International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-2 Issue-5, June 2013

Concept & Proposed Architecture of Hybrid Intrusion Detection System using Data Mining

Sandeep Kumar Singh, Nishant Chaurasia, Pragya Sharma

Abstract*— now day's security is the primary concerned in the field of computer science. Intrusion detection system provides stronger security services with the help of rules. This paper has developed a security model based on hybrid intrusion detection system using data mining approach. Proposed security model is the combing network based and host based with efficient data mining approach to detect any type of intrusion which coming from public network or occurring in computer system. Basically this model work on misuse and anomaly detection mode, it will use an approach to extract features from arriving data packets and will apply data mining algorithm to get the rule for match normal and abnormal behavior.

Index Terms*— Intrusion Detection System, Security, Network System, Host System, Data mining, association.

I. INTRODUCTION

A network IDS is one that monitors network traffic and will raise an alert on suspected intrusion packets or act on those packets such as blocking them or resetting concerned connections. A host
Based IDS is one that monitors the activities and the data residing on a particular host. An IDS achieves its function using different detection models. The misuse detection model is based on defined signatures of known attacks. Known characteristics of attacks (such as certain contents in packets) are coded as signature and the IDS match packets against them to detect possible intrusions. This method is good for detecting known attacks if the signatures are well written. Poorly written signatures may result in large number of false positives. The signature database may also grow to a point that will affect the effectiveness of detection. This nevertheless is the most popular detection method currently. Another model is anomaly detection, or statistical detection. In this model a baseline is first established by observing "normal" activities for a "training" period. Intrusion is then achieved by observing behaviors that depart from the normal profile. The advantage of this model is that it is capable of detecting new or unknown intrusion. However most real systems will have variants in their so called normal profiles and this detection method can be prone to high false positive rates.

Manuscript published on 30 June 2013.
***** Correspondence Author (s)
Sandeep Kumar Singh*, Department of Computer Science and Engineering, RGPV Bhopal, India.
Nishant Chaurasia, Department of Computer Science and Engineering, RGPV Bhopal, India.
Pragya Sharma, Department of Computer Science and Engineering, RGPV Bhopal, India.

II. RELATED WORK

In [1] two technique of intrusion detection system have discussed one is host based and other is network-based intrusion detection techniques.

In this they demonstrate how the two can work together to provide additionally effective intrusion detection and protection.

Basically here an hybrid IDS suggested, which combines network and host IDS, with anomaly and misuse detection mode, utilizes auditing programs to extract an extensive set of features that describe each network connection or host session, and applies data mining programs to learn rules that accurately capture the behavior of intrusions and normal activities. The researchers proposed a number of techniques such as (firewall, encryption) to prevent such penetration and protect the infrastructure of computers, but with this, the intruders managed to penetrate the computers. IDS has taken much of the attention of researchers, IDS monitors the resources computer and sends reports on the activities of any anomaly or strange patterns. In [2] they explain the stages of the evolution of the idea of IDS and its importance to researchers and research centers, security, military and to examine the importance of intrusion detection systems and categories, classifications, and where can put IDS to reduce the risk to the network. In [3] a host based intrusion detection system for Microsoft Windows XP environment has suggested. This method that had used in the study was applying intrusion detection pattern matching technique on the Security Event Log File for Microsoft Windows XP. The intrusion had identified when there was matching of intrusion pattern that is create with Security Event Log in Microsoft Windows XP. In [5] this paper, classifications of intrusion detection and methods of data mining applied on them were introduced. Then, intrusion detection system design and implementation of based on data mining were presented. Such a system used APRIORI algorithm to analyze data association, which is the most influencing algorithm in mining Boolean association rules continuity item muster, with recurrence arithmetic based on idea of two period continuity item muster as core. Intrusion detection is the process of identifying and responding to suspicious activities targeted at computing and communication resources, and it has become the mainstream of information assurance as the dramatic increase in the number of attacks. Intrusion detection system (IDS) monitors and collects data from a target system that should be protected, processes and correlates the gathered information, and initiates responses when evidence of an intrusion is detected. In [6] designed and implemented a host-based intrusion detection system, which combines two detection technologies, one is log file analysis technology and the other is BP neural network technology.

