International Journal of Scientific & Engineering Research, Volume 4, Issue 12, December-2013
ISSN 2229-5518

Review: Ensemble Neural Network and KNN Classifiers for Intrusion Detection

Shalinee Chaurasia ,Prof. Anurag Jain

ABSTRACT-In this paper; we represent the ensemble algorithm to improve the intrusion detection precision. Intrusion detection system monitors system activities for malicious supervise and produces reports to a management system. Intrusion system is a software application to update represents a new Ensemble Technique. Security is becoming a major problem for all networks, the main reason for using Data Mining Classification Methods for Intrusion Detection Systems due to the huge amount of existing and newly appearing network data that needs analysis. This paper proposes a combining classification approach for intrusion detection. In this paper we are using a neural network model; K- Nearest Neighbors method .These data mining classification Intrusion detection has proven to be useful for a variety of knowledge gathering. Work , implemented in two phases with the first phase of the neural network for better results and improve the KNN classifiers and we both bagging using it for the second stage of the well is to find a class of classification systems.

KEYWORDS - Intrusion Detection System, Bagging, Neural Network, K- Nearest Neighbour (KNN).

1. INTRODUCTION

Intrusion Detection problem is one of the most promising research issues of knowledge Security. The problem provides very good opportunities in terms of providing host and network security [1]. An Intrusion Detection System (IDS) itself can be formed as the tools, methods, and resources to help identify, assess and report unauthorized or unapproved network activity [2].An Intrusion Detection System (IDS) is a program that gets at the details of what happens or has happened during an execution and tries to get ideas of indications that the computer has been misused. An intrusion detection system monitors the activities of a given environment and decides whether these activities are malicious (intrusive) or legitimate (normal) based on system integrity, confidentiality and the availability of information resources[3].IDS is very important part of any Security architecture.

Shalinee Chaurasia , Department of Computer Science Engg, RGTU University, Radharaman Institute of Technology and Science,INDIA,
8989487067 (e-mail: shalinee0501@gmail.com).
Anurag Jain, HOD, Department of Computer Science, RGTU University, Radharaman Institute of technology and science, INDIA, (e-mail: anurag.akjainr@gmail.com)

2. INTRUSION DETECTION SYSTEM

IDSs are software system designed to prevent and identify the misuse of computer network and system. An intrusion detection system (IDS) is a device or software application that network or system activities for malicious activities 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 [4].

3. TYPES OF THE INTRUSION DETECTION

There are two main types of IDS

3.1 Network intrusion detection system

Network intrusion detection is designed to detain outsiders. Using packet sniffing .Looking at IP header as well as data parts. A Network Intrusion Detection System (NIDS) is an intrusion detection system that attempts to discover unauthorized access to a computer network by analysing traffic on the network for signs of malicious activity.

3.2 Host-based intrusion detection system

Host-based systems are designed more to detain insiders, but can't effectively detain outsiders. Host-based systems maintain a large database of behavioral information that can be mined for trends indicative of misuse.[5]

4. INTRUSION DETECTION MODEL TECHNIQUES

4. 1. Signature-Based intrusion detection-

Signature-based is also known as misuse detection or knowledge-based systems. They follow the same principle as most anti-virus software and rely on the knowledge accumulated about previous attacks and vulnerabilities to detect intrusion attempts. Misuse detection systems compare current activities of the host or the network monitored with "signatures" of known attacks. If the current activities match any of the known signatures, an alarm is triggered.[6]

Misuse intrusion detection-uses well-defined patterns of the attack that exploit weaknesses in system and application software to identify the intrusions. These patterns are encoded in advance and used to match in opposition to the user behavior to detect intrusion [7]. A signature-based or so-called misuse detection system searches for known malicious patterns in the payload. A pure signature-based IDS uses only single events for the analysis [8].

4.2 .Novelty intrusion detection

The anomaly detection technique canters on the concept of a baseline for network behavior. This baseline is a description of accepted network behavior, which is learned or specified by the network administrators, or both. Events in an anomaly detection engine are caused by any behaviors that fall outside the predefined or accepted model of behaviour. A behaviour-based IDS, also known as an anomaly or novelty detection system, analyses in the first instance the traffic data [9].

