# Management of Distributed Denial of Service Attack in Cloud Computing Environment

*Thesis submitted in partial fulfillment*

*of the requirements for the degree of*

**Master of Technology**

*in*

**Computer Science and Engineering**

*by*


**Abhinav saxena**

**(Roll No: 213CS2170)**

*under the guidance of*

**Prof. Bibhudatta Sahoo**



**Department of Computer Science and Engineering**
**National Institute of Technology, Rourkela**
**Rourkela-769 008, Odisha, India**


**June, 2015.**

Department of Computer Science and Engineering
**National Institute of Technology Rourkela**
Rourkela-769 008, Odisha, India.

# Certificate

This is to certify that the work in the thesis entitled " ***Management of Distributed Denial of service attack on cloud computing environment***" submitted by ***Abhinav Saxena*** is a record of an original research work carried out by him under our supervision and guidance in partial fulfillment of the requirements for the award of the degree of Master of Technology in Computer Science and Engineering, National Institute of Technology, Rourkela. Neither this thesis nor any part of it has been submitted for any degree or academic award elsewhere.

**Prof. Bibhudatta Sahoo**
Assistant Professor
Department of CSE
National Institute of Technology
Rourkela-769008

Place: NIT,Rourkela-769008
Date: 25 - 05 - 2015

# Acknowledgment

First of all, I would like to express my deep sense of respect and gratitude towards my supervisor Prof. Bibhudatta Sahoo, who has been the guiding force behind this work. I want to thank him for introducing me to the field of Cloud Computing and giving me the opportunity to work under him. His undivided faith in this topic and ability to bring out the best of analytical and practical skills in people has been invaluable in tough periods. Without his invaluable advice and assistance it would not have been possible for me to complete this thesis. I am greatly indebted to him for his constant encouragement and invaluable advice in every aspect of my academic life. I consider it my good fortune to have got an opportunity to work with such a wonderful person.

I wish to thank all faculty members and secretarial staff of the CSE Department for their sympathetic cooperation.

During my studies at N.I.T. Rourkela, I made many friends. I would like to thank them all, for all the great moments I had with them.

When I look back at my accomplishments in life, I can see a clear trace of my family's concerns and devotion everywhere. My dearest mother, whom I owe everything I have achieved and whatever I have become; my beloved father, for always believing in me and inspiring me to dream big even at the toughest moments of my life; and my brother and sister; who were always my silent support during all the hardships of this endeavor and beyond.

*Abhinav saxena*

# Abstract

Cloud Computing is a recent technology, it provides a simple and unambiguous taxonomy of three service models available to cloud consumers: software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). There are several security issues with the delivery model of cloud. Our work is to dealing with management of Distributed Denial of Service attack on SaaS model of cloud computing environment. If DDoS attack is capable enough to violate the Service Level Agreement (SLA) on availability it can cause huge financial claim and it will affect the reputation of industries in a market. So our basic aim is to design a management model that will avoid the SLA violation on availability due to a DDoS attack. Our model works in three stages (1) Detection of DDoS attack (2) Avoidance of DDoS attack and (3) prevention of DDoS attack.

Feedforward Neural Network method for detection of DDoS attack. Sigmoid function is used as Neural modal for obtaining the desire output. The Supervised learning model adjusts the connection weight and bias value of ANN model. Using predefined datasets to train the ANN model. For the Avoidance of DDoS attack data center dynamically allocate the resources on virtual machines. A new virtual machine will be clone based on the image file of the original. Replicate the resources on a virtual machine in order to avoid the SLA violation (Availability issues). Message Authentication Code (MAC) is used for prevention of DDoS. The message Authentication code increases the overhead on the network. The design goal is to decrease the overhead of MAC on the network so we are using Router Packet Filtering method that reduces that MAC overhead on packet over the network. This lower overhead increases the speed of authentication and reduces the amount of dynamically allocated resources that will prevent the violation of the SLA on cloud computing.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

**Cloud Computing**

**DDoS Attack**

**Motivation**

**Problem Statement**

# CHAPTER 1

# Introduction

## 1.1 Introduction

Cloud computing is a new prominent technology that making a huge impact on software market. Cloud computing provides different service to Cloud users. Cloud is a model that provide on-demand services and provide different resources. On the basis of demand, it serves sources like network, server, storage and applications. These services are regularly stipulated provided with least management effort and minimum provider interaction. NIST(National Institute of Standard and Technology) publish definition in its publication SP 800-145, "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.".

The basic services provided by Cloud computing are SaaS(Software as a Service), PaaS(Platform as a Service), IaaS(Infrastructure as a Service). The growth of Cloud computing also gives a wide range to attack and affect the vulnerability of the system.

**Infrastructure-as-a-Service** provides a capability to the consumer of processing, storage, network and allow access to deploy and run other software. These software are allow to access operating system and application.

| Security | Description |
|---|---|
| Security Standard | Describe the security standard and precaution measure in Cloud computing. |
| Network | Includes attacks issues in network such as Availability, Denial of Service (DoS), DDoS, internet protocol vulnerability, etc. |
| Access Control | Involves authentication and control on services. It includes the privacy issues of users. |
| Cloud Infrastructure | Covers attack aspects specific to the Cloud infrastructure. |
| Data | It includes data related security issues such as integrity, confidentiality, data ware housing, etc. |

Table 1.1: Cloud Security Issues

**Platform-as-a-service** provides capability to the customer to deploy his own created software or acquired application on Cloud.

**Software-as-a-service** provides capability to the customer to use the provider's application that runs on Cloud infrastructure. These applications are accessible by using various client infrastructure such as web browser.

Clouds bring out a wide range of benefits including configurable computing resources, economic savings, and service flexibility. However, security and privacy concerns are shown to be the primary obstacles to a wide adoption of clouds.

### 1.1.1 Cloud security Categories

We classify cloud computing security related issues into the following five categories.

## 1.2 Distributed Denial of Service attack in Cloud Computing

A Denial of Service (DoS) attack is a one of the major network attack in Cloud computing. The purpose of DoS is to prevent a legitimate user from a specific resource of the Cloud. Network resources such as computer system, web server, or website are blocked by DoS attack. A Distributed Denial of Service attack which is formally known as a coordinated attack. This has an impact on the availability of services that target only particular system or network that is launched indirectly through many compromised computing systems. The services compromised by the attack are those of "primary victim". And the systems that are used to launch DDoS attack are called as "Secondary victims". The Secondary victims are used by an attacker in DDoS attack to increase the impact of the attack. Secondary victims also make it difficult to track down the identity of the attacker [17,36]. DDoS attack will cause huge impact on availability in Cloud computing which can causes violation of Service Level Agreement.

## 1.3 Service Level Agreement in Cloud Computing

Recently, more and more enterprises are also investigating how to leverage on the Cloud computing advantages such as the pay per use model and rapid elasticity. However, major challenges have to be faced in order for enterprises to trust Cloud providers with their core business applications. These challenges are mainly related to QoS, in our view covering dependability, performance and security, and a comprehensive Service Level Agreement (SLA) is needed to cover all these aspects [8,32].

## 1.4 Motivation

Considering more and more resources being shared in Cloud platforms, especially in an elastic Cloud environment which could nearly provide unlimited capabilities, the effect of DDoS attacks can be not only much more powerful and influential, but also in much wider range. DDoS attack serves as a weapon for hackers. The growing dependency on the internet also increases the impact of DDoS attack on the Cloud environment.

1. In October 2002, a herbinger of future large-scale DDoS attacks was introduce that affected 8 of 13 root DNS server, critical server that serving as a roodmap for virtually all internet communication.

2 The impact of DDoS attack is so powerful so that can resulting in frustrated customer and even violate the Server Level Agreement(SLA). According to the Yankee Group, a DDoS attack against Amazon, Yahoo, eBay and other major

sites in February 2000 causes a huge loss of US $ 1.2 billion. Microsoft faced the lose of US $ 500 million in a few days under DDoS attack on its site. According to Total Server Solution, total financial loss from sustaining DDoS attack in 2010 is around US $ 295,000.

3 According to the survey of SC magazine Distributed Denial of Service is the second most security challenge in Cloud computing environment. The Notorious nine: Cloud computing top threats give fifth place to DDoS attack. Cisco put Distributed Denial of Service attack in top security concern in Cloud computing.

## 1.5 Problem Statement

A Distributed Denial of Service attack is one of the availability attacks. A DDoS attack generates a large amount of traffic to the specific system on the network. Therefore, it depletes the system resources, and end users do not receive reliable services. While it is possible to attack various types, it is difficult to detect this attack or to find its source.

In Cloud computing, it is very vulnerable to an availability attack due to the structural features of Cloud system. When physical resources are provided as logical resources via a virtualization layer, availability attacks on one virtual machine can affect other virtual machines which share the resources between them. Due to this feature, availability attacks on infrastructure are very intimidating on those virtual machines which share the physical resources. A Cloud service provider must ensure the certain level of service quality, according to a service level agreement.

Service Level Agreement (SLA) is an agreement between Cloud provider and client. SLA also contain availability issues. A Cloud provider must ensure the availability of service under any scenario (even under DDoS attack).

## 1.6 Research Contribution

This thesis contains a model that deals with DDoS attack on SaaS model of Cloud computing. This model deals DDoS attack in three parts: Detection of DDoS attack, Avoidance of SLA violation and Prevention of DDoS attack.

1. **Artificial Neural Network** (Feedforward Neural Network) computational model is used for detection of DDoS attack. The Artificial Neural Network is trained with predefined dataset. The Mean Square Error (MSE) for classification is compared as per epochs.

2. **Mathematical formulation** that express the number of Virtual machines required to reduce the total execution time to prevent the SLA violation.

3. **Signature based** authentication model with **Router-based filtering** that is use for authentication of users in order to avoidance of DDoS attack.

## 1.7 Thesis Organization

In this chapter, a brief introduction of Cloud computing, Distributed Denial of Service attack, Service Level Agreement, Motivation towards DDoS attack and Problem statement is discussed. The rest of Thesis is organized as follow:

In **Chapter 2**, Cloud computing, DDoS attack in Cloud computing, Impact of DDoS attack in Cloud computing and Proposed method.

In **Chapter 3**, DDoS detection system in Cloud, Related work, ANN based DDoS detection and simulation.

In **Chapter 4**, Software as a service model in Cloud computing, Avoidance of SLA's violation, proposed method and simulation.

In **Chapter 5**, Signature based Authentication, Router based filtering, signature-based authentication, simulation.

# Chapter 2

# Distributed Denial of Service Attack in Cloud Computing

# CHAPTER 2

# Distributed Denial of Service Attack in Cloud Computing

## 2.1 Cloud Computing

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management efforts or service provider interaction.The NIST definition also provides a unifying view of five essential characteristics that all Cloud services exhibit: on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service. The Cloud contain some essential components to provide service to the customer.

### 2.1.1 Cloud Components

In general, A Cloud computing architecture contains several major parts: distributed server, the datacenter and clients. As shown in figure 1.1, these are the major components that build Cloud computing solution. Each element has own specific role. So, every element delivers their specific role on Cloud applications. There is brief description about elements.

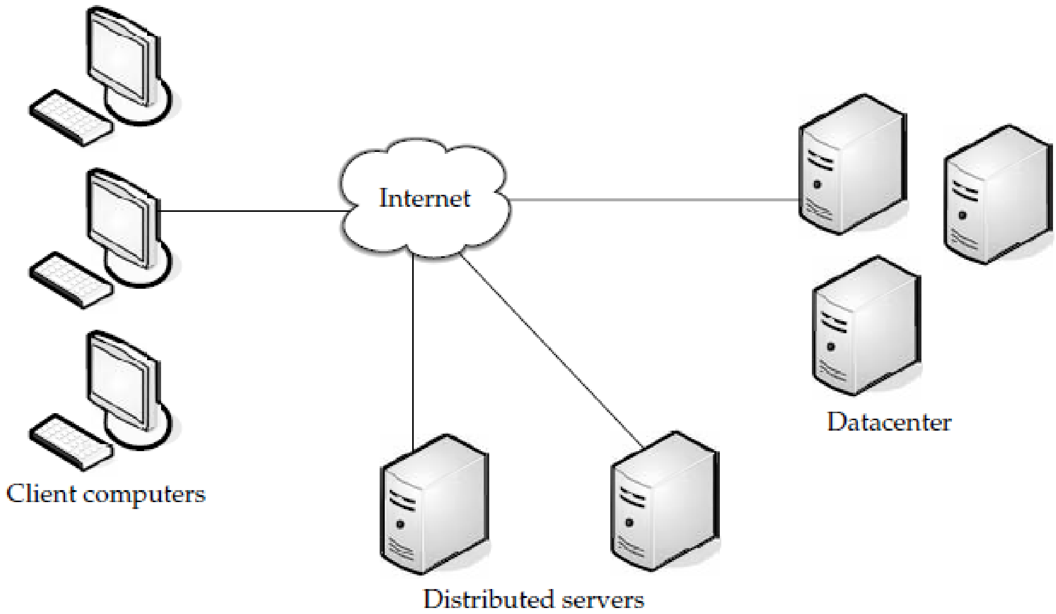


Figure 2.1: Cloud Computing Components

**Client**

The device that interacts with end user on Cloud to manage their information is called Client. We can categories clients into tree types:

- **Mobile** : Mobile devices include PDAs or smartphones, like a Blackberry, Windows Mobile Smartphone, or an iPhone.
- **Thin** : Thin client are computer device which do not contain internal hard drives, in that case server process the entire request and return all the information to display.
- **Thick** : Thick client uses regular computer and connected to Cloud using a web browser like Mozilla firefox, goggle chrome etc.

**Datacenter**

A cluster of servers which provided an application to the subscriber is called Datacenter. Now a day virtualized servers concepts are growing in the market. So that a installed software can have a multiple instances that can be used by different users. A physical server can have multiple virtual servers running the state.

**Distributed Servers**

In distributed server, Servers are situated at different locations. But for a client it seems like it is right away next to each other. A client is able to access data from all the servers even they are located at different places. This scenario gives more options and flexible to provided a different level of security. As like Amazon who has many Cloud solution in the server which is distributed in every part of the world, many are adopting this method. So it is easy to distribute workload if some site will damage, the load balancer distribute that load to other sites and service will still available for customers. Adding more hardware in a Cloud are easily achievable, we can simply add hardware at another site.

### 2.1.2 Issues in Cloud Computing

According to the , we can classify the current Cloud security issues and various security solution. We classified security issues into five categories:

- **Category 1:** The first category is security standard that deals with governing bodies and authorities that define Cloud security policies. It includes service level agreement(SLA),auditing and different agreements between user, provider and stakeholder.
- **Category 2:** Second is network category refers to the medium by which user and Cloud are connected. It includes browsers, connections on network and exchange of information.
- **Category 3:** Access control category include authentication, identification and authorization.
- **Category 4:** Cloud infrastructure category deals security issues with in SaaS, PaaP, IaaS,etc; and it only deal with virtualization.
- **Category 5:** The fields related to data integrity and confidentiality are comes under data category.

## 2.2 Distributed Denial of Service Attack in Cloud Computing

The Distributed Denial of Service attack disturbs the availability of services in Cloud computing. When an attacker induce a large amount of packets to the target server of Cloud the 'Response Time' of service will increase. In DDoS attack, an attacker is hiding their identity by changing its source address.

### 2.2.1 DDoS attack architectures

Two types of DDoS attack on network is in focus:
The Agent-Handler model and the Internet Relay Chat (IRC)-based model.

**Agent-Handler model** of a DDoS attack consists of agents, handlers and clients. Agent software is running on a compromised system that affected by DDoS attack. Basically, handlers are the software packages which is selected over internet. The client is a system from where an attacker handles the DDoS attack on the system. The attacker can easily communicate with handlers and identify which agents in working phase. While performing attack users and owners have no idea about that their systems are participating in DDoS attack and it security has been compromised. Agents are instructed to communicate either with multiple users or with single users. Its depends on the configuration of DDoS attack on the network by an attacker.

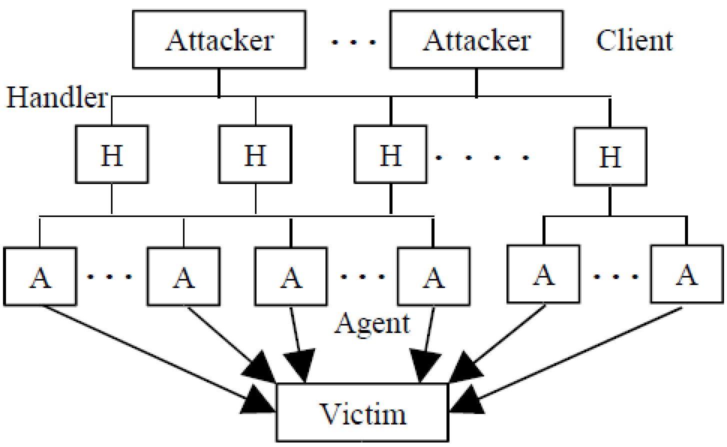


Figure 2.2: Agent-Handler mode

Usually, the attacker chooses router or server of the network where handles a large amount of network traffic. The attacker placed the handler software on that routers or servers. This Software makes harder to differentiate the client messages, handler messages and agent messages. On describing DDoS, the term “handler” can replace with “master” and “agents” can replace with “demons”, respectively.

**IRC-Based Attack Model:** The architecture of IRC is almost same as Agent-Handler model the only difference is that the handler program does not install on a network server instead of that, to connect any client to an agent communication channel is used. The IRC channel having some drawbacks like, an attacker can use “legitimate” IRC ports for sending commands to agents. So in this condition makes difficult to track down the DDoS command packets. The IRC server has a large amount of traffic that helps an attacker to hide their identity. Another disadvantage is that the attacker does not need to generate a database of number of agents since the attacker can access all log file of IRC server, the list of available agent is easy to see. Software installed on IRC network of an IRC channel to perform agent activities and that give an alert message when the agent is in running condition.

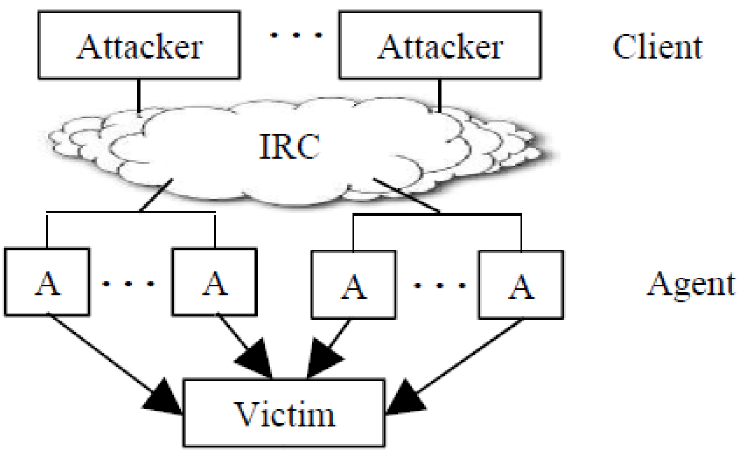


Figure 2.3: IRC-Based Attack Model

In IRC-based DDoS attack architecture, the agents are regularly alluded to as ”Zombie Bots” or ”Bots”. In both IRC-based and Agent-Handler DDoS attack models, we allude to the operators as ”secondary victims” or ”zombies”, and the objective of the DDoS attack as the ”primary victim”. Very much planned agent software uses just a little extent of assets (memory and transmission capacity) so that the clients of optional casualty frameworks experience negligible execution sway when their framework takes an interest in a ”DDoS assault”.

### DDoS attack classification

We propose a classification of DDoS attack. We can mainly divide DDoS attack target in resource consumption attack and bandwidth consumption attack.

**Bandwidth depletion** DDoS attacks can be classified as flooding attack and amplification attack.

In Flooding attack by DDoS attacker sends a large amount of packet to the victim system by selecting agents (Zombies). This attack causes saturation in bandwidth. Flooding attack can be performing by UDP (User Datagram Protocol) and ICMP (Internet Control Message Protocol) packets.

An attacker performs UDP Flooding attack by sending large amount of UDP packet to the victim system. On receiving these UDP packets the victim system start processing data. If requested application is not running on the target port the victim system send ICMP packet to indicate that destination port is unreachable.

ICMP flooding attack is performed by an attacker by sending a huge amount of ICMP ECHO REPLY packets to victim system. Victim system replies these packets and that causes bandwidth depletion. Most of the time ICMP packets send zombies have spoofed IP address, so it is hard to trace the attacker.

In **Amplification attack** agents sending message to a broadcast IP address, it causes all system sends reply of the broadcast address to the victim system. In this attack broadcast IP address amplify the attack traffic and that cause’s victim system’s bandwidth. Some example of amplification attacks are:

DDoS smurf attack preforms by ICMP ECHO REQUEST and REPLY packets. In that attacker sends ICMP REQUEST packet and it amplifies by network by broadcasting it. These packets having spoofed return address (IP address) of victim system.

so each system reply these request by sending ICMP ECHO REPLY packets. So, this attack amplifies the attacking packets and causes bandwidth consumption.

Attacker uses UDP ECHO packets to perform DDoS Fraggle attack. The UDP ECHO packet amplifies over network. In this attacker uses broadcast address that targeted the victim system. This attack will cause huge damage to the victim system.

Resource consumption DDoS attack is performs by attacker by capturing resources of server. Attacker performs the attack by misemploy the protocol of network. By that resource are not available for legitimate users. Some examples of resource consumption attack are:

DDoS TCP SYN attack is performed by exploiting the three-way handshake between sender and receiver by sending bogus TCP SYN request to victim server. By zombies attacker sends large amount of TCP SYN packets to the victim with spoofed source IP address. By receiving these packets the victim server sends ACK+SYN response. When the large amount of SYN request is being sends server runs out of processor and resource of server.

The agents send PUSH+ACK TCP packets to the victim server. When the victim server found that packets it instruct to unload all the data in the TCP buffer. It sends the acknowledgment when completed. By sending the large amount of these packets by agents the server cannot process that much amount of data and system will crash.

## 2.3 Impact of DDoS Attack in Cloud Infrastructure

A Cloud usually possesses profound resources and has full control and dynamic allocation capability of its resources. Therefore, Cloud offers us the potential to overcome DDoS attacks. However,individual Cloud hosted servers are still vulnerable to DDoS attacks. An attacker can perform DDoS attack for a particular service so it will not available for legitimate users.

## 2.4 Proposed Method

In this thesis, a model is proposed which uses practical dynamic resource allocation mechanism to confront DDoS attacks that target individual Cloud customers(service). The Intrusion Prevention System (IPS) is placed at the location to monitor incoming packets. When a Cloud-hosted server is under a DDoS attack, broker will dynamically allocate extra resources from the available cloud resource pool, and new virtual machines will be cloned based on the image file of the original using the clone technology. The IPS filters the attack packets and guarantees the quality of service (QoS) for benign users at the same time. When the volume of DDoS attack packets decreases, our mitigation system will automatically reduce the number of the image file, and release the extra resources back to the available Cloud resource pool.

### 2.4.1 Architecture of Cloud Network

The architecture of Cloud network can model as shown in fig (2.4). Users are connected through the internet, that accesses cloud resource. For performing DDoS attack on Cloud an attack, attacker choose agents(based on the security of nodes). The attacker uses these agents to perform DDoS attack.

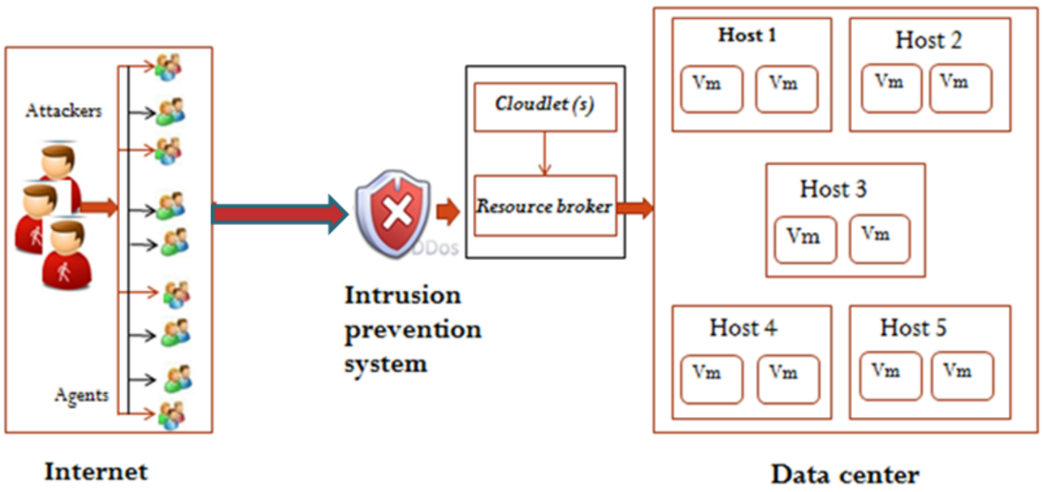


Figure 2.4: cloud architecture

These incoming packets are passing through Intrusion Prevention System. Intrusion Prevention System (IPS) is deployed on the network before broker. This intrusion prevention system is for preventing the Cloud from DDoS attack. The IPS uses the SaaS task (cloudlet) information so that it can detect and prevent the system from DDoS attack according to services. The source broker is the major entity of the system, it acts between datacenter and users. For datacenter the broker is the user, it provide service to the user via a broker.

The broker placed cloudlet on a virtual machine of one of the hosts of datacenter. The cloutlets will run on a virtual machine until the service time of user for that cloudlet not expired. The service time can be by extended by broker.

## 2.5 Conclusion

This chapter contained the brief discussion of Cloud computing and security issues in Cloud computing. It also includes the detail of the impact of Distributed Denial of Service attack in Cloud computing and our proposed method's architecture to deal with DDoS attack in Cloud infrastructure. The IPS deals DDoS attack in three steps: Detection of DDoS attack, Avoidance of SLA's violation and prevention of DDoS attack.

# Chapter 3

# Artificial Neural Network Based DDoS Detection in Cloud

CHAPTER 3

# Artificial Neural Network Based DDoS Detection in Cloud

Recently, many enterprises are interested in a facility provided by Cloud computing like charges as pay per use and rapid elasticity of resources. As the area of Cloud computing increasing the scope of the attack also goes bigger. Distributed Denial of Service is one of the major attacks in Cloud computing. DDoS attack is very common these scenarios, but its impact can crash the Cloud's server. DDoS attack is a very general attack in a Cloud computing environment. DDoS attacks break down the system and make resources unavailable to legitimate users, which can cause violations of availability in " Service Level Agreement (SLA)". And violation of the SLA is a big issue for any corporation. Many algorithms and approaches have been proposed in order to deal with DDoS attack in a Cloud computing environment. We have shown some of the existing approaches.

## 3.1 Related Work

Distributed Intrusion Detection System (DIDS) [34] is proposed by (Steven R., et al.). The author proposed open source network analyzer, which generates log files using snort. In a network, packets is collected by IDS and types of a packet is observed, packets are dropped by comparing the packet information and if not analyzes the packets form a serious attack. If any anomaly is identified the alert messages send to all other IDS and their other modules, by a voting method packet is considered as a bad packet and add it to block list. This basic aim of this method is to provide prevention from the bad packet, but it can cause malfunctioning on Cloud computing .This system works on the network layer of OSI model. This mechanism has a limitation on Cloud computing.

In application layer DDOS attack [10] (by Durcekova) is a detection mechanism that uses SQL server 2005 to store detail of all users into a database and hidden Marokva Modal is used to check browsing behavior of the client. Users will be blocked if known anomaly is detected.

Semantic rule based approach [27](Patel, Ahmed, et al.) is used to detect anomaly behavior on Application layer in cloud. A deterministic finite automaton (DFA) is used for representing characteristics of different malicious activity. This method shows six malicious characteristics are represented by finite automation (FA) processing and the rules are given for example what is the probability of each user traversal if the user does not follow the rule it had given. If anomaly occurs the score of each page is calculated and the two algorithms are used to calculate final threshold value. The final threshold value is compared with scores to determine malicious user.

In Dempster Shafer Theory [22](Lonea, Alina Madalina, et al.) is use for detection of DDoS attack on Cloud computing. In this approach evidence are collected and combined in attack condition. This state of space is defined as a TCP, UDP, Normal and ICMP types of packets. The value in state space element is assigned probability bases and the turted value is generated. This method is lack by its complex behavior of DST and conflicting beliefs.

### 3.1.1 Problem Background

Traditional detection methods such as Distributed IDS is Easy to implement and give better result but classification technique is missing in that method. Hidden Markova Method is easy to implement but number of state on this method is unpredictable. Semantic Rule Based detection is application layer based approach. If anomaly occurs the score of each page is calculated and the two algorithms are used to calculate final threshold value. The final threshold value is compared with scores to determine malicious user. But it may not be predictable some times. Dempster Shafer Theory is applied to detect DDoS. This approach is based on the data collected at the time of attack. This approach is defeated by complexity and its conflicting beliefs.

### 3.1.2 IDS Techniques for DDoS Detection

Signature-based intrusion detection system (SIDS) is use identified that intruders using given a pattern. This pattern is made by some set of rule or we can call signature. The signature-based intrusion detection system able to provide high level of security with accuracy and give very less false positive in order to identifying intruders. A little variation of configuration detection system will affect the whole system (Brown et al., 2002) [3]. So, it will unable to detect an unknown attack and even little variation in the known attack. Signature based attack is preferable because of its easily maintained property and its fast updating configuration system. The system uses many attributes to generate a signature that identifies the traffic. As like, in SNORT (2011, https://www.snort.org) signature is generated by header (e.g. source address, destination address, ports) and some optional values (e.g. payload, metadata). Stiawan et al [35]. (2010) shown signature-based IDS issues and its various framework. In Cloud, signature-based intrusion detection technique is used for detection of a known attack. It can be used either by attaching signature at front-end of Cloud to detect external intrusions or at back-end of Cloud to detect intrusions either internally or externally. Like any other network, the unknown attack cannot be detected in Cloud infrastructure. Some other approaches are also given by Roschke et al. (2009), bakshi and Yogesh (2010), Lo et al. (2008) [25, 29], and Mazzariello et al. (2010) [24] that uses signature based IDS in order to detect intruder on virtual machines.

Anomaly-based (or behavioral-based) detection compares behavior of packets and classifies it into anomalous and the normal flow of packets. Various techniques have been used like data mining, statistical model and Hidden Markov Model (HMM) to detect anomaly behavior of packet flow. Anomaly-based approach collect statistical data of user's behavior, on that basis, a legitimate user will differentiate with a malicious user. This technique has able to detect attacks. This approach is more efficient because it generates a rule that decreases the rate of false alarm for known and unknown attack.

Dutkevyach et al. (2007) gives an approach for a real time system using anomaly based solution to detect intruders. This approach analyzes protocol under attack and traffic in multidimensional order. Zhengbing et al. (2007) give approach of lightweight IDS by optimizes and reduce the number of IDS. This approach can be used in a real-time system. So, it is an efficient and effective technique of an IPS. This technique is based on behavior profile and data mining that periodically update the system for new attack. By using Anomaly detection techniques, a Cloud is able to detect anomaly behavior at a different level. A Cloud is a full of unstable traffic occur on network level or system level occurs, which makes difficult to monitor (or control) by intrusion detection techniques. Garfinkel and Rosenblum (2003), Vieira et al. (2010), Dastjerdi et al. (2009) and Guan and Bao (2009) [12,18,38,42] proposed an approach that works on different level of Cloud to detect intruders. The Anomaly-based detect have ability of soft computing that can deals with uncertain and partially true data that attract to apply for intrusion detection (Moradi and Zulkernine, 2004) [26].

SVM (Han and Kamber, 2006) [15]is an intrusion detection technique (IDS) that works more efficiently when it works on limited amount of data. In Chen et al. (2005) [6], introduces a SVM based approach which gives improved results on false positive rate when it compared with Artificial Neural Network (ANN). The Artificial Neural Network (ANN) required large amount of sample data for training the modal to give better classification compare to SVM. The SVM uses data only in binary form, so this approach can apply in limited specification. Still, SVN approach will improve the detection accuracy result by combining with different techniques (Li and Lu, 2010) [41]. Li and Lu in 2010 purposed a model for network intrusion prevention system. This model is a combination of SNORT and configured firewall. SVM classifier uses SNORT tool in order to reduce reduce the false rate alarm SVM classifier and improve accuracy of IPS. The result of SVM approach is not effective in Cloud.

Genetic algorithms (GAs) (Dhanalakshmi and Ramesh Babu, 2008 and Li, 2004) [7,21] are designed to generate an optimize solution by taking selected features for input. By selecting these optimal parameters the accuracy of Intrusion Detection System will be improved. Lu and Traore (2004) [23] proposed a GP based method that plans outline form feature of network. The effective classification of network intrusions are done by using confidence based fitness function. But in this approach fitness function takes more time in training period. Gong et al. (2005) [13] used some features like Source port, Destination port, Duration protocol, Attack name etc. of packets. For fitness function confidence based framework is used. This approach is more flexible and simple. The designed set of rules used for intrusion detection in network. The feature of categorizing the input increases the accuracy of detection rate. However, best fit is a problem in this approach. Xiao et al. in 2005, introduced Genetic Algorithm and information theory based approach to detect abnormal activities. The information of features of network and various types of intruders is analyzed with network attack. The limitation is, it only works with discrete feature. In Cloud computing infrastructure, the accuracy of Intrusion Detection System will improve by selecting parameters that results optimized output. So, the GA approach can also feasible in Cloud.

Hybrid techniques are a combination of two or more different techniques. NeGPAIM is a hybrid technique that is a combination of two low-level components. For misuse detection, NeGPAIM uses fuzzy logic and for anomaly detection it uses the neural network. For high-level component, it includes one component that is generated by analyzing central engine of two low-level components. Dynamic updating is not required, it results an effective output without a regular update of the model. Some improved model IDS are proposed by, Katar (2006) which generate an approach that is the combination of Naïve Bayes and Decision tree classifiers on different sets of data input. Each classifier generates independent output that is a combination of multiple fusion techniques. The overall performance of this approach is improved because it takes the advantages of a different classifier. It uses that advantage of using soft computing on traditional IDS. A drastic improvement on flexibility is achieved by combining fuzzy logic to data mining techniques. A genetic algorithm (GA) enhance the performance of IDS by using fuzzy logic for classification since GA selects best-fit rules for IDS. For a specific pattern, GA gives better results for matching patterns rather than other general methods (Beg et al., 2010). Some models prefer SVM in case of handling network that have large number of features. Association rule-based IDS is efficient for only correlated attacks. However, associated rule-based IDS uses the predefined data to classifies the data so its efficiency depends on used knowledge base.

| Proposed by | Method | Limitation |
|---|---|---|
| Signature based detection | • Identifies intrusion by matching captured patterns with preconfigured knowledge base.<br>• High detection accuracy for previously known attacks.<br>• Low computational cost. | • Computational Complexity increases exponentially<br>• High false alarm rate for unknown attacks. |
| Anomaly detection | • Uses statistical test on collected behavior to identify intrusion.<br>• Can lower the false alarm rate for unknown attacks. | • Detection accuracy is based on amount of collected behavior or features. |
| SVM based IDS | • Data classification is possible if data is limited.<br>• It include large amount of features. | • It is limited for only discrete features and preprocessing is required. |
| GA based IDS | • It select best feature from given dataset.<br>• Able to handle large amount of dataset. | • Implementation is complex.<br>• Not suitable for general data. |
| Hybrid technique | • It have efficiency to classify accurately. | • Cost required for computation is high. |

Table 3.1: Review on DDoS Detection System in Cloud Computing

### Observations

## 3.2 Artificial Neural Network Based DDoS Detection

Artificial Neural Network's are massively parallel computing systems consisting of an extremely large number of simple processors with many interconnections. This method is inspired by the nervous system of the human brain which have the ability to learn by training and by example. As like human brain which is connected by a lot of interconnected neurons, our model contain hidden layers node which are interconnected with each other and, input and output layers. Input is passed to the neuron of ANN. Activation function generates output by using these signals as input. These results are used for classification.

The ANN is like a black box. This Black box is works as a function which takes multiple inputs and generates multiple outputs. It is not easy to give definition of ANN. Artificial Neural Network (ANN) in machine learning process that design on basis of human brain and consist number of artificial neurons. ANN is used to solve a variety of problems in pattern recognition, optimization, prediction, associative memory and control. The ANN is built by large number of independent connection. The ANN gives better result if data is non-linear dependent between Input/Output.

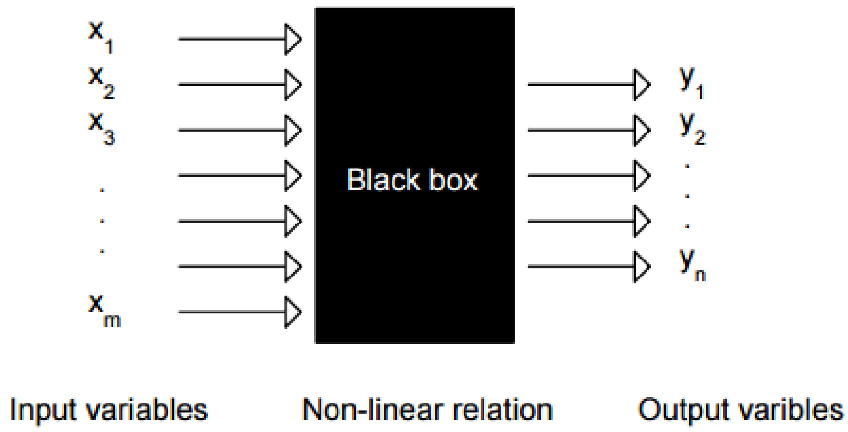


Figure 3.1: Neural network as a black-box featuring the non-linear relationship between the multivariate input variables and multi-variate responses

### 3.2.1 Basic concept of ANN

The neurons of Artificial Neural Network (ANN) work as biological neurons. It also works as biological network like receiving data from its neighbors $x_i$ and process them. The generated output will than forwarded to its neighboring. Neuron “j” decides either to forward output $y_j$ or drop it. Generated output signal is in a form of 0 or 1. It can also generate any real values between 0 and 1, depends on input in binary on real values.

Principally from the authentic perspective the capacity which ascertains the yield from the m-dimensional info vector X, f(X), is viewed as being made out of two sections. The main part assesses the supposed ’net information’, Net, while the second one ”transfer” the net data in a non-direct way to the output value y.

First of all the ANN model takes the input $x_1, x_2, ...x_i, ...x_m$, these inputs are linearly multiplied with weight coefficients $w_ji$. And than transfer these values to the new neuron's axon. The Net value at j-th neuron is calculated as:

$$Net_j = \sum_{i=1}^{m} w_j i x_j \tag{3.1}$$

and next, $Net_j$ , is put as the argument into the Activation function,

$$Net_j = f(Net_j) \tag{3.2}$$

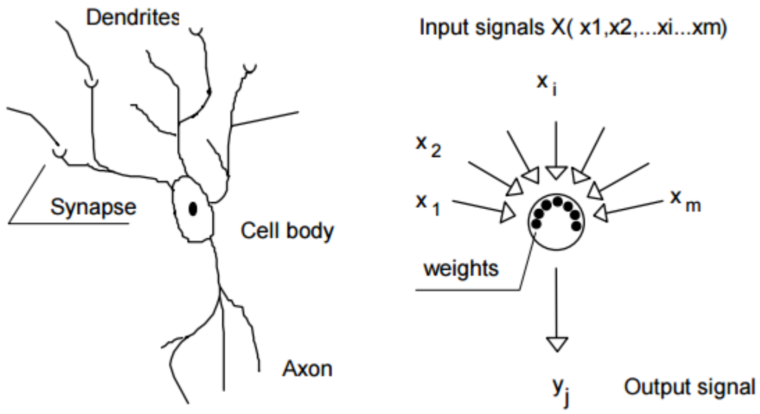


Figure 3.2: Comparison between the biological and artificial neuron. The central body of artificial neuron generate an output Y by taking multiple inputs X.

### 3.2.2 Why Artificial Neural Network

The ANN consist network of Artificial Neurons. That includes

- parallel connection,
- self learning,
- adaptability,
- fault tolerance and
- low energy consumption.

The ANN method is used for pattern classification, classify data into different classes by training the model. The ANN also include some other advantages like:

1. Self-organization: ANN is able to generates its own organization and information collated at the time of learning will used for representation.

2. Adaptive learning: ANN have functionality to learn it self at the time of training using given data.

3. Fault tolerance: Network failure can damage the whole system and consistent decrease in performance. ANN have ability to deal with these failures of network by the use of redundant information coding.

4. Real time operation: ANN have capability to computes the task in parallel. but it required some special hardware for this purpose.

### 3.2.3 Background

Artificial Neural Network [20] (Jin Li, Yong Liu) is used to detect DDoS attack in network. The Linear Vector Quantization (LVQ) model is used for detection of DDoS detection. LVQ network achieve a good classification of the linear network effects. It effectively avoid the requirement of linear network data must be linearly separable. However, LVQ model of Artificial Neural Network is complex in nature.

The Feed Forward Neural Network model is simplex model of ANN. Feed Forward neural network performs well with linearly separable data. In SaaS model of Cloud computing an client need to have authority to use a service provided by Cloud. Using this feature we can classify the user into attacker or a legitimate user.

### 3.2.4 Feed Forward Neural Network Architecture

Artificial Neural Network's are massively parallel computing systems consisting of an extremely large number of simple processors with many interconnections. This method is inspired by the nervous system of the human brain which have the ability to learn by training and by example. As like human brain which is connected by a lot of interconnected neurons, our model contain hidden layers node which are interconnected with each other and, input and output layers. Input is passed to the neuron of ANN. Activation function generates output by using these signals as input. These results are used for classification.Feed Forward Neural Network is built by interconnecting independent neurons. Based on the type of connection the architecture of ANN can be divided into various types. The model of Feed Forward neural network is shown in fig 3.2.

It shows three layer neuron network. This architecture do not give feedback to the network and input flows in a forward direction that is why is called as Feed Forward neural network. This neural network has the ability to learn any nonlinear mapping.

A mathematical operation performed by neuron on receiving input from previous layers. The mathematical formula used for calculation is:

$y_i = f_i(w_i j.x_i + b_j)$

The output y is generated by neural network is given by:

$y_k = f_{ko}(\sum_{j=1}^{N} W_{jk}^{h} f_k(\sum_{j=1}^{P} W_{ji}^{I} + b_0) + b_j)$

Where

$W_{jk}^{h}$ is a connection weight between hidden layer and the output layer.
$W_{ji}^{I}$ is a connection weight between the input layer and hidden layer.
$b_0$ is bias value of output layer.
$b_j$ is bias value of hidden layer.
$f_k$ is the hidden layer activation function.
$f_{ko}$ is the output layer activation function.

### 3.2.5 Neuron Model

An activation function describe the neuron model. Log-sigmoid function (also known as logistic function) uses hyperbolic tangent function. The Sigmoid function is similar to step function, however the addition of region is uncertain. Sigmoid function is very similar to input/output relationships of biological neurons.

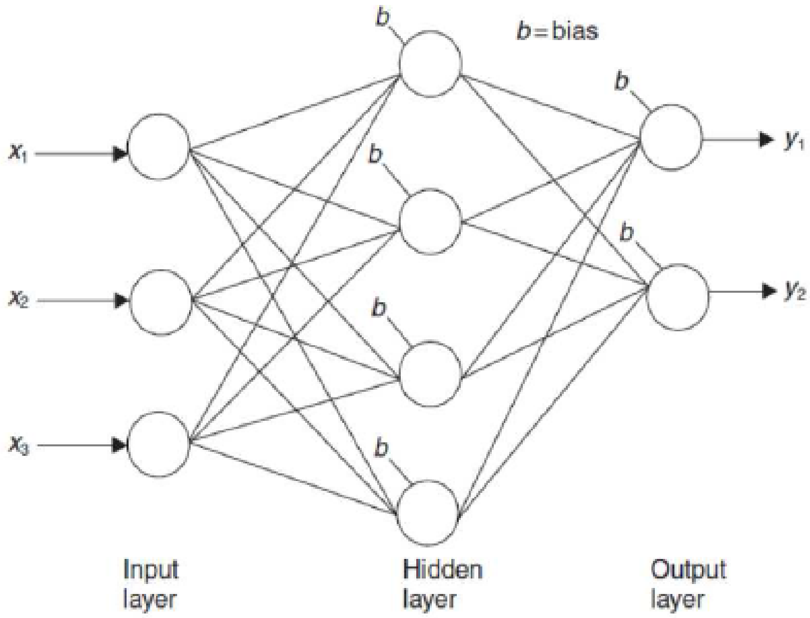


Figure 3.3: Three layer feed forward network

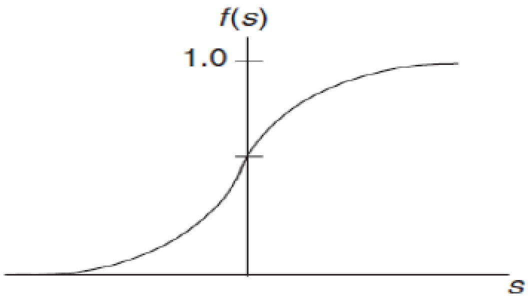


Figure 3.4: Sigmoid function

The activation function (f) used to calculate the desire output, with x as an input, b is a bias value and w is the weight given to each input. The sigmoid activation function is used for neural network. The activation function is given by:

$f_i(s) = \frac{1}{1+e^{-s_j}}$

### 3.2.6 Learning in Neural Networks

Feed Forward Neural Network learning process of adjusting the connection weight and bias value of neural network to generate a desired output. There are two basic training modes.

**1.** Supervised learning - In this learning process of neural network a give sample data is used to trained to network. Give sample contain a set of input and the desire output. The difference of actual output and desire out is used to manage the value of connection weight.

**2.** Unsupervised learning - It is an self learning process, it do not takes any feedback to manage the weight.

### 3.2.7 Training Algorithm: Backpropagation

The Backpropagation is a learning algorithm. It search for the weight values that minimize the total error of ANN model. The learning algorithm have two steps: *Forward Pass* is the first step where network is activated and error of output layer is calculated. In *Backward Pass* the computed error is used for update the weights of network. This is done recursively until weights of network is stable. Backpropogation weight update is based on gradient decent method. The Backpropogation method is terminated when summation of error of all output data is lesser that some threshold value in that epoch.

Now a day Back propagation is famous training algorithm of ANN. A Gradient decent optimization approach is uses to train a neural network. The objective function is given as the summation of squared error.

$$J = \frac{1}{M}(\sum_{j=1}^{M}(d_j - y_j)^2)$$

where,
$d_j$ are the desired output,
$y_j$ are the actual network outputs.
For minimizing the total error, weight matrix has to be updated. The updated formula is given as:

$$W_{new} = W_{old} + \Delta W$$

where$\Delta W = -\mu \frac{\vartheta J}{\vartheta w}$

## 3.3 FFNN Algorithm for DDoS Detection in Cloud Computing

Feed Forward Neural Network model is uses for detection of normal and abnormal activity on Cloud computing network. The input values are taken from IP packet information and output value will classify the data into "Attack or Normal" data flow. The predefined dataset is uses to train Feed Forward Neural Network. Using learning algorithm to manage the weight metrics to minimize the mean square error.

**Algorithm 1:** FFNN learning Algorithm

**Data**: Predefine dataset of DDoS attack, Max iteration.

**Result**: A stable weight matrix.

k = 1;

Initialize **weights matrix** with random value;

**while** *k* $\leq$ *Maximum iterations* **do**

  **for** *each Inputs* **do**

    for each network input x the output value y is calculate. calculate the error and update the weights in backward order: $W_{new} = W_{old} + \Delta W$.

  **if** *Threshold value satisfied* **then**

    **return**

  k = k+1;

### 3.3.1 Input and Output

In Feed Forward Neural Network, desired data will be collect from IP packet information. The attributes which are linearly independent use for input in Feed Forward Neural Network. We are selected seven attribute of IP packets for Input data is following:

- Source IP address,
- Destination IP address,
- Type of packet,
- Port number,
- Flow rate of different Packet type (ICMP, UDP, TCP, etc),

- Application ID,
- Packet number.

**Hidden Layers**

Deciding number of hidden layers in an important part of neural network. Number of neural network will affect the overall performance of neural network. According to some 'rule-of-thumb' method the number of hidden layers use for neural network should be 2/3 the size of the size of the input layer. So we are taking 6 hidden layers in neural network.

## 3.4 Simulation

Feed Forward Neural Network is simulated for DDoS detection by using real DDoS dataset, DDoS LLS DDoS-1.0 dataset, available from MIT [11]. The dataset is modified according to SaaS model under DDoS attack. 6 hidden layer are used for simulation. 30 percent of total data is a DDoS data. This experiment was performed on Intel Pentium dual-core, 2.6 GHz, 2 GB Ram and 500 GB SATA hard drive. The tool using for implementation is MATLAB R2012b. The graph is plotted between number of epochs and Mean Square Error when performing Training, Validation and Testing. Simulation is performing in a manner so that 5 SaaS task on Cloud infrastructure out of which 2 SaaS task are affected by DDoS attack.

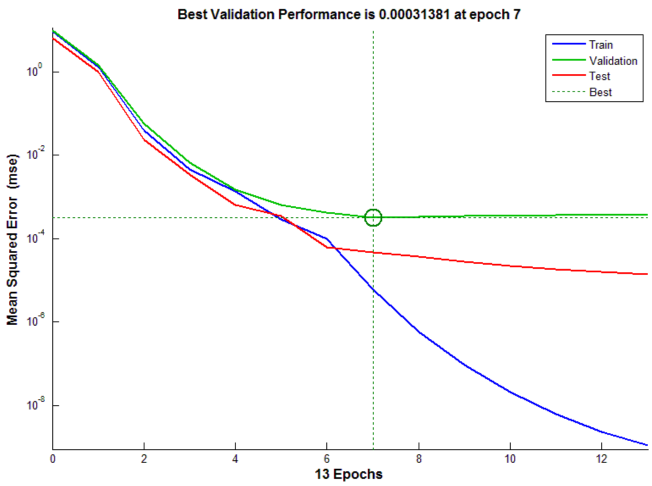


Figure 3.5: Performance measure of Mean Square Error (MSE)

In this simulation 60 percent of dataset for Training, 20 percent of dataset for Validation and 20 percent of dataset for Testing. From the graph we can see that best value of Mean Square Error (MSE) is 0.00031381 at epoch 7. This Mean Square Error is error in classification of packet flow into ”Attack and Normal”.

## 3.5 Conclusion

In this chapter contains the description about Feed Forward Neural Network and simulate the result using Matlab tool. The predefined dataset (DDoS LLS DDoS-1.0 dataset, available from MIT) for training, validation and testing. Dataset has been modified as per SaaS model of Cloud computing. On analyzing the graph of performance measure, we can see the mean square error of classification of data is consistently decreasing. So the performance of DDoS detection in Cloud computing in simultaneously increasing using Feed Forward Neural Network.

# Chapter 4

# Avoidance of Service Level Agreement's Violation

**Service Level Agreement in Cloud Computing**

**Software as a Service model in Cloud Computing**

**Proposed method**

**Simulation**

**Conclusion**

# CHAPTER 4

# Avoidance of Service Level Agreement's Violation

## 4.1 Service Level Agreement in Cloud Computing

In cloud computing, the customer has less control on service delivery model. To prevent themselves from getting low performance a customer need to have some guarantees document. An SLA is an agreement between service client and service provider. An SLA has the detail of the level of service provided to the customers. SLA also contains the information of how to measure the quality of service, where a client will report of service and how violation should handle [4,16,37]. Our focus is on availability attribute of SLA. Availability can be defined as "the readiness of correct service". We can say availability of service is a probability of receiving service that defines in agreement. To avoid down times that affect service availability, fault avoidance and fault tolerance are used [2,14,28].

#### SLA on Availability

Among available cloud offerings, storage service reveals an increasing level of market competition. According to iSuppli [33] global cloud storage revenue is set to rise to $5 billion in 2013, up from $1.6 billion in 2009.Management of computing resources as a service by a single company implies the risk of single point of failure such as financial difficulties (bankruptcy), software or network failure, etc.. However, even if the vendor runs data centers in various geographic regions using different network providers, it may have the same software infrastructure. Therefore, a failure in the software in one center will affect all the other centers, hence affecting the service availability. In July 2008, for instance, Amazon storage service S3 was down for 8 hours because of a single bit error [1].

Cloud Service Providers (CSP) usually offers two resource plans to the customers: Short term on-demand, in this plan a customer will charge on the basis of what he uses. This business model is not feasible for Economic Denial of Service (EDoS) attack. The service will be affected if cloud allocates resource based on number of instances.
Long-term reservation, in this scheme a customer occupies or reserved maximum usage limit of resource. There is a limit in reserved resource for application, so DDoS is possible in this situation.
Most of the cloud providers like Go Grid, amazon, etc. offers both plan.

## 4.2 Software as a Service model in Cloud Computing

The SaaS model of cloud computing comes popular in these days. Most of the business models and researchers have great attention on the Software-as-a-service model in cloud computing. The SaaS model has a functionality of on-demand service to users. So it depends on computing utility rather than the stand alone software itself.
To provide service to the Cloud customers, different SaaS task runs on a virtual machines. Assuming each SaaS task have a specific size. Every SaaS task runs on a virtual machine of Cloud infrastructure that are situated in Datacenter. Every virtual machine has different execution speed, so execution time will vary on virtual machines.

## 4.3 DDoS's Impact in Cloud Computing

A public cloud environment has a large amount of resource. It is easy to handle or maintain the quality of service even in rapid growth in service demands. Therefore " it is almost impossible for a DDoS attack to shut down a whole Cloud ". However, individual cloud customers (those parties hosting service in Cloud) is unable to escape from DDoS attack. These individual cloud customers do not have the advantage to avoid the DDos attack impact. So, When an attacker performs DDoS attack with targeting specific customer, the service of that customer will be affected. On generating a large amount of service request packet by an attacker, the Cloud provider will not be able to reply all packets on a limited interval of time. The large amount of request for a service will the 'Response Time' of a service. If the 'Response Time' exceeded a 'Maximum Response time' value in Service Level Agreement (SLA) will be violated.

## 4.4 Proposed Method

In approach is to dynamically allocate the resource on a virtual machine in order to avoid SLA violation by DDoS attack that targets individual cloud customers. So we can say there will be the several resources accessible point for customers in cloud data center. The IDS monitors the incoming traffic of packets. When the IDS detect DDoS attack, extra resources will be dynamically allocated from an available large pool of resources. The new machine allocation will be clone based image file of original resources.

### 4.4.1 Problem Formulation

Let there are 'n' virtual machines which is donated by a set:

VM = $\{vm_1, vm_2, vm_3, ..., vm_n\}$,

and a SaaS composed of 'm' task:

ST = $\{st_1, st_2, st_3, ..., st_m\}$;

Our objective :

$ET_{total} = \Sigma ET_i \leq MRT_{sla}$

where, i $\forall$ SaaS task and

$MRT_{sla}$ = Maximum Response Time in sla.

### 4.4.2 Formula Generation

Assuming that all SaaS tasks are fixed size. So, under DDoS attack an individual service will affect and it required more SaaS task copies to fulfill customers demand in a limited interval of time. We consider that increase the number of virtual machines also increases the performance.

⇒ Initial Execution Time, $ET_{init} = \frac{SaaS\ task\ size(st)}{MIPS_{init}}$;

Where, MIPS (Million Instruction per Second) is performance of a virtual machine,

⇒ Number of task Executed by $i^{th}$ vm in $ET_{init}\ time = \frac{MIPS_i}{MIPS_{init}}$.

⇒ Total number of task executed in $ET_{init}$ time = $\Sigma \frac{MIPS_i}{MIPS_{init}}$.

⇒ As we assume performance of virtual machine in linearly increasing.

So for 'm' virtual machines, $\Sigma \frac{MIPS_i}{MIPS_{init}} = \frac{m \times \{MIPS_{init} + MIPS_{max}\}}{2}$.

⇒So, Total Execution time for 'm' tasks $ET_{total} = \frac{ET_{init} \times Total\ number\ of\ SaaS\ \ task(m)}{Number\ of\ tasks\ executed\ at\ ET_{init}\ times}$.

⇒ Putting values into equation, $ET_{total} = \frac{2 \times n \times ST}{m \times \{MIPS_{init} + MIPS_{max}\}}$.

This formula is used to identified the number of virtual machines needed to reduce the total Execution Time ($ET_{total}$) and SLA maximum response time will not violated.

## 4.5 Simulation

We simulate our model on java platform. This experiment was performed on cloudsim 3.0 using NetBeans IDE 7.4 . The is plotted between execution time and number of cloudlets (SaaS tasks). In this graph (fig. 4.1) we are showing execution will decrease by increasing virtual machine according to this formula.

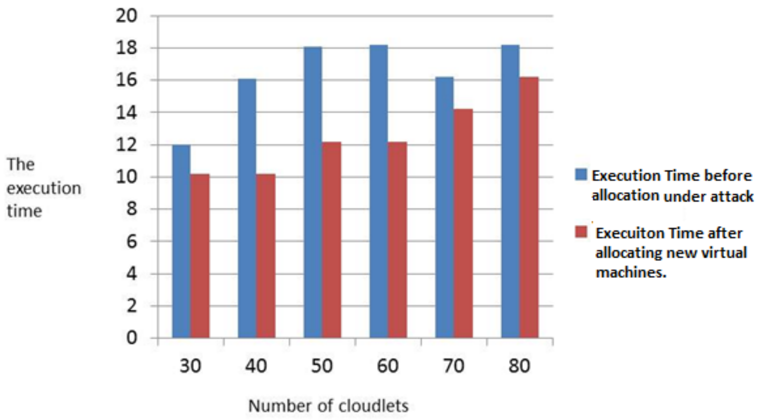


Figure 4.1: Decrease in execution time by allocating new virtual machines

## 4.6 Conclusion

In this chapter, we are given a mathematical formula, on that basis a cloud provider can replicate it resources to maintain the quality of service. By this formula, a service provider will ensure that the number replica of original resource need generate in order to avoid SLA violation.

# Chapter 5

# Signature Based DDoS Prevention

**Related Work**

**Router-Based Packet Filtering**

**Signature Based Authentication**

**Simulation**

CHAPTER 5

# Signature Based DDoS Prevention

Till now we discussed how our purposed model detects and avoids the distributed denial of service attack. Now we use Router-Based packet filtering and Message Authentication Code(MAC) to authenticate the sender, that ensures the sender is an legitimate user or a attacker. So, it ensure the originator of the message.

## 5.1 DDoS Prevention System

We are allocating extra resource from public cloud resource pool for preventing Serves Level Agreement (SLA) between client and service providers. But only generating clone of resources is not a solution of DDoS attack, we also need to discard all the packets those are involve in DDoS attack. To filter the packets client need to show their authenticity to the server. For checking their authority we are using Message Authentication Code (MAC) in IP packets.

## 5.2 Related Work

Various schemes has been introduced to prevent from Distributed Denial of Service attack. In this section we shown detail of some prevention technique and their observations.

Collaborative Peer-to-Peer Architecture(2008), Saad Radwane et.al [31] proposed a peer to peer architecture for a collaborative defense against DDoS attacks. This architecture include an Intrusion Detection System which gives the information about DDoS attack. This IDS is deploy on a network and detect the attack based on rate of flow of network traffic. The basic drawback of this architecture is the fix threshold value not enough to decide whether the flow of packet is attack or not.

Intrusion Detection in the Cloud (2009), Roschke et.al in 2009 [30] proposed a deployment architecture of Intrusion Detection Systems in the Cloud.He proposed the deployment of IDS on each layer of the cloud to gather the alerts from the sensors deployed in the cloud.The challenges faced during the implementation of this IDS are: the output of different sensors is not standardized, communication between sensors and the management component is not much flexible, complex architectures with multiple sensors.

DDoS Defense as Network Service (2010), Du Ping et.al [9] proposed Cloud Based attack Defense System (CLAD). It runs on cloud infrastructure in order to protect web servers. The protection of CLAD system is based on network layer defense system which depends on execution environment of cloud infrastructure. The limitation of this system is that, it is only applicable for small infrastructure.

CBF: Confidence Based Filtering method introduced by Chen Qi et.al [5] in 2011 for cloud environment. It used to distinguish attack packet from legitimate packet on the bases of correlation pattern. CBF in real time filtering give enough accuracy. However, this method is fast but maintaining database is a costly.

CIDS : Intrusion Detection in Cloud Systems (2012) Kholidy A. Hisham et.al [19] in 2012 proposed a method called Cloud based Intrusion Detection System (CIDS). In this method IDS is classified: Host based IDS, Network based IDS and Distributed IDS. CIDS has no central coordinate and has a elastic architecture wit P2P solution. The main advantage in this method is that, it distribute the load at several cloud servers.

But in order to apply CIDS methodology we need to adapt an appropriate dataset.A cloud user can have more than one instance of its data in a multiple virtual machine and that currently not available in exited dataset.

Defense of DDoS attack for Cloud Computing, Yang Lanjuan et.al [40] in 2012 introduced a defense system for cloud based on Software Oriented Architecture (SOA).It is used to trace down the source of DDoS attack and a differentiate legitimate packet from attack packet by packet filtering approach. SBTA (SOA based traceback approach) is used to mark tag on header of the message packet and send it to web server. The drawback of this system is that as number of attack packet increase the efficiency of this system consecutively decrease.

## 5.3 Observations

| Proposed by | Method | Limitation |
|---|---|---|
| Saad Radwane et.al [31] | DHT (Distributed Hash Table) algorithm | Fixed threshold value to detect DDoS |
| Roschke et.al [30] | Network-based sensors and host based sensors for each layer | complex architectures with multiple sensors |
| Du Ping et.al [9] | Runs CLAD on execution environment | applicable on small sized companies only |
| Chen Qi et.al. [5] | Confidence Based Filtering method(CBF) | fast but costly method |
| Kholidy A. Hisham et.al [19] | CIDS architecture distributes the processing load at several cloud locations | currently not available for the existing datasets |
| Yang Lanjuan et.al [40] | Software Oriented Architecture (SOA) to traceback the source of DDoS attack | attack packet increases defensive performance decreases progressively |

Table 5.1: DDoS Prevention Reviews

## 5.4 Signature Based Authentication in Cloud Computing

According to [39], if Pentium 300 MHz machine spends all its processing time even than it can generate only 80 512-bit RSA signatures and 2000 signature per second will be verifies. This signature based scheme also makes 64 bytes of overhead per packet on communication. So we can say this approach is not possible.

So, on behalf of signing every packet, when our ANN detection algorithm detects DDoS attack we sends request to senders to generate a Message Authentication Code (MAC) to verify its identity.And we also put Route-Based packet filtering to reduce amount of MAC need to generate.

## 5.5 Router-Based Packet Filtering

The goal of router-based filtering can be classify in two groups: (1) Proactive filtering block IP spoofed packet from reaching the server, and (2)Reactive is identification of spoofed IP flow i.e IP traceback.

We are using proactive and reactive scheme in Router-based distributed packets filtering (DPF) techniques. The router information is used by Router-based distributed packet filtering when the packet is arriving at the router. A border router filter at Autonomous System (AS) this border router drop the packet based on IP address of source IP address. Single AS has limited impact in order to filter IP spoofed packets.

#### Assumption

According to [11], internet Autonomous topology is based on power-law connectivity. It is also consider in router topology. A 'power-law' is a topology where the connectivity of graph is concentrate in few "center" or nodes.

### 5.5.1 Router-Based Detection and Discarding of Spoofed IP Packets

Let us consider the sample network graph as shown in fig 5.1. All these nodes shown are Autonomous Systems (AS). This is an undirected network graph. Now consider the host of AS6 is trying to perform DoS attack, where the server is at AS7. The attacker on AS6 using a spoof IP address of node AS3. But when generated from AS6 will reach boarder router AS4, the router topology will identify that the packet having IP address of AS3 cannot pass through the link (4, 7). On that basis router-based filter will identify that the packet is a spoofed IP packet. So this packet will be discarding at AS4. The process of dropping the packet is a proactive filtering. Note that to identifying a spoofed IP packet for proactive filtering, AS only need to inspect the source IP of the packet from router topology table. The packet will be drop on the basis of the path of the packet from the source.

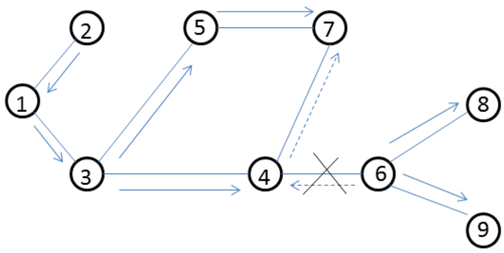


Figure 5.1: Illustration of route-based packet filtering executed at node 4.

### 5.5.2 Performance Measure for Packet Filtering

Let T is a subset of vertex (V). i.e T $\subseteq$ V. T is the set of AS or node on which filters are deployed. Where $\gamma$ = |T| /|V| is known as a coverage ratio to measure the filtering performance. Two metrics are used for measuring performance.

**Proactive** is scalar values that varies from 0 to 1. It shows the fraction value of node which cannot be spoofed by an attacker. The proactive performance matric is denoted by $\Phi_2(1)$.

**Reactive** Reactive is a variable value that also varies from 0 to 1. In this case the value of $\tau \geq 1$. This value denote the fraction of Autonomous system where original source can be identifies within $\tau$ sites. This is denoted as $\Psi_1(\tau)$.

We are using two sets S$_{a,t}$ and C$_{s,t}$ (a,s,t e V). which is used to measure performance metrics. S$_{a,t}$ are the family of nodes or more precisely we can say group of IP addresses. So that an attacker (a host in AS) can send spoofed IP address that will reach the t node without being dropped by any router-based filtering on AS (member of T). As the member of $S_{a,t}$ set increase, attacker have more option for generating spoofed packets which will be undetected from set of filter T.

$C_{s,t}$ is a set of node which define in victim's perspective. It is a set of node those are able to send IP spoofed packet, having IP address s and t is target (destination) address. As the set member of $C_{s,t}$ increase, the probability of receiving spoofed IP packet will also increase.

Note for any value of s and t, $|C_{s,t}|$ = 1 means no attacker can use s as a source IP address of spoofed packet to attack at target t.

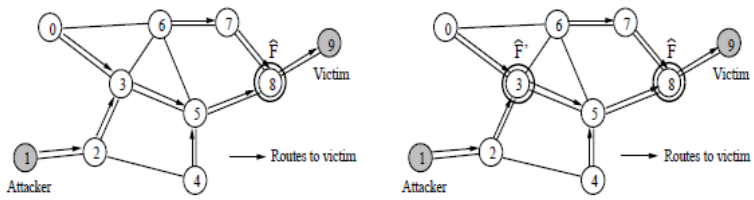


Figure 5.2: Left: Withroute-based filtering executed at node 8, Right: Distributed filtering with filter F at AS 3.

Figure 5.2 illustrates the impact of route-based distributed filtering on curtailing the attacker's ability to engage in spoofing. Without route-based filtering, an attacker residing at AS 1 can disguise himself with undetectable spoofed IP addresses belonging to AS 0−8, i.e., $S_{1,9}$ = 0, 1, . . . , 8, when attacking a server in AS 9. With route-based filtering at AS 8, the spoofable address range shrinks to 0, 1, . . . , 5. With distributed filtering at AS 8 and AS 3, $S_{1,9}$ = 1, 2.

### Proactive Filtering Measures

The proactive filtering performance measure is given by formula:

$$\Phi_1(\tau) \quad = \quad \frac{|\{t : \forall\ a \in V,\ |S_{a,t}| \leq \tau\}|}{n}.$$

The value of $\tau$ is greater than equal to 1 $(i.e\tau \geq 1)$. Basically $\tau$ is a fraction of node which cannot reached to target from any source. It look meaningless to calculate value of $\Phi_1(\tau)$ where $\tau \geq 2$. So, we focus on another proactive measure:

$$\Phi_2(\tau) \quad = \quad \frac{|\{a : \forall\ t \in V,\ |S_{a,t}| \leq \tau\}|}{n}.$$

### Reactive Filtering

Every IP spoofed packet cannot be filters by proactive filtering, thus the remaining part will measured by reactive filtering metric $\Psi_1(\tau)$. That is measure the trace back of spoofed IP packets.

$$\Psi_1(\tau) \quad = \quad \frac{|\{t : \forall\ s \in V,\ |C_{s,t}| \leq \tau\}|}{n}.$$

Here, $\tau \geq 1$ have a sense if value of $\tau$ is greater than 1. That represents the ambiguity to trace back source of spoof IP packet.
For reactive filtering we use message authentication code to verify the user.

## 5.5.3 Selection of Filters

There are a lots of Autonomous system (AS) in a network. The select of a limited amount filters from these AS is a NP-hard problem. So, the traditional way is to chose random AS for filter. The Greed approach for selecting filters from AS that improves the results.

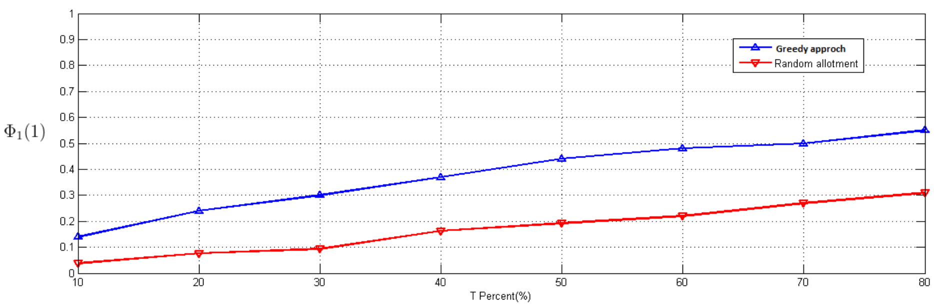


Figure 5.3: Proactive filtering performance

## 5.6 Message Authentication Code for Authentication

Massage Authentication Code is used to authenticate the user. As shown in fig 5.4, the sender generate a random private key and send it to receiver. Sender uses a hash function to create a MAC from the concatenation of the key and the message, h(K|M), then send this packets. Receiver separates the message from the MAC. Receiver than compares the newly created MAC with the one received. If the two MACs match, the message is authentic.

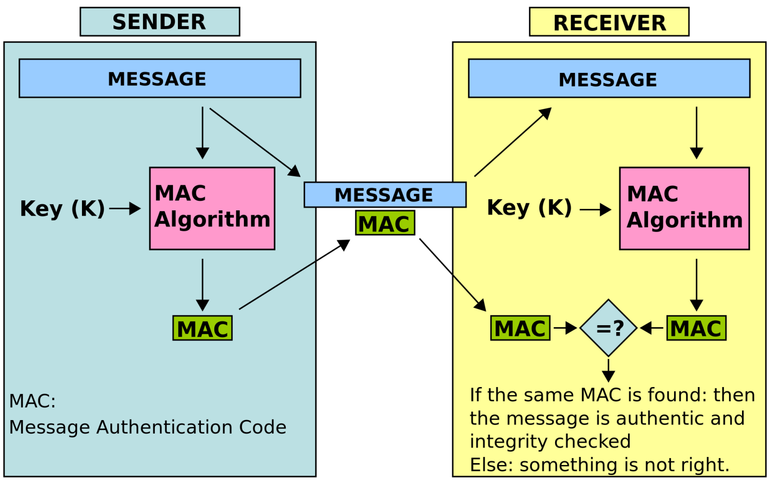


Figure 5.4: Message Authentication Code

Hashing algorithm is use for generation of 'Message Authentication Code'. Any cryptographic hash function, such as MD5 or SHA-1, may be used in the calculation of an HMAC; the resulting MAC algorithm is termed HMAC-MD5 or HMAC-SHA1 accordingly. Observation of some hashing algorithms are follow:

| Algorithm | Message Size (bits) | Block Size (bits) | Word Size (bits) | Message Digest Size (bits) |
|---|---|---|---|---|
| SHA-1 | $< 2_{64}$ | 512 | 32 | 160 |
| SHA-256 | $< 2_{64}$ | 512 | 32 | 256 |
| SHA-384 | $< 2_{128}$ | 1024 | 64 | 384 |
| SHA-512 | $< 2_{128}$ | 1024 | 64 | 512 |

Table 5.2: Message Authentication Code Size

## 5.7 Message Authentication Code in Cloud Computing

The private key is given by the Cloud providers when a user pays for the service. This key is use to generate Message Authentication Code is Hash Function (H). This MAC is attached with the packets when the server demands verification.

When IDS (Intrusion detection system) detected DDoS attack on our system the server demand Message Authentication Code (MAC) from the users. We are assuming that only those users who are not authorized to use that service is performing the attack. On receiving MAC request from the server, Spoof Defense System (SDS) is used to generate and add MAC to each packet.

The source put its signature and IPS validate that attached signature using source private key. The approach of attaching and verifying the signature on packets is not feasible because of it expensive computation. On merging with Router-Based filtering, a lot of packets will be discarded at a router on the basis of the source path. So the overhead of MAC will reduce in a significant amount. This makes possible to use MAC for authentication.

## 5.8 Simulation

We have simulated DDoS prevention using Signature Based authentication with router-based filtering. We have simulated our result on C compiler. We generated ”Power-law” network with 3015 nodes. The graph is generated between MAC overhead on network and number of Agents chooses by an attacker to perform DDoS attack. The Message Authentication Code of 64 Bytes (512 bits) is added to the packet.

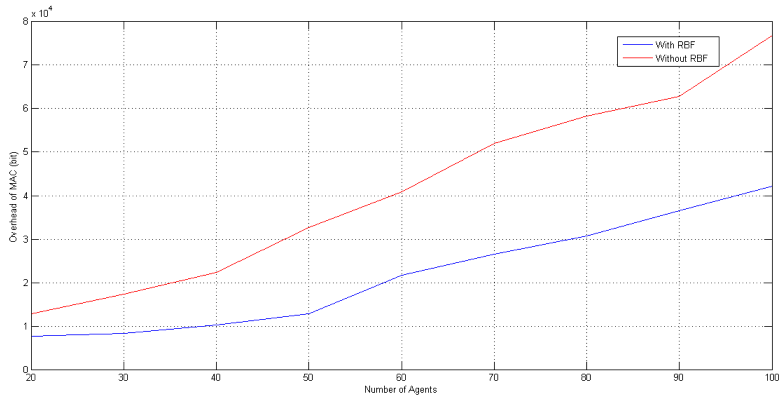


Figure 5.5: Overhead on network with 20 percent of victim

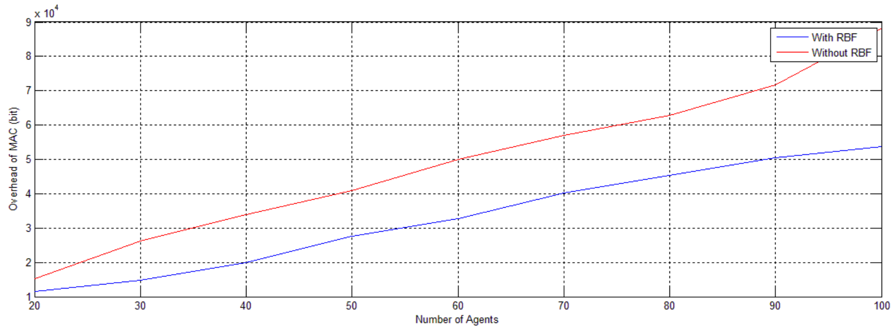


Figure 5.6: Overhead on network with 30 percent of victim

## 5.9 Conclusion

In this chapter consist an approach to the authenticating user using signature-based model. The use of router-based distributed packet filtering approach reduces the overhead on a network of MAC attach with packets. The valid signature is only generated by an authentic user (authorize to use the service in Cloud). This Message Authentication Code is used to verify that whether the user is authorized to use this service or not in Cloud.

# Chapter 6

# Conclusion and Future work

CHAPTER 6

# Conclusion and Future Work

In this thesis work, we have discussed the Distributed Denial of Service attack in Cloud computing. There are so many traditional ways to deal with Distributed Denial of Service attack in Cloud computing. Our focus is to deal with DDoS attack in such a way so that Service Level Agreement (SLA) between customers and Cloud provider will not violate. The research model work in three steps: Detection of DDoS attack, Avoidance of SLA's violation and prevention from DDoS attack. Feed Forward Neural Network (FFNN) is used for detection of DDoS attack. The inputs are selected from IP packet's information. Matlab tool is used to calculate the Mean Square Error (MSE) to the number of Epochs. The consistent decrements in Mean Square Error will show the reduction of error in detection of DDoS attack in Cloud using ANN.The dynamic allocation of a resource on virtual machines to ensure the avoidance of SLA's violation. The generated formula is use to verifies that number of virtual machine needed to reduce the response time. Our objective is to keep the Response time lesser than the Maximum response time mentioned in Service Level Agreement. The Signature based scheme is used for authentication of users. The user who pay for the service in Cloud maintain a private key which is use to generate Message Authentication Code by the hash function. Use of Router-based filter reduces the overhead on the network. The use of Distributed Router-based filter reduce the cost of signature-based filtering and makes it feasible to use in Cloud computing.

However, there has been few constraints associated with our work which we want to sort out in future. Firstly, our Intrusion Prevention System is analyzes the flow of network packets. In future, we want establish a communication system between these IPS so DDoS will be detected earlier. Moreover, the dynamic allocation of IPS based on Queuing Theory will improve the performance by allocating resources in optimize way.

# Bibliography

[1] Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., and others, "A view of cloud computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.

[2] Ayadi, I., Simoni, N., and Aubonnet, T., "Sla approach for cloud as a service," pp. 966–967, 2013.

[3] Brown, D. J., Suckow, B., and Wang, T., "A survey of intrusion detection systems," *Department of Computer Science, University of California, San Diego*, 2002.

[4] Chauhan, T., Chaudhary, S., Kumar, V., and Bhise, M., "Service level agreement parameter matching in cloud computing," pp. 564–570, 2011.

[5] Chen, Q., Lin, W., Dou, W., and Yu, S., "Cbf: A packet filtering method for ddos attack defense in cloud environment," pp. 427–434, 2011.

[6] Chen, W.-H., Hsu, S.-H., and Shen, H.-P., "Application of svm and ann for intrusion detection," *Computers & Operations Research*, vol. 32, no. 10, pp. 2617–2634, 2005.

[7] Dhanalakshmi, Y. and Babu, I. R., "Intrusion detection using data mining along fuzzy logic and genetic algorithms," *International Journal of Computer Science and Network Security*, vol. 8, no. 2, pp. 27–32, 2008.

[8] DOULIGERIS, C. and MITROKOTSA, A., "Ddos attacks and defense mechanisms: classification and state-of-the-art," *Computer Networks*, vol. 44, no. 5, pp. 643–666, 2004.

[9] DU, P. and NAKAO, A., "Ddos defense as a network service," pp. 894–897, 2010.

[10] DURCEKOVA, V., SCHWARTZ, L., and SHAHMEHRI, N., "Sophisticated denial of service attacks aimed at application layer," in *ELEKTRO, 2012*, pp. 55–60, IEEE, 2012.

[11] FALOUTSOS, M., FALOUTSOS, P., and FALOUTSOS, C., "On power-law relationships of the internet topology," vol. 29, no. 4, pp. 251–262, 1999.

[12] GARFINKEL, T., ROSENBLUM, M., and OTHERS, "A virtual machine introspection based architecture for intrusion detection.," in *NDSS*, vol. 3, pp. 191–206, 2003.

[13] GONG, R. H., ZULKERNINE, M., and ABOLMAESUMI, P., "A software implementation of a genetic algorithm based approach to network intrusion detection," in *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, 2005 and First ACIS International Workshop on Self-Assembling Wireless Networks. SNPD/SAWN 2005. Sixth International Conference on*, pp. 246–253, IEEE, 2005.

[14] GONZALEZ, A. J. and HELVIK, B. E., "System management to comply with sla availability guarantees in cloud computing.," pp. 325–332, 2012.

[15] HAN, J., KAMBER, M., PEI, J., HAND, D. J., MANNILA, H., SMYTH, P., KLOESGEN, W., ZYTKOW, J., THURAISINGHAM, B., TAN, P., and OTHERS, "Eksploracja danych," 2005.

[16] JUNG, S.-M., JUNG, J.-K., KIM, T.-K., and CHUNG, T.-M., "A study on the resource management against availability attacks in cloud computing," pp. 479–485, 2014.

[17] KAR, S. and SAHOO, B., "An anamaly detection system for ddos attack in grid computing," *International Journal of Computer Applications in*, 2009.

[18] KESHTELI, A. H., DASTJERDI, M. S., REZVANIAN, H., AMINORROAYA, A. A., AMINI, M., HASHEMIPOUR, M., and OTHERS, "Zinc status in goitrous school children of semirom, iran," *Journal of Research in Medical Sciences*, vol. 14, no. 3, pp. 165–170, 2009.

[19] KHOLIDY, H. A. and BAIARDI, F., "Cids: A framework for intrusion detection in cloud systems," pp. 379–385, 2012.

[20] LI, J., LIU, Y., and GU, L., "Ddos attack detection based on neural network," in *Aware Computing (ISAC), 2010 2nd International Symposium on*, pp. 196–199, IEEE, 2010.

[21] LI, W., "Using genetic algorithm for network intrusion detection," *Proceedings of the United States Department of Energy Cyber Security Group*, pp. 1–8, 2004.

[22] LONEA, A. M., POPESCU, D. E., and TIANFIELD, H., "Detecting ddos attacks in cloud computing environment," *International Journal of Computers Communications & Control*, vol. 8, no. 1, pp. 70–78, 2012.

[23] LU, W. and TRAORE, I., "Detecting new forms of network intrusion using genetic programming," *Computational Intelligence*, vol. 20, no. 3, pp. 475–494, 2004.

[24] MAZZARIELLO, C., BIFULCO, R., and CANONICO, R., "Integrating a network ids into an open source cloud computing environment," 2010.

[25] MODI, C. N., PATEL, D. R., PATEL, A., and RAJARAJAN, M., "Integrating signature apriori based network intrusion detection system (nids) in cloud computing," *Procedia Technology*, vol. 6, pp. 905–912, 2012.

[26] MORADI, M. and ZULKERNINE, M., "A neural network based system for intrusion detection and classification of attacks," in *Proceedings of the 2004 IEEE international conference on advances in intelligent systems-theory and applications*, 2004.

[27] PATEL, A., TAGHAVI, M., BAKHTIYARI, K., and JÚNIOR, J. C., "An intrusion detection and prevention system in cloud computing: A systematic review," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 25–41, 2013.

[28] PATEL, K. S. and SARJE, A., "Vm provisioning method to improve the profit and sla violation of cloud service providers," pp. 1–5, 2012.

[29] ROSCHKE, S., CHENG, F., and MEINEL, C., "An extensible and virtualization-compatible ids management architecture," in *Information Assurance and Security, 2009. IAS'09. Fifth International Conference on*, vol. 2, pp. 130–134, IEEE, 2009.

[30] ROSCHKE, S., CHENG, F., and MEINEL, C., "Intrusion detection in the cloud," pp. 729–734, 2009.

[31] SAAD, R., NAIT-ABDESSELAM, F., and SERHROUCHNI, A., "A collaborative peer-to-peer architecture to defend against ddos attacks," pp. 427–434, 2008.

[32] SCHNJAKIN, M., ALNEMR, R., and MEINEL, C., "A security and high-availability layer for cloud storage," pp. 449–462, 2011.

[33] SCHNJAKIN, M., ALNEMR, R., and MEINEL, C., "A security and high-availability layer for cloud storage," pp. 449–462, 2011.

[34] SNAPP, S. R., BRENTANO, J., DIAS, G. V., GOAN, T. L., HEBERLEIN, L. T., HO, C.-L., LEVITT, K. N., MUKHERJEE, B., SMAHA, S. E., GRANCE, T., and OTHERS, "Dids (distributed intrusion detection system)-motivation, architecture, and an early prototype," in *Proceedings of the 14th national computer security conference*, vol. 1, pp. 167–176, Citeseer, 1991.

[35] STIAWAN, D., IDRIS, M., ABDULLAH, A. H., and OTHERS, "The prevention threat of behavior-based signature using pitcher flow architecture," *IJCSNS International Journal of Computer Science and Network Security*, vol. 10, no. 4, pp. 289–294, 2010.

[36] SWAIN, B. R. and SAHOO, B., "Mitigating ddos attack and saving computational time using a probabilistic approach and hcf method," in *Advance Computing Conference, 2009. IACC 2009. IEEE International*, pp. 1170–1172, IEEE, 2009.

[37] UNDHEIM, A., CHILWAN, A., and HEEGAARD, P., "Differentiated availability in cloud computing slas," pp. 129–136, 2011.

[38] VIEIRA, L. E. A. and DA SILVA, L. A., "Geomagnetic modulation of clouds effects in the southern hemisphere magnetic anomaly through lower atmosphere cosmic ray effects," *Geophysical research letters*, vol. 33, no. 14, 2006.

[39] WONG, C. K. and LAM, S. S., "Digital signatures for flows and multicasts," pp. 198–209, 1998.

[40] YANG, L., ZHANG, T., SONG, J., WANG, J., and CHEN, P., "Defense of ddos attack for cloud computing," vol. 2, pp. 626–629, 2012.

[41] YUAN, R., LI, Z., GUAN, X., and XU, L., "An svm-based machine learning method for accurate internet traffic classification," *Information Systems Frontiers*, vol. 12, no. 2, pp. 149–156, 2010.

[42] ZHANG, Z., XIE, W., ZHOU, S., GAO, S., and GUAN, J., "Anomalous behavior of trapping on a fractal scale-free network," *EPL (Europhysics Letters)*, vol. 88, no. 1, p. 10001, 2009.