ISSN:0975-9646

Secure Cloud Environment Using Hidden Markov Model and Rule Based Generation

Harsha Banafar

M Tech CSE ,
OIST BHOPAL MP

Sanjay Sharma

Asst. Professor Mtech (CSE)
OIST BHOPAL MP, India.

Abstract—**Cloud computing nowadays growing rapidly and is an innovative computing model that delivers services on the Internet for fulfilling computing demand of the users. There are different integrated technologies in Cloud. Each has some loopholes, which raises several security and privacy concerns. One of the major security concerns in Cloud computing is to protect against network intrusions that affect confidentiality, availability and integrity Cloud resources and offered services.**

Hidden Markov model is a state transition based model in which the transition of states can be predicted using the probability distribution of one state to another. Here in this paper this model can be used for the intrusion detection and prevention. But the detection and prevention of intrusion is done at the cloud environment where the security issues and privacy are main concerns. The proposed methodology implemented here is efficient as compared to the other existing technique for intrusion detection. This model also improve true positive rate and reduce false positive rate of intrusion detection systems.

Keywords— Cloud Computing, Challenges in Cloud Computing, Hidden Markov Model

I. Introduction

All data storing in the cloud has become a trend. Due to an increasing number of clients store their important data in remote servers in the cloud, without leaving a copy in their local computers. To greater extent data owners start choosing to host their data in the Cloud [4]. Cloud Computing has been thinks about as the next-generation structural design of IT Enterprise (IT) design for enterprises, due to its long list of unprecedented improvements in the IT record on past: ever-present network access, locality of network independent resource pooling, quick resource elasticity, on-demand identity-service, convention-based pricing and transference of risk [2]. Mainly it moves towards the application software and databases to the centralized huge data centers. A cloud storage system, consisting of all collected works on storage servers, offers long-term storage services over the Internet. Storing data in a third party's cloud method origins serious anxiety over data secrecy. Due to sometimes the data stored in the cloud is so essential that the users must ensure it is not lost or fraudulent. At the same time as it is simple to check data integrity after comprehensively downloading the data to be checked, downloading huge amounts of data just for ensuring data integrity is a dissipate of communication bandwidth. For this reason, a lot of works have been done on intending remote data integrity ensuring protocols, which allow data integrity to be checked not including entirely downloading the data. Consequently, Storage-as-a-Service recommended by cloud service providers (CSPs) become known as a result to take the edge off the load of huge restricted data storage and shrink the maintenance cost by means of outsourcing data storage [1],[3].

As Cloud services are delivered through the Internet; security and privacy of Cloud resources and offered services are the biggest concerns. International Data Corporation (IDC) survey showed that security is the greatest challenge of Cloud L. Martin Cyber Security division shows that the major security concern after data security is intrusion detection and prevention in cloud. At network layer, Cloud suffers from traditional attacks such as IP spoofing, Address Resolution Protocol (ARP) spoofing, Routing Information Protocol (RIP) attack, DNS poisoning, man-in-the-middle attack, port scanning, Insider attack, Denial of Service (DoS), Distributed Denial of Service (DDoS) etc. E.g., an internal DoS attack demonstrated against the Amazon Elastic Compute Cloud (EC2). DoS attack on the underlying Amazon Cloud caused BitBucket.org, a site hosted on Amazon Web Services (AWS) to remain unavailable for few hours. These attacks affect the confidentiality, integrity and availability of Cloud resources and offered services. To address such issues, major Cloud providers (like Amazon, Window Azure, Rack Space, Eucalyptus, Open Nebula etc.) use the firewall. Firewall protects the front access points of system and is treated as the first line of defense. As firewall sniffs the network packets only at the boundary of a network, insider attacks cannot be detected. Few DoS or DDoS attacks are too complex to detect using traditional firewall. E.g., if there is an attack on port 80 (HTTP server) or port 25 (Mail server), firewall cannot differentiate normal traffic from attack traffic. Also, ThreatMetrix, a company providing Cybercrime Defender Platform stated that "You Can't Fight the Fire from Behind the Firewall" in [www.threatmetrix.com/docs/Whitepaper- ybercrime-Defender.pdf]. Therefore, use of only traditional firewall to block all the intrusions is not an efficient solution. Another solution is to integrate network based intrusion detection system (NIDS) in Cloud computing. NIDS performs the role of alert system and adds the next preventive layer of security by detecting network attacks that penetrate our system. There are two types of techniques used in NIDS. One is signature based detection that can detect known attacks efficiently and another is anomaly detection that determines whether a given behavior is malicious or not.

The efficiency of NIDS depends on parameters like used detection technique (signature based or anomaly based), it's positioning within network (front end or back end), its configuration (centralized or distributed).

Types of Cloud Services:

In this Software as a Service (SaaS) Service Users are make available to right to use the software applications and Databases and this is also called as On Demand Software services. The cloud client requires paying to use the cloud software applications. In Infrastructure as a Service (IaaS) the Cloud Service supplier materials the resources on require basis from their Data centers. The resources are Software packages, unrefined, Virtual local area networks, load stability's, firewalls, IP addresses, Virtual machine disk image library, file based storage spaces. In Platform as a Service (PaaS) of the Cloud Service supplier a computing stage to the course developers. Computing stage comprises database, and web server, operating system, programming language execution environment. And in Network as a Service (NaaS) the Cloud Service provider make available network/transport connectivity services and/or inter-cloud network connectivity services to the clients.

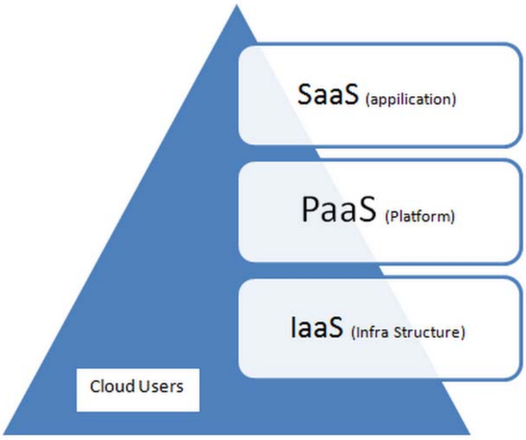


Fig. 1.1 Cloud Architecture

***Cloud Computing Models*:**

Private Cloud: Mainly this communications have possession of a particular group.

