ABSTRACT

The incremental increase in the usage of technology has led to an increase in the amount of data that is being processed over the Internet significantly over the time period. With the huge amount of data that is being flown over the Internet, comes the scenario of providing security to the data, and this is where an Intrusion Detection System (IDS) comes into the picture and helps in detecting any virtual security threats. Intrusion Detection System (IDS) is a system that monitors and analyzes data to detect any intrusion in the system or network. Intruders find different ways to penetrate into a network. The IDS which is being proposed is being implemented using latest technologies such as Machine Learning Algorithms to classify the attacks and detecting them whenever an attack happens and also to find which machine learning algorithm is best suitable for identifying the attack.

Keywords:

Intrusion, Intrusion Detection System, Denial of service, User to Root attacks, Remote to User attacks, Local Area Network, Principal Component Analysis, Support Vector Machine, Random Forest, Decision Tree, KNN Algorithm, Logistic Regression, Alerts, False Positives, False Negatives.

LIST OF FIGURES

FIGURE NUMBER	FIGURE NAME	PAGE NUMBER
3.1.1.1	System Architecture	16
3.1.2.1.1	Data Visualization for Intrusion Detection Rate	18
3.1.2.2.1	Flowchart of Logistic Regression Algorithm	23
3.1.2.2.2	Flowchart of Decision tree Algorithm	25
3.1.2.2.3	Flowchart of Random Forest Algorithm	26
3.1.2.2.4	Flowchart of KNN algorithm	28
4.1.1.1	Use Case Diagram for the Proposed System	30
4.1.2.1	Class Diagram for the Proposed System	31
4.1.3.1	Sequence Diagram of the Proposed System	32
4.1.4.1	Activity Diagram for the Proposed System	34
4.1.5.1	State Chart Diagram for the Proposed System	36

4.1.6.1 Component Diagram for the Proposed System 38

4.1.7.1 Deployment Diagram for the Proposed System 39

5.3.1 Scatter Plot of the Data Set 62

5.3.2 Data Histograms of the Data Set 63

5.3.3 Output of the Program in Command Prompt 64

5.3.4 MSE Values 64

5.3.5 MAE Values 65

5.3.6 RMSE Values 65

5.3.7 Accuracy Rate of the Algorithm 66

5.3.8 Detection Rate of the Algorithm 66

5.3.9 Running Snort by Configuring Rules 67

5.3.10 Snort Checking for Alerts 67

5.4.1 Accuracy Rate Comparison of the Algorithms 68

LIST OF TABLES

TABLE NUMBER	TABLE NAME	PAGE NUMBER
3.1.2.1.1	Basic Features of individual TCP Connections	19
3.1.2.1.2	Content Features within a connection suggested by domain knowledge	20
3.1.2.1.3	Traffic features computed using a two second time window	21
5.4.1	Experimental Analysis Result of the Proposed System	68

TABLE OF CONTENTS

Abstract **i**

List of Figures **ii**

List of Tables **iv**

1. INTRODUCTION 1

1.1 Introduction 1

1.2 Motivation for the work 5

1.3 Problem Statement 5

2. LITERATURE SURVEY 7

2.1 Network Intrusion Detection System Based On Machine Learning Algorithm 7

2.2 Detecting web based Ddos attack using attack using mapreduce operations in cloud computing environment 8

2.3 A Pattern Recognition Scheme for Distributed Denial of Service (DDOS) Attacks in Wireless 9

2.4 Analyzing Log Files for Post-Mortem Intrusion Detection 9

2.5 Network Traffic Analysis and Intrusion Detection using Packet Analyzer 11

2.6 Existing System 12

3. METHODOLOGY 13

3.1 Proposed System 13

3.1.1 Architecture of the System 16

3.1.2 Implementation of the System 17

3.1.2.1 DataSet Details 17

3.1.2.2 Implementation of Machine Learning Algorithms 21

4. DESIGN 29

4.1 UML Diagrams 29

4.1.1 Use Case Diagram 30

4.1.2 Class Diagram 31

4.1.3 Sequence Diagram 31

4.1.4 Activity Diagram 33

4.1.5 State Chart Diagram 35

4.1.6 Component Diagram 36

4.1.7 Deployment Diagram 38

5. EXPERIMENTAL ANALYSIS AND RESULTS 40

5.1 System Configurations 40

5.1.1 Software Requirements 40

5.1.2 Hardware Requirements 46

5.2 Source Code 47

5.3 Screen Shots 62

5.4 Experimental Analysis/Testing 68

6. CONCLUSION AND FUTURE SCOPE 69

6.1 Conclusion 69

6.2 Future Scope 69

REFERENCES 70

1. INTRODUCTION

1.1 Introduction

An Intrusion Detection System (IDS) is a device or software application that monitors a network or systems for malicious activity or policy violations. Any intrusion activity or violation is typically reported either to an administrator or collected centrally using a security information and event management (SIEM) system. A SIEM system combines outputs from multiple sources and uses alarm filtering techniques to distinguish malicious activity from false alarms.

