International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering Vol. 2, No. 1, 2015, pp. 11-18.

ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

The Survey of Artificial Neural Networks-Based Intrusion Detection Systems

Masoomeh Sharifzadeh[a]

[a] *Young Researchers and Elite Club, Marvdasht Branch , Islamic Azad University, Marvdasht , Iran.*

Abstract

Artificial Neural Networks (ANN) is methods which can help Intrusion Detection Systems (IDS) through their Learning opportunities. There are different kinds of IDS based on various artificial intelligence methods; one of these methods is ANN which some scholars work on. Two important characters of IDS are speed and accuracy. Our goal in this paper is to survey IDS methods based on ANN; a brief description of ANN, its working operation and a description of IDS are expressed firstly; later in Methods section, some methods are briefly described and then in comparison section these methods compered which cause a better understanding of benefits and defects of each one. At last consequences are obtained by considering the whole mentioned methods.

Keywords: Artificial Neural Networks (ANN), Intrusion Detection Systems (IDS)

Introduction:

Nowadays networks play an essential role in any organization and society. With the rapid development of internet, the security of internet becomes a vital problem. In order to prevent profiteers from intruding internet systems, intrusion detection becomes extremely critical. Intrusion detection systems detect and prohibit the abnormal actions [4]. So we should use some methods to safe our networks from invalid accesses. In the context of computer networks, an IDS can roughly be defined as a tool designed to detect suspicious patterns that may be related to a network or system attack. Intrusion detection (ID) is then a field focused on the identification of attempted or ongoing attacks in a computer system (host IDS—HIDS) or network (network IDS—NIDS) [7]. Traditional intrusion prevention techniques, such as firewalls, access control or encryption, have failed to fully protect networks and systems from attacks [1]. In spite of being some methods which help us to achieve this goal some attacker can defeat these methods and inter to the network, so we need another system to recognize them. The neural network can be trained to classify the behavior of system into normal and abnormal [4].

In this paper we have introduced some methods of intrusion detection systems using different kinds of artificial neural networks and compare them with each other. The remainder of this review is organized as follows. Section 2 defines a background of artificial neural networks and intrusion detection. Section 3 describes some IDS based on ANN and in section 4 these methods are compared with each other, and finally in section 5 we conclude.

Background

2-1. Artificial Neural Network

An Artificial Neural Network (ANN) is one of the supervised machines learning technique which consists of a collection of processing units called neurons that are highly interconnected in a given topology. ANNs have the ability of learning-by-example and generalization from limited, noisy data; they have, hence, been successfully employed in a broad spectrum of data intensive applications. [1] Artificial neural networks have got widespread parallel structures. Networks adopt based on correspondence between input and target, until the network output and target become quite alike. After training network, applying a specific input to it, leads to a specific output. Each neuron has a bias and a transfer function. We have different kinds of ANN such as MLP, RBF, ELMAN, RNN which are used in the intrusion detection.

2-2. Intrusion Detection

To create a perfect security for networks, in addition to use of Intrusion prevention equipment such as fire walls, other systems called intrusion detection systems (IDS) are needed to recognize attackers if they pass fire walls, antivirus and … and enter the system and find a solution to deal with them.

IDS have the duty of reconnoiter and distinguish any unauthorized use with external and internal users. Now there are different kinds of IDS as software and hardware systems. Speed and accuracy are the advantages of hardware systems, but easy use and adaptability with different operating systems of software systems are the abilities that cause it to be more general. Intrusion is a set of unlawful acts which endangers accuracy and confidentiality or access to a source. Attackers usually benefit from software defects, break passwords, eavesdropping network traffic and weaknesses in network design, services or network computers to percolate systems and networks.

IDS can be classified due to three aspects, detection method, architecture and how it responds to attackers. Different methods of IDS are: misuse detection, anomaly detection and signature-based detection.

Methods: Intrusion detection systems based on ANN

3-1. Method1: Neural networks-based detection of stepping-stone intrusion

Ching et al. [2] introduced a method based on neural networks to detect stepping-stone intrusion. When network intruders launch attacks to a victim host, they try to avoid revealing their identities by indirectly connecting to the victim through a sequence of intermediary hosts, called stepping-stones. They tried to detect such a long connection chain by estimating the number of stepping-stones. They proposed two schemes, in the first schemes they used eight packet variables and in the second they bunches a sequence of consecutive packet round-trip times. A rightful user accesses the target host through a direct connection and seldom through a long chain with more than two connections. So by estimating the downstream length from the monitoring computer to the victim computer, could recognize whether an attack happens. Connection c_i is a connection from computer host H_i to host H_{i+1}. An attacker in computer H1 launches an attack to the victim computer H_n via the stepping-stones H_2, . . ., H_i,. . . ,H_{n-1}.The sensor is the host where the detecting program is loaded to monitor the connection chain. In this approach they collected the TCP (Transmission Control Protocol) packets and added the computed packet round-trip times from the time stamps for neural network development. When a packet is send in a network, there is a time distance between sending the packet and receiving its Echo. Which is called RTT. Four factors impress RTT.

