Shraddha Verma, 2021, 9:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

A Review on Secure Data Search Based on Encryption and Decryption Techniques in Cloud Data

M. Tech. Scholar Shraddha Verma, Asst. Prof. Dr. Neha Singh
Department of Computer Science Engineering,
IASSCOM Fortune Institute of Technology
Email-id – ifitbhopal@gmail.com

Abstract- Cloud computing has created a paradigm shift in the way information technology resources are provided and consumed by increasing agility and lowering costs. However, cloud-security concerns appear at the top of almost all surveys. Traditional security solutions are having a hard time trying to keep up with the growing security needs of the ever-changing threat landscape in today's cloud environment. This research paper aims to explore the concept of cloud security and its impact on the present day IT environment. We try to classify the security risks in the cloud environment, assess the cost of cloud security, delve into the types of attacks, look for currently available security solutions and arrive at a roadmap for cloud security. Cloud Server provides storage and search services. To perform efficient searches, the cloud uses verification keys to maintain privacy protection or meet authentication requirements and provide equivalent proof of encrypted documents based on tokens. Most security issues are caused by people deliberately creating malicious or malicious purposes. This Paper reviews and examines some Encryption and Decryption technologies. As a result, the better solution to the symmetric key encryption and the asymmetric key encryption is provided.

Keywords: - RSA, Cloud, Encryption, Decryption, Cloud Server, Security.

I. INTRODUCTION

Today, organizations generate a large quantity of sensitive data, such as financial data, individual information, or electronic health records. As a result, the generation of large amounts of digital data has improved congruently, or data storage volume for many organizations has generally increased significantly. It is very tough to succeed such a large amount of data in a local storage system, and high costs are caused by the need for high capacity storage organization or supplies of experts who manage them.

Although storage costs in recent years, the amount of hardware has been significantly reduced, or about 75% of government costs are still related to data storage management (Magalhaes et al. 2015).

Cloud computing is a distributed community that provides calculating or storage space as services to end users. The architecture/model of cloud computing is that all servers, networks, presentations or other basics connected to the facts centre are accessible to the end users. Cloud computing is upward in attention of technology and business organizations, but this is useful for solving social problems. It can also be beneficial. Cloud computing refers to online operation, configuration and access to applications. It provides online data storage, infrastructure or submissions.

Cloud computing allows individuals and businesses to shift the burden by managing large amounts of data or performance processes that require computing for powerful servers. Due to the growing approval of cloud figuring, more or more data proprietors are being encouraged to subcontract

their data to cloud attendants in order to provide great convenience and reduce data management costs. Data tenants provide services to many businesses and companies, and they insist on improving data security standards by following a covered method, including following: data encryption, key organization, strong admission controls, or security intellect. The cloud attendant performs query or returns encrypted papers with an additional proof according to the token generated by Data owners. The Data users will receive the result with the corresponding proof so they can verify the correctness and decrypt scrambled leaflets after verification is accurate.

1. Concepts of Cloud Computing:

Cloud computing is an advantage of information technology / business applications. Any organization can gain this benefit by paying or renting usage. Storage, servers and applications belong to the cloud computing area and are prerequisites for on-demand access. Therefore, unlike traditional methods of building data centers, hardware, applications and applications can be executed in a secure way before concentrating on building / transmitting business solutions. Cloud computing eliminates the need for expensive data centers and management because cloud vendors provide, manage and monitor the health and accessibility of the framework.

Registering a cloud is an event on the network that allows administrators to provide versatility, quality of service (QoS) and, in most cases, to ensure custom on-demand and low-cost computing infrastructure. These infrastructures can be simple and access in a universal way. Cloud computing is a model used to authorize expedient, on-demand network admission to a public pool of configurable computing value (such as systems, servers, storage, function, or management). These resources can be managed by negligible or cloud Service-fast configuration and release. The term "cloud" for vendor interaction is built from the network and its schematic representation is cloud. It refers to various specific types of services or submission that have been communicated in Internet cloud, and in many cases the devices used to get these products and applications require no special applications.[1][2]

2. Data Integrity and Availability:

Cloud users (i.e. data owners) must always be able to check honour of files at all times. To improve storage efficiency, it is best to copy metadata required for file and data integrity checks (such as authentication codes) at the same time. Considering a malicious user on the cloud server, the cloud server must verify that user really owns file before produce a link to his file; user must also confirm that cloud really stores the file in its storage or must be in the entire lifecycle Review the integrity of the file.

