Artificial Neural Network Approaches to Intrusion Detection: A Review

IFTIKHAR AHMAD1,2, AZWEEN B ABDULLAH2, ABDULLAH S. ALGHAMDI1
Department of Software Engineering1, College of Computer and Information Sciences,
King Saud University, Riyadh, KINGDOM OF SAUDI ARABIA
Department of Computer and Information Sciences2
Universiti Teknologi PETRONAS, Tronoh, Perak, MALAYSIA
{wattoohu@gmail.com, azweenabdullah@petronas.com.my, abdksu@gmail.com}

Abstract: - Intrusion detection systems are the foremost tools for providing safety in computer and network system. There are many limitations in traditional IDSs like time consuming statistical analysis, regular updating, non adaptive, accuracy and flexibility. It is an Artificial Neural Network that supports an ideal specification of an Intrusion Detection System and is a solution to the problems of traditional IDSs. Therefore, An Artificial Neural Network inspired by nervous system has become an interesting tool in the applications of Intrusion Detection Systems due to its promising features. Intrusion detection by Artificial Neural Networks is an ongoing area. In this paper, we provide an introduction and review of the Artificial Neural Network Approaches within Intrusion Detection, in addition to make suggestions for future research. We also discuss on tools and datasets that are being used in Artificial Neural Network Intrusion Detection Systems. This review may help the researcher to develop new optimize approach in the field of Intrusion Detection.

Key-Words: - Artificial Neural Network, Intrusion Detection System, Anomaly Detection, False positive, False Negative, ROC, Detection Rate, RMSE, IDA, MLP

1 Introduction
The rapid expansion of computer networks and mostly of the Internet has created many stability and security problems [1]. During recent years, number of attacks on network has dramatically increased and consequently interest in network intrusion detection has increased among the researchers [2]. The reliance of private and government organizations is increasing on their computer networks and defending theses system from attack is serious. Intrusion detection systems are the foremost tools for providing safety in computer and network system. Because a single intrusion of a computer network can cause a heavy loss or the consistency of network became insecure [3]. Therefore, accurate detection of network attack is very important. For half a century, developers have protected their systems using rules that identify and block specific events. However, the nature of current and future threats urgently requires the development of automated and adaptive IDS [4]. Therefore, An Artificial Neural Network inspired by nervous system has become an interesting tool in the applications of Intrusion Detection Systems. It supports an ideal specification of an Intrusion Detection System and is a solution to the problems of traditional IDSs. Application of ANN in intrusion detection is an ongoing area [5]. In the following sections, we briefly introduce the areas of IDSs, Artificial Neural Networks, and ANN approaches to intrusion detection. Furthermore, research, development and implementation is presented in terms of NN, dataset, system implementation and testing parameter details. At last an overview of this research area is provided, in conjunction with indications for future areas of study.

2 Intrusion Detection Systems
An illegitimate user that can access network assets and play some thing disaster is known as intruder. An IDS is used to detect illegal access to a computer or network system. There are various methods of responding to a network intrusion, but they all require the exact and suitable recognition of the attack [6]. Dr. Dorothy Denning proposed an intrusion detection model in 1987 which became a landmark in the research in this area. The model which she proposed forms the basic core of most intrusion detection methodologies in use today [7]. The intrusion detection systems can be classified into three categories as host based, network based and vulnerability assessment based [8]. There are many approaches that are being used to accomplish the desirable elements of an intrusion detection system like anomaly detection, misuse detection, combined anomaly/misuse detection, pattern recognition and networking monitoring.

3 Artificial Neural Networks
The first artificial neuron was formed in 1943 by the neurophysiologist Warren McCulloch and the logician Walter Pits [9]. An artificial neuron is a processing element with many inputs and one output. An artificial neural network consists of a group of processing elements that are greatly interconnected and
convert a set of inputs to a set of preferred outputs [10]. The result of the transformation is determined by the characteristics of the elements and the weights associated with the interconnections among them. By modifying the connections between the nodes the network is able to adapt to the desired outputs [11]. It offers the potential to resolve a number of the problems encountered by the other current approaches to intrusion detection. Artificial neural networks are alternatives. The first advantage in the use of a neural network in the intrusion detection would be the flexibility that the network would provide. A neural network would be capable of analyzing the data from the network, even if the data is incomplete or unclear. Similarly, the network would possess the ability to conduct an analysis with data in a non-linear fashion. Further, because some attacks may be conducted against the network in a coordinated attack by multiple attackers, the ability to process data from a number of sources in a non-linear fashion is especially important. The natural speed of neural networks is another advantage [3].