$$RTT=t_1+t_2+t_3+t_4 \quad (1)$$

Propagation delay (t_1), processing delay (t_2), queuing delay (t_3) and transmission delay (t_4). Physical distance determines t_1. Problem of TCP packet matching is needed for computing RTTs. When we send some packet and receive some echo, each send should be matched with one echo. This is hard when there is a many to many relation. For solving this problem they use a Queue. Observing connection chain had been required for times ago, but in this method this problem has been resolved.

They do some test to collect information as input to neural network. They used two experimental groups. a) a three LAN connection chains , b) a three Internet connection chains. RTTs related to group a are small and in contrast in group b are longer. They have improved two scheme ,I and G, with neural networks .In scheme I, they used eight packet variables of an particular packet each variable is an input to ANN so there are eight nodes in the input layer to estimate the number of stepping-stones .

Scheme G: they collected a sequence of n RTTs which are sequential and gave them as input to an ANN. They used three different transfer functions (1) sigmoid: is a continuous, monotonic mapping of the input into a value between 0 and 1. (2) TanH (Hyperbolic Tangent): is a continuous, monotonic mapping into the range _1 to 1. (3) Sine: takes the trigonometric sine of the input modified by Gain.

3-2. Method2: Intrusion detection using ANN and fuzzy clustering

Wang et al. [3]proposed a method based on ANN and fuzzy clustering which improved ANN-based IDS and try to solve these two problem.1)low recognition accuracy 2)stability of recognition. Their technique has three stages. Stage 1) fuzzy clustering technique which produce different training subset. 2) An ANN for each of training set. 3) For fending errors of different ANN, they introduce a fuzzy aggregation module to learn later and combining different ANN results.

Using ANN-based techniques alone as IDS do not act well especially on data with low abundance. So they try to use a hybrid method.

There are two phase, learning phase and testing phase.
Learning phase consist of three main stage: 1) For each arbitrary data set ,DS, at first it disparts to two set, learning set TR and testing set TS. And then TR is divided to different subset TR_1, TR2,…,TR_n by using fuzzy clustering Module. 2) For each training set TR_i, an ANN model, ANN_i is trained with a special learning algorithm for formation of k ANN model with different bases. 3) To reduce errors each ANN_i is simulated with the whole TR and results are received. And then they used membership grades produced by fuzzy clustering to combine the results.
In test phase, they gave the test data as input to k different ANN_i, and based on these outputs final results were gained by using fuzzy aggregation.

Fuzzy clustering module

This method should creates groups in a way that data in a category have the same kind and data center but there should be dissimilarity between different category. There are two clustering technique: hard clustering and soft clustering. They used a soft clustering technique named as fuzzy c-means clustering.
To train they used feed forward neural network. And the activation function was bipolar sigmoid. Each ANN_i trains each of TR_k by using a back propagation algorithm. The goal of fuzzy aggregation module is to collect the results of different ANNis and reduce error detections. Since each ANN_i in module ANN just learn from TR_i. And because errors are nonlinear, they used a new ANN as follow:

1) The whole TR is as an input to each of ANN_i and results are received.
2) The input for new ANN is constructed by using the results of previse stage and membership grade of each cluster as follow:
3) The new ANN is trained with Y_{input} as input and Y^{TR} as output.

They considered four intrusions:

1) Denial of service (DOS)
2) Probe(PRB)
3) Remote to Local (R2L): unallowable access from a remote device for exploitation of machine vulnerability.
4) User to Root (U2R): unallowable access to root for using system's abilities.

There are four items to evaluate the accuracy of IDS. True positive, shows that IDS can recognizes a special attack has occurred. A true negative, shows that IDS does not make mistake in recognizing common errors. False positive, shows that an attack is recognized with IDS but in fact does not occurred. A false negative show that IDS cannot recognize intrusion after a special attack has occurred and this is because of lack of information about an intrusion type.' However as the number of instance for the U2R, PRB, and R2L attacks in the training set and test set is every low, these quantities is not sufficient as a standard performance measure' [2]. So they introduce precision, recall and f-value which do not related to the size of test and training samples. That were defined as follow:

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FP} \quad (3)$$

$$F - Value = \frac{(1+\beta^2)\times Recall\times Precision}{\beta^2\times(Recall\times Precision)} \quad (4)$$

And also they computed percentage of training successfully for measuring detection constancy for ANN-based IDS.