The new cloud computing paradigm provides an on-demand purchase of shared configurable computer resource pools and a convenient way to pay for on-demand network access. It requires minimal interaction or management work between service providers. Businesses can now choose to subcontract their data to cloud stowage to reduce the burden of local data storage or condense conservation. [3][4]

3. Secured File Transfer:

Although the cloud provides limited assistances to data owners, subcontracting data to remote servers and providing data organization to entrusted cloud service providers can pave the way for losing physical control of data (Wang et al., 2009). Cloud is inherently insecure and reliable for clients, which stances new challenges for integrity, concealment and data obtain ability of cloud computing. Delete e.g. less commonly received data to provide usable disk space or hide injured or corrupted data to shield organization's standing. Some organizations report that the data on the servers of main cloud substructure providers is corrupted and that there are several cloud service interruptions, such as massive deletion of Gmail emails, Side kick Cloud Disaster, Amazon S3breakdown and Amazon EC2 service interruptions (Crossbow, etc.). 2010). The Clearinghouse for Privacy (PRC) described that more than 535 data cracks occurred in 2011, including breaches of cloud-based email service providers based on Epsilon, Sony Play Station Network, Sony Online Performing or Sony Pictures' Damages, 3, 3 million theft patients' medical data of Sutter Surgeons Services, theft of customers' information on EMC's RSA.[5]

When adversaries get regulator over cloud server, they have capacity to launch replay attack or forge dose which aims at contravention linear individuality between data which is determined in the corrupted cloud server by replacing the stored ones with old encoded data. Then, honour of user data stored on exterior cloud servers is highly susceptible to internal

or outside occurrences. If there is no local copy of data, traditional integrity confirmation methods (such as signatures and hash functions) are not suitable for cloud storage.

In addition, it is difficult to transfer great files from cloud storage. When users use their mobile devices to access data, the situation gets worse. [6]

3.1 Integrity-Based: Only the auditors of this group are allowed to check veracity of outsourced data unswervingly or through a third part.

3.2 Recovery Based: In addition to carefully checking the integrity of the datasheet, this type of technology also uses Reed-Solomon deletion codes to support forward error correction codes (FEC) codes.

3.3 Deduplication-Based: It confirms data truth or efficiency by eliminating data duplication or growing data storage optimization. To project or implement an effective method of remote data inspecting, following related attributes must be considered:

3.4 Efficiency: When auditing data, it saves storage costs, calculation costs and communication costs between the client and the server.

3.5 Public/Private Verifiability: the data owner can only check integrity of subcontracted data in the private verifiability mode, while complex confirmation tasks are deputized to a third party in the public verification function to minimize data loss Calculate the cost and review the owner of electronic data.

3.6 Frequency: Repeat the verification process often with different challenge information.

3.7 Possibility of Detection: It represents possibility of perceiving potential data corruption.

3.8 Dynamic Update: This means that operations such as inserting, deleting, modifying and adding outsourced data can be performed without downloading all the data. The current RDA method now requires frequent review and involves many procedures or normal data transmission. Therefore, the RDA method will increase the auditor's computational and statement costs, which is a major problem for several data owners, particularly when using mobile devices with imperfect data processing resources (Fernando et al., 2013). On other hand, due to dynamic nature of data, the most important design aspect of RDA method is to facilitate dynamic data update operations for dissimilar presentations. Therefore, diverse types of data structures (such as binary trees) are used in such methods to achieve this goal. Though, the data structure used in the RDA method cannot effectively support dynamic data update operations and thus cannot handle large data efficiently. The most important thing is to update data frequently (Knauthe et al. 2008). This is because the auditor has to rebalance a great amount of data blocks in data structure several times, which will result in higher calculation costs for the auditor. Therefore, a new data structure must be designed to promote the dynamic updating of big data.[7]

II. PROBLEM FORMULATION

To ensure the secure data transmission and storage at minimal cost and searching time. The central goal of cloud computing is to improve computational capacity of the cloud system and to enhance the access levels to the services and resources of the cloud cheaply.

Cloud computing defines a remote server that is accessible via Internet, facilitating the use of business applications and features and computer software. This can save users money spent on annual or monthly subscriptions. Due to the benefits of cloud services, more and more personal information is concentrated on cloud servers, such as private videos and photos, individual health records, emails, government documents, company financial data, etc.[8]

III. LITERATURE REVIEW

The following sub-sections give information mined from technical books and IEEE papers. There are many papers related to cloud computing, cloud security, ECC algorithm and Shamir secret distribution.

Following the review, the following documents appear to be relevant to the current work of this paper:

The notion of Multi-Key Searchable Encryption (MKSE) enables data owners to outsource their data into a cloud server, while supporting fine-grained data sharing with the authorized users. Note that the traditional MKSE is vulnerable to data leakage. That is, the malicious data owner may collude with the server and recover the search queries of authorized users.

Recently, **Hamlin et al. (**PKC'18) presented a new MKSE construction that can ensure data privacy between data owner and authorized users, where the share key is generated depending on data owner, authorized user and the specific document. However, their scheme cannot support verifiable search in the case of the malicious cloud server. In this paper, we propose a new verifiable MKSE (VMKSE) scheme by leveraging Garbled Bloom Filter, which can simultaneously support verifiability of search result and secure data sharing in multi-user setting. Compared to the state-of-the-art solution, the proposed scheme is superior in efficiency and verifiability. The experiment results demonstrate the efficiency of our scheme

Yaping Su et. al. The notion of Multi-Key Searchable Encryption (MKSE) enables data owners to outsource their data into a cloud server, while supporting fine-grained data sharing with the authorized users. Note that the traditional MKSE is vulnerable to data leakage. That is, the malicious data owner may collude with the server and recover the search queries of authorized users. Recently, Hamlin et al. (PKC'18) presented a new MKSE construction that can ensure data privacy between data owner and authorized users, where the share key is generated depending on data owner, authorized user and the specific document. However, their scheme cannot support verifiable search in the case of the malicious cloud server. In this paper, we propose a new verifiable MKSE (VMKSE) scheme by leveraging Garbled Bloom Filter, which can simultaneously support verifiability of search result and secure data sharing in multi-user setting. Compared to the state-of-the-art solution, the proposed scheme is superior in efficiency and verifiability. The experiment results demonstrate the efficiency of our scheme

Ayoub ELMAARADI et. al (2019) Recently, we have seen a real digital revolution, and all companies prefer to use cloud computing because of its ability to provide the easiest way to deploy the required services. However, due to the vulnerability of privacy to cyber attacks, this digital transformation poses various security challenges. In this work, we will propose a new hybrid access system for virtual private cloud, which combines an online and host-based access management system to overcome the boundaries between each other. . . Based on an IDS network and gaining access to a host, the goal is to improve security in a private cloud environment. We recommend using non-traditional mechanisms in the concept of IDS (Detection Engine). Machine learning, machine learning algorithms will be used to build IDS in two regions to detect malicious traffic in the network area, as an additional protection to the network, and can also detect anomalies in in the local area to provide more privacy and confidentiality to the virtual machine. The training of the "ANN" artificial neural network is not in our field but instead offers a new strategy for IDS -based ANN. In our future work we will present all the details related to the ANN architecture and parameters, as well as the results of some experiments.[1]

Zina Chkirbene et al. (2019) Cloud computing is considered to be one of the most effective technologies for hosting and delivering Internet services. However, even though it is widely used, cloud security is the only major issue with cloud computing. To protect communications in this world, several security systems have been proposed, most of which are based on the nature of the attack. These systems are generally ineffective at detecting all types of attacks. Recently, machine learning technology was introduced. This means that if there are not enough examples in a particular class of training series, the decision may be incorrect. In this article, we present a new firewall solution called Enhanced Intrusion Detection and Classification (EIDC) system for secure computing environments. EIDC uses a new aggregation technology called "Frequency Decision" to identify and classify the received traffic packets, which contain the location of node 11. In this document we will use the words "node" and "alternately user "Learning algorithm to evaluate the classification of final attack categories. This strategy improves the learning efficiency and accuracy of the system. To generate our results, the UNSW-NB-15 public data set was used. Our results show that EICD improves anomalous completion by up to 24% compared to complex trees.[2]

Li Zhixin et al (2018) With the development and implementation of cloud computing platforms, the learning and prediction methods of the cloud computing sector provide a reliable solution for cloud platforms. However, due to the excitement and uncertainty of the cloud space environment, its study and prediction of the security status of the cloud space will be affected. Therefore, the method in this paper combines the internal security features with the features of the visible cloud, builds a transitional

model of the security of the cloud platform, and implements the ongoing development of the Ada Boost predictive algorithm the cloud field observation state. Then, based on the observable and predictive results, the Markov latency model is used to study and predict the security situation in the cloud platform, and the potential trends in the internal security situation future cloud platforms are counted.

The experimental results show that, compared with HMM, this method is able to predict the likelihood of a hidden safe state in the next two pre-existing time periods, as well as the agreed and working time of the employment is growing by about 16% and 1%, respectively.[3]

T. Nathezhtha et.al (2018) Security has always been a major issue in the cloud. The source is the most valuable and sensitive information that an attacker intends to steal. If data is lost, the privacy and security of all cloud users will be compromised. Even if the cloud network is protected externally, there are still threats from internal attackers. Internal attackers will destroy vulnerable user nodes and gain access to the system. They connect to an internal cloud network and launch attacks pretending to be trusted users. Machine learning methods are widely used for cloud security issues. Existing security methods based on machine learning classify nodes as nodes that do not carry irregular behavior based on short-term behavior data.

These systems cannot distinguish whether a point with abnormal behavior is a dirty node or a damaged point. To solve this problem, this paper proposes a long-term memory model (ILSTM), which can study a user's behavior and train them for it and store behavioral data. The model can easily classify user behavior as normal or abnormal.

The proposed ILSTM not only detects an invalid node, but also uses the calculated trust element to determine if a node with irregular behavior is a node, a new user node or a node. The proposed model not only detects the actual attack, but also reduces false alarms on the cloud network. [4]

Nimmy Krishnan et al. (2018) More recently, technology has evolved from traditional software models to cloud technologies. The significant increase in the number of applications using cloud - based infrastructure requires the protection of their security systems. An intrusive intelligence system is one of the most effective security solutions for protecting cloud environments. Although there are many methods of finding illegal access, such as signature-based and anomalous methods, the machine-based (ML) method has become the ultimate search area. With a robust learning model and a data-driven approach, ML-based solutions have proven to be effective for cloud environments. The attack performance is retrieved from the network and logged in to the application. The existence of attacks can be verified by performing machine learning techniques (such as logical reversal and propagation of beliefs). Use performance indicators as the average research time to evaluate the effectiveness of the method.[5]

Naiji Zhang et al. (2019) Denial of service attacks are one of the most common attacks, which are difficult to mitigate, but become more difficult when faced with low -level DoS (LDoS) attacks. LDoS exploits a hole in the TCP congestion control system by sending malicious traffic at a constant rate and affecting the victim's computer. Recently, machine learning methods have been implemented to detect complex DDoS attacks and improve the efficiency and robustness of hacking systems. In this research, the algorithm aims to balance the detection rate and its effectiveness. The detection algorithm combines entropy-density (PSD) function and supports a vector machine to detect LDoS traffic in normal traffic. In the solution, the detection rate and its efficiency can be adjusted depending on the parameters of the decision algorithm. To improve its efficiency, this detection method will always detect the attack by first calculating the PSD entropy and comparing it to the two adaptive thresholds. The threshold can effectively filter a sample of around 19% with a search rate. In order to reduce computational costs and find only the most relevant patterns for the search, a machine learning model based on the supporting vector machine is used to study the traffic patterns and select the appropriate features the search algorithm. The experimental results show that this method can detect 99.19% of LDoS attacks, and in the best case has the difficulty of time O (n log n). [6]

Sushant Sharma et. al (2020) Various studies were conducted to determine whether it is possible to detect web attacks through machine learning techniques. Negative and negative feedback is said to be the main problem that needs to be addressed

in order to make machine visibility and prevention of web attacks reliable and trustworthy. In our research, we try to identify and address the root causes of negative and negative responses. During our experiment, we used the CSTP 2010 HTTP database, which contains traffic generated for e-commerce web sites. Our experimental results show that for all tested machine learning algorithms, the application of well-defined feature mining results improves the detection and classification of web-based attacks. The effectiveness of the machine learning algorithm in attack output evaluates the Precision, Recall, Accuracy and F-ref ratings. Among the three algorithms tested, the J48 resolution tree algorithm gives the highest rates of "true validation", "accuracy" and "memory rate".[7]

Kritika Soni et.al (2019) Protecting resources from unauthorized users is a major challenge in cloud computing. The data access control model (RBAC) is a common method of data access control. It is an access control applied to many cloud platforms. As the number of characters increases, so does the complexity. To circumvent the limitations of RBAC, a character-based access control model is introduced. Here, licensing is directly related to responsibility. By combining RBAC and ABAC, a hybrid access control plan can be developed, which is more flexible and dynamic. This article presents various data entry control models and compares their characteristics. Compare the characteristics of all these security models.[8]

Gavini Sreelatha et al. (2020) Cloud concepts such as resource sharing, outsourcing, and multitasking pose significant challenges to the security community. In addition, the provision of cloud services based on trusted third parties and Internet technologies will create new security threats in the cloud environment. Cloud security research still faces shortcomings in improving the accuracy of detection and launching new or unknown attacks into the cloud. Machine learning technology plays an important role in automatically finding the discrepancies between legitimate and destructive data with real ones. Therefore, intelligent security systems need to be developed to study, adapt, and detect attacks or illegal situations in a diffused and dynamic cloud environment. This work is about a security solution designed for cloud environments through a new mechanism called mobile learning technology. The transfer learning model exploits the detection of known or unknown attack types by using knowledge from the source field. Transfer learning does not learn to attack from the beginning, but transfers knowledge to a trained attacker as a source of information.

This work provides a platform to identify and resolve new attacks against the target, simply training them and attaching them to the source and maintaining their effectiveness.[9]

Puja Ghosal et al. (2018) The Internet of Things (IoT) has attracted attention in the last few years because of its potential for significant advances in healthcare systems. With the rapid growth of cloud computing, it has been used in almost all medical fields. This is because IoT solutions use sensors that can be used to collect large amounts of data, and the cloud computing environment is ideal for processing and analyzing this large amount of data. However, cloud-based IoT solutions face many security issues. Therefore, this article conducted an in-depth analysis of cloud-based IoT systems that use machine learning to analyze data and review the various machine learning methods used in online security. Therefore, this article presents a typical model for future cloud technologies, which combine cloud technologies, IoT solutions and machine learning. A few use cases are given in this article to better understand our model.[10]

Yigit Sener et. al (2020) Cloud technology allows developers and organizations to focus on their products without having to deal with issues such as local server capabilities, infrastructure changes, data protection, licensing or human capital. This article attempts to explain the installation of machine learning software using Amazon Web Services (AWS) tools. In it, it illustrates the reasons for choosing a cloud work environment rather than a local resource. On the other hand, it should be noted that the implementation of this experience was produced in a hybrid way: use the local infrastructure for development, and then move to the cloud environment during the process installation only. In this respect it can be considered as a PaaS experiment. This research is considered to be a useful guide for entrepreneurs and startups with limited budgets who aim to deliver their products in a fast and measurable manner. [11]

Hou et.al (2019) A study by Size Hou et.al (2019) established Alibaba ECS's simulation of a home-based system. The structure of the device is based on computational technology. The whole method would be to develop a clear category to find the boundary between the common code and the mutation code. It can be applied to the search for network mutation codes. The project uses the set content to divide them into two types, good and bad. The final results show that the RBF method of SVM work is the most effective for this task. This research has gained good network security detection in the Internet of Things system and increased the application of machine knowledge [12].

