March - April 2020
ISSN: 0193-4120 Page No. 9122 - 9127

A Survey on Anomaly Detection Methods to Secure Data in Cloud Computing Environment

B. Raja Rao[1], Dr. M. Sreenivasulu[2]
[1]Research Scholar, Dept of CSE, JNTUA, Anatapuramu, Center of Research KSRM College of Engineering, Kadapa
[2]Professor & HOD CSE, KSRM College of Engineering, Kadapa
b.rajarao1207@gmail.com[1], mesrinu@rediffmail.com[2]

Article Info
Volume 83
Page Number: 9122 - 9127
Publication Issue:
March - April 2020

Article History
***Article Received**: 24 July 2019*
***Revised**: 12 September 2019*
***Accepted**: 15 February 2020*
***Publication**: 09 April 2020*

Abstract

Cloud computing is the latest technology in use and under research. Cloud computing is totally related subject to internet. Cloud is huge memory storage where clients can manage their data in a cloud service provider. Lack of security the intruders can easily access data. The integrity, security and availability of cloud clients need to be protected against various threats. The security is challenging issue in cloud computing. An intrusion Detection and prevention system is shortly known as IDPS is employed into cloud environment. This is due to knowing about any harmful behaviour of the systems in the network. To provide security to data stored in cloud, various intrusion detection techniques can be used by cloud service providers. Anomaly detection is one of popular intrusion detection technique. The concept of anomaly detection system to detect and prevent the abnormal activities in the cloud computing environment was discussed in this paper.

Keywords; *Anomaly Detection Systems, Cloud Computing, Cloud Security, IDS*

I. INTRODUCTION

Cloud computing is recent trend in computer science and engineering which has ‘n’ number of users who interact with the resources via internet using web-based applications and tools. There are 4 cloud deployment models defined by NIST are private, public, hybrid and community clouds. In the cloud categories public is the first one, the general public can utilize the services of cloud. The different user’s, individuals, large organizations, small medium enterprises and governments are allowed share the cloud resources. Second category of cloud is private; we can have this type of cloud for a single organization use. Third category of cloud is hybrid and which is like a coin whit two sides, one with public and prvate. A community cloud is the art of cloud computing which shares common concerns between several organizations.

Cloud computing offers services to users in different ways: Infrastructure as a service (IaaS), Software as a service (SaaS), and Platform as service (PaaS). Infrastructure as a service is supply related to physical components like hardware, memory like storage, computer servers, data centres to various users. Above said resources are allocated to the cloud users as a virtual machine instances and storage. PaaS supply interface to the users to store, design and develop applications using various development tools, application programming interfaces and use software libraries. The SaaS provides the users a complete software application over internet and this eliminates to install the software in local machines.

Cloud computing has a nature of sharing data openly with other users and hence there is a high possibility of intruders. It requires intrusion detection mechanisms to safe guard data from intruders. The intrusion detection mechanisms are mainly divided two categories. They are misuse detection techniques and anomaly detection techniques.

Published by: The Mattingley Publishing Co., Inc.

Ultimately paper coined to be: The various security threats in cloud computing is mentioned in section 2. The intrusion detection technique and various categories are described in section 3. Section 4 deals with methods and techniques of cloud which deviates from what we expect. Finally the conclusion of the paper is given in section 5.

II. SECURITY ATTACKS IN CLOUD COMPUTING NETWORKS

According to the present scenario, almost all organizations are taking transition to cloud and share the confidential data among its users. Many intruders like hackers, violate the principles of security in cloud computing, hackers are controlled by various ways in access the data are given below:

2.1. Denial of Service (DOS) attack: Attackers are very intelligent in sense, they prevent the original users of the cloud by sending bulk of messages to the server for verifying requests. At the same time intruder so called attackers pause the server before terminating the connection [2]. When the server prepares to close the connection then hacker sends many messages with invalid addresses. By doing so the server may be visible to be very busy. This incident brings the network traffic and services not to accessible by other users.

2.2. Malware injection Attack: In this technique, malicious code or services will be inserted by attacker in cloud network. After inserting the malicious code or service, the system looks like the existing one providing normal services executing in the cloud. The users were forced to download the malicious software without their knowledge. With this, the attackers are stealing the users information without their knowledge.

2.3. Wrapping attack: The attacker uses XML signature wrapping method. In this wrapping method, the malicious code in the message format of Simple Object Access Protocol (SOAP) will be inserted by the attacker. Harmful code is embedded with fake content of the message and is transferred to the server. Then the cloud server may be interrupted with injected malicious code by the attacker [3].

