©NIJEP
Publications

OPEN ACCESS
ISSN: 2992-5487

NEWPORT INTERNATIONAL JOURNAL OF ENGINEERING AND PHYSICAL SCIENCES (NIJEP)

Volume 3 Issue 2 2023

A review on Cloud computing and related technologies

[1]Emeka Okorie, [2]Yateghtegh. S. and [1]Okoh Chris

[1]Department of Computer Science Tansian University Umunya, Nigeria
[2]CEO Skyhub Nigeria Limited
*Corresponding Author E-mail: emeka.okorie@tansianuniversity.edu.ng

ABSTRACT
Cloud computing is a model for enabling network users on-demand access to a shared pool of configurable computing resources that can be rapidly provisioned and released to the client without direct service provider interaction. It can also be defined as the use of computer technology that harnesses the processing power of many inter-networked computers while concealing the structure that is behind it. Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network. The name comes from the use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. It is essentially the use of virtual servers made available over the internet. Cloud computing entrusts remote services with a user's data, software and computation.

Keywords: Cloud computing, IT, Technologies and Infrastructures

INTRODUCTION

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 [1]. 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 [2-4]. 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 [5-6]. 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 [7-8].

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 [9-10], 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 [11-13], signature based approach [14-17], anomaly detection approach [18], 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 [19-24] 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 [25-27], 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.

Review of Relevant Literatures

[28], researched on a fully automated deep packet inspection and verification system with machine learning. The study used machine learning algorithm such as the K-Nearest Neighbour (K-NN), Support Vector Machine (SVM), logistic regression and Naive bayes, K-mean clustering. The result showed that machine learning solution to cyber security are promising with high accuracy, however the study never considers flood attack. [29], researched on supervised machine learning bot detection system to identify social twitter bots. The study engaged multiple ML algorithms like K-Nearest Neighbour (K-NN), Support Vector Machine (SVM), logistic regression and Naive bayes, K-mean clustering for the detection of botnets. The solution when tested showed high detection accuracy, but flood attack was not considered in the study. [30], presented a research on threat detection based on ML approaches. The solution considered K-Nearest Neighbour (K-NN), Support Vector Machine (SVM), logistic regression and Naive bayes, K-mean clustering. The result showed good threat detection performance of comparative threat datasets, however the performance can be improved with artificial neural network.

[31], applied ML predictive analysis to SQL injection detection and prevention of web based application systems. The study developed a ML based security solution to protect cloud based platforms. The result when tested showed good result, but flood attack was not considered. [32], presented a research on malicious URL detection system using ML. The study considered the various ML algorithms which have been adopted to solve the problem of intrusion detection system such as Artificial neural network, K-Nearest Neighbour (K-NN), Support Vector Machine (SVM), logistic regression and Naive bayes, K-mean clustering. The comparative study showed that they all achieved good and acceptable threat detection accuracy, but can be improved with more training dataset and better configuration. The study also revealed that neural network achieved the best result when compared to others. [33], used ML to monitor attacks on cloud. The study reviewed the various impacts of ML on cloud based cyber security. They also performed a qualitative analysis of the ML algorithm considered for the study such as K-Nearest Neighbour (K-NN), Support Vector Machine (SVM), logistic regression and Naive bayes, K-mean clustering and then recommended its adopted for cyber security solutions.

Cloud computing Overview

Cloud computing is a model for enabling network users on-demand access to a shared pool of configurable computing resources that can be rapidly provisioned and released to the client without direct service provider interaction [34]. It can also be defined as the use of computer technology that harnesses the processing power of many inter-networked computers while concealing the structure that is behind it [35]. Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network [36]. The name comes from the use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams [37]. It is essentially the use of virtual servers made available over the internet. Cloud computing entrusts remote services with a user's data, software and computation. A slow migration towards this has been going on for several years, mainly due to the infrastructure and support costs that go into standalone hardware. It can simply be broken down to a browser based application that is hosted on a remote server. Cloud computing as it relates to Internet technology is all around us. Accessing mails and searching for information online are all due to the power of processing technology that exists at a distant location without the user knowing about it.

Explanations of common terminologies used in cloud computing

Some of the terms associated with cloud computing include the following [38].

i. **Back office:** Making technology easier for customers saves companies money, by taking the technical issues out of the equation so that businesses can focus their energy on creating a superior product or service. This is commonly known as back office tasks. They are generally rudimentary data parsing procedures that are time consuming as well as tedious. Back office applications are software that an organization uses to administer operations that are not related to any direct sales effort and interfaces that are not seen by consumers. An example of a back office service that is out today is Amazon's Web Services platform.

ii. **Web 2.0:** Web 2.0 cloud computing is a blanket term, but it is usually associated with some type of social networking technology – that is, a large number of social users that are interconnected via their relations with the people and things they find interesting. It describes a second generation of the World Wide Web focused on the ability for people to collaborate and share information online. It basically refers to the transition from static HTML web pages to a more dynamic web that is more organized. One of the biggest Web 2.0 companies today is Facebook.

iii. **On-demand computing:** On demand computing is a business terminology and would refer to back office processing power, for example a remote data processing center that processes payroll functions for a company located thousands of miles away. It is an enterprise-level model of technology and computing in which resources are provided on an as-needed and when-needed basis. This type of model was created to overcome the challenge to enterprises in being able to meet fluctuating demands.

iv. **Thin client:** A thin client is a term used for a terminal that connects to the cloud. This could be a computer, a cell phone or even an mp3 player. It can be referred to as software as well. As long as the device can connect to the cloud, it is known as a thin client for all intents and purposes. The meaning behind its being "thin" is that it does not require much processing power to be a client to the cloud itself.

v. **Workload migration:** Workload migration is the concept of optimizing server farm technology to be data and energy efficient (Gens, 2008). With so much processing ability coupled with an enormous amount of power consumption, companies managing server farm technology are finding that they need IT people who are well versed in workload migration technology to be able to manage all that this entails. Some cloud computing companies tout services to help companies with work load migration, offering services that assist their clients with the "internal cloud" process.

vi. **Server farm:** A server farm is a cluster of computers whose sole purpose is to provide processing power greater than what a single machine would be able to do on its own [39]. A perfect example of this would be what companies use for web hosting of individual websites. Even though there is one website the server farm provides failover capability in case something was to happen to any single machine hosting the website. It is ideal for server farms to be located near a reliable source of power.

Key characteristics of cloud computing

The National Institute of Standards and Technology's (NIST) definition of cloud computing identifies five essential characteristics [40]:

- **On-demand self-service:** A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.
- **Broad network access:** Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).
- **Resource pooling:** The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.
- **Rapid elasticity:** Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.
- **Measured service:** Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

Cloud computing can also be said to exhibit the following characteristics:

- Agility improves with users' ability to re-provision technological infrastructure resources [41].
- Application programming interface (API) accessibility to software that enables machines to interact with cloud software in the same way the user interface facilitates interaction between humans and computers. Cloud computing systems typically use REST (representational-based transfer)-based APIs [42].

- Cost is claimed to be reduced and in a public cloud delivery model capital expenditure is converted to operational expenditure [43]. This is purported to lower barriers to entry, as infrastructure is typically provided by a third-party and does not need to be purchased for one-time or infrequent intensive computing tasks.
- Device and location independence enable users to access systems using a web browser regardless of their location or what device they are using (e.g., PC, mobile phone)
- Virtualization technology allows servers and storage devices to be shared and utilization to be increased. Applications can be easily migrated from one physical server to another.
- Multi tenancy as shown in the work in [44], enables sharing of resources and costs across a large pool of users thus allowing for:
  - Centralization of infrastructure in locations with lower costs (such as real estate, electricity, etc.)
  - Peak-load capacity increases (users need not engineer for highest possible load-levels)
  - Utilisation and efficiency improvements for systems that are often only 10–20% utilised
- Reliability is improved if multiple redundant sites are used, which makes well-designed cloud computing suitable for business continuity and disaster recovery
- Scalability and elasticity via dynamic ("on-demand") provisioning of resources on a fine-grained, self-service basis near real-time without users having to engineer for peak loads [45].
- Performance is monitored and consistent and loosely coupled architectures are constructed using web services as the system interface according to the work in [46].
- Security could improve due to centralization of data, increased security-focused resources, etc [47]. But concerns can persist about loss of control over certain sensitive data, and the lack of security for stored kernels. [48], stated that security is often as good as or better than other traditional systems, in part because providers are able to devote resources to solving security issues that many customers cannot afford. However, the complexity of security is greatly increased when data is distributed over a wider area or greater number of devices and in multi-tenant systems that are being shared by unrelated users. In addition, user access to security audit logs may be difficult or impossible. Private cloud installations are in part motivated by users' desire to retain control over the infrastructure and avoid losing control of information security.
- Maintenance of cloud computing applications is easier, because they do not need to be installed on each user's computer and can be accessed from different places.

Cloud computing was chosen for the deployment of this application due to these characteristics and benefits.

Related Technologies to Cloud Computing

Cloud computing is often compared to the following technologies, each of which shares certain aspects with cloud computing:

- **Grid Computing:** Grid computing is a distributed computing paradigm that coordinates networked resources to achieve a common computational objective [49-50]. It is the collection of loosely coupled, heterogeneous, and geographically dispersed computer resources from multiple locations to reach a common goal. It is a distributed system with non-interactive workloads that involve a large number of files. The development of Grid computing was originally driven by scientific applications which are usually computation-intensive. Cloud computing is similar to Grid computing in that it also employs distributed resources to achieve application-level objectives. However, cloud computing takes one step further by leveraging virtualization technologies at multiple levels (hardware and application platform) to realize resource sharing and dynamic resource provisioning.
- **Utility Computing:** it involves the provision of resources on-demand and customers are charged based on usage rather than a flat rate. Cloud computing can be perceived as a realization of utility computing. It adopts a utility-based pricing scheme entirely for economic reasons. With on-demand resource provisioning and utility based pricing, service providers can truly maximize resource utilization and minimize their operating costs.
- **Virtualization:** Virtualization is a technology that abstracts away the details of physical hardware and provides virtualized resources for high-level applications. A virtualized server is commonly called a virtual machine (VM). Virtualization forms the foundation of cloud computing, as it provides the capability of pooling computing resources from clusters of servers and dynamically assigning or reassigning virtual resources to applications on-demand.
- **Autonomic Computing:** Originally coined by IBM in 2001, autonomic computing aims at building computing systems capable of self-management, i.e. reacting to internal and external observations

without human intervention [51]. The goal of autonomic computing is to overcome the management complexity of today's computer systems. Although cloud computing exhibits certain autonomic features such as automatic resource provisioning, its objective is to lower the resource cost rather than to reduce system complexity.

In summary, cloud computing leverages virtualization technology to achieve the goal of providing computing resources as a utility. It shares certain aspects with grid computing and autonomic computing but differs from them in other aspects. Therefore, it offers unique benefits and imposes distinctive challenges to meet its requirements. This work chose cloud computing for the application integration based on these features.

Cloud computing architecture

This section describes the architectural, business and various operation models of cloud computing.

i. A layered model of cloud computing

Generally speaking, the architecture of a cloud computing environment can be divided into 5 layers:

a. **Facility Layer:** Heating, ventilation, air conditioning (HVAC), power, communications, and other aspects of the physical plant comprise the lowest layer, the facility layer.

b. **The hardware layer:** This layer is responsible for managing the physical resources of the cloud, including physical servers, routers, switches, power and cooling systems. In practice, the hardware layer is typically implemented in data centres. A data centre usually contains thousands of servers that are organized in racks and interconnected through switches, routers or other fabrics [52]. Typical issues at hardware layer include hardware configuration, fault tolerance, traffic management, power and cooling resource management.

c. **The infrastructure layer:** Also known as the virtualization layer, the infrastructure layer creates a pool of storage and computing resources by partitioning the physical resources using virtualization technologies such as Xen, KVM and VMware. The infrastructure layer is an essential component of cloud computing, since many key features, such as dynamic resource assignment, are only made available through virtualization technologies.

d. **The platform layer:** Built on top of the infrastructure layer, the platform layer consists of operating systems and application frameworks. The purpose of the platform layer is to minimize the burden of deploying applications directly into VM containers. For example, Google App Engine operates at the platform layer to provide API support for implementing storage, database and business logic of typical web applications.

e. **The application layer:** At the highest level of the hierarchy, the application layer consists of the actual cloud applications. Different from traditional applications, cloud applications can leverage the automatic-scaling feature to achieve better performance, availability and lower operating cost. Compared to traditional service hosting environments such as dedicated server farms, the architecture of cloud computing is more modular. Each layer is loosely coupled with the layers above and below, allowing each layer to evolve separately. This is similar to the design of the OSI

ii. Service Models

f. **Software as a Service (SaaS):** The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a program interface [52]. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings. Examples of SaaS include: Google Apps, Microsoft office 365, Onlive, GT Nexus, etc.

g. **Platform as a Service (PaaS):** The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.

h. **Infrastructure as a Service (IaaS):** The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components (e.g., host firewalls). Offering virtualized resources (computation, storage, and communication) on demand is known as Infrastructure as a Service (IaaS) [53]. Infrastructure services are considered to be the bottom layer of cloud computing systems. Amazon Web Services mainly offers IaaS, which in the case of its EC2 service means offering virtual machines with a software stack that can be customized similar to how an ordinary physical server would be customized. Users are given privileges to perform numerous activities to the server, such as: starting and stopping it, customizing it by installing software packages, attaching virtual disks to it, and configuring access permissions and firewalls rules [54].

The Cloud and Types of Intrusion Threats

There are several of types of cloud security and each of them has its own field of problems to tackle and solve. In this section, we present a couple of them that are most relevant for this thesis. Network security is, in simple terms, the practice of protecting a network from being accessed by unauthorized actors. The main point is to deflect malicious activity in the network and the connections between the devices [55]. The most commonly used network protocol today for transmitting traffic flow is Internet Protocol (IP). The network consists, of course, of many layers and all of the layers are vulnerable to attacks. The two most frequently occurring types of malicious activity on a network are network intrusion and distributed denial of service (DDoS) attacks [56]. Network intrusion can be exploited by introducing unwanted packages to the network whose purpose is to consume resources of the network, interfere with the functions of the resources, and gain information and knowledge about the system that can be used later for more severe attacks [57]. The main objective of application security is to secure its functionalities and to ensure they are implemented in a way that protects them from attacks. Software updates and testing are ways to enhance the security of the application [58]. In recent studies, the area in cloud security with most breaches has particularly been web applications. Testing, as mentioned, is the key ingredient in preserving application security and this should be repeated throughout the lifespan of the application. The three phases of an application are generally development, quality assurance, and production [59]. While most organizations perform diligent testing and checking for vulnerabilities during the first and second stage, the actual problems occur in the production stage. Continuous testing post-production is important to maintain integrity concerning the security aspects of applications [60]. As vast amounts of our data and information are stored in digital environments, such as clouds, the protection of these storage spaces is as important as protecting any physical storage units [61].

The last aspect of cloud security to describe is endpoint security. Endpoint security in short is the practice of securing the entry point of the network from being exploited or attacked. By end- and entry points we mean various devices such as laptops, mobile phones, printers, desktops, smart watches and any other "smart" device that could be accessed via a network [62].

Endpoint security is not a new concept but sometimes an overlooked one. Today, the average antivirus software might not be sufficient to address the complexity of malicious attacks encountered every day. Data is the most valuable asset in most markets and companies, which means there need to be tools to protect it. Since the number of entry points in a 15 network can be several thousands, automatic detection systems and endpoint protections platforms (EPP) are the way in the future to keep up with the pace [8].

There are many types of cyber threats which includes malware, denial of service, phishing, web based attack, data breach, web application attack, man in the middle attack, distributive denial of service attack, botnet attack, blackhole attack among others as they are selected and discussed below;

i. Malware

Malware is an often-used method in cyber criminality. Malware is malicious software used to accomplish identity theft, cyber espionage, and disruptions in systems. Malware appears in the form of viruses, Trojan horses, and ransomeware [10]. In contrast to a typical bug, malware is an attempt designed to cause harm. The usage of malware has seen a shift from consumer targets to business targets. Further statistics show that 50% of malware attacks were designed to steal personal information and that 71% of businesses which were targeted saw malicious software spreading between employees [13].

A new alarming trend in malware is the concept of Malware-as-a-Service (MaaS). Contrary to Software-as-a-Service (SaaS), MaaS is a black-market business for criminals that sell and rent malware to other criminals [20]. This enables people without the technical skills to perform cyber attacks. The ease in how almost anybody can launch an attack via MaaS organizations has led to an increase in botnets [23].

ii. Web-based Attacks

Web-based services are increasing in the world and this offers an appealing opportunity to be exploited by malicious actors. Web-based attacks are a way to inject malicious script of false URLs to users, and by that redirect the user to a desired webpage [40]. This method can also be used to have the victim download malicious files and to inject harmful content into web pages that are trusted but have compromised features. An overload of login requests with usernames and passwords, such as brute-force attacks, affects the key features of cloud security, such as availability of web sites as well as the confidentiality and integrity of the information, accessible in the web services.