Rupesh Raj Karn et al. (2019) Cloud network monitoring data is dynamic and distributed. Over time, cloud surveillance signals may appear, disappear or change their importance and clarity. Therefore, a modified machine learning (ML) model of a given data may become insufficient. The model may be very accurate at one time, but due to changes in the input data and its characteristics, it may lose accuracy at a later time. Therefore, it is often necessary to use an active model option in distributed learning. Under this option, less efficient models will be deleted (even if they are actively modified from the old data), or placed in the standing state when a new model or model is planned to replace it.

The popular Ensemble (EML) method can be used to improve overall performance of the accuracy of the ML model series. Unfortunately, EML has many disadvantages, including the need for ongoing training, too much computing workload, and the need for too much training documentation, too much risk of overuse and the process of building a time -consuming model. In this article, we present a new cloud method for ML model selection and modification, which can perform model building and selection, and has a competitive advantage over existing methods. Before generating the targeted controlled learning model in an automated way, we use non-controlled learning to better analyze the data space. Specifically, we built a cloud DevOps architecture for automatic editing and selection based on container delivery and messages passing between containers, and used new automated processing methods to create and evaluate the examples, algorithm ML. The proposed methods and tools are represented in the cloud network security data set. [13]

Marouane Hachimi et.al (2020) In 5G networks, cloud radio access (C-RAN) networks are considered promising in providing real-time cloud infrastructure, shared radio collaboration and processing. Centralized data to reduce energy consumption and share future resources intelligently. More recently, the security of the C-RAN network has attracted public attention for its vulnerability to malignant attacks. Among the various technologies for intrusive intrusion, one of the most important is the intrusion notification machine, as it can learn and correct the corresponding behavior without the intrusion of the intruder hand. In this guide many solutions have been proposed, but they show low inaccuracy in classifying attacks or simply provide an attack detection layer. The focus of this research is to implement multi-machine intrusion detection (ML-IDS) in 5G C-RAN, which can detect and classify four types of jamming attacks: continuous jamming, sudden harassment, and deceitful harassment and reactive interference. This installation improves protection by amplifying false alarms in the C-RAN architecture. The proposed solution was verified using the WSN-DS (Wireless Sensor Network Data Set) data set, which is data reserved for the anonymous. The final classification rate was 94.51%, and the adverse event rate was 7.84%.[14]

Song Xia et al. (2019) In this article, we present a system based on validation studies to protect network users from malicious network traffic. By training two educational assistants: a cyber attack activist and a network security operator, and based on a less sensitive network environment, the system is not designed to surpass machine learning algorithms. Traditional (such as CNN and LSTM), but can also identify opponents. For example, This is one of the biggest challenges faced by today's intervention detection systems based on machine learning.[15]

Sumanth Gowda et al. (2018) Active application security testing is carried out through an automated tool with a built-in scanner that can automatically crawl all web pages and report in accordance with pre-defined rules on security breaches . Such pre-defined laws cannot fully define their vulnerability, and it is often necessary to verify these results in order to clear up inaccuracies. Eliminating the effects

of this can be a very painful and difficult task. This article offers a way to eliminate the negative aspects through machine learning. Based on historical data that can be used for false positives, set up an appropriate machine learning model to predict whether the stated defects are really weak or negative.[16]

Dharitri Tripathy et al. (2020) Software as a Service (SaaS) by Dharitri Tripathy et al. (2020) was quickly adopted to enable applications and services to run on the cloud software platform. However, the success of SaaS in cloud deployments cannot hide the security challenges faced by Internet applications deployed on cloud SaaS. Like other web -based systems, cloud applications are also vulnerable to the most common web attacks. SQL injection attacks are one of the biggest threats to SaaS applications. May result in loss of sensitive and important data (e.g. financial, personal). Through this attack, an attacker can steal confidential and confidential information from a company or organization, which can have a significant impact on both tangible objects (such as data) and intangible objects. (Such as, fame). The purpose of this research was to investigate the feasibility of using machine learning techniques for SQL injection detection at the application level. The algorithm to be tested is a trained category for various dirty and micro wages. They treat it as a paycheck and determine if the entry has a dirty code. The results show that these algorithms are able to distinguish between normal and risky wages, and the search rate reaches 98%. This article also compares the effectiveness of various machine learning models in detecting SQL injection attacks.[17]

Linda Joseph et.al (2018) With the constant change of cloud services, large cloud computing is evolving at a much faster pace. Compared to public cloud services, the cloud's flexibility is delivered over a large area of it, which is often less secured. Therefore, these cloud services need to be protected and protected, which is critical to cloud infrastructure. Therefore, in this research work, we have identified the real -time sensors that can be represented by Infrastructure as a Service (IaaS) in a private cloud environment. We examined attack models from virtual machine simulations and analyzed them based on the techniques learned by the control machine. The virtual machine snapshot API call sequence is considered to be the integration of controlled and uncontrolled machine learning algorithms to classify the invaded and unattended virtual machine memory images. The images of the memory cameras of the attacked virtual machine are used as input to the self -healing algorithm, which can replicate the operation of the virtual machine. The way we detect malware can be up to 93% compared to virtual machine screens.[18]