5. LITERATURE SURVEY

Two most significant motives to launch attacks as described in are, either to force a network to stop some services that it is providing or to steal some information stored in a network. An intrusion detection system must be able to detect such anomalous activities. However, what is normal and what is anomalous is not defined, i.e., an event may be considered normal with respect to some criteria, but the same may be labelled anomalous when this criterion is changed. Hence, the objective is to find anomalous test patterns which are similar to the anomalous patterns which occurred during training. The underlying assumption is that the evaluating criterion is unchanged and the system is properly trained such that it can reliably separate normal and anomalous events.

Subbulakshmi, A. Ramamoorthi et al.[1] In this paper uses Misuse and Anomaly detection using SVM , NBayes, ANN and ensemble approach. Ensemble approach outperforms all the other approaches with high classification rate.

V. Bapuji et al.[2] In this paper The implementation of Soft Computing and Artificial Intelligence methods are used widely in Intrusion Detection System are gaining its ability to learn and evolve which makes them more accurate and efficient in facing the enormous number of unpredictable attacks. The two major techniques for machine learning were highlighted, with the use of Genetic Algorithm and Artificial Neural Network providing intrusion system with extra intelligence.

Jyotiprakash Sahoo et al.[3]In this review, Intrusion Detection System overview is presented and various approaches for applying genetic algorithms for network intrusion detection are discussed. GA as evolutionary algorithms was successfully used in different types of IDS.

Harley Kozushko et al.[4] In this paper that combined network-based and host-based intrusion detection systems effectively prevent attacks from insider as well as outsider sources. While there are new methods of intrusion detection, most systems

utilize signatures to search for patterns of misuse and either report to the security officer or automatically respond to the misuse.

Shaik Akbar et al. [5] proposed in this paper we discussed a brief overview of Intrusion Detection System (IDS), related detection techniques and about the KDD Cup 99 Intrusion data. We are sure this brief survey is useful for all researchers those who want to investigate more efficient methods against intrusions.

Kapil Kumar Gupta et al. [6] in this thesis, we explored the suitability of conditional random fields for building robust and efficient intrusion detection systems which can operate, both, at the network and at the application level. The three issues are: 1. Limited attack detection coverage 2.Large number of false alarms and, 3.Inefficiency in operation. We performed a range of experiments which show that, in order to detect intrusions effectively, it is critical to model the correlations between multiple features in an observation. Our framework is highly scalable, easily customizable and can be used to build efficient network intrusion detection systems which can detect a wide variety of attacks.

6. PROPOSED METHODE

In order to gain data access an attacker performs a sequence of malicious events. An experienced attacker can also disguise attacks within a number of normal events in order to avoid detection. Hence, to reduce the false alarms and increase attack detection accuracy, intrusion detection systems must be capable of analysing entire sequence of events rather than considering every event in isolation.

There is a set of records (*training set)* each record contains a set of *attributes*; one of the attributes is the *class*. The given data set is divided into training sets and test sets, with training set used to build the model and test set used to validate it. We train different layers in our framework for select different features.

