ISSN: 2321-7782 (Online)
Impact Factor: 6.047

Volume 4, Issue 6, June 2016

International Journal of Advance Research in Computer Science and Management Studies

Research Article / Survey Paper / Case Study

Available online at: www.ijarcsms.com

Design & Development of a new hybrid system to Prevent Intrusion at cloud using genetic algorithm

Prashant Singh[1]
Department of Computer Science & Engineering
Amity School of Engineering & Technology
Amity University, Lucknow – India

Bramah Hazela[2]
Department of Computer Science & Engineering
Amity School of Engineering & Technology
Amity University, Lucknow – India

Abstract: ***Data are at the core of IT security concerns for any organization, whatever the form of infrastructure that is used. Cloud computing does not modified this, but cloud computing does bring an added focus because of the distributed nature of the cloud computing infrastructure and the shared responsibilities that it involves. Security considerations enforce both to data at rest (held on some form of storage system) and also to dynamic data (being transferred over some form of communication link), both of which may required specific consideration when using cloud computing services. Essentially, the questions relating to data for cloud computing are about several forms of risk: risk of theft or unauthorized disclosure of data, risk of tampering or unauthorized modification of data, risk of loss or of unavailability of data. It is also worth remembering that in the case of cloud computing, data assets, may well incorporate things such as application programs or machine images, which can have the similar risk considerations as the contents of databases. Cloud computing provides a framework for supporting end users easily attaching powerful services and applications through Internet. Denial of Services (DoS) attack or Distributed Denial of Services (DDoS) are significant security issues in cloud environment. Moreover, we show a distributed architecture for providing intrusion detection in Cloud Computing environment, which enables Cloud service providers to offer security solutions as a service using genetic and neural network. It is a hierarchical and multi-layer architecture designed to accumulate information in the Cloud environment, using multiple distributed security components, which can be used to perform complex event correlation analysis.***

Keywords: Security, Cloud computing, Threats, Genetic Algorithm, Neural Networks.

I. INTRODUCTION

Cloud computing is a service distribute over the internet for computing, data access and cloud storage by create scalability, elasticity and less cost. Second invention platform for division which suggests [1] a variety of services and applications to the user not including actually obtain them. Cloud computing has become the rising technology for computing these days. A cloud computing is generally includes models like Infrastructure-as-a-service [1], Platform-as-a-service and Platform-as-a-service. Cloud Computing is mainly focuses on allocating data and computations over a scalable information centers of network. The cloud computing is basically a type of computing that rely on allocation the computing resources and rather than local servers and applications. In cloud computing word cloud used as a symbol for internet, so phrase the cloud computing means "a type of internet based computing [2]. The web device services, storage and applications are delivering to company computers. Example: where the larger collections of the system are connected in private networks/public networks, the energetically scalable are environment for the application, data and file storage. This technology are cost of estimation, application hosting, substance storage and delivery reduced the considerably. Within cloud computing has used four characteristics: scalability, elasticity, standardization, cost effectiveness.

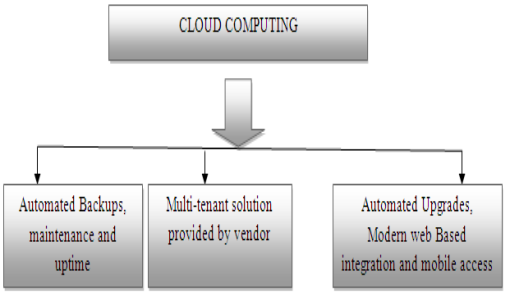

Figure 1: Cloud Computing

They are computing refers to computing collection of the virtualized computer property. As a cloud computing system ought to have the properties that are:

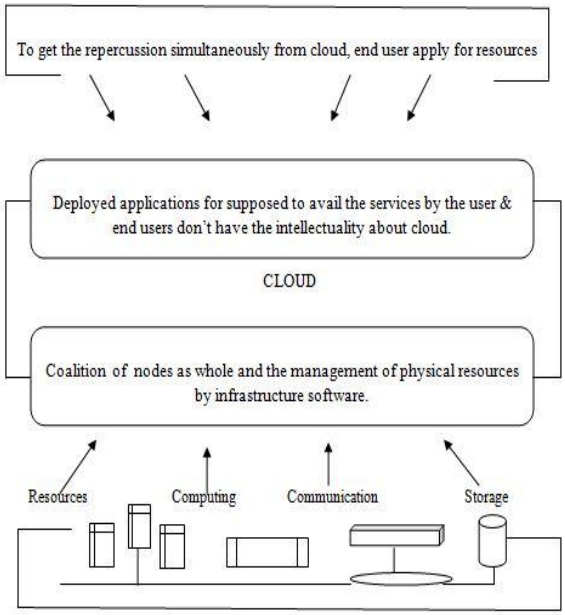

Figure 2: Cloud computing pattern

- To impart the service to end users speedily via interface on Internet.
- To allocate and organize physical resources properly and rapidly.
- To obtain the characteristics of scalability.
- To conform the rule of on demand function and distribution.
- To deal with gigantic quantity of requests and examine needs gracefully.

The pattern of the model carry out the relation $m \leftrightarrow 1 \leftrightarrow n$ .The above figure shows the cloud computing system arrangement.

1.1 CLOUD MODELS

Basically cloud models are two types: deployment model and Service model Cloud computing.

A. Deployment Model: Cloud deployment models correspond to a specific type of cloudy environment, primarily divided by size and access. Deployment model are three types

Public cloud computing: It is specifically rely on third party to advice services by paying them on monthly basis according to the procedure. Public Cloud environment is made accessible to all unrestrained users who can subscribe the desired services [3]. The security issues apparently will be decided by the service provider and so it is very essential to choose the service provider.

Private Cloud Computing: The organization itself has command over the services. Usually organizations leave for private cloud when the need of sensible data occurs. Scaling can be done very efficiently by appending hardware and thus the environment can be expanded. The security will be more due to the control of contained by internal structure and thereby data will be safe.

Hybrid Cloud Computing: It is the amalgamation of both public and private cloud computing. A limited(less) sensible data will be storing in public and all others in Private Cloud.

Table 1: Comparison of Public and Private Cloud Computing

Public Cloud Computing	Private Cloud Computing
Can be used by more customers	Only a single customer
Suitable for no sensible information	Suitable for sensible information
Less security	Highly Secure
Utilizes shared infrastructure.	Utilize shared infrastructure.

B. Service Models: A cloud is a computing process in which services are scattered above network by computing processes. The cloud symbols concealing for complex environment it has in organization arrangement. Service models consist of three main classifications:

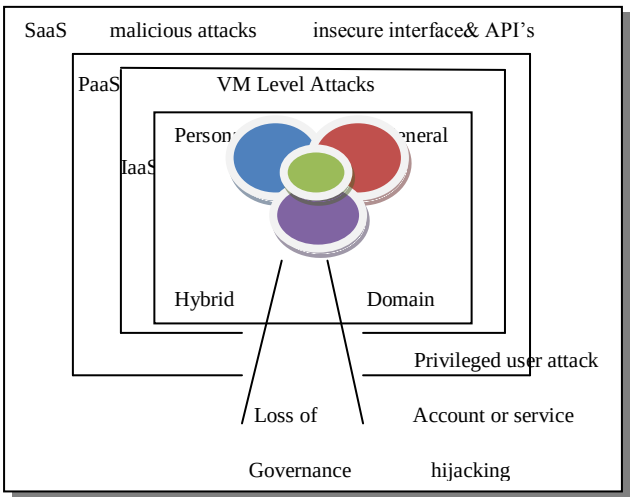

Figure 3: Service Models

SaaS (Software as a Service)

- The web benefits are assigned to commercial software.
- From a middle location, the software is operated.
- One to many is the process for assigning the software.
- The users don't require managing software up gradations and patches.
- Among number of software's, Application Programming Interfaces allows the integration.

PaaS (Platform as a Service)

- To permit the services to expand, experiment, organize, host and safeguard the application in the same integrated improved atmosphere and the analogous services desired to accomplish the application development mechanism.
- The web designed UI formation tools assists to make, adapt, test and organize dissimilar UI framework.
- Multi-tenant strategy that has numerous simultaneous users use the homogeneous growth application.
- Constructed in scalability of deployed software counting load balancing and failover.
- In addition with the web services and databases of repeated standards.
- Sustain for growth team collaboration – some PaaS solutions comprises of project planning and communication tools.

- Tools to curb(handle) billing and subscription management.

IaaS (Infrastructure as a Service)

- The resources are scattered as a service.
- It allows for legitimate scaling.
- It has a patchy cost, usefulness pricing model.
- Usually has multiple users on a solitary piece of hardware.

1.2 CLOUD ATTACK

Attacks performed with the help of tools or exploit scripts that target vulnerabilities existent in cloud protocols, services and applications. They may emerge in the form of denial-of service attacks, probes, and worms, and may leave their trails at several locations of cloud‘s infrastructure. [3, 5]. The Data Security of "Cloud" is stored in different physical locations, in various parts of the Earth, in the absence of corresponding technical and regulatory constraints; data security is difficult to get protection. [2, 3].

In this paper, we created a series of rules to illustrate security policies that IDS can monitor. The method increases resource availability of Cloud Computing system and handle the potential threats by deploying Multi-layer IDS. We can suppose that VMs have equal quantity of resource, then host OS can assign less guest OS with IDS, because IDS use much resource based on genetic algorithm.

II. GENETIC ALGORITHM

Genetic algorithms (GA's) are search algorithms that work via the process of natural selection [1], shown in figure 9. They begin with a sample set of potential solutions which then evolve to a set of best solutions. Inside the sample set, solutions that are poor tend to die out while the better solutions mate and propagate their advantageous traits, thus introducing more solutions into the set that boast better potential (the total set size remains constant; for each new solution added, an old one is removed) [25]. A little random mutation help that a set won't stagnate and simply fill up with numerous copies of the same solution [27].

In general, genetic algorithms tend to work better than traditional optimization algorithms because they're less likely to be led astray by local optima. This is because they don't make use of single-point transition rules to move from one single instance in the solution space to another. Instead, GA's take advantage of an entire set of solutions spread throughout the solution space, all of which are experimenting upon many potential optima. Every person in a population is taken as fixed-length binary string. The size of the string is the parameters domain and the precision [28].

The decoding starting from binary string $<b_{22}b_{21}\ldots b_0>$ into a real number is straightforward. It is explained in two steps:

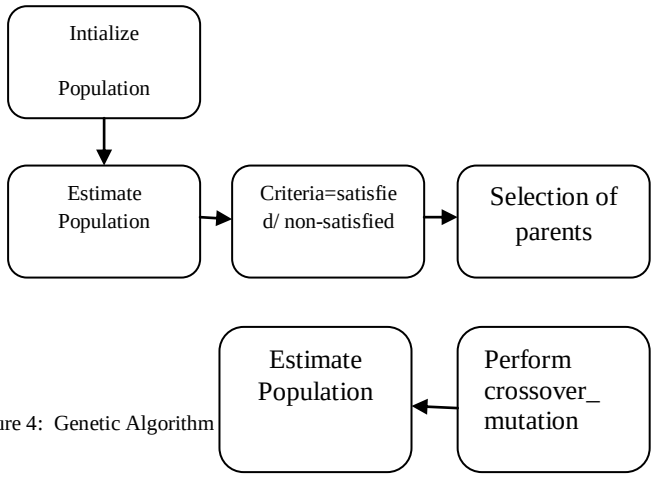

Figure 4: Genetic Algorithm

1. Conversion of the binary string $<b_{22}b_{21}\ldots b_0>$

$$x' = \sum_{i=0}^{22} b_i 2^i$$

2. Calculation of a real number x as

$$x = -2.0 + x' \frac{7}{2^{23} - 1}$$

Genetic Algorithm is shown below in the form of an algorithm:

{

Initialize_population;

Estimate_population;

While Termination_Criteria_NotSatisfied

{

Select parents for reproduction;

Perform crossover_mutation;

Repair ();

Estimate population;

}}

III. NEURAL NETWORK

The basic aim of neural network is to work like human brain works [34]. Neural network consists of various no. of neurons and their working is similar to the brain neuron structure. There are various types of neural networks but commonly used neural network is Back Propagation Neural network [35]. Two types of structure has been found in neural network model:

- Cyclic;
- Acyclic

The normal Back propagation is the mainly applied to train Multilayer FNN. The linear and nonlinear outputs are correspondingly given by:

The net input has given by:

$$n1^{k1+1}(i) = \sum_{j=1}^{s1k1} w1^{k1+1}(i,j) a1^{k1}(j) + b1^{k1+1}(i) \quad (12)$$

The unit i is given by

$$a1^{k1+1}(i) = f1^{k1+1}(n1^{k1+1}(i)) \quad (13)$$

This recurrence relation is executed at the final layer

$$-F1^{M1}(n1^{M1})(t1_{q1} - a1_{q1}) \quad (14)$$

The structure of neuron can be shown as below [34]:

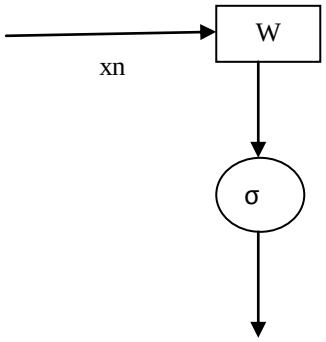

Fig. 5: Neuron Structure

It can be achieved that the function *s* has zero threshold and the actual threshold with the opposite sign is understood as a further weight, *bias* w_0= -*h* of additional formal input x_0=1 with constant unit value.

IV. PROPOSED WORK

Proposed Feature Selection Technique Using GA Genetic based feature selection algorithm has been used in this work in order to select suitable subset of features so that they are potentially useful in classification. Another advantage of GA based feature selection in this work is that it finds and eliminates the redundant features if any because these redundant features may misguide in clustering or classification. The reduction in number of features reduces the training time and ambiguousness, thus a weighted sum genetic feature selection algorithm has been proposed which has increased global search capability and is better in attribute interaction when compared to other algorithms such as the greedy method.Fig.4 shows the architecture of Genetic Cloud IDS.

We propose mix of Cloud Intrusion Detection System and two types of chromosomes based on different criteria. The first type is created based on the job length, The second type is created based on the bandwidth of the resources. In each type of chromosomes and represented as chromosomes form; Then the computational resources are assigned to the chromosomes randomly and then the algorithm calculates the fitness value of every chromosome of each type of chromosomes. The fitness value is achieved by neural network. Then, the algorithm selects two chromosomes individuals from the mentioned two types of chromosomes according to the fitness value. Then the algorithm performs the crossover operation with the aid of neural network for these two chromosomes. At the end of this step, a new chromosome will be created which is the best chromosome of the first generation. For cross over step are homolog. The detail of the proposed approach is as follow.

1-The main purpose of our algorithm is: assigning the most suitable resources to the jobs based on the bandwidth and computational capacities of resources and the job length.

2-The algorithm tries to assign the jobs with high length to the resources with high bandwidth and high computational resources. We evaluate the performance of our approach with some famous cloud scheduling models.

Algorithm:

Upload data.

predicts the rows and cols of the data.

new data is a variable which would not contain any nan number

new_data(j,2)=new_val2; taking data form 1 to total number of sessions

Data uploading done.

Get clusters

Apply size command provides you the way to find the number of rows and cols of the matrix


```
Apply Genetic algorithm
For cluster1 and similarly 2
for i=1:loop_val
   for s=1:loop_val2
   Fs=optp(i,s);
   Ft=mean(optp(i,:));
   FitnessFunction = @(e)fitness_fn(e,Fs,Ft);    %calling fitness function
   numberOfVariables = 1;
   [x(i) fval] = ga(FitnessFunction,numberOfVariables,[],[],[],[],[],[],[],options);     reduced_index(i)=round(x(i));
   if reduced_index(i)==1
      GareducedFeatures(s1,p1)=Fs;
      Gapos1_x(s1,p1)=i;
      Gapos1_y(s1,p1)=s;
      s1=s1+1;
      p1=p1+1;
Get suspected attacks
Optimise the values using NN as shown below;
for i=1:numel(GareducedFeatures11)
   Training_set(i) =GareducedFeatures11(i);
   group(i)=1;
   totalvalues=totalvalues+1;
net=newff(Training_set(1:numel(group)),group,20);
net.trainparam.epochs=50;
net=train(net,Training_set(1:numel(group)),group);
```

V. RESULTS AND ANALYSIS

A Genetic Algorithm (GA) is a optimization technique for generating new rules in cloud Intrusion detection system. The evolution usually starts from a population of randomly generated individuals. Here the individuals are Intrusion detection rules. In each generation, the fitness of every rules in the population is evaluated, multiple rules are selected from the current population based on their fitness, and modified by recombination and mutation to form a new rules. The new rules is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the rules.

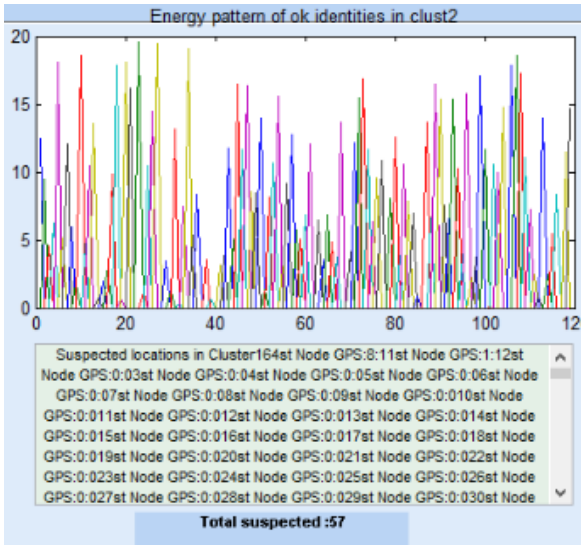

Figure 5: Result Evaluation

Above figure displays the suspected threat numbers using genetic and neural network in proposed work and it has been seen that 57 number of threats has been suspected.

VI. Conclusion and Future Scope

Cloud intrusion detection datasets are able to detect cloud attacks. In proposed work, Cloud based IDS is able to detect 57% of Random sets of cloud attacks. IDS was able to detect the same percentage of attacks and no false positive alarm is raised while filtering background traffic using neural network. The efficiency of cloud IDS is determined by injecting attacks. Result show that latency is increasing according to background traffic. This does not have effect on cloud intrusion datasets.

The work presented in the paper has fulfilled some gaps. Further work, in this area could be carried out in order to make universal cloud intrusion datasets. The efficiency of cloud based IDS can be improved by implementing multiple IDS over the cloud and installing Multiple Management servers.

Acknowledgement

I would like to acknowledge and give credits to the following people who helped me for the completion of this research paper. I am very thankful to the Head of the Department of Computer Science, Dr. Deepak Arora, to encourage me and provided all kinds of facilities for the completion of this thesis. I also like to thanks my guide, Mr. Bramah Hazela, whose support and guidance helped me to complete the work. Finally a special thanks to Wg. Cdr. (Dr.) Anil Kumar (Retd.), who provide best research facilities and made everything possible.

References

1. Cong Wang, Qian Wang, Kui Ren and Wenjing Lou “Ensuring Data Storage Security in Cloud Computing.” IEEE, 2009.
2. Balachandra Reddy Kandukuri, Rama Krishna Paturi and Dr. AtanuRakshit, “Cloud security issues” In Services,Computing, 2009. IEEE International Conference on, page 517520, 2009.
3. Guan, Qiang, Chi-Chen Chiu, and Song Fu. "Cda: A cloud dependability analysis framework for characterizing system dependability in cloud computing infrastructures." Dependable Computing (PRDC), 2012 IEEE 18th Pacific Rim International Symposium on. IEEE, 2012.
4. Amanatullah, Yanuarizki, et al. "Toward cloud computing reference architecture: Cloud service management perspective." ICT for Smart Society (ICISS), 2013 International Conference on. IEEE, 2013.
5. Luo, Shengmei, et al. "Virtualization security for cloud computing service."Cloud and Service Computing (CSC), 2011 International Conference on. IEEE, 2011.
6. Sabahi, Farzad. "Virtualization-level security in cloud computing."Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on. IEEE, 2011.
7. Luo, Shengmei, et al. "Virtualization security for cloud computing service."Cloud and Service Computing (CSC), 2011 International Conference on. IEEE, 2011.

8. Duan, Qiang, Yuhong Yan, and Athanasios V. Vasilakos. "A survey on service-oriented network virtualization toward convergence of networking and cloud computing." Network and Service Management, IEEE Transactions on 9.4 (2012): 373-392.
9. Sabahi, Farzad. "Virtualization-level security in cloud computing."Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on. IEEE, 2011.
10. Tan, Yuesheng, Dengliang Luo, and Jingyu Wang. "Cc-vit: Virtualization intrusion tolerance based on cloud computing." Information Engineering and Computer Science (ICIECS), 2010 2nd International Conference on. IEEE, 2010.
11. Zhong, Liang, et al. "A Virtualization-based SaaS Enabling Architecture for Cloud Computing." Autonomic and Autonomous Systems (ICAS), 2010 Sixth International Conference on. IEEE, 2010.
12. Dong, Hanfei, et al. "Formal discussion on relationship between virtualization and cloud computing." Parallel and Distributed Computing, Applications and Technologies (PDCAT), 2010 International Conference on. IEEE, 2010.
13. Zhu, Qixuan, and Xi Zhang. "Game-theory based power and spectrum virtualization for maximizing spectrum efficiency over mobile cloud-computing wireless networks." Information Sciences and Systems (CISS), 2015 49th Annual Conference on. IEEE, 2015.
14. K. Deb, "An Efficient Constraint Handling Method for Genetic Algorithms," Comput. Methods Appl. Mech. Eng., vol. 186, no. 2-4, pp. 311-338, 2000.
15. Guo Pengfei, Han Yingshi, “Chaotic genetic algorithm for structural optimization with discrete variables”, Journal of Liaoning Technical University, 2007, 26(1), pp. 68-70.
16. Guo Pengfei, Han Yingshi, “Chaotic genetic algorithm for structural optimization with discrete variables”, Journal of Liaoning Technical University, 2007, 26(1), pp. 68-70.
17. Ma Lixiao, Wang Jiangqing. The application of genetic algorithm in the combinatorial optimization problem. Computer engineering and science, 2005, 27 (7) 72-73.
18. Ge Jike etc. Genetic algorithm research review. Computer application research. 2008, 25 (10): 2911-2916.

