COMPUTERS & SECURITY 60 (2016) 177–192

Available online at www.sciencedirect.com

ScienceDirect

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

Computers & Security

MVPSys: Toward practical multi-view based false alarm reduction system in network intrusion detection

Wenjuan Li [a], *Weizhi Meng* [a,b,*], *Xiapu Luo* [c,d], *Lam For Kwok* [a]

[a] *Department of Computer Science, City University of Hong Kong, Hong Kong, SAR, China*
[b] *Infocomm Security Department, Institute for Infocomm Research, Singapore*
[c] *Department of Computing, The Hong Kong Polytechnic University, Hong Kong, SAR, China*
[d] *PolyU Shenzhen Research Institute, China*

ARTICLE INFO

Article history:
Received 9 December 2015
Received in revised form 25 March 2016
Accepted 21 April 2016
Available online 28 April 2016

Keywords:
Network security
Intrusion detection
False alarms
Alarm filter
Multi-view
Semi-supervised learning
Practical applications

ABSTRACT

Network intrusion detection systems (NIDSs) have been developed for over twenty years and have been widely deployed in computer networks to detect a variety of network attacks. But one of the major limitations is that these systems would generate a large number of alarms, especially false alarms (positives) during the detection. To address this issue, many machine learning approaches have been applied to reduce NIDS false positives. However, we notice that multi-view based approach is often ignored by the literature, which uses one function to model a particular view and jointly optimizes all the functions to optimize and improve the learning performance. In addition, most existing studies have not implemented their algorithms into practical alam systems. In this paper, we thus develop *MVPSys*, a practical multi-view based false alarm reduction system to reduce false alarms more efficiently, where each view represents a set of features. More specifically, we implement a semi-supervised learning algorithm to construct two-view items and automatically exploit both labeled and unlabeled data. That is, this system can automatically extract and organize features from an incoming alarm into two feature sets: destination feature set and source feature set, where the former contains the features related to the target environment and the latter contains the features about the source environment. In the evaluation, we deploy our system into two real network environments besides using two datasets. Experimental results indicate that our system can achieve a stable filtration accuracy of over 95%, offering a significant improvement as compared with the state-of-the-art algorithms.

1. Introduction

With the rapid development of computer networks, network intrusions (e.g., worms, spamware, trojan, denial of service, etc) have become a major threat (Symantec Corp., 2012), which can cause big losses without timely detection. To address this problem, network intrusion detection systems (NIDSs) have been widely deployed to defend against a variety of network attacks, and these detection systems have already become an essential component for current defense mechanism (Scarfone and Mell, 2007).

Traditionally, network intrusion detection systems can be classified into two general categories based on their detection methods: signature-based NIDS like Vigna and Kemmerer (1998) and anomaly-based NIDS like Ghosh et al. (1998). The

* *Corresponding author.* Tel.: +6564082842.
E-mail address: yuxin.meng@my.cityu.edu.hk (W. Meng).
http://dx.doi.org/10.1016/j.cose.2016.04.007
0167-4048/

former detects an attack by monitoring incoming traffic or program behaviors and comparing them to known signatures. The signature (or called *rule*) is a kind of description for known attacks and exploits. If an accurate match is identified, an alarm will be produced. In contrast, the latter detects a potential attack by identifying great deviations between incoming events and the pre-defined normal profile. The normal profile can be used to represent a normal behavior or network events. If the deviation exceeds a threshold, an alarm will be generated.

But one of the major limitations for these detection systems is that a lot of alarms, especially *false positives*, would be generated during the detection (Hubballi and Suryanarayanan, 2014; McHugh, 2000). False positives indicate situations identifying a normal event as an intrusion. As network traffic is increasing, false positives have become overwhelming for the IDS analysts. In fact, several studies such as Bloedorn et al. (2000) and Pietraszek (2004) have demonstrated that nearly 99% of alarms reported by an intrusion detection system are not related to security issues. Therefore, false positives are regarded as a key limiting factor to impede the effectiveness and development of NIDSs (Axelsson, 2000).

1.1. Motivations

In the literature, many machine learning approaches such as Alharby and Imai (2005), Law and Kwok (2004), Meng and Kwok (2011) and Pietraszek (2004) have been applied to reducing NIDS false positives. These efforts prove that the construction of an appropriate alarm filter is a promising way to reduce the number of false alarms produced by an intrusion detection system. However, we identify that state-of-the-art work gives little attention to multi-view based approach where each view represents a set of features. For example, in Web page classification, a Web page can be described by the document text itself and by the anchor text attached to hyperlinks pointing to this page (Sun, 2013).

In machine learning area, it is identified that conventional learning algorithms, such as support vector machines, kernel machines, etc, can concatenate all multiple views into one single view to adapt to the learning settings. However, this concatenation of single-view learning may cause over-fitting problems, when only a small training dataset is available (Xu et al., 2013). In contrast to single-view learning, many research studies indicate that multi-view learning can help improve and optimize learning performance (Sun, 2013; Xu et al., 2013; Zhou et al., 2007). Thus, multi-view learning should be given more attention in intrusion detection.

In addition, most existing studies did not implement their algorithms into a practical alarm reduction system. Motivated by these, in this work, we aim to develop a practical multi-view based false alarm filter to reduce NIDS false positives. Our work attempts to stimulate more research in the context of multi-view learning and piratical alarm system development.

1.2. Contributions

With the motivations above, we develop a practical false alarm reduction system based on multi-view based semi-supervised learning and investigate its performance using public datasets and live data in real network environments. To the best of our knowledge, our work is an early effort to develop a practical multi-view based system for NIDS false alarm reduction. Besides using datasets, we also deploy our system in real networks for evaluation. The contributions of this work can be summarized as below:

- First, we develop a multi-view based false positive reduction approach, which can particularly extract two feature sets (two-view) from an incoming NIDS alarm. The former contains the features related to the target environment (destination feature set), while the latter contains the features about the source (source feature set). Our work is an early study to apply multi-view technique in reducing NIDS false alarms.
- We further describe a prototype implementation called *MVPSys*, which can conduct false alarm reduction in either off-line mode or real-time mode. Specifically, our implementation employs a semi-supervised learning algorithm that can automatically exploit unlabeled data and labeled data without human intervention. This characteristic makes semi-supervised learning methods more suitable for real networks as they could require a small quantity of labeled data (Mao et al., 2009).
- In the evaluation, we explore the performance of *MVPSys* on Snort false alarm reduction (Roesch, 1999) with two datasets and in two real network environments. Experimental results demonstrate that our system can reduce false alarms at high and stable accuracy as compared to other supervised algorithms and relevant semi-supervised learning algorithms. Taking DARPA dataset as an example, our system can achieve an accuracy of 96.2%, while the other best similar algorithms can only reach around 91.2%.

The remaining parts of this paper are organized as follows. In Section 2, we introduce the background of NIDS alarms and related challenges, and describe research efforts about the use of machine learning techniques in reducing NIDS false alarms. Section 3 describes our proposed system of *MVPSys* in detail and Section 4 describes a prototype implementation of *MVPSys*. In Section 5, we present our experimental methodology and analyze experimental results. Later, we give a discussion in Section 6 and conclude our work with future directions in Section 7.

2. Background

In this section, we introduce the background of NIDS alarm types, analyze the reasons for causing false positives, present challenges and review related work regarding the use of machine learning in reducing NIDS false alarms.

2.1. NIDS alarms

As shown in Table 1, we summarize four situations regarding the inputs and outputs for an intrusion detection system,