Classification Of Attacks In Network Intrusion Detection System

[1] Shwetambari Ramesh Patil, [2]Dr.Pradeep Deshmukh ,
[1]Department of Computer Engineering, Rajarshi Shahu College of Engineering
[2]Department of Computer Engineering, Rajarshi Shahu College of Engineering
University of Pune
Tathawade, Pune 411033. India.
[1] Shwet19@gmail.com , [2] pkdeshmukh9@gmail.com

***Abstract*—We have designed a network intrusion detection system based on the artificial neural networks using Multi Layer Perceptron (MLP) and Modified signature Apriory algorithm ; and the testing results of the prototype system proved the validity of the method and the advantages over other methods suggested. In the present study a more general problem is considered in which the attack type is also detected such as smurf, teardrop,etc. This feature enables the system to suggest proper actions against possible attacks.**

Keywords: ***Multi Layer Perceptron (MLP), Signature Based Apriory Algorithm, Intrusion Detection System, Artificial Neural Network***

I. INTRODUCTION

In today computer network Internet is a global public network. With the growth of the Internet and its potential, there has been subsequent change in business model of organizations across the world. More and more people are getting connected to the Internet every day to take advantage of the new business model popularly known as e-Business. Internetwork connectivity has therefore become very critical aspect of today's e_business. There are two sides of business on the Internet. On one side, the Internet brings in tremendous potential to business in terms of reaching the end users. At the same time it also brings in lot of risk to the business. There are both harmless and harmful users on the Internet. While an organization makes its information system available to harmless Internet users, at the same time the information is available to the malicious users as well. We have designed a network intrusion detection system based on the artificial neural networks using Multi Layer Perceptron (MLP) and Modified signature Apriory algorithm ; and the testing results of the prototype system proved the validity of the method and the advantages over other methods suggested. In the present study a more general problem is considered in which the attack type is also detected such as smurf, teardrop,etc. This feature enables the system to suggest proper actions against possible attacks.

A. .Intrusion Detection System

An intrusion detection system (IDS) is a device or software that monitors network or system activities for malicious activities or policy violations and produces reports to a Management Station. Some systems may attempt to stop an intrusion attempt but this is neither required nor expected of a monitoring system. Intrusion detection and prevention systems (IDPS) are primarily focused on identifying possible incidents, logging information about them, and reporting attempts. In addition, organizations use IDPSes for other purposes, such as identifying problems with security policies, documenting existing threats and deterring individuals from violating security policies. IDPSes have become a necessary addition to the security infrastructure of nearly every organization. Some existing Intrusion Detection System is explained below.

1) A host-based intrusion detection system (HIDS) - is an intrusion detection system that monitors and analyzes the internals of a computing system as well as (in some cases) the network packets on its network interfaces (just like a network-based intrusion detection system (NIDS) would do). This was the first type of intrusion detection software to have been designed, with the original target system being the mainframe computer where outside interaction was infrequent. This system has strong deterrence for insiders, strong insider detection and weak outsider detection. Weak real-time response. It has good for long-term attacks. Excellent for determining extent of compromise. Good at trending and detecting suspicious behaviour patterns.

2) Network intrusion detection system (NIDS)- In computer security, a Network Intrusion Detection System (NIDS) is an intrusion detection system that attempts to discover unauthorized access to a

computer network by analyzing traffic on the network for signs of malicious activity.

3)Hybrid Systems: A hybrid system is nothing but simple combination of two or three intrusion detection systems put together. For e.g. host based and network based .

II. ARTIFICIAL NEURAL NETWORK

An Artificial Neural Network, is a mathematical model inspired by biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases a neural network is an adaptive system that changes its structure during a learning phase. Neural networks are used to model complex relationships between inputs and outputs or to find patterns in data. 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 [10]:

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

1)Multi Layer Perceptron (MLP)

A multilayer perception (MLP)[1] is a feed forward artificial neural network model that maps sets of input data onto a set of appropriate output. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. MLP utilizes a supervised learning technique called back propagation for training the network. MLP is a modification of the standard linear perceptron and can distinguish data that is not linearly separable.

A .Architecture

The benefit of the artificial neural networks[1] includes the ability of faster information processing, the ability of classification and the ability of learning and self organization.

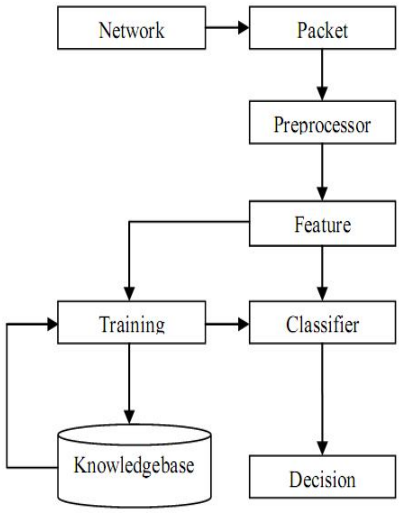

Fig1.System Overview Diagram

Referring the figure.1. modules are

- Packet Monitor: This module monitors network stream real time and capture packets to serve for the data source of the NIDS
- Preprocessor: This module, network traffic collected and processed for use as input to the system.
- Feature Extractor: This module extracts feature vector from the network packets (connection records) and submits the feature vector to the classifier module.
- Classifier:This module is to analyze the network stream and to draw a conclusion whether intrusion happens or not.
- Decision: When detecting that intrusion happens, this module will send a warning message to the user.
- Knowledgebase: The function of this module is to serves for the training samples of the classifier phase. As you know, the artificial neural networks can work effectively only when it has been trained correctly and sufficiently. The intrusion samples can be perfected under user participation, so the capability of the detection can improve continually.

All the above modules together make the NIDS architecture system based on the artificial neural networks(MLP), and the function of each module has been introduced briefly.

B. Datasets

The KDD is the data set used for The Third International Knowledge Discovery and Data Mining Tools Competition, which was held in conjunction with KDD-99 [1].The Fifth

International Conference on Knowledge Discovery and Data Mining. The competition task was to build a network intrusion detector, a predictive model capable of distinguishing between ``bad'' connections, called intrusions or attacks, and "good" normal connections. This database contains a standard set of data to be audited, which includes a wide variety of intrusions simulated in a military network environment. It contains attack of category DOS(e.g. sync flood); R2L(guessing password); U2R(`buffer overflow); probing(port scanning).

C. Feature Extraction

Feature selection and extraction is important problems in implementing the intrusion detection system. Network stream itself is not suitable for the input of the Classifier module, so it is necessary to extract some features from the network stream. The features extracted from the network stream are a feature vector which serves for the description of the packet. A complete description of all 41 features is available [4], [15]. We select all 41 features as an input in training phase first.

D. Classifier module

Classifier module is the widely used neural network model, which has made a good figure in the field of Pattern Recognition. We select two hidden layers in the Classifier module , consequently we have a three layers in neural network. (The input layer is not counted ;however, the output layer is counted). The dimension of the input layer is the number of the features selected, and the dimension of the output layer is the number of sorts that can be classified by the Classifier. The transfer function. The function *Logsig(x)* = 1/ (1+ exp (-x)) can be used in the Classifier model. The learning function. *trainGD* function can be used in the Classifier which works based on Back-propagation algorithm.

WEKA Tool is used for the implementation of MLP algorithm for intrusion detection and classification of attacks.

III. MODIFIED SIGNATURE BASED APRIORI ALGORITHM

By the observation of the attack signatures, we find that there are some attack signatures dependent on other previous attack signatures. This is due to the new attack is a derivative from the previous attack. So far as we know, there are at least two kinds of attacks have this property. For the first case, a new attack is variation of an existing attack. The steps of how to find out frequent k-item sets will be as follow. At the first step, all of the frequent items will be found. And then we use a simple way to scan the database in order to find the frequency of occurrence of each item, and decide which one meets the minimum support. Secondly, we generate the candidate n-item sets by checking all of the possible combinations of the frequent items with already known signatures, if they meet the minimum support requirement. Then, append this n item sets from right. We can first append the backward, until the minimum support is unsatisfied. Then, we append forward, and stop when the same condition occurred. Finally, the maximum length of frequent-item set can be mined by our method.

When the minimum support decreases, the processing times of algorithms increase because of the total number of candidate item sets increases. Our algorithm is faster than the Signature Apriori no matter what the minimum support is. The reason is that the number of candidate 1-itemsets is not very large. Therefore, in the real environment, there are not too much candidate item sets to be generated during each pass of finding signatures.

One of the approaches of developing a network safety is to describe network behaviour structure that point out offensive use of the network and also look for the occurrence of those patterns. While such an approach may be accomplished of detecting different types of known intrusive actions, it would allow new or undocumented types of attacks to go invisible. As a result, this leads to a system which monitors and learns normal network behaviour.

B.Snort

We use SNORT TOOL for implementation Modified Signature Based Apriory Algorithm and for detection of attacks.

C. Algorithm Pseudo code

The pseudo code for the algorithm is given below for a transaction database T, and a support threshold of ϵ. Usual set theoretic notation is employed, though note that T is a multiset. C_k is the candidate set for level k. Generate() algorithm is assumed to generate the candidate sets from the large item sets of the preceding level, heeding the downward closure lemma. $count[c]$ accesses a field of the data structure that represents candidate set c, which is initially assumed to be zero. Many details are omitted below, usually the most important part of the implementation is the data

structure used for storing the candidate sets, and counting their frequencies.

$$\texttt{Apriori}(T, \epsilon)$$
$$L_1 \leftarrow \{ \texttt{ large 1-itemsets } \}$$
$$k \leftarrow 2$$
$$\texttt{while } L_{k-1} \neq \emptyset$$
$$C_k \leftarrow \{c | c = a \cup \{b\} \wedge a \in L_{k-1} \wedge b$$
$$\texttt{for transactions } t \in T$$
$$C_t \leftarrow \{c | c \in C_k \wedge c \subseteq t\}$$
$$\texttt{for candidates}$$
$$c \in C_t$$
$$count[c] \leftarrow count[c] + 1$$
$$L_k \leftarrow \{c | c \in C_k \wedge \ count[c] \geq \epsilon\}$$
$$k \leftarrow k + 1$$
$$\bigcup L_k$$
$$\texttt{return } k$$

IV .TESTING RESULT

A. MLP:- The implemented intrusion detector was a three layer MLP. In this database 381738 samples are exist for evaluating in testing and training phase, 21000 of which are used for training, and the 300 are used for testing. Table I gave the testing results for the unseen data (test set) in intrusion detection system. From Table.I, we can see that the Classifier has the ability to detect the unseen intrusions This is mainly because that the artificial neural networks have the ability of classification and self organization. For these six kinds of attacks, the testing results are satisfactory.

Attacks Types	Detection Rate
Smurf	92.1
Teardrop	89.6
Satan	93.22
Guest	90.6
Buffer Overflow	83.9
Warezclient	88.2

TABLE I. TESTING UNSENN DATA

BModified Signature Based Apriori Algorithm:- Table II gave the testing results for the unseen data (test set) in intrusion detection system. From Table II, we can see that the algorithm has the ability to detect the unseen intrusions.

Attacks Types	Detection Rate
Smurf	89.2
Teardrop	87.7
Satan	94.2
Guest	89.6
Buffer Overflow	81.9
Warezclient	86.2

TABLE I. TESTING UNSENN DATA

V . CONCLUSIONS AND FUTURE WORK

Neural Network based intrusion detection system intended to classify the normal and attack patterns and the type of attack. The signature Apriory algorithm is used for real traffic analysis with more closely accurate. Here We are dealing with Kdd cup dataset file only. And for real time detection we are assuming incoming packets are only icmp packets. In future we can classify more attack type and proper action can be taken against attacks types detected.

REFERENCES

[1]"Classifying Attacks in a Network Intrusion Detection System Based on Artificial Neural Networks", Mohammad Reza Norouzian, Sobhan Merati,Feb. 13~16, 2011 ICACT2011, ISBN 978-89-5519-155-4

[2] James Cannady,"Arificial neural network for misuse detection,"Proceedings of the 1998 National Information Systems Securityconference(NISSC'98),Arlington,VA,1998

[3]Srinivas Mukkamala,"Intrusion detection using neural networks and support vector machine,"Proceedings of 2002 IEEE International Honolulu,HI2002

[4]J.Ryan,M.Lin and R. Miillulainen,"Intrusion Detection with Neural Networks,,'AI Approaches toFraud Detection and Risk management:Papers from the 1997 AAAI Workshop,Providence,RI,pp.72-79,1997

[5]Kabiri P,Ghorbani A,"A Research in intrusion detection and response-a survey". International Journal of Network Security ,2005

[6]Helman ,P.,Liepins ,G.,and Richards,"Foundation of Intrusion Detection ",In Proceedings of the Fifth Computer Security Foundations Workshop pp.114-120,1992
[7]Anderson, D., Frivold,T. & Valdes ,"Next Generation Intrusion Detection Expert System(NIDES):A Summary".SRI International Technical Report SRI-CSL-95-07,1995
[8]P.Garcia-Teodoroa,J.Diaz-Verdejao,G.Macia-Ferna'ndeza,E.,Va'zquezb, "Anomaly based network intrusion detection:Techniques,systems and challenges," Elsevier,2009
[9] Sergios Theodorios and Konstantinos Koutroumbas,"Pattern Recognition, "Cambridge: Academic Press, 1999
[10]H.Debar,M.Becker,and D.Siboni,' A neural network component for an intrusion detection system, "Proceedings of 1992 IEEE Computer Society Symposium on Research in Security and Privacy,Oakland,California,pp.240-250,1992.
[11] Archive, U. K. "KDD Cup 1999 Data." Retrieved April 20, 2009, from http://www.sigkdd.org/kddcup/index.php
[12] R Lippmann, JW Haines, DJ Fried, J Korba, K Das,"The 1999 DARPA off-line intrusion detection evaluation"- Elsevier, 2000.
[13] Kristopher Kendall, "A Database of Computer Attacks for the Evaluation of Intrusion Detection
[14] Salvatore J. Stolfo, Wei Fan, Wenke Lee, Andreas Prodromidis ,"Costbased modeling and evaluation for data mining with application to fraud and intrusion detection, Results from the JAM Project", pp. 20- 27, 1999.
[15]http://www.sigkdd.org/kddcup/index.php?section=1999 Systems,"Master Thesis,MIT,1999