ISSN 2063-5346

CLOUD INTEGRATION OF ENCRYPTED HETEROGENEOUS DATABASE ASSISTED BY TRANSPARENT CIPHERTEXT RETRIEVAL SYSTEM

A Yogaraj B.E,M.Tech[1], Jalaja S M.E[2], Sneha S[3], Zoe Saral D[4] , Challa Raja Kumari [5],Avanthika[6], Prathima[7]

Article History: Received: 01.02.2023 Revised: 07.03.2023 Accepted: 10.04.2023

Abstract

In the Internet of Things with cloud assistance, recovery of ciphertext under diverse data sets is a critical challenge (IoT). The older systems for retrieving ciphertext are not capable of supporting encrypted data sets that are diverse, which results in reduced security risks and computational requirements to some degree.The focus of our research is to address this issue by presenting a transparent technique for retrieving ciphertext that is programming language agnostic, platform independent, and compatible with various types of data sets.. By integrating our suggested access and integration middleware in a unique way with a pre-existing encrypted cloud database system, we have created the retrieval mechanism. This middleware facilitates queries across different platforms and programming languages in the database system. The proposed plan includes a data integration strategy for diverse databases, aiming to facilitate cross-database searches of encrypted cloud data sets. Effective data sharing between various IoT apps and various database systems is made possible by the provided ciphertext retrieval solution. By preventing malevolent hackers, we demonstrate that the system ensures data security. Thorough testing demonstrates that our design is effective and practicable.

Keywords: *Ciphertext retrieval, cross-database query, cloud-assisted Internet of Things (IoT), heterogeneous data sets, transparent access*.

[1],[2] Associate Professor, Department of Electronics and Communication Engineering,
[3].[4],[5],[6],[7] Final Year UG Students, Department of Electronics and Communication Engineering,
(Vel Tech High Tech Dr Rangarajan Dr Sakunthala Engineering College Chennai, India)

DOI:10.31838/ecb/2023.12.s1-B.225

I. INTRODUCTION

Research interest in ciphertext retrieval technologies has grown over time. As a result the suggested approach proves to be more effective for the CRHM system, where the Public Cloud Server and the Private Cloud Server collaborate to execute the Ciphertext Retrieval process. While maintaining relatively high retrieval accuracy CRHM has high index generation and retrieval efficiency. A network of physical devices possessing distinct IP addresses and the capability to sense and gather information from the external environment is referred to as a cloud network. Transparent ciphertext retrieval systems, among other distributed applications, have grown significantly as a result of heterogeneous cloud computing. Users typically encrypt sensitive information before transmitting it to cloud servers since the public cloud services cannot be entirely trusted. The utilization of actual computing devices is made possible by the cloud, which offers resources like abundant storage space and powerful processing at little cost. Cloud technology proves beneficial for hosting websites and mobile applications by allowing resource scalability at a minimal expense.

In the cloud, retrieving ciphertext from heterogeneous datasets is a major challenge. The cloud increases the value of the data by analyzing and sharing it. Prior ciphertext retrieval methods can reduce the workload and ease security worries to some extent, but they are unable to handle encrypted diverse datasets. This work suggests a transparent ciphertext retrieval method that is unrestricted by programming language, accessible platforms, or dataset heterogeneity to address this problem. The ciphertext retrieval system makes it easier for many applications and database systems to share data. Hence, the technology guarantees data security by keeping out risky hackers. Yet, the plan is workable and effective.

II. LITERATURE SURVEY

[1] **Li Yong et al [2021]** To implement data storage and sharing in government cloud systems, he suggested ensuring privacy protection is a crucial concern that requires resolution. This research proposes a Top-k cyphertext retrieval scheme with keyword semantic extension in a hybrid government cloud environment. The aim is to fulfill the privacy protection requirements of document security sharing and data security storage between authorized data access users and also data publishers over the government cloud platform.

[2] **Jung-Shian Li et al [2020]** proposed that the demand for secure cloud storage, search over encrypted datasets, and the quick development of cloud services have all become crucial issues. Images of identification and driver's licenses that have recently leaked have drawn a lot of attention. Asymmetric scalar-product-preserving encryption (ASPE) and holomorphic encryption are two instances of the growing trend in secure computing (HE).