2.4. Flooding attack: In flooding attack, fake data will be created and tries bring the server down by sending the fake data in all directions. The large volumes of fake data will appear in the cloud. Server can be overloaded sometimes, at this scenario with the process that is allocated to the nearest server [4]. When it chooses nearest server then it allocates capable and quicker processing request. While handling the requests, server duty is to find it is according to the principle and rules of the requested request. If it is invalid request then there will be a request for authentication by the server. It will go through the CPU utilization and memory allocation and based on that system will respond.

2.5. Side Channel attack: In this method, a machine said to be malicious virtual machine will be placed in closer to target cloud server by the attacker [9]. This technique is used to target the server machine and observes the cryptographic algorithms to violate the security mechanisms. So the attacker steals the information by analyzing the cryptographic algorithms used by the servers and steals the information. This is major security threat for the system.

2.6. Data Authentication attack: Data Authentication attack [5] is the process of collecting user login and password details unauthorized. The authentication information is frequently targeted by the attackers. There are so many ways to authenticate users which can be based upon user behaviour, what type of user, i.e. based on his search history. The attackers were targeting the authentication process.

III. INTRUSION DETECTION SYSTEM (IDS) TECHNIQUE

Intrusion means trespassing into the cloud without permission. Users can send information to the cloud or receive information from the cloud. If more bulk

Published by: The Mattingley Publishing Co., Inc.

of information then there is a problem for security. An intrusion detection system (IDS) is a system that is used for the purpose of finding and restricting the malicious activities.

There are two categories of Intrusion Detection (ID) methods. They are Misuse based Detection systems and Anomaly Detection systems.

I. Misuse-based Detection or Signature based: Misused based intrusion detection or signature based uses information that is already identified by attack patterns. Which are having deviated behaviour from what we expect are identified and they are matched with the patterns that are in database. Misuse based intrusion detection is very powerful in detecting of intruders and may be failed when there is nothing pattern on attacks is available in stored database [7]. There are four components in misuse detection model. Components are data related set modules, user related profiles, misuse detecting related data and response engine. The Data is drawn from different sources such as network traffic, log files, and etc. The set of huge data should be cleaned and cleaning process is part of pre-processed and it is transformed to easy understanding analyses. User profile is studied for the normal and abnormal behaviours [8]. Profiles are examined in the cloud and intrusion is said to be occur because of deviations of activities from the profiles.

2. Misuse based detection approach: Cloud is shared with in different user, there is a possibility of attacks for information or data. Attackers cannot be identified before and they can be identified dynamically. They are detected based on their abnormal behavioural patterns towards data or network traffic [9]. When it is found to be abnormal case then it is treated as malicious and message is notified by alarm sound to the administrator.

IV. METHODS AND TECHNIQUES OF ANOMALY DETECTION IN CLOUD BASED NETWORKS

Something that is deviating from what we expect is called detection, detection methods and those methods are used to find anomalous patterns in network traffic [10]. It also provides the extra information to the network administer to monitor the network behavior and the reasons for faults in a network. There are 3 categories of Anomalies. They are Point, Contextual and Collective Anomalies.

Point Anomalies

A single instance of data instance is considered to be anomalous with the rest of data instances. If single instance of data deviates from its normal activity, then it is known as anomalous. The anomalous activity appears to be found outside the boundaries of the normal region. The concept of the point anomalies is shown in Fig 1.

Here N1 and N2 are the normal behavior regions. The Points O1 and O2 are said to be anomalies and Points in region O3 are also said to be anomalies. So the anomalous points O1, O2and points in O3 are falls outside the normal regions N1 and N2.

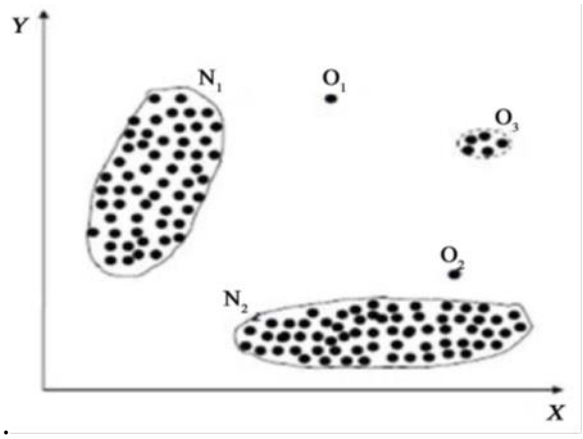

Fig:1 Point Anomalies

Contextual Anomalies

The information is different with respect to the context is said to contextual anomaly. So the contextual anomaly is said to occur when the occurrence of information is in an precise context. The concept of Contextual Anomaly is shown in fig 2.

