Scientific Journal Impact Factor (SJIF) : 3.518
International Journal of Modern Trends in Engineering and Research
ISSN (ONLINE) : 2349 - 9745 ISSN (PRINT) : 2393 - 8161

A Comparative Analysis of Decision Tree Based Intrusion Detection System

Shikha Singh[1]and Simmi Jain[2]
[1]*Department of Computer Science and* [2]*Engineering,*
NRI Institute Of Science and Technology, Bhopal

Abstract— In this paper, a complete review is discussed regarding the recent scenario of the intrusion-detection-system in the domain of the security of the networks based on the decision tree approach. Here the various intrusion-detection-system techniques and their application on the basis of decision tree also discussed that are available and on which various researches have made. By using the approaches ofintrusion-detection-systems for protecting the computers and the networks from the malignant attacks through implementing the traditional statistical approaches also described in this paper. Some detection approaches that are applied for the intrusion detection are focused on some specific methods. So it is required that the abilities of the approaches of intrusion-detection get updated according to the development of the latest intrusion.In this paper various intrusion detection approaches are analysed for detection of intrusion bythe use of decision tree algorithm.

Keywords— Decision Tree, Intrusion Detection, J48, WEKA, Attacks.

I. INTRODUCTION

With the overwhelming growth of thecomputer networks, and various technologieshaving theinterconnection among them has now suffered several types of security issues. There are various kinds of attackers or hackers are also raising in number accordingly with the increase in the failures in the security mechanism that are found in daily life of networking and those failures are proved beneficial for the attackers. The attacks are made by the intruders or the attackers when they are accessing the system externally without having any privilege, and by the un-certified users that are trying to achieve the access control over the network system in order to violate the security of the system and damaging the information or data present over the network while transmission [1].Within the world of technologies there are more risk have been introduced over the security of the information and for protecting the information from being attacked in the form of any type of cyber attack. So for resolving these issues an approach have been introduced for providing the protection to the system or information from violation or any damage, this approach is termed as the Intrusion-detection-system. In order to measure the performance of the intrusion-detection-system, its capability to determine the various attacks with accuracy have been observed that is classified as the accuracy, classification and scalability. For detecting the intrusions the user activities and their patterns are determined in the intrusion-detection-system. The intrusion-detection-system is classified on the basis of misuse detector, its functionality and the anomaly detector. To the detection of the misuse, the patterns of the attacks are searched and identified. As the system and the technology are growing rapidly hence the complexity for detecting the intrusions have also been raised so it is also becomes very crucial to introduce a best intrusion-detection-system for the recent types of intrusions [2].

Hencethis is affecting the system and their corresponding organization adversely. Various types of host-based intrusion-detection-systems have been introduced for the organizations in order to detect the intruders and finding the patterns for the detection of the intrusion and preventing the system or organization by introducing the intrusion prevention mechanism by analysing and comparing the detected behaviour with the normal behaviour of the pattern [3].

All these types of scenarios increase the necessity of the intrusion-detection-system to monitor the malicious activity within the network and the traffic of the network in order to find the

intrusion. As today is the world of the customer's satisfaction hence the suppliers are continuously trying to introduce better mechanism for intrusion detection and provide it to their customers to resolve various issues regarding security. Many of the IDS are available as an open source over the free of cost for implementation. Thus there are various ways to implement a wide range of IDS applications for any organization [4].

Hence in this paper, the various reviews on the intrusion system and their approaches in various fields are described and the review on its implementation with the decision tree algorithm have also discussed in detail based on the several existing researches in this domain.

II. INTRUSION-DETECTION-SYSTEM

Intrusion-detection-system recognized the known and the unknown patterns of the attacks over the network after which this system performs the required actions according to the detected intrusion for theconnections of network. IDS include the several approaches and the set of methods that are implemented for detecting the malicious activities that are found on the network and at the host. The attacker troubles the networks and the host without informing them in order to damage the information or violate the security. The huge organizations like E-bay, E*TRADE are the best examples of the firms that have got effected by this types of intrusion [5].