Xia Xinxin et al (2018) With the rapid development of the information sector, the application of the Internet of Things, computing and artificial intelligence has had a major impact on people's lives. , And network equipment showed a significant increase in growth. Meanwhile, the more complex network environment has become a more serious security issue. In new circumstances, traditional security solutions have become ineffective. Therefore, it is an important task for the security industry to seek advances in security technology and improve security visibility and security capabilities. Botnets have become one of the biggest challenges in online security. Especially in the last year or two, botnets have become one of the most vulnerable countries. Therefore, botnets have a huge impact on the world. Botnets have made a splash. The issue once again caught the attention of the public. This paper takes botnet detection as a topic, and focuses on the latest research results on botnet detection based on machine learning technology. First, it explains the process of applying machine learning technology to space security search, introduces the characteristics of botnets, and then introduces machine learning to botnet detection. It focuses on the analysis and summarization of security operations and the common machine learning algorithms used in these solutions. Finally, it summarizes the challenges of existing solutions, as well as future development paths and challenges of machine learning technology in Internet security.[19]

Sanjay Kumar et al. (2020) The purpose of Sanjay Kumar et al. (2020) are to understand the development path of the big data sector and focus on incorporating the latest trends into big data. Regardless of the granularity of the data, the human brain can process many different forms of data efficiently and instantly. The idea behind the field of big data processing is to mimic the way the natural brain works and handles big data. All fields and areas of machine learning are based on recognizing and emulating the natural methods used by nature to process large amounts of data. The current trend is

to use cloud infrastructure and machine learning technology to handle large amounts of data. In general, the goal of intelligence is to develop humanoid robots that can work, talk and think like humans.

The future trend of big data is being applied to technologies such as microprocessors that process big data in the brains of robots to plan their behavior, express emotions and think like humans only. [20]

IV. CRYPTOGRAPHIC SYSTEMS

Cryptographic Systems can be divided into deterministic and probabilistic encryption scheme [7]. Deterministic encryption scheme allows the plaintext is encrypted by using keys that always provide the same cipher text, but the encryption process is repeated many times. In this scheme, every plaintext has one to one relationship with the keys and cipher text otherwise it will produce more than one output of particular plaintext during the decryption process. Probabilistic Encryption Scheme shows the plaintext has different cipher text with the different keys.

The probabilistic encryption scheme is significantly secure than the deterministic encryption scheme because it makes difficult for a cryptanalyst to access any sensitive information regarding plaintext that is taken from cipher text and corresponding key. Furthermore, the cryptographic algorithms can be further divided into two main categories like keyless cryptosystem and key-based cryptosystem as shown in Fig. 1. In the keyless cryptosystem, the relationship between the plaintext and cipher text having a different version of the message is exclusively depend on the encryption algorithm [8].

The keyless cryptosystem is generally less secure than key-based systems because anyone can gain access to the algorithm will be able to decrypt every message that was encoded using keyless cryptosystem such as Caesar cipher [9].

The key based cryptosystem can be further categories into symmetric key (secret key) encryption and asymmetric key (public key) encryption based on the type of security keys utilized for the encryption or decryption process [10]-[13].

The detail of the cryptosystems is explained as follows:

1. RSA Algorithm:

RSA consists of Public-Key and Private-Key. In our Cloud environment, Pubic- Key is known to all, whereas Private-Key is known only to the user who originally owns the data. Thus, encryption is done by the Cloud service provider and decryption is done by the Cloud user or consumer. The RSA algorithm (named after the inventors Rivets, Shamir and Adleman) was one of the first cryptographic algorithms that met the requirements for public key systems as stated by Diffie and Hellman [5]. Since then it has reigned supreme as the only widely accepted and implemented general purpose approach to public key systems.[10]

2. Key Schedule Algorithm:

Key schedule algorithm is employed to generate secret keys and plays an important role in the development of encryption and decryption key. The insignificant key generation algorithm generates weak keys that are used for encryption process can easily attack using brute force attack because cryptanalyst continuously trying all possible combinations to get original text using this attack [27]-[29]. All cryptographic algorithms follow the consideration of Advanced Encryption Standard (AES) that must support the key lengths include 128 bits, 192 bits and 256 bits [19]. The number of the round for that key length is 10, 12, 14 respectively and the round keys are taken from the cipher key using key schedule algorithm and utilized in the construction of block cipher. For the development of fully secure block cipher, the multiple numbers of rounds ensure the high diffusion and employed invertible transformation.

3. Symmetric Key Encryption:

The symmetric key (secret key) encryption is employed similar key for the encryption and decryption of a message. Encryption and decryption keys are keeping secret and only known by authorized sender and recipient who want to communicate. The allocation of different keys to the different parties increases the overall message security. The strength of the symmetric key encryption is depending on the secrecy of encryption and decryption keys. The symmetric encryptions algorithms can be classified into block and stream cipher on the basis of the grouping of message bits

[14], [15]. In a block cipher, a group of messages characters of a fixed size (a block) is encrypted all at once and sent to the receiver. Moreover, the block cipher can be further divided into binary and non-binary block cipher based on the final results of the message, keys and cipher text.