It shows temperature changes during the months of

Published by: The Mattingley Publishing Co., Inc.

March - April 2020
ISSN: 0193-4120 Page No. 9122 - 9127

year. Normally the temperature gradually reduced or increased, when passed from one month to next month, but during jun, the temperature suddenly decreased and increased. It show anomaly behavior in the context of jun month compare remaining months of the year. The temperature changes in the June month not related to temperature changes in the remaining months.

There is drastic drop in temperature just before June, this value is not an indication of something normal value found during this time. When there is sense that it occurs, it is treated as anomaly and is spelled as conditional anomalies.

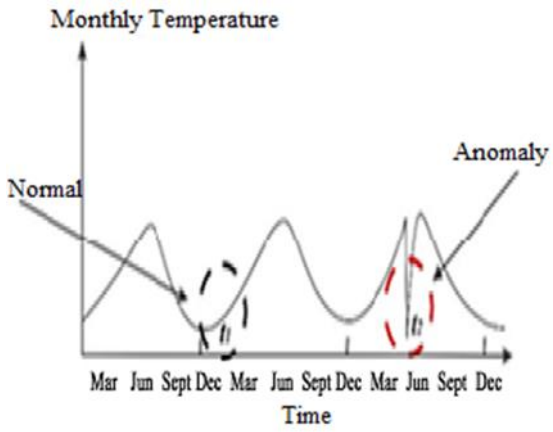

Fig:2 Contextual Anomalies

Collective Anomaly

A set of data occurrences that is deviating with from entire data set. So it is said to collective anomaly. Here the individual data occurrence is not anomalous, but when the data instances are combined, then anomaly may occur. The concept of collective anomaly is shown in Fig 3.

ECG of human pulse role is shown in fig 3.There is a linear line, in the ECG work which is not related to other parts of work and it is on un wanted occurrence. It shows anomaly behavior compared to other parts. This says that no response conditions in pulse rate observation.

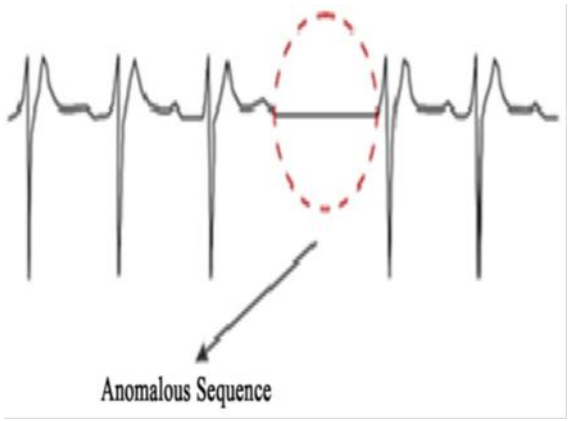

Fig:3 Collective Anomaly

Under cloud networks there are different techniques in identifying anomaly are machine learning, data mining, statistical analysis and Adaptive Anomaly Detection Systems.

4.1 Statistical Anomaly Detection Systems

Statistical method can identify anomaly by clearly noting the computations in the network traffic. This creates a profile for generated values in resenting their behavior.

There are two methods; the first method carries with normal or anomaly rules or signatures. Second one carries with the updates to the profile which is created with regular intervals. At the time of updates anomaly scores are calculated. Then the current and inception values are compared, if the present values are greater than the inception values, it is treated as anomalous and detected.

The advantage of this technique is, no prior knowledge of anomaly patterns or security risks is required.

4.2 Data Mining Based Anomaly Detection Systems

The data mining techniques are used for finding the patterns in data set whose behavior is abnormal or not expected. If it finds the patterns in the data set are abnormal, then malicious activity is said to occur in the system.

So the data mining techniques can be used to extract

Published by: The Mattingley Publishing Co., Inc.

or analyze the patterns in large data sets and checks the patterns belong to normal or abnormal activity. To detect anomalies in cloud computing environment, it uses the techniques namely clustering, classification, and association rules.

The clustering is a task of grouping the data based on certain parameters. Each cluster or group consists of patterns in the data sets that are similar to one another and different from other group or cluster's patterns. The clustering technique can be used for detecting intrusion in cloud data base.

The classification is the task of identifying the category of new instance. The category is also reported as label that will be data sets. This will help in identifying normal patterns and abnormal patterns in large data sets.

4.3 Machine Learning Based Anomaly Detection Systems

The system can improve its performance of task of detecting anomaly by learning the system behavior regularly. Initially, the system stores the information about normal behavior. When abnormal behavior is detected by the system, the machine learns new behavior and it stores the anomaly behavior pattern. Above technique generates a system that is used to improve performance of the program by past results. This technique is used on present system to improve performance over past results and extract information. Machine learning anomaly is based on different categories like Bayesian network, genetic algorithm and neural network.

Bayesian Network detect anomaly with the prior knowledge or signature of the data. This technique is combined with the statistical mechanism for detecting anomaly in efficient and effective manner.

Neural Networks has the capability to understand the patterns that are not visible and can detect the abnormal patterns in the network traffic. The Neural network not only identifies the known attacks but also new patterns or unknown attacks. Genetic algorithms are used as early algorithm techniques such as mutation, selection and others. These processes are framed as bunch of rules from the information on network analysis by IDS.

4.4 Adaptive Anomaly Detection Systems

The cloud operators maintain the log of detected failure records. This log will help the AAD system to detect anomalies in cloud environment. It will also update its log records whenever new threat or anomaly has been identified. So the AAD system changes the system behavior from new results or detection from the cloud provider. This will help in detecting new anomalies or abnormal activities in the system.

A comparative analysis of cloud based intrusion detection system as shown in table 1.

Detection Technique	Data Source	Limitations/ Challenges	Benefits
Statistical ADS	Profiles creations form the network traffic.	High computational Overhead because of two profiles to be created and update anomaly scores to be calculated at regular activities.	No need of knowledge related to security risks.
Data Mining Based ADS	Network traffic, known attacks patterns.	Privacy for the information can be violated.	Meaningful(normal) information will be extracted from the cloud, that reduces infrastructure the storage cost.
Machine Learning Based ADS	User behaviour, Known attack patterns.	Can't detect the unknown attacks.	Simple to implement because attacks can be Detected based on prior knowledge.
Adaptive ADS	Network traffic, known attacks and audit data.	Complexity in this method for by learning new types of attacks.	Can detect known and unknown attacks.

V. 5. CONCLUSION

Security is an important issue related cloud computing environment. An efficient IDS will safeguard user's data in cloud. The two types of IDS mechanisms namely Misuse related Detection and which is detection based on actual deviations. Anomaly based detection methods were presented in this paper. An efficient IDS technique, an anomaly-based intrusion detection system was presented

along with the three taxonomy anomalies namely, point, contextual and collective anomalies and various anomaly based methods such as statistical anomaly, data mining process gives clarity of machine learning based actual deviations from the adaptive anomalies.

REFERENCES

[1]. Parveen Kumar, "Cloud Computing: Threats, Attacks and Solutions". International Journal of Emerging Technologies in Engineering Research (IJETER) Volume 4, Issue 8, August (2016).

[2]. M. H. Sqalli, F. Al-Haidari, and K. Salah, "Edos-shield-a two-steps mitigation technique against edos attacks in cloud computing," in Utility and Cloud Computing (UCC), 2011 Fourth IEEE International Conference on, pp. 49–56, IEEE, 2011

[3]. ApurvaShitoot, Sanjay Sahu, Rahul Chawda, "Security Aspects in Cloud Computing", IJETT, Volume 6 number 3 - Dec 2013.

[4]. KaziZunnurhain and Susan V. Vrbsky, "Security Attacks and Solutions in Clouds".

[5]. Dr.Nedhal A. Al-Saiyd, Nada Sail, "Data Integrity in Cloud Computing Security", Journal of Theoretical and Applied Information Technology, 31st December 2013. Vol. 58 No.3 [10] Rushikesh Vilas Belamkar, "Challenges and Security Issues in Cloud Computing", ISRJ, ISSN 2230-7850, Volume-4, Issue-2,March-2014.

[6]. Koushal Kumar. "Intrusion Detection and Prevention System in enhancing Security of Cloud Environment" International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 6, Issue 8, August 2017, ISSN: 2278 – 1323.

[7]. W. Voorsluys, J. Broberg, and R. Buyya, "Introduction to cloud computing," Cloud Computing, pp. 1-41, 2011.

[8]. Z. Mahmood, ―Cloud Computing: Characteristics and Deployment Approaches‖, 11th IEEE International Conference on Computer and Information Technology, pp. 121-126, 2011.

[9]. J. Weng and G. Qin, ―Network Intrusion Prevention Systems‖, JTB_Journal of Technology and Business, pp. 37-49, October 2007.

[10]. Arif Sari, "A Review of Anomaly Detection Systems in Cloud Networks and Survey of Cloud Security Measures in Cloud Storage Applications". Journal of Information Security, 2015, 6, 142-154

[11]. Cisco Network Solutions, 2015. http://www.cisco.com/go/ips

[12]. Hand, D.J., Mannila, H. and Smyth, P. (2001) Principles of Data Mining. The MIT Press, Cambridge.