iii. Phishing

The COVID-19 pandemic caused a massive increase in regard to phishing attacks. ENISA reports [8], that the number of phishing scams in one month rose with 667%. Phishing attacks are most often seen attached to emails. The email is meant to look trustworthy but contains malicious attachments and links. These emails try to tap into human emotions and cause human errors as a consequence. In order to achieve that, keywords such as "payment" are used to evoke rushed and ill-advised decisions from the victim [7]. Phishing emails with fake error messages and fake update suggestions are typical examples of phishing attacks [4]. With usage of AI, mitigation is possible with pattern seeking and learning abilities. For instance, AI can recognize typical dialog patterns, interaction characteristics, and grammar and IPtax anomalies and it can scan images attached to emails to detect fraudulent links and login requests [9].

iv. Web Application Attack

The utilization of the internet has led to more web applications and these play a vital role in companies' ability to provide services. Web applications rely heavily on databases that have the task of storing and transmitting the requested data to the user (ENISA, 2020). The typical attack methods used to threaten databases are SQL injections (SQLi) and Cross-Site Scripting (XSS). SQLi attacks exploit vulnerabilities in web security by injecting malicious code to database queries, thus gaining access to data and altering it in a harmful way [10]. XSS attacks work with the same principle, but here the malicious code is planted in web based applications and websites that can then redirect the end user to malicious websites [17].

v. Distributed Denial of Service (DDoS) Based Flood Attack

DDoS is the phenomenon where a user is not able to access certain data or resources in a system. To accomplish this, the attackers flood the host network or target with requests and traffic, resulting in the system crashing for not being able to respond (CISA, 2009). DDoS attacks are not unfamiliar to cloud security experts, but the techniques used by malicious actors are becoming more advanced. Trends in DDoS attacks show that over half of the attacks last for less than 15 minutes and that multiple attack vectors are used at the same time [20].

vi. Data Breach

A data breach is a type of incident where data and/or parts of an information system is accessed without authorization. Once having accessed the breached system, the attacker can misuse and destroy the data. Data breaches and human errors are closely linked, since many vulnerabilities and unintentional exposure of data are due to insufficient implementation and configuration of the system itself [40].

Flood Attack on Cloud Infrastructure

Cloud computing is one of the high-demand services and prone to numerous types of attacks due to its Internet based backbone. Flooding based attack is one such type of attack over the cloud that exhausts the numerous resources and services of an individual or an enterprise by way of sending useless huge traffic. The nature of this traffic may be of slow or fast type. Flooding attacks are caused by way of sending massive volume of packets of TCP, UDP, ICMP traffic and HTTP Posts. The legitimate volume of traffic is suppressed and lost in traffic flooding traffics. Early detection of such attacks helps in minimization of the unauthorized utilization of resources on the target machine. Various inbuilt load balancing and scalability options to absorb flooding attacks are in use by cloud service providers up to ample extent still to maintain QoS at the same time by cloud service providers is a challenge.

IP Flood is a common form of Denial of Service Attack (DDoS) that can target any system connected to the Internet and providing Transmission Control Protocol (TCP) services (e.g. web server, email server, files transfer). A IP flood is a type of TCP state exhaustion attack that attempts to consume the connection state tables present in many infrastructure components, such as load balancers, firewalls, Intrusion Prevention Systems (IPS), and the application servers themselves. This type of DDoS attack can take down even high-capacity devices capable of maintaining millions of connections. Unlike other types of DDoS attacks, IP flood DDoS attacks are not intending to use up all of the host's memory, but rather, to exhaust the reserve of open connections connected to a port, from individual and often phony IP addresses. IP floods are often called "half-open" attacks because this type of DDoS attack intends to send a short burst of IP messages into the ports, leaving insecure connections open and available, and often resulting in a complete server crash.

Machine Learning for Wireless Network Security

Wireless network anomaly detection systems are human-independent. They detect anomalies by revealing abnormal characteristics in a wireless network over a period of time. The effectiveness of this technique is its capabilities to differentiate between normal and abnormalities within a network. Machine Learning (ML) can provide wireless network threat prevention methods to detect current, new and subtle attacks without extensive human-based training or intervention. It is defined as a set of methods that can automatically detect patterns to predict future data trends [50]. Whilst a large number of machine learning techniques exist, the fundamental operation of all of them relies upon optimal feature selection. These features are the metrics which will be used to detect patterns and trends. For example, one feature of a network is the packet size: machine learning techniques may monitor the packet size over time and generate distributions from which conclusions may be drawn regarding an intrusion. This section reviews some of the numerous machine learning techniques which can be used to train the feature extracted vectors for wireless network security.

Artificial Neural Networks (ANN)

The field of neural networks is a subarea of machine learning. The human brain has about 100 billion nerve cells. We humans owe our intelligence and our ability to learn various motor and intellectual capabilities to the brain's complex relays and adaptivity. For many centuries biologists, psychologists, and doctors have tried to understand how the brain functions. Around 1900 came the revolutionary realization that these tiny physical building blocks of the brain, the nerve cells and their connections are responsible for awareness, associations, thoughts, consciousness, and the ability to learn. Human brain is one of the best 'machines' we know for learning and solving problems. Within the machine learning fields, there is an area often referred to as brain-inspired computation. The brain-inspired technique is indeed inspired by how our human brain works. It is believed that the main computational element of our brain is neuron. The complex connected network of neurons forms the basis of all the decisions made based on the various information gathered. This is exactly what Artificial Neural Network technique does [40].

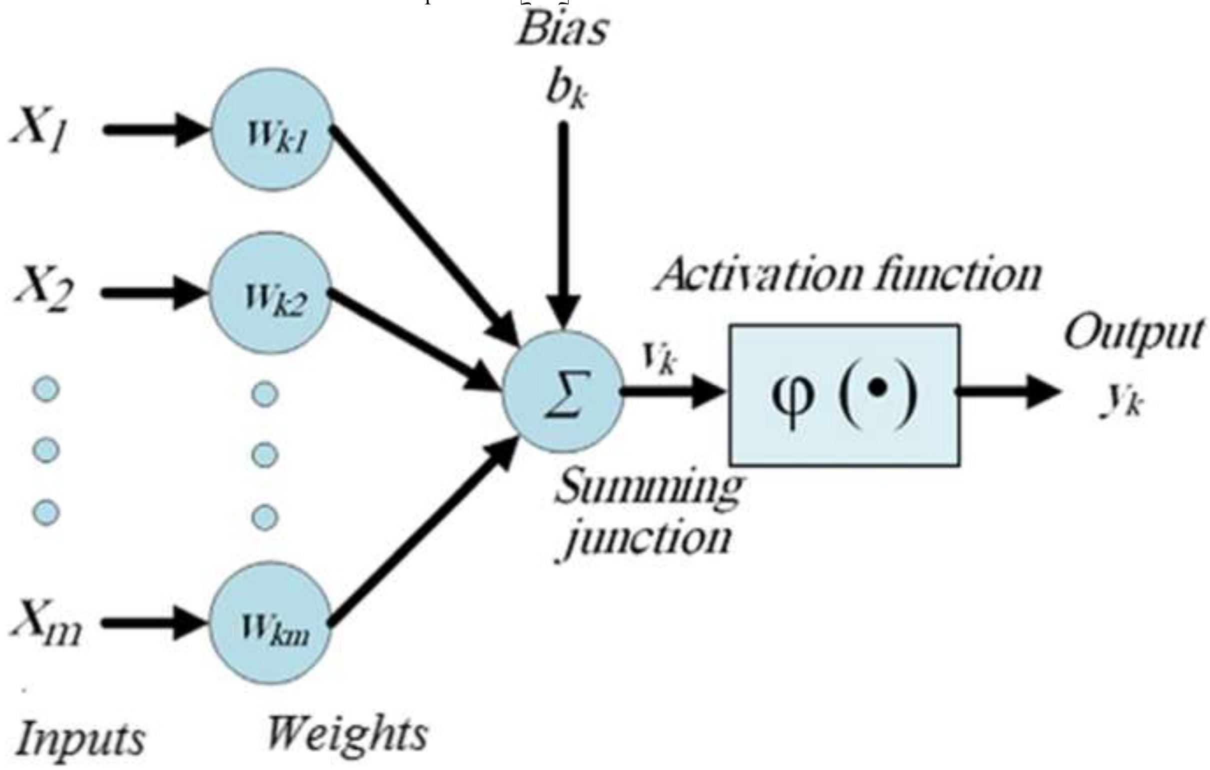

Figure 1: Structure of a Biological Neuron .

An artificial neural network is a system of hardware or software that is patterned after the workings of neurons in the human brain and nervous system. Artificial neural networks are a variety of deep learning technology which comes under the broad domain of Artificial Intelligence [60]. However, the tools used for modeling, namely mathematics, programming languages, and digital computers have very little in common with the human brain. With artificial neural networks, the approach is different. Starting from knowledge about the function of natural neural networks, we attempt to model, simulate, and even reconstruct them in hardware. According to [20], the efficiency of a neural network is a function of how extremely adaptive it is to learning very quickly. Each node weighs the importance of the input it receives from the nodes before it. The inputs that contribute the most towards the right output are given the highest weight. There are many different types of neural networks which function on the same principles as the nervous system in the human body. Howard Rheingold said, "The neural network is this kind of technology that is not an algorithm, it is a network that has weights on it, and you can adjust the weights so that it learns.

A neural network has a large number of processors. These processors operate parallel but are arranged as tiers. The first tier receives the raw input similar to how the optic nerve receives the raw information in human beings. Each successive tier then receives input from the tier before it and then passes on its output to the tier after it. The last tier processes the final output. Small nodes make up each tier. The nodes are highly interconnected with the nodes in the tier before and after. Each node in the neural network has its own sphere of knowledge, including rules that it was programmed with and rules it has learnt by itself [40].

Defence on Cloud-Hosted Machine Learning Model Attacks

In this section, the summary of different defensive strategies against attacks on cloud-hosted ML models as described above in thematic analysis.

MiniONN

DNNs are vulnerable to model inversion and extraction attacks. [6], proposed that without making any changes to the training phase of the model it is possible to change the model into an oblivious neural network. They make the nonlinear function such as tanh and sigmoid function more flexible, and by training the models on several datasets, the authors demonstrated significant results with minimal loss in the accuracy. In addition, they also implemented the offline precomputation phase to perform encryption incremental operations along with the SIMD batch processing technique.

ReDCrypt

A reconfigurable hardware-accelerated framework is proposed by [50], for protecting the privacy of deep neural models in cloud networks. The authors perform an innovative and power-efficient implementation of Yao's Garbled Circuit (GC) protocol on FPGAs for preserving privacy. The proposed framework is evaluated for different DL applications, and it has achieved up to 57-fold throughput gain per core.

Arden

To offload the large portion of DNNs from the mobile devices to the clouds and to make the framework secure, a privacy-preserving mechanism Arden is proposed by [34]. While uploading the data to the mobile-cloud perturbation, noisy samples are included to make the data secure. To verify the robustness, the authors perform rigorous analysis based on three image datasets and demonstrated that this defense is capable to preserve the user privacy along with inference performance.

Image Disguising Techniques

While leveraging services from the cloud GPU server, the adversary can realize an attack by introducing malicious created training data, perform model inversion, and use the model for getting desirable incentives and outcomes. To protect from such attacks and to preserve the data as well as the model [13], proposed an image disguising mechanism. They developed a toolkit that can be leveraged to calibrate certain parameter settings. They claim that the disguised images with block-wise permutation and transformations are resilient to GAN-based attack and model inversion attacks.

Homomorphic Encryption

For making the cloud services of outsourced MLaaS secure, [14], proposed a privacy-preserving framework using homomorphic encryption. They trained the neural network using the encrypted data and then performed the encrypted predictions. The authors demonstrated that by carefully choosing the polynomials of the activation functions to adopt neural networks, it is possible to achieve the desired accuracy along with privacy-preserving training and classification. In a similar study, to preserve the privacy of outsourced biomedical data and computation on public cloud servers, [15], built a homomorphically encrypted model that reinforces the hardware security through Software Guard Extensions. They combined homomorphic encryption and Software Guard Extensions to devise a hybrid model for the security of the most commonly used model for biomedical applications, that is, LR. The robustness of the Secure LR framework is evaluated on various datasets, and the authors also compared its performance with state-of-the-art secure LR solutions and demonstrated its superior efficiency.

Pelican

[17], proposed three mutation-based evasion attacks and a sample-based collision attack in white-, gray-, and black box scenarios. They evaluated the attacks and demonstrated a 100% success rate of attack on Google's phishing page filter classifier, while a success rate of up to 81% for the transferability on Bitde fender Traffic Light. To deal with such attacks and to increase the robustness of classifiers, they proposed a defense method known as Pelican.

Rounding Confidences and Differential Privacy

[20], presented the model extraction attacks against the online services of BigML and Amazon ML. The attacks are capable of model evasion, monetization, and can compromise the privacy of training data. The authors also proposed and evaluated countermeasures such as rounding confidences against equation-solving and decision tree path finding attacks; however, this defense has no impact on the regression tree model attack. For the preservation of training data, differential privacy is proposed; this defense reduces the ability of an attacker to learn insights about the training dataset. The impact of both defenses is evaluated on the attacks

for different models, while the authors also proposed ensemble models to mitigate the impact of attacks; however, their resilience is not evaluated.

Increasing Entropy and Reducing Precision

The training of attack using shadow training techniques against black box models in the cloud-based Google Prediction API and Amazon ML models are studied by [40]. The attack does not require prior knowledge of training data distribution. The authors emphasize that in order to protect the privacy of medical-related datasets or other public-related data, countermeasures should be designed. For instance, restriction of prediction vector to top k classes, which will prevent the leakage of important information or rounding down or up the classification probabilities in the prediction. They show that regularization can be effective to cope with overfitting and increasing the randomness of the prediction vector.

Dropout and Model Stacking

In the study by [19], the authors created three diverse attacks and tested the applicability of these attacks on eight datasets from which six are similar as used by [20], whereas in this work, news dataset and face dataset is included. In the threat model, the authors considered black box access to the target model which is a supervised ML classifier with binary classes that was trained for binary classification. To mitigate the privacy threats, the authors proposed a dropout-based method which reduces the impact of an attack by randomly deleting a proportion of edges in each training iteration in a fully connected neural network. The second defense strategy is model stacking, which hierarchically organizes multiple ML models to avoid overfitting. After extensive evaluation, these defense techniques showed the potential to mitigate the performance of the membership inference attack.

Randomness to Video Analysis Algorithms

Hosseini et al. designed two attacks specifically to analyze the robustness of video classification and shot detection [30]. The attack can subtly manipulate the content of the video in such a way that it is undetected by humans, while the output from the automatic video analysis method is altered. Depending on the fact that the video and shot labels are generated by API by processing only the first video frame of every second, the attack can successfully deceive API. To deal with the shot removal and generation attacks, the authors proposed the inclusion of randomness for enhancing the robustness of algorithms. However, in this article, the authors thoroughly evaluated the applicability of these attacks in different video setting, but the purposed defense is not rigorously evaluated.

Neuron Distance Threshold and Obfuscation

Transfer learning is an effective technique for quickly building DL student models in which knowledge from a Teacher model is transferred to a Student model. However, [50] discussed that due to the centralization of model training, the vulnerability against misclassification attacks for image recognition on black box Student models increases. The authors proposed several defenses to mitigate the impact of such an attack, such as changing the internal representation of the Student model from the Teacher model. Other defense methods include increasing dropout randomization which alters the student model training process, modification in input data before classification, adding redundancy, and using orthogonal model against transfer learning attack. The authors analyzed the robustness of these attacks and demonstrated that the neuron distance threshold is the most effective in obfuscating the identity.

Table 1 Systematic Summary of Related Literatures

Author	Title	Technique	work done	Research gap/limitation
Uday and Manal (2016)	Fully automated deep packet inspection and verification system with machine learning.	Machine learning	The study used machine learning algorithm such as the K-Nearest Neighbour (K-NN), Support Vector Machine (SVM), logistic regression and Naive bayes, K-mean clustering.	The result showed that machine learning solution to cyber security are promising with high accuracy, however the study never considers flood attack
Phillip et al. (2018)	Supervised machine learning bot detection system to identify social twitter bots	Machine learning	The study engaged multiple ML algorithms like K-Nearest Neighbour (K-NN), Support Vector Machine (SVM), logistic regression and Naive bayes, K-mean clustering for the detection of botnets	The solution when tested showed high detection accuracy, but flood attack was not considered in the study.
Shailendra et al. (2017)	Threat detection based on ML approaches	Machine learning	The solution considered K-Nearest Neighbour (K-NN), Support Vector Machine (SVM), logistic regression and Naive bayes, K-mean clustering.	The result showed good threat detection performance of comparative threat datasets, however the performance can be improved with artificial neural network.
Solomon et al. (2017)	ML predictive analysis to SQL injection detection and prevention of web based application systems.	Machine learning	The study developed a ML based security solution to protect cloud based platforms.	The result when tested showed good result, but flood attack was not considered.
Doyen et al. (2019)	Malicious Url Detection System Using Ml	Machine learning	The study considered the various ML algorithms which have been adopted to solve the problem of intrusion detection system such as Artificial neural network, K-Nearest Neighbour (K-NN), Support Vector Machine (SVM), logistic regression and Naive bayes, K-mean clustering	The study also revealed that neural network achieved the best result when compared to others.

CONCLUSION

From the review of relevant literatures, many works have been developed over the years to solve the problem of intrusion on cloud based collaborative platforms, however despite their success on the packet penetration to the cloud was successfully secured, but the infrastructures remain vulnerable to attack and has remained a major gap waiting to be addressed over the years. This problem will be solved in this research, developing a multi level intrusion detection system which employed machine learning technique to secure the cloud log management server and guarantee data confidentiality and integrity.

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. Anthony T. Velte, Toby J. Velte and Robert Elsenpeter (2010). Cloud computing — A Practical Approach, 2010.
3. Ashigwuike E.C, A. R. A. Aluya, J. E. C. Emechebe and S. A. Benson (2020) " Medium Term Electrical Load Forecast Of Abuja Municipal Area Council Using Artificial Neural Network Method " *Nigerian Journal of Technology (NIJOTECH) Vol. 39, No. 3, July 2020, pp. 860 – 870;* Print ISSN: 0331-8443, Electronic ISSN: 2467-8821
4. Axelsson S, (1999). Research in Intrusion-Detection Systems: A Survey,tech. report TR-98-17, Dept. Computer Eng.,Chalmers Univ. of Technology, 1999.
5. 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.
6. 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.
7. Binita S. (2021) "Comparative Analysis of classification algorithms for intrusion detection"North Dakota State University of Agriculture and Applied Science; Master's Thesis.
8. Buyya, R., Broberg, J., and Goscinski, A.M. (2011). Cloud computing: Principles and paradigms. *John Wiley and Sons, Inc.*
9. Cloud Security Alliance, *Security Guidance for Critical Areas of Focus in Cloud computing,* Dec. 2009 Enisa, *Cloud computing Risk Assessment,* Nov. 2009
10. 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/
11. Cryptoclarity (2009). Encrypted storage and key management for the cloud. Retrieved from
12. 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.
13. deep learning," in Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, (ACM, Toronto, Canada), 2291–2293
14. 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.
15. Doyen Sahoo, Chenghao Liu, and Steven CH Hoi. (2019) malicious url detection using machine learning: A survey. arXiv preprint arXiv:1701.07179.
16. ENISA. (2020). ENISA Threat Landscape 2020 – Web-based Attacks. DOI: 10.2824/552242 [Online] Available at: https://www.enisa.europa.eu/publications/webbased-attacks; Retrieved on 2/19/2022
17. Farber, D., (2008). The new geek chic: Data centers. *CNET News.* Retrieved from http://news.cnet.com/8301-13953_3-9977049-80.html
18. Foley, J. (2008). Private clouds take shape. *InformationWeek.* Retrieved fromhttp://www.informationweek.com/news/services/business/showArticle.jhtml?articleID=209904474
19. Gens, F. (2008). Defining "cloud services" and "cloud computing". Retrieved from http://blogs.idc.com/ie/?p=190
20. Gruschka N. and M. Jensen (2010). "Attack Surface: A Taxonomy for Attacks on Cloud Services," IEEE 3rd International Conference on Cloud computing, pp. 276-279, 2010.
21. Hesamifard, E., Takabi, H., Ghasemi, M., and Jones, C. (2017). "Privacy-preserving machine learning in cloud," in Proceedings of the 2017 on cloud computing security workshop, 39–43
22. Hodo, X. Bellekens, A. Hamilton, P.-L. Dubouilh, E. Iorkyase, C. Tachtatzis, and R. Atkinson, "Threat analysis of IoT networks Using Artificial Neural Network Intrusion Detection System," in 2016 3rd International Symposium on Networks, Computers and Communications (ISNCC), 2016, pp. 1–6.
23. http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf
24. http://www.cryptoclarity.com/CryptoClarityLLC/Welcome/Entries/2009/7/23_Encrypted_Storage_and_Key_Management_for_the_cloud.html

25. Inyiama H.C and Agbaraji Chukwudi Emmanuel; "A Survey Of Controller Design Methods For A Robot Manipulator In Harsh Environments" European Journal of Engineering and Technology Vol. 3 No. 3, 2015 ISSN 2056-5860
26. Irfan Gul and M. Hussain (2011). Research on Distributed Cloud Intrusion Detection Model appear in International Journal of Advanced Science and Technology Vol. 34, September, 2011
27. JaeHyuk Jang (2010). Cisco, *Cloud computing: Drive Business Paradigm Shift*, 2010.
28. Jiang, Y., Hamer, J., Wang, C., Jiang, X., Kim, M., Song, Y., et al. (2018). Securelr: secure logistic regression model via a hybrid cryptographic protocol. *IEEE ACM Trans. Comput. Biol. Bioinf* 16, 113–123. doi:10.1109/TCBB.2018.2833463
29. Jolera. (2020). 3 Ways AI Prevents Phishing Attacks. [Online] Available at: https://www.jolera.com/3-ways-ai-prevents-phishing-attacks/; Retrieved on 3/2/2022
30. Jun Ho Lee, Min Woo Park and Jung Ho Ecom (2011).Multi-level Intrusion Detection and Log Management in Cloud computing IEEE computer society, pp 552-555, Feb.2011.
31. Kenny S. and B. Coghlan (2005). Towards a Grid-Wide Intrusion Detection System, Proc. European Grid Conf. (EGC 05),Springer, pp. 275–284,2005.
32. Kento S, Hitoshi. S, Satoshi. M, (2009). "A Model-based Algorithm for Optimizing I/O Intensive Applicationsin Clouds using VM-Based Migration", 9th IEEE/ACM International Symposium, Cluster Computing and Grid, 2009.
33. Lei, Y., Chen, S., Fan, L., Song, F., and Liu, Y. (2020). Advanced evasion attacks and mitigations on practical ml-based phishing website classifiers. *arXiv*
34. Liu, J., Juuti, M., Lu, Y., and Asokan, N.. (2017). "Oblivious neural network predictions via minionn transformations," in Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, October 2017, 619–631
35. Mazzariello C., Bifulco R. and Canonico R. (2010) "Integrating a Network IDS into an Open Source Cloud Computing Environment", Sixth International Conference on Information Assurance and Security, pp. 265-270.
36. Mills, E. (2009). Cloud computing security forecast: Clear skies. *CNET News.* Retrieved from http://news.cnet.com/8301-1009_3-10150569-83.html
37. Mkuzhalisai and G.Gayathri (2012). Research on Enhanced Security In Cloud With Multi-Level Intrusion Detection System, appear in International Journal of Computer and Communication Technology (IJCCT) ISSN (ONLINE): 2231 - 0371 ISSN (PRINT): 0975 –7449 Vol-3, Iss-3, 2012.
38. Mohd Helmy Abd Wahab, Mohd Norzali Haji Mohd, Hafizul Fahri Hanafi, Mohamad Farhan Mohamad Mohsin (2008)" Data Pre-processing on Web Server Logs for Generalized Association Rules Mining Algorithm" PROCEEDINGS OF WORLD ACADEMY OF SCIENCE, ENGINEERING AND TECHNOLOGY VOLUME 36 DECEMBER 2008 ISSN 2070-3740
39. National Institute of Standards and Technology (2012).*The NIST definition of cloud computing*. Retrieved from
40. Nils van noort (2020) "machine learning for monitoring attacks in the cloud"; 33rd twente student conference on it july 3rd, 2020, enschede, the netherlands.
41. Olofin B.B (2021a) "Key Scheduling Algorithm Assessment of Cryptographic Strength" IJESAT; Vol 8; Issues 3; pp 15-22.
42. Olofin B.B. (2021b) "A review of recent trends in intelligent agent technology" IJICST; Vol 6; Issue 2; pp. 90-98.
43. 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.
44. Ramgovind S, Eloff MM, and Smith E (2010). Research on The Management of Security in Cloud computing appear in 2010 IEEE
45. Rebecca Bace and Peter Mell (2001). *NIST Special Publication on Intrusion Detection Systems*, 16 Aug. 2001.
46. Roberto Di Pietro and Luigi V. Mancini (2008). *Intrusion Detection Systems*, Springer, Jan. 2008.
47. Rouhani, B. D., Hussain, S. U., Lauter, K., and Koushanfar, F. (2018). Redcrypt: real-time privacy-preserving deep learning inference in clouds using fpgas. *ACM Trans. Reconfigurable Technol. Syst.* 11, 1–21. doi:10.1145/3242899.
48. Shailendra Rathore, Pradip Kumar Sharma, and Jong Hyuk Park. Xssclassi_er: An effcient xss attack detection approach based on machine learning classi_er on snss. JIPS, 13(4):1014{1028, 2017.

49. Solomon Ogbomon Uwagbole, William J Buchanan, and Lu Fan. Applied machine learning predictive analytics to sql injection attack detection and prevention. In 2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), pages 1087{ 1090. IEEE, 2017.
50. Sonu D. (2020)"Understanding of Intrusion Detection System for Cloud Computing with Networking System" International Journal of Computer Science and Mobile Computing, Vol.9 Issue.3, March-2020, pg. 19-25
51. Soumya Mathew and Ann Preetha Jose (2012). Securing Cloud from Attacks based on Intrusion Detection System, International Journal of Advanced Research in Computer and Communication Engineering Vol. 1, Issue 10, December 2012
52. Strachey C. (1959). Time sharing in large fast computers. *Proceedings of the International Conference on Information processing, UNESCO* paperB.2.19, P. 336–341.
53. Subramanian K. (2009). Recession is good for cloud computing – Microsoft Agrees. Retrieved from http://www.cloudave.com/link/recession-is-good-for-cloud-computing-microsoft-agrees
54. Tramèr, F., Zhang, F., Juels, A., Reiter, M. K., and Ristenpart, T. (2016). "Stealing machine learning models via prediction APIs," in 25th USENIX security symposium (USENIX Security 16), 601–618
55. U.S. Patent (1994). A network [...] is shown schematically as a cloud", U.S. Patent 5,485,455, column 17, line 22, filed Jan 28, 1994. Retrieved from http://www. google. com/ patents?vid=5790548
56. Uday Trivedi and Munal Patel. A fully automated deep packet inspection veri_cation system with machine learning. In 2016 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), pages 1{6. IEEE, 2016.
57. Vieira K., Schulter A., Carlos B. Westphall, and C. Westphall M. (2010), "Intrusion Detection for Grid and Cloud Computing", IEEE Computer Society, pp. 38-43.
58. Vieira, K. Schulter, A. Westphall, C.B. and Westphall, C.M. (2010). Intrusion Detection for Grid and Cloud computing IEEE computer society,vol 12, issue 4, pp. 38 – 43,2010.
59. Wang, J., Zhang, J., Bao, W., Zhu, X., Cao, B., and Yu, P. S. (2018). "Not just privacy: improving performance of private deep learning in mobile cloud," in Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining London, United Kingdom, January 2018, 2407–2416.
60. WhatIsMyIPAddress. (2020). MaaS Chaos. Malware-as-a-Service is Growing. [Online] Available at: https://whatismyipaddress.com/maas; Retrieved on 3/11/2022
61. Wikipedia,*http://en.wikipedia.org/wiki/Cloud_computing*
62. Zhang, Q.,Cheng, L., and Boutaba, R. (2010). Cloud computing: state-of-the-art and research challenges". *J Internet Serv Appl 1. The Brazilian Computer Society*, p. 7–18.

Emeka Okorie, Yateghtegh. S. and Okoh Chris (2023). A review on Cloud computing and related technologies. *NEWPORT INTERNATIONAL JOURNAL OF ENGINEERING AND PHYSICAL SCIENCES (NIJEP)* 3(2):107- 120.