can only detect known attacks. On the other hand, the anomaly-based IDS [10,12] detects an attack by identifying great deviations between current events and its normal profiles.[2] Based on the detection method, the advantage of an anomaly-based IDS is that it has the capability of identifying unknown attacks. However, in real settings, the signature-based approach is more widely used than the anomaly-based method since false alarm rate (FAR), which indicates the possibility of detecting an intrusion when there are no intrusions, of the signature-based method is much lower as compared to the method of identifying anomalies [14].

Problem. Although intrusion detection systems is effective in identifying attacks, a big suffering problem of these systems (both signature-based detection and anomaly-based detection) is that a large number of alarms, especially non-critical alarms, will be generated during their detection procedure [5,8]. This issue stems

☆ A preliminary version of this paper appears in Proceedings of International Conference on Computational Intelligence and Security (CIS 2011), pp. 978–982, 2011 [1].

* Corresponding author.
E-mail addresses: yuxin.meng@my.cityu.edu.hk (Y. Meng), cslfkwok@cityu.edu.hk (L.-F. Kwok).

[1] A signature (or called *rule*) is created by a specific IDS to describe an attack or an exploit.

[2] A normal profile is pre-established to describe the normal behavior of a user, a host or a network connection.

0140-3664/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.comcom.2013.11.001

primarily from the fact that current IDSs detect not only the intrusions, but also unsuccessful attack attempts (e.g., failed remote-to-local attacks [18]). In real deployment, it is hard for an IDS to understand the situation of an attack attempt [15], it thus has to report all detected attack attempts so as to mitigate security risks (i.e., reducing the chance of missing an attack). The large number of non-critical alarms can greatly decrease the effectiveness of such systems and heavily increase the burden of analyzing the IDS alarms [7].

The definition of a non-critical alarm can be defined as below [13,43]:

- *Definition of a non-critical alarm.* A non-critical alarm is neither related to a malicious activity nor related to a successful attack. That is, a non-critical alarm is either a false positive or a non-relevant true positive.

In other words, *non-critical alarms* contain both false alarms and non-relevant true alarms. A *non-relevant true positive* means that a generated alarm is a true positive but is not relevant to the real settings. The decision of this non-relevant true positive is subject to security experts who configure IDSs and analyze alarms. The non-critical alarms can adversely affect the results of analyzing the generated alarms and it is very difficult to filter out these alarms during the detection process [5]. To solve the problem of non-critical alarms, a lot of efforts have been made in literature such as *signature improvement*, *alert verification* and *alert correlation*.

Contributions. In this work, we advocate that combining intrusion detection signatures with contextual information is a promising solution to mitigate the above problem. But the traditional signature matching suffers from an issue that the processing burden is at least linear to the size of an input payload string [49], which can greatly decrease the matching performance. To mitigate these issues, based on our previous work [1], we propose a novel scheme of *hash-based contextual signatures* that combines the original intrusion detection signatures with not only contextual information but also hash functions, with the purpose of improving the performance of contextual signature matching and filtering out the non-critical alarms more efficiently.

The proposed scheme of *hash-based contextual signatures* is compatible with different representations of intrusion detection signatures (e.g., Snort signature-format or Bro signature-format). Specifically, by combining with hash functions, our scheme can be used to construct an *adaptive reputation-based contextual and hashed non-critical alarm filter*, which can help improve the performance of existing contextual signatures in filtering out non-critical alarms. For this filter, checking whether an alarm is non-critical depends on the stored contextual information. For example, even for an unsuccessful attack, an IDS can produce several alarms. But only those alarms which are exactly matched to the stored contextual information would be treated as non-critical alarms.

To the best of our knowledge, our work is an early work in constructing a non-critical alarm filter by using contextual signatures in practice. The contributions of our work can be summarized as below:

- We propose a scheme of *hash-based contextual signatures* that can further improve the performance of contextual signature matching and non-critical alarm reduction by combining both contextual information and hash functions. *Contextual information* here generally means that knowledge about the current state of the network such as networking features and target configuration. For non-critical alarms, *contextual information* can also refer to any alarm contents (e.g., descriptions) that can be used to identify a non-critical alarm.
- By using a hash function, our scheme can be utilized to construct an *adaptive reputation-based contextual and hashed non-critical alarm filter*, which is able to be adaptive to IP sources, in better improving the performance of contextual signature matching. The *adaptive* here means that our proposed filter can intelligently update *IP-based Index Hash Table* and *Table of Matched Contextual Information*.
- We evaluated our proposed non-critical alarm filter with a real dataset and in a network environment. The experimental results positively demonstrate that our scheme is encouraging and efficient in filtering out non-critical alarms.

The rest of the paper is organized as follows. We introduce the background of two types of IDS signatures and review some related work in mitigating false alarms or non-critical alarms in Section 2. Section 3 summarizes the original contextual signatures and its applications, describes our proposed scheme of hash-based contextual signatures in detail and introduces its application in constructing a non-critical alarm filter. In Section 4, we present how to choose a hash function, describe our experimental methodology and analyze the experimental results. Finally, we conclude our work with future directions in Section 5.

2. Background and related work

In this section, we aim to introduce the background of IDS signatures and review some related work and approaches in mitigating the problem of false alarms or non-critical alarms.

2.1. Background

In this section, we briefly introduce two types of IDS signatures: *Snort* signatures and *Bro* signatures. Snort [9,11] is an open-source lightweight signature-based network intrusion detection system while Bro [2,16] is an open-source anomaly-based network intrusion detection system. Both IDS-tools are very important and widely used in intrusion detection.

Snort usually uses *fixed strings* to describe an attack while Bro uses *regular expressions* to express an attack. The signatures of Snort are free and comprehensive, whereas the signatures of Bro are more flexible than those of Snort. The conversion between these two signatures is not very difficult. In Fig. 1, we show an

```
alert udp $EXTERNAL_NET any -> $HOME_NET 10080:10081
    (msg:"SCAN Amanda client-version request"; flow:to_server;
    content:"Amanda"; fast_pattern:only; classtype:attempted-recon;
    sid:634; rev:5;)
```

(a) Snort

```
Signature sid-634 {
ip-proto == udp
dst-ip == $HOME_NET    # The destination ip should be defined
                       #    in advance
dst-port == 10080, 10081
payload /.* Amanda/
event "SCAN Amanda client-version request"
}
```

(b) Bro

Fig. 1. An example: representing a signature with Snort signature-format and Bro signature-format respectively.