ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

A Survey - Comparative Study on Intrusion Detection System

Vinutha H.P.[1] , Dr.Poornima B[2]

Assistant Professor, Dept. of CS&E, BIET, Davangere, India[1]

Professor& Head, Dept. of IS&E, BIET, Davangere, India[2]

Abstract: This paper presents a survey on intrusion detection system. As the tremendous growth and usage of internet is increased the number of intruders and hackers have increased. Therefore the security on the network becomes the major issue. To overcome this many intrusion detection approaches are used. Intrusion detection gives the confidentiality, integrity and security of a resource. Data mining is the one of technology applied to intrusion detection to detect the network attacks, to reduce the complexities and to get normal behavioral pattern.

Keywords: Intrusion Detection System, Data set, Data Mining Algorithms, Classification.

I INTRODUCTION

In a recent year, usage of internet is increased in all the fields. As the usage of internet is increasing in our daily life, the network security is becoming necessary in order to obtain security, integrity and confidentiality of a resource. Along with the firewalls, intrusion detection system (IDS) has become a main component of the security system. The role of IDS is to trap the hacker's presence on the network. As the large number of incidents is increasing in our daily life IDS's are used with improved techniques. IDS plays an important to secure the network and its main goal is to view the network activities automatically to indentify the malicious attacks. Intrusion detection system is becoming critical component to secure the network in today's world. By using data mining in IDS can improve the detection rate, managing the false alarm rate and reduce false positive rate. Intrusion Detection technology identifies and deals with the malicious network of computer and computer network resources. In order to detecting data target IDS has been classified as into two categories:

- Host-based intrusion detection systems
- Network-based intrusion detection system

Host based IDS's are designed to monitor, detect and response to activity and attacks on the given host. Network based IDS's capture network traffic for their intrusion detection operations.

II MODES OF INTRUSION DETECTION SYSTEM

IDS can operate in different modes. The modes of operation are to have a basis for analysis of network packets. These metrics can be used to deduce whether a particular network or a system has been compromised or not. In most of the cases, the information collected indicates that whether the further action needs to be taken or not. The important two modes are:

- Anomaly detection
- Misuse detection

A. Anomaly detection

Anomaly detection is a process of scanning for abnormal activity that is encountered in the network. It does not required prior knowledge of attacks and it can also detect new attacks in the network. Anomaly detection is an audit data collected at the time of normal operation. Sometimes anomaly detection is referred as behavior based detection because it associates with variation from user behavior. The main advantage of anomaly detection is it has ability to detect novel attack or unknown attacks based on audit data. Anomaly detection can be divided into static and dynamic anomaly detection. Static detector just tackles the software portion of the system.

B. Misuse Detection

Misuse Detection is another method of attack employed by IDSs. This system compares the activities with the pre generated signature. In this method, IDS inspects tries to detect abnormal behavior by analyzing the given traffic based on several rules and by comparing these rules the system can detects type of attacks. Some times its is called as Signature-based detection technique because alarms are generated based on particular attack signatures. The main advantages of misuse detection are it has ability to give accurate result and having lesser false alarms. The disadvantage of misuse detection is that it will detect only the known attacks.

III BACKGROUND

Network intrusion detection is a process of monitoring and analyzing the data and events occurring in the computer network in order to detect attacks, vulnerability and other security problems. Networks security problems can vary widely and can affect different security requirements including authentication, integrity, authorization, and availability. Intrusion Detection System plays an important role to keep our network secure. Data mining based IDS can efficiently improve variants detection rate, manage false alarm rate and decrease false dismissals. Fig.1 shows

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

the general framework of the IDS. Initially all the incoming packets over the network are captured. These collected data are sending for preprocessing to eliminate the noise; irrelevant and misused attributes are replaced. The preprocessed data are analyzed and classified according to their severity measures. If the record is ordinary, it does not require any more change or else it sends for report generation to raise alarms. Based on the severity the alarms are raised to handle the state.

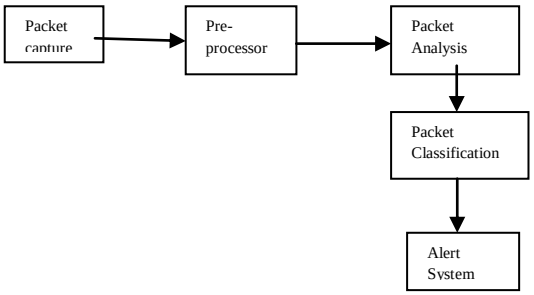

Fig 1. General Framework

IV DETECTION ISSUES

IDS can have both anomaly detection and misuse attacks, can be categorized based on the type of alarm raised by the IDS. The fallowing types of detection result are possible:

- True positive: Occurs when an actual attack occurs and the IDS responds to it by raising the appropriate alarm.
- True negative: When no attack happens, the intrusion detection system does not raise alarm.
- False positive: This occur when an IDS reads no attack. This is very serious drawback in intrusion detection systems. This is also known as false alarm.
- False negative: This occurs when the potential or genuine attack is missed by IDS.

V DATA MINING

Data mining techniques are used vastly in various fields. At the time of designing Network Intrusion Detection System, it is necessary to detect correct attacks within less time and raise the appropriate alarm. To do this data mining techniques are the most advantages and efficient techniques that can be used to design the intrusion detection system. Data mining based intrusion detection techniques generally fall into any of the two categories; anomaly detection and misuse detection. Generally data mining refers to the process of extracting the descriptive models from large storage of data. Use of data mining algorithms in IDS provides good performance and security. These systems are capable of detecting known and unknown attacks from the network. Different data mining techniques like summarization, clustering, classification can be used for analyzing and detecting the intrusion. Some of the beneficial steps of data mining are taken to solve the network intrusion detection problems like:

- Used to process larger amount of data
- It is more suitable to discover the hidden and ignored information
- It has a supervised learning methodology
- It can perform data summarization and visualization that can help the security analyst

VI DATA SET

Various data sets are available for the purpose of research. To test the effectiveness and feasibility of intrusion detection system, researches can make use of various available data sets. Some of the important available data sets are DARPA (Defense Advanced Research Project Agency), KDD Cup'99, NSL-KDD. The network intrusion detection system can also capture the packets in real time. The packet capturing can be achieved by using various available tools for packet capturing. These tools can capture the packets that are received from the network and convert them into readable form and then those packets are used for detection purpose. Most efficient packet capturing tools are available for capturing packets from network. Use of popular data set or data set captured from the network is needed to be pre-processed.

VII NETWORK ATTACKS

Attacks fall into fallowing four main categories;[1]

A. Denial of service(DoS) attacks
In this attack attackers disrupt a host or network services they make the some computing or memory resources too busy. For example: ping of death, SYN flood etc.

B. Probing Attack
In this attack the attackers scans the network to gather the information and then uses it to exploit the system. For example: port scanning etc.

C. Remote to Local Attack(R2L)
This occurs when an attacker who does not have an account on a remote machine sends packet to that machine over a network and exploit some vulnerability to gain local access. For example: password guessing etc.

D. User to Root Attack(U2R)
In this an attacker starts out with access to a normal user account on the system and is able to exploit vulnerability to gain root access to the system. For example: buffer overflow attack etc.

VIII RELATED WORK

Ayei ET. al. [1] has enhanced efficiency of intrusion detection by proposing a hybrid technique using both misuse and anomaly detection approaches. This is achieved by combining features of J48,Boyer Moore and K-NN algorithms. The HYBRITQ-4 performs well against four different attacks with high detection rate and low false positive rate. The experimental results have shown on different iteration.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

Ankita ET. al.[2] in this paper IDS is built using ensemble technique: Bagging and Boosting. They have implemented classification using SVM and Decision tree with both ensemble techniques. They have applied both techniques individually to the different classifier and results are compared.

Yogitha et.al. [3] Proposed intrusion detection system using Support Vector Machine (SVM). Verification is done by conducting experiments on NSL-KDD Cup'99 data set which is improver version of KDD Cup'99 data set. By using this NSL-KDD Cup'99 data set they have reduced extensive time required to build SVM model by performing proper pre-processing on data set. In this classification is done by using SVM. By doing proper kernel selection attack detection rate is increased and false positive rate (FPT) is decreased. In this proposed work author has used Gaussian Radial Basis Function.

Rowayda A. Sadek et.al. [4] proposed a new hybrid algorithm NNIV-RS(Neural Network with Indicator Variable using Rough Set for attribute reduction) algorithm is used to reduce the amount of computer resources like memory and CPU time required to detect the attack. In this approach feature reduction is done by using Rough Set Theory. Indicator Variable is used to represent the data set in more efficient way. Network packet classification has been achieved by Neural Network; neural network consist of a collection of preprocessing elements that are highly interconnected and transform a set of input to a set of desired outputs. With this hybrid approach they have achieved detection rate of 96.7% with false alarm rate of 3%.

Ahmed et.al.[5] in this paper intrusion detection system is papered with PSO-Discritize-HNB is used. This is the combination of Particle Swarm Optimization (PSO) and Information Entropy Minimization (IEM) descriptive method with the Hidden Naïve Bayes (HNB) classifier. Experiment is conducted on NSL-KDD data set. This proposed network IDS leads to high detection accuracy (98.2%) and speed up the time to 0.18sec after reducing the number of features from 41 to 11.

Hesham et.al.[6] has developed intrusion detection system using Bayesian probability because they wanted to improve the accuracy of the R2L attack. Used Bayesian method to classify the data accordingly. They have achieved better result than Chou's PhD result, where Chou has achieved a DR of 69.82% for the R2L. But the author has achieved better result for R2L attack with a DR of 85.35% by using features like Count, Srv count and Srv_diff_host_rate with a threshold value 0.6. But the CR considerably low then Chou because they have used a low threshold value which reduces the accuracy of detection of normal record but increases DR for R2L attack.

Renuka et. al. [7] has proposed an Artificial Neural network based NIDS by using a concept of ensemble binary classification and multi-boosting. Using these two concepts simultaneously it efficiently detects the attack with the low false alarm rate even at the high traffic. With the use of dynamic multi-boosting and database storage the time taken to detect the attack has been decreased efficiently.

Naveen N C et.al. [8] Has analyzed that designing the IDS for real time has become more challenging. Whenever a new thread is occurred a new knowledge map has to be built, to achieve this they have used SLFN (Single-Hidden Layer Feed Forward Neural Network). SLFN can detect attack faster compared to other methods. As a learning technique, SLFN demonstrated good potential in resolving Regression and Classification problems. Finally they have concluded that IDS using soft computing technique may prevent time consuming trials of other algorithm.

Tao Peng et. al.[9] have considered DARPA 2000 data set for Intrusion Detection Scenario to train and test the NIDS. To achieve this it has been implemented with the architecture of the data mining-based network intrusion detection system in real time. This framework is a distributed architecture consists of sensor, data pre-processor, extractors of features and detectors. To improve the efficiency they have adopted a novel FP-tree structure and FP-growth mining methods based on the FP-tree without candidate generation. Apriori candidate generation algorithm has been integrated into FP-growth method.

FP-growth adopts a divide-and-conquer strategy that compresses the database representing frequent item into a frequent-pattern tree, and proceeds mining of the FP-tree. The method is highly compressed and frequent item sets generation is integrated so repeated scanning of the item sets is not necessary. As they have adopted FP-growth for feature extraction the resource consuming and efficiency are satisfied.

Wenke Lee ET. al. [10] has first tried to mine the system audit data to study consistent use full patter of program and user behavior. They have also used the set of relevant system features presented in the patterns to compute inductively learned classifiers that can recognize anomalies and known intrusions. In order to make the classifier an effective model they should have a sufficient audit data for training and a set of predictive system features. To guide the audit data and feature selection they have proposed the association rule and frequent episodes from the audit data, which is used in classification model. They have incorporated domain knowledge into these basic algorithms using the axis and reference attributes.

Eleazar, Matthew et. al. [11] presented an adaptive model generation, a method for automatically building detection model for data mining based intrusion detection system. The data collected by intrusion detection sensor models are used to achieve this. The detection models are updated by the systems automatically as more data is collected. Adaptive model generation may significantly reduce the cost of deploying an IDS system because it removes the need for manually creating training set.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

TABLE 1 COMPARITIVE STUDY

Author(s)	Year	Paper name	Technique	Result
Ayei E. Ibor	2015	A Hybrid Mitigation Technique for Malicious Network Traffic based on Active Response	HYBRITQ-4(J48, Boyer Moore, K-NN)	High detection rate and low false positive rate.
Annkita Patel,Risha Tiwari	2014	Bagging Ensemble Technique for intrusion Detection System	Bagging & Boosting: SVM & Decision Tree	Combination of two algorithms is better than other ensemble technique.
Yogita B. Bhavsar & Kalyani C. Waghmare	2013	Intrusion Detection System Using Data Mining Technique Support Vector Machine	Data Mining , SVM,	Detection rate is increased & False positive rate is decreased.
Rowayda A. Sadek, M. Sami Soliman & Hagar S. Elsayed	2013	Effective Anomaly Intrusion Detection System based on Neural Network with Indicator Variable and Rough set Reduction	NNIV-RS	High detection rate and low false positive rate
Ahemd A Elngar, Dowalt, Fayed	2013	A Real Time Anomaly Intrusion Detection System with High Accuracy	PSO-Discritize-HNB	High detection accuracy and speed up the time
Hesham Altwaijry, Saeed Algarny	2012	Bayesian based Intrusion Detection System	Bayesian Probability	Achieved better detection rate with low threshold value
Renuka Devi Thanasekaran	2011	A Robust & Efficient Real Time Network Intrusion Detection System Using Artificial Neural Network in Data Mining	ANN: Binary Classification and Multi boosting	Time taken to detect the attack is decreased
Naveen N. C., R. Srinivasn, S. Natarajn	2010	A Unified Approach for Real Time Intrusion Detection using Intelligent Data Mining Techniques	SLFN	Faster attack detection compared to others
Tao Peng, Wanli Zuo	2006	Data mining Intrusion Detection System in real time	Data Mining: FP Tree & FP Growth	Resource consuming for feature extraction and efficiency are satisfied.
Wanke Lee, Salvotore J Stolfa	2000	Adaptive Intrusion Detection: A Data Mining Approach	Association Rule & Frequent Episode	Made classifier as an effective model

Daniel et. al.[12] In this paper the author has proposed a technique called pseudo-Bayes estimator to enhance an anomaly detection system's ability to detect new attack while reducing the false alarm rate as much as possible. They have worked on audit data analysis and mining (ADAM). ADMA is an anomaly detection system. It is composed of three modules: a preprocessing engine, a mining engine and a classification engine.

ADAM applies mining association rules technique to observe the abnormal events in network traffic data, and then it uses a classification algorithm to classify the abnormal events into the normal and abnormal instances. But the normal instance and attacks that the classifier is able to detect is limited, to overcome this pseudo-Bayes estimator method is used.

ADAM works in two phases one is training phase which is an offline phase were data stream is used to locate the attacks and another phase is detection phase. Attack free part of the streams is fed into a module.

IX CONCLUSION

This paper shows the implementation performed based on various data mining algorithms. There are several intrusion detection tools are available with competing feature to detect four different type of attacks. After the overall survey we can conclude that using single algorithm does not gives accurate result. Therefore to increase the detection rate and to reduce false alarm combination of algorithm gives more accuracy.

REFERENCES

[1] Ayei E. Ibor , Gregory Epiphaniou "A Hybrid Mitigation Technique for Malicious Network Traffic based on Active Response" 2015 International Journal of Security and its Application vol 9, No. 4(2015), pp 63-80.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 7, July 2015

[2] Sivaranjani S, Mr. Ravi Pathak, Vaidehi.V "Network Intrusion Detection using Data Mining Technique" 2014 International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)Volume 3 Issue 6, June 2014
[3] Yogita B. Bhavasar, Kalyani C. Waghmare "Intrusion Detection System Using Data Mining Technique: Support Vector Machine" 2013 International Journal of Emerging Technology and Advance Engineering volume 3, Issue 3, March 2013.
[4] Rowayda A. Sadek, M. Sami Soliman and Hagar S. Elsayed "Effective Anomaly Intrusion Detection System based on Neural Network with Indicator Variable and Rough set Reduction" IJCI international Journal of Computer Science Issue, Vol,10, Issue 6, No 2, 2013
[5] Ahmed A. Elngar, Dowlat A. El A. Mohamed and Fayed F. M. Ghaleb "A Real-Time Anomaly Network Intrusion Detection System with High Accuracy" 2013 Inf. Sci. Lett. 2, No. 2, 49-56 (2013)
[6] Hesham Altwaijry , Saeed Algarny "Bayesian based intrusion detection system" 2012 Journal of King Saud University 2012
[7] Renuka Devi Thanasekarn "A Robust and Efficient Real Time Network Intrusion Detection System Using Artificial Neural Network In Data Mining" 2011 International Journal of Information Technology Convergence and Services(IJITCS) Vol. 1, No. 4, 2011
[8] Naveen N C, Dr. R Srinivasan , Dr. S Natarajan "A Unified Approach for Real Time Intrusion Detection Using Intelligent Data Mining Techniques" 2011, IJCA Special Issue 2011
[9] Tao Peng, Wanli Zuo "Data Mining for Network Intrusion Detection System in Real Time" IJCSNS International Journal of Computer Science and Network Security, VOL.6 No.2B, February 2006
[10] Wenke Lee ,Salvatore J. Stolfo "Adaptive Intrusion Detection: a Data Mining Approach" 2000
[11] Huy Anh Nguyen, Deokjai Choi "Application of Data Mining to Network Intrusion Detection: Classifier Selection Model"
[12] S. A. Joshi, Varsha S. Pimprale "Network Intrusion Detection System based on data Mining" International journal of Engineering Science and Innovative Technology(IJESIT) Volume 2, Issue 1, 2013
[13] Prakash S.P., Madan B.S., Tugnayat R.M. "Approach for Intrusion Detection System Using Data Mining" Journal of Data Mining and Knowledge Discovery, 2012
[14] Ahmed youssef and Ahmed Emam "Network Intrusion Detection Using Data Mining and Network Behavior Analysis" IJCST , Vol 3,No 6, 2011
[15] Mathew G Schultz and Eleazar Eskin "Data Mining Methods for Detection of New Malicious Executables"
[16] Dr. S Vijayarani, Ms. Maria Sylviaa.S, "Intrusion Detection System-A Study" international Journal of Security, Privacy and Trust Management(IJSPTM) Vol 4, No 1, 2015
[17] Wenke Lee, Salvatore J. Stolfo,Kui W Mok "A Data Mining Framework for Building Intrusion Detection Models"
[18] Paul Dokas, Levent Ertoz, Vipin Kumar, Aleksandra Lazarevic, Jaideep Srivastava,Pang-Nig Tan "Data Mining for Network Intrusion Detection"