$$Percentage\ of\ training\ successfully = \frac{The\ numbers\ of\ training\ successfully}{The\ number\ of\ training\ result\ \&\ discussion} \quad (5)$$

3-3. **Method3: A hybrid RBF/Elman neural networks for IDS**

Tong et al. [4]. They used a hybrid RBF/ELMAN neural network for detecting both anomaly and misused detection. RBF network are used as a Real-time classification and ELMAN network is used for memory reconstruction of past events. RBF networks use a local exponential function and needs less time than MLP. They used a BSM audit data format, which consist of seven fields, and these fields contain information about each software behavior. They gained the network input through following steps

1) Get the trace of privileged processes' system calls.
2) Slide a window of size "L" across the trace, recording each unique sequence of length "L" that is encountered.
3) Encode each unique sequence into a string consisted of "0" and "1", recording each string.
4) Pick up several "exemplar" strings from the strings, each "exemplar" string is different from each other, its value of the distance metric with other strings is bigger.
5) To encode a string of data, the distance metric is used to measure the distance from the data string to each of several "exemplar" strings.

They used a three layer RBF neural network which each hidden parts execute a radial activated function. 'The Gaussian activation function is used as activation function [3]. The input of RBF is nonlinear and its output is linear.

Each ELMAN network has a set of context nodes. Each context node received its input from a node in hidden layer and gets its output to each node in the same layer. Since content nodes depend on activations on hidden layer from previous input, content nodes maintain the state s information between inputs. In a hybrid RBF/ELMAN network the output of RBF is used as the input of ELMAN so the number of RBF s output and ELMAN input should be equal. When the RBF classifies an input the result is stored with ELMAN. In implementation, content values are set to zero. Weights of recurrent connections between content nodes and hidden nodes are constant values, 0 and 1.' Both the input nodes and context nodes activate the hidden nodes, and then the hidden nodes feed forward to activate the output nodes' [4]. Content nodes at time t+1 contain values which are obtained from time t.

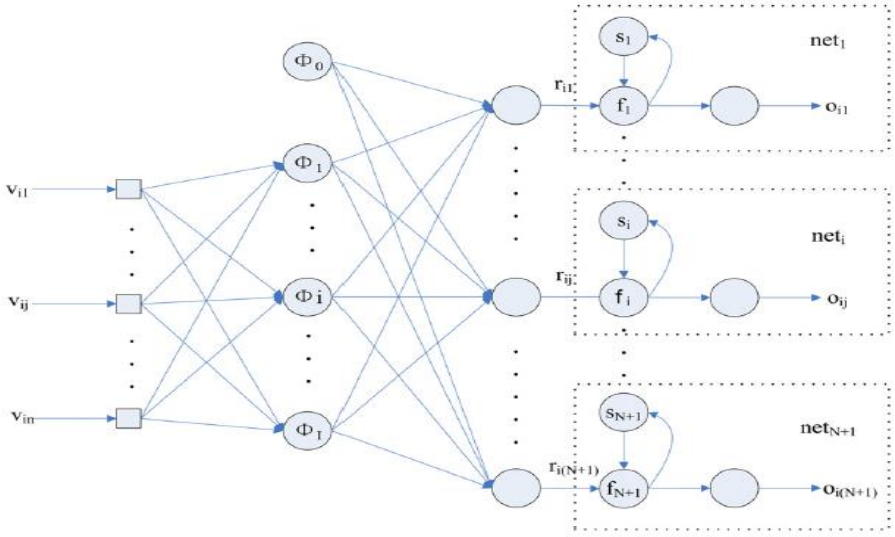

Figure (1). The hybrid RBF/Elman neural network model. [4]

When RBF network recognized a misuse intrusion, the ELMAN network can collect a large value in its context unit quickly [3]. And if the output of RBF is a normal action, ELMAN network reduces the value of content nodes to zero. If the value of content nodes decrease from a threshold, the hybrid network recognize that an attack is occurred. The advantage of ELMAN network is that allows random misuse behavior which is expected through system operation. A hybrid model can be extended and updated easily. If a new attack event occurred, the pattern classification module that adapts the new event is trained and a new interface for the event is produced. A new memory module is used to alert this intrusion events connect to the interface. [3] The system sensitivity is changed with changing recurrent connection weight and the threshold of the output nodes.

3-4. Method4: Neural Networks for Intrusion Detection in MANET

Moradi et al. [5] .They used a feed forward neural network for detecting DOS attacks. They proposed a model through 4 steps:

1) They had studied DOS attacks and check the parameters in network which were impressed by this attack. They found four parameters, Packet loss (PL): The numbers of packets are lost in every time. Packet sending (PS): The average of packets is sent to a host node from the whole network. Packet receiving (PR): number of packets received to the host node. Energy consumption (EC): DOS attacks consume a lot of energy of host node.
2) They simulated a MANET network with 6 nodes and applied DOS attacks to that.
3) They collected different values for above four parameters as inputs
4) They design a feed forward neural network to learn to states of nodes. They used TANSIG and LOGSIC as transfer function.

3-5. Method5: Intrusion detection using reduced-size RNN based on feature grouping

Sheikhan et al. [6].They used a reduced-size structure of Recurrent Neural Network (RNN), based on the grouping of features to detect misuse behaviors. The speed is improved by reducing the size of RNN. They used International Knowledge Discovery and Data mining group (KDD) to collect data set. Their input features were divided into four categories. Basic features (B-F), content features (C-F), time-based traffic features (TT-F), and host based traffic features (HT-F). The RNN had five outputs, one of them showed normal class (no attack). The others represent the type of attacks which was detected: Denial-of-Service (DOS), Probe, Remote-to-Local (R2L), and User-to-Root (U2R). The connections between the input nodes and first hidden layer nodes were based on the kind of categorization of features.

Comparison

Now we compare the described methods as for table (1).In the first method different kinds of ANN were used and the best results had been achieved by means of sigmoid function as transfer function and Delta rule as learning rule. And they used correct rate for improving the results. Shame G in comparison with I had better results because factors such as network bandwidth and network congestion did not affect RTTs so much. An important benefit of this method is that does not need to monitor the network. In the second method trying to improve the results of neural network in a way that total inputs were classified and this cause more special input, so the ANNs became smaller. The topology of each ANN became more particular for their special inputs and can learn the inputs better, quicker and more accurate. At last by aggregating results of all ANNs, a unit output has been. Advantages of the second method in compere with the first one is that in addition that each ANN learn it's input in terms of a property but the final result was achieved by synthesizing of all results and was unique. The precision of this method is high particularly for Dos attacks. In the third method was tried to strengthen the used neural network by connecting an Elman network. An RBF network was used for both anomaly detection and misuse

detection which its speed is more than MLP and has the ability of modeling complicated maps. The RBF networks should learn three parameters, the center of radial basis function, the variance of radial basis function and weights. The self-organized algorithm was used for choosing center of radial basis function and LMS for training the network. Important benefits of Elman network is that can detect a sequence of misuse occurrences and then promulgate an attack. So did not know an accidental misuse as an attack. Another Substantial property is the power of expansion the network. But the problem of this method is that the sensitivity of network change by changing the weights of recurrent connection and threshold. If Dos attack is consider this method has high detection power and its false positive rate is zero.

The fourth method pays to detecting Dos attacks in MANET networks. Good results were provided despite the simplicity.

The fifth method is like the other ones and the important advantage is categorization of features as inputs and this caused the results better. In compare with MLP and ELMAN networks it is quicker.

Table (1): comparison of various ANN based IDS

methods	**ANN**	**Transfer function**	**Learning rule**	**attacks**
method1	Feed forward	sigmoid	Delta rule	stepping-stone
method2	Feed forward and fuzzy clustering	-	-	Dos
method3	RBF/ELMAN	Radial basis function	LMS algorithm	anomaly detection ,misuse detection (Dos)
method4	Feed forward	TANSIG &LOGSIC	-	DOS
method5	Recurrent Neural Network (RNN)	-	-	DOS, Probe, R2L,U2R

Conclusion

From comparing all the methods by considering the inputs to ANN we understand when we specific the inputs, the results of ANN are better. In each method they tried to create ANNs for a group of inputs which had the most similar properties.

References:

S. Xiaonan and W. Banzhaf," The use of computational intelligence in intrusion detection systems: A review", Applied Soft Computing 10 (2010) 1–35.

H. Ching and S. Hsuan Stephen Huang, "Neural networks-based detection of stepping-stone intrusion", Expert Systems with Applications 37 (2010) 1431–1437.

G. Wang and J. Hao and J. Mab and L.Huang, "A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering", Expert Systems with Applications 37 (2010) 6225–6232.

X.Tong and Z. Wang and H. Yu," A research using hybrid RBF/Elman neural networks for intrusion detection system secure model", Computer Physics Communications 180 (2009) 1795–1801.

Z. Moradi and A.Rahmani and M.Teshnehlab," Implimantaion of Neural Networks for Intrusion Detection in MANET", PROCEEDINGS OF ICETECT 2011.

M.Sheikhan and Z. Jadidi and A. Farrokhi ," Intrusion detection using reduced-size RNN based on feature grouping", Neural Comput & Applic DOI 10.1007/s00521-010-0487-0

E. Corchado and A. Herrero, "Neural visualization of network traffic data for intrusion detection ", Applied Soft Computing 11 (2011) 2042–2056