HOSTED BY
ELSEVIER
CrossMark
Available online at www.sciencedirect.com
ScienceDirect
Karbala International Journal of Modern Science 1 (2015) 15–25
http://www.journals.elsevier.com/karbala-international-journal-of-modern-science/

On the designing of two grains levels network intrusion detection system

Safaa O. Al-mamory [a,*], Firas S. Jassim [b]

[a] College of Business Informatics, University of Information Technology and Communications, Iraq
[b] College of Sciences, University of Dyala, Iraq

Received 5 June 2015; revised 27 July 2015; accepted 29 July 2015
Available online 26 September 2015

Abstract

Despite the rapid progress of the information technology, protecting computers and networks remain a major problem for most authors. In this paper, two grains levels intrusion detection system (IDS) is suggested (*fine-grained* and *coarse-grained*). In normal case, where intrusions are not detected, the most suitable IDS level is the *coarse-grained* to increase IDS performance. As soon as any intrusion is detected by *coarse-grained* IDS, the *fine-grained* is activated to detect the possible attack details. Very fast decision tree algorithm is used in both of these detection levels. In order to ensure efficiency of the proposed model, it has been tested on KDD CUP 99 offline dataset and a real traffic dataset. Experimental results demonstrate that the proposed model is highly successful in detecting known and unknown attacks, and can be successfully adapted with packets' flow to increase IDS performance. This article explains how we got a detection rate greater than 93% with an average processing time equals to 3×10^{-6} s per example.

Keywords: Network security; Intrusion detection system; Classification; Very fast decision tree algorithm

1. Introduction

The frequency of computer intrusions has increased rapidly during the last two decades. Intrusion Detection Systems (IDSs) are an essential component of a complete defense-in-depth architecture for network security. They collect and inspect packets, looking for evidence of intrusive behaviors. As soon as an intrusive event is detected, an alarm is raised giving the security analyst an opportunity to react promptly. Unfortunately, most of designed IDSs cannot cope with fast networks.

Although several IDS systems are available, the common objectives of these systems are to reduce the amount of false alarms [1], and to recognize new attacks in order to increase detection ratio. In this paper, the concentration is on detecting known and unknown attacks in fast networks in order to mitigate the influence of the attack by shrinking the time gap between the real attack and its detection.

This paper contribution is to build two grains levels IDS in order to detect abnormal behavior of network traffic and cope with fast networks. It is well known that the intrusion occurrence in networks with respect

* Corresponding author.
E-mail address: salmamory@uoitc.edu.iq (S.O. Al-mamory).
Peer review under responsibility of University of Kerbala.

http://dx.doi.org/10.1016/j.kijoms.2015.07.002

to normal traffic is rare. These motivate us to build the proposed two grains levels IDS. These detection levels are *fine-grained* and *coarse-grained*. In normal case, where intrusions are not detected, the most suitable IDS level is the *coarse-grained* to increase monitoring performance. At the moment of intrusion is detected by *coarse-grained* IDS, the *fine-grained* IDS is activated to detect as most as possible of attack details. Fig. 1 shows the main idea. The *coarse-grained* IDS focuses on five packet features while *fine-grained* IDS works on 20 features. Very Fast Decision Tree (VFDT) [2] algorithm is selected as a fast classifier. The advantages of the proposed system are processing and analyzing of high-speed network traffic, discovering and accurately identifying new attacks to reduce the false alarms to the maximum extent, and detecting the intrusion in real time.

DARPA KDD CUP 99 dataset is used as a benchmark for the proposed IDS, which contains 41 features. As a preprocessing step, we analyzed these features and have selected 20 features having information gain ratio over the average of the dataset. Then, we trained and tested the proposed system. This gave us a detection rate greater than 93% with an average processing time equals to 3×10^{-6} s per example.

This paper is organized as follows: Section 2 reviews related work. Section 3 describes very fast decision tree algorithm. Section 4 states the proposed system. Section 5 presents the experiments and results. Finally, Section 6 concludes this paper.

2. Related work

Nowadays, authors have designed numerous IDSs to detect computer and network intrusions. Several data mining techniques have been used to make networks' intrusions detectable. The first class of approaches uses decision trees (DT) to build attack model. Several variations of decision trees were used such as partial decision tree [3], C4.5 [4], random forest [5], ID3 decision tree [6], and J48 [7]. These decision trees models vary in the splitter measure (i.e. information gain, gain ratio, gini index), pruning technique, branching types, dataset types, etc. The common objective of these decision trees is to iteratively partition the given dataset into subsets where all elements in each final subset belong to the same class. These models have been built from network packets to detect network intrusions with high precision. The main issue with these methods is that they cannot be adaptive with distribution variation in network packets while the proposed system solved this problem by selecting algorithm which works with concept drift.

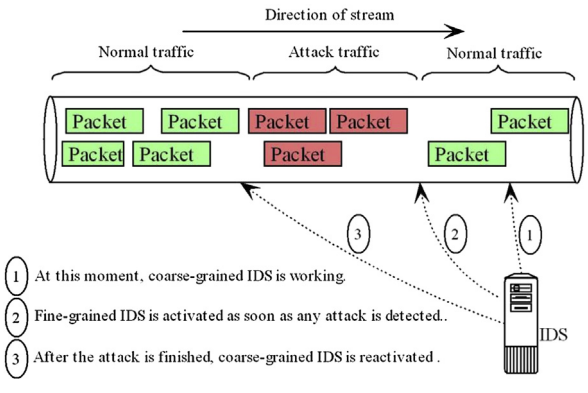

Fig. 1. States the working of the proposed system.

Another class of these approaches has used evolutionary computation [8]. Self-Organizing Map [7] and Multilayer Perceptron (MLP) [7] were trained to recognize normal from abnormal traffic. In addition, genetic programming [9] is achieved very high detection ratio combined with slow model. However, these techniques have performance issues and cannot work in online mode. One of the main goals of this paper is to enhance IDS performance.

Different class of efficient data mining approaches is used to differentiate malicious traffic from normal ones. Bayes network classifier by Staniford et al. [10] is used to calculate the conditional probabilities of several connection features with respect to other connection features. The anomalous connection is determined using these probabilities. SVM is used by Eskin et al. [11], and Honig et al. [12] in addition to their clustering methods for unsupervised learning. The achieved performance was as good as or better than both of their clustering methods. In addition, Fuzzy logic rules by Luo [13] attempted to classify network data. The author verified that the combination of fuzzy logic with association rules and frequency episodes generates more abstract and flexible patterns for anomaly detection. The author approach utilizes fuzzy association rules and fuzzy frequency episodes to extract patterns for temporal statistical measurements at a higher level than the data level.

An additional class of approaches proposed Multi-level IDSs to achieve highest attack detection rate. Multi-level IDS designed by Chen et al. [14] is composed of IDS, firewall, and a report system in order to present a unified report format to the end user. This multi-level IDS supports specific types of these