IJARCCE

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

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

Intrusion detection system to provide security in a network against denial of service (dos) attacks

Amogh Mahesh[1], Nikhil M[2],SathyaPramod D S[3], Ajith L[4], Neelaja K[5]

Student, Computer Science & Engineering, NIE Institute of Technology, Mysore, India[1,2,3,4]

Assistant Professor, Dept. Of Computer Science & Engineering, NIE Institute of Technology, Mysore, India[5]

Abstract: In the era of internet, the network security has become the key foundation for many web applications. Intrusion detection is one which resolves this kind of a problem of internet security. Improperness of intrusion detection system (IDS) has given an opportunity for data mining technique to make several contributions to the field of intrusion detection. In the recent years, many researchers are using data mining technique for building IDS. Here, we propose a new approach by utilizing the data mining techniques such as neuro-fuzzy logic and support vector machine(SVM) helping IDS to attain high detection rate. The proposed technique has four major steps: primarily, k-means clustering is used to generate different training subsets. Then, based on obtained trained data subset, different neuro-fuzzy models are trained. Subsequently, a vector for SVM is formed and in the end, classification using radical SVM is to detect intrusion has happened or not. Experimental results show that out proposed approach do better than BPNN, multiclass SVM and other well-known major methods such as decision tree and Columbia model in terms of sensitivity, specificity and in particular detection accuracy.

Keywords: Intrusion detection system, Fuzzy-Neural networks, Support Vector Machine (SVM), K-means clustering.

I. INTRODUCTION

Internet services are indispensable and yet, vulnerable to Denial of Service (DoS) attacks, and especially to Distributed Denial of Service (DDoS) attacks. DDoS attacks, which many attacking agents cooperate to cause excessive load to a victim host, service, or network. DDoS attacks have increased in importance, number and strength over the years, becoming a major problem. Furthermore, significant growth in size of attacks and in their sophistication is reported.

To identify Bandwidth Distributed Denial of Service (BW-DDoS) attacks, which disrupt the operation of the network infrastructure by causing congestion or an excessive amount of traffic. BW-DDoS attacks can cause loss or severe degradation of connectivity, between the Internet and victim networks or even whole autonomous systems, possibly disconnecting whole regions of the Internet.

BW-DDoS attacks are usually generated from a large number of compromised computers (zombies or puppets). Bandwidth Distributed Denial of Service are the most frequently used DoS method. Most BW-DDoS attacks use few simple ideas, mainly, flooding, i.e., many agents sending packets at the maximal rate, and reflection, i.e., sending requests to a server with fake (spoofed) sender IP address, resulting in server sending (usually longer) packet to the victim. [1]

In the current detection approaches a number of IP trace back approaches have been suggested to identify attackers and there are two major methods for IP trace back, the probabilistic packet marking (PPM) and the deterministic packet marking (DPM). Both of these strategies require routers to inject marks into individual packets.

The DPM strategy requires all the Internet routers to be updated for packet marking. Moreover, the DPM mechanism poses an extraordinary challenge on storage for packet logging for routers. Further, both PPM and DPM are vulnerable to attacks, which are referred to as packet pollution.

PPM strategy can only operate in a local range of the Internet (ISP network), where the defender has the authority to manage. ISP networks are generally quite small, and cannot trace back to the attack sources located out of the ISP network.Because of the vulnerability of the original design of the Internet, we may not be able to find the actual attackers at present.

II. K-MEANS CLUSTERING

The k-means clustering algorithm is used to group unlabeled data [2]. In our proposed technique, we are intended to group our input data set into different clusters based on intrusion. Since our input data set consists of thenormal data and trained data, training data set is grouped into clusters using k-means clustering techniques. Examining and learning the behaviour and characteristics of the single data point within a cluster can give hints and clue on all other data points in the same cluster. This is because of the fact that all data points inside a cluster differ only by a small amount and usually follow a more or less similar structure. Hence, the data and then classifying is a simpler method and is less time consuming.

K means clustering is well known for solving clustering problem and it has the simplest procedure and easy way of

Copyright to IJARCCE

IJARCCE

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

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

classifying data. [3] It clusters the given set of data through certain number of k clusters fixed a priori. Main idea behind the k means clustering is to define k centroid and it should be placed in a cunning way because of different location causes different results. Next step is to associate to the nearest centroid to the particular given data set, when no point is pending first step is completed. Now we need to recalculate k new centroid from the previous result. After this k new centroid, binding has been between the nearest centroid and the same data set points, loop has been generated and it changes it location step by step and aim of algorithm is to minimise an objective function.

III. FUZZY-NEURAL NETWORKS

K-means clustering results in the formation of 'K' clusters where each cluster will be a type of intrusion or the normal data[4]. For every cluster, we have Neuro-fuzzy classifiers associated with it, i.e., there will be 5 number of Neuro fuzzy classifiers is trained with the data in the respective cluster. Neuro-fuzzy makes use of back propagation learning to find out the input membership function parameters and least mean square method to find out the consequent parameters.

Computers are not able to make their own decisions that are computers cannot think to provide thinking capability to computer or providing an artificial intelligence can be done by using neural networks.

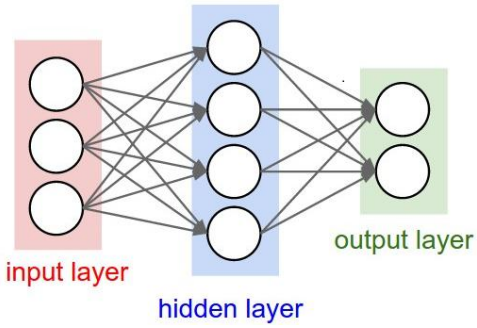

Fig 1 Neural networks layout

Neural network has three layers input, hidden and output as shown in Fig 1. Where input layer gets the clustered data, hidden layer contains the trained data and output layer gives the result by comparing the clustered data with the trained data. Fuzzy logic is used to train the data in the hidden layer of neural networks.

IV. SVM CLASSIFIERS

Classification of the data point considering all its attributed is a very difficult task and takes much time for the processing, hence decreasing the number of attributes related with each other of the data point is of paramount importance. The main purpose of the proposed technique is to decrease the number of attributes associated with each data, so that classification can be made in a simpler and easier way. Neuro-fuzzy classifier is employed to efficiently decrease the number of attributes. [5]

This classifier is used as it produces better results for binary classification when compared to the other classifiers. But use of linear SVM has the disadvantages of getting less accuracy result, over fitting results and robust to noise.

These short comings are effectively suppressed by the use of the radial SVM where nonlinear kernel functions are used and the resulting maximum margin hyper plane fits in a transformed feature space. In our proposed technique, nonlinear kernel functions are used and the resulting margin hyper plane fits in a transformed feature space. When the kernel used is a Gaussian radial basis function, the corresponding feature space is a Hilbert space of infinite dimensions.

V. PROPOSED WORK

BW-DDoS attack, where the attacker sends as many packets as possible directly to the victim, or from an attacker controlled machines called 'zombies' or 'bots'.The simplest scenario is one in which the attacker is sending multiple packets using a connectionless protocol such as UDP. In UDP flood attacks, the attacker commonly has a user-mode executable on the zombie machine which opens a standard UDP sockets and sends many UDP packets towards the victim. For UDP floods, and many other BW-DDoS attacks, the attacking agents must have zombies, i.e., hosts running adversary-controlled malware, allowing the malware to usethe standard TCP/IP sockets. The first attempts to avoid detection, and the second tries to exploit legitimate protocol behaviour and cause legitimate clients/server to excessively misuse their bandwidth against the attacked victim.

The four modules are

A. **Construction of normal Dataset:**The data obtained from the audit data sources mostly contains local routing information, data and control information from MAC and routing layers along with other traffic statistics. The training of data may entail modelling the allotment of a given set of training points or characteristic network traffic samples.

B. **Local Data Collection:**A normal profile is an aggregated rule set of multiple training data segments. New and updated detection rules across ad-hoc networks are obtained from normal profile. The normal profile consists of normal behaviour patterns that are computed using trace data from a training process where all activities are normal. During testing process, normal and abnormal activities are processed and any deviations from the normal profiles are recorded.

C. **Training normal data using cluster mechanism:**It calculates the number of points near each point in the feature space. In fixed width clustering technique, set of clusters are formed in which each cluster has fixed radius also known as cluster width in the feature space. K-means and fuzzy logic algorithms are used.

IJARCCE

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

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

D. **Testing Phase:** The testing phase takes place by comparing each newtraffic samples with the cluster set to determine the anonymity. The distance between a new traffic sample point and each cluster centroid is calculated. If the distance from the test point s to the centroid of its nearest cluster is less than cluster width parameter , then the traffic sample shares the label as either normal or anomalous of its nearest cluster. If the distance from s to the nearest cluster is greater than w, then s lies in less dense region of the feature space, and is labelled as anomalous. SVM classifier are used.[6]

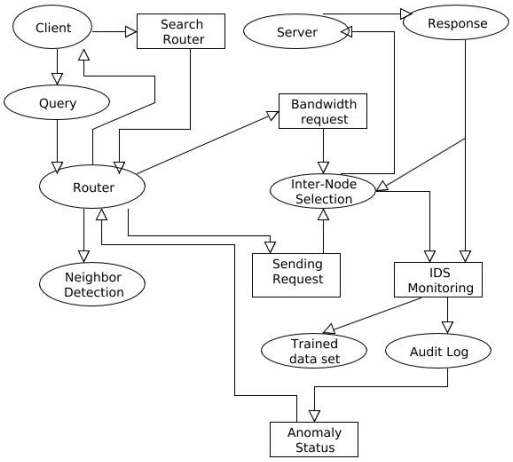

Fig 2 System Architecture

TheFig 2 consists of four main components: client, server, router and Intrusion Detection system (IDS). First, the client search for the router that are nearby as well as the few distance apart from the client, router sends its information back to the client. The client sends the query to the router, and then the router sends that information to the neighboring nodes, also requests for the bandwidth along with its bandwidth.

Server gets the query from the client through router. Then the server sends the query to the response phase. Then the response phase send those information to the IDS monitoring system/phase, where malicious activities are detected and alert message is sent to the user/administrator. IDS monitoring system/phase has two phases of data, trained dataset and audit log.

Trained dataset is a set of data used to discover potentially predictive relationships. A test set is a set of data used to assess the strength and utility of predictive relationships. The test and training sets are used in intelligent system, machine learning, genetic programming and statistics.

An Audit log is a security-relevant chronological record, set of records, and/or destination and source of records that provide documentary evidence of the sequence of activities that are affected at any time of specific operation, procedure, or event. Audit log contains anomaly status. Supervised anomaly detection techniques require a data set that has been labeled as "normal" and "abnormal" and involves training a classifier. These are sent back to the router, then to the client.

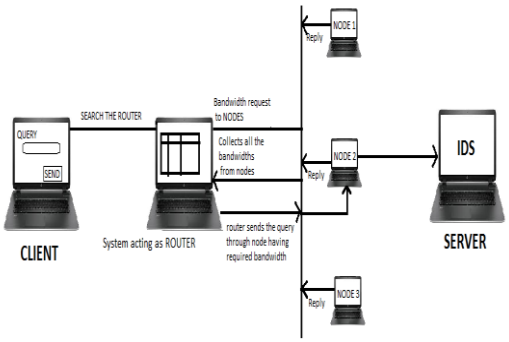

Fig 3 Layout

VI. MODULES USED

A. Client Form: It contains design part of the client, it calls J progress bar, finally it calls .vo package when send button is clicked.

B. .vo package (serialization): It has get and set methods to initializing data.

C. IDS desginTab Form: It is combination of IDS monitoring form and IDS graph form.

D. Node Form: It is combination of node tab form and routing table form.

E. Server Form: It contains server design.

F. Java Component Used:

- JLabel: Swing component used to create a text label.
- JTextField: This component creates small area to insert text.
- Jbutton: Used to create clickable buttons with labels on it.
- JTextArea: This creates big area to display the text along with scroll bars.
- JTabbedPane: Used to create tabs to display different pages.
- JRadioButton: Creates a menu where only one option can be selected.
- JTable: This swing component creates table along with the column headers.
- ActionListener: Used to receive the action events, to process that actionPerformed() is invoked.
- MouseListener: It receives mouse events(press, release, click, enter and exit), to track mouse moves MouseMotionListener() is invoked.
- java.net.socket: Socket is an end point for communication between two machines. This class implements client side sockets.
- java.net.serversocket: This class implements server side socket and waits for the request to process it, and finally returns the result.

IJARCCE

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

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

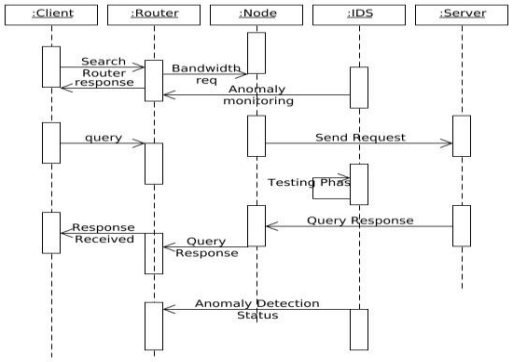

Fig 4 Sequence diagram

Fig 4 shows what interactions are possible within the functional components of the system.The client machine searches the router.The router then requests the node for Bandwidth.Router responds to the client.Anomaly monitoring is informed to router by the IDS.Clients sends query to the router.Node sends request to the server.Testing phase is done in IDS.Server sends the query response to the node, and then sends it to the client.Client receives the response.IDS send the anomaly status to the router.

VII. CASE STUDIES

Normal Scenario: In the normal scenario, all the nodes in the network will be in normal state. The router chooses the node which has highest bandwidth and sends the request through that node. Then the request is sent to the server and server replies back result of the requested query. Then we perform data aggregation, and we get to know that the node is not intruder. Attacked Mode: In the attacked scenario, one node in the network will be in attack state. The router chooses the node which has highest bandwidth and sends the request through that node. Then the request is sent to the server and server replies back result of the requested query. We are only attacking the node by sending more number of packets. We enter the number of times to send

CONCLUSION

So far, bw-ddos attacks employed relatively crude, inefficient, brute-force mechanisms. However, several known attacks, which aren't commonly used, let attackers launch sophisticated attacks, which are difficult to detect and might considerably amplify attackers' strength. Deployed and proposed defenses might struggle to meet these increasing threats; therefore, we need to deploy more advanced defenses.

This might involve proposed mechanisms as well as new approaches. Some proposed defenses raise operational and political issues; these are beyond the scope of our article but should be considered carefully. Finally, for a defense mechanism to be practical, it must be easy to deploy and require minor changes, if any, especially to the internet's core routers.

REFERENCES

[1]. S. Wei, J. Mirkovic, and M. Swany, "Distributed Worm Simulation with a Realistic Internet Model," Proc. Workshop Principles of Advanced and Distributed Simulation (PADS 05), IEEE CS, 2005, pp. 71–79.
[2]. S. kumar, "classification of K-means clustering in intrusion detection", Purdue University, 1995, pp. 3-6.
[3]. http://home.deib.polimi.it/matteucc/clustering/kmeans.html
[4]. Y. Dhanalakshmi and Dr. I. Ramesh Babu, "Intrusion Detection Using Data Mining Along Fuzzy Logic", International Journal Computer Science & Network Security, Vol. 8, No.2, pp.22-31,2008.
[5]. S. Kumar, E. Spafford, "Support Vector Machine Concept: A rule-based Intrusion Detection", IEEE Transaction on Computer Science, pp.4-8,2008.
[6]. T. Xia, G. Qu, S. Hariri," An Effective Network Intrusion Detection", 2013.