Incremental Hybrid Intrusion Detection Using Ensemble of Weak Classifiers

Amin Rasoulifard, Abbas Ghaemi Bafghi, and Mohsen Kahani

Department of Computer Science and Engineering, Faculty of Engineering
Ferdowsi University of Mashhad, Mashhad, Iran
am_ra84@stu-mail.um.ac.ir, {ghaemib,kahani}@um.ac.ir

Abstract. In this paper, an incremental hybrid intrusion detection system is introduced. This system combines incremental misuse detection and incremental anomaly detection. It can learn new classes of intrusions that do not exist in the training dataset for incremental misuse detection. As the framework has low computational complexity, it is suitable for real-time or on-line learning. Also experimental evaluations on KDD Cup dataset are presented.

Keywords: Hybrid intrusion detection system, Ensemble of Weak Classifiers, Incremental learning, KDD Cup 99 Dataset.

1 Introduction

Misuse detection systems use patterns of well-known attacks or weak spots of the system to identify intrusions. The main shortcoming of such systems[1-3] are the necessity of hand-coding of known intrusion patterns and their inability to detect any future(unknown) intrusions not matched with the patterns stored in the system. Anomaly detection systems, on the other hand, firstly establish normal user behavior patterns (profiles) and then try to determine whether deviations from the established normal profiles can be flagged as intrusions. The main advantage of anomaly detection systems is that they can detect new types of unknown intrusions [4-6].

Weak classifiers are those that obtain 50 percent classification accuracy on its own training data [7]. *Ensembles* are combinations of several models whose individual predictions are combined in some manner (e.g., averaging or voting) to form a final prediction [8].

Several hybrid intrusion detection systems have been proposed for combining misuse detection and anomaly detection [9-17]. We propose a hybrid intrusion detection system which combines the incremental misuse intrusion detection and incremental anomaly detection. In addition, when the intrusion detection dataset is so large that whole dataset can't be loaded into the main memory, the original dataset can be partitioned into several subsets, and then the detection model is dynamically modified according to other training subsets after the detection model was built on first subset.

The rest of the paper is organized as follows: related works is presented in section 2, the proposed incremental IDS is presented in section 3, KDD Cup 99 Dataset is presented in Section 4, experimental evaluation is presented in section 5, comparison

H. Sarbazi-Azad et al. (Eds.): CSICC 2008, CCIS 6, pp. 577–584, 2008.