International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 07 | July 2020 www.irjet.net p-ISSN: 2395-0072

NETWORK SECURITY AND INTRUSION DETECTION SYSTEM USING DATA MINING TECHNIQUES

[1]UMASANKAR .K, [2]RAJESWARI.P,

[1] Assistant Professor, Department of Computer Science.
[2] Master of Philosophy, Department of Computer Science.
PRIST University Kumbakonam Campus, Thanjavur,Tamil Nadu, India

---

Abstract - *with the outstanding increase in the use of computers and computer resources over the network and the development of applications on different areas, the main focus is on computer networks security. The intrusion detections of various networks attacks are actually an important element of network security systems. The main role of the intrusion detection system is to detect and prevent unauthorized access or damage to network resources. This paper develops secure data mining based intrusion detection system on both Network Intrusion Detection System to monitor all network traffic passing on segment, where a alert system is installed to alert the administrator of any illegal signature based activity or anomaly users, and Host Intrusion Detection to monitor inbound and outbound packets from a network device, and will alert the user or network administrator of suspicious user behavior detected. The framework designed addresses negative effects of its weaknesses so as to enhance operational effectiveness. The most importance of intrusion detection systems and the old techniques, type, characteristics and limitations would be given special attention in this research.*

Keywords: Anomaly Detection, Network Security, Data mining Intrusion Detection System, Misuse Detection, Clustering

I. INTRODUCTION

With the rapid growth of the Internet, there has been an improvement in the way people live but because of threats, perpetrated by individuals or an organization. They are used to violate network security. Security means the level of protection provided by a network or system. The main objectives of security are confidentiality, integrity and data availability [1]. Network attacks can be called Intrusion. Intervention means any set of malicious objects that attempt to compromise information security purposes. In the early days, the only common methods used for networks such as encryption, logs, private network etc but were not sufficient to fully protect the network. It is difficult to rely entirely on fixed defense strategies. This increases the need for a robust system, which can be a system of monitoring and directing illegal activities. The development of a powerful network protection system has therefore been introduced and is known as the Intrusion Detection System. The intrusion detection system collects information online from the network after which it monitors and analyzes this information and incorporates it into normal and risky activities, providing results to the system administrator [2]. IDS is a place, where Data mines are widely used, this is due to unequal limits, flexibility and alidity.

TYPES OF ATTACKS

A. Dos attack

A denial-of-service attack or distributed denial-of-service attack is an effort to make a computer resource out of stock to its indented users [32].This type of attack slows down the system or

shut down the system that interrupt the service and deny the authorized user. Due to this attack high network traffic occurs.

B. User to Root Attack (U2R)

In this attack, the attacker steals the password, dictionary and then finds the root of accessing the system.

C. Probing

That report allocation process of using through to gather information or discover well-known vulnerabilities. An attacker who has a record, of which machines and services are available on a notorious system, can make use of this information to look for delicate points.

D. Remote to User Attack (R2U)

Remote attack, an aggressor has the competence to send packet to a machine over arrangement is not supported have and report on that machine, make use of some vulnerability to some small techniques process

E. some dropping attack

Common dropping is a complex layer hit consisting of capturing packets from the network transmitted by others' computers and reading the sensitive information like passwords, session tokens, or any kind of confidential information.

II. LITERATURE SURVEY

Intrusion and attack methods have now evolved with large amounts of network traffic information and complex behavior beyond traditional intrusion (IDS). High identity reliability, high false positive rates and high runtime benefits from existing cloud IDs. This paper introduces a structure that detects distribution intrusions focused on machine learning for cloud environments. The proposed system is intended to be implemented in the cloud with network providers on the edge of the service provider. Incoming network traffic interrupts the network router at the physical layer. Enables base classification system to pre-process traffic on cloud routers using Naval Sliding Windows Algorithm (SWA). Irregular network traffic data is synchronized to the central storage database on each side of the router per session. The next step in determining each type of attack is to conduct a final classification phase based on the random forest. Security systems can detect abnormal activity or behavior on networked systems. However, in very rare cases, traditional security systems, including firewalls and anti-viruses, may not work properly. An accurate and smart intrusion detection system (IDS) is needed to solve this problem. In recent decades, various methods and strategies have been introduced to overcome the shortcomings of IDS such as high false alarm rates, poor reliability and time consuming assets. The paper proposes a hybrid infiltration detection (HII) method based on DT and KNN . A functional selection method is used to extract structured information from the NSL-KDD data set to improve the performance of the proposed approach.

