Increasing the Rate of Intrusion Detection based on a Hybrid Technique

Khattab M. Ali Alheeti
Computer Sciences, College of Computer
University of Anbar -Iraq
khattabheeti@yahoo.com

Laith Al-Jobouri
Ministry of Science and Technology-Iraq
lamoha@essex.ac.uk

Professor Klaus McDonald-Maier
University of Essex Technologies
Colchester, United Kingdom
kdm@essex.ac.uk

Abstract – This paper presents techniques to increase intrusion detection rates. Theses techniques are based on specific features that are detected and it's shown that a small number of features (9) can yield improved detection rates compared to higher numbers. These techniques utilize soft computing techniques such a Backpropagation based artificial neural networks and fuzzy sets. These techniques achieve a significant improvement over the state of the art for standard DARPA benchmark data.

Keywords: Intrusion Detection, Soft Computing, Fuzzy set, Neural Networks.

I. INTRODUCTION

Intrusion detection systems (IDS) have been an active area of research and development for the past few decades [1]. An intrusion can be defined as "an act of a person of proxy attempting to break into or misuse a system in violation of an established policy" [2]. IDS are hardware and/or software systems for monitoring and detecting data traffic or user behavior to identify attempts of illegitimate access and system manipulation through a network by malware and/or attackers (crackers, or disgruntled employees). IDS have been used to protect information systems along with prevention ased mechanisms such as authentication and access control [3]. This paper essentially addresses the issue of classifying vital input features for intrusion detection (ID). The ability to identify the important inputs and redundant inputs of a classifier lead directly to reduced size, faster training and possible more precise results, it is critical to be able to identify the important features of network traffic data for intrusion detection in order for the IDS to achieve best performance. The data we used in our experiments originated from MIT's Lincoln Lab. It was developed for intrusion detection system evaluations by DARPA and is considered a benchmark for intrusion detection evaluations [4]. We executed experiments to rank of importance of input features for each of the five classes (Normal, Probe, DOS, U2R, and R2L) of patterns in the DARPA data. It is shown that using only the significant features for classification gives good accuracies and, in certain cases, reduces the training time and testing time of the neural network based intelligent classifier.

With respect to the research on IDS's, the intrusion detector neural networks attracted a growing number of computer scientists and they have proposed several different intelligent systems. Various types of classifier were used by the researchers to detect intrusions; this work is concerned with using a combination of artificial neural networks (ANNs) with fuzzy systems as classifiers. Related works are:

Chandrashekhar et al - proposed a new hybrid technique by utilizing data mining techniques such as fuzzy C means clustering, Fuzzy neural network / Neurofuzzy and radial basis function support vector machines for strengthening the intrusion detection system. The proposed technique has five major steps in which, the first step is to perform the relevance analysis, and then input data is clustered using Fuzzy C-means clustering. This achieved good results with 98.94% accuracy in the case of DOS intrusion and in other cases such as PROBE, RLA and URA; 97.11%, 97.78% and 97.80% respectively [5].

Mostaque et al presented several research papers outlining the foundations of intrusion detection systems, suitable methodologies and good fuzzy classifiers using genetic algorithms which are the focus of current development efforts and the solution of the problem of Intrusion Detection System to offer a real world view of a intrusion detection [6].

Kaur et al presented an approach that utilizes fuzzy if-then rules to detect known and unknown attacks, i.e. sequential multilevel misuse along fuzzy if-then rules. The performance of this algorithm has utilized the KDD'99 data set [7].

EL Kadhi et al, suggested Artificial Neural Networks (ANN) architecture for decision making within intrusion detection systems. The idea is to manage generating a huge test set including sort events as inputs and the corresponding signatures as possible outputs [8].

A new model was designed and implemented to increase the detection rate. Intrusion depends on series of 41 different factors, in this paper we will use a reduce set of only 8 factors, which employs a combination of neural networks and fuzzy sets [9].

The presented work detects attacks (Intrusion) through building artificial detection system using feed forward neural networks to detect attacks with a low false negative rate (which is the most important point), and low false positive rate. To do so, two feed forward neural network architectures (one for non fuzzified data, the other for fuzzified data) are suggested, and their behaviors in detecting the attacks studied. Here, the suggested IDS not only has the ability to distinguish if the access is normal or an attack, but also capable of distinguishing the attack type [10].