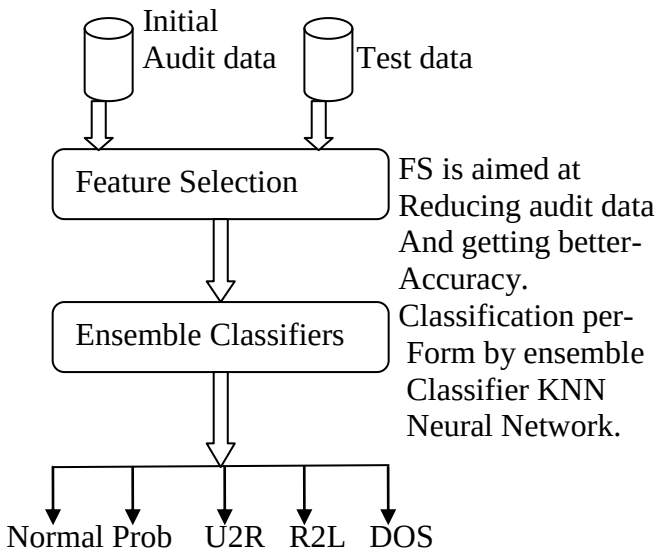

Fig 1 process

The Feature Selection Stage is aimed at reducing more data and getting better accuracy**.** We was trained each of the classifiers by using the same training data. Classification is performed with the help of ensemble classifiers. Here we ensemble Neural Network and KNN by using ensemble classifier**.** Training data was presented to the combine of classifiers. This training dataset has five classes and they are *Normal*, *U2R*, *R2L*, *Probe* and *Do's*.

7. FEATURE SELECTION

Feature selection is a very efficient way to reduce the dimensionality of a problem. Redundant and irrelevant variables are removed from the data before being fed to the machine learning algorithm used as a classifier. Feature selection is a pre-processing step which can be independent of the choice of the learning algorithm or not. It can be used in order to improve the computational speed with minimum reduction of accuracy. Other advantages include noise reduction and robustness against over-fitting. Generally, automatic selection of features works much better than manual selection because the algorithm is able to find correlations between the features that are not always obvious even for a human expert.

7.1Intrusion Detection using Neural Network and K-NN

7.1.1 NEURAL NETWORK

Neural networks have been used both in anomaly intrusion detection as well as in misuse intrusion detection. For anomaly intrusion detection, neural networks were modelled to learn the typical characteristics of system users and identify statistically significant variations from the user's established behaviour. In misuse intrusion detection the neural network would receive data from the network stream and analyse the information for instances of misuse.

In the first approach of neural networks for intrusion detection, the system learns to predict the next command based on a sequence of previous commands by a user. Here a shifting window of w recent commands is used. The predicted command of the user is compared with the actual command of the user and any deviation is signalled as intrusion. If w is too small, there will be many false positives and if it is too big some attacks may not be detected. NNID (Neural Network Intrusion Detector) identifies users based on the distribution of commands used by the user. This system has three phases. In the first phase it collects the training data from the audit logs for each user for some period and constructs a vector from that data to represent the command execution by each user. In the second phase, neural network is trained to identify the user based on these command distribution vectors. In the final phase the network identify the user for each new command distribution vector. If the networks identified user is different from the actual user, it signals anomaly intrusion [11].

7.2.2 K- NEAREST NEIGHBORS

The first machine-learning algorithm we'll look at is k-Nearest Neighbors (KNN).K-NN is a type of instance based learning. K-NN algorithm is amongst the simplest of all Machine learning algorithms. Object is classified by a Majority vote of its neighbors. There is main two parts Training Sets and Test Sets.

KNN PROCEDURE:

1. Store all input data in the training set.
2. for each pattern in the Test set.
3. Search for the K nearest patterns to the input pattern using a Euclidean Distance Measure.
4. For classification compute the confidence for each class as Ci/K.where Ci is no of patter and K is nearest pattern.

8. ENSEMBLE TECHNIQUE

There are a number of methods for generating classifiers in the ensemble. In order to be effective, there must be diversity between each of the classifiers. Multiple Base models (classifiers ,regresses) each covers a different region of the input space.Each base model is trained on a slightly different train set.

GOAL- Improve the accuracy of the Base model

8.1.BAGGING-Corporatedecision-making analogy.

- Managers seek advice of experts in areas that s/he does not have expertise.
- The skills of the advisers should complement each other rather than being duplicative.

Bagging Procedure:

Classifier generation

Step 1. Create t data sets from a database applying the sampling with replacement scheme.

Step 2 Apply a learning algorithm to each sample training set data.

Classification

Step 3. For an object with unknown decision, make predictions with each of the t classifiers.

