Available online at www.scienceparks.in

ISSN: 2321-8045
Volume-1, Issue-3, 8th August-2013
Journal DOI: 10.9780/23218045/122013/9

Intrusion Detection System using Self Organizing Map (SOM): A Review

Samarjeet Borah Anindita Chakraborty

Department of Computer Science & Engineering
Sikkim Manipal Institute of Technology
Majitar, Sikkim

Abstract: As there is a rapid expansion in computer usage and its networks the security of computer system has become very essential. Everyday there are new kinds of attacks which are faced by the industries. Many algorithms are been proposed for the intrusion detection system development using artificial intelligence technique. One such algorithm is Self Organizing Map. The aim of intrusion detection system is to identify attacks with a high detection rate and low false alarm rate. The neural network which are capable of supervised learning after the characteristics of the user behavior are able to identify the abnormalities present, but they have their own disadvantages –they are not able to detect new intrusions. Consequently unsupervised learning methods were given a closer look. In the field of intrusion detection system the anomaly detection aspects is very important and thus there are many approaches that are addressing these security issues. The usage of Self Organizing Map (SOM) along with its different SOM algorithm is applied to the problem of host based intrusion detection networks.

Keyword: Intrusion detection system, Neural networks, Self organizing map

I.INTRODUCTION

Intrusion is a sequence of set of related activity which performs to provide unauthorized access to the useful information and the unauthorized file modification which are the cause of harmful activity. Intrusion Detection System (IDS) is a system that identifies the attacks on a network and takes the corrective action in preventing them at real time. These are the set of techniques that are used in detecting suspicious activity both at network and host level and in hybrid level. Host based IDS are used to deploy in an individual host machines and can monitor audit data in a single host. Network based IDS monitors the traffic data during the process of sending and receiving the data by the hosts, Hybrid IDS are used in both the cases. There are two main approaches to design an IDS. 1) Misuse Based IDS 2) Anomaly Based IDS. In misuse based intrusion detection system intrusion are detected by analyzing and looking for the activities that corresponds to known signatures of intrusion or vulnerabilities. A major disadvantage of misuse detection IDS is that there is a lag time between a new treat being discovered and the signature for detecting that treat being applied to our IDS. During this lag time the IDS won't be able to detect new threat. Anomaly based detection is used to identify the activities that may vary from established pattern for user or group[6]. The Self-Organizing Map is a neural network model for analyzing and visualizing high dimensional data. It belongs to the category of competitive learning network. It defines a mapping from high dimensional input data space onto a regular two-dimensional array of neurons. The first use of self organizing map in misuse detection was Kohonen self organizing map as described by Cannady. In the hybrid neural network the output of a Kohonen map provided input to a conventional feed forward neural network and was prototyped to address temporally dispersed and possibly collaborative attacks in a simulated data stream. Temporally dispersed attacks are those which are conducted by a single attacker over an extended period of time, collaborative attackers are those which in concern to achieve single intrusion. Both these types of attacks are very complex. In order to address this, enhanced pattern recognition was designed and was tested to identify series of events that constitute an attack[4].

II.SELF ORGANIZING MAP ARCHITECTURE

The SOM architecture consists of following two levels. The first level is made up of three maps, each map summarizes one feature. The first map in the first level is based on the location from where the connection is made from. The second map is based on the user account from which the log in connection is made. The final map in the first level is based on the connection type. The first level maps provide a concise summary of features in the three input domains with respect to time. The second level map is composed from a single network, receiving input from the three preceding layer maps, and is therefore responsible for combining results into a unified summary of user activity. Based on this result, network the managers gets into the conclusion whether a particular session represents activity that could be considered "abnormal" or not.[2]