[3] **Ion-Dorinel Filip et al [2019**] proposed that one of the most promising technologies is robotics, and that non-humanoid robots, such as self-driving automobiles, have a variety of processing needs.The specifications stem from the robots' need to function in constantly changing surroundings and swiftly implement intricate decision-making algorithms in real-time. Additionally, there are requirements for processing large datasets that are compatible with batch processing.

[4] **ANAT GOLDSTEIN [2022]**

He talked on how the adoption of precision agriculture practises and the massive data collection that has resulted from the IT and big data revolution in agriculture. It is necessary to have the ability to combine data from various sources, analyse the data quickly, and deduce intelligent conclusions in order to use these data for effective decision support.

[5] **Heng He et al [2021]** proposed in recent years as more and more individuals and organizations entrust the storage of their data and software to cloud servers. Users typically encrypt sensitive information before transmitting it to cloud servers since the public cloud services cannot be entirely trusted. As a result, cypher text retrieval technology has steadily grown in popularity among researchers. The present related schemes have shortcomings, such as their poor accuracy, security, and retrieval effectiveness, as well as their inability to handle "many owners" mode and multi-keyword retrieval.

[6] GAI Keke et al [2016]hypothesized that the real-time performance of telehealth systems has significantly improved thanks to the broad adoption of heterogeneous cloud computing. Healthcare practitioners can obtain medical data from a range of sources that are supported by various cloud service providers using a cloud-based telehealth system. An alternative approach to enable multiple data users to share data is to utilize data duplication in distributed cloud databases.

III. PROPOSED SYSTEM

The proposed model for retrieving ciphertext in cloud assisted IoT consists of four main entities, the IoT operator the encrypted cloud database system the encrypted public cloud database and the information system with multiple subsystems.

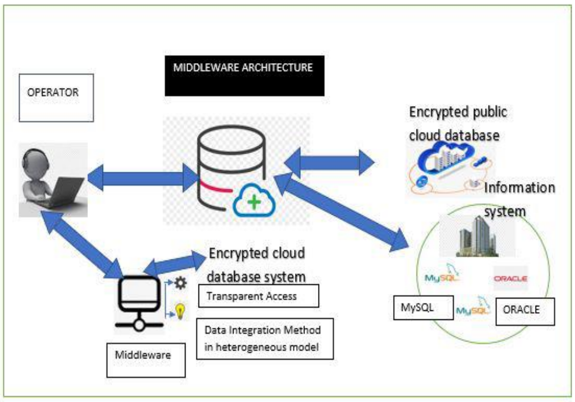

OPERATOR
MIDDLEWARE ARCHITECTURE
Encrypted public cloud database
Information system
Encrypted cloud database system
Transparent Access
Data Integration Method in heterogeneous model
Middleware
MySQL
ORACLE
MySQL
ORACLE

Fig 3.1 Cipertext Retrival System

To encrypt IoT data, the information system with diverse datasets communicates with encrypted cloud database system and transmit the encrypted data to the encrypted public cloud database for storage. The encrypted data is dispersed over different database platforms using the secured public cloud database.The IoT operator requests information via the encrypted public cloud database in order to obtain data from the information system.The information system's data must first be encrypted within the encrypted public cloud database system before the ciphertext data can be decrypted.

A.Access and Integration Middleware Architecture

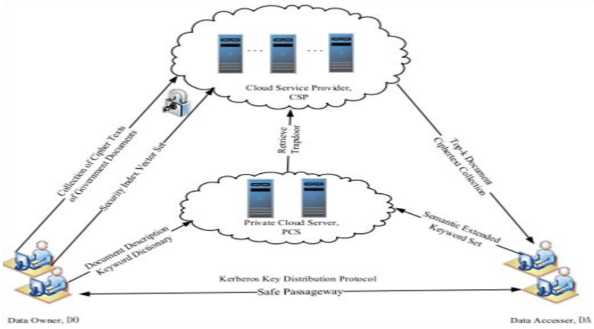

Cloud Service Provider, CSP
Collection of Cipher Texts of Government Documents
Security Index Vector Set
Retrieve Trapdoor
Top-k Document Ciphertext Collection
Private Cloud Server, PCS
Document Description Keyword Dictionary
Semantic Extended Keyword Set
Kerberos Key Distribution Protocol
Safe Passageway
Data Owner, DO
Data Accesser, DA

Fig 3.2 Hybrid government cloud system architecture

I. It allows transparent access to relational MSSQL, Oracle, SQL Server, and MySQL databases that are encrypted. Additionally, it provides data transfer security on the open channel and prevents hackers and cloud service providers from accessing important information.

II.Encrypted heterogeneous databases are supported for cross-database searches. It is possible to perform joint queries and integrate data from multiple databases containing encrypted data across different tables.

B. Architecture of a Ciphertext Attribute-Based System for hybrid cloud storage

The system's architecture enables users to log in and upload files. Upon uploading, the files undergo encryption, following which a key is generated and stored in a private cloud. The process then verifies whether the files are duplicates. Only once the user's credentials have been validated against the data sets are they granted access. The system starts with authorization before establishing control, generating keys, and storing uploaded files in a hybrid cloud for subsequent access. To retrieve a file, the system first checks for duplicates of the uploaded file. Once the user's credentials are authenticated, a key is exchanged, and the file can be retrieved.

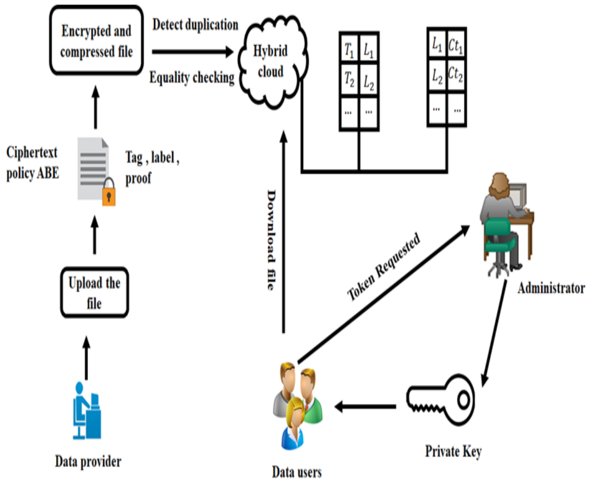


Fig 3.3 Hybrid cloud storage system architecture for the CipherText Attribute-Based System

C. Generation Of Authorization Controls And Keys

An authorized user can generate queries for specific files and the rights they own using a private cloud by using their unique private keys. Only those users who have the necessary log-in information can upload files to the cloud storage. Unauthorized users who lack the necessary permissions or access to a file can be prevented from doing so. When receiving a request from users, the S-CSP executes a duplicate check in the system. The system authenticates the user, and authorization to upload a file or data is only granted after concluding that they are legitimate. Any user that accesses the private cloud server without first requesting a file token must be unable to obtain any valuable information from the token.

a)Adding a File to a Hybrid Cloud

The introduction of hybrid cloud architecture addresses the issue of de-duplication and verifies authorized users. The encryption and key generation processes are done on a private cloud. The encrypted data will be managed, stored, and deduplication checked in the public cloud. Authorized users private keys are managed by the hybrid cloud and are not directly provided to them..

b)Detect Deduplication

The tag cannot be used to determine the key and jeopardize the security of the data because the key and tag are separately derived. The server side will keep a copy of the encrypted data as well as the tag that goes with it. Now that the tags are being compared using an equality checking technique, if they match, the user will be informed and the previously submitted file will be denied access.

c)Exchanging keys

In the private cloud, encrypted files and the private keys of authorized users are handled and kept. File token requests from users are considered only after their identity is confirmed, and the requested file is taken into consideration. Users can submit data and queries through the private cloud's interface for safe storage and computation. After confirming the user's identity, the private cloud server gives them the associated file token or key. Before uploading this file, the user may execute an authorized duplicate check with the public cloud.

d)Verification And File Retrieving

A set of keys is chosen as a symmetric key for each user and also sent to private cloud. An identification protocol is defined, which