MAD-IDS: Novel Intrusion Detection System using Mobile Agents and Data Mining Approaches

Imen Brahmi[1], Sadok Ben Yahia[1], and Pascal Poncelet[2]

[1] Faculty of Sciences of Tunis, Tunisia
sadok.benyahia@fst.rnu.tn
[2] LIRMM Montpellier, France
Pascal.Poncelet@lirmm.fr

Abstract. Intrusion Detection has been investigated for many years and the field reached the maturity. Nevertheless, there are still important challenges, *e.g.*, how an Intrusion Detection System (IDS) can detect distributed attacks. To tackle this problem, we propose a novel distributed IDS, based on the desirable features provided by the mobile agent methodology and the high accuracy offered by the data mining techniques.

Keywords: Intrusion Detection System, Mobile Agents, Data Mining Techniques.

1 Introduction

As an important gatekeeper of network, *Intrusion Detection Systems* (IDS)s must have the ability to detect and defend intrusions more proactively in shorter period. However, most current IDSs are centralized and thus a central analyzer presents a favorable target to the attackers.

In this paper, we investigate another way of tackling this problem. Moreover, according to the advantages of mobile agent technology, which includes: *reducing network overload*, *overcoming network latency*, *system scalability*, etc, this technology seems to be very suitable to solve intrusion detection in a distributed environment [2]. Thus, we introduce a new distributed IDS, called **MAD-IDS** (*Mobile Agent using Data mining based Intrusion Detection System*). The MAD-IDS system integrates the data mining techniques and the mobile agent methodology in order to detect both known and unknown attacks.

2 The MAD-IDS system

Figure 1 provides an overall architecture of MAD-IDS. Its distributed structure comprises different agents which are able to move from one station to another, called respectively: *Sniffer, Filter, Misuse Detection, Anomaly Detection, Rule Mining* and *Reporter Agent*.

Each of these agents will be individually described in the following subsections.

2.1 The Sniffer Agent

The Sniffer Agent collects the network packets and stores them in a "*sniffing file*". The benefits of this kind of agent include: *i*) the cloning and the distribution throughout the network; and *ii*) the duplication in order to lighten the network charge.

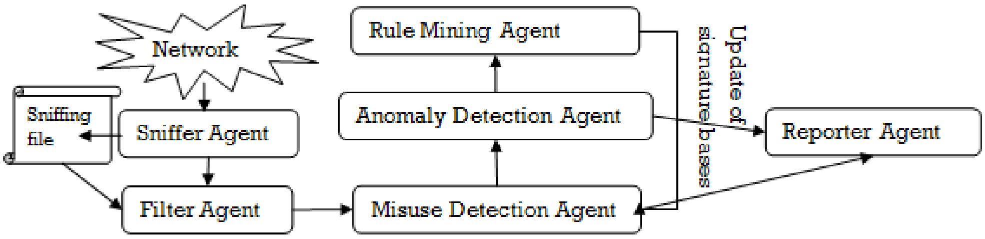

Fig. 1. The architecture of MAD-IDS

2.2 The Filter Agent

The Filter Agent aggregates and merges events stored in the sniffing file. It performs its tasks as a pretreatment phase, which precedes the analysis phase.

2.3 The Misuse Detection Agent

The Misuse Detection Agent detects known attacks in network connections. Hence, if there is a similarity between the filtered packets and attacks signatures, then the agent raises an alert to the Reporter Agent, and then removes these anomalous packets from further analysis.

Although the known attacks are detected, it remains nevertheless the problem of the new attacks detection. One answer to the problem could rely on data mining techniques.

2.4 The Anomaly Detection Agent

The Anomaly Detection Agent provides the combination of distributed IDS with clustering techniques. The clustering-based anomaly detection algorithm is based on the steps described as follows:

Step 1 (Initialization) : Partition the training data into k clusters;
Step 2 (Assignment) : Assign each instance to its closest center;
Step 3 (Updating) : Replace each center with the mean of its members;
Step 4 (Iteration) : Repeat Steps 2 and 3 until there is no more updating;
Step 5 (Anomaly finding) : For each test instance Z:

1. Compute the Euclidean distance between Z an initial cluster C_i;
2. Find cluster C_i that is closest to Z;
3. Classify Z as an anomaly or a normal instance using a pre-defined *Threshold*.

2.5 The Rule Mining Agent

The Rule Mining Agent provides the construction of a summary of anomalous connections detected by the Anomaly Detection Agent. To mine association rules, we apply the *Informative Generic Basis* ($\mathcal{IGB}$) [1]. In addition to the elimination of redundancy, the application of the $\mathcal{IGB}$ basis during an intrusion detection process provides: *the increase of the overall coverage of detectable attacks* and *the maximum convey of useful knowledge*, while being *the information lossless* [1]. Therefore, the database of signatures of the Misuse Detection Agent can be updated regularly by the addition of the extracted rule set.

2.6 The Reporter Agent

The Misuse and the Anomaly Detection Agents send their findings as alerts to the Reporter Agent which transmits to the system administrator.

3 Experimental results

During experiments, we partly use the traffic data DARPA [3]. Table 1 shows the distributions of record types in training and testing datasets, used during our experiments.

Record Type	Training Set	Testing Set
Normal	48886	27322
Intrusion	37804	23009

Table 1. The considered datasets at a glance.

In fact, to evaluate the performance of an IDS, two interesting metrics are usually of use: the *Detection Rate* (DR), which indicates the number of correctly detected intrusions; and the *False Positive Rate* (FR), which calculates the number of instances that were incorrectly considered as attacks.

Figure 2 plots the DR and FR for the considered datasets, using our clustering-based anomaly detection algorithm.

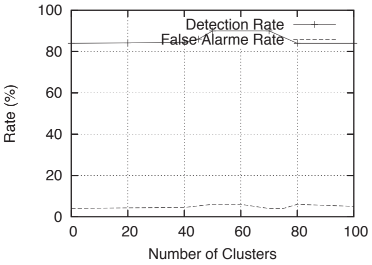

Fig. 2. DR and FR using the clustering-based anomaly detection algorithms.

As shown in Figure 2, the best performance of our anomaly clustering-based algorithm was obtained when DR = 89.89% and FR = 1.00%.

4 Conclusions

In this paper, a novel distributed multi-agent IDS architecture, called MAD-IDS was presented. MAD-IDS integrates the mobile agent methodology and the data mining techniques to accommodate the special requirements in distributing IDS. The preliminary experimental results indicated that the data mining algorithms used in MAD-IDS are feasible for detecting attacks within a distributed environment.

References

1. S. Ben Yahia, G. Gasmi, and E. Mephu Nguifo. A New Generic Basis of Factual and Implicative Association Rules. *Intelligent Data Analysis (*IDA*)*, 13(4):633–656, 2009.
2. N. Jaisankar, R. Saravanan, and K. D. Swamy. Intelligent Intrusion Detection System Framework using Mobile Agents. *International Journal of Network Security and its Applications (IJNSA)*, 1(2):72–88, 2009.

[3] Available at: http ://www.ll.mit.edu/IST/ideval/data/data_index.html