II. INTRUSION DETECTION SYSTEMS

As a result of increases of intruders on computers and networks, improved and essentially automated surveillance has become a necessary addition to information technology security. The main difficulty is a distinguishing between natural connections and abnormal connections in computer networks due to the significant overlap in the monitoring data. This detection process can generate false alarms resulting from the use of intrusion detection based on the (Anomaly Intrusion Detection Systems) [4]. The Fuzzy Set can be employed to reduce the rate of false alarm, where the degree of relationship to the use of any process for the separation of this overlap could be used to define normal and

abnormal behavior in distributed systems. For that data fuzzification is needed before classification.

IDS were proposed to complement prevention-based security measures. An intrusion is defined to be a violation of the security policy of the system; intrusion detection thus refers to the mechanisms that are developed to detect violations of system security policy [11]. Intrusion detection is an important part of a network security system. It complements existing security technologies, such as firewalls, by providing crucial information to the network administrators about attacks and intrusions that may be undetected by existing security technologies. IDS can be divided into two types' 1) *anomaly* detection and 2) *misuse-* or signature-based detection. Misuse detection systems match incoming network traffic to a database of known intruder signatures to detect intrusions. While a misuse detection system enjoys a high rate of success at detecting known attacks, they are ineffective in detecting new or unknown attacks. On the other hand, anomaly detection systems create a normal profile of the network or host under observation and flag deviations from the normal profile as probable intrusions. As these systems predict anomalous behavior, they have the advantage of being able to detect new and novel attacks [12].

We can, of course, obtain labeled data by simulating intrusions in a network. However, then we would be limited to the set of known attacks and we would not be able to detect new attacks. As a result, it has been seen that currently available commercial solutions to detect intrusions in gigabit networks can detect less than half of the attacks directed at them [13] at gigabit speeds. The motivation behind our intrusion detection structure is simple: we are using a sampling bas technique to reduce the number of features that needs to be processed, thereby enabling anomaly detection in high-speed networks. In typical cases, sampling would lead to loss of information, leading to inaccurate predictions and/or false alarms. In order to avoid such a state, the proposed predicative model needs to detect the intruder with a low number of the features at the same time high rate accuracy. Hence, the system proposed here will employ the *best nine features* select from the data set, that contains 41 attributes that describe the different features of the corresponding connection (22 of these features describe the connection itself and 19 of them describe the properties of connections to the same host in the last two seconds).

III. Artificial Neural Networks in Intrusion Detection

The ability of soft computing techniques for dealing with uncertain and partly true data makes them attractive to be applied in intrusion detection. Some research has used soft computing techniques other than ANNs in intrusion detection. For example, Fuzzy logic, and genetic algorithms have been used along with decision trees to automatically generate rules for classifying network connections [14]. However, ANNs are the most commonly used soft computing technique in IDS [1], [15], [16], [17], and [18]. The learning process is essentially an optimization process in which the parameters of the best set of connection coefficients (weighs) for solving a problem are found and includes the following basic steps [19]:

- Present the neural network with a number of inputs (vectors each representing a pattern)
- Check how closely the actual output generated for a specific input matches the desired output.
- Change the neural network parameters (weights) to better approximate the outputs.

Some IDS designers exploit ANNs as pattern recognition technique. Pattern recognition can be implemented by using a feed-forward neural network that has been trained accordingly. During training, the neural network parameters are optimized to associate outputs (each output represents a class of computer network connections, like normal and attack (DOS, Prob, U2R, R2L)) with corresponding input patterns (every input pattern is represented by a feature vector extracted from the characteristics of the network connection record). When the neural network is used, it identifies the input pattern and tries to output the corresponding class. When a connection record that has no output associated with it is given as an input, the neural network gives the output that corresponds to a taught input pattern that is least different from the given pattern [16]. Once the net is trained on a set of representative command sequences of a user, it constitutes (learns) the profile of the user and when put into action, it can discover the variety of the user from its profile [19], [20].

A. Back propagation technique for intrusion detection

The use of neural networks in intrusion detection is not new because there are at least two works that were developed during the last decades. The first model is used in a hyper view [16] for a user behavior modeling. The second one is that discussed in previous studied [21]. While these works used neural networks for either user anomaly detection or misuse detection, we use them here for both network misuse and anomaly detection particularly over the different KDD 1999 data sets [22].

B. Feature Groups

To assess the performance of the proposed IDS, a standard set of data KDD (knowledge Discovery in Database) proposed by Massachusetts Institute of Technology's (MIT) Lincoln Labs is used. The dataset contains about 311029 connection records which can be divided mainly into five categories

- Normal data part: (60593) records.
- Probing attack part (surveillance and other probing): (4166) records.
- DOS attack part (denial-of-services): (229853) records.
- U2R attack part (unauthorized access to the local super user (root) privileges): (230) records.
- R2L attack part (unauthorized access from a remote machine): (16187) records.

The suggested neural networks were trained with the reduced feature set (9 out of 41 features) using a data set that consists of 311029 connection records. A five-class binary classification was performed. The Normal data belongs to class (5), and 39 attack types that could be classified into four main categories (summarized in Table (1)): class (1) Probe, class (2) DOS, class (3) U2R belongs to, class (4) R2L.

Table (1) The 39 attacks and their categories [23]:

Table 1 Show the types of the attacks

Probing	DOS	U2R	R2L
Ipsweep, mscan,nmap, portsweep, saint,satan	Apache2, back, land, mailbomb, Neptune, pod, processtable, smurf,teardrop,udpstorm	Buffer-overflow, httptunnel, loadmodule,perl, ps,rootkit, sqlattack, xterm	ftp-write, guess-asswd, imap,multihop,named ,phf, send-mail, snmpgetattack,snmpguess, spy, arzclient, arezmaster,worm,xlock,xsnoop

IV. SELECT THE IMPACT OF FEATURES INPUTS

Feature selection and ranking is an important issue in intrusion detection systems. There are features that can be monitored for intrusion detection purposes, the elimination of useless features (or audit trail reduction) enhances the accuracy of detection while speeding up the computation, thus improving the overall performance of an ID. In cases where there are no useless features, concentrating on the most important ones may improve the real-time performance of IDS without affecting the accuracy of detection in statistically significant ways:-

- ***Data filtering***: The main purpose of data filtering is to reduce the amount of data processed by the IDs. Data that may not be useful can be eliminated before processing. This has the benefit of decreasing storage space requirements, reducing processing time and improving the detection rate. However, as filtering may move useful data, it must be done with care.
- ***Feature selection***: In difficult classification domains, some data may hinder the classification process (through false correlations or redundancy). Extra features can increase computation time, and can impact the accuracy of IDS [14]. Feature selection improves classification by error and trial for the subset of features, which best classifies the training data. The features under consideration depend on the type of IDS, for example, network-based IDS will analyze network related information such as a packet destination IP address, logged in time of a user, type of protocol, duration of connection etc. It is not known which of these features are redundant or irrelevant for IDS and which ones are relevant or essential for IDS.
- ***Evaluate***: To evaluate the performance of any neural network recognition system, the accuracy of the system result should be calculated as follows:

$$\frac{\text{Number of correctly classified patterns}}{\text{Total number of patterns}} \quad \ldots\ldots (1)$$

Also the four alarms will be calculated as follows [24,25,26]:

Let

TP= # normal connection record classified as normal (TP)
TN= # attack connection record classified as an attack (TN)
FP= # normal connection record classified as an attack (FP)
FN= # attack connection record classified as normal (FN): Then

TP_Rate (sensitivity) = $TP/(TP+FN)$ ……(2.1)

TN_Rate(specificity)= $TN/(TN+FP)$ … (2.2)

FN_Rate=(1-sensitivity)= $FN/(FN+TP)$ ..(2.3)

FP_Rate=(1-specificity)= $FP/(FP+TN)$ …(2.4)

V. ACON STRUCTURE CLASSIFIERS

The all-class-in-one-network (ACON) structure is adopted in case that all classes are lumped into one super-network. Two Back propagation feed forward neural networks are used as ACON IDS. One trained with normal data (with out fuzzification) and other of the fuzzification data.

A. ANN ARCHITECTURE WITH NONFUZZIFIED DATA

The NN used in this work consists of three layers, an input, a hidden, and an output layers. An input layer consisting of ***9*** neurons equal to features vector that have been selected from KDD dataset. The hidden layer consists of ***22*** neurons. In addition the output layer consists of ***5*** neurons, based on trial and error. Finally, the network training is stopped when the Least-square-error *E* between the desired d_i and actual output y_i is less than *Emax* or when number of sweeps=500 or more. Here, *Emax* is chosen to be 0.000001.

$$E = \frac{1}{2P}\sum_{p=1}^{P}\sum_{i=1}^{m}(y_i - d_i)^2 \quad \ldots (3)$$

Where *P* is the total number of training patterns, and

$$d_i = \begin{cases} 1 & \text{If the training pattern} \in \text{i-th texture} \\ -1 & \text{Otherwise} \end{cases}$$

The activation function (nonlinearity function), *tanh*(*x*). For all experiments, the learning rate α was set to 0.00001, each training yielding differents, of which the best result is selected.

B. ANN ARCHITECTURE WITH FUZZIFIED DATA

In this type of feed-forward neural networks data should be fuzzified before training the NN. The proposed NN consists of ***45*** neurons in the input layer (number of features (***9***) × fuzzy linguistic values (***5***)), since for each feature there are ***5*** membership values. The training and stopping conditions used in this net are the same of the parameters used in the ANN with nonfuzzified data.

VI. THE PROPOSED MODEL OF IDS

We selected the vital features using the ANNs. After training and testing of each property we note the total number of features that ***nine*** of which have more influence in the accuracy of intrusion detection, from the ANN's of the class node as explained in Section '*Importance of data reduction for intrusion detection systems*'. These *9*-variables are *C, E, F, L, W, X, Y, AB and AD*. Furthermore, the back propagation network classifier was constructed using the training data and then the classifier was used in the test data set to classify the data as an attack (Five classes) or normal data.

Process with actual data and fuzzified data:

i. Collect Data Set 1999 from DRAPA.
ii. Data Set encodes.
iii. Uniform Selection.
iv. Normalization.
v. Fuzzify data
vi. Training and test data, ANNs, FNNs
vii. Then determine what features are the most effect.

The following figure (1) shows the proposed model:

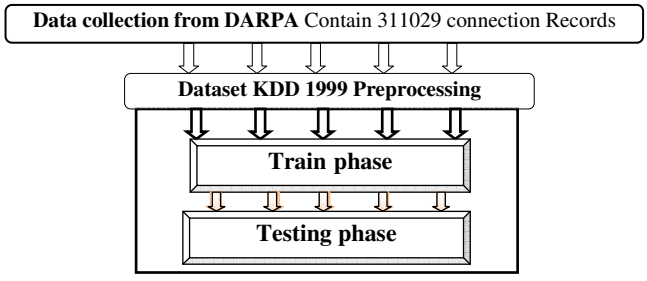

Figure 1 illustrates the work of the IDS model

VII. RESULTS

After the training and testing of all *41* features the performance of classification and back propagation networks shows a table (2) below:

Table 2 Performance of classification and backpropagation networks

Attack Class	41 – Variable data set
	Accuracy (%)
Normal	99.64
Prob	97.85
DOS	99.47
U2R	48.00
R2L	90.58

On the other hand, after the training and testing the reduced set of *9* features show Performance of classification and back propagation networks in Table (3) below:

Table 3 Performance of classification and back propagation networks

Attack Class	9 – Variable data set
Normal	99.95 %
Prob	99.42 %
DOS	99.95%
U2R	100 %
R2L	98.13 %

Table (4) overviews the number of records the data set used in the train and test:

Attack Class	**9 – Variable data set**				
	Real Record	Neural Networks	Match Records	Miss Records	Accuracy
Normal	**7656**	**7662**	**7652**	**10**	**99.95**
Prob	**3944**	**3928**	**3921**	**7**	**99.42**
DOS	**50040**	**50029**	**50014**	**15**	**99.98**
U2R	**384**	**384**	**384**	**0**	**100**
R2L	**1391**	**1368**	**1365**	**3**	**98.13**
Unknown	**0**	**45**	**0**	**45**	**Nan**

The rates of the alarms are calculated by equations (2.1-2.4) as shown Table 5.

Alarm Type	Accuracy
True positive	99.81 %
True negative	98.70 %
False negative	0.19 %
False positive	1.30 %

VIII. CONCLUSIONS

The main purpose of the research presented here was to enhance the detection rate of the R2L and reduce the proportion of the false negative. This research has investigated new techniques for intrusion detection and performed data reduction. From the practical results, it is seen that by using the Neural Networks Normal, U2R and R2L with high accuracy, respectively. Our future research will be directed to another improvement of the specific intrusion system developed is the use of real-time intrusion detection.

[1] J. Cannady, "Artificial neural networks for misuse detection," Proceedings of the 1998 National Information Systems Security Conference (NISSC'98), Arlington, VA, 1998.

[2] S. Malik," Network Security Principles and Practices", chapter 14, November 15, 2002.

[3] K. M. Ali Alheeti and et al "Affect Fuzzication on Neural Networks Behavior as Intrusion Detector", the 4th IEEE Conference on Industrial Electronics and Applications (ICIEA 2009), ISBN: 978-1-4244-2799-4, to be held on May 25-27, Xi'an Chine.

[4] Http://kdd.ics.uci.edu/databases/kddcup99/task.htm

[5] A. M. Chandrashekhar and K. Raghuveer," Fortification of hybrid intrusion detection system using variants of neural networks and support vector machines**",** International Journal of Network Security & Its Applications (IJNSA), Vol.5, No.1, January 2013**.**

[6] M. Md. M. Hassan," Current studies on intrusion detection system, Genetic algorithm and fuzzy logic:, International Journal of Distributed and Parallel Systems (IJDPS) Vol.4, No.2, March 2013.

[7] P. Kaur et al,**"** Mingle Intrusion Detection System Using Fuzzy Logic ", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-2, Issue-3, February 2013**.**

[8] N. KADHI etal," A Mobile Agents and Artificial Neural Networks for Intrusion Detection", Journal of software, VOL. 7, NO. 1, JAN. 2012.

[9] K. M. Ali Alheeti et al, "Artificial intelligence and its role in building an intelligent system has a high capability in tracking", Published in the Conference of the Almaref College University, ISNB (1815-3364), Vol. 19 Thirteenth Year 1433 A.H.-2012 A.D. Iraq.

[10] K. M. Ali Alheeti et al," Application of a Fuzzy Neural Network Combined with an Expert Petri Net System to Intrusion Detection System" The 13th International Arab Conference on Information Technology ACIT'2012 Dec.10-13 ISSN: 1812-0857.

[11] S. Chebrolu, A. Abraham, JP. Thomas" Feature detection and ensemble design of intrusion detection systems". Compute Secur; 24: 2005, pp.295–307.

[12] A. Patcha *, J. Park." Network anomaly detection with incomplete audit data ".Bradley Department of Electrical and Computer Engineering, Elsevier: 2007

[13] B. Yocom, R. Birdsall, D. Poletti-Metzel, Gigabit intrusion detection systems,http://www.nwfusion.com/reviews/2002/ 1104rev.html, 2002.

[14] C. Sinclair, L. Pierce, and S. Matzner, "An application of machine learning to network intrusion detection," Proceedings of 15th Annual Computer Security Applications Conference (ACSAC '99), Phoenix, AZ, pp. 371-377, 1999.

[15]. K. Fox, R. Henning, J. Reed, and R. Simonian, "A neural network approach towards intrusion detection," Proceedings of 13th National Computer Security Conference, Baltimore, MD, pp. 125-134, 1990.

[16] H. Debar, M. Becker, and D. Siboni. "A neural network component for an intrusion detection system". In Proceedings of the 1992 IEEE Symposium On Research in Computer Security and Privacy, Oakland, CA, May 1992.

[17] S. Mukkamala, "Intrusion detection using neural networks and support vector machine," Proceedings of the 2002 IEEE International Honolulu, HI, 2002.

[18] R. Cunningham and R. Lippmann, "Improving intrusion detection performance using keyword selection and neural networks," Proceedings of the International Symposium on Recent Advances in Intrusion Detection, Purdue, IN, 1999.

[19] S. Theodorios and K. Koutroumbas, Pattern Recognition, Cambridge: Academic Press, 1999.

[20] K. Kendall, "A Database of Computer Attacks for the Evaluation of Intrusion Detection Systems," Masters Thesis, MIT, 1999

[21] J. Cannady. "Artificial Neural Networks for Misuse Detection". In Proceedings of the 1998 National Information Systems Security Conference (NISSC'98), Arlington, VA, USA, October 5-8 1998.

[22] KDD 99 Task. Available at: http://kdd.ics.uci.edu/ databases/kddcup99/task.html, 1999.

[23] V. G. Jecheva, E. P. Nikolova," Learning Problem and BCJR Decoding Algorithm in Anomaly-based Intrusion Detection Systems" Journal of software, VOL. 2, NO. 6, December 2007, pp,48-50.

[24] S. M.Al-naqshabandi, "Simulation system for computer network intrusion detection", a thesis submitted in partial fulfillment of the requirements for the degree of doctor of philosophy in computer science, Al-Nahrain University, Baghdad, Iraq, 2007,pp.61-66.

[25] R. Chang, Liang-Bin Lai , W. Su, J. Wang, Jen-Shiang Kouh "Intrusion Detection by Back propagation Neural Networks with sample-Query and Attribute-Query" , International journal of computational intelligence research, ISSN 0973-1873, Vol.3, No.1,2007 , pp,6-10.

[26] S. J. Stolfo, Wei Fan, W. Lee, A. Prodromidis, & P. K. Chan, "Cost-based modeling for Fraud and Intrusion Detection: Results from the JAM Project", 2000, http://www1.cs.columbia.edu/jam/recent-project-papers.html, last access day 28-july-2007, pp.1-15.