The message bit size for the binary block cipher is 64, 128, 192, and 256 and the non-binary block cipher has not defined the standard that depends on the cipher implementation.

4. Asymmetric Key Encryption:

The asymmetric key encryption is commonly referred to as public key encryption in which different keys are employed for the encryption and decryption of the message. The encryption key is also said as the public key and can be utilized to encrypt the message with the key. The decryption key is said to as secret or private key and can be used to decrypt the message. The strength of the asymmetric key encryption is utilized with digital signature then it can provide to the users through message authentication detection. The asymmetric encryption algorithm includes RSA.

5. Public Key:

The public key consists of the modulus n and the public (or encryption) exponent e. The private key consists of the private (or decryption) exponent d, which must be kept secret. p, q, and λ(n) must also be kept secret because they can be used to calculate d. In fact, they can all be discarded after d has been computed. In the original the Euler totient function φ(n) = (p − 1)(q − 1) is used instead of λ(n) for calculating the private exponent d. Since φ(n) is always divisible by λ(n) the algorithm works as well. That the Euler totient function can be used can also be seen as a consequence of Lagrange's theorem applied to the multiplicative group of integers modulo pq. Thus any d satisfying $d \cdot e \equiv 1 \pmod{\varphi(n)}$ also satisfies $d \cdot e \equiv 1 \pmod{\lambda(n)}$. However, computing d modulo φ(n) will sometimes yield a result that is larger than necessary (i.e. d > λ(n)). Most of the implementations of RSA will accept exponents generated using either method (if they use the private exponent d at all, rather than using the optimized decryption method based on the Chinese remainder theorem described below), but some standards such as FIPS 186-4 may require that d < λ(n). Any "oversized" private exponents not meeting that criterion may always be reduced modulo λ(n) to obtain a smaller equivalent exponent.

6. Data Encryption Standard (DES):

DES is the earliest symmetric encryption algorithm developed by IBM in 1972 and adopted in 1977 as Federal Information Processing Standard (FIPS) by the National Bureau of Standard (NBS). The NBS is currently the National Institute of Standards and Technology (NIST) that evaluate and implement the standard encryption algorithm. It includes 64 bits key that contains 56 bits are directly utilized by the algorithm as key bits and are randomly generated. The remaining 8 bits that are not used by algorithm because it is used for the error detection as set to make a parity of each 8- bit byte [17], [37], [38]. DES utilized the one secret key for encryption and decryption process and key length is 56 bits and performs the encryption of message using the 64 bits block size. Similarly, the decryption process on a 64 bits cipher text by using the same 56 bits key to produce the original 64 bits block of the message

7. Key Distribution:

Suppose that Bob wants to send information to Alice. If they decide to use RSA, Bob must know Alice's public key to encrypt the message and Alice must use her private key to decrypt the message. To enable Bob to send his encrypted messages, Alice transmits her public key (n, e) to Bob via a reliable, but not necessarily secret, route. Alice's private key (d) is never distributed.

8. Encryption:

After Bob obtains Alice's public key, he can send a message M to Alice. To do it, he first turns M (strictly speaking, the un-padded plaintext) into an integer m (strictly speaking, the padded plaintext), such that $0 \le m < n$ by using an agreed-upon reversible protocol known as a padding scheme. He then computes the cipher text c, using Alice's public key e, corresponding to this can be done reasonably quickly, even for very large numbers, using modular exponentiation. Bob then transmits c to Alice.

9. Decryption:

Alice can recover m from c by using her private key exponent d by computing given m, she can recover the original message M by reversing the padding scheme.

V. CONCLUSION

Cloud Computing is still a rapidly evolving landscape; and one that requires us to stay current or fall behind. Still, we must not be complacent. Just as security professionals have done for ages, we must continue to evolve our processes, methods, and techniques in light of the opportunities that Cloud Computing brings to our industries. This evolution is critical to our long-term success as we find new ways to improve the efficacy and efficiency of our security enforcement and monitoring capabilities. Cloud Computing isn't necessarily more or less secure than a traditional environment. As with any new technology, it creates new risks and new opportunities.

In some cases moving to the cloud provides an opportunity to re-architect older applications and infrastructure to meet or exceed modern security requirements. At other times the risk of moving sensitive data and applications to an emerging infrastructure might exceed our tolerance.

REFERENCES

[1] Ayoub ELMAARADI; Abdelouahid LYHYAOUI; IKRAM CHAIRI New security architecture using hybrid IDS for virtual private clouds 2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS) Year: 2019 DOI: 10.1109/IEEE Marrakech, Morocco.

[2] Zina Chkirbene; Aiman Erbad; Ridha Hamila A Combined Decision for Secure Cloud Computing Based on Machine Learning and Past Information 2019 IEEE Wireless Communications and Networking Conference (WCNC) Year: 2019 DOI: 10.1109/ IEEE Marrakesh, Morocco.

[3] Zhixin Li; Lei Liu; Yanli Zhang; Bin Liu Learning and Predicting Method of Security State of Cloud Platform Based on Improved Hidden Markov Model 2018 3rd International Conference on Smart City and Systems Engineering (ICSCSE) Year: 2018.

[4] T. Nathezhtha; V. Yaidehi Cloud Insider Attack Detection Using Machine Learning 2018 International Conference on Recent Trends in Advance Computing (ICRTAC) Year: 2018 DOI: 10.1109/ IEEE Chennai, India.

[5] Nimmy Krishnan; A. Salim Machine Learning Based Intrusion Detection for Virtualized Infrastructures 2018 International CET Conference on Control, Communication, and Computing (IC4) Year: 2018 DOI: 10.1109/IEEE Thiruvananthapuram, India.

[6] Naiji Zhang; Fehmi Jaafar; Yasir Malik Low-Rate DoS Attack Detection Using PSD Based Entropy and Machine Learning 2019 6th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/ 2019 5th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom) Year: 2019 DOI: 10.1109/CSCloud/ IEEE Paris, France.

[7] Sushant Sharma; Pavol Zavarsky; Sergey Butakov Machine Learning based Intrusion Detection System for Web-Based Attacks 2020 IEEE 6th Intl Conference on Big Data Security on Cloud (Big Data Security), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS) Year: 2020 DOI: 10.1109/ IEEE Baltimore, MD, USA.

[8] Kritika Soni; Suresh Kumar Comparison of RBAC and ABAC Security Models for Private Cloud 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon) Year: 2019

[9] Gavini Sreelatha; A. Vinaya Babu; Divya Midhun Chakkarvarthy Ensuring Anomaly-Aware Security Model for Dynamic Cloud Environment using Transfer Learning 2020 5th International Conference on Communication and Electronics Systems (ICCES) Year: 2020 DOI: 10.1109/ IEEE Coimbatore, India.

[10] Puja Ghosal; Debanjan Das; Indrajit Das Extensive Survey on Cloud-based IoT-Healthcare and Security using Machine Learning 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN) Year: 2018 DOI: 10.1109/ IEEE Kolkata, India.

[11] Yigit Sener; Hasan Fahri Yetim; Selami Bagriyanik Delivering Machine Learning Applications via Cloud Platforms: An Experience Report 2020 Turkish National Software Engineering Symposium (UYMS) Year: 2020.

[12] Size Hou; Xin Huang Use of Machine Learning in Detecting Network Security of Edge Computing System 2019 IEEE 4th International Conference on Big Data Analytics (ICBDA) Year: 2019.

[13] Rupesh Raj Karn; Prabhakar Kudva; Ibrahim Abe M. Elfadelv Dynamic Auto selection and Auto tuning of Machine Learning Models for Cloud Network Analytics IEEE Transactions on Parallel

International Journal of Science, Engineering and Technology

An Open Access Journal

and Distributed Systems Year: 2019 DOI: 10.1109/TPDS.2018.2876844.

[14] Marouane Hachimi; Georges Kaddoum; Ghyslain Gagnon; Poulmanogo Illy Multi-stage Jamming Attacks Detection using Deep Learning Combined with Kernel zed Support Vector Machine in 5G Cloud Radio Access Networks 2020 International Symposium on Networks, Computers and Communications (ISNCC) Year: 2020.

[15] Song Xia; Meikang Qiu; Hao Jiang An adversarial reinforcement learning based system for cyber security 2019 IEEE International Conference on Smart Cloud (Smart Cloud) Year: 2019 DOI: 10.1109/ IEEE Tokyo, Japan.

[16] Sumanth Gowda; Divyesh Prajapati; Ranjit Singh; Swanand S. Gadre False Positive Analysis of Software Vulnerabilities Using Machine Learning 2018 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM) Year: 2018.

[17] Dharitri Tripathy; Rudrarajsinh Gohil; Talal Halabi Detecting SQL Injection Attacks in Cloud SaaS using Machine Learning 2020 IEEE 6th Intl Conference on Big Data Security on Cloud (Big Data Security), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS) Year: 2020.

[18] Linda Joseph; Rajeswari Mukesh To Detect Malware attacks for an Autonomic Self-Heal Approach of Virtual Machines in Cloud Computing 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM) Year: 2019.

[19] Xiaxin Dong; Jianwei Hu; Yanpeng Cui Overview of Botnet Detection Based on Machine Learning 2018 3rd International Conference on Mechanical, Control and Computer Engineering (ICMCCE) Year: 2018.

[20] Sanjay Kumar; Priyanka Gupta; Sachin Lakra; Lavanya Sharma; Ram Chatterjee The Zeitgeist Juncture of "Big Data" and its Future Trends 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon) Year: 2019.

[21] Yaping Su 1, Jianfeng Wang 1, Yunling Wang 1, and Meixia Miao2 Efficient Verifiable Multi-Key Searchable Encryption in Cloud Computing Digital Object Identifier 10.1109/ACCESS. 2019. 2943971