ISSN: 2321-7782 (Online)
Volume 2, Issue 2, February 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Paper / Case Study
Available online at: www.ijarcsms.com

A Review on attacks classification using decision tree algorithms

Bhura Parul[1]
Computer Science and Engineering
Parul Institute of Technology
Vadodara – India

Rajiv kumar Gurjwar[2]
Computer Science and Engineering
Parul Institute of Technology
Vadodara – India

Abstract: ***As the e-market growing day by day which is totally depends on security and major field which need security is network. The system that tries to detect malicious activity is NIDS. NIDS uses classification of malicious activities. There are different algorithms in data mining to classify the attacks done by NIDS such as J48, SVM, Naïve Bayes, Random Forest, Random Tree, Rotation Forest, kstar, KNN, Multiboosting, Random Subspace etc. In existing system broadly four types of attack discussed but there are some problems mentioned below related to classifier algorithms: In layered approach more than one algorithm used for each layer. As we discussed RFA gives good classification rate (i.e. 99.16%, 99.4%, and 99.25% for PROBE, R2L and DOS layer respectively) but for U2R layer it is 53.79%. So there is need to modify RFA to get higher classification rate.***

Keywords: ***Layered Approach, R2L, PROBE, U2R, DOS, NORMAL, Classification, Classification Tree, Multiclass Attributes, IDS.***

I. INTRODUCTION

In today's world internet has become an indispensable part of one's everyday life. Most of the routine transactions are available online, either it be information regarding a subject or other service like reservation, online shopping also known as e-shopping. Network security became a one of the most important task for every organization. Security is concerned with people trying to access remote services for what they are not authorized too. It also deals with the problems of legitimate messages being captured and replayed, and with people trying to deny that they sent certain messages. A brief description of Network Security, Intrusion Detection System, Host based Intrusion Detection, Network Intrusion Detection Based System, Misuse Detection Based, Anomaly Detection Based are explored in the following sections:

1.1 Intrusion Detection System (IDS)

IDS are the process to analyze the network traffic. It can be treated as pattern recognition problem-distinguishing between different categories of attack. According to S.A.Joshi, there are two types techniques: first is Signature based & second is anomaly based. [1][8]

1.1.1 Signature-based detection techniques:

In this technique, IDS inspects the monitored packets on the basis of proof of attacks based on existing model & predefined rules for specific known attacks. This technique detects only known & predefined attacks only.

1.1.2 Anomaly based detection techniques:

In this technique, does not require the predefined knowledge of attacks so this also detects the new attacks. Here IDS inspects the system activities on the basis of detection any deviation from existing model of normal & expected behavior through the

system. According to system, the source of input information may be on a host, network, host & network. Based on this IDS again classified into three categories such as Network based, Host based & Hybrid based IDS. [1][8]

II. Related Work of Intrusion Detection Techniques

Juan Wang et al. (2009) [2] in paper has studied about C4.5 decision tree algorithm for intrusion classification. This decision tree convert into rule is used to classify the attack. Here technique of creating rules, information gain ratio used instead of information gain. This rule is used to classify the attack. This technique gives the accuracy rate over 90%.

Dr.T.Subbulakshmi et al. (2013) [3] in paper has discussed intrusion detection is based on data mining algorithms. Attack classification based on multiple learning algorithm such as applied on different layer for classification the R2L, U2R, Probe, DOS attack. Here four classification rate obtained by this experiment. Subbulakshmi used ten data mining algorithm & gives best classification rate using Random forest & Rotation forest algorithm. The classification rate for R2L & DOS Attacks is good using feature selection. But for U2R classification rate is very law. This result obtained using KDD CUP'99 dataset as input. Proposed system used KDD CUP'99 dataset as input dataset but according to Mahbod Tavallaee [4], KDD CUP'99 data set not performance good according to real world application.

Zhao Yongli et al. (2013) [5] in paper has described that improved feature selection algorithm to identify most appropriate subset of feature for certain attack. MAHALANOBIS distance feature ranking based method used & enhance exhaustive search to select a better combination of features. Input is KDD CUP'99 dataset. SVM & KNN algorithm is used as classifier. Experiment results give high classification rate & low misclassification with reduce feature subsets. But proposed approach on two class classification problem based on labeled data. The future work refers to multi-class classification problem, large amount of unlabeled data for training & incremental learning problems.

A.Kumaravel et al. (2013) [6] in paper presented multi layer system to achieve high efficiency, also improve the attack classification accuracy. In layered approach two stages defined. In first stage input data is classified in either normal or attack. In second stage, classify the attacks in four classes & type of attack. Proposed approach gives Classification rate for DOS layer, PROBE layer, U2R layer, R2L layer are high. But this result is based on rules such as JRip.

Poonam Gupta et al. (2013) [7] in paper they introduced Decision tree C4.5 algorithm, NSL-KDD dataset & WEKA tool to improve accuracy of attack detection. This experiment gives better performance for known attack but cannot detect unknown attack.

S.A.Joshi et al. (2013) [8] in paper they use Data mining, Feature selection, Multiboosting algorithm techniques for attack detection. Here Feature selection contributes to improve the overall accuracy, reduces the number of false alarms & improve the detection rate of attacks in training data. Multiboosting gives better performance than bagging. This experiment gives 94% accuracy using gain ratio & high detection rates for U2R & R2L attacks. Proposed system used KDD CUP'99 dataset as input dataset but according to Mahbod Tavallaee [4], KDD CUP'99 data set not performance good according to real world application.

Jaisankar N et al. (2013) [9] in paper the system consists of four modules: clustering module, outlier detection module, classification module & decision module. Here KDD-cup99 dataset as input, SVM & fuzzy EC4.5 algorithms are used. The proposed system reduces the time complexity, improves the overall detection accuracy & minimizes the false alarm rate considerably.

III. Conclusion and Future Work

Nowadays Intrusion Detection System is very necessary. The Layered approach is very effective system & gives good result. But in layered approach different algorithm used for each layer. Random forest algorithm gives good result for every layer but for U2R layer, classification rate very law. Modified the random forest & can improve the result for U2R layer. It is possible to merged layer & used only one algorithm without affecting the result. Future work of this system are implementing in

programming language like C & java. In this system used KDDcup99 dataset but we also can try improve this system using real world data.

References

1. P.S.Prabhu, "Network Intrusion Detection Using Enhanced Adaboost Algorithm", International Journal of Communication and Engineering, Volume o3-No.3,Issue:02,2012.
2. Juan Wang, Qiren Yang, Dasen Ren, "An Intrusion detection algorithm based on decision tree technology", Asia-paific conference on Information Processing,2009.
3. Dr.T.Subbulakshmi, Ms.A.Farah Afroze, "Multiple Learning based Classifiers using Layered Approach and Feature Selection for Attack Detection", IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology,2013.
4. Mahbod Tavallaee,et. al, "A Detailed Analysis of the KDD CUP 99 Data Set" IEEE Symposium on Computational Intelligence in Security and Defence Applications(CISDA),2009
5. Zhao Yongli et. al, "An Improved Feature Selection Algorithm Based on MAHALANOBIS Distance for Network Intrusion Detection System 2013.
6. A.Kumaravel, M.Niraisha, "Multi-Classification Approach for Detecting Network Attacks" IEEE Conference on Information and Communication Technologies ,2013.
7. Poonam Gupta, S.R.Tandan, Rohit Miri, "Decision Tree Applied For Detecting Intrusion", International Journal of Engineering Research & Technology(IJERT),Vol.2 Issue 5, 2013.
8. S.A.Joshi, Varsha S. Pimprale, "Network Intrusion Detection System(NIDS) based on Data Mining", International Journal of Engineering Science and Innovative Technology(IJESIT),Volume 2,Issue 1, 2013.
9. Jaisankar N, Swetha Balaji,Lalita S, Sruthi D, "Intrusion Detection System Using K-SVMeans Clustering Algorithm",International Journal of Computational Engineering Research(IJCER),2013.