Log file analysis is an approach of misuse detection, and BP neural network is an approach of anomaly detection. By combination of these two kinds of detection technologies, the HIDS that have implemented can effectively improve the efficiency and accuracy of intrusion detection.

Problem Formulation: it is already known that intrusion detection system is important security mechanism for network whatever it's network based or host based because in some manner network based system have produced batter result and some manner host based system has produced batter result it will depend on the type of need if system are working in network then network based IDS required and if system are working stand alone then host based IDS will required There is no reason why both NIDS and HIDS can not be used in conjunction as a strong IDS complimentary strategy. One common disadvantage of most data mining techniques is the extensive amount of time required for training and learning the model being inspected. that association rule mining has number of drawback like generate too many rules and sometimes these are even trivial rules; The association rule are not expressions of cause/effect rather they are descriptive relationships in particular databases, so there is no formal testing to increase the predictive power of these rules some more points is also necessary for the study that attackers can hide an attack in two fundamentals ways. First, they can change the way the attack is delivered, for example, by splitting the attack into many network packets. Second, they can alter the attack payload so that it no longer matches the NIDS signature, for example, by using a different encoding for URL, Eliminating false positives. And eliminating false negatives, Understanding what constitutes a security relevant event and how to report it, Testing of Intrusion Detection Systems, Losses during attacks and recovering from the attack, Scalability of Intrusion Detection System.

III. PROPOSED WORK

Proposed Work: Proposed hybrid IDS will have following component.

- Network Analyzer and Network admin
- Host Analyzer and Host admin
- Network database and Host database
- Anomaly detector and misuse detector
- Mining algorithms
- Rule matching
- Alarm Generator

Analyzer: it will relate with the network operating system, usually to discuss Windows families or UNIX/Linux families

Network Analyzer and Network Admin: : Network analyzer will analysis packet which are traveling over network and collect the network packets. Network administrator will pass collected packets as passed from Network analyzer . Network analyzer works to analyze the packet information monitor the information and capture the information from the packet. Find the similar kind of packets and passed the packet to n/w administrator.

Network admin works to extract information from packets which passed from network analyzer and collect the information in to network DB. It's main work to extract the information from packets.

Host Analyzer and Host Admin: host analyzer works with security log .Basically there are three types of log –security log, event log, and application log .Its works analysis of security files

Periodically system log filtering .Host analyzers pass the packet to host administrator.

Host Administrator extracts information from packets and creates a training data set. By the training data set packets are compared and find out the abnormal packets.

Network database and Host database: Basically network database and host database contain the data which comes from network admin & host admin. Its work to save & contain the hole information.

Anomaly detector and misuse detector: Anomaly detector works to detect the unknown type of data find out the anomaly in the data which comes from network admin.

misuse detector works with host administrator, works to detect the unknown type of data find out the anomaly in the data which comes from Host administrator.

Mining algorithms: Here we used two kind of mining algorithm Association Rule Mining &Clustering Mining.

An associational rule is an implication form like X□Y, Where X□I, Y□I, and X□Y=□. The support of rule X□Y in the transaction D is the ratio of the number of transactions Contained X and Y in a transaction set to the number of all Transactions, denoted by Support (X□Y), that is:

Support (X□Y) =|{T: X□Y□T, T□D}□□|D|

The confidence level of rule X□Y in the transaction D is the ratio of the number of transactions contained X and Y in a Transactions set to the number of transactions contained X, Denoted by Confidence (X□Y), that is:

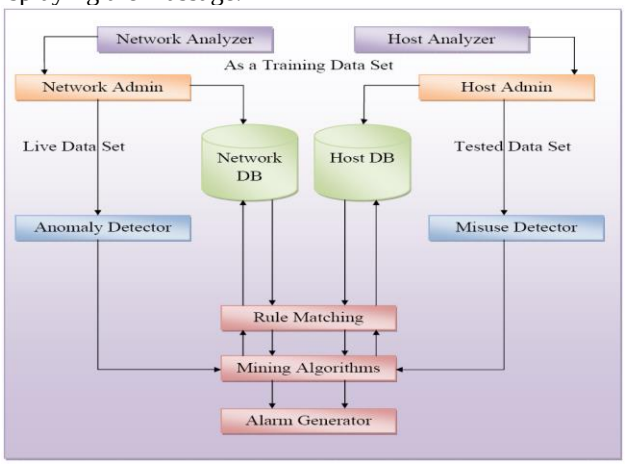
Confidence (X□Y)=|{T: X□Y□T, T□D}|□|{T: X□T, T□D}|

Given a transaction set D, the tasks of association analysis are to create the associational rules that support and confidence Level from mining data are respectively greater than the Minimum support (minsupp) and the minimum confidence (minconf) given by the users.

In Clustering Mining find the similar kind of pattern and make the clusters from similar kind of data over the network. Make the different type of clusters.

Rules Matching: It contains the different type of matching rules. Its work to find out the abnormal packets (which are not authorized) and passed to mining algorithm.

Alarm Generator: Give the alarm when the abnormal activities are perform by displaying the massage. When different type of intruders are find it warn to user by displaying the massage.



Analyzer: it will relate with the network operating system, usually to discuss Windows families or UNIX/Linux families.

IV. CONCLUSION

The proposed model of hybrid IDSs offers several advantages over alternative systems. First of all it will provide higher security, it will support high availability and scalability, and most important thing it will produced good results in terms of normal and abnormal behaviors of arrived packet. The proposed model will include integration of individual components to produced batter results. It will support to a system/network administrator the privileges for finding the intrusions which is reliable, secure and fast. The proposed model of hybrid IDS will be implement short time and at a low cost. It will also provide a best user interface.

REFERENCES

[1] Duanyang Zhao, Qingxiang Xu, Zhilin Feng "Analysis and Design for Intrusion Detection System Based on Data Mining" 2010 Second IEEE International Workshop on Education Technology and Computer Science.

[2] Asmaa Shaker Ashoor and Prof. Sharad Gore " Importance of Intrusion Detection System (IDS)" International Journal of Scientific & Engineering Research, Volume 2, Issue 1, January-2011

[3] Firkhan Ali Bin Hamid Ali and Yee Yong Len "Development of Host Based Intrusion Detection System for Log Files" IEEE symposium on business, engineering and industrial application(ISBEIA) langkawi, malaysia 2011

[4] Anthony Chung "On Testing of Implementation Correctness of Protocol Based Intrusion Detection Systems " Ninth IEEE International Conference on Software Engineering Research, Management and Applications 2011.

[5] Chunyu Miao and Wei Chen "A Study of Intrusion Detection System Based on Data Mining" published in IEEE conferences 2010

[6] Lin Ying, ZHANG Yan and OU Yang-Jia "The Design and Implementation of Host-based Intrusion Detection System" Third IEEE International Symposium on Intelligent Information Technology and Security Informatics 2010.

Mr.Sandeep Kumar Singh born in 1987, Mau (U.P.), India. He awarded B.E in Computer Science and Engineering from the Rajiv Gandhi Technical University, Bhopal, India in 2005-2009; He has been M.Tech scholar in Computer Science and Engineering(Software Engineering), at the Rajiv Gandhi Technical University, Bhopal, India

Mr.Nishant Chaurasia is an Assistant Professor of Computer Science Department in Laxmi Narayan Institute and Technology, Gwalior, Madhya Pradesh, India. He obtained his M.Tech degree from RGPV University. His research interest is in the field of MANETs.

Miss.Pragya Sharma born in 1990, Gwalior, India. She awarded B.E in Computer Science and Engineering from the Rajiv Gandhi Technical University, Bhopal, India in 2007-2011; She has been M.Tech scholar in Computer Science and Engineering, at the Rajiv Gandhi Technical University, Bhopal, India.