3 ANN approaches to ID

3.1 Approach-1

One of the first works to intrusion detection by NN was performed by Ryan et al. in 1998 [12]. They trained and tested an offline NNIDS on a system of ten users. They used 2-Layer MLP architecture for their system and backpropagation for training purpose. The data source for training and testing was operating system logs in UNIX environment. The result parameters to evaluate the performance of the system were false positive and false negative. They implemented their system in the PlaNet Neural Network simulator. A systematic review of their work is given in the table.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data Source</th>
<th>NN Structure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ryan et al.</td>
<td>1998</td>
<td>Operating System Logs</td>
<td>2-Layer MLP</td>
<td>7% FP, 4% FN</td>
</tr>
</tbody>
</table>

Table 1

3.2 Approach-2

Another attempt in the same field was made by Cannady in 1998 [6]. He also used the 2-Layer MLP architecture for his system and backpropagation for training purpose. The data source for training and testing was Network Packets collected by Real Secure. Nine of the packet characteristics of network data were selected and presented to the MLP network which has four fully connection layers. He used RMSE parameter for training and testing data for performance measuring. A systematic review of his work is given in the table.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data Source</th>
<th>NN Structure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cannady</td>
<td>1998</td>
<td>Network Packets Collected by Real Secure Network Monitor Software</td>
<td>2-Layer MLP</td>
<td>RMSE of 0.0582 for Training Data, RMSE of 0.069 for Test Data.</td>
</tr>
</tbody>
</table>

Table 2

3.3 Approach-3

Ghosh et al. in 1999 [13] presented a host based IDS that focused on building program profiles and used these program profiles to identify normal software behavior and malicious software behavior. The system was trained and tested on SUN platform and use Basic Security module (BSM) as source of data. Input data were extracted from BSM and a distance metric, which constituted input vectors of the NN. The IDS presented was a single hidden layer MLP. The number of input nodes was equal to the number of exemplar strings. Lucky Bucket algorithm is used to capture the temporal locality of anomalous events. Performance analysis was done with DARPA database. Ghosh et al. in 1999 [14] also used Elman Networks and results of their works are shown in the table below.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data Source</th>
<th>NN Structure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ghosh et al.</td>
<td>1999</td>
<td>Sun’s BSM</td>
<td>2-Layer MLP</td>
<td>Anomaly Detection: 2.2% FP, 22.7% FN Misuse Detection: 18.7% FP, 9.1% FN</td>
</tr>
<tr>
<td>Ghosh et al.</td>
<td>1999</td>
<td>Sun’s BSM</td>
<td>Elman Networks</td>
<td>No FP, 22.7% FN</td>
</tr>
</tbody>
</table>

Table 3

3.4 Approach-4

Another work is one by Rhodes et al. in 2000 [15], they proposed to use of self-organizing neural networks to recognize anomalies in network data stream. Unlike from other approaches which use self organizing maps to process entire state of a network or computer system to detect anomalies, proposed system breaks down the system by using collection of more specialized maps. A monitor
stack was constructed and each neural network become kind of specialist to recognize normal behavior of a protocol and raise an alarm when a deviation from normal profile occurs. The test intrusions were buffer overflow attempt. The review of their works is given in the table.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data Source</th>
<th>NN Structure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhodes et al</td>
<td>2000</td>
<td>Network Packets [buffer overflow]</td>
<td>SOM</td>
<td>D.R (57%) BIND server &amp; rotsh exploit</td>
</tr>
</tbody>
</table>

Table 4

3.5 Approach-5
Lippmann and Cunnigham of MIT Lincoln Laboratory in 2000 [16] conducted a misuse detection model with neural networks, by searching attack specific keywords in the network traffic. They used a MLP network to detect Unix-host attacks, and attacks to obtain root-privilege on a server. The data that they presented to the neural network consisted of attack-specific keyword counts in network traffic. Two neural networks were used in the system, one for providing an attack probability and one for classifying attacks. A two-layer perceptron was designed with k input nodes, 2k hidden nodes and 2 outputs (normal and attack) and the training algorithm used in the system was backpropagation. The results of their work are in given below.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data Source</th>
<th>NN Structure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lippman et al</td>
<td>2000</td>
<td>Network Packets</td>
<td>2-Layer MLP</td>
<td>One False Alarm per Day 20% false Negative D.R 80%</td>
</tr>
</tbody>
</table>

Table 5

3.6 Approach-6
In another study Zhang et al. 2001 [17], statistical analysis was used in conjunction with MLP networks. System is a distributed hierarchical application in the sense that system consists of hierarchy of Intrusion Detection Agents (IDAs) at multiple tiers where each tier corresponds to different network scope. IDAs are IDS components that monitor the activities of a host or a network. An IDA, which consists of the following components: the probe, the event pre-processor, the statistical processor, the neural network classifier and the post processor. Probe collects network traffic and abstracts it into statistical variables. Event pre-processor collects data from probes and other agents and formats it for the statistical analyzer. Statistical model compares the data to the previously compiled reference model which describes the normal state of the system. A “stimulus vector” is formed and forwarded to the NN. Neural network analyzes the vector and decides whether it is anomalous or normal. Post processor generates reports for the agents at higher tiers or it may display the results through a user interface. Backpropagation, perceptron, perceptron-backpropagation hybrid, fuzzy ART MAP, radial-basis function networks with 2-8 hidden nodes were tested. The experimental test bed consisting of 11 workstations and 1 server was built by using OPNET network simulation software. UDP flooding attack was simulated within the test bed. The review of their work is given in the table.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data Source</th>
<th>NN Structure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang et al</td>
<td>2001</td>
<td>Network Packets Generated by OPNET Network Simulation Software</td>
<td>Backpropagation, Perceptron, Perceptron-Backpropagation Hybrid, Fuzzy ART MAP, Radial Basis Function Networks</td>
<td>BPROP &amp; HPBPROP performed better than Perceptron, Fuzzy ART MAP, Radial Basis Function networks RMSE&lt;0.05 Statistical Analysis</td>
</tr>
</tbody>
</table>