***Public Cloud*:** In this cloud they are open for publically free, that accessible in public networks. It may need various cloud structure are required and provide cloud security deliberations.

Group of people Cloud: It may belong to quite a lot of groups or hybrid; they will distribute among this kind of infrastructure. They will deal with inside or by a third party.

Cross Cloud: Combination of two or more Private, Public and group of people clouds.

In private cloud, cloud is possessed by private group and they keep up their own auditing standards and development. More often than not private cloud will not join with the public networks like internet, so the possibilities of external attacks is very small .The private cloud client are edged and easily can do observing. The public cloud join with public network like internet and there is unconstrained number of client can join with public cloud depends upon the service providers competence. The contract SLA between cloud service provider and cloud client is not visible to all Cloud clients, so there is a possibility of concurrence infringement between Cloud Service Provider (CSP) cloud clients.

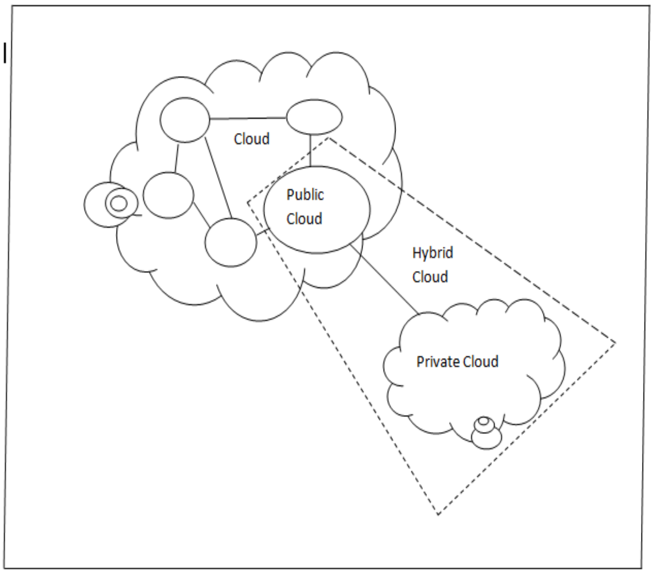


Fig. 1.2 Cloud Computing Types

Characteristics of Cloud Computing

Wherever Cloud computing is the stipulation of enthusiastically scalable and frequently virtualized stores as a services over the internet clients need not have acquaintance of, proficiency in, or organize over the skill communications in the "cloud" that sustains them. Cloud computing signifies a major transform in how we accumulate information and run applications. As an alternative of hosting apps and data on an each entity of desktop computer, the whole thing is hosted in the "cloud"—a grouping of computers and servers right of entry passing through the Internet.

Cloud computing demonstrates the following key characteristics:

1. ***Alertness****:* It progress with clients' talent to prerequisite technological infrastructure resources.
2. ***Multi occupancy:*** It facilitates distribution of resources and costs transversely a big pool of clients consequently allowing for all clients to use those resources.
3. ***Consumption and Competence:*** Generally it enhancements for structures that are frequently only 10–20% make use of that given resources.
4. ***Consistency:*** It is getting better if multiple unnecessary sites are used, which makes well-planned cloud computing appropriate for business stability and failure revival.
5. ***Performance***: is observed and dependable and insecurely combined structural designs are assembled using web services as the system interface.

6. ***Security***: could get better due to centralization of data, enhanced security-focused resources, etc., but apprehensions can continue about failure of manage over assured susceptible data, and be deficient in security for accumulated kernels. Security is frequently as good as or improved than other conventional systems; in some measure because suppliers are able to allocate resources to explaining security concerns that many consumers cannot have enough money. On the other hand, the complication of security is to a huge extent enhanced when data is allocated over a wider area or greater number of devices and in multi-occupant methods that are being shared by unconnected clients. As well, client right of entry to safety measures inspection logs may be complicated or not viable. Private clouds installations are in part inspired by clients' aspiration to keep hold of manage over the infrastructure and keep away from bringing up the rear organize of information security.
7. ***Protection***: Protection of cloud computing applications is easier, because they do not require to be installed on each client's computer and can be right of entries from different positions, etc.

CLOUD COMPUTING CHALLENGES

Cloud computing implementation is facing several challenges in different aspect and According to the survey conducted in 2008 by IDC survey about cloud computing challenges, cloud computing security challenges is on the top most threat to cloud, and became well known that cloud computing paradigm is a kind of virtualization environment with another technology such as grade and clusters and distributed computing, all these technologies has his own security disadvantages. In addition to the threats coming from cloud component, therefore securing issues is not related to cloud just but also related to other technologies and most dangers threat to cloud is vitalization security. And one of the potential attacks to cloud virtualization system is neighbor attacks , which by any virtual machine can attack its neighbor in same physical infrastructures and thus prevent it from providing its services or, which has been known as denial of service attack DoS attack as has been existed in AWS Amazon, that kind of attack can effect on cloud performance in general and can cause financial Losses and can cause harmful effect in other servers in same cloud infrastructure.

Issues in Cloud Data Storage:

Cloud Computing moves the application software and databases to the large data centers, where the management of the data and services may not be fully trustworthy. This unique attribute, however, poses many new security challenges which have not been well understood. In this, we focus on cloud data storage security, which has always been an important aspect of quality of service. To ensure the correctness of users data in the cloud.

A. ***Trust:*** Trust is defined as reliance on the integrity, strength, ability and surety of a person or thing. Entrusting your data on to a third party who is providing cloud services is an issue. Recent incidents like In April of 2012 Amazon's Elastic Compute Cloud service crashed during a system upgrade, knocking customers' websites off-line for anywhere from several hours to several days. That same month, hackers broke into the Sony PlayStation Network, exposing the personal information of 77 million people around the world. And in June a software glitch at cloud-storage provider Dropbox temporarily allowed visitors to log in to any of its 25

Million customers' accounts using any password or none at all. These issues have certainly created doubts in mind of cloud consumers and damaged the trust ability of Consumers.

B. Privacy: Different from the traditional computing model, cloud computing utilizes the virtual computing technology, users' personal data may be scattered in various virtual data center rather than stay in the same physical location, even across the national borders, at this time, data privacy protection will face the controversy of different legal systems. On the other hand, users may leak hidden information when they accessing cloud computing services. Attackers can analyze the critical task depend on the computing task submitted by the users.

C. ***Security:*** Cloud service providers employ data storage and transmission encryption, user authentication, and authorization. Many clients worry about the vulnerability of remote data to criminals and hackers. Cloud providers are enormously sensitive to this issue and apply substantial resources to mitigate this problem.

D. ***Ownership***: Once data has been relegated to the cloud, some worry about losing their rights or being unable to protect the rights of their customers. Many cloud providers address this issue with well-skilled user-sided agreements. According to the agreement, users would be wise to seek advice from their favourite legal representative.

E. ***Performance and Availability***: Business organizations are worried about acceptable levels of performance and availability of applications hosted in the cloud.

F. ***Legal:*** There are certain apprehensions for a cloud service provider and a client receiving the service like location of the cloud provider, infrastructure and physical location of the data and outsourcing of the cloud provider's services etc.

G. ***Multiplatform Support:*** More an issue for IT departments using managed services is how the cloudbased service integrates across different platforms and operating systems, e.g. OS X, Windows, Linux and thinclients. Usually, some customized adaption of the service takes care of any problem. Multiplatform support requirements will ease as more user interfaces become web-based.

H. ***Intellectual Property:*** A company invents something new and it uses cloud services as part of the invention. Is the invention still patentable? Or there can be issues like cloud service provider can make claim for that invention or leak the information to the competitor.

I. ***Data Backup:*** Cloud providers employ redundant servers and routine data backup processes, but some people worry about being able to control their own backups. Many providers are now offering data dumps onto media or allowing users to back up data through regular downloads.

J. ***Data Portability and Conversion*****:** Some people have concerns like, switching service providers; there may be difficulty in transferring data. Porting and converting data is highly dependent on the nature of the cloud provider's data retrieval format, particular in cases where the format cannot be easily revealed. As service competition grows and open standards become established, the data portability issue will ease, and conversion processes will become available supporting the more popular cloud providers. Worst case, a cloud subscriber will have to pay for some custom data conversion.

These are certain areas in which cloud computing requires to excel and solve problem related to it. Out of all the

Problems Security, Privacy and Intellectual property put the major threats on growth of cloud computing that are needed to be worked upon.

AttacksHandle by our System*:* As we know Cloud computing services needs internet connection. That's why we choose attacks for intrusion detection and prevention which is most harmful as my concerns. These attacks are:

Denial of serice **:** Denial of service attack poses as one of the most networks famous attacks, by which one victim machine can receive more than its capacity and so other end users requests cannot be served by server, and in the cloud environment, that kind of attack can be most harmful than unclouded environment because of VMs neighboring and resource sharing in cloud computing environment, so one virtual machine can be used as a source of denial of service attack to another virtual machine in same infrastructure and for overcome this security threat , several kind of flooding DoS attacks detecting and prevention approach has been suggested, and every one propose a method to detect or prevent this attack .

SYN Flooding: A SN Flood is a denial of service attack to which every TCP/IP implementation is vulnerable. Each half-open TCP connection made to a machine causes the 'tcpd' server to add a record to the data structure that stores information describing all pending connections. This data structure is of finite size, and it can be made to overflow by intentionally creating too many partially-open connections. The half-open connections data structure on the victim server system will eventually fill and the system will be unable to accept any new incoming connections until the table is emptied out. Normally there is a timeout associated with a pending connection, so the half-open connections will eventually expire and the victim server system will recover. However, the attacking system can simply continue sending IP-spoofed packets requesting new connections faster than the victim system can expire the pending connections. In some cases, the system may exhaust memory, crash, or be rendered otherwise inoperative.

Smurfing: Another common way of bringing down a host is known as smurfing. This exploits the Internet Control Message Protocol (ICMP), which enables users to send an echo packet to a remote host to check whether it's alive. The problem arises with broadcast addresses that are shared by a number of hosts. Some implementations of the Internet protocols respond to pings to both the broadcast address and their local address (the idea was to test a LAN to see what's alive). So the protocol allowed both sorts of behavior in routers. A collection of hosts at a broadcast address that responds in this way is called a smurf amplifier.

The attack is to construct a packet with the source address forged to be that of the victim, and send it to a number of smurf amplifiers. The machines there will each respond (if alive) by sending a packet to the target and this can swamp the target with more packets than it can cope with. Smurfing is typically used by someone who wants take over an Internet relay chat (IRC) server, so they can assume control of the chatroom. Innovation was to automatically harness a large number of "innocent" machines on the network to attack the victim.

Ftp_write ***Attack:*** The Ftp-write attack is a Remote to Local User attack that takes advantage of a common anonymous ftp misconfiguration. The anonymous ftp root directory and its subdirectories should not be owned by the ftp account or be in the same group as the ftp account. If any of these directories are owned by ftp or are in the same group as the ftp account and are not write protected, an intruder will be able to add files (such as an rhosts file) and eventually gain local access to the system .

Attack Signature An intrusion detection system can monitor for this attack by watching all anonymous ftp sessions and assuring that no files are created in the ftp root directory.

Anomaly Detection: Anomaly (or behavioral) detection is concerned with identifying events that appear to be anomalous with respect to normal system behavior. A wide variety of techniques including data mining, statistical modeling and hidden markov models have been explored as different ways to approach the anomaly detection problem. Anomaly based approach involves the collection of data relating to the behavior of legitimate users over a period of time, and then apply statistical tests to the observed behavior, which determines whether that behavior is legitimate or not. It has the advantage of detecting attacks which have not been found previously. The key element for using this approach efficiently is to generate rules in such a way that it can lower The false alarm rate for unknown as well as known attacks provided anomaly based solution to prevent intrusion in real time system, which analyzes protocol based attack and multidimensional traffic.

Intrusion Detection System

An intrusion-detection system (IDS) can be distinct as tools, solution, and resources used to help identify, assess, and to claim unconstitutional or unapproved network action. Intrusion detection is characteristically one part of an overall fortification system that is installed around a system or device it is not a stand-alone protection measure. An intrusion detection system (IDS) is an indispensable part in a good network security background. It enables detection of suspicious packets and attacks. With the help of IDs, all network traffic can be observed.

Conventional intrusion detection systems (IDS) in wired networks analyse the behaviour of the elements in the network trying to identify anomalies produced by intruders and, once identified, start a response against the intruders. These detection systems are usually placed in those elements with more confluent traffic such as routers, gateways, and switches. Unfortunately, in ad-hoc networks, those elements are not uses, and it is not possible to guess which nodes will route more traffic from its neighbours and install IDS systems only in those nodes.

Intrusion detection techniques are traditionally categorized into two methodologies: anomaly detection and misuse detection. Anomaly based intrusion detection systems base their decisions on anomalies, belongings that do not usually occur. Misuse detection seizes intrusions in terms of the characteristics of known attacks or system vulnerabilities; any action that conforms to the pattern of a known attack or vulnerability is considered intrusive.

The components of Intrusion Detection System can be of three types.

Network intrusion detection: Network intrusion detection systems listen to network communications. They are acquainted with intrusions which come during the networking environment. Essentially a network intrusion detection system (NIDS) is a service which listens on a network interface looking for suspicious traffic. Network intrusion detection systems are mostly signature based.

Host Based Intrusion Detection: Host intrusion detection systems (HIDS) inhabit on a resource supervised. This resource is mostly a computer server or workstation. HIDS seem at produced log files, changes in the file system or check for changes in the process table. Their objective is to identify intrusions into a host.

Signature based intrusion detection**:** Signature based intrusion is based on signatures of known attacks. These signatures are accumulated and evaluated against events or received traffic. If a pattern matches, an alert is generated.

HIDDEN MARKOV MODELS

A Hidden Markov Model (HMM) is a double stochastic model. The model is denoted by (A, B, ┐), where A is the set of observables, B is the set of hidden states, and ┐ is the set of transition probabilities, i.e., the probabilities from going to one hidden state to another. This model is known as double stochastic since there is a hidden layer that contains some hidden states. This hidden layer follows the principles of Markov process. The other layer contains the states of the observables in a particular time t of the model construction. This is also a Markov process where the observable outputs can be seen, unlike the hidden layer.

The HMM algorithm works in two steps. The HMM is trained in the first step using the training sequences. At the initial state (at time t0), the state transition probabilities and the observable output probabilities are randomly assigned. However, assigning these probabilities according to prior knowledge of the system, instead of the random assignment, can improve the performance of HMM. At this point, the model is denoted with 0. Then, applying the Baum-Welch algorithm, the HMM 0 is adjusted according to the input training sequences and construct the new model.After every adjustment of , the probability difference of the previous model and the adjusted model is calculated. If the difference is below the preset probability difference threshold, the model is known to be the final HMM. Otherwise, further adjustment is required. In the next step, the unknown sequences are applied to the model and the likelihood of the sequences (i. e., the probability of how much a sequence conforms the HMM) are determined. If the probability is above the predefined acceptable probability, the sequence is concluded as a non-anomalous sequence. Otherwise, it is concluded as an anomalous one. The HMM algorithm has very accurate prediction of anomaly and has been used for complex sequence analysis. However, the model training time is very high in HMM algorithm.

II. LITERATURE REVIEW

Krunal Patel [2]- One of the major security concerns in Cloud computing is to protect against network intrusions that affect confidentiality, availability and integrity Cloud resources and offered services. To address this issue, they would design hybrid network intrusion detection system (NIDS) in Cloud.NIDS consists of snort and classifiers , which aims to detect network intrusions in Cloud computing environment with low false alerts and affordable computational cost. To ensure feasibility of their NIDS module in Cloud, they will evaluate performance and quality results of NIDS on the KDD'99 and the NSL-KDD experimental datasets. For that here they surveyed different Threats to cloud computing and different IDS/IPS techniques.

Chirag Modi et.al [5] they propose a framework which integrates a network intrusion detection system (NIDS) in the Cloud infrastructure. They use snort and decision tree (DT) classifier to implement this framework. It aims to detect network attacks in Cloud by monitoring network traffic, while maintaining performance and service quality. To validate our approach, they evaluate the performance and detection efficiency by using the freely available NSL-KDD and KDD experimental intrusion datasets. The results show that the proposed framework has a higher detection rate with low false positives at an affordable computational cost.

Cloud computing provides a new computing services but the main shortage is security of cloud services used, especially in cloud infrastructure aspects. In cloud infrastructure service, several users can share the same infrastructure, which can cause DoS attack. In this paper a novel model [6] to detect flooding based DoS attack in cloud environment has been proposed and consisting three phases has been proposed a covariance matrix move towards to distinguish this kind of attack. The given phrases are: (1) The first-phase is to model the usual traffic prototype for baseline profiling and (2) the second phase is the intrusion detection processes and (3) finally anticipation phase.

According the initial experiments conducted, it can be concluded that this approach can detect this kind of attack it is clear that the two experiments which implemented on normal traffic are similar in the flags average regardless of the amount of traffic and in other hand, the average of flags in abnormal traffic is totally different from normal traffic as in third experiment. Therefore the covariance matrix approach can be effectively used to detect abnormal traffic (DoS attack). From the give experimental result, shows that detecting the flooding attack efficiently used in cloud computing environment.

Emil Kuriakose John and Sumaiya Thaseen suggested An Efficient Defense System for IP Spoofing in Networks [7]. They propose a promising solution for the detection of malicious IP packets under real time environment. IP Spoofing happens in various ways through Distributed Denial of Service (DDoS) attack, Replay attack and Blind Spoofing. These are the main techniques by which spoofing attempts occur in the Internet. The application can be used for spoofing any type of malicious packets with the key advantage of consuming fewer resources. It is an integrated tool comprising of packet analyzer, active ports and machines on LAN. The system is resistant to Distributed Denial of service attack (DDoS), Replay attack and blind spoofing [7].

In this paper [8], we explored the problem of providing simultaneous public auditability and data dynamics for remote data integrity check in Cloud Computing. Their construction is deliberately designed to meet these two important goals while efficiency being kept closely in mind. To accomplish efficient data dynamics, here they get better the offered confirmation of storage models by influencing the classic Merkle Hash Tree (MHT) building for block tag confirmation. To sustain efficient management of multiple auditing assignments, then they extra investigate the method of bilinear aggregate signature to expand most important outcome into a multi-user setting, wherever TPA can execute multiple auditing tasks at the same time. Widespread security and concert investigation of their proposed manufacture and give explanation for the routine of their method demonstrate that the proposed scheme to sustain scalable and efficient public auditing in Cloud Computing is extremely efficient and provably make safe.

Mahsa Khosronejad, Elham Sharififar, Hasan Ahmadi Torshizi and Mehrdad Jalali [12] – In today's communication system computer and information security are a major concern as these are vulnerable to potential attackers. Because to increase the potential of advanced computer communication and distributed systems leads to attack on the data flow over the network which affects integrity and availability of information greatly. Therefore, the security of Web applications is a key topic in computer security. This paper presents two hybrid approaches for modeling IDS. C5.0 and HMM are combined as a hierarchical hybrid intelligent system model (C5.0-HMM). Empirical results with KDD Cup 99 Intrusion data illustrate that the proposed hybrid systems provide more accurate intrusion detection systems.

III. PROBLEM STATEMENT

The techniques implemented for the intrusion detection system is particularly based on the rules that are generated on the basis of which True and false alarm ration are predicted. But the techniques implemented so far for the intrusion detection are not efficient in terms of detection rate and alarm generation rate .As well as the efficient techniques foe intrusion Detection and Prevention are available for network. We design technique for Cloud Computing Environment.

IV. PROPOSED METHODOLOGY

Here in the code we are creating cloud environment. As soon as the cloud environment establishes the users of the cloud, brokers and data centers are initialized. Now the process of data transmission starts from the user to data center, but in the cloud environment broker decides from which user the data is stored to which datacenter. Hence the proposed work we have maintained IDS to each of the brokers in the cloud. The simulation of how the rules are created and how these rules are used for the detection of intrusions in the packets from the user to data centers. First of all brokers creates rules for each of the dataset such as IDS or My Dataset. After these rules are established the number of nodes in the cloud environment is established i.e. based on users and datacenters. Now HMM is applied here on these users and datacenters. As soon as the packet sends from user to data center a transition state is observed according to HMM and a probability is calculated.

Each and after transition of packets a probability is calculated and if the probability is generated than threshold value, hence from the packets rules are generated and if these rules are matched with the trained rules, then the intrusion is detected.

Distinct Feature: My Thesis work, works like a protocol in which define some rule and according to this rules we can do necessary function. here it works to find Attacks in selected file, it define type of attack in particular file , apply the prevention on particular file as well as generate graphs which shows the performance of algorithm.

This protocol can be applied on NIDS, HIDS according to our requirement in cloud environment. The rules defined in this system are for Behaviour based Intrusion Detection and Prevention System.

Model Diagram:

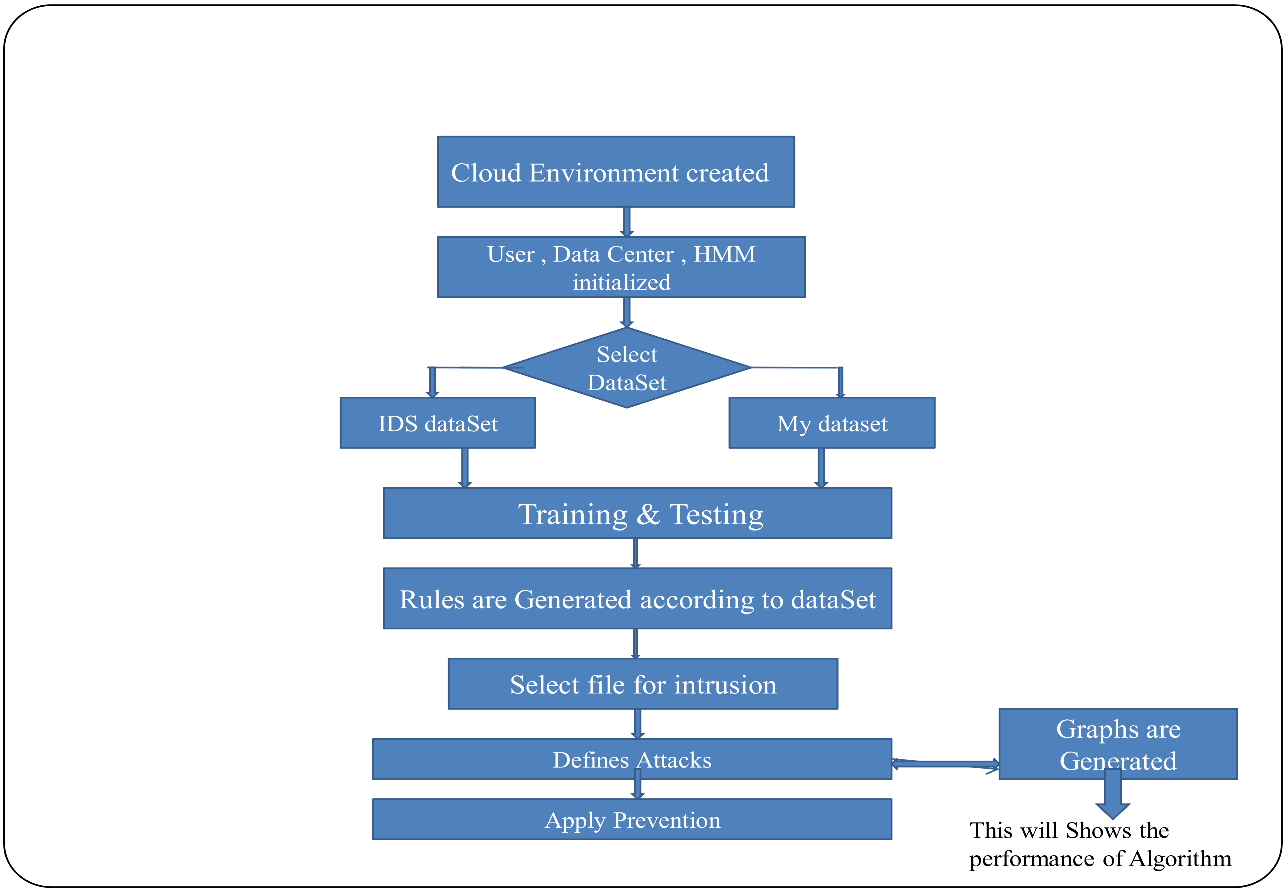

Cloud Environment created
User , Data Center , HMM initialized
Select DataSet
IDS dataSet
My dataset
Training & Testing
Rules are Generated according to dataSet
Select file for intrusion
Defines Attacks
Graphs are Generated
Apply Prevention
This will Shows the performance of Algorithm

Fig. 1.3 main working diagram of proposed system

4.1 *Main Steps of working method*

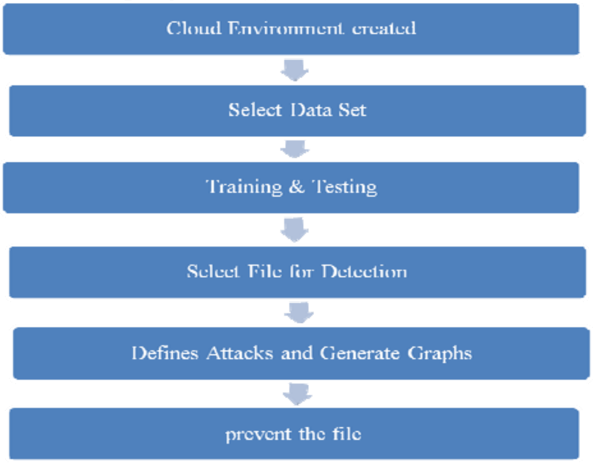

Cloud Environment created
Select Data Set
Training & Testing
Select File for Detection
Defines Attacks and Generate Graphs
prevent the file

Fig 1.4 Steps in Methodology

METHODOLOGY

1. Create cloud simulation environment consisting of dynamic number of users and datacenters.
2. As soon as the cloud is established the sender can send packets to the datacenter.
3. Now initialize the HMM with certain parameters at server or broker level of the cloud.
4. The number of states in the model will depends on the users in the cloud.
5. As soon as the packets start sending from user to datacenter, HMM starts calculating the probability of each of the packet in the transition state.
6. If the probability of packet exceeds threshold value then a rule system is developed based on the attribute values of the dataset IDS.
7. The rules are then learned and can be prevented by changing the rules of the test dataset.

4.2 ALGORITHM

4.2.1 DETECTION OF INTRUSION

A Hidden Markov Model contains five tupples:
N – Is the number of states in the model Q {Q1, Q2, Q3........}.
M – Is the number of observation symbols V {V1, V2, V3…..}.
A – State transition Probabilities
B – is the distribution of each of the states
Π – Is the initial state distribution

1. The initial transition probability from one state Q1 to another state Q2 at a particular instance of time t+1 depends on the state at time t according to the assumption of markov i.e.

$$\mathbf{a_{ij} = p(q_{t+1} = s_j \mid q_t = s_i)}$$

2. The probabilities of the transition of the states is independent of the actual time where the transition takes place according to the assumption of stationarity i.e.

$$p(q_{t+1} = s_j \mid q_{t1} = s_i) = p(q_{t2+1} = s_j \mid q_{t2} = s_i)$$

3. Lets 'n' is the number of packets 'pkt' send at a particular transition at a particular instance of time.
4. Calculate each step of the transition the state which is most probable $q_t, 1 \le t \le T$ for the observation $z_t, 1 \le t \le T$, probability of state transition δt can be computed using viterbi algorithm.
5. After each step of the transition calculate the general probability of the packet to be transmitted at each step Q.
6. The average probability can be computed using

$$\delta_{avg}=(\sum_{k=1}^{T}\delta_k^{(i)})/_{T}$$

7. The condition is checked i.e. if the average probability is less than the threshold value then the intrusion is detected in the packet.

$$\delta_{avg} < \text{(initial threshold value)}$$

8. Let use for example the dataset contains a number of attributes such as source, destination, port, server error rate and time then on the basis of the values of the dataset values certain rules are developed and these rules are used for the detection of intrusion.

4.4.2 PREVENTION OF INTRUSION

The prevention of the intrusion in the cloud environment can be prevented by increasing average threshold of the transition states of the packets of the packets 'pkt'.

1. The initial transition probability from one state Q1 to another state Q2 at a particular instance of time t+1 depends on the state at time t according to the assumption of markov i.e.

$$a_{ij} = p(q_{t+1} = s_j \mid q_t = s_i)$$

2. The probabilities of the transition of the states is independent of the actual time where the transition takes place according to the assumption of stationarity i.e.

$$p(q_{t+1} = s_j \mid q_{t1} = s_i) = p(q_{t2+1} = s_j \mid q_{t2} = s_i)$$

3. Lets 'n' is the number of packets 'pkt' send at a particular transition at a particular instance of time.
4. Calculate each step of the transition the state which is most probable $q_t, 1 \le t \le T$ for the observation $z_t, 1 \le t \le T$, probability of state transition δt can be computed using viterbi algorithm.
5. After each step of the transition calculate the general probability of the packet to be transmitted at each step Q.
6. The average probability can be computed using

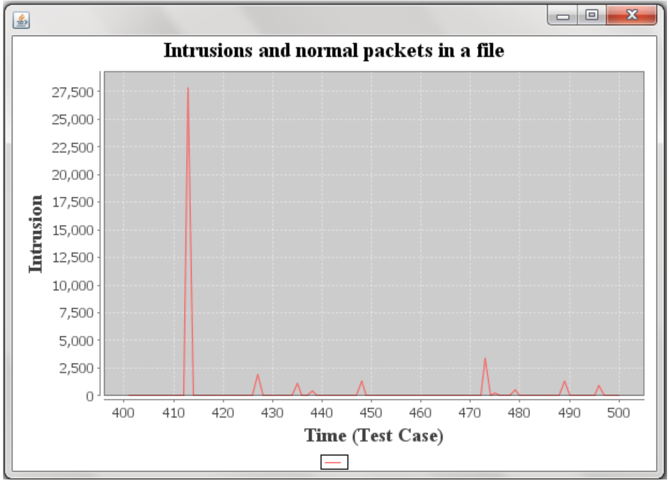
$$\delta_{avg}=(\sum_{k=1}^{T}\delta_k^{(i)})/_{T}$$

$$\delta_{avg}=(\sum_{k=1}^{T}\delta_k^{(i)})/_{T}+1$$

7. The condition is checked at each node of the cloud means the next state and the previous state of the transition i.e. if the average probability is less than the next average threshold value then the packets is transferred from the next other transition state.
8. The prevention is simply done by checking the rule set of the data and changing the abnormal behavior characteristics to the normal behavior.

V. RESULT ANALYSIS

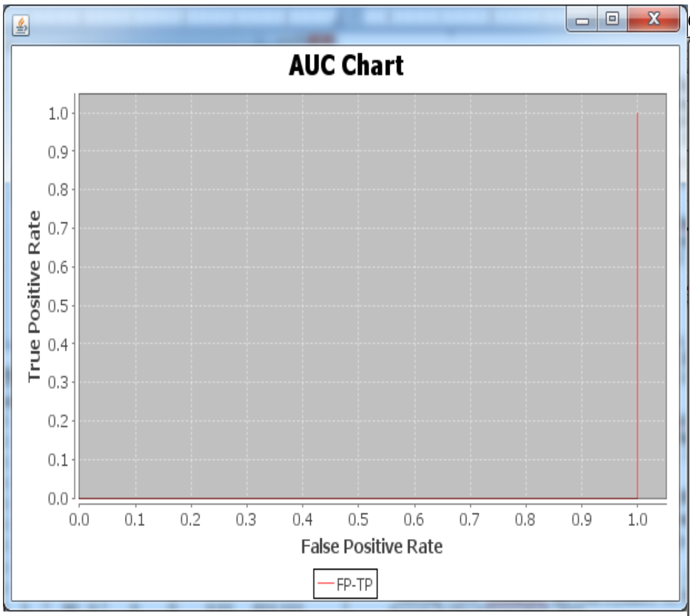
***5.1 Score Chart*:** This chart shows that how much intrusion system will detect with increase in time as packets is scans. That means it shows that if we have 500 packets in a file and we know intrusion is in 200 packets. It shows number of packet contain intrusion attack.



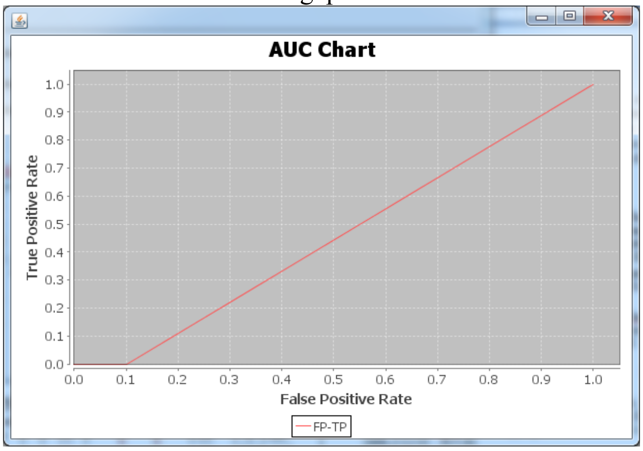
***5.2 AUC Chart*:** This chart shows that how accurate system will detect intrusion . That means it shows the false positive and true positive rate. As well as it shows the Accuracy of Proposed Algorithm.

AUC Chart Based on False positive

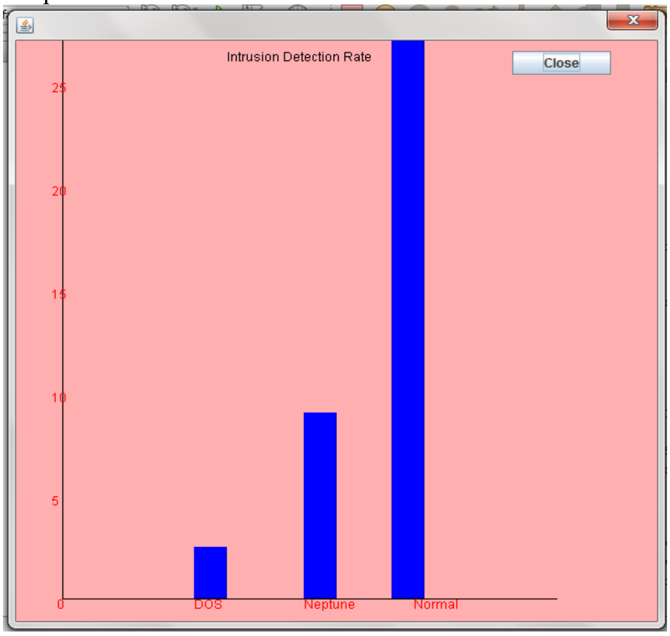
TP	TN	FP	FN	FP-Rate	TP-Rate	AUC	Accuracy	Threshold
0	9	0	91	0	0	0	0.09	27870.667
0	9	0	91	0	0	0	0.09	27869.667
0	9	0	91	0	0	0	0.09	27868.667
0	9	0	91	0	0	0	0.09	27867.667
0	9	1	90	0.1	0	0	0.09	27866.667
0	8	2	90	0.2	0	0	0.08	3333.333
0	7	3	90	0.3	0	0	0.07	1866.667
0	5	5	90	0.5	0	0	0.05	1333.333
0	4	6	90	0.6	0	0	0.04	1066.667
0	3	7	90	0.7	0	0	0.03	933.333
0	2	8	90	0.8	0	0	0.02	533.333
0	1	9	90	0.9	0	0	0.01	400
0	0	10	90	1	0	0	0	266.667
90	0	10	0	1	1	0	0.9	0



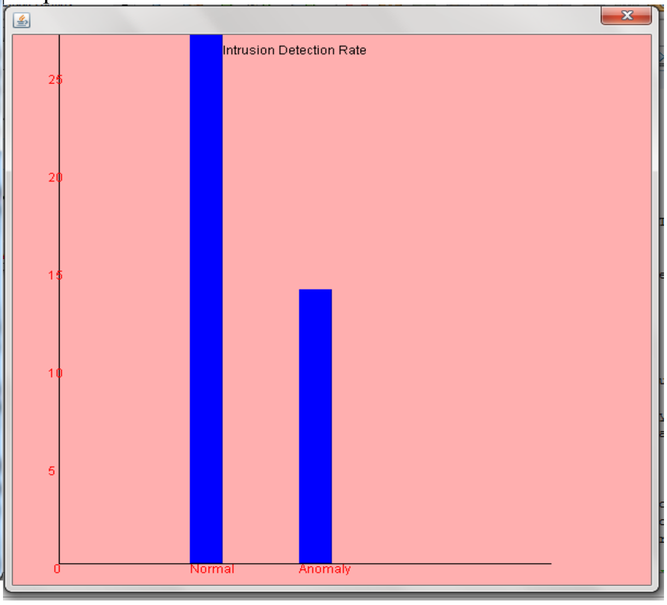
AUC Chart Based on Throughput



Graph1:



Graph2:



VI. EXPERIMENTAL RESULTS

TABLE 1 : COMPAIRISON OF PROPOSED METHOD WITH EXISTING TECHNIQUE

No. of Packets	Detection Ratio Existing	Detection Ratio Proposed
10	87%	98%
20	89%	98%
50	92%	99%
100	93%	99%
200	95%	99%
500	96%	99%

VII. CONCLUSION

Our research indicates that that Security and Privacy are the major issues that are needed to be countered, efforts are being made to develop many efficient System That can Provide Security and privacy at the user level and maintain the trust and intellectual property rights of the user. We are currently developing a secure computation platform based on a simple Hidden Markov Model. Cloud computing is currently the latest trend when it comes to online computing, it may help the enterprise and the end user by providing their needs, but the provider has to make sure that they are valuable and customer data is safe.

The experimental results show the performance of the proposed methodology. Here the framework for detection and prevention of intrusions is implemented over the cloud environment, such that data storage over cloud can be made secure. Intrusion Detection using HMM is used here for the detection and prevention.

VIII. FUTURE

Although the technique implemented here for the detection and prevention of intrusion in the cloud environment using Hidden Markov Model is efficient in terms of detection and prevention rate and also the cost effective, but further enhancements can done in terms of bandwidth, power consumption.

We can also add some more attacks intrusion detection and prevention methods according to need of system requirement.

ACKNOWLEDGMENT

The making of the paper needed co-operation and guidance of a number of people. I therefore consider it my prime duty to thank all those who had helped us for making it successful. It is our immense pleasure to express our gratitude to Professor Mr. Sanjay Sharma (head of department- Computer Science and Engineering, OIST) as a guide who provided us constructive and positive feedback during the preparation of this paper.

REFERENCES

[1] Ayad Barsoum, Anwar Hasan, Ontario, Canada"Enabling Dynamic Data and Indirect Mutual Trust for Cloud Computing Storage Systems" Digital Object Indentifier 0.1109/TPDS.2012.337 1045-9219/12/$31.00 © 2012 IEEE.

[2] KRUNAL PATEL "SECURITY SURVEY FOR CLOUD COMPUTING: THREATS & EXISTING IDS/IPS TECHNIQUES" International Conference on Control, Communication and Computer Technology, 24th, March 2013, Chandigarh,

[3] P. Mell and T. Grance, "Draft NIST Working Definition of Cloud Computing," http://csrc.nist.gov/groups/SNS/cloud-computing/index.html, June 2009.

[4] F. Seb´e, J. Domingo-Ferrer, A. Martinez-Balleste, Y. Deswarte, and J.-J. Quisquater, "Efficient remote data possession checking in critical information infrastructures," IEEE Trans. on Knowl. And Data Eng., vol. 20, no. 8, 2008.

[5] Amazon elastic compute cloud (Amazon EC2), http://aws.amazon.com/ec2/.

[6] Chirag Modi, Dhiren Patel, Bhavesh Borisanya, "A Novel Framework for Intrusion Detection in Cloud" SIN'12, October 25-27, 2012, Jaipur, India. Copyright 2012 ACM 978-1-4503-1668-2/12/10 ...$15.00.

[7] Mohd Nazri Ismail, Abdulaziz Aborujilah, Shahrulniza Musa, AAmir Shahzad, "Detecting Flooding based DoS Attack in Cloud Computing Environment using Covariance Matrix Approach" 2013.

[8] Emil Kuriakose John and Sumaiya Thaseen "Efficient Defense System For IP Spoofing In Networks", computer Science & Information Technology (CS & IT), pp. 185–193, 2012.

[9] Qian Wang, Cong Wang, Kui Ren, Wenjing Lou, Jin Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing", 2010.

[10] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[11] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[12] Mahsa Khosronejad, Elham Sharififar, Hasan Ahmadi Torshizi and Mehrdad Jalali " Developing a hybrid method of Hidden Markov Models and C5.0 as a Intrusion Detection System" International Journal of Database Theory and Application

[13] Afroza Sultana and Abdelwahab Hamou-Lhadj, Mario Couture "An Improved Hidden Markov Model for Anomaly Detection Using Frequent Common Patterns"

[14] Mohd Nazri Ismail, Abdulaziz Aborujilah, Shahrulniza Musa & AAmir Shahzad "New Framework to Detect and Prevent Denial of Service Attack in Cloud Computing Environment" International Journal of Computer Science and Security (IJCSS), Volume (6) :Issue (4) 226