IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 3, May 2014.
www.ijiset.com
ISSN 2348 – 7968

A Survey on Intrusion Detection System with Data Mining Techniques

Ms. Ruth D [1], Mrs. Lovelin Ponn Felciah M [2]

[1]*M.Phil Scholar, Department of Computer Science, Bishop Heber College (Autonomous), Trichirappalli, Tamilnadu, India*

[2]*Asst.Professor, Department of Computer Science, Bishop Heber College (Autonomous), Trichirappalli, Tamilnadu, India*

Abstract

In recent years to maintain the network security is a main problem in the world. The system or network is hacked by the unauthorized users. There are many methods to increase the security such as encryption, firewall. But these methods are failed to detect the intrusions. For that a new technology can be used is called Intrusion detection system (IDS). The intrusion detection is the process to monitor the network or system and identify the intrusions. The IDS used data mining techniques for the network security, because to protect the network from various attacks and malicious traffic that originates from the internet. Data mining is used to extract the large amount of data from the database and also it is applied in many fields like Biological, Banking, Medical, Management, etc. This survey paper describes the Data mining approaches which are used to the detect intrusion in a network.

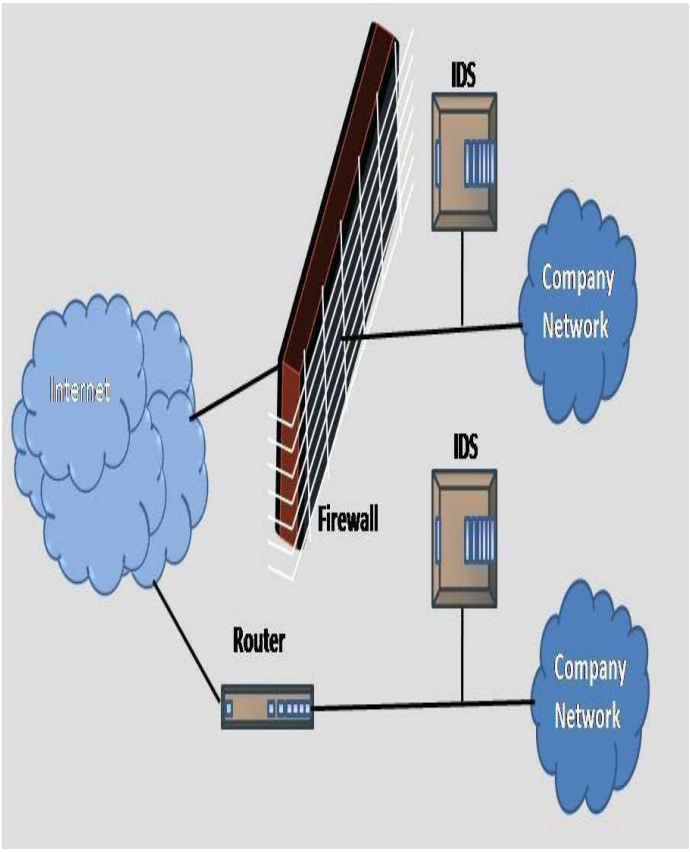
Key words: *Anomaly detection, Data mining, Intrusion detection system, Misuse detection.*

1. Introduction

Data mining is the process of discovering interesting patterns (or knowledge) from large amounts of data. The data sources can include databases, data warehouses, the Web, any other information repositories or data that are streamed into the system [5].

Data mining is also called KDD (Knowledge Discovery in Databases). The goal of data mining is process is used to extract information from the dataset and it is changed into an understandable structure.

Fig. 1 Typical locations for an intrusion detection system



The *intrusion detection* is software that automates the intrusion detection process [3]. The intrusion has many types namely viruses, worms, Trojan horse, etc. The normal detection system like firewall, virtual private network are failed to detect some critical intrusions from the network. The IDS finding the threats give response to the network administrator or user of the system and also raises alarms or signals when the security violations are occurred. In figure 1 describes how the intrusion detection can be processed. The information can be retrieved from the internet, then it is checked by the firewall and finally it protected by the IDS after that it send the information to the corresponding network.

2. Methodology

The Intrusion detection systems are protecting the network from unknown attacks. The methods are used in here are C4.5, support vector machine (SVM), genetic algorithm, Apriori association rules algorithm, clustering and classification.

Table. 1 Algorithms and data mining techniques used in intrusion detection systems

Intrusion detection algorithms	**Algorithms tasks**
C4.5	This algorithm is a supervised learning. The attribute values can be mapped and it is applied to classify new unseen instances. This algorithm is suitable for continuous and discrete value attributes.
Support Vector Machine (SVM)	This algorithm is related with supervised learning and used for classification and Prediction.
Genetic algorithm	This algorithm set the rules et randomly and used for finding optimal solutions to a specific problem.
Apriori Association Rules Algorithm	The association rule is unsupervised learning. Apriori algorithm is a frequent item sets for Boolean association rule. It is a level wise search
Clustering	This method is collecting the data sets based on similarity or distance between them.

2.1. Detection Techniques

The components of intrusion detection system are: data source, analysis engine and response manager. The data source contains two categories namely host based and network based. The analysis engine gets the information from the data source and analyzes the attacks. The response manager detects the attacks and gives response to the corresponding users. The Intrusion detection has two categories for detecting attacks in the network or host. They are:

a) Anomaly or Statistical detection

b) Misuse or Signature detection

Most of the researchers use these two techniques for detection rate and false alarm rate.

Anomaly Detection

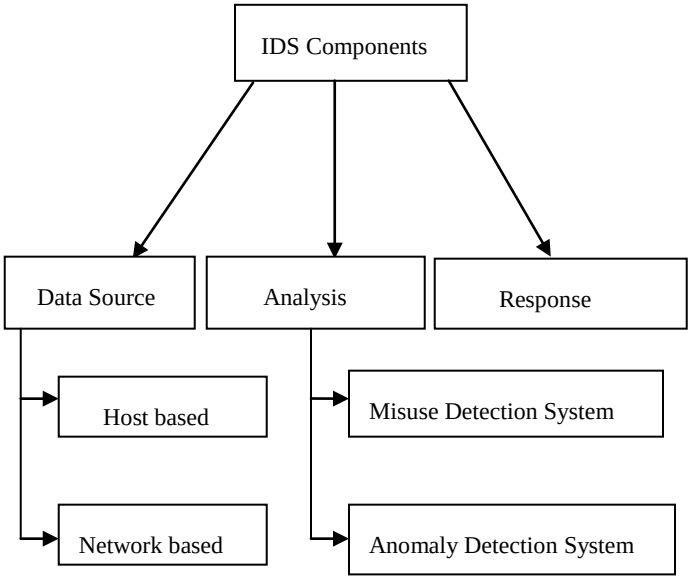
This detection technique is comparing user's current behavior with usual behavior which is already stored in the database. It used to detect the unknown attacks. By using statistical techniques to find patterns of activity that appears to

be abnormal [3]. It is failed in high detection rate. But the goal of anomaly detection system is to find the intrusion in the system timely.

Misuse Detection

This type of detection technique system use patterns of well-known attacks of the system to match and identify known intrusions [3].The misuse detection system is ability to detect only the known attacks which already stored in the database and generate the fewer false alarm rates. The disadvantage of this system is unable to detect the newly invented attacks.

Fig. 2 Components of Intrusion Detection System



3. A Review of Literature

This section discusses about various detection algorithms for network security.

[1] *Network Intrusion Detection System based on Data Mining – S.A. Joshi, et. al.,*

In this paper the author discuss about the data mining algorithms and Intrusion detection system to detect the unknown attacks from the dataset. There different kinds of attacks but the authors of this paper discuss the few kinds of attacks. They compares the four types of attacks are:

a) Probing attack
b) Denial of service
c) User to root
d) Remote to local

Then the author listed out the various data mining techniques and intrusion detection techniques which is used for the detecting the attacks like signature based detection, anomaly based detection, network- based intrusion detection system, host-based detection system. Comparing these types of attacks and finding the high detection rates.

pattern capturing algorithm has high detection rate. Finally find out the percentage for detection rate and false alarm rate.

[2] *Anomaly Detection in Network using Data mining Techniques – Sushil Kumar Chaturvedi, et. al.,*

The main work of this compares the two types of algorithms C4.5 and Support Vector Machine (SVM). First the given dataset is pre-processing and then the data can be partition into training and testing. The third stage the dataset is applied in C4.5 and SVM algorithm. The author of this paper compares these two algorithms and find out the detection rate comparison and false alarm rate comparison. By using these two data mining techniques they justify the C4.5 algorithm is better than the SVM.

[3] *Application of Genetic Algorithm in Intrusion Detection System – Omprakash Chandrakar, et. al.,*

This paper describes about basic concepts of network intrusion detection system, components and types of attacks. The IDS contains the three types of components namely data source, analysis engine, response manager. This paper gives the overview of genetic algorithm. The genetic algorithm randomly selected the input (chromosome) and calculates the fitness value for each generated initial chromosome. The iteration has performed some specific operations namely sorting, selection, crossover, mutation and finally calculates the fitness value for chromosome.

[4] *Anomaly Detection System by Mining Frequent Pattern using Data Mining Algorithm from Network Flow – A.R. Jakhale, et. al.,*

This paper describes an anomaly detection system and its two phases namely training and testing. The sliding window and clustering is used to monitoring the network traffic by mining the frequent patterns using algorithms. The algorithms are so effective and used in real time monitoring. The frequent multi-pattern capturing algorithm has high detection rate. Finally find the percentage for detection rate and false alarm rate.

[5] *A Survey on Intrusion Detection using Data Mining Techniques - R. Venkatesan, et al.,*

This paper describes the overview of the intrusion detection system and its each technique. The authors discuss pros and cons of anomaly detection and misuse detection. By combining these two categories and data mining approaches, then include the Apriori association rule algorithm for calculating the confidence levels. Apriori algorithm employs an iterative approach known as a level wise search, where k-item sets are used to explore (k + l)-item sets [5].

[6] *A Review of Intrusion Detection System in Computer Networks - Abhilasha A Sayar, et.al.,*

In this paper the author discuss about the classification of Intrusion detection system, advantageous and disadvantageous and its types. In this the IDS uses the artificial intelligence, fuzzy logic and neural network. The techniques are used to detect the intrusions in the images. For example, in military the original information's are changed into images and then send to another location. By using the artificial intelligence with IDS the user can easily identify the unknown attacks. This paper is useful for beginners to study the basic concepts of Intrusion detection system and also detect all kind of images.

4. Conclusion

This survey paper study on various techniques which are used to detect the attacks from unknown users. The intrusion detection system components are useful to know about the process of detection. The IDS is combined with the data mining techniques and algorithms detect the threats and give immediate response to the user, and also find the percentage of detection rate, false alarm rate, and confidence level.

References

[1] S.A.Joshi, Varsha S.Pimprale, "Network Intrusion Detection System (NIDS) based on Data Mining", International Journal of Engineering Science and Innovative Technology, Vol. 2, No. 1, January 2013, ISSN. 2319-5967.

[2] Sushil Kumar Chaturvedi, Prof. Vineet Richariya. Prof. Nirupama Tiwari, "Anomaly Detection in Network using Data mining Techniques", International Journal of Emerging Technology and Advanced Engineering, Vol. 2, No. 5, May 2012, ISSN. 2250-2459.

[3] Omprakash Chandrakar, Rekha Singh, Dr. Lal Bihari Barik, "Application of Genetic Algorithm in Intrusion Detection System", International Institute for Science, Technology and Education, Vol. 4, No. 1, 2014, ISSN. 2224-5774.

[4] A.R. Jakhale, G.A. Patil, "Anomaly Detection System by Mining Frequent Pattern using Data Mining Algorithm from Network Flow", International Journal of Engineering Research and Technology, Vol. 3, No.1, January 2014, ISSN. 2278-0181.

[5] R. Venkatesan, Dr. R. Ganesan, Dr. A. Arul Lawrence Selvakumar., "A Survey on Intrusion Detection using Data Mining Techniques", International Journal of Computers and Distributed Systems, Vol. 2, No. 1, December 2012, ISSN. 2278-5183.

[6] Abhilasha A Sayar, Sunil. N. Pawar, Vrushali Mane., "A Review of Intrusion Detection System in Computer Network", International Journal of Computer Science and Mobile Computing, Vol. 3, No. 2, February 2014, pp. 700 - 703.

ISSN 2348 – 7968

Authors

First Author- Ruth D, Pursuing M.Phil (CS), Bishop Heber College (Autonomous), Trichirappalli, Tamilnadu, India.

Second Author- Lovelin Ponn Felciah M, Assistant Professor, Bishop Heber College (Autonomous), Trichirappalli, Tamilnadu, India.