ISSN:0975-9646

Optimized Intrusion Detection System

Ragini Sen, Navdeep Kaur Saluja, Pratibha sahu

Department of Computer Science & Engineering
Infinity Management and Engineering College (IMEC)
Sagar, India

Abstract- In the communication system when sender wants to communicate with the system or communication is between sender and receiver then we have to trust to intermediate user but the problem is intruders or security. So It is necessary to detect the intruders and have to remove it in the communication system for the protection of the information. In this paper optimized intrusion detection system (OIDS) has been proposed to detect the intruders in which ANN base classification is used. To test the performance of the algorithm , implementation has been done. Outcome of the result is that OIDS performs better than the existing techniques

Keywords -Intrustion detection system , KDD, Learning system, ANN.

I. INTRODUCTION

An intrusion detection system (IDS) is a device or software application that monitors network or system activities for malicious activities or policy violations and produces reports to a management station. Some systems can try to stop an intrusion attempt but this is neither required nor expected of a monitor system. Intrusion detection and prevention systems (IDPS) are primarily focused on identifying possible incidents, logging information, reporting attempts etc. In addition, organizations use IDPSes for other purposes, 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**.**

IDPSes typically record information related to practical events, inform security administrators of important observed events and produce reports. Many IDPSes can also counter to a detected threat by attempting to prevent it from succeeding. They use a number of response techniques, which involve the IDPS stopping the attack itself, changing the security environment (e.g. reconfiguring a firewall) or changing the attack's content

1.1 Challenges in Intrusion Detection System

Intrusion detection systems in theory looks like a defense tool which every organization needs. However there are some challenges the organizations face while deploying an intrusion detection system. These are discussed below.

1. IDS technology itself is undergoing a lot of enhancements. It is therefore very important for organizations to clearly define their expectations from the IDS implementation. IDS technology has not reached a level where it does not require human intervention. Of course today's IDS technology offers some automation like notifying the administrator in case of detection of a malicious activity, shunning the malicious connection for a configurable period of time, dynamically modifying a router's access control list in order to stop a malicious connection etc.
But it is still very important to monitor the IDS logs regularly to stay on top of the occurrence of events. Monitoring the logs on a daily basis is required to analyze the kind of malicious activities detected by the IDS over a period of time. Today's IDS has not yet reached the level where it can give historical analysis of the intrusions detected over a period of time. This is still a manual activity. It is therefore important for an organization to have a well-defined Incident handling and response plan if an intrusion is detected and reported by the IDS. Also, the organization should have skilled security personnel to handle this kind of scenario

2. The success of an IDS implementation depends to a large extent on how it has been deployed. A lot of plan is required in the design as well as the implementation phase. In most cases, it is desirable to implement a hybrid solution of network based and host based IDS to benefit from both. In fact one technology complements the other.However, this decision can vary from one organization to another. A network based IDSis an immediate choice for many organizations because of its ability to monitor multiplesystems and also the fact that it does not require a software to be loaded on a production system unlike host based IDS. Some organizations implement a hybrid solution. Organizations deploying host based IDS solution needs to keep in mind that the host based IDS software is processor and memory intensive. So it is very important to have sufficient available resources on a system before installing a host based sensor on it.

3. It is important to take care of sensor to manager ratio. There is no thumb rule as such for calculating this ratio. To a large extent it depends upon how many different kinds of traffic is being monitored by each sensor and in what environment. Lot of organizations deploy a 10:1 ratio. Some organizations go for 20:1 and some others 15:1.It is very important to design the baseline policy before starting the IDS implementation and avoid false positives. A badly configured IDS sensor may send a lot of false positives to the console and even a 10:1 or even better sensor to console ratio can be inadequate.

4. The IDS technology is still reactive rather than proactive. The IDS technology works on attack

signatures. Attack signatures are attack patterns of previous attacks. The signature database needs to be updated whenever a different kind of attack is detected and the fix for the same is available. The frequency of signature update varies from vendor to vendor.

