Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 2, Issue. 8, August 2013, pg.76 – 81

RESEARCH ARTICLE

ANOMALY INTRUSION DETECTION SYSTEM USING NEURAL NETWORK

Ms. Vrushali D Mane[1], Ms. Abhilasha Sayar[2], Prof. Sunil Pawar[3]

[1]Department of E&TC, India

[2]Department of E&TC, India

[3]Department of E&TC, India

Abstract— As human body remain secure from various environmental anomalies so as our computers also have security from the network intrusions. This paper proposes a new technique to intrusion detection system that can be effectively find and classify different afflictions. This system has to identify dynamic behaviour of systems. There are several techniques to help IDS to identify and get the changing behaviour of system. In this paper we are going to use artificial neural network ANN to get system changes. To get desired result we can use KDD'99 data set in our IDS system.

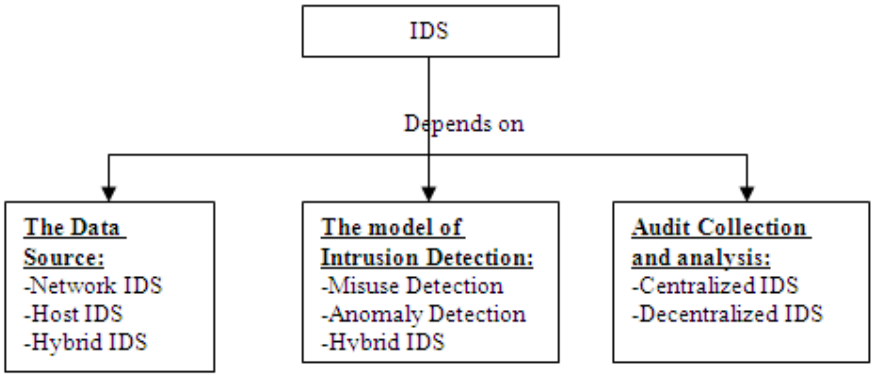
Key Terms: - Anomaly detection system; Neural Network; KDD; IDS

I. INTRODUCTION

Today internet is a very popular medium to collect information of any type of domain. Personal user, educational institute, business, corporate world depends on internet for different type of data. Anybody can access it and easily get desired output. With these entire advantages internet comes with big subject called security. This security can be of any type. Many companies faced lot of crisis due to intrusions and hackers. Many methods came in market to make system secure from such disastrous concepts like installing firewalls, encryption techniques, intrusion detection systems [1],[2].

In network and in systems connected in networks have some behavior with an external events. Intrusion detection system helps to find out such events in addition to the signatures of the intrusion. Its goal is to secure system for various network critical issues like confidentiality, integrity, and availability [3],[4]. Intrusion detection system (IDS) collects information from various parts of network and system, analyze it to find intrusion affected component. The IDS has a number of techniques to detect the danger.

IDS can be characterized by depending on three main aspects [5]:



1) The Data Source: - Network IDS examine networks (ex. Traffic), Host IDS examines system components (Ex: Operating system), Hybrid IDS supports both.

2) The model of intrusion detection: - Misuse detection verifies signature on data, anomaly detection verifies behavior of system, Hybrid IDS monitors both.

3) The audit collection and analysis: - Centralized IDS controlled by central resource. Decentralized IDS controlled from a local control node with hierarchical reporting to one or more central location(s).

In this paper we have developed anomaly based IDS to verify the behavior of the system by using ANN for attack sensing and categorization. As we know ANN must be trained and tested under some input to work properly for such types of input and complete desired objective. To give such input we have used the KDD'99 data set for training and testing our system.

The proposed system can work in three main styles: detection (for examining normal and abnormal actions), classification (if any abnormal action found categorize it in four main attack types: DOS, PROB, U2R, or R2L), and detailed classification (for detailed classification of abnormal events into 29 sub attack types).