Step 4. Select the most frequently predicted decision

9. CONCLUSION

In this paper, we describe neural network and knn classifier for detecting intrusions. This paper provides the details of data mining ensemble techniques used to identifying intrusion and general working of ensemble techniques. There are Different Data

mining techniques like classification, clustering and association rule are very helpful in analysing the network data but basic data mining techniques are not sufficient for sensing unknown attacks so we are ensemble classification method to detect unknown intrusions.

REFERENCES

[1] T. Subbulakshmi, A. Ramamoorthi , " Ensemble Design For Intrusion Detection System",Department of Computer Science and Engineering, Thiagarajar College of Engineering, Madurai,Vol 1, No 1, August 2009,pp 1-2.

[2] V. Bapuji ,"Soft Computing and Artificial Intelligence Techniques for Intrusion Detection System", Department of Informatics, Kakatiya University, Warangal, India, Vol 2, No.4, 2012,pp 2-30.

[3] Jyotiprakash Sahoo, "A Survey on Evolutionary Approaches to Intrusion Detection Systems", Dept. of IT, C V Raman College of Engineering .

[4] Harley Kozushko, "Intrusion Detection: Host-Based and Network-Based Intrusion Detection Systems",September 11 (2003).

[5]Shaik Akbar, Dr.K.Nageswara Rao, "Intrusion Detection System Methodologies Based on Data Analysis", International Journal of Computer Applications (0975 – 8887) Volume 5 No.2, August 2010.

[6]Kapil Kumar Gupta,"Robust and Efficient Intrusion Detection Systems", Department of Computer Science and Software Engineering,January 2009.

[7] M.Govindarajan and RM.Chandrasekaran, "Intrusion Detection using an Ensemble of Classification Methods" , Proceedings of the World Congress on Engineering and Computer Science 2012 Vol I.

[8]http://www.cis.syr.edu/~wedu/Teaching/cis758/LectureNotes/Intrusion_Detection.pdf.

[9]http://searchsecurity.techtarget.com/tip/IDS-Signature-versus-anomaly-detection.

[10]. Sandhya Peddabachigari, Ajith Abraham, Johnson Thomas, "Intrusion Detection Systems Using Decision Trees and Support Vector Machines", Department of Computer Science, Oklahoma State University, USA.

[11] Anazida Zainal,"Ensemble of One classifier for Improved Network Intrusion Detection System", Security 4 (2009) 217-225.

[12] Shrinivas Mukkamala, Andrew H. Sung," Intrusion detection using an ensemble of intelligent paradigms", Department of Computer Science, New Mexico Tech, Journal of Network andComputer Applications 28 (2005) pp167–182.

[13] Thakre S.P., Ali M.S."Network Intrusion Detection System & Fuzzy Logic.

[14].http://en.wikipedia.org/wiki/Intrusion_detection_system

[15] Alan Bivens, Chandrika Palagiri," Network-Based Intrusion Detection Using Neural Networks", Department of Computer Science, Rensselaer Polytechnic Institute Troy, New York, ANNIE-2002,St. Louis, MO, vol. 12, ASME Press, New York, NY, 2002, pp. 579-584.

[16] S. Xiaonan Wu*, W. Banzhaf,"The Use of Computational Intelligence in Intrusion Detection System: A Review," springer – 2009.

[17] Benjamin Thirey and Christopher Eastburg," Increasing Accuracy Through Class Detection:Ensemble Creation Using Optimized Binary Knn Classifiers",IJCSEA Vol.1, No.2, April 2011

[18] Hui Zhao," Intrusion Detection Ensemble Algorithm based on Bagging and Neighborhood Rough Set",School of Mathematics and Computer Science, International Journal of Security and Its Applications Vol.7, No.5 (2013), pp.193-204.

[19] Dae-Ki Kang, Doug Fuller," Learning Classifiers for Misuse and Anomaly Detection Using a Bag of System Calls Representation" ,IEEE(2005).

[20] Deepika P Vinchurkar, Alpa Reshamwala," A Review of Intrusion Detection System Using Neural Network and Machine Learning Technique",IJESIT Volume 1, Issue 2, November 2012.