III. DATA MINING ASSISTS IN INTRUSION DETECTION

The suggested proposed scheme is illustrated in figure. 1.

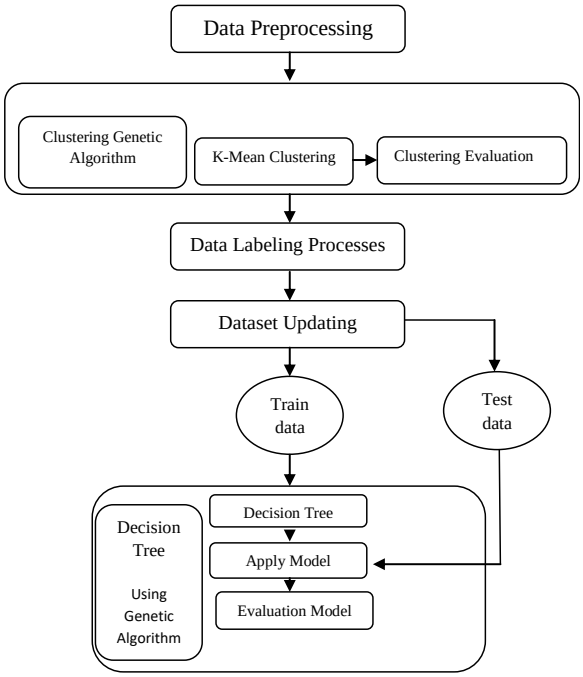

Figure 1. Datamining Based IDS Architecture

(i) Data Preprocessing

Upload Data Set: The desired data set will be uploaded at this level used for data mining. Properties Selection: The data set properties in step 1 are selected. The basis of my choice is a simple feature and sub-set of features. Feature characters are described at this level: 'character' determines the identity of an attribute whether it is casual, unique or labeled Converting nominal data to numerical data: In clustering method, nominal data must be translated into numbers. Normalization: Feature values are normalized at this level based on the Z-transition method.

(ii) Categorization

Sorting is the charge of pleasing each and all Specifies simple and unusual methods for a particular class of examples and new instructions for the dataset under consideration. It is effective for both abuse detection and irregular detection, but is often used to detect abuse. Classification The dataset is classified as a predetermined set. It is less capable of detecting intrusions than clustering. Various classification methods are used in IDS such as Decision Tree, Innocent Bay Classification, K-Near Neighbor Classification, and Support Vector Machine.

(iii) K-Means Enhancement

The K-Means is a simply and wide range of using clustering technique proposed by James McCain. In this algorithm, the user specifies the cluster K number, i.e. classifies the number of pre-defined clusters. K- Means the first step of clustering is to select k examples as the center of the clusters. For example Assignment, measure the distance between centroid and each case using Euclidean distance and assign each data point to the cluster according to the minimum distance. The K-Means algorithm runs low for small datasets. Maximum execution time when data point is maximum. It's a fast repeating algorithm, but it's external and noise sensitive.

(iv) Data Labeling & Dataset Updating

The structure proposed in this report applies to both monitored and unselected policies. The previous section uses an optimized clustering approach for cluster data. At this point the cluster member is tagged, and then the set data classification pattern and systems that detect system intrusions are subject to the labels specified for each data set. Cluster members are called by cluster names.

(v) Sub-Dataset Train & Test

Following changes in the primary data collection, the train and test subgroups were removed. Due to this, the updated data sets of train and test subcommittees will increase by 70 and 30 percent, respectively. This approach is used to select different types of data, e.g. Sample without placement.

Broke through the data of eq(3-5).

D=Dataset, tr= train, ts= test

$$D = tr_{data} \cup ts_{data} \quad (3)$$

$$\{tr_{data}\} = 70\% \; of \; D \quad (4)$$

$$\{ts_{data}\} = D - \{tr_{data} \mid tr_{data} \cap ts_{data} = \emptyset$$

(vi) Enhancement of Decision Tree

DT analyzes the data applied to the system to identify the data from above. Each node has a DT feature label. When a new node is created at a given time, a function is selected to increase its partition strength according to the records assigned to a particular sub-tree. The Gini index determines the strength of this division. Gini coefficient is a method of calculating the impurity of nodes in the form of an equation (6) shows:

$$GINI\, x = 1 - (k|x)_{2k}$$

$(k|x)_2$ is the sum of class k of node x record numbers to all node data. The worst division occurs when the less data in every category of the specified node is obtained in equal amounts. The more ideal and homogeneous is the larger the Gini coefficient of the node.

$$FF = Accuracy - Clasification\ Error$$

IV RESULTS AND DISCUSSION

Many devices have safety violations that can quickly make them vulnerable and unable to be addressed. In addition, extensive work has been carried out on intrusion detection methods, which are still considered incomplete and not a complete method for intrusion. The role of network administrators and security experts also has become a highly priority and challenge So more stable systems cannot be replaced. So better performance is need for each intrusion detection methods. The proposed Data mining Based Techniques provides better performance compared to other existing methods.

V. CONCLUSION

Through this research which is focusing on optimizing Kmean's clusters and DT classification algorithms are used to find out the intruders in a precise manner. For this feature the GA and the higher accuracy of the used classifier optimize total runs and trust. One of the most important characteristics is the applicability of this method for both supervised and unsupervised approaches. The idea is to use the suggested model to test its generalization for future research in the field of network intrusion monitoring in different datasets.

REFERENCES

[1]. Daniel A. Keim, "Intrusion Prevention Systemd" IEEE Trans. Visualization and Visual Data Mining, vol. 8,Jan-Mar 2002. *(references)*

[2[. FuzailMisarwala, KausarMukadam, and KiranBhowmick, "Applications of Data Mining in Fraud Detection", vol. 32015.

[3[. Esther Nowroji., Vanitha., "Detection Of Data Using IP Address Recognition Technique", vol. 4, International Journal for Research in Applied Science & Engineering Technology, 2016.

[4[. Ahmad FIRDAUS, Nor Badrul ANUAR, Ahmad KARIM, MohdFaizalAb RAZAK, "Discovering optimal features using static analysis and a genetic search based method for malware detection" Frontiers of Information Techonology and Electronic Engineering, 2018.

[5[. JavvajiVenkataramaiah, BommavarapuSushen, Mano. R, Dr. GladispushpaRathi, "An enhanced mining leading session algorithm for fraud app detection in mobile applications" International Journal of Scientific Research in Engineering., April2017.

[6]. Avayaprathambiha.P, Bharathi.M, Sathiyavani.B, Jayaraj.S "To Detect Fraud Ranking For Mobile Apps Using SVM Classification" International Journal on Recent and Innovation

Trends in Computing and Communication, vol. 6, February2018

[7]. Suleiman Y. Yerima, SakirSezer, Igor Muttik, "Android Malware Detection Using Parallel Machine Learning Classifiers", 8th International Conference on Next Generation Mobile Applications, Services and Technologies,Sept.2014.

[8]. SidharthGrover,"Malware detection: developing a system engineered fair play for enhancing the efficacy of stemming search rank fraud", International Journal of Technical Innovation in Modern Engineering &Science,Vol. 4, October2018

[9]. PatilRohini, Kale Pallavi, JathadePournima, KudaleKucheta, Prof.PankajAgarkar,"MobSafe: Forensic Analysis ForAndroid Applications And Detection Of Fraud Apps Using CloudStack And Data Mining", International Journal of Advanced Research in Computer Engineering &Technology,Vol. 4, October2015

[10]. Neha M. Puram, Kavita R. Singh,"Semantic Analysis of App Review for Fraud Detection using Fuzzy Logic", International Journal of Computer & Mathematical Sciences,Vol. 7, January2018

[11]. VivekPingale, LaxmanKuhile, Pratik Phapale, Pratik Sapkal, Prof. Swati Jaiswal,"Fraud Detection & Prevention of Mobile Apps using Optimal Aggregation Method", International Journal of Advanced Research in Computer Science and Software Engineering,Vol. 8, March2016.

[12]. D.Janet, Vikrant Chole ,"A Review on Ranking Based Fraud Detection in Android Market", International Journal of Science and Research,Vol. 6, January2017.

[13]. Monika Pandey, Prof. TriptiSharma,"Fraud App Detection using Fuzzy Logic Model Based on Sentiment of Reviews", International Research Journal of Engineering and Technology,Vol. 5, Sep2018.

[14].MahmudurRahman, MizanurRahman, BogdanCarbunar, and DuenHorngChau,"Search Rank Fraud and Malware Detection in Google Play", IEEE Transactions on Knowledge and Data Engineering,Vol. 29, June2017.

[15]. Tahura Shaikh#1, Dr. DeepaDeshpande,"Feature Selection Methods in Sentiment Analysis and Sentiment Classification of Amazon Product Reviews", International Journal of Computer Trends and Technology (IJCTT),Vol. 36, June2016.