ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Issue 10, October 2016

Secrypt in IDPRS with Mining Techniques

[1]Thejeswari.C.K, [2]Sumathy.V
[1]Assistant Professor, [2]Assistant Professor (SG),
Department of Computer Science and Engineering,
Rajalakshmi Engineering College, Thandalam, Chennai – 602105, India

Abstract: The data packets sent across network are intruded by many attackers, prone through failures and packet losses. Those packets that are attacked by intruders affect the basic components of network security such as confidentiality, integrity and availability. Sensors in general uses alarms to indicate intrusion which brings the problem of inaccuracy and need for alarms increased in large complex network. To improve accuracy and performance, we proposed the novelty technique for data mining techniques like supervised learning and unsupervised learning. We make use of classification and clustering techniques to solve the problem of intrusion. Classification is one of the supervised learning techniques and clustering is one of the unsupervised learning techniques which are used to detect the type of attack. The intrusions are detected by using sensors in the network that compromises the basic three network security components, prevention technique like authentication and information protection across network systems are also used along with an alert message to the user. Finally, the detected intrusion is recovered by using the recovery techniques.

Keywords: Data Mining, Classification, Clustering, Intrusion, Authentication

I. INTRODUCTION

Network security consists of the provisions and policies adopted by a network administrator to prevent and monitor unauthorized access, misuse, modification, or denial of a computer network and network-accessible resources. Network security involves the authorization of access to data in a network, which is controlled by the network administrator. A secure network must provide the following three basic components of network security:

1. ***Data confidentiality***: Data that are being transferred through the network should be accessible only to those that have been properly authorized.
2. ***Data integrity*:** Data should maintain their integrity from the moment they are transmitted to the moment they are actually received. No corruption or data loss is accepted either from random events or malicious activity.
3. ***Data availability***: The network should be resilient to Denial of Service attacks.[1]

There are different types of attacks that occur in networks, viz,.

- *External break ins:* When an unauthorized user tries to gain access to a computer system.
- *Masquerader (internal) attacks*: When an authorized user makes an attempt to assume the identity of another user. This attacks are called also internal because they are caused by already authorized users.
- *Penetration attack*: In this attack a user attempts to directly violate the system's security policy.
- *Leakage:* Moving potentially sensitive data from the system.
- *Denial of Service*: Denying other users the use of system resources, by making these resources unavailable to other users.
- *Malicious use:* In this category fall miscellaneous attacks such as file deletion, viruses, resource hogging etc.[2]

Intrusion Detection (ID) is the process of monitoring events occurring in a system and signaling responsible parties when interesting (suspicious) activity occurs. Intrusion Detection Systems (IDSs) consist of 1) an agent that collects the information on the stream of monitored events, 2) an analysis engine that detect signs of intrusion, and 3) a response module that generates responses based on the outcome from the analysis engine.[2]

Intrusion detection can be defined as technology designed to observe computer activities for the purpose of finding security violations or we can say Intrusion detection is the process of identifying and responding to malicious activity targeted at computing and networking sources. Intrusion detection provides the following:

a) Monitoring and analysis of user and system activity.
b) Checking and comparing vulnerabilities.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Issue 10, October 2016

c) Availability of critical data files
d) Statistical analysis of activity patterns based on the matching to known attacks
e) Abnormal behavior analysis
f) Operating system analysis and comparison with stable state.[3]

This section is an overview of the four major categories of networking attacks using IDS's. Every attack on a network can comfortably be placed into one of these groupings

1. *Denial of Service (DoS):* A DoS attack is a type of attack in which the hacker makes a computing or memory resources too busy or too full to serve legitimate networking requests and hence denying users access to a machine e.g. apache, smurf, neptune, ping of death, back, mail bomb, UDP storm etc. are all DoS attacks.
2. *Remote to User Attacks (R2L):* A remote to user attack is an attack in which a user sends packets to a machine over the internet, which s/he does not have access to in order to expose the machines vulnerabilities and exploit privileges which a local user would have on the computer e.g. xlock, guest, xnsnoop, phf, sendmail dictionary etc.
3. *User to Root Attacks (U2R):* These attacks are exploitations in which the hacker starts off on the system with a normal user account and attempts to abuse vulnerabilities in the system in order to gain super user privileges e.g. perl, xterm.
4. *Probing***:** Probing is an attack in which the hacker scans a machine or a networking device in order to determine weaknesses or vulnerabilities that may later be exploited so as to compromise the system. This technique is commonly used in data mining e.g. saint, portsweep, mscan, nmap etc.[4]

Intrusions Detection can be classified into two main categories. They are as follow:

1. *Host Based Intrusion Detection***:** HIDSs evaluate information found on a single or multiple host systems, including contents of operating systems, system and application files .
2. *Network Based Intrusion Detection:* NIDSs evaluate information captured from network communications, analyzing the stream of packets which travel across the network.[4]

The techniques for the intrusion detection can be divided into two categories:

1. *Anomaly Intrusion Detection*: Anomaly detection technique assumes that all the intrusive activities are anomalous. Anomaly Detection Techniques includes Statistical, Neural Network, Immune System, file checking and Data Mining based approaches for the detection of attacks.
2. *Misuse Intrusion Detection:* Misuse Intrusion Detection uses the pattern of known attacks or weak spots of the system to match and identify the attacks. Misuse Detection Techniques includes genetic algorithm, expert system, pattern matching, state transition analysis and keystroke monitoring based approaches for the detection of attacks.[5]

Intrusion Prevention Systems are network security applications, that monitor and change network and system activities if found suspicious. The main functions of IPSs are to identify malicious activity, log information about it, and attempt to block or stop and report that activity. IPSs can be considered extensions of IDSs. The main differences that should be figured out between them are, that IDPs are placed in-line and are able to actively prevent or block.[3]

Intrusion Prevention Systems can be classified into four different types viz,.

1. *Network Based Intrusion Prevention System (NIPS):* In this kind of IDPS, it analysis the traffic of entire network by analyzing protocol activities and take appropriate actions.
2. *Wireless Intrusion Prevention System (WIPS):* In this kind of IDPS, it analysis the traffic of Wireless network by analyzing protocol activities and take appropriate actions.
3. *Network Behavior Analysis (NBA):* This type of IDPS examines traffic to identify threats that generate unusual traffic flow, such as DDOS attack, malware and Policy Violation.
4. *Host Based Intrusion Prevention (HIPS):* This type of IDPS monitors single host for suspicious activity by analyzing events occurring within that host. [3]

Data mining (DM), also called Knowledge-Discovery and Data Mining, is the process of automatically searching large volumes of data for patterns using association rules.[1] To accomplish the tasks of data mining, data miners use one or more of the following techniques:

1. *Data summarization:* summarizing data with statistics, including finding outliers

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Issue 10, October 2016

2. *Visualization*: presenting a graphical summary of the data
3. *Clustering:* Cluster the data into natural categories
4. *Association rule discovery*: defining normal activity and enabling the discovery of anomalies
5. *Classification*: predicting the category to which a particular record [6]

II. LITERATURE SURVEY

Intrusion detection starts with instrumentation of a computer network for data collection. Pattern-based software 'sensors' monitor the network traffic and raise 'alarms' when the traffic matches a saved pattern. Security analysts decide whether these alarms indicate an event serious enough to warrant a response.

A response might be to shut down a part of the network, to phone the internet service provider associated with suspicious traffic, or to simply make note of unusual traffic for future reference.[6]

An alarm shows that the system has been attacked. There were large number of alarms used when the organization have a large complex network.

This made the human analysts overwhelmed for the need to review when all alarm rings at once. Sometimes, alarm gave a false positive response [7] even when there was no intruders detected. This brought an inefficiency and inaccuracy in the intrusion detection.

In the current architecture for intrusion detection as shown in figure 1, Sensors analyze the network traffic and the set of events were stored in the sensor events database.

These events from database were further sent to the HOMER (Heuristic for Obvious Mapping Episode Recognition) which is a filter that filters out the sensor data before sending it to the classifier and clustering analysis.

Data mining (DM), also called Knowledge-Discovery and Data Mining, is one of the hot topic in the field of knowledge extraction from database [8].

The heuristic operates on aggregations by source IP, destination port, and protocol and then check to see if a certain threshold of destination IPs were hit within a time window. If the threshold is crossed, an incident is generated and logged to the database.

These Data mining techniques filter and detect the false alarms along with the anomalous behavior.

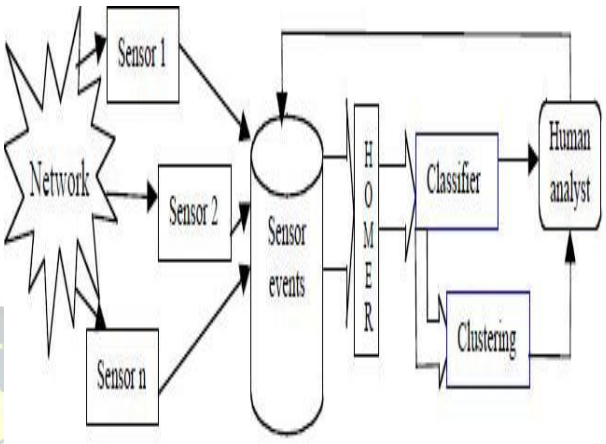

Fig.1: Intrusion Detection – Architecture Diagram

The Web server acts as a front end to the database and the human analysts can send Queries through the interface. So far, the intrusion was only detected and the response is indicated by an alarm and no method to recover the data that has been attacked. New attacks are not detected using the earliest technique of alarms. Hence new techniques for intrusion detection and Prevention with various algorithms are created manually.

III.PROPOSED WORK

In our paper, we will propose the following architecture for IDPRS (Intrusion Detection Prevention Recovery System). The act of intrusion can be detected, prevented and recovered by using the various mining techniques across networks. Here, we have used a new architecture for IDPRS which is given in the figure 2.

The data packets which are sent across the network are sensed by the means of various sensors, say,

sensor1, sensor2, sensor3, Sensor n

and if any intrusion is detected, it is recorded in the database named Intruded events. The database stores all types of intrusions in the data.

These events from the database are later sent to the HOMER (Heuristic for Obvious Mapping Episode Recognition) based on the type of the intrusion. This is achieved by using one of the mining techniques called the classification.

The classifiers take the input from the database and classify it based on the type of intrusion. If the classified

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com

International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Issue 10, October 2016

data assumes the data is intruded by someone, and then the attack is detected and based on the attack like anomaly or misused, they are clustered via a filter called HOMER.

The intruded data after detection is sent to recovery system for recovery. The data is recovered by applying complex algorithms and resend the data to the desired user. The prevention system is used to alert the user when the data is going to be attacked.

The alert of intrusion is given by considering various parameters like buffer size, data packet size, time delay, etc,. Based on the type of intrusion, the data is clustered and analyzed by the human analyst. The main approach in our paper is bringing the recovery system if the data intruded is detected.

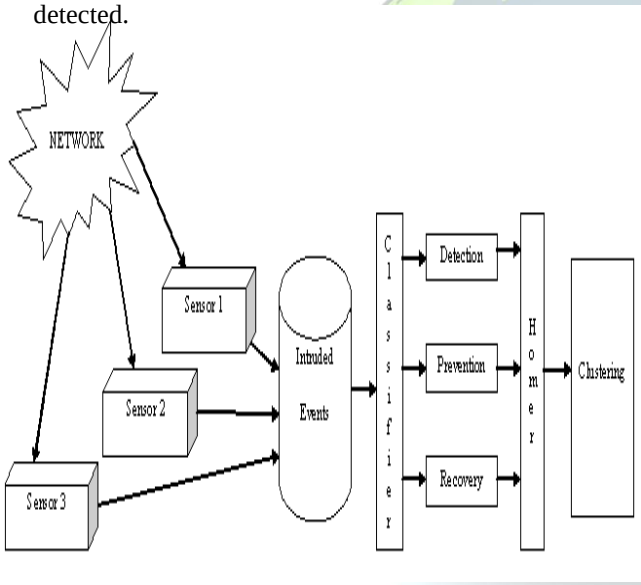

Fig. 2: IDPRS Architecture Diagram

The working system is clearly explained in the figure 3. The sensor senses the data packets across the network and stores the intruded events in the database. The intruded events are classified based on detection or prevention. If the intruded event is I the prevention stage, an alert message is given to the user.

Else if it is the detection stage, detect the type of attack and send it to the recovery system. Based on the type of recovery, they are grouped by the clustering technique.

The data is recovered by resending the data in a secured manner by the means of complex algorithms like DDDS. The algorithm is chosen based on the priority and the level of intrusion. Finally the secured data is sent to the intended user.

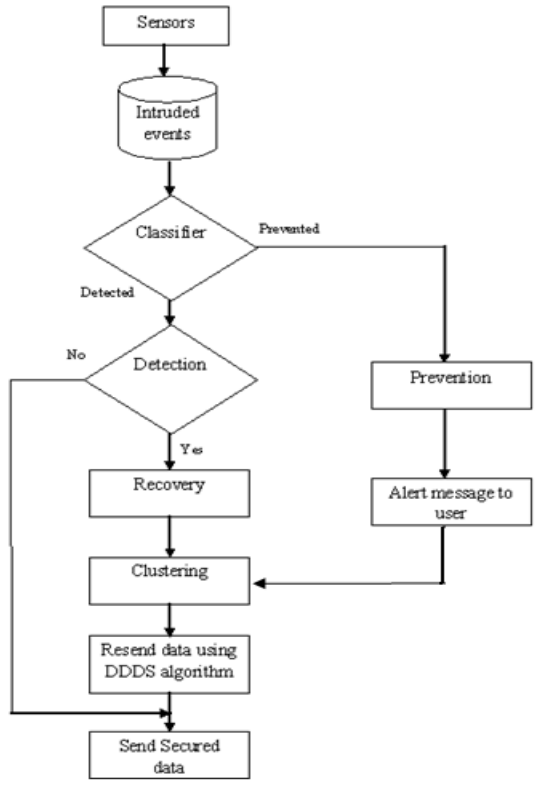

Fig. 3: Flow Chart - Steps in IDPRS

A. Intrusion Detection – Data Mining Techniques: Classification:

Classification consists of assigning a class label to a set of unclassified cases. The objective of classification is to analyse the input data and to develop an accurate description or model for each class using the features present in the data. This model is used to classify test data for which the class descriptions are not known. Decision tree is one of the classification techniques.

Decision tree learning is the construction of a decision tree from class-labeled training tuples. A decision tree is a flow-chart-like structure, where each internal (non-leaf) node denotes a test on an attribute, each branch represents the outcome of a test, and each leaf (or terminal) node holds

a class label. The topmost node in a tree is the root node. The decision tree consists of nodes that form a rooted tree, meaning it is a directed tree with a node called "root" that has no incoming edges. All other nodes have exactly one incoming edge. A node with outgoing edges is called an internal or test node. All other nodes are called leaves (also known as terminal or decision nodes). In a decision tree, each internal node splits the instance space into two or more sub-spaces according to a certain discrete function of the input attribute values.

C4.5 chooses the attribute of the data that splits its set of samples into subsets enriched in one class or the other. The splitting criterion is the normalized information gain. The attribute with the highest normalized information gain is chosen to make the decision.

Clustering:

Clustering is one of the unsupervised learning techniques and the most useful tasks in data mining process for discovering groups and identifying interesting distributions and patterns in the underlying data. Clustering problem is about partitioning a given data set into groups (clusters) such that the data points in a cluster are more similar to each other than points in different clusters. In the case of unsupervised learning, training data are not pre-defined. K-mode algorithm is one of the clustering techniques which is used to group the similar type of problems into a separate cluster.

IV. CONCLUSION

In this paper, we have given the approach of detecting the intrusion followed by the recovery of the data along with the prevention system by using the techniques of data mining like classification and clustering.

This is an overview of the security mechanism. We can further implement these different data mining techniques for the intrusion detection and recovery of the data by the updated complex algorithms.

REFERENCES

[1]. Theodoros Lappas , Konstantianos Pelecchrinis,"Data Mining Techniques for (Network) Intrusion Detection Systems"

[2]. Tyrone Grandison , Evimaria Terzi, "Intrusion Detection Technology", 2007

[3]. Bilal Maqbool Beigh, Prof.M.A.Peer, "Intrusion Detection and Prevention System: Classification and QuickReview", ARPN Journal of Science and Technology, VOL. 2, NO. 7,August 2012, ISSN 2225-7217

[4]. Mohammad Sazzadul Hoque, Md. Abdul Mukit and Md. Abu Naser Bikas, "An Implementation of Intrusion Detection System Using Genetic Algorithm", International Journal of Network Security & Its Applications (IJNSA), Vol.4, No.2, March 2012

[5]. Sandip Sonawane , Shailendra Pardeshi and Ganesh Prasad, "A survey on intrusion detection techniques" World Journal of Science and Technology 2012, 2(3):127-133, ISSN: 2231 – 2587

[6]. Eric Bloedorn, Alan D. Christiansen, William Hill,Clement Skorupka, Lisa M. Talbot, Jonathan Tivel," Data Mining for Network Intrusion Detection: How to Get Started"

[7]. Ashok Chalak Naresh D Harale Rohini Bhosale "Data Mining Techniques for Intrusion Detection and Prevention System" IJCSNS International Journal of Computer Science and Network Security, VOL.11 No.8, August 2011

[8]. Manikandan R, Oviya P and Hemalatha C, "A New Data Mining Based Network Intrusion Detection Model" Journal of Computer Applications ISSN: 0974 – 1925, Volume-5, Issue EICA2012-1, February 10, 2012

[9]. Meera Gandhi, S.K.Srivatsa, "Detecting and preventing attacks using network intrusion detection systems" International Journal of Computer Science and Security, Volume (2) : Issue (1) 49

[10]. SandipSonawane , ShailendraPardeshi and Ganesh Prasad, "A survey on intrusion detection techniques"World Journal of Science and Technology 2012, 2(3)pp:127-133, ISSN: 2231 – 2587

[11]. Manish Kumar, M. Hanumanthappa, T. V. Suresh Kumar, "Intrusion Detection System Using Decision Tree Algorithm", IEEE Conference, Nov. 2012, pp 629-634.

[12]. Sahilpreet Singh, Meenakshi Bansal, "A Survey on Intrusion Detection System in Data Mining", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 3, no. 3. Jun. 2013, pp 2190-2194.

[13]. Raj Kumar, Rajesh Verma, "Classification Algorithms for Data Mining: A Survey", International Journal of Innovations in Engineering and Technology (IJIET), vol. 1, no. 2. Aug. 2012.

[14]. Krunal Khurana, Priti.S.Sajja,Zankhana Bhatt, "Fuzzy Based Research Techniques for Intrusion detection and Analysis:A Survey" International Research Journal of Engineering and Tehnology (IRJET), vol.3 ,Issue: 5,May 2016

[15]. http://www.cs.ccsu.edu/~markov/ccsu_courses/DataMining-1.html

[16]. http://www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/issues.htm