Table 6

3.7 Approach-7
Lee and Heinbuch in 2001 [18] utilized an experimental IDS with a hierarchy of neural networks. Each of the neural networks in the hierarchy focused on different portions of nominal TCP behavior. Portions of these observed TCP behaviors are connection establishment, connection termination and port usage. System was trained to detect three kinds of attack, which are SYN flood, fast SYN port scan, and stealth SYN port scan. Backpropagation learning algorithm was used to train system. Input vectors to each of the neural networks were generated randomly. The review of their work is given in the table.
3.8 Approach-8
In 2002 Jirapummin et al. [19] presented an alternative methodology for both visualizing intrusions by using SOM and classifying intrusions by using Resilient Propagation, Neptune attack (SYN flooding), Portswep and Satan attacks (port scanning) were selected from KDD Cup 1999 data set. For RPROP, 3 layer NN is utilized with 70 nodes in first hidden layer, 12 neurons in second hidden layer and 4 neurons in the output layer. The transfer functions for the first hidden layer, second hidden layer and the output layer of RPROP were tan-sigmoidal, log-sigmoidal and log-sigmoidal respectively. IDS results are shown below.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data Source</th>
<th>NN Structure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jirapummin et al.</td>
<td>2002</td>
<td>KDD Cup 1999 [TCP SYN &amp; Port Scanning]</td>
<td>3-Layer RPROP with SOM</td>
<td>D.R 90% 5% FP 10% FN</td>
</tr>
</tbody>
</table>

Table 7

3.9 Approach-9
Bivens et al. [20] in 2002 proposed a neural network model for a network-based intrusion detection system. Their anomaly detection system had used MLP network for detection. System read tcpdump data. The review of their work is given in the table.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data Source</th>
<th>NN Structure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bivens et al.</td>
<td>2002</td>
<td>DARPA 1999 [DOS,DDOS &amp; Port attacks]</td>
<td>SOM for Clustering MLP for Detection</td>
<td>76% FP No FN</td>
</tr>
</tbody>
</table>

Table 8

3.10 Approach-10
Another study was made by Mei-Ling Shyu et al. in 2003 [21]. They used KDD Cup 1999 as a data source for training and testing of their system. The neural network used by them was PCC. A review of their work is given in the table.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data Source</th>
<th>NN Structure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mei-Ling Shyu et al.</td>
<td>2003</td>
<td>KDD Cup 1999</td>
<td>PCC</td>
<td>DR → 95%  FA → 5%</td>
</tr>
</tbody>
</table>

Table 10

3.11 Approach-11
Yao Yu et al. in 2004 [22] worked on FTP brute force attacks. They used samples that were collected from local network traffic. They used Hybrid BP/CNN as neural network architecture. A ROC curve is used to evaluate the system performance by them. A review of their work is given in the table.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Data Source</th>
<th>NN Structure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yao Yu et al.</td>
<td>2004</td>
<td>FTP brute force attacks samples from LAN</td>
<td>Hybrid BP/CNN</td>
<td>ROC</td>
</tr>
</tbody>
</table>

Table 11

3.12 Approach-12
Morteza Amini et al. in 2005 [23] worked on intrusion detection by using ART1 and ART2. They compare both NN and showed that ART-1 is better in performance wise but ART-2 is better in response wise. They also used standard data set KDD Cup 1999.In another work in 2006 [24], they worked on IP, TCP, UDP, and ICMP packets in the LAN environment. A review of their works is given in the table.
Different researcher used different neural network architecture to implement their proposed systems in the field of intrusion detection. Mostly parameters for testing their results were false positive, false negative, detection rate and ROC. They used MATLAB, PlaNet, OPNET, JOONE, URANO, NeuralWorks simulators and some are developed in a personalized way. No doubt NN minimize various flaws in traditional IDSs like time consuming statistical analysis, regular updating, non adaptive, efficiency, accuracy and flexibility. But it also suffers many problems in the research of intrusion detection. There are two types of training/learning supervised and unsupervised that are used in NIDS. The first involves training overheads (time consuming, regular update and unable to detect novel attack) while the second one is not much more optimized in performance (false positive, false negative, detection rate and MSE). Application of ANN in intrusion detection is an ongoing area and is limited to academic research till now.

5 Future suggestion

So presently a research is required that will develop an adaptive, flexible and optimize neural network intrusion detection system that will provide the potential to identify network activity in a robust way. By taking advantages of both training/learning supervised and unsupervised networks, we might be able to develop new optimized approach that may help and guide the security implementer and researcher in the field of intrusion detection in the future.

References:


