Secure Data Search in Cloud Services based on Encryption Scheme

Shivanik Chaturvedi
Swami Vivekanand College of Engineering, Indore (MP), India

Abstract— In recent days, Cloud storage has become good entrant for organizations that suffer from resource limitation. Cloud computing is a procedure that surveys internet founded computing. The most challenging task in the cloud is to ensure availability, integrity, and secure file transfer Searchable Symmetric Encryption (SSE) has been extensively explored in cloud storage, enabling cloud services directly search for encrypted data. Most SSE solutions are only suitable for honest but curious cloud services and will not differ. Data protection or file protection in the cloud environment is one of the biggest problems in the cloud environment. The accuracy of the data means that if the information is accessible to a remote system or local system it should not be altered. Verification is an effective way to authenticate users who are trying to access information. The availability of data indicates the availability of data if necessary. Confidentiality is usually through encryption technology. The confidentiality of data and keywords is the most important privacy requirements in SSE. It ensures that users' plaintext data and keywords cannot be revealed by any unauthorized parties, and an adversary cannot learn any useful information about files and keywords through the proof index and update tokens used in GSSE.

Keywords: Encryption, Decryption, Cloud Server, GSSE, SSE, Cloud, Data User

I. INTRODUCTION

Today, cloud computing is a major technology that is considered a great tool to improve the IT economy. The history of cloud computing can be traced back to the 1950s, when a large computer appeared. In the mainframe, many users can access the central computer through a mute terminal. Large computers provide ample storage space with high processing power. In the 1970s, IBM developed a virtual machine that allowed multiple operating systems to run together to perform different tasks. A host operating system that has a frequent occurrence in the visitor system of a virtual machine is called a virtual machine instance. Each example has its own infrastructure and memory. In the 1990s, telecom operators introduced virtual private networks (VPNs) to share a single physical infrastructure. VPN is mainly used to transmit data through dedicated network lines. This VPN uses encryption to provide more secure security than traditional Internet. In 1999, Salesforce.com came up with the idea of launching enterprise software through a website. The next major development is Amazon Web Services, which allows rented computers for users to run their own businesses. Then, Amazon launched Elastic Cloud (EC2) in 2009 to provide SaaS (software as a service). The significant contribution of Microsoft and Google to cloud computing. They introduced a killer app. Other important technologies that make up computer science are high-speed communication channels, software-related standards, and hypervisors. According to its purpose, cloud computing can perform a variety of functions, such as hosting a cloud host, a cloud server serving as a server, and cloud storage providing storage space. Using cloud computing can be improved with less complexity. The most important contribution of this article is to verify the availability and reliability of cloud storage, and to ensure the transmission of reliable data with the time and cost of research.

To ensure the security of data stored and transmitted, information in the cloud environment is more important. The idea of a cloud is to transfer data from your own computer and exchange data in the cloud in a secure way (El-Booz et al., 2016). It is very important to understand the security requirements of cloud providers. In general, the following challenges must be met while ensuring cloud security: the availability of data storage. Unauthorized attack by a third party. Encryption method. File search and sharing. The real tool that stores the data. Cloud computing is a kind of custom computing process based on Web computing. To reduce the cost and time of data processing, cloud computing technology is used (Joshi et al., 2010). Here, the data owner will transmit the data through the cloud server through the user. When multiple users share a central cloud server, this will affect the security of data transmission. In order to securely send data through the cloud server, encryption or deletion settings are used. Challenges in searching of data in cloud computing environment: Data owners use all encryption algorithms to hide the data, and then forward the data to the cloud server. Organized data is stored in a cloud server, and the user retrieves its data through a decryption program (Goh et al., 2013). The encryption algorithm includes public key, private key and key production. Public keys are distributed to cloud servers to retrieve and transmit data. The private key is shared by the user, otherwise the user will use your own key to extract the data. The encryption algorithm will generate keys such as private key or public key. The data is deleted and shared on the cloud server, where the data will not be accessible, but can be shared with each user. Users apply full technology to search for data from the cloud. There are many search methods that can be used for data search

II. RELATED WORK

In this article, the author has used extensive research to conclude that the safety and privacy of physical, environmental and virtual security is the responsibility of the seller. This article argues that organizations can control the three major phases of physical, logical, and business processes to address central threats, cyber security and internal security. The author proposes an automated management plan that aims to address security concerns in the calculation of threat threats by establishing relationships between CSPs based on existing attacks. The author also recommends that users should be responsible for ensuring security of data center in the cloud, enforcing policy monitoring, improving users' understanding of security practices. Extending, resolving disputes, managing legal matters, evaluating capabilities and providing solutions. we talked about purpose of this emerging cloud-based technology, which provides common sources and services for

lower prices and software when using cloud services, and some security issues. In addition, while emphasizing the multi-tasking, the CSA's cloud-based data security issues were also addressed. The authors conclude that security issues with cloud computing can be mitigated by modifying or designing a proprietary architecture for many cloud applications [3][8][10]

A. 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.

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.

B. 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. The accuracy of the ML model series. Unfortunately, EML has many disadvantages, including the need for ongoing training, too much computing workload, 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 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.

C. 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 behaviour 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%.

D. 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.

E. 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.[2]

III. PROPOSED SYSTEM

We aim to develop a verifiable SSE scheme, i.e., GSSE that allows the index used for search result verification to be separated from the one used for the SSE operations.

Therefore, GSSE is decoupled from the existing SSE schemes. In particular, data owner will builds an encrypted index based on the Merkle Patricia Tree (MPT) and upload it to cloud services, which enables data users to verify the integrity of search results. Meanwhile, data owner will also upload a timestamp-chain based on the root of MPT to ensure data freshness across multiple users. The confidentiality of data and keywords is the most important privacy requirements in SSE. It ensures that users' plaintext data and keywords cannot be revealed by any unauthorized parties, and an adversary cannot learn any useful information about files and keywords through the proof index and update tokens used in GSSE. 2) Verifiability: A verifiable SSE scheme should be able to verify the freshness and integrity of the search results for users. 3) Efficiency: A verifiable SSE scheme should achieve sub linear computational complexity

The data owner first citations keywords of each article or build a keyword directory. He/she encrypts papers as well as keyword index. The data owner blocks the file and password folders locked in the cloud. Data users can obtain individual results, evidence or key credentials in public, and they or others can verify the accuracy, reliability and validity of the search results without decryption. The advantage of a cloud computing equity service is that it provides reliable returns on investment, but the losses are even greater. Compared to traditional computer technology, cloud computing has various advantages. Cloud computing provides customers with supercomputing and high-end devices at affordable prices

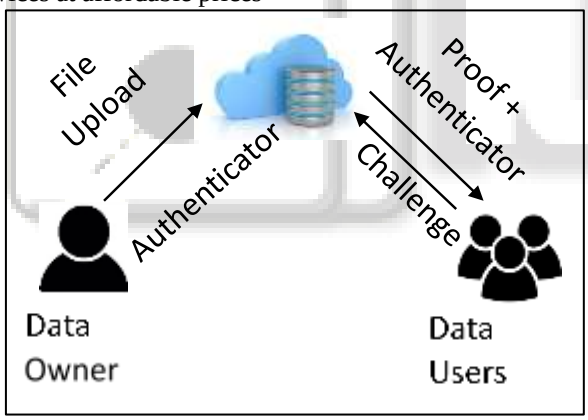

Fig. 1: Proposed Architecture Diagram

A. *Registration*

This is the process of registering or registering to cloud. To take benefit of cloud documents, all data owners and data users must register. During this process, your basic information (such as email, contacts, etc.) will be collected and stored in the cloud. During the registration process, a particular user's cloud ID is generated automatically.

B. *Cloud ID*

Each user must produce a Cloud ID or use it to classify an identifier with near security. The identifier does not repeat the identifiers that have been or will be twisted to identify other identifiers. Therefore, the information marked with Cloud ID by liberated parties can later be collective into a single folder or transmitted on same channel without the need to tenacity struggles among identifiers

C. *Data Owner*

Data Owner extracts keywords of each article or also figures a keyword Index. Data Owner encrypts documents r keyword Index using a key and outsources in Cloud.Data Owner provides the Public Verification Key and Proof Index to the Data User via Cloud for document verification. Data Owner is the only authorized person to add, modify, or delete the document(s) from the cloud.

D. *Cloud Service Provider*

The cloud service provider can see all uploaded or transferred documents in cloud. CSP obtains document request from the data user, verifies identity before granting permission, and then CSP executes query or revenues encrypted document based on search token, or also returns document with other evidence on document to confirm search results.

E. *Public Verification Key*

Public verification key is a safety quantity planned to make sure that your document outsourced in cloud doesn't get hacked. By confirming public key, the Data Owner and the Data User adding added cover of defence to documents or files in the cloud by authorising each other's identities.

F. *Data User*

Data User send a appeal to the cloud server.After request granted from the Cloud, the Data User receiving the Public Verification Key from the Cloud generated by Data Owner. The Data User now decrypts and downloads the encrypted documents, after verifying with the Public Verification Key. After receiving verification from cloud, the data user will download the file within a particular time limit.

G. *Verification with Proof Index*

It is a proof generating system for verifying cloud search by Public Verification Key; here data users or others can confirm accuracy of search result by Verification key

Fig. 2: User Registration

Fig. 3: Upload File

Fig. 5: Upload Data

Fig. 6: Search Keywords for Upload File

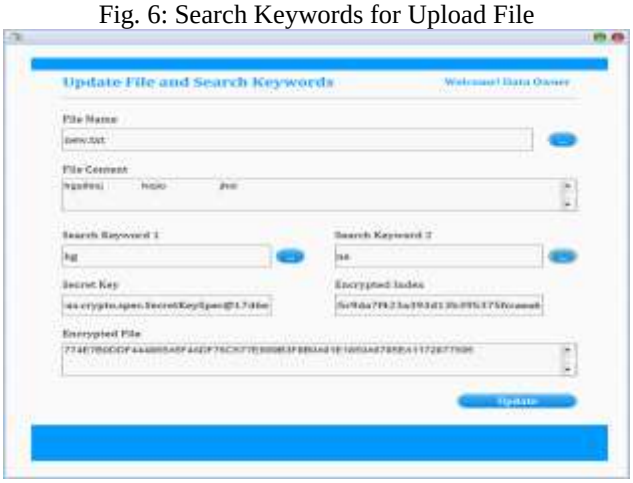

Fig. 7: Upload File and Search Keywords

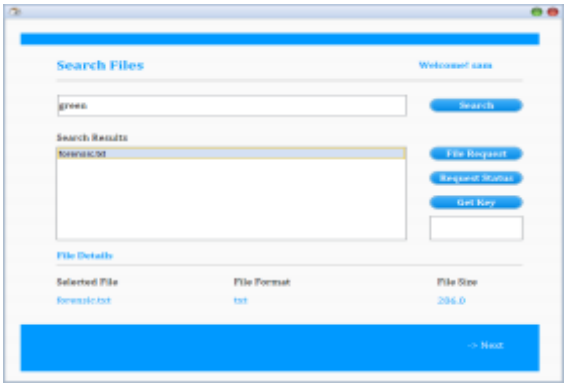

Fig. 9: Search File

Fig. 10: Cloud Server

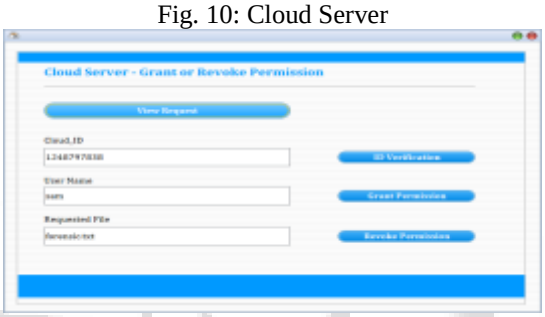

Fig. 11: Cloud Server Grant Revoke

Fig. 12: Authentication key generation

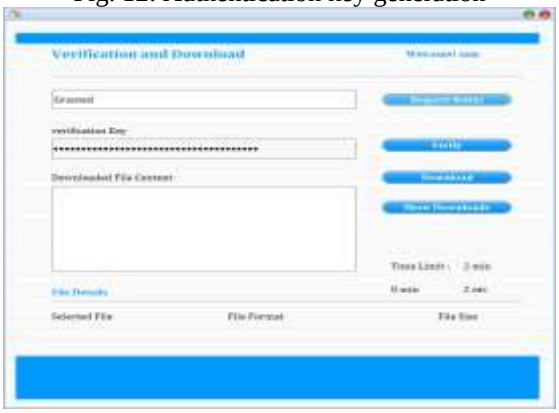

Fig. 13: Data User – Received Public verification Key from CSP

Fig 14 Cloud Server

In this work, we focus on the problem of verifiable searchable encryption under a multi-user setting. A GSE scheme is proposed which can support verifiability of search result even when both the data owner The experimental results denote that our scheme can achieve security goals for data owner-server collusion while maintaining a comparable performance.

	Previous Work	Proposed Work
Technique	Multi-Key Searchable Encryption (MKSE)	GSE
Language	Java language	Java language
searching for documents	Token Gen phase	Keyword generation
Time	1400ns	2 sec

Table 1: Comparison Table with Existing Technique

IV. CONCLUSION

In this paper, we focus on privacy, security and access to the cloud computing environment. While cloud security services can be well-designed and succeeded by experts, they can provide effective organisation or threat valuation services. Though, threats we are discussing here show that the implementation of present security mechanisms in the cloud should be carefully considered. In order to accelerate the development of cloud computing, many improvements to existing mechanics are needed, and new innovation systems need to be established. we plan to cover the planned work to other parts of cloud. Cloud computing brings various tasks for structure or submission developers, engineers, system administrators and service providers. Collected on the web servers. It helps in reducing allocation of geographical area required for storing records and also promotes paperless work. Time consumed for searching required documents is less. Every organization prefers computerization as well as remotely accessible web services Hence data security and protection comes in highest priority so recent developments will be on securing and protecting data collection on web server The steps of the algorithm are also determined to ensure that the operation works efficiently. More and more people and communications companies are integrating data into remote servers or reducing problem of storing and maintaining local data. Uncertainty as to how to ensure security in the external computer environment has led to security issues such as authentication, licensing, existence, trust, confidentiality and anonymity.

REFERENCES

[1] 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
[2] 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
[3] Size Hou;Xin Huang Use of Machine Learning in Detecting Network Security Linda Joseph;Rajeswari Mukesh To Detect Malware attacks for an Autonomic Self-Heal Approach of Virtual Machines in Cloud Computing 2019 Fifth of Edge Computing System 2019 IEEE 4th International Conference on Big Data Analytics (ICBDA) Year: 2019
[4] Rupesh Raj Karn;Prabhakar Kudva;Ibrahim Abe M. ElfadelvDynamic Autoselection and Autotuning of Machine Learning Models for Cloud Network Analytics IEEE Transactions on Parallel and Distributed Systems Year: 2019 DOI: 10.1109/TPDS.2018.2876844
[5] Marouane Hachimi;Georges Kaddoum;Ghyslain Gagnon;Poulmanogo Illy Multi-stage Jamming Attacks Detection using Deep Learning Combined with Kernelized Support Vector Machine in 5G Cloud Radio Access Networks 2020 International Symposium on Networks, Computers and Communications (ISNCC) Year: 2020
[6] Song Xia;Meikang Qiu;Hao Jiang An adversarial reinforcement learning based system for cyber security 2019 IEEE International Conference on Smart Cloud (SmartCloud) Year: 2019 DOI: 10.1109/ IEEE Tokyo, Japan
[7] 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
[8] Dharitri Tripathy;Rudrarajsinh Gohil;Talal HalabiDetecting SQL Injection Attacks in Cloud SaaS using Machine Learning 2020 IEEE 6th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS) Year: 2020
[9] International Conference on Science Technology Engineering and Mathematics (ICONSTEM) Year: 2019
[10] 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
[11] 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