5. While deploying a network based IDS solution, it is important to keep in mind on every important aspect of the network based IDS in switched environment. Unlike a HUBbased network, where a host on one port can see traffic in and out of every other port in the HUB, in a switched network however, traffic in and out of one port can not be seen by a host in another port, because they are in different collision domains.A network based IDS sensor needs to see traffic in and out of a port to detect any malicious traffic. In a switched environment, port mirroring or spanning is required to achieve this. One entire VLAN can be spanned to one port on which the network based IDS sensor is installed.Although this is a solution, there may be performance issues for a busy network. If all the 10/100 Mbps ports in a VLAN are mirrored to another 10/100 Mbps port in theVLAN, the IDS sensor may drop traffic, as the combined traffic of all the ports could be more than 100 Mbps. Now, Gigabit port speed being available, this becomes an even more difficult challenge. Cisco systems has an IDS module for Catalyst 6000 series switch which can sit on the switch back plane and can monitor traffic right off the switch back plane. But this solution is yet to scale to Gigabit speed. This module supports traffic only up to 100 Mbps as of now. The portability of network based IDS in as witched environment is still a concern

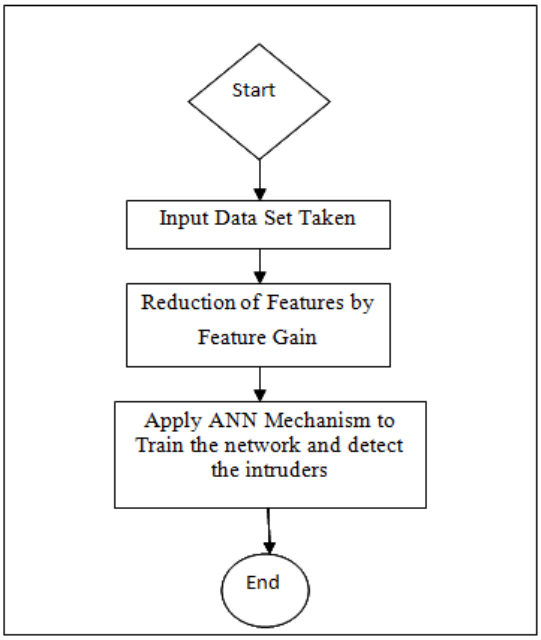
II. PROPOSED WORK

2.1 Optimized Intrusion Detection System

In this approach we have achieved the subset of 24 features by reducing NSL-KDD [26] dataset of 41features for intrusion detection on the basis of feature's vitality. The vitality of feature is determined by considering three main performance criteria the classification accuracy, TPR and FPR of the system. We used feature gain to identify the important set of features: starting with the set of all features, on efeature was removed at a time until the accuracy of the classifier was below a certain threshold. In otherwords, the feature selection of is "leave-one-out" remove one feature from the original dataset, redo the experiment, then compare the new results with the original results. Since there are 41 features in NSLKDDdata set, the experiment is repeated 41 times to ensure that each feature is either important,unimportant or less important. By deletion of a feature if performance decreases then feature is important,if performance increases then feature is unimportant and if no changes found in performance then feature is less-important. Here we have explained the algorithm for OIDS. First, we apply ANN classifier on dataset with 41 features and its performance output like classifier's accuracy, RMSE, averageTPR value and set F is input to this algorithm.

2.2 Flow Chart and Algorithm of OIDS

2.2.1 Flow Chart of OIDS



Description:

Initially, data set taken as input, apply feature gain method to reduce the data set then apply Artificial neural network to train the network to detect the intruders , repeat the mechanism again and again to achieve the accuracy. Algorithm is shown in 2.2.2

2.2.2 Algorithm of OIDS

Input:

F= Full set of 41 features of NSL-KDD dataset

Ann= Artificial Neural Network

err= RMSE

avg_tpr= average TPR

//ann, err and avg-tpr resulted from invocation of NBC on full dataset, these values used as threshold values for feature selection

//OIDS Algorithm:

Begin

Initialize: S={F}

For each *feature {f} form*

(1) T=S-{f}

(2) Invoke ANN Classifier on dataset with T Features

(3) If CA>=ann And RMSE<=err And A_TPR>=avg_tpr then

S=S-{f}

F=S// Set F with reduced features

End

2.3 Advantage of OIDS

1. Accuracy to Detect the Intruders is increased
2. Time is reduced to train the network

III. Result And Analysis

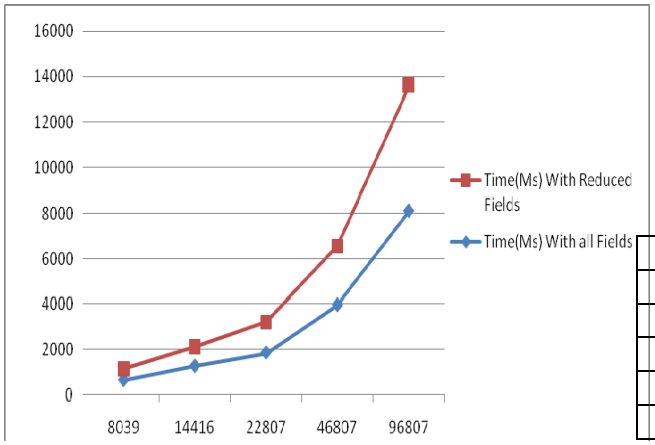
3.1 The data set to be used in our experiments is NSL-KDD labeled dataset. The number of records in the NSLKDDtrain and test sets are reasonable. This advantage makes it affordable to run the experiments on the complete set without the need to randomly select a small portion. Consequently, evaluation results of different research work will be consistent and comparable. NSL-KDD dataset contains one type of normal data and 22 different types of attacks which falls into one of four categories. These are DoS, probe, R2L ,and U2R. we extracted only 62,986 records out of 1,25,973 NSL-KDD dataset connections for training and testing.

3.2 Experimental Analysis

3.2.1 For the Experimental Analysis No. of test Record is taken 8039, 14416, 22807,46807 and 96807 And timing with all field and timing with reduced fields has been calculated in which timing of reduced filed is decreased as compared to timing with all Fields as shown in table 3.2.1 and graph 3.2.1

Test Record	Time(Ms) With all Fields	Time(Ms) With Reduced Fields
8039	656	477
14416	1279	854
22807	1872	1347
46807	3982	2570
96807	8128	5510

Table 3.2.1 Time with all Fields and Reduced Fields



3.2.1 Graphical Analysis of all Fields and Reduced Fields

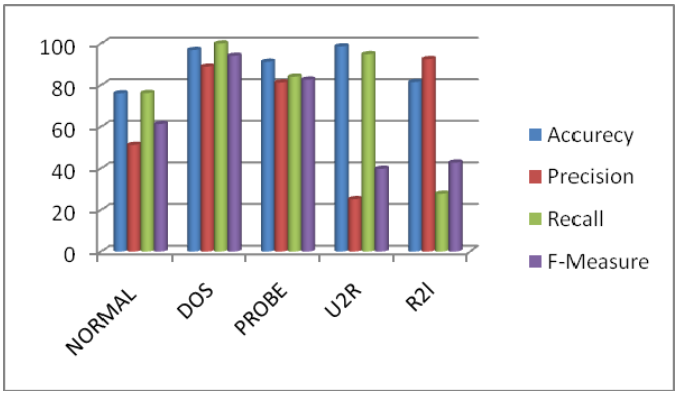
3.2.2 Data set with all fields has been taken as shown in the confusion matrix 3.2.2.1 and accuracy , precision , recall and F-Measure has been calculated as shown in table 3.2.2.3 and graph 3.2.2.4

Confusion Matrix					
CLASS/CLASS	**NORMAL**	**DOS**	**PROBE**	**U2R**	**R2L**
NORMAL	1524	2	388	49	37
DOS	0	2000	0	0	0
PROBE	7	248	1681	56	8
U2R	2	0	0	37	0
R2l	1439	0	0	5	556

Table 3.2.2.1 : Input Data Set

CLASS/CLASS	Accuracy	Precision	Recall	F-Measure
NORMAL	76.07	51.28	76.2	61.3
DOS	96.89	88.89	100	94.12
PROBE	91.21	81.25	84.05	82.62
U2R	98.61	25.17	94.87	39.78
R2l	81.48	92.51	27.8	42.75

Table 3.2.2.2 accuracy , Precision , Recall and F-Measure Calculation with All Fields



Graph 3.2.2.3: Accuracy , Precision , Recall and F-measure Calculation for all Fields

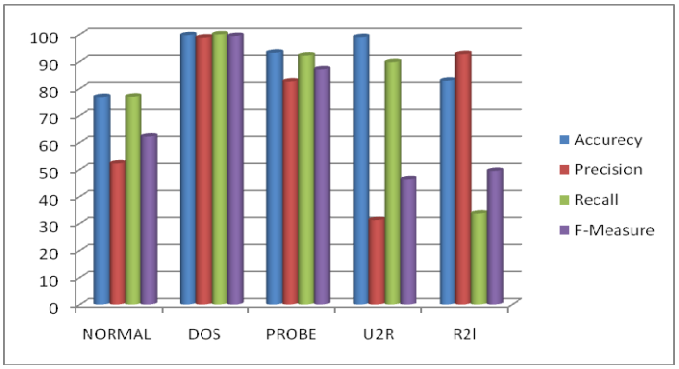
3.2.3 Data set with all fields has been taken as shown in the confusion matrix 3.2.3.1 and accuracy, precision, recall and F-Measure has been calculated as shown in table 3.2.3.2 and graph 3.2.3.3

DOS	0	2000	0	0	0
PROBE	81	24	1843	42	10
U2R	2	0	0	35	2
R2l	1323	0	0	3	674

Table 3.2.3.1 : Input Data Set

CLASS/CLASS	Accurecy	Precision	Recall	F-Measure
NORMAL	76.76	52.24	76.9	62.22
DOS	99.7	98.81	100	99.4
PROBE	93.21	82.57	92.15	87.1
U2R	98.99	31.25	89.74	46.35
R2l	82.85	92.71	33.7	49.43

Table 3.2.3.2 accuracy , Precision , Recall and F-Measure Calculation with reduced Fields



Graph 3.2.3.3: Accuracy , Precision , Recall and F-measure Calculation for reduced Fields

3.4 Comparison and Analysis

3.4.1 Data Set of Normal , Dos , Probe , U2R and R2l has been taken with all fields and reduced fields in which accuracy of reduced field has been increased as compared to all fields.

3.4.2 Data Set of Normal , Dos , Probe , U2R and R2l has been taken with all fields and reduced fields in which precision of reduced field has been increased as compared to all fields.

3.4.3 Data Set of Normal , Dos , Probe , U2R and R2l has been taken with all fields and reduced fields in which recall of reduced field has been increased as compared to all fields.

3.4.4 Data Set of Normal , Dos , Probe , U2R and R2l has been taken with all fields and reduced fields in which F-Measure of reduced field has been increased as compared to all fields

IV. CONCLUSION AND FUTURE WORK

4.1 CONCLUSION

This thesis examines the intrusion detection problem by characterizing intrusion detection possibility with respect to three standard feature selection methods using Correlation-based Feature Selection,Information Gain and Gain. A proposed model for intrusion detection is applicable to real life time applications. To reduce the input here we have used ANN Method. Empirical results show that selected reduced attributes give better performance to design IDS that is efficient and effective for network intrusion detection

4.2 FUTURE WORK

In future, we want to implement our algorithm in hardware for the better speed , and use different classifier to check which better suit for the detection of intrusion

REFRENCES

[1]. S. mukherjee and N. Sharma , “Intrusion Detection using Naive Bayes Classifier with Feature Reduction”, Sciverse Science Direct , Elesvier , 2012 , pp. 119-128.

[2]. S.B. Chao, “Incorporating Soft Computing Techniques Into a Probabilistic Intrusion Detection System” ,IEEE Transaction on System, Man and Cybernetics-Part C: Application and Reviews, Vol. 32, NO. 2, MAY 2002, pp. 154-160.

[3]. Z. Yu , J.J.P. Tsai , “An Automatically Tuning Intrusion Detection System” , IEEE Transaction on System, Man and Cybernetics-Part B: Cybernetics, Vol. 37, NO. 2, APRIL 2007, pp. 173-182.

[4]. J. Hu and X. Yu , “A Simple and Efficient Hidden Markov Model Scheme for Host-Based Anomaly Intrusion Detection”, IEEE Network , January/February 2009, pp. 42-48.

[5]. L. Chen , J. Leneutre , “A Game Theoretical Framework on Intrusion Detection in Heterogeneous Networks” ,IEEE Transaction on Forensics and Security, Vol. 4, NO. 2, JUNE 2009, pp. 165-174.

[6]. Y.D. Lin , H.Y. Wei and S.T. Yu, “Building an Integrated Security Gateway: Mechanism Performance Evaluations, Implementations, and Research Issues”, IEEE Communications Surveys, pp. 1-14.

[7]. E. Pontes, A. Guelfi and E. Alonso, “Forecasting for Return on Security Information Investment: New Approach on Trends in Intrusion Detection and Unwanted Internet Traffic” , IEEE Latin America Transactions, Vol. 7, NO. 4, AUG. 2009, pp. 438-446.

[8]. Z. M. Fadlullah, T. Taleb, A. V. Vasilakos, M, Guizani, and N. Kato, “DTRAB: Combating Against Attacks on Encrypted Protocols Through Traffic-Feature Analysis” , IEEE/ACM Transactions On Networking, Vol. 18, NO. 4, AUGUST 2010, pp. 1234-1244.

[9]. J.H. Cho, I.R. Chen and P.G. Feng, “Effect of Intrusion Detection on Reliability of Mission-Oriented Mobile Group Systems in Mobile Ad Hoc Networks” , IEEE Transaction On Reliability, Vol. 59, NO. 1, March 2010, pp. 231-242.

[10]. F.M. Pérez, F. J.M. Gimeno, D.M. Jorquera, J.A.G. M. Abarca, H.R. Morillo, and I.L. Fonseca, “Network Intrusion Detection System Embedded on a Smart Sensor” , IEEE Transactions On Industrial Electronics, Vol. 58, NO. 3, MARCH 2011, pp. 722-731