International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307, Volume-3, Issue-2, May 2013

# To Improve Security in Cloud Computing with Intrusion detection system using Neural Network

**Richa Sondhiya, Maneesh Shreevastav, Mahendra Mishra**

***Abstract— Cloud computing is a new type of service which provides large scale computing resource to each customer. Cloud Computing Systems can be easily threatened by various cyber-attacks, because most of Cloud computing system needs to contain some Intrusion Detection Systems (IDS) for protecting each Virtual Machine (VM) against threats. In this case, there exists a tradeoff between the security level of the IDS and the system performance. If the IDS provide stronger security service using more rules or patterns, then it needs much more computing resources in proportion to the strength of security. So the amount of resources allocating for customers decreases. Another problem in Cloud Computing is that, huge amount of logs makes system administrators hard to analyze them. In this paper, we propose a method that enables cloud computing system to achieve both effectiveness of using the system resource and strength of the security service without trade-off between them. In this paper, we propose a soft Computing technique such as MLP Algorithm for detecting the unknown intrusion in network intrusion detection in cloud computing environment.***

***Index Terms— Cloud Computing, Neural Network, Intrusion detection system***

## I. INTRODUCTION

Designing Intrusion Detection System in Cloud Computing Environment is one of the most critical tasks. The development of the technology and hardware, we can imagine that it is possible to get rid of the great mass of the spending for fixed assets, such as expensive networks server and software. Just a web browser can help us complete our common business applications online from the provider on internet. This is the idea the cloud computing providers de-scribe to us. But security in cloud computing environment is of major concern. Intrusion Detection Systems (IDSs) are amongst the main tools for providing security in networks, cloud and grid. Network traffic data, provided for the design of intrusion detection system, which includes rich information about system and user behavior, but the raw data itself, can be difficult to analyze due to its large size. The major issue or key challenges in network security is to being able to reduce data size. Due to this large data set, classification techniques require a huge amount of memory and CPU resource.

To re-duce the number of computer resources, both memory and CPU time, Dimensionality reduction is the key concept.

The feature space having reduced features truly contributes to classification that cuts pre-processing costs and minimizes the effects of the 'peaking phenomenon' in classification. Thereby improving the overall performance of classifier based intrusion detection systems.

## II. CLOUD COMPUTING ENVIRONMENT

The Cloud Computing is one of the emerging technologies in the world. It is an Internet-based computing technology, where shared resources such as software, platform, storage and information are provided to customers on demand. Cloud computing is a technology by which dynamically scalable and virtualized resources are provided to the users over the Inter-net. Cloud computing customers do not own the physical infrastructure, thereby avoiding capital expenditure. They rent resources from a third-party provider, consume them as a ser-vice and pay only for resources that they use [10]. Three primary models for Cloud Computing have emerged:

### *A. Software as a Service (SaaS)*

Applications (word processor, CRM, etc.) or application ser-vices (schedule, calendar, etc.) execute in the ‚cloud' using the interconnectivity of the internet to propagate data [4].

### *B. Platform as a Service (PaaS)*

Applications are built in the ‚cloud' on the platform using a variety of technologies [4].

### *C. Infrastructure as a Service (IaaS)*

Compute resources (processors, memory, storage, bandwidth, etc.) are provided in an as-needed, pay-as-you-go model. IaaS is able to provide from single server up to entire data centers [6].

In this paper, we propose a soft Computing technique such as MLP Algorithm for detecting the unknown intrusion in network intrusion detection in cloud computing environment. Fig1 shows a basic cloud network. Fig shows cloud containing servers which provide services to user. Cloud provide various types of services. It provide storage, infrastructure, platform as a service. Cloud works in pay-as-you-go basis, that means we have to pay according to services which we use. Cloud covers broad category of services.

**Manuscript received May, 2013.**

**Richa Sondhiya**, Department of Information Technology, L.N.C.T. Bhopal, India.

**Maneesh Shreevastav**, Department of Information Technology, L.N.C.T. Bhopal,India.

**Mahendra Mishra**, Department of Information Technology, L.N.C.T. Bhopal, India.

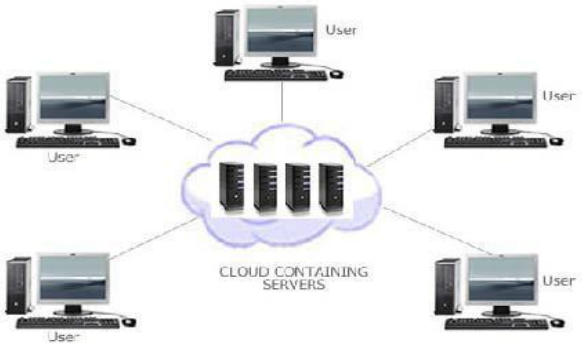



**Fig 1: A Basic Cloud Network**

## III. INTRUSION DETECTION SYSTEM

An Intrusion Detection System (IDS) constantly monitors actions in a certain environment and decides whether they are part of a possible hostile attack or a legitimate use of the environment. The environment may be a computer, several computers connected in a network or the network itself. The IDS analyzes various kinds of information about actions emanating from the environment and evaluates the probability that they are symptoms of intrusions. Such information includes, for example, configuration information about the current state of the system, audit information describing the events that occur in the system (e.g., event log in Windows XP), or network traffic. Several measures for evaluating IDS have been suggested (Debar *et al.* 1999; Richards 1999; Spafford and Zamboni 2000; Balasubramaniyan *et al.* 1998). These measures Include accuracy, completeness, performance, efficiency, fault tolerance, timeliness, and adaptively. The more widely used measures are the True Positive (TP) rate, that is, the percentage of intrusive actions (e.g., error related pages) detected by the system, False Positive (FP) rate which is the percentage of normal actions (e.g., pages viewed by normal users) the system incorrectly identifies as intrusive, and Accuracy which is the percentage of alarms found to represent abnormal behavior out of the total number of alarms. In the current research TP, FP and Accuracy measures were adopted to evaluate the performance of the new methodology. There are IDS that simply monitor and alert and there are IDS that perform an action or actions in response to a detected threat. We'll cover each of these briefly.

## IV. TYPE OF INTRUSION DETECTION SYSTEM

### *A. NIDS*

Network Intrusion Detection Systems are placed at a strategic point or points within the network to monitor traffic to and from all devices on the network. Ideally you would scan all inbound and outbound traffic; however doing so might create a bottleneck that would impair the overall speed of the network.

### *B. HIDS*

Host Intrusion Detection Systems are run on individual hosts or devices on the network. A HIDS monitors the inbound and outbound packets from the device only and will alert the user or administrator of suspicious activity is detected

### *C. Signature Based*

A signature based IDS will monitor packets on the network and compare them against a database of signatures or attributes from known malicious threats. This is similar to the way most antivirus software detects malware. The issue is that there will be a lag between a new threat being discovered in the wild and the signature for detecting that threat being applied to your IDS. During that lag time your IDS would be unable to detect the new threat.

### *D. Anomaly Based*

An IDS which is anomaly based will monitor network traffic and compare it against an established baseline. The baseline will identify what is "normal" for that network- what sort of bandwidth is generally used, what protocols are used, what ports and devices generally connect to each other- and alert the administrator or user when traffic is detected which is anomalous, or significantly different, than the baseline.

## V. LITERATURE SURVEY

Multiple research activities were introduced to address the issue of intrusion detection within cloud computing environments.

*Dastjerdi et. al.* [7] implemented applied agent-based IDS as a security solution for the cloud. The model they proposed was an enhancement of the DIDMA [8]. The system is mainly designed to protect the networks' resources and cannot be customized as a service.

*Bakshi et. al.* [9] proposed another cloud intrusion detection solution. The main concern was to protect the cloud from DDoS attacks. The model uses an installed intrusion detection system on the virtual switch and when a DDoS attack is detected. Despite being reported as effective, the model helps to protect the cloud itself, not the cloud clients who in turn don't have any kind of authority over the intrusion detection system being used.

*Mazzariello et. al.* [10] Another recent and significant contribution to this field is the work of where they proposed a model for detecting DoS attacks against Session Initiation Protocol(SIP). The model is limited to detecting SIP flooding attacks and falls largely with the category of intrusion detection systems designed to protect the cloud itself.

*Lo et. al.* [11] proposed a framework that is mainly designed to create cloud networks that are immune against the Distributed Denial of Service (DDoS) attacks [12]. The utilized IDS implementation was the Open Source Snort IDS and the framework itself is designed as a Distributed Intrusion Detection System (DIDS) [13] [14]. The proposed framework supports the idea of cooperative defense by the IDS sensors in the cloud network. Within the framework, an IDS sensor is deployed in each network region. Any sensor will send out the alert to the other sensors while they are suffering from a severe attack defined in its block table. Each sensor exchanges its alerts and has a judgment criterion to evaluate the trustworthiness of these alerts. After evaluation, the new blocking rule is added into the block table if the alerts are regarded as a new kind of attack.

*Roschke et. al.* [15] have proposed an intrusion detection framework based on the VM-based IDS [16]. In their work, they have developed a general framework for intrusion detection. It consisted of separate IDS sensors for each virtual host. The IDS sensors can be of different vendors. To enable the collection and correlations of alerts from the different IDS implementations, an Event Gatherer was made to work as a medium to standardize the output from the different sensors as well as realize the logical communication. The cloud user can

have access to both the applications and the IDS sensors. The users can access the sensors, configure, modify rule sets, and modify detection thresholds. Additionally, users can review the alerts generated when attacks that target their virtual hosts or services are spotted. The framework also includes the IDS Management module which is responsible for orchestrating the message passing and alert transfer among the different IDS sensors and the main storage unit whether it was a file system, a network database, or a shared folder. This approach of separating the IDS from the protected hosts is of great advantage. But it is criticized for requiring the large consumption of computing resources since every virtual application, platform, or host needs a separate VM-Based IDS.

## VI. TRADITIONAL TECHNIQUE USED IN IDS

### A. *Clustering Techniques*

Cluster analysis [5] is the process of partitioning data objects (records, documents, etc.) into meaningful groups or clusters so that objects within a cluster have similar characteristics but are dissimilar to objects in other clusters. In clustering, unsupervised classification of unlabeled patterns (observations, data items or feature vectors), is performed. In training set there are no predefined category labels are associated with the object. In result of clustering we get compact representation of large data sets by a very small number of cluster centroids, example of this set is collections of visited Web page. There are numbers of applications of clustering like data mining, document retrieval, pattern classification and image segmentation. Because of this clustering when internet users try to retrieve web information documents from internet they can reveal collections of documents belonging to the same topic. As shown by Sequeira and Zaki (2002), clustering is also useful for anomaly detection: With the assumption that all clusters are based on 'normal' data only, the normality of a new object is calculated by its distance from the most similar cluster.

A good clustering method will produce high quality clusters in which similarity is high known as intra-classes and inter-classes where similarity is low. The quality of clustering depends upon both the similarly measure used by the method and it, s implementation. The concept of clustering algorithms is to build a finite number of clusters, each one with its own center, according to a given data set, where each cluster represents a group of similar objects. Each cluster encapsulates a set of data and here the similarities of the surrounded data are their distance to the cluster center.

Generally speaking, clustering techniques can be divided into two categories pair wise clustering and central clustering. The pair wise clustering is also known as similarity-based clustering; it groups similar type of data instances together which are based on a data-pair wise proximity measure. Graph partitioning-type methods are an examples of this category. The central clustering is also known as centroid-based or model-based clustering. Model based clustering is called central or centroid clustering because it represents each cluster by a model, and that model known as its centroid".

The Central clustering algorithms [3] are sometime works more efficiently as compared to similarity-based clustering algorithms. We prefer centroid-based clustering in place of similarity-based clustering. Sometimes we cannot efficiently get a desired number of clusters, e.g., 100 that is set by users. Similarity-based algorithms usually have a complexity of at least $O(N2)$ (for computing the data-pair wise proximity measures), where $N$ is the number of data. There are following types of partitioning.

#### *i. Hard Partitioning*

These kind of methods are based on classical set theory and defines the presence or absence of a data point in a partition subset on strict logic, that is the object either belong to a subset or not.

#### *ii. Soft Partitioning*

A soft clustering algorithm partitions a given data set not an input space. Theoretically speaking, a soft partition not necessarily a fuzzy partition, since the input space can be larger than the dataset.

##### *a. K-means algorithm:*

The K-means clustering is a classical clustering algorithm. After an initial random assignment of example to K clusters, the centres of clusters are computed and the examples are assigned to the clusters with the closest centres. The process is repeated until the cluster centres do not significantly change. Once the cluster assignment is fixed, the mean distance of an example to cluster centres is used as the score. Using the K-means clustering algorithm, different clusters were specified and generated for each output class

***Input***: The number of clusters K and a dataset for intrusion detection

***Output*:** A set of K-clusters that minimizes the squared –error criterion.

Algorithm:

1. Initialize K clusters (randomly select k elements from the data)
2. While cluster structure changes, repeat from 2.
3. Determine the cluster to which source data belongs Use Euclidean distance formula. Add element to cluster with min (Distance (xi, yj)).
4. Calculate the means of the clusters.
5. Change cluster centroids to means obtained sing Step 3.

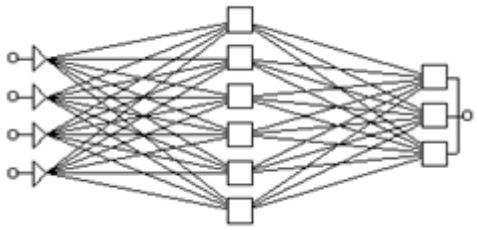
The Main Disadvantage of K-Mean algorithm is that algorithm may take a large number of iterations through dense data sets before it can converge to produce the optimal set of centroids. This can be inefficient on large data sets due to its unbounded convergence of cluster centroids.

## VII. PROPOSED TECHNIQUES

### A. *Multilayer perceptron algorithm*

We present the architecture of a feed-forward neural network that is used to detect the intrusion in cloud computing environment. Multilayer-perceptron algorithm (MLP) [7, 8] is a learning algorithm of Artificial Neural Networks. The Feed-Forward Neural Network architecture can efficiently of approximate large number of problems with very high accuracy. It is also having good generalization ability. The working of Multilayer-perceptron algorithm is based on the error-correction learning rule. Error propagation has two passes which are through the different layers of the network, one of them is known as forward pass, and another known as backward pass. In working of forward pass there is an input

vector which is applied to the sensory nodes of the network and the effect of input vector propagates through the network in one after another layer or in layer by layer fashion. The actual response of the network comes in form of a set of output. The synaptic weights of the networks are all fixed during the forward pass. And during the backward pass the synaptic weights are all adjusted. These weights are adjusted according to error-correction rule. Then to find out error the actual response of the network is subtracted from the desired response and then error signal is produced. This error signal is then propagated backward through the network which is against the direction of synaptic conditions. Fig2 shows the architecture of Multilayer Perceptron.



**Input Layer Hidden Layer Output Layer**

**Fig 2: Multilayer Perceptron**

## VIII. CONCLUSION

Cloud computing is a "network of networks" over the internet, therefore chances of intrusion is more with the erudition of intruders attacks. Different IDS techniques are used to counter malicious attacks in traditional networks. .In this work, we study the possible use of the neural networks learning capabilities to classify the intrusion data set. the proposed approach is to identify the unseen or unknown attach using neural network technique.

## REFERENCES


[1] Sebastian Roschke, Feng Cheng, Christoph Meinel," Intrusion Detection in the Cloud", Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing, 2009.

[2] Chi-Chun Lo, Chun-Chieh Huang, Joy Ku, "A Cooperative Intrusion Detection System Framework for Cloud Computing Networks", 39th International Conference on Parallel Processing Workshops, 2010.

[3] Andreas Haeberlen," An Efficient Intrusion Detection Model Based on Fast Inductive Learning", Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 19-22 August 2007.

[4] Richard Chow, Philippe Golle, Markus Jakobsson, "Controlling Data in the Cloud: Outsourcing Computation without Outsourcing Control", ACM Computer and Communications Security Workshop, CCSW 09, November 13, 2009.J.

[5] Kleber, schulter, "Intrusion Detection for Grid and Cloud Computing", IEEE Journal: IT Professional, 19 July 2010.

[6] Irfan Gul, M. Hussain, "Distributed cloud intrusion detection model", International Journal of Advanced Science and Technology Vol. 34, September, 2011.

[7] http://en.wikipedia.org/wiki/Cloud_computing.

[8] I. Foster, Yong Zhao, I. Raicu, and S. Lu, "Cloud Comuting and Grid Computing 360-Degree Compared," in Grid Computing Environments Workshop, Austin, Texas, 2008, pp. 1 - 10.

[9] F. Liu, W. Guo, Z. Q. Zhao, and W. Chou, "SaaS Integration for Software Cloud," in IEEE 3rd International Conference on Cloud Computing, Miami, FL, 2010, p. 402.

[10] McAfee Security. Security as a Service. [Online]. http://www.mcafee.com/us/small/security_insights/security_as_a_service.html

[11] HackerTarget.com. (2008, April) Security from the Cloud: Remote Vulnerability Scanning. Whitepaper.

[12] K. Balakrishnan, S. Roy, and M. Holt. (2009, April) Email and Web Security SaaS. Whitepaper.

[13] Panda Security. (2009) Switching from Anti-Virus to Security as a Service (SaaS). Whitepaper.

[14] A. V. Dastjerdi, K. Abu Bakar, and S. Tabatabaei, "Distributed Intrusion Detection in Clouds Using Mobile Agents," in Third International Conference on Advanced Engineering Computing and Applications in Sciences, 2009, pp. 175-180.

[15] P.Kannadiga and M.Zulkernine, "Distributed Intrusion Detection System Using Mobile Agents," in Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, 2005.



**Richa Sondhiya** has done her bachelor of engineering in computer science from Takshshila Institute of Engineering &Technology, Jabalpur in 2010. She has pursuing her M.Tech in Information Technology from L.N.C.T. Bhopal. She has great interest in field of Soft Computing. She is currently working in field of cloud computing.

**Maneesh Shreevastav** has done his bachelor degree from UIT. He has done his master degree from Maulana Azaad National Institute of Technology. He has great interest in field of Networking.

**Mahendra Mishra** has completed his bachelor degree from Shri Vaishnav Institute Indore. He has done his Master Degree from RITS Bhopal. He is having his special interest in field of Data Mining and Mobile Adhoc Networks.