II. LITERATURE SURVEY

1) INTRUSION DETECTION SYSTEM

Intrusion Detection system is now vital component of security system. IDS is different from other technologies by detecting and providing administration data of new attack which are unexpected to components where other only detect attacks. This can happen in IDS by endlessly observing and studying the events generated internally and externally in system or network. IDS perform three operation to find out intrusion, firstly it observe and study the network traffic and system, second it recognizing any abnormal activity, evaluating level of intensity and generating warning bell [7]. So the IDS is consider as a best protection system which observe traffic on network and study the state of the system for any external generated event. Using this phenomenon IDS recognize all abnormal state of system and make it secure from vulnerabilities. In addition to these IDS also stores the data records, which would be important proof when you file a case against a attacker. Hence IDS is essential elements which complete a total cycle of information security and become a logical component of firewall [5]. IDS can be used with different kinds of security elements to provide best security mechanism for system or network like home security. In home security various tools collaborate with each other like wired walls, sensors attached to doors, cameras to capture images etc [6]. So IDS and Home security both have same sequence of action observation, recognition, reporting. Some impressive IDS have the facility to report any danger to the system administrator who take care all security issues. The ability of the IDS system is not only to detect the certain behavior of the system but also it is self-tolerable to handle problems (with three wrong password lock user account). Many people endlessly working on IDS to improve its tools and performance in network and system. But still many users require lot of study of IDS, its tools, and how it works. Intrusion attacks can be internal to organization or external to organization, IDS must understand and differentiate these attacks [9].

The IDS however is not an answer to all your Security related problems. User must know what you can, and cannot expect of your IDS. In the following subsections shows what an Intrusion Detection Systems are capable of, but each network environment varies and each system needs to be tailored to meet your enterprise environment needs.

The IDS can:

1) Add a greater degree of integrity to the rest of you infrastructure
2) Trace user activity from point of entry to point of impact
3) Recognize and report alterations to data
4) Automate a task of monitoring the Internet searching for the latest attacks
5) Detect when your system is under attack
6) Detect errors in your system configuration
7) Guide system administrator in the vital step of establishing a policy for your computing assets
8) Make the security management of your system possible by non-expert staff

The IDS cannot:

1) Compensate for a weak identification and authentication mechanisms
2) Conduct investigations of attacks without human intervention
3) Compensate for weaknesses in network protocols
4) Compensate for problems in the quality or integrity of information the system provides
5) Analyze all the traffic on a busy network
6) Always deal with problems involving packet – level attacks
7) Deal with some of the modern network hardware and features

The IDS can be used at following positions in network and also it depends on the environment of users system

1) Between users internal network and external network
2) In the perimeter network before the firewall to detect the attacks on your servers in perimeter network
3) Between the firewall and user network, to detect an intrusion in case of the firewall penetration
4) In the remote access.
5) Between servers and user community, to identify the attacks from the inside
6) On the intranet, ftp, and database environment.

When the intrusion detection system get installed in users system it has members to be involved like Information security officer, Network administrators, Database administrators, Senior management, Operating system administrators, Data owners. The resources will not be utilized effectively without all these members. Vulnerably and risk assessment must be done prior to implementing IDS.

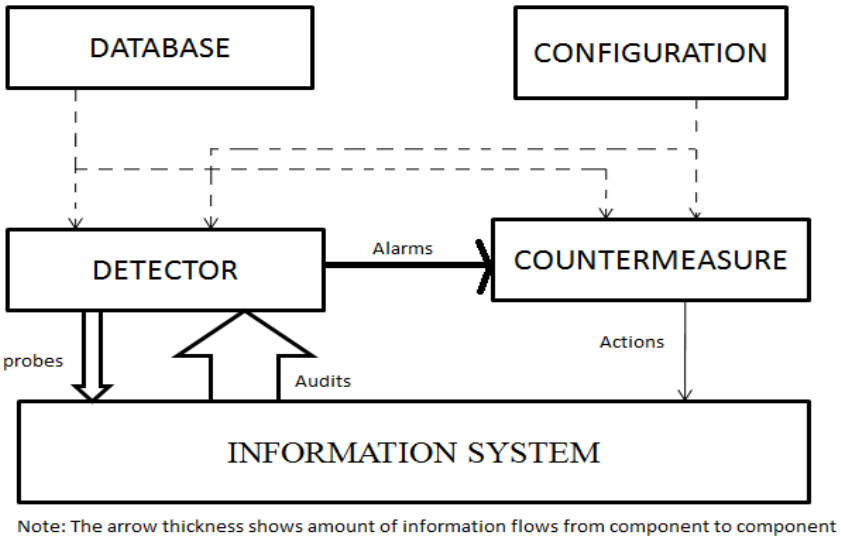


Figure: Simple Intrusion Detection System

The detector work is to remove unnecessary information from the audit trail. It then presents either a synthetic view of the security-related actions taken during normal usage of the system, or a synthetic view of the current security state of the system. A decision is then taken to evaluate the probability that these actions or this state can be considered as symptoms of an intrusion or vulnerabilities. A countermeasure component can then take corrective action to either prevent the actions from being executed or change the state of the system back to a secure state.

Properties of intrusion-detection systems

IDS have following properties []:

Accuracy: IDS has an accuracy which deals with the correct recognition of attacks and generation of false alarms. Inaccuracy occurs when an intrusion-detection system flags a legitimate action in the environment as anomalous or intrusive.

Performance: The performance of an IDS deals with rate at which observation events are processed. If the performance of the IDS is poor, then real-time detection is not possible.

Completeness: IF IDS detects all attacks then we can say it is complete. Incompleteness occurs when the IDS not able to detect an attack. This measure is much more difficult to evaluate than the others because it is impossible to have a global knowledge about attacks or abuses of privileges.

Fault tolerance: An intrusion-detection system should itself be resistant to attacks, especially denial-of service type attacks, and should be designed with this goal in mind. This is particularly important because most intrusion-detection systems run above commercially available operating systems or hardware, which are known to be vulnerable to attacks.

Timeliness: An intrusion-detection system has to perform and propagate its analysis as quickly as possible to enable the security operation to react before much damage has been done, and also to prevent the attacker from subverting the audit source or the intrusion-detection system itself. This implies more than the measure of performance because it not only encompasses the intrinsic processing speed of the intrusion-detection system, but also the time required to propagate the information and react to it.

2) ARTIFICIAL NEURAL NETWORKS (ANNs)

An Artificial Neural Network is a soft computing technique to process information that is almost same as that of biological nervous systems, like brain, which process information. We know that brain is composed of a complex interconnected information processing elements (neurons) working in collaboration to solve certain problems. Each information processing element (neuron) is fundamentally a summation element of an activation function. The Neurons may have multilayer architecture where input of each neuron (with weight parameter of the connection line) is given as the input to all of the neurons in the next layer. Each neural network must be trained for correct working in system for which parameter of connection coefficient (weighs) are found to solve the problem and has following basic steps [8]:

- Assign number of inputs to the Neural Network (vectors which creates a pattern)
- Check for specific input whether the actual output generated closely matches the desired output.
- Changing weights as per the output generated.

Some IDS developers changes ANN for a pattern recognition technique. To implement pattern recognition technique one can use a feed-forward neural network that has been trained accordingly. During training, the neural network parameters are modified to match outputs (each computer connection generates an output, normal and abnormal event) for input patterns (every input pattern is selected by a vector properties extracted from the symptoms of the network connection evidence). The neural network is used to classify the input pattern by identifying its class. The most usually found application of neural networks in IDSs is to train the neural network on a sequence of information units, each of which may be audit information or a sequence of commands. The input to the network consists of the large KDD data set. Once the network is trained on a set of representative data, it learns the profile of the attack and when put in action, it can detect the attack from its profile.

To get correct result network need to be trained. Generally, learning is a process in which the parameters like connection weights and bias levels of the ANN are modified through a continuing process of stimulation by the environment in which the network is involved. The parameter changes take place in network determines type of learning. In a general, the learning process may be classified as supervised or unsupervised [14], [15]. Using ANNs in implementing IDSs is an important to collect the dynamic information of system or network.

III. THE PROPOSED APPROACH

The proposed approach for ANN based intrusion detection system contains four main modules. These stages are examination, recognition, categorization, and reporting. All these modules are explained in following subsections.

A. Examination Module:

The Examination module refers to system administrator by giving an interface to him. It has various tools to capture actual tools to analyse and administer processes internal to the system like packet traffic, packet capturing. It also has some administrative tools that let the user to analyse profiles and the transfer of data. Examination module comprises understanding the system activity to find traffic, process, open, delete, create file, and different other operations. It gathers the valuable data and set them in profile to decide whether data is normal or abnormal.

The first activity performed by this module is to understand traffic by capturing packets. The host system connected in network can receive many packets from different workstations. All these packets are captured by examination module to understand traffic on network. We can create application programming interface in examining packet capturing.

B. Recognition Module:

The anomaly is recognized in the system by the recognition module by following five main steps. These steps are defining characteristics by understanding the source, dividing with header priorities, Encoding, Normalization and De-Normalization, and finally anomaly recognition phase.

1) Defining characteristics: This is the initial step for recognition. It starts after the examination module finish to understand traffic by packet capturing on the network. Each packet sent on network has source and the destination addresses specified. Also it has the source ports and the destination ports .The examination module capture packet either at the network layer (IP) and/or the transport layer

(TCP, UDP, ICMP).The characteristics of KDD'99 dataset determine values for each packet header field.

2) Diving: The very important operation of IDS is to select essential properties and dividing them. In examination module large number of properties and characteristics are used and all of them are not needed or useful. Some of them are difficult to get which contain drawback for IDS. By using divide principle all properties are categorized into preliminary, secondary, and less important divisions where preliminary properties are important properties, secondary properties are less important and less important properties has small amount of effect on recognition module. This classification is necessary to improve the system result. In this system if 22 dataset out of 41 KDD dataset are given input and divide them into 5 preliminary, 7 secondary, 10 less important properties user can get good result.

3) Encoding: the rule to change the form or convert the part of information into another form is coding. And bringing to its normal form is decoding. Encoding is done without any private or public key. The non-numeric data in table is converted to serial numbers. This process examines data to check whether it is normal or abnormal. If it is normal it assigns 1 digit to it and if it is abnormal it assigns 0 digits to it

4) Normalization and de-normalization: Normalization is the operation to sort out data easily in database. Normalization es0tablish relationship between columns of the tables in database and make them supple. In normalization repetitive data are removed, partial dependency and transitive dependency occurs. After division de-normalization occurs in which multiplication of the resulting values of the data and weights takes place. This provides a property of importance in the process of recognition so that one can divide properties into the three levels that is preliminary, secondary, less important.

5) Anomaly detection: In this phase actual intrusion is recognized by using ANN technique. Here ANN analyses activities of user and groups of users from pattern changes .It typically creates knowledge which contain profiles of all activities like examination.

C) Classification Module:

The classification operation can be performed by various techniques like ANNs, statistical methods, genetic algorithms, and others. In this system we are using Artificial Neural Network (ANN) to classify the attacks. The result is correct if and only if ANN is properly trained and tested. There are four types of attack for classification is probing, denial-of-service (DoS), remote-to-local (R2L), and user-to-root (U2R) attacks. Each type has sub type of various attacks.

D) Alert Module

This is the last operation of the proposed IDS. In this stage events are identified to check its status that is normal or abnormal. For any abnormal event IDS must give signal to system administrator to take corrective action.

The proposed IDS have been developed by developing all above modules using particular technology. The system requires two more programs need to be download and install. These are the Win cap and Sharp cap programs. WinPcap is an open source library application for packet capturing and network analysis. It can be used directly or indirectly. In our development we will use it indirectly. Sharp cap is an open source program.

Following diagram shows flow of operations in system.

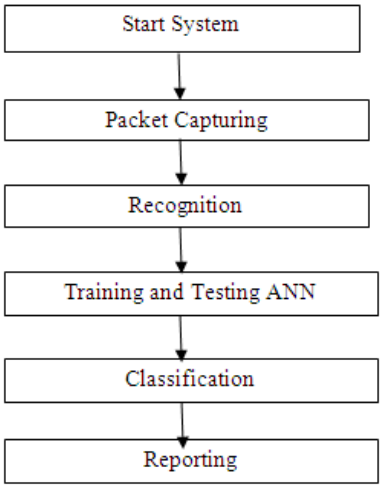


Figure 3: Flow of operations in system

IV. CONCLUSION

By training and testing Artificial neural network the performance of IDS improved at high level to recognise anomaly attacks in network or system. To make any future changes or extending system is easy because of modularity is maintained in system. However with lots of advantages ANN require lots of time and data to train it to give correct result. One more problem is that ANN can assure classification and recognition at various levels. So user requirements are necessary to define level of detail. For better result we can use 41 features input instead of 22 feature, test ANN performance without normalisation.

REFERENCES

[1] D. Herrmann, "A practical guide to security engineering and information assurance", 2002, www.auerbach-publications.com.
[2] S. Kiran, "Exploring a novel approach for providing software security using soft computing systems", International Journal of Security and Its Applications, Vol. 2, No. 2, pp. 51- 58, 2008.
[3] S. Alexander, "An anomaly intrusion detection system based on intelligent user recognition", Ph.D. Thesis, Faculty of Information Technology, University of Jyväskylä, Finland, 2002.
[4] S. Mansour and A. Sha'bani, "Fast neural intrusion detection System Based on Hidden Weight Optimization Algorithm and Feature Selection", World Applied Sciences Journal, No. 7 (Special Issue of Computer & IT), pp. 45-53, 2009.
[5] Hamdan.O.Alanazi, Rafidah Md Noor, B.B Zaidan, A.A Zaidan "Intrusion Detection System: Overview ",JOURNAL OF COMPUTING, VOLUME 2, ISSUE 2, FEBRUARY 2010, ISSN 2151-9617.
[6] Ondrej Linda, Todd Vollmer, Milos Manic, Member, IEEE "Neural Network Based Intrusion Detection System for Critical Infrastructures", Proceedings of International Joint Conference on Neural Networks, Atlanta, Georgia, USA, June 14-19, 2009.
[7] W. Jeffery, "Information Security Policy", California State University, Sacramento Information Security Office, July 2009.
[8] E. H. Spafford and D. Zamboni, "Intrusion detection using autonomous agent", Computer Networks, vol. 34, pp. 47-570, October 2000.
[9] V. Theuns and H. Ray, "Intrusion detection techniques and approaches",Journal of Computer Communications, Vol. 25, No. 15, pp. 1356 - 1365,2002.[10] W. Mahoney and W. Sousan," IDEA: A new intrusion detection data source", The 2nd International Conference on Information Security and Assurance, Korea, April 2008.
[10] L. Theodoros and P. Konstantinos, "Data mining techniques for (network) intrusion detection systems" Department of Computer Science and Engineering, UC Riverside, CA, USA, 2005.
[11] R. Ghosh, "A novel hybrid learning algorithm for artificial neural networks", Ph.D. Thesis, School of Information Technology, Griffith University, 2002.
[12] S. Jonas, "Neural network computations using Mathematics", A Tutorial by Wolfram Research, http://www.mvs.chalmers.se/~sjoberg/ September 2005.
[13] M. Hagan, Neural Network Design. Cengage-Nelson, Canada, 2008.
[14] V. Konstantinos, "Machine learning approaches to medical decision making ", PhD Thesis, Department of Computer Science, University of Bristol. March 2001 2215

