Intrusion Detection System Using Data Mining

Paresh Goliwale, Vishal Gupta, Atish Johre, Sneha Bendale

Abstract— *Security of any organization is the primary concern nowadays. But internal intrusion is the big problem as the person knows all the internal information of the organization, so the person can launch the attack from internally without firewall detection. And as the people in organization has a tendency to share the passwords to the colleagues it is very easier for the internal person to launch the attack from inside of the organization. Tracking this user becomes really difficult for firewall because it is mainly focused on the attack happening through other network. To detect this attack anomaly based detection is one of the technique which is low cost as compared to other detection/prevention methods. Anomaly based method creates the image of behavior of each user and if user does activities which are not in that image of behavior it is considered as a malicious activity and threat to the system.*

Keywords: Data mining, insider attack, intrusion detection , system call (SC), users' behaviors.

I INTRODUCTION

We have a tendency to area unit currently living in an exceedingly borderless world wherever nothing is on the far side reach. The profound and fast technology growth has given rise to new vulnerabilities and threats to the mechanization era. Additionally, the dependency on network and net amenities to support the growing want for on-line services has positively will increase the cyber rate. Within the late 70's and early 80's, watching user activities for any malicious or uncommon behaviors were done manually exploitation the written audit logs, but this has evolved considerably. Threats and attacks are getting a lot of frequent and should be handled in an exceedingly a lot of economical and effective manner. these days we have a tendency to see new attack technique in an exceedingly daily, therefore, there should be a mechanism to observe and management these activities. There's a precise want for brand new sort of protection against this new hazard. there'll ne'er be enough or an excessive amount of security enforced, particularly with all the web services created offered through internet; but the safety enforced ought to be reliable and at an equivalent time won't jeopardize the performance of a network or system. IDS offer a second layer of defense before typical security technique equivalent to authentication and access management. As a result of the importance of maintaining confidentiality, accessibility and therefore the integrity of our Worthiest assets that is that the data, IDS has become a necessity. Nonetheless before investment in associate degree IDS, it's vital to grasp the present infrastructure and therefore the actual desires of the businessman. There are a unit many sorts of IDS with its own set of classification out there within the market; host or network primarily based, signature or anomaly primarily based, active or passive watching, time period or interval process and at last can the implementation be centralized or distributed. All the on top of classifications have their own blessings and downsides. Moreover, several researches are dispensed within to perceive the topic higher and within the long haul to supply a more practical and economical IDS.

A. MOTIVATION

The aim of the system is to sight bound well-known intrusion attacks on the host system and show warnings to the user and additionally store data concerning the IP addresses and permit the traffic supported that data.

B. PROBLEM STATEMENT

Basically, this project seeks to answer the question: "Is it sensible and acceptable to mix intrusion detection and response with rhetorical management of collected knowledge among one IDS in today's networks?" The difficulty we are going to address during this analysis is three-fold. First, will associate degree IDS gather helpful rhetorical proof throughout associate degree attack while not impacting its primary mission of sight and respond? Second, what's needed to produce an appropriate case file of rhetorical information? And, finally, in an exceedingly sensible implementation, will associate degree IDS be enforced that may accomplish each its primary mission and, at an equivalent time, collect and manage forensically pure proof which will be utilized in a legal setting? There are a unit many difficulties in addressing these problems. First, the theoretical needs of associate degree IDS in terms of acting its primary mission could also be at odds with the wants of assembling and protective rhetorical proof. The first mission of associate degree IDS is to sight and answer security incidents. The definition of a security incident ought to be, a minimum of partially, determined by the organization's security policy. Therefore, the careful definition of the IDS' primary mission is part determined by the safety policy, not by some overarching customary or generic procedure. The result's that there is a large inequality among needs for associate degree IDS from organization to organization. That contrasts considerably with the comparatively static set of needs for developing and managing proof to be used in an exceedingly due process of law. A second problem is that the IDS, by design, doesn't manage its data within the sense that a forensics system will. there's a demand among a rhetorical system (automated or not) for, among different things, the upkeep of a sequence of custody whereby all proof is accounted for and its integrity authenticated to from the time of its assortment to the time of its use in an exceedingly due process of law. The third problem deals with the design of the IDS. The flexibility of a program to perform wide disparate tasks (in this case detection associate degreed response moreover as rhetorical management of data) implies associate degree design that will or might not be gift presently in an IDS. Thus, there develops the necessity for a regular design for intrusion detection systems that are also capable of rhetorical knowledge management.

C. SCOPE

The system frames bound rules primarily based upon the input given by the user. It then permits traffic inwards or outward primarily based upon the principles. The system additionally detects bound well-known attacks and offers warnings to the user.

II LITERATURE SURVEY

A. OVERVIEW

a) Chris Clifton Gary Gengo explains, one aspect of constructing secure networks is identifying unauthorized use of those networks. Intrusion Detection systems look for unusual or suspicious activity, such as patterns of network traffic that are likely indicator of unauthorized activity. However, normal operation often produces traffic that matches likely “attack signatures”, resulting in false alarms. We are using data mining techniques to identify sequences of alarms that likely result from normal behavior, enabling construction of filters to eliminate those alarms. This requires cheap cost for some platforms, by constructing an anomaly based intrusion detection system. The proposed approach has preliminary results identifying common patterns in alerts from a particular platform.

b) Fang-Yie Leu, Kun-Lin Tsai elaborates, currently, most computer systems use user IDs and passwords as the login patterns to authenticate users. However, many people share their login patterns with coworkers and request these coworkers to assist co-tasks, thereby making the pattern as one of the weakest points of computer security. Insider attackers, the valid users of a system who attack the device internally, are tough to detect when you consider that maximum intrusion detection systems and firewalls pick out and isolate malicious behaviors launched from the outdoor international of the machine simplest. in addition, a few studies claimed that reading machine calls (SCs) generated by way of instructions can perceive these commands, with which to as it should be stumble on assaults, and attack patterns are the features of an attack. consequently, in this paper, a security machine, named the inner Intrusion Detection and protection device (IIDPS), is proposed to hit upon insider attacks at SC stage via the usage of statistics mining and forensic strategies. The IIDPS creates users’ private profiles to maintain song of customers’ utilization habits as their forensic features and determines whether or not a legitimate login consumer is the account holder or no longer with the aid of evaluating his/her modern-day computer usage behaviors with the styles collected within the account holder’s personal profile.

c) Krishna Kant Tiwari, Susheel Tiwari, Sriram Yadav explains, in these days an increasing number of public and com-mercial services are used through the Internet, so that security of information becomes more important issue in the society information Intrusion Detection System (IDS) used against attacks for protected to the Computer net-works. On another way, some data mining techniques also contribute to intrusion detection. Some data mining techniques used for intrusion detection can be classified into two classes: misuse intrusion detection and anomaly intrusion detection. Misuse always refers to known attacks and harmful activities that exploit the known sensitivity of the system. Anomaly generally means a generally activity that is able to indicate an intrusion. In this paper, comparison made between 23 related papers of using data mining techniques for intrusion detection. Our work provides an overview on data mining and soft computing techniques such as Artificial Neural Network (ANN), Support Vector Machine (SVM) and Multivariate Adaptive Regression Spine (MARS), etc.

B. EXISTING SYSTEM

Network Security has become the key foundation with the tremendous increase in usage of network-based services and information sharing on networks. Intrusion poses a serious risk to the network security and compromise integrity, confidentiality & availability of the computer and network resources. Data mining technique has been widely applied in the network intrusion detection system by extracting useful knowledge from large number of network data. In this paper a hybrid model is proposed that integrates Anomaly based Intrusion detection technique with Signature based Intrusion detection technique is divided into two stages. In first stage, the Signature based IDS SNORT is used to generate alerts for anomaly data. In second stage, data mining techniques the hybrid IDS model is evaluated using KDD Cup Dataset. The proposed assemblage is introduced to maximize the effectiveness in identifying attacks and achieve high accuracy rate as well as low false alarm rate.

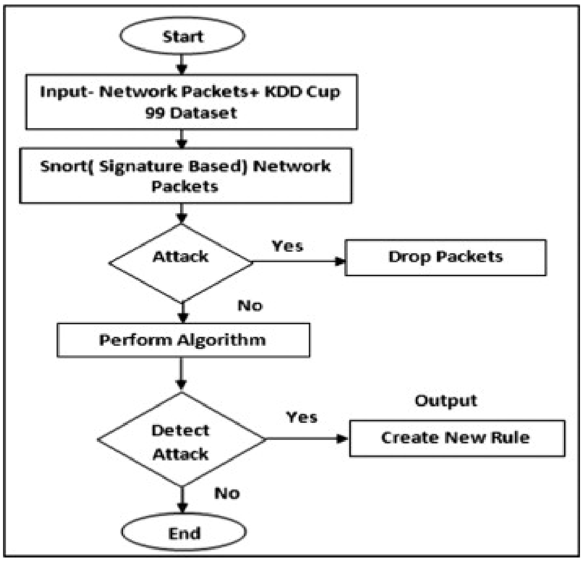

FIGURE: SIGNATURE BASED DETECTION

III PROPOSED SYSTEM

A. OVERVIEW

Intrusions are the activities that violate the security policy of system. Intrusion Detection is the process used to identify intrusions. An intrusion detection system (IDS) is a device or software application that monitors network or system activities for malicious activities or policy violations and produces reports to a management station. In our proposed system we focus on banking scenario to detect the intrusion or malicious activities. We want the network log data that contain the all information like the transaction, the status of intrusion. Then we upload that data to system to detect the percentage of intrusion and generate the report. Many time the inner person attack on system through network then we can't find that some attack happen on network. But use of ids we can detect any inner and outer attacks from network. The proposed System is introduced to maximize the effectiveness in identifying attacks and achieve high accuracy rate. Data mining techniques is evaluated using KDD Cup Dataset. The proposed assemblage is introduced to maximize the effectiveness in identifying attacks and achieve high accuracy rate as well as low false alarm rate.

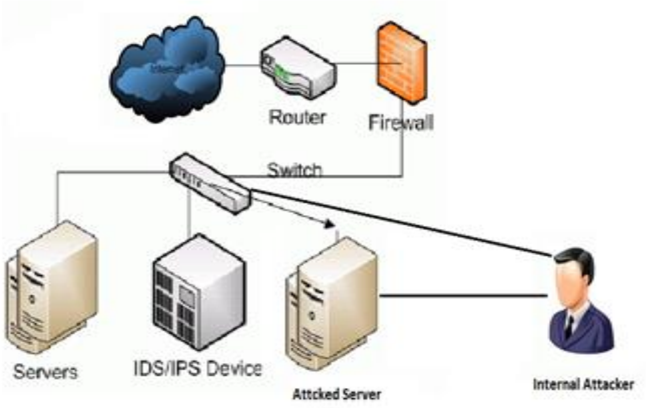

FIGURE: ARCHITRCHTURE OF IDS

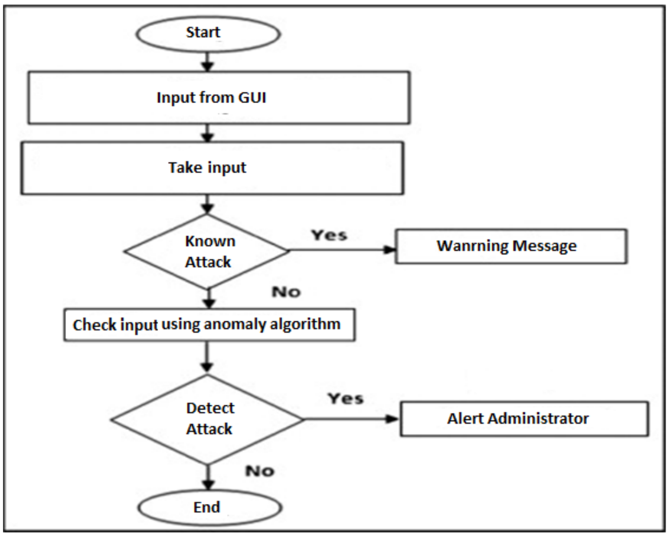

FIGURE: IDS ALGORITHM

The architecture of the system shows that if the attack is happened from the inside of the network the firewall will not be able to detect it the IDS server will detect it as it is employed within the network of the system and it will also use the K-means data clustering method to separate the abnormal activity from the normal activity.

B. HARDWARE AND SOFTWARE SPECIFICATIONS

Hardware Specification:

Recommended Requirements: -

Processor: Intel i3/i5/i7 /AMD FX Series

Ram: 4 GB or higher

Software Specification:

1. Operating System: Windows

2. Frontend: Web Application

IV APPLICATION

The proposed system is a device or software application that monitors a network or systems for malicious activity or policy violations.

V SUMMARY

We've conferred the main points of a replacement approach known as Outlier Detection approach to sight the intrusion within the network. Our coaching model consists of huge datasets with distributed surroundings that improves the performance of Intrusion detection system. The projected approach is additionally being tested with the KDD datasets that area unit received from planet. The machine learning approaches sight the intrusion within the network with large execution time. Storage to predict the compared to the projected IDS system that takes less execution time and storage to check the dataset. Here during this study, the performance of projected IDS is healthier than that of different existing machine learning approaches and may considerably sight most anomaly knowledge within the network. In future, the projected work is presumably used for varied distance computation perform between the trained model and testing knowledge. Our analysis work is thought-about to enhance the potency of IDS in an exceedingly higher manner.

VI CONCLUSION

We have presented the details of a new approach called Internal Detection approach to detect the intrusion in the computer network. The performance of proposed IDS is better than that of other existing machine learning approaches and can significantly detect almost all anomaly data in the computer network. In future, the system can be made more intelligent that it will distinguish which is active and passive attack on the system. Our research work can be considered to improve the efficiency of IDS in a better manner.

VII REFERENCES

[1] William Stallings, “Cryptography and Network Security”, Principles and Practices, Third Edition.

[2] D. E. Denning, &quot; An intrusion-detection model & quot;. IEEE Transactions on Software Engineering, Vol. SE-13(No. 2):222-232, Feb. 1987.

[3] Stephen Northcutt, Judy Novak, “Network Intrusion Detection”, Third Edition, Pearson Education 2003.

[4] Kaining Lu Zehua Chen Zhigang Jin Jichang Guo.” An Adaptive Real-Time Intrusion Detection System Using Sequences of System Call”, CCECE 2003.

[5] Shingo Mabu, Ci Chen, Nannan Lu, Kaoru Shimada, and Kotaro Hirasawa,” An Intrusion- Detection Model Based on Fuzzy Class Association-Rule Mining Using Genetic Network Programming”, IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, Vol. 41, No. 1, January 2011.

[6] R Rangadurai Karthick, Vipul P. Hattiwale, Balaraman Ravindran,” Adaptive Network Intrusion Detection System using a Hybrid Approach ”, IEEE 2012.

[7] Vincent F. Mancuso, Dev Minotra, Nicklaus Giacobe, Michael McNeese and Michael Tyworth ” ids NETS: An Experimental Platform to Study Situation Awareness for Intrusion Detection Analysts” ,IEEE International MultiDisciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, New Orleans, LA, 2012.

[8] Gholam Reza Zargar, Tania Baghaie, “Category-Based Intrusion Detection Using PCA”, Journal of Information Security, 2012.

[9] Neethu B, “Classification of Intrusion Detection Dataset using machine learning Approaches”, International Journal of Electronics and Computer Science Engineering, 2012.