Intrusion Detection and Prevention in Cloud Computing using Genetic Algorithm

Umar Hameed, Shahid Naseem, Fahad Ahamd, Tahir Alyas, Wasim-Ahmad Khan

Abstract— High level security is an essentially required in the communication and information sharing on the network clouds. Intrusion detection system (IDS) is being used to detect violations and malicious behavior over networks and hosts. Purpose of our paper is to provide an Intrusion detection system to detect and prevent the malicious behavior on the cloud computing. We propose an intrusion detection system which is based on the cloud computing to reduce the risk of intrusion on the cloud networks and cover up the deficiency of already in use intrusion detection systems. Our design is based on cloud computing Software-as-a service (Saas) model for detection and prevention of intrusion cloud based users. Our System set up a hardware layer which forwarded the request to a service layer which analyzes the request patterns and further forwards to the intrusion detection layers which checks the request and provides the alerts against the malicious requests by applying genetic algorithm. Already existing Intrusion detection systems provides the detection in a single unit and use of genetic algorithm is rear in the already existing systems..

Index Terms: Intrusion detection, Prevention, Cloud Computing, Genetic Algorithm, Knowledge Base repositor.

1 Introduction

Intrusion detection systems (IDSs) are software or hardware mechanisms which helps the system to detect attacks against computer systems. IDS are composed different components such as sensors, console, and central engine. Sensors detect security events, console monitors events and Central Engine records events logged by the sensors in a database and use a system of rules to generate alerts from security events received.

It is normally designed and used to monitor and detect the invalid activities of the connected nodes and the users which can harm the system resources, for the security of the system. It can keep a track record of user applications, networks or combination of activities to tack the known and unknown attacks [1].

Mainly Intrusion detection system provides the following features,

- Monitoring and analyzing the user's activities and abnormal activity.
- Auditing the system configuration and vulnerabilities.
- Critical System and data files integrity assessment and generating alarms.
- Operating System activities analysis.

- *Umar Hameed is with computer science department as a research fellow at NCBA&E , Lahore, Pakistan his area of intrest is Artificial Intelligence. He is working as Deputy Controller Examination at UCEST (Lahore Leads University), Lahore, Pakistan. Cell +92-300-4253292 E-mail: umfer@yahoo.com*
- *Shahid Naseem* is *is with computer science department as a PHD research fellow at NCBA&E , Lahore, Pakistan his area of intrest is Artificial Intelligence. He is working as Assistant professor (IT) at UCEST (Lahore Leads University), Lahore, Pakistan. Cell +92-300-5305627 E-mail: shahid.naseem@gmail.com*
- *Fahad Ahmad* is *is with computer science department as a PHD research fellow at NCBA&E , Lahore, Pakistan his area of intrest is Artificial Intelligence. He is working as Hardware Enginee at PIMS, Lahore, Pakistan. Cell +92-333-0969489 E-mail: fahadahmad84@gmail.com*
- *Tahir Alyas is with computer science department as a PHD research fellow at NCBA&E , Lahore, Pakistan his area of intrest is Artificial Intelligence. He is working as Lecturare (IT) at Garrision University, Lahore, Pakistan. Cell +92-333-6106500 E-mail: tahir.alyas@gmail.com*
- *Waseem Ahmad Khan is with computer science department as a research fellow at NCBA&E , Lahore, Pakistan his area of intrest is Artificial Intelligence. He is working as Faculty Member at NCBA&E Lahore, Pakistan. Cell +92-322-4645667 E-mail: wasimahmad.ucitr@gmail.com*

2. Approaches for Intrusion Detection

Anomaly Detection approach is based on the study of the "normal" behavior and creating a attributed model from this behavior. As abnormal behavior indicates the illegitimate use of the system so any behavior which deviates from the normal behavior is known as the anomalous [2]

Main drawback of Anomaly based approach are it often requires extensive training in order to find out the normal behavior patterns and in many cases it provides the false alarms as particular normal user and system behavior can differ widely.

Misuse Detection approach is based on the principle that there exists a well defined record of already attempted intrusion attacks. These could be matched with the current preceding and an exact type of attack can be detected easily. So in Intrusion detection systems that use Misuse or Signature based detection looks for a certain match from a predefined pattern of events to track the known attacks [3].

Main advantages of the misuse detection approach are that false alarms are not often occurs and it does not require extensive training to detect attacks.

Intrusion detection system is mainly classified in two categories based on the system monitoring.

Network Intrusion Detection System (NIDS) that monitors the complete network traffic to track and detect proposed attacks on the network and inform administrator about the attack.

Another category is called Host Based Intrusion Detection System (HIDS) that monitors a single host system. HIDS involves installing an agent on the local host that monitors and reports on the system configuration and application activity.

ISSN 2229-5518

3. CLOUD COMPUTING

Cloud computing is the mechanism of computing services delivery over the Internet. In cloud services individuals and businesses can use software and hardware facilities which are managed by third party at some remote locations. Some main examples of cloud services are online file storage, social networking sites (Facebook, Twitter etc), webmail, and online business applications etc. Cloud computing model provides the facility of accessing information and sharing the computer resources from anywhere through a remote network connection. Cloud computing provides a shared pool of resources, including data storage space, networks, computer processing power, and specialized corporate and user applications [4]

Main characteristics of the cloud computing are

• On-demand self-service User / Organizations can request and manage their own computing resources, like network resources, storage, processing as per their requirements as choice.

• Broad network Access: Provides the capabilities to access the services on the internet and private networks thorough a mechanism enables access to client platforms (e.g., mobile phones, tablets, laptops, and workstations).

• Resource pooling: Provides the capabilities of accessing and using the multiple pooled computing resources, from different networks and remote data access points as required.

• Rapid elasticity: Services can be scaled and appropriated according to the requirements in any quantity at any time.

• Measured Services: Cloud systems always provide a automatic control of resources usage metering and billing accordingly (eg. user accounts, storage, bandwidth usage, processing). Resources usage are monitored and controlled in accordance with the services provided and availed.

Fundamental service models of cloud computing include software-as-a-service (Saas), platform-as-a-service (Paas) and infrastructure-as-a-service (Iaas). Software as a service model provides the access of complete services like operating software, database etc. to the user managed by the cloud providers. Cloud provider installs operating application software and provides the complete infrastructure platform. Cloud user can access the facilities without any installation of any kind. In business oriented models Saas provides pricing applications to apply the subscription fee. Most commonly used Saas applications in use are salesforce.com, Google apps etc.

In Platform as a service cloud provider manage complete hardware, network and a platform encapsulating a layer of software and provides it as a service that can be used to build higher-level services. Cloud user can build and run their own software solutions and applications according their own requirements without bothering about the basic requirements of hardware and operating layers and their maintenance. PaaS offerings can provide for every phase of software development and testing, or they can be specialized around a particular area such as content management.

In Infrastructure as a service model provides computers Physical or (more often) virtual machines (including storage and compute capabilities as standardized services over the network. Servers, storage systems, switches, routers, and other systems are pooled and made available to handle workloads to the cloud users. Could users run their own operating systems and applications to perform their tasks [5].

Moreover, cloud services can be made available mainly by three ways that include public clouds where cloud provider provides the storage, server and network resources to use for users. These are normally used to deploy the services which may be made available to all the general (possible) cloud users. Email services, social networking sites and online photo storages are normally deployed through public cloud. Private cloud are built and deployed for some particular client or organization providing the utmost control over data, security, and quality of service to the user. Weather the infrastructure is maintained in own premises or by the cloud provided, the organization have the significant control on the resources of the Private cloud. Hybrid cloud combines the characteristics of both public and private models. The ability to supplement a private cloud with the resources of a public cloud can be used to maintain service levels for rapid workload fluctuations.

4. GENETIC ALGORITHM

Genetic algorithm is the most popular type of evolutionary algorithm. In a genetic algorithm, a population of candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem is evolved toward better solutions. Each candidate solution has a set of properties (its chromosomes or genotype) which can be mutated and altered; traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible.

The evolution usually starts from a population of randomly generated individuals, and is an iterative process, with the population in each iteration is called a generation. In each generation, the fitness of every individual in the population is evaluated; the fitness is usually the value of the objective function called "fitness function" in the optimization problem being solved. During the process of evaluation "Crossover" is used to simulate natural reproduction and "Mutation" is used to mutation of species. The more fit individuals are stochastically selected from the current population, and each individual's genome is modified (recombined and possibly randomly mutated) to form a new generation. The new generation of candidate solutions 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 population.

Although crossover and mutation are known as the main genetic operators, it is possible to use other operators such as regrouping, colonization-extinction, or migration in genetic algorithms [6].

A typical genetic algorithm requires:

1. A genetic representation of the solution domain,
2. A fitness function to evaluate the solution domain.

A standard representation of each candidate solution is as an array of bits. Arrays of other types and structures can be used in essentially the same way. The main property that makes these genetic representations convenient is that their parts are easily aligned due to their fixed size, which facilitates simple crossover operations. Variable length representations may also be used, but crossover implementation is more com-

plex in this case. Tree-like representations are explored in genetic programming and graph-form representations are explored in evolutionary programming; a mix of both linear chromosomes and trees is explored in gene expression programming.

Once the genetic representation and the fitness function are defined, a GA proceeds to initialize a population of solutions and then to improve it through repetitive application of the mutation, crossover, inversion and selection operators. General working of genetic algorithm is shown figure 1.

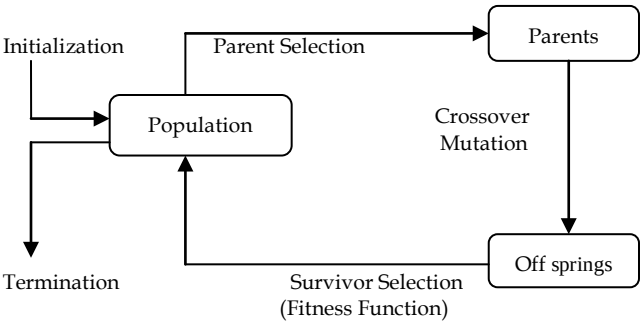

Figure 1: General Scheme of Genetic Algorithm

5. Current System

Currently used system is shown in Figure 2, which elaborates that in the current system users groups and cloud users, information is monitored by the agent groups and then forwarded to the cloud computer service components which further forwards information to the collectors which are responsible to forward the information to the appropriate analysis engine for detection.

Intrusion detection service component is mainly responsible for detection with its three main components. Detection Engine is a sophisticated decision and pattern matching mechanism. It analysis data came from the collector and matches it to known patterns of activity stored in the signature database. Event Publisher provides the results reports to the users. IDS

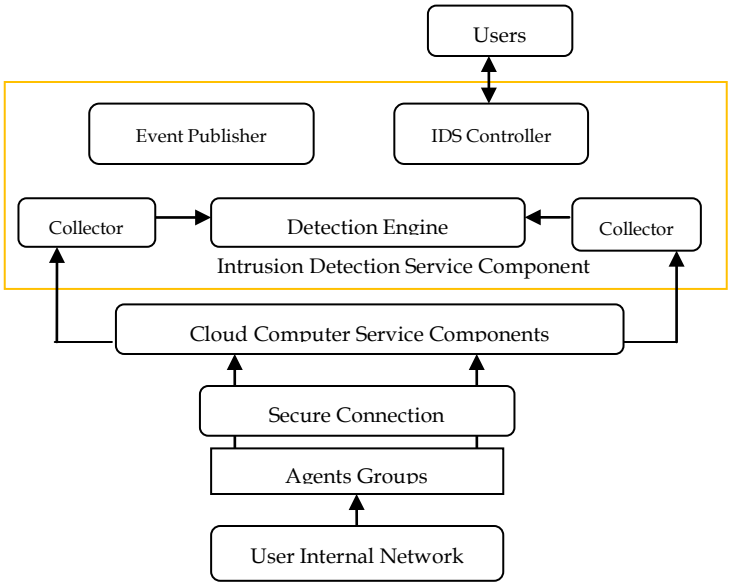

Figure 2: Existing Cloud IDS Model

6. Proposed System

Developing an Intrusion detection mechanism for cloud environment is quit challenging and motivated for both cloud user and cloud providers. It should cover all the normal features of a network based IDS there must exist the security features for fast and secure access to applications and data.

We propose a three layers method to detect the intrusion. The design is relaying on Software-as-a service model for providing the security to the cloud based users.

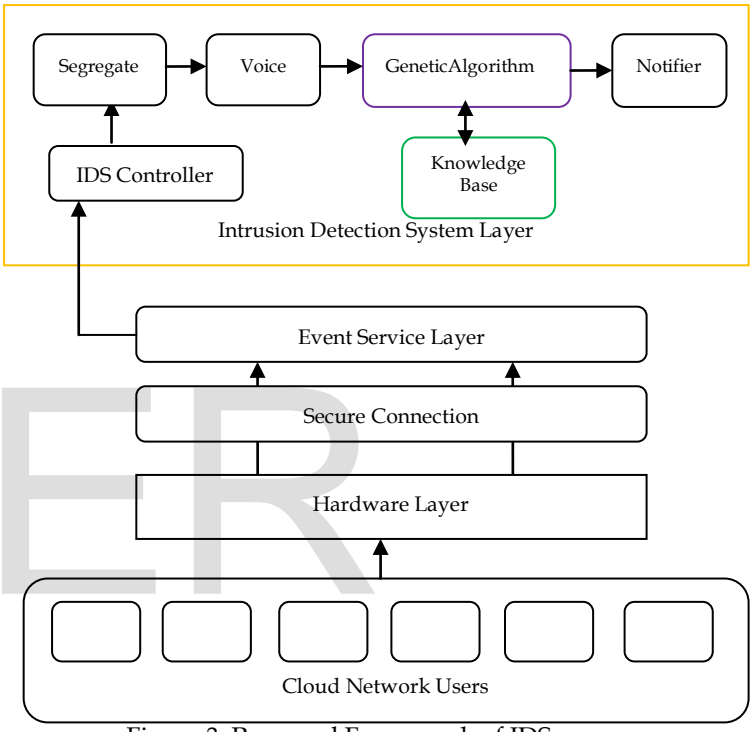

Figure 3: Proposed Framework of IDS

• **Cloud Network Users:** Presents the users on the cloud accessing the cloud facilities.

• **Hardware Layer:** Our System deploys a hardware layer which is integrated in the cloud network to collect the necessary information. The hardware layer protects the whole network. It is incorporated in cloud based on the rules set or threshold to improve the efficiency and provide the protection flexibility. It forwards the request to the event service layer through the secure connection. The hardware is just like an agent.

• **Secure Connection**: Is a secure connection path established by event service layer to absorb the information from hardware layer otherwise the system behavior can be tainted by external intrusion.

• **Event Service Layer:** This layer works as an intermediate layer; it receives the messages from the hardware layer through a secure connection (established to eliminate the ex-

ternal intrusion) checks message and forwards these messages to the intrusion detection layer.

• **Intrusion detection System Layer:** This layer is the main layer responsible for the intrusion detection. IDS layer consists of sub components to for controlling the intrusion detection. All these sub components have specific work to elaborate.

IDS Controller receives the message it is responsible for reading the details and then forwards the items of the interests to the segregator for further segregation of voice data and forwarding it to the Genetic Algorithm Unit.

Segregate separates the data based on its content and thus in our proposed framework, it segregates the voice data and forwards it to the voice unit.

Voice section picks the voice records and forwards it to the Genetic Algorithm unit.

Genetic Algorithm Unit: Genetic Unit is a pattern matching and decision making unit performs these tasks by using the genetic algorithm technique. It analysis data (voice data) in details and matches the same with known behaviors stored in the existing knowledge base. Fitness Function "X=δX0 " is the key for checking the match.

First of all it checks the existence by matching the location in knowledge base. If the location is not matched with any existing details in the Knowledge base it provides an exit from the system. If existence check is passed it analysis and matches the voice records for the proper recognition based on the fitness function X=δX0.

If the fitness is proved then different authentication checks are implemented to reach the result at each step it apply genetic algorithm.

If the desired record is not matched it reproduces the new record by mutation, performs fitness test X=δX0 and rechecks from the data in knowledge base. Same is continued for three generations as per termination criteria. Based on the said results it identifies the malicious behavior, provides access or exit and generates the alerts through the notifier.

Knowledge Base is the stored knowledge about the prospect cloud service users. It retains the knowledge required for security checks and can be in partial form.

Notifier is a interface to provide the result reports for analysis to users and the system gathered from the results provided by the genetic algorithm unit.

Working of the genetic unit is displayed in figure 4.

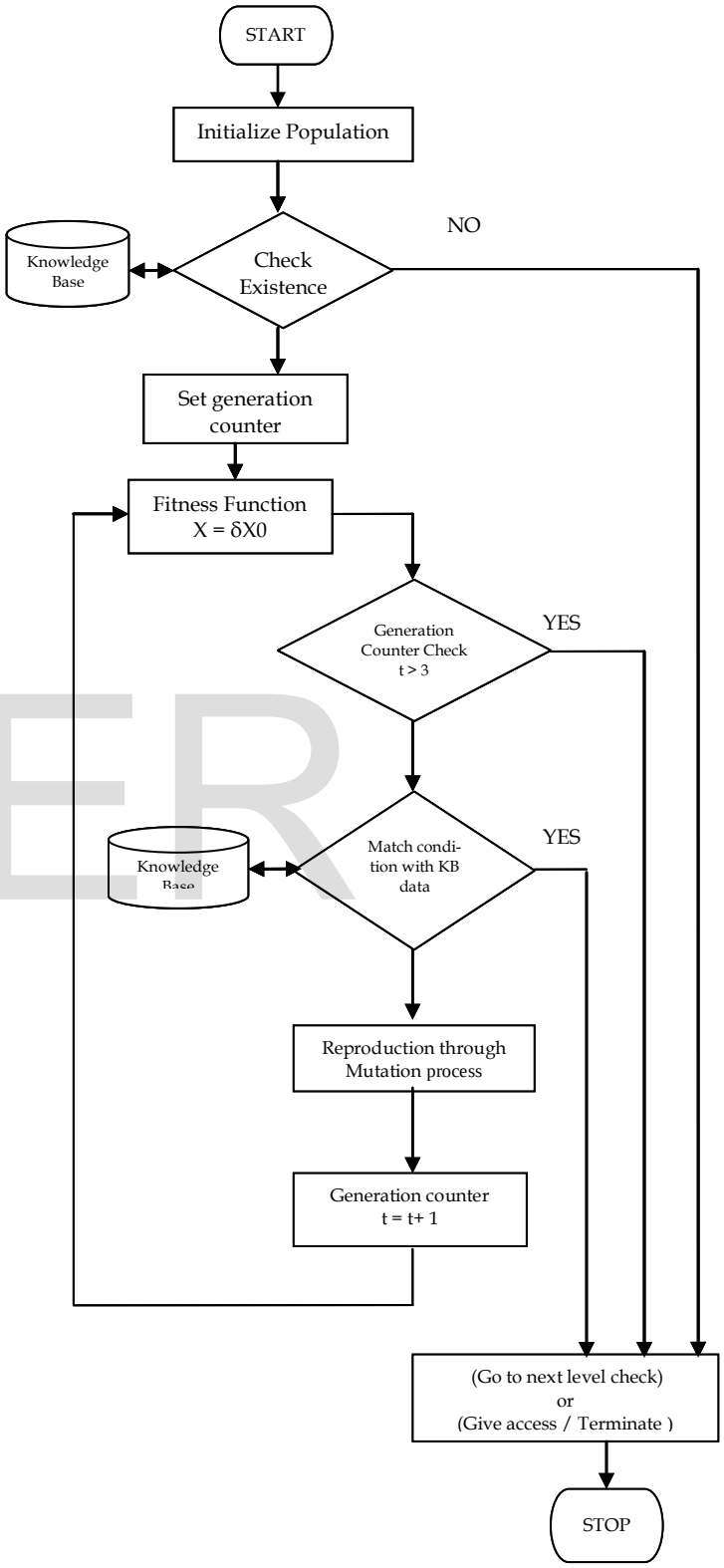

Figure 4: Working of Genetic Algorithm Unit

7. Conclusion

It is concluded that use of genetic algorithm with fixed generational iteration in intrusion detection system helps in enhancing the security features in cloud computing. As genetic unit checks and restricts intrusion attacks for certain threshold level of generations and thus do not provides access to the data unless security checks are completed. The proposed model is focused on voice data validity, genetic algorithm is a significant approach to detect, prevent and thus reduce the intrusion in this regard.

8. Future Work

Our current proposed framework for IDS is handling voice intrusion only. An enhanced implementation of model to take care of all other data aspects for IDS, like high definition visual or biometric data and simulation test is prospective goals. IDS system can be enhanced further by encapsulating the adaptive data mining features with the efficient evolutionary computing methods for effective intrusion detection model based on the large adaptive knowledge base.

9. References

[1] S. INSTITUTE, "Understanding Intrusion Detection System," SANS INTITUTE INFO SECTION READING ROOM, 2001.

[2] A. Zarrabi and A. Zarrabi, "Internet Intrusion Detection System Service in a Cloud," IJCSI International Journal of Computer Science, vol. 9, no. 5, 2012.

[3] C. Lawrence, "Intrusion Prevention Systems: The Future of Intrusion Detection," in Intrusion Prevention Systems: The Future of Intrusion Detection, Auckland, 2004.

[4] Office of Privacy Commissioner Of CANADA, "www.priv.gc.ca," Introduction to Cloud Computing. [Online].

[5] M. Boniface, B. Nasser, J. Papay and S. C. Phillips, "Platform-as-a-Service Architecture for Real-Time Quality of Service Management in Clouds," in Internet and Web Applications and Services (ICIW), 2010 Fifth International Conference on, Barcelona, 2010.

[6] A. Ziarati, ""A multilevel evolutionary algorithm for optimizing numerical functions"," IJIEC, vol. 2, 2011.