Although Intrusion Detection Systems monitor networks for potentially malicious activity, they are also disposed to false alarms. Hence, organizations need to fine-tune their IDS products when they first install them. It means properly setting up the intrusion detection systems to recognize what normal traffic on the network looks like as compared to malicious activity.

Classification of Intrusion Detection System

IDS are classified into 5 types

1. **Network Intrusion Detection System (NIDS)**

   Network intrusion detection systems (NIDS) are set up at a planned point within the network to examine traffic from all devices on the network. It performs an observation of passing traffic on the entire subnet and matches the traffic that is passed on the subnets to the collection of known attacks. Once an attack is identified or abnormal behavior is observed, the alert can be sent to the administrator. An example of an NIDS is installing it on the subnet where firewalls are located in order to see if someone is trying crack the firewall.

2. **Host Intrusion Detection System (HIDS)**

   Host intrusion detection systems (HIDS) run on independent hosts or devices on the network. A HIDS monitors the incoming and outgoing packets from the device only and will alert the administrator if suspicious or malicious activity is detected. It takes a snapshot of existing system files and compares it with the previous snapshot. If the analytical system files

- User to Root (U2R)

An attacker has local access to the victim machine and tries to gain super-user privilege. For example, buffer overflow attacks.

- Remote to Local (R2L)

An attacker tries to gain access to victim machine without having an account on it. For example, password guessing attack.

- Probe

An attacker tries to gain information about the target host. For example, port-scan and ping-sweep.

1.2 Motivation for the Work

Motivation for the work is to propose a security system, which detects malicious behaviors launched toward a system at SC level. The IDS uses data mining approaches namely Decision Tree, Random Forest, Logistic Regression and KNN is used to identify attack. The attack features are learned by the machine learning algorithm.

The contributions of proposed work are: 1) Identifying attack class by applying machine learning algorithm, 2) Identifying which algorithm is best suitable for IDS problem to effectively resist insider attack.

Intrusion detection system uses classification techniques to make decision about every packet pass through the network whether it is a normal packet or an attack. Our objective is to classify the attack into multiple attack types namely DOS, U2R, R2L, PROBE packet.

1.3 Problem Statement

Intrusion detection begins where the firewall ends. Preventing unauthorized entry is best, but not always possible. It is important that the system is reliable and accurate and secure. Intrusion detection is defined as real-time monitoring and analysis of network activity and data for potential Vulnerabilities and attacks in progress.

One major limitation of current intrusion detection system (IDS) technologies is the requirement to filter false alarms. IDS is defined as a system that tries to detect and alert of attempted intrusions into a system or a network. IDSs are classified into two major approaches. Intrusion detection is the process of monitoring the events occurring in a computer system or network and analyzing them for signs of possible incidents, which are violations or imminent threats of violation of computer security policies, acceptable use policies, or standard security practices. Intrusion prevention is the process of performing intrusion detection and attempting to stop detected possible incidents.

Intrusion Detection and Prevention Systems (IDPS) are primarily focused on identifying possible incidents, logging information about them, attempting to stop them, and reporting them to security administrators. In addition, organizations use IDPSs for other purposes, such as identifying problems with security policies, documenting existing threats and deterring individuals from violating security policies. IDPSs have become a necessary addition to the security infrastructure of nearly every organization. Thus there is necessary to propose a strong detection mechanism to identify attacks.

2.3 A Pattern Recognition Scheme for Distributed Denial of Service (DDoS) Attacks in Wireless

Sensor Networks, Baig, Zubair & Baqer, M & Khan, Asad. (2006), 1050 - 1054. 10.1109/ICPR.2006.147.

This paper defines distinct attack patterns depicting Distributed Denial of Service (DDoS) attacks against target nodes within wireless sensor networks for three most commonly used network topologies. A Graph Neuron (GN)-based, decentralized pattern recognition scheme is proposed for attack detection. The scheme does analysis of internal traffic flow of the network for DDoS attack patterns. We stipulate that the attack patterns depend on both the current energy levels, as well as the energy consumption rates of individual target nodes. The results of varying pattern update rates on the pattern recognition accuracies for the three network topologies are included in the end to test the effectiveness of our implementation.

Techniques used

The Graph Neuron (GN) is an in-network, distributed, pattern recognition algorithm which can form an associative memory overlay on the physical sensor network by interconnecting sensor nodes in a graph-like structure called the GN array. The Graph Neuron (GN) as a light-weight pattern recognition application was used to detect DDoS attack patterns inWSNs.

2.4 Analyzing Log Files for Post-mortem Intrusion Detection

Gamboa, Karen & Monroy, Raúl & Trejo, Luis & Aguirre Bermúdez, Eduardo & Mex-Perera, Carlos. (2012), IEEE Transactions on Systems Man and Cybernetics Part C (Applications and Reviews). 42. 10.1109/TSMCC.2012.2217325.

Upon an intrusion, security staff must analyze the IT system that has been compromised, in order to determine how the attacker gained access to it, and what he did afterward. Usually,