Intrusion is defined as the any type of group of the actions which are trying to affect theconfidentiality, integrity or the availability of the system and the intrusion-detection-system is the device or the software application which monitors thetraffic of network for detecting the suspicious activity, if any activity of this type found then an alert is generated for the system or the network administrator in order to warn them. The major three modules of the intrusion-detection-system are:

- Monitoring
- Analyses
- Response

The Intrusion-Detection-Systems is capable of finding the attacks within theexisting environments and it resolve the uncertainties within the monitoring of network through implementing the detection systems over the area of attacks.The intrusion-detection-system iscapable to provide the prevention commands also to the firewalls and the access control that changes to the routers which may be considered as anenhancement over the technologies of firewall. This may provide the decisions for access control which are dependent on the application content, no on the IP address or the ports similar to the basic types of firewalls.

2.1Various TypesOf IDS

Here are described some types of Intrusion-detection-systems, which are of three types mainly termed as:

- Hostbased intrusion-detection-system (HIDS),
- Network basedintrusion-detection-system (NIDS),
- Distributed intrusion-detection-system (DIDS).

A. Host based intrusion-detection-system(HIDS):

This type ofintrusion-detection-system monitors the information and thenanalyse it which is gathered from the particular hostsystem. The host-based IDS areexecuting over the host machine that detects theintrusion through gathering the information like the file system which is used, the events on the network and the system calls etc for detecting the intrusion. This type of IDS analyse the modification found within the host kernel, in the behaviour of program and with the host file system [6].

B.Network based intrusion-detection-system (NIDS):

This type of intrusion-detection-systemis attempt to find the malignant activity over the network like DoS attacks, scan the ports or may try to crack the computers through monitoring the

networktraffic even. And the collectedinformation from the network is then got compared with the known patterns of intrusion for the detection of intrusion.

C.Distributed intrusion-detection-system (DIDS):

This type of intrusion-detection-systemincludes various types of other intrusion-detection-systems likeNetwork based intrusion-detection-system, Hostbased intrusion-detection-system, etc that are found over the huge network, and these all may interact with each other, or through the centralserver which provides the monitoring of network.

2.2 Techniques of Intrusion Detection

There are various techniques that have been applied for the Intrusion-detection-systemin order to detect the intrusion and provide the security to the system. Here are described some of the techniques of the IDS that provides the security in some way [7]:

A. Statistical Techniques

In this technique, dependent on some particular conditions the statistical comparison have made on the data which is gathered from the network or the system in order to find the intrusion.

B. Pattern recognition Techniques

Within this technique, the major advantage is by applying the sequence of the penetration scenarios which are embedded within the system, decreases the requirement to analyse the huge amount the data.

C. Rule based Techniques

In this the set of "if-then" rules are implemented to detect the attacks in which every rule is linked with a particular operation within the system.

III. TREE BASED ALGORITHMS FOR IDS

3.1 Decision Tree Techniques

In the decision tree, since there is no need for any type of domain knowledge or the parameter settings to be implemented for the decision tree classifiers, hence it is considered as the most efficient tool for the knowledge discovery process. [8] Within a decision tree, the topmost node is called the root node.This decision trees may manage the data that is having highdimension along with their demonstration of theobtained knowledge in the form of tree which is simple and also easy to resemble with the humans behaviours. The most simple and rapid phases of the decision-tree induction are learning and theclassification which is having the more accuracy.

3.2. Decision Tree Algorithm:-

Step1: Initially the connection in between the Client and Server is established.
Step2: Then the intrusion-detection-system may accept the Input Data from the Client system.
Step3: Then Apriori Algorithm is implemented.
Step4: In case the Training Data or the Attacks that are obtained from the dataset got matched with Tested-Data,then the generated output is found to be similar.
Step5: Exit.

Herefor the decision tree described the J48 algorithm to the auditdata in this paper.

3.3J48 Algorithm

For generating the pruned and unpruned C4.5 Decision-tree the J48 algorithm is implemented. The decision tree which is made by using the J48 created the group of the training data through using the information entropy concept. This J48 may also be implemented for the purpose of classification. It utilized the concept of every attribute from data may be utilized to prepare the decision through dividing the data intosome tiny sub-sets. This algorithm determines the normalizedinformation which is obtained from the dividing the data. Then prepare the leaf nodefor informing to select that class. So the J48 generated a decisionnode which is at top in the tree by using some class expectedvalue.

IV. LITERATURE REVIEW

In the paper [9], the data set ofKyoto 2006+ is used for the intrusion detection over the network by implement the J48 decision tree algorithm. This concept is implemented within the one of the tool of WEKA 3.6.10 for generating the decision tree in order to detect the intrusion. This method provides the accuracy of 97.23 percent approximately with the 99 percent of high true positive rate for the attacked and the normal both the packets. This type of simulation is proved helpful for the detection of the unknown attacks also. The result obtained by this analysis may provide the generation of the more instance then the previous one within the generated tree.

Within this paper [10], the analysis of the imbalance dataset is done by applying the tree-based real-time IDS for finding their effects. As thissuggested system may include the imbalanced data-set for the real-time IDS therefore this type of analysis is now proved necessary. In this paper, the accuracy and the reliability of the classification have been described for providing the optimal solution for intrusion detection.This approach will be enhanced for applying it for the multiclass classifiers is considered as the future enhancement of this paper.

[11] After implementation we are able to increase thedetection rate and decrease false alarm rate of Intrusion-detection-system by combining two data miningalgorithms C4.5 Decision Tree and Support VectorMachine. Comparison is done using various parameterslike Accuracy, Detection rate, False Alarm Rate.Building an effective intrusion detection models with goodaccuracy and real-time performance are essential.An attempt will be made in futureto classify types of attack into different categories likeDOS, Probe, U2R and R2L. A more efficient featureselection algorithm can be used in future.

This suggested paper [12], provides the implementation of the decision tree for the intrusion detection for detecting the several attacks with higher accuracy and thus this approach provides the 96.9 percent of the accuracy and less false alarm rate. In this simulation the KDD data-set is used to perform the analysis for obtaining the higher rate of detection of different intrusion over several networks. For this here used the Decision tree and Stratified weighted Sampling approach and also the Apriori algorithm is applied for the pre-processing of the data-set that is classified in three different phases such as data pre-processing phase as the first phase, fusion decision phase is the second phase and the data call back phase is the final phase of this approach.

Within this paper [13], for finding the known petterns or attack the decision tree and the Genetic Algorithm approaches have been applied. In the generalization and the detection of the latest patterns of attack thesetechnique have been provedefficient. The approach DTGA provides better accuracy and better performance than the Decision Tree and Genetic Algorithm's accuracy. The Memtic algorithms have been also successfully implemented along with DT-GAalgorithm and thus provide the new ways for the detection of the intrusions over the network.

In this paper [14], the most famous approach of the Network Intrusion-detection-systems which is called as the Snort rule-checking have been suggested.Here, suggested that the Snort prioritiesof thereal attacksmay befound accurate within the real-time regarding thenetworks having the high speed, through the decision-tree classifier by implementing this on the information for the extracted features such assource port, protocol, and the destinationport that is having the 99 percent accuracy. The issues of thisapproach of Snort mayinform the priorities which are dependent on the classes of the set of attackthat are utilized through the default rules. This approach proved beneficial for the anomaly detection intrusion-detection-system as a complement for this approach.

Within this paper [15], the hybrid machine learning approach is implemented and analysed for obtaining the better mechanism for the intrusion-detection-system by using the clustering and classification. In this analysis the DARPA KDD‟99 data-set is used for the detection of the rate of attackclasses such as Normal, DOS,Probe, U2R that are ranges from nearly 66 percent up to the 100 percent also providing the false alarm rate which is reduced till 0.102. The selection of the number of attributes for the processing may also influence theperformance of the system, less attributes provides the betterperformance of the system. For the future work of this approach may provide the

detection rate up to 100 percent and zero false alarm rate thus it will proved to be more efficient and effective system.

V. CONCLUSION

Since the intrusion-detection-system is the concept of protection operations which provide benefit to the system in the form of firewalls,UTM etc as the protection mechanisms. Thus the intrusion-detection-system is generally find the signatureof the attacks and then generates an alert regarding that attack.This paper has described the overview over the several intrusion-detection-systems to find the malignant nodes and also offers the security from those attacks. Regarding the objective of security the most appropriate approach of Decision tree have been described which is implemented for the intrusion-detection-system for detecting the intrusions in the better way. This paper is concluded that the various approached analysis have done by applying the decision tree algorithm in several ways in order to obtain the accuracy in detection.

REFERENCES

[1] WenyingFeng, Qinglei Zhang, Gongzhu Hu, Jimmy Xiangji Huang,"Mining Network data for intrusion detection through combining SVM"s with ant colony networks", Future Generation Computer Systems 37 (2014) 127-140", Elsevier 2014.

[2] LeventKoc , Thomas A. Mazzuchi, Shahram Sarkani,2012. The study of A network intrusion detection system based on a Hidden Naïve Bayes multiclass classifier. In Proceeding of the Elsevier Science Direct in Expert Systems with Applications.

[3] Gisung Kim, Seungmin Lee and Sehun Kim, 2014. The study of A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. In Proceeding of the Elsevier Science Direct in Expert Systems with Applications.

[4] Parveen Kumar,Nitin Gupta "A Hybrid Intrusion Detection System Using Genetic-Neural Network" International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 National Conference on Advances in Engineering and Technology (AET- 29th March 2014)

[5] SamanehRastegari, Chiou-Peng Lam, Philip Hingston2015 ,"A Statistical Rule Learning Approach to Network Intrusion Detection", IEEE.

[6] S. Gupta, P. Kumar and A. Abraham, "A Profile Based Network Intrusion Detection and Prevention System for Securing Cloud Environment," International Journal of Distributed Sensor Networks, Research Article: Hindawi Publishing Corporation, 2013, Article ID 364575

[7] Deepthy K Denatious and Anita John. "Survey on data mining techniques to enhance intrusion detection".In Computer Communication and Informatics (ICCCI), 2012 International Conference on Digital Object Identifier, p. 1–5.IEEE, 2012.

[8] Abid Khan, Prof. Kavita Burse, Prof. KavitaRawat, "An Efficient Intrusion Detection using J48 Decision Tree in KDDCUP99 Dataset", International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, Volume 6, Issue 2, February 2016.

[9] ShailendraSahu, B M Mehtre, "Network Intrusion Detection System Using J48 Decision Tree", Advances in Computing, Communications and Informatics (ICACCI), International Conference, 2015, IEEE.

[10] Dr.R.Balasubramanian, S.J.Sathish Aaron Joseph, "Intrusion Detection on Highly Imbalanced Big Data using Tree Based Real Time Intrusion Detection System: Effects and Solutions", International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 2, February 2016.

[11] VaishaliKosamkar, Sangita S Chaudhari, "Improved Intrusion Detection System using C4.5 Decision Tree and Support Vector Machine", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 1463-1467.

[12] Ms. TruptiPhutane, "Intrusion Detection System Using Decision Tree And Apriori Algorithm", International Journal of Computer Engineering and Technology (IJCET) Volume 6, Issue 7, July 2015, pp. 09-18.

[13] K.P.Kaliyamurthie , D,Parameswari , DR. R.M. Suresh, "Intrusion Detection System using Memtic Algorithm Supporting with Genetic and Decision Tree Algorithms", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 3, March 2012, ISSN (Online): 1694-0814.

[14] Adel Ammar, "A Decision Tree Classifier for Intrusion Detection Priority Tagging", Journal of Computer and Communications, 2015, 3, 52-58.

[15] Purushottam R. Patil, Yogesh Sharma and ManaliKshirasagar,"Performance Analysis of Intrusion Detection SystemsImplemented using Hybrid Machine Learning Techniques", nternational Journal of Computer Applications (0975 – 8887)Volume 133 – No.8, January 2016.