©NIJEP Publications
OPEN ACCESS
ISSN: 2992-5487
NEWPORT INTERNATIONAL JOURNAL OF ENGINEERING AND PHYSICAL SCIENCES (NIJEP)
Volume 3 Issue 2 2023

Cloud based log management system and its vulnerabilities

Emeka Okorie , Iloka B. C., Nwaukwa Johnwendy and Anyaragbu Hope.

Department of Computer Science Tansian University Umunya, Nigeria.
***Corresponding Author E-mail: emeka.okorie@tansianuniversity.edu.ng**

ABSTRACT
Cloud computing service is a new computing paradigm in which people only need to pay for use of services without cost of purchasing physical hardware. For this reason, cloud computing has been rapidly developed along with the trend of Information Technology (IT) services. It is efficient and cost economical for consumers to use computing resources as much as they need or use services they want from cloud computing provider (Hodo et al., 2016). One major factor that made cloud computing most attractive is its robust capacity to manage and store unlimited size of resources. This as a result provoked the engagement of countless numbers of enterprises in both the public and private sectors on the platform for data and other resource management. This IT infrastructure is made up of three major models which are the software as a service, infrastructure as a service and platform as a service which collaborated for the effective management of data. However, due to the huge volume of confidential information it contains has remained a centre of attraction for hackers and criminals. Today, the advancement in technology has upgraded the cloud base platform as a multi-mesh distribution and service oriented paradigm, multi domain, multi-tenancies, and multiple user autonomous administrative infrastructures; but unfortunately these characteristics exposed the cloud based platform to many threats which queries the integrity, confidentiality and availability of its resources.
Keywords: Cloud, Computing, IT, Hackers and platform

INTRODUCTION

Over the years, cloud based infrastructures have recorded countless number of attack models ranging from wormhole, black hole, denial of service, man in the middle, IP spoofing, among others, and employed the models for ransomeware [1, 2, 3, 4], which subjects the organization to pay a given ransom to restore service. This problem has dominated the global cyber security studies over the years and has remained a major not to crack. Many solutions proposed over the years to solve this problem such as cryptographic technique [5] signature based approach [6, 7, 8], anomaly detection approach [9], focused on the security of the packet, without considering the cloud base server infrastructure, hence leaving it vulnerable to intruders. To solve this problem, there is need for intrusion detection systems which monitors the server and ensure that hackers do not gain access to it. Recently the use of Machine Learning (ML) has gain increased attention towards cyber security. ML is a branch of artificial intelligence [10] which employed series of mathematical algorithms to learn from data and solve regression or pattern recognition problems. This solution was used to solve the problem of cloud intrusion detection in [11], but despite the success flood attack was never considered. Flood attack is a type of denial of service attack which has threatened the integrity of cloud based infrastructures over the years and has remained a major challenge. This problem will be addressed in this research developing a multi-level intrusion detection algorithm which will employ machine learning technique to monitor, detect and control flood attack on targeted cloud based infrastructures [12].

© Okorie *et al*
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited

OPEN ACCESS
ISSN: 2992-5487

Aim of the Study

The aim of this research is to perform an empirical study of a cloud based log management system and identify its vulnerabilities.

Significance of the Study

The importance of this study cannot be overemphasized as it will benefit stakeholders both in the public and private sectors. Some of these benefactors are all cloud based management companies like the Alfresco cloud service limited, Researchgate, Destinet, Github among other collaborative platforms.

METHODS AND SYSTEM ANALYSIS

Research Methodology

The methodology used for the development of the new system is the Top down Design Methodology (TDDM). The TDDM was adopted to guide the system development starting from the problem down to the solution. This allowed an empirical study of the testbed to find out the problem and then propose a solution to solving it within a specified time frame.

The Case Study Cloud Platform

The case study cloud log management system considered for this research is the Alfresceo catering service limited with registration number 950624 and located at Garki, Abuja, Nigeria. The company is responsible for the management of multiple enterprise information over there cloud log management platform with services such as document management, enterprise collaboration, analytics and insights on process management, open source enterprise solutions for content management capabilities, etc.

This company was selected as a case study for this research due to the huge volume of information they management and hence made them a primary target for criminals and attackers. The geographical coordinates of the company is Lat 6.55717N; Long 3.36400E as shown in the figure 1:

Figure 1: Location of the Case study (Courtesy: Google Map)

Research Methods

Data collection

The main source of data collection used for this research is the Afresco Service Limited. The company provided 12gigabyte sample of their log management server records from 4th to 27th April, 2022. This data were used as the training dataset and also for the empirical study conducted shortly to find vulnerabilities on the server. The description of the data collected was presented in the table 1, while the data samples are presented in table 2;

© Okorie *et al*
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited

OPEN ACCESS
ISSN: 2992-5487

Table 1: Data attributes of Log files

Attributes	**Description**
Client IP	The IP address of the client machine
Client name	The name of the client
Data	Data of user access
Time	Time of user log
Server site	The name of the internet service as appeared on the server
Server computer name	Server name
Server IP addresses	The IP of the internet service provider
Serer port	Server port used for the data transmission
Client server URL stream	Targeted default web page of the site
Client server URL query	The client query which begins with "?"
Server client status	the status code returned by the user link
Server client size	The size of data transmitted in bytes
Client server method	Client methods of request like Get, POST or HEAD
Latency	Total time taken for client to perform action
Client server version	Protocol version like the HTTP
Client server host	Host header name
User agent	Browser type used by the client
Cookies	Contents with cookies
Referrers	Link from where clients jump to the site
Server client win32 status	The window status code

© Okorie *et al*
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited

OPEN ACCESS
ISSN: 2992-5487

Table 2: Server Log Files Data (Source: Afresco Services)

Server Name	Address	Res.	User	Date/Time	Latency
heladmsrv	172.16.3.147	0	Root	4/04/22 8:47:34 AM	00:06:15
heladmsrv	172.16.3.88	0	Root	4/04/22 8:54:02 AM	00:06:06
Admsrv	172.16.3.147	0	Root	4/04/22 9:17:40 AM	00:00:36
Pcshare	172.16.3.88	0	nobody	4/04/22 9:23:11 AM	00:00:12
Pcshare	172.16.3.88	0	nobody	4/04/22 9:23:23 AM	00:00:11
Pcshare	172.16.3.88	0	nobody	4/04/22 9:23:49 AM	00:00:11
Pcshare	172.16.3.88	0	nobody	4/04/22 9:27:10 AM	00:00:12
heladmsrv	172.16.3.88	0	Root	4/04/22 9:01:20 AM	00:27:04
Pcshare	172.16.3.88	0	nobody	4/04/22 9:43:34 AM	00:00:11
Afpsrv	172.16.3.147	0	hendrik	4/04/22 9:19:02 AM	00:44:22
heladmsrv	172.16.3.88	0	Root	4/04/22 9:28:25 AM	00:36:35
Pcshare	172.16.3.88	0	nobody	4/04/22 10:03:20 AM	00:00:11
heladmsrv	172.16.3.88	0	-1	4/04/22 10:08:04 AM	00:00:22
heladmsrv	172.16.3.88	0	Root	4/04/22 10:12:44 AM	00:02:19
heladmsrv	172.16.3.88	0	Root	4/13/22 10:22:01 AM	00:10:08
heladmsrv	172.16.3.147	0	Root	4/04/22 10:47:34 AM	00:06:15
heladmsrv	172.16.3.88	0	Root	4/04/22 10:54:02 AM	00:06:06
Admsrv	172.16.3.147	0	Root	4/04/22 11:17:40 AM	00:00:36
Pcshare	172.16.3.88	0	nobody	4/04/22 11:23:11 AM	00:00:12
Pcshare	172.16.3.88	0	nobody	4/04/22 11:23:23 AM	00:00:11
Pcshare	172.16.3.88	0	nobody	4/04/22 11:23:49 AM	00:00:11
Pcshare	172.16.3.88	0	nobody	4/04/22 11:27:10 AM	00:00:12
heladmsrv	172.16.3.88	0	Root	4/04/22 12:01:20 PM	00:27:04
Pcshare	172.16.3.88	0	nobody	4/04/22 12:43:34 PM	00:00:11
Afpsrv	172.16.3.147	0	hendrik	4/04/22 13:19:02 PM	00:44:22
Pcshare	172.16.3.88	0	nobody	4/04/22 14:03:20 PM	00:00:11
heladmsrv	172.16.3.88	0	Root	4/04/22 14:28:25 PM	00:36:35
heladmsrv	172.16.3.88	0	-1	4/04/22 15:08:04 PM	00:00:22
heladmsrv	172.16.3.88	0	Root	4/04/22 16:12:44 PM	00:02:19
heladmsrv	172.16.3.88	0	Root	4/13/22 16:22:01 PM	00:10:08

Data processing

Due to the huge volume of the data managed by the server and collected for the research, the need for processing was vital to remove bugs, malware and other unnecessary features which might compromise the new system integrity. The data processing was done using associate rule based mining algorithm adopted from [13] to sorts and removes data formats with jpg, gif, bmp, etc to ensure quality data for training.

Machine learning algorithm

The nature of the problem been solved is a pattern recognition type and hence Machine Learning (ML) algorithm is the appropriate solution to it. The ML is set of mathematical algorithm which can learn data and then solve pattern recognition problems. The algorithm is trained to learn and then develop the desired model for the detection of the user log information. The machine learning algorithm adopted to solve this problem is the artificial neural network.

© Okorie *et al*
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited

Artificial Neural Network (ANN) based threat detection algorithm

ANN is a machine learning algorithm biologically inspired from the model of the human brain. These ANN are built with neurons which have weights, bias function and activation functions to learn data and then solve problems. In this research, the neural network model in [12-14] was adopted and used to develop a security algorithm which identified threats from the incoming data request to the log server intelligently.

Threat control algorithm

The previous section discussed the threat detection algorithm which was developed by the neural network. Now that the threat was detected, the control algorithm was used to ensure that the threat do not penetrate to the sever and cause problem. This was achieved using simple rule based logic control algorithm which makes decision from the output of the threat detection classification to allow throughput or deny it.

Multi –Level Intrusion Detection System

The security of the log management server involved two steps which are the detection to detect the problem and then the control to isolate it from the targeted severs. This was achieved using the ANN and rule base logic control algorithm to train all incoming log request to detect threat and prevent it from the targeted server. The detection algorithm monitors all data logs to the server and when abnormally is detected, the control enable access denial to the log and hence protect the server.

System Analysis

The system analysis provided the analysis of the existing system and then the analysis of the new system as highlighted and discussed below;

The existing system analysis (Empirical Study)

To analyze the existing system, an empirical study of the testbed was conducted via data collection and systematic analysis. The workflow of the testbed was presented in the figure 3.2;

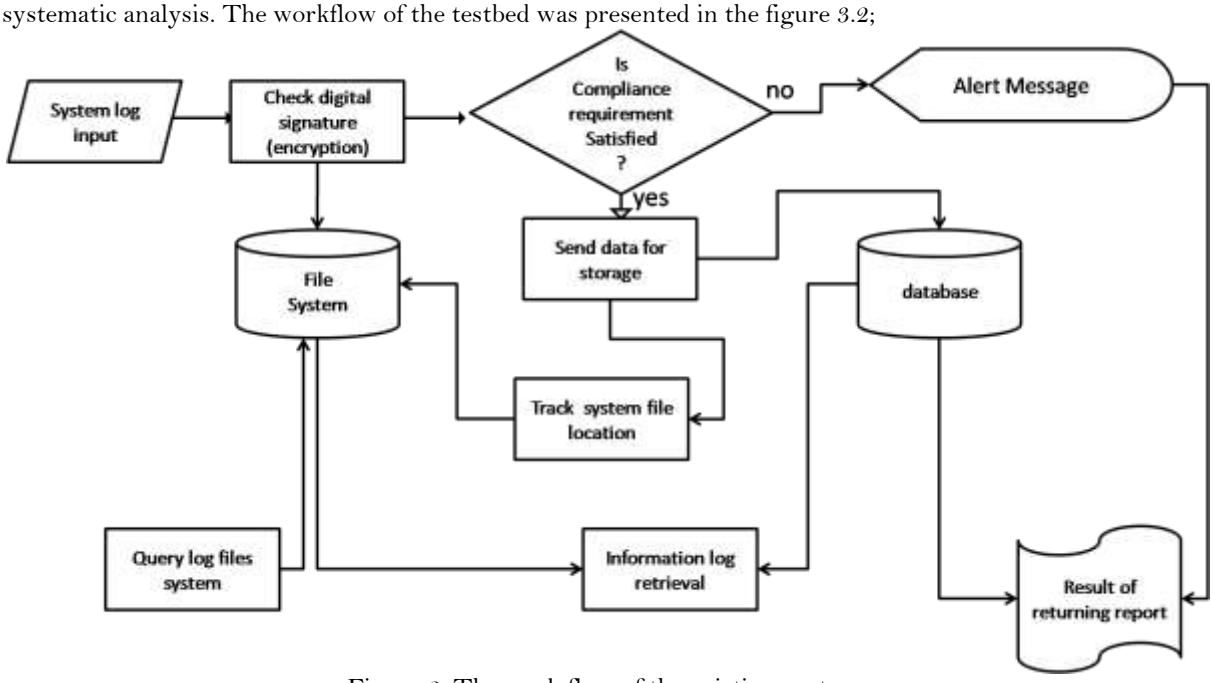

Figure 2: The work flow of the existing system

From the workflow diagram in figure 2, the systems log information from the enterprise users logs into the server which authenticates the access using standard encryption algorithm to check the digital signature for compliance and then access to the cloud. The encryption algorithm makes use of cipher key to encode the packet whose the key is used by the server to check and then allow access to server, but is the signature did not match, then alert is activated for threat notification on the server. The technical problems of standard encryption algorithms however are lack of intelligence, the key can be guessed correctly, the security focused on the data and not the cloud infrastructure, etc.

Analysis of the new system

Having analyzed the behaviour of the log management server in the previous section and identifying the technical problems with it, the new system developed was analyzed using the logical flow chart in figure 3;

© Okorie *et al*
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited

OPEN ACCESS
ISSN: 2992-5487

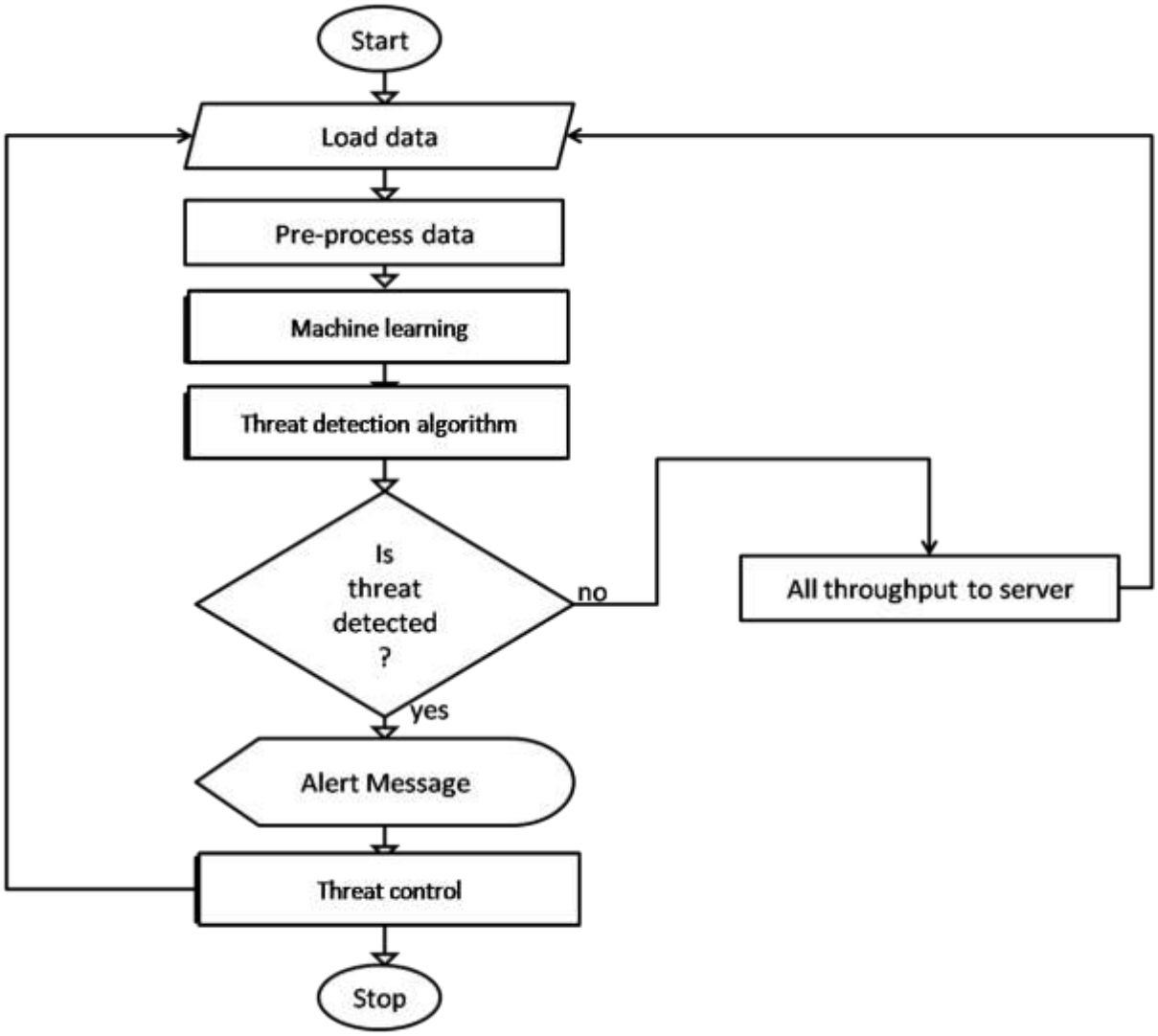

Figure 3: The process flow chart

The figure 3 presented the system flow chart which showed the logical flow between each of the processes used to achieve the new system. The data collected was loaded for processing to remove noise attributed with image formats. Then the neural network identified the data and train for the detection algorithm which was used to detect any intrusion to the server and control via access denial to the server.

CONCLUSION

Cloud computing technology provides human many advantages such as economical cost reduction and effective of IT based resources. Unfortunately the expansion of the technology equally expands its vulnerabilities to threats and has remained a very big challenge over the years.

REFERENCES

1. Amirreza Zarrabi, (2012). Research on Internet Intrusion Detection System Service in a Cloud, appear in International Journal of Computer Science Issues, Vol. 9, Issue 5, No 2, September 2012, ISSN (Online): 1694-0814
2. Bakshi A., Yogesh B, (2010)"Securing cloud from DDOS Attacks using Intrusion Detection System in Virtual Machine", Second International Conference on Communication Software and Networks, pp. 260-264.
3. Bharadwaja S., Sun W., Niamat M., Shen F., (2011) "Collabra: A Xen Hypervisor based Collaborative Intrusion Detection System", Eighth International Conference on Information Technology: New Generations, pp. 695-700.
4. Binita S. (2021) "Comparative Analysis of classification algorithms for intrusion detection"North Dakota State University of Agriculture and Applied Science; Master's Thesis.
5. Corbató, F.J., Daggett, M.M., Daley, R.C. (1962). An experimental time-sharing system. *SJCC Proceedings*. MIT. Retrieved from http://larch-www.lcs.mit.edu:8001/~corbato/sjcc62/
6. Cryptoclarity (2009). Encrypted storage and key management for the cloud. Retrieved from

© Okorie *et al*

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited

7. Debar H., M. Dacier, and A. Wespi (1999). Towards a Taxonomy of Intrusion Detection Systems, Int'l J. Computer and Telecommunications Networking, vol. 31, no. 9, pp. 805–822,1999.
8. Dhage S., Meshram B., Rawat R., Padawe S., Paingaokar M., Misra A., (2011) "Intrusion Detection System in Cloud Computing Environment", International Conference and Workshop on Emerging Trends in Technology (ICWET), pp. 235-239.
9. Doyen Sahoo, Chenghao Liu, and Steven CH Hoi. (2019) malicious url detection using machine learning: A survey. arXiv preprint arXiv:1701.07179.
10. National Institute of Standards and Technology (2012).*The NIST definition of cloud computing.* Retrieved from
11. Nils van noort (2020) "machine learning for monitoring attacks in the cloud"; 33rd twente student conference on it july 3rd, 2020, enschede, the netherlands.
12. Olofin B.B (2021a) "Key Scheduling Algorithm Assessment of Cryptographic Strength" IJESAT; Vol 8; Issues 3; pp 15-22.
13. Olofin B.B. (2021b) "A review of recent trends in intelligent agent technology" IJICST; Vol 6; Issue 2; pp. 90-98.
14. Phillip George Efthimion, Scott Payne, and Nicholas Proferes. Supervised machine learning bot detection techniques to identify social twitter bots. SMU Data Science Review, 1(2):5, 2018.

Emeka Okorie, Iloka B. C., Nwaukwa Johnwendy and Anyaragbu Hope.(2023). Cloud based log management system and its vulnerabilities. *NEWPORT INTERNATIONAL JOURNAL OF ENGINEERING AND PHYSICAL SCIENCES (NIJEP)* 3(2): 100-106.

© Okorie *et al*
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited