INTRUSION DETECTION SYSTEM BY USING SVM & ANN

Meenakshi Sharma
Associate Professor(C.S.E)
S.S.C.E.T, Badhani
Pathankot, India
sharma.minaxi@gmail.com

Shubam Sharma
Research Scholar(C.S.E)
S.S.C.E.T, Badhani
Pathankot, India
shubam.sharma67@gmail.com

Abstract- **Security and privacy of a system is compromised as information security work as a shield gor information society, when an intrusion happens. Intrusion Detection System (IDS) plays vital role in network security as it detects various types of attacks in network. So here, we are going to propose Intrusion Detection System using data mining technique for helping IDS to attain a higher maximum detection rate.**

Keywords -Classification, Intrusion Detection System (IDS), Kernel Function, Pre-processing, Support Vector Machine (SVM).

I. INTRODUCTION

The network security is comely a crucial need of new society to assure the secrecy information flowing over the networks. Revelation of interference over network is the utmost severely critical path to prevent their illegal use by the assaulter.

The adequate intrusion detection is needed as a shield of the patchwork system to detect the incursion over the network. A object choice and classification based Intrusion Detection model is conferred by executing feature selection, the dimensions of NSL-KDD data set is diminished and then by applying machine learning approach, we are able to frame Intrusion detection system to find barrage on system and boost the intrusion detection using the grabbed data. With the increasing number of new invisible attacks the purpose of this model is to develop a system for intrusion detection, and the design will be capable of detecting new and previously invisible attacks using the basic signatures and the features of known attacks. The influential and worthwhile report always captivates traducers and is always amenable to supreme attacks over the network.

Intrusion is getting into the system or system server by the attacker by sending the malevolent packet to the user system and then pirate, corrupting or modifying any confidential instruction or paramount knowledge, the dispatching of network packet over the network for misuse determination is admitted as attack. The obtrusion can ensue over the system or server by virtue of any extant system weakness or vulnerability, such as system misconfiguration, user misuse or program defects. An inventive intrusion can also be contrived by inserting heterogeneous susceptibilities together.

In a global network there are millions of big servers and large number of on-line values are executing in the system while such networks attract more attackers and need intelligent intrusion detection model as a defense for their network system.

A. INTRUSION DETECTION

The intrusion detection methods are based on network and hot based.

Network based intrusion detection system (NIDS) gives the control over the data and the information which are travelling on the networks that is over internet and also detects if there is any type of bug or faults in network.

In this type of system each packet that is frames travels by the path of network are firstly analyzed before sending that packet over a network, NDS is used in examining of these data packets.

In case of hot based system detection of intrusion, it also control in same manner by detecting the attacks over the network but the only difference is that it is based on networks events, so this type based detects the all operating system events and other way by which all the actions and working of each computes and host is examined, therefore it is individual based intrusion detection system.

B. PATH OF ANALYSIS

There are two ways of detecting intrusion while analyses are misuse and anomaly detection.

- In case of misuse or any unauthorized access instruction or the frame matching is applied to detect the intrusion or bug in the system software, So this based of system depends upon the matching processes.
- Abnormal system is different approach of detection. It follows different steps by collaborating the intrusion and type of attacks and then give a brief account on its behavior's afterward it start matching with the existing previous recorded
- Attacks or in other words we say that it uses the signature method in intrusion based on the previous and already been created signatures.

There is lack of security in these types of detection methods. As it own detect the known attack with an precision. So the updating of these database attacks is required. Because anomaly intrusion detection is only for the detection of computer system bugs and misuse.

C. DATASET CHARACTERIZATION

To perform the intrusion detection the NSL-KDD data set is used, this data set was prepared in March, 2009 by Tavallaee,M, Bagheri, E, W, Lu, and Ali A, Ghorbani and was the modified and improved version of old KDD data set. It consist of chosen records from KDD which is a complete data set and has been the most riotously used data set used for the study and assessment anomaly based intrusion detection. KDD data set is created by S. J. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. K. and is made based on the data captured in IDS evaluation program KDD is the most preferred data set for intrusion detection as compared to other available data set because it is well labelled and contain several attack types and shows the multiple attack scenarios, whereas the other data set are limited. KDD is composed of about 4 gigabytes of raw compressed data of 7 weeks of network traffic collected with the help of TCP dump which is packet analyzer runs under the command line used to intercept network packets flowing under the network. This data was processed into around 5 million connection records, each with about 100 bytes. KDD dataset consists of around 4,900,000 data packets with each packet containing 41 features and one class indicating packet either normal or attack with specifying the attack type.

The NSL-KDD data set that has been used for intrusion detection is refined data set of original KDD data set. KDD data set was having the problem of duplicate or redundant data that was removed in NSL-KDD data set. Redundant data is having the disadvantage of biasing the learning algorithms that is removed in NSL-KDD data set, this make data set more realistic for attack detection.

NSL-KDD data set has two separate sets, one is train data and other is test data set. Train data consist of 24 attack types and on the other hand test data contains 14 additional attacks including previous 24 attacks, this makes the detection more realistic and accurate because now the accuracy of learning algorithms are also

checked for previously unseen attacks.

II. RELATED WORK

Mrutyunjaya et al.(2012) developed the detection system with the help of knowledge discovery of database sets. In this use of different hybrid techniques are applied with comparing their performance.

Sabhnani et al.(2003) has used an different operation in process of machine learning implementation to detection of intrusion in KDD data sets while number of algorithms are used in measuring the performance. Model shows an appropriate results to detect the user to root and remote to local attacks accurately.

***Meera Gandhi et al. (2010*)** described number of guide lines, laws and many other decision trees. Techniques of machine learning is used for the selection and training of data sets. As several attacks define their alternative characteristic this is the best choice of using machine learning techniques and algorithm. Model of machine learning shown an improvement in the classification and performance. Author in this scenario used tenfold cross approach for validation to make enhancement. Attack detection method are dependent upon the several algorithm approaches as a result of using this tenfold cross method for obtaining accuracy for the familiar intrusions.

Rowayda A. et al. (2013) proposed the technology and the implementation of the algorithm known as the artificial neural network using fluctuate value. The methods in this had used powerful attributes for the reduction in faults and correction of the results with precise accuracy algorithm used here are reliable.

Hafiz Muhammad et al. (2012) As in computer security this detection plays a vital role so that the efficiency of the system remains accurate. Selection and conversion of features are performed on individual method. But in this case the use of complex approach LDA and GA with hybrid combination involved for the transformation. Principal component analysis is a approach of feature attraction which is selected over LDA technology. Another radial bases function approach is there for classification of network influx into the activities of intrusion.

Dabas et al. (2013) proposed the techniques of intrusion detection system for the protection of computers and the patch networks according to the requirement. These techniques which provide security and reliability. Methods involved work on the presumptions. As these anomaly detection model difference in the activities with compare to the anomaly based intrusion detection system.

III. STANDARDS

An aim is set in the intrusion detection system for the selection of particular data sets. Then algorithm approach of machine learning is applied for collection of features and the new Weka approach is defined for many data values which are observed in the fold cross for validations. By the investigation of the whole system method report shows that Recall, Accuracy and Precision three of the subsets are derived from the feedback system data sets and the intrusion detection model.

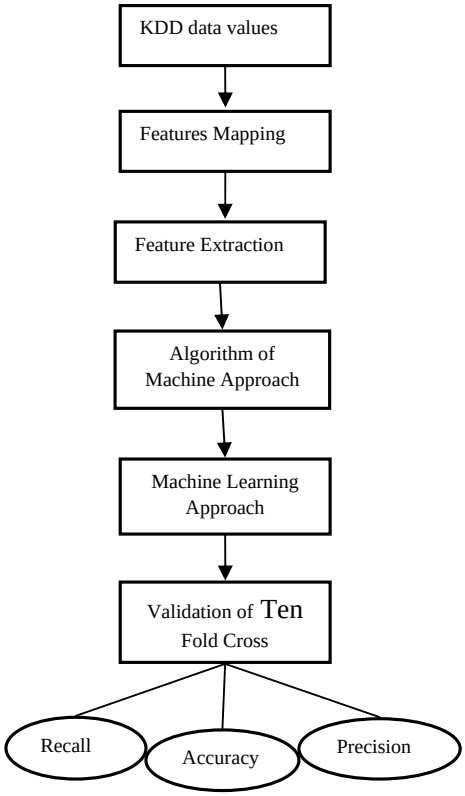

Figure 3.1 Flowchart of IDS [5]

This area starts up by taking an standard for intrusion detection system. All the data values are mapped by describing the features after the study and analysis. Weka is introduced for gathering and selection of the features for machine learning design including various data sets and values for this Weka.

The whole development strategy is done by the intrusion detection system. The three term analysis, recall and decision are the output we received from the data sets which are valid. As recall is the parameter of the feature originate from the observation and evaluation of system where as accuracy is obtained from the intrusion detection system.

Machine learning is a model which describes the feature extraction methodology. NSL-KDD is a data set on which feature extraction is performed by Principal Component Analysis(PCA).PCA is the important term for determination of feature extraction and useful feature of evaluated data for intrusion detection. So by testing and analyzing these feature the help in safe detection of irregular and abnormal activities.

This feature extraction scales down the number of members from the selected features.

TOOL DESCRIPTION

Weka is the one of the most extensible and critical popular tool used in the field of data mining and purpose of classification. In case of application term of algorithms, collection of open source of data mining as well as pre-processing, clustering and extraction of data is done by Weka as a tool.

FEATURE EXTRACTION

Feature Extraction is the term used for the selection of most important useful attribute features from the total overall features which deals with the critical role in intrusion detection. Feature extraction helps in reducing the features from data set without affecting the effective indicators of system intrusions. As principle component analysis algorithm is implemented to find the procedure feature in NSL-KDD data values in detection system.PCA type of feature extraction is used for reducing the complexity of the system. PCA algorithm is used for the extraction by transforming the data space into feature space.

IV. RESULTS OBSERVATION

Classifier	Precision	Recall	Accuracy
Linear	64.98	69.45	70.60
Quadratic	71.13	72.57	77.13
Polynomial	53.83	53.92	65.21
RBF	72.36	75.97	65.54
Multilayer Perception	66.45	65.69	80.11
Adaptive Boost with SVM	99.53	99.50	99.57

Table 1 - Accuracy, Precision and Recall comparison.

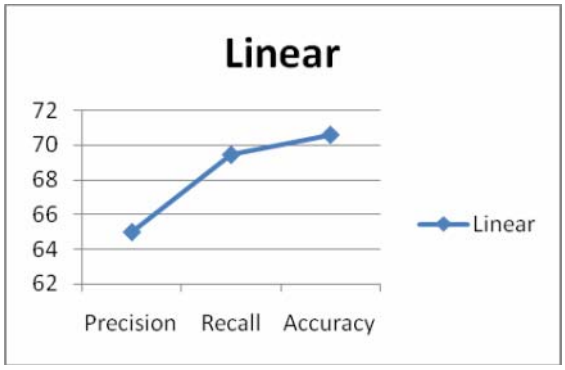

Figure 4.1 Linear graph.

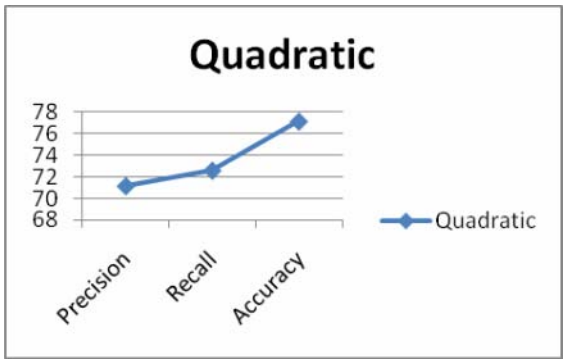

Figure 4.2Quadratic graph

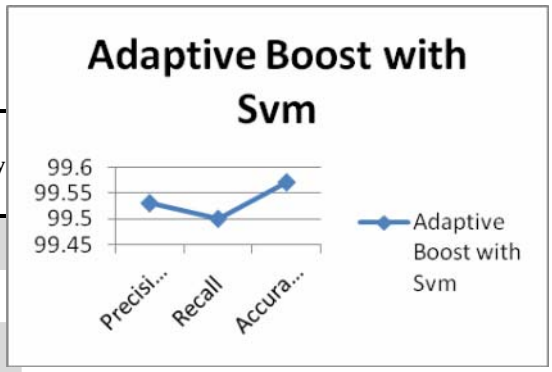

Figure 4.3 Adaptive Boost graph

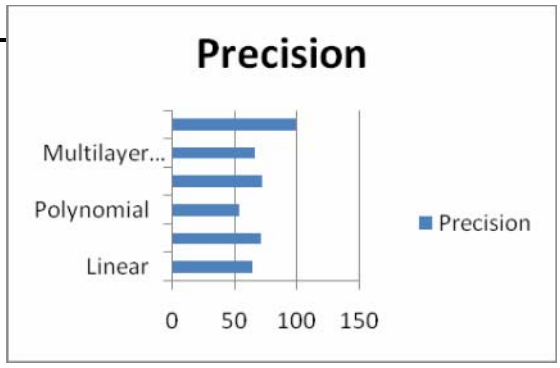

Figure 4.4 Precision Bar graph

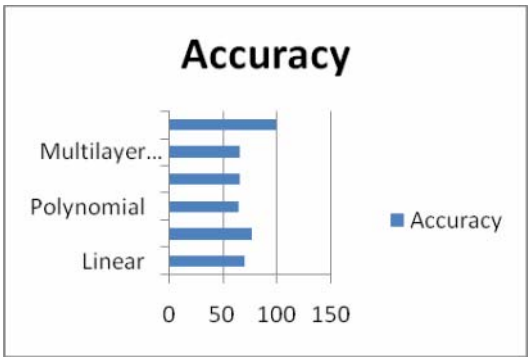

Figure 4.5 Accuracy Bar graph

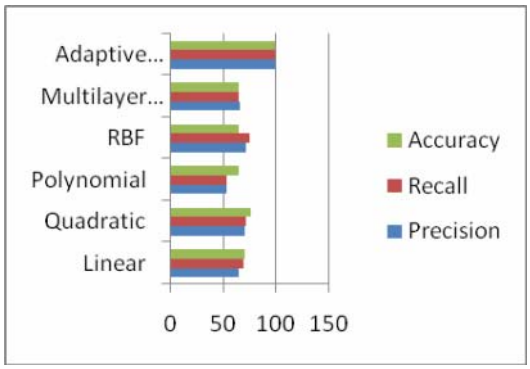

Figure 4.6 Comparison of Classifiers on SVM machine

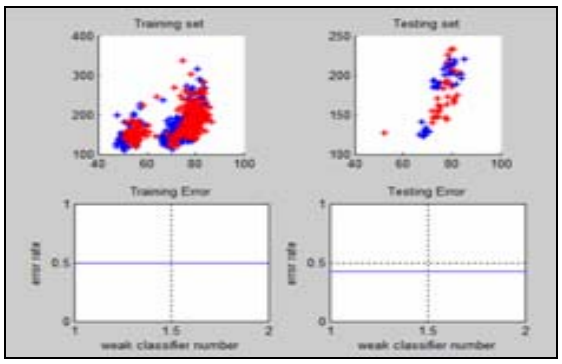

Figure 4.7 Training Boost SVM-RBF Training & Testing Error

V. CONCLUSION AND FUTURE SCOPE

In these work two models for intrusion detection is presented, that suggest, for the detection of intrusion it is not necessary to perform the test on all the 41 features of NSL-KDD data set. Second by using feature Extraction the features of the training and test set are reduced to 33 features and further by removing them, the biasing of learning algorithms towards the frequent and easily detectable records in the data set is reduced.

Future Scope: The proposed models can be checked on different Intrusion detection system, to establish the new benchmarks on Intrusion detection.

In future work, new hybrid model for intrusion detection can be built by optimizing the different machine learning algorithms.

More parameters can be set for network features to improve the rate of intrusion detection, by further applying more techniques on proposed model this model can be laid as basic foundation for real life intrusion detection in future.

REFERENCES:

[1] M. P. Agrawal, V. C. Pandey, and S. P. Keshri, “Importance of Intrusion Detection System with its Different approaches,” vol. 2, no. 5, pp. 1902–1908, 2013.

[2] T. Ahamad and A. Aljumah, “Hybrid Approach using intrusion Detection System,” vol. 2, no. 2, pp. 87–92, 2014.

[3] H. O. Alanazi, R. Noor, B. B. Zaidan, and A. A. Zaidan, “Intrusion Detection System : Overview,” vol. 2, no. 2, pp. 130–133, 2010.

[4] P. Banerjee, P. Banerjee, and S. S. Dhal, “International Journal of Advanced Research in Computer Science and Software Engineering,” *Int. J.*, vol. 2, no. 9, pp. 62 – 70, 2012.

[5] Y. B. Bhavsar and K. C. Waghmare, “Intrusion Detection System Using Data Mining Technique : Support Vector Machine,” *Int. J. Emerg. Technol. Adv. Eng.*, vol. 3, no. 3, 2013.

[6] N. Branch and I. Azad, “An Efficient Hybrid Intrusion Detection System based on C5 . 0 and,” vol. 7, no. 2, pp. 59–70, 2014.

[7] R. A. Calix and R. Sankaran,

"Feature Ranking and Support Vector Machines Classification Analysis of the NSL-KDD Intrusion Detection Corpus," no. 2006, pp. 292–295, 2009.

[8] A. M. Chandrashekhar and K. Raghuveer, "fortification of hybrid intrusion detection system using variants of neural," vol. 5, no. 1, 2013.

[9] M. G. Feshki, "Managing Intrusion Detection Alerts Using Support Vector Machines," no. 9, pp. 266–273, 2015.

[10] S. Ganapathy, K. Kulothungan, S. Muthurajkumar, and M. Vijayalakshmi, "Intelligent feature selection and classification techniques for intrusion detection in networks : a survey," pp. 1–16, 2013.

[11] N. F. Haq, "Application of Machine Learning Approaches in Intrusion Detection System : A Survey," vol. 4, no. 3, pp. 9–18, 2015.

[12] L. M. Ibrahim, "Anomaly networkintrusiondetection system based on distributed time-delay neural network (dtdnn)," vol. 5, no. 4, pp. 457–471, 2010.

[13] L. Khan, M. Awad, and B. Thuraisingham, "A new intrusion detection system using support vector machines," pp. 507–521, 2007.

[14] V. Kosamkar and S. S. Chaudhari, "Improved Intrusion Detection System using C4 . 5 Decision Tree and Support Vector Machine," vol. 5, no. 2, pp. 1463–1467, 2014.

[15] P. G. Kumar, "INTRUSION DETECTION USING ARTIFICIAL NEURAL NETWORK WITH REDUCED INPUT FEATURES."

[16] H. Lee, J. Song, and D. Park, "Based on Multi-class SVM," pp. 511–519, 2005.

[17] S. Mukkamala, G. Janoski, and A. Sung, "Intrusion Detection : Support Vector Machines and Neural Networks," 1998.

[18] S. Mukkamala and A. H. Sung, "Feature Selection for Intrusion Detection using Neural Networks and Support Vector Machines 3 . RANKING THE SIGNIFICANCE OF INPUTS," pp. 1–17.

[19] S. Mukkamala, A. H. Sung, and A. Abraham, "Intrusion Detection Using Ensemble of Soft Computing Paradigms."

[20] S. Mukkamala, A. H. Sung, and A. Abraham, "Intrusion detection using an ensemble of intelligent paradigms," vol. 28, 2005.

[21] P. P. Naik, "An Approach for Building Intrusion Detection System by Using Data Mining Techniques," vol. 2, no. 2, pp. 112–118, 2014.

[22] R. R. Reddy, "A Survey on SVM Classifiers for Intrusion Detection," vol. 98, no. 19, pp. 38–44, 2014.

[23] K. K. Tiwari, S. Tiwari, and S. Yadav, "Intrusion Detection Using Data Mining Techniques."

[24] G. Wang, J. Hao, J. Ma, and L. Huang, "Expert Systems with Applications A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering," *Expert Syst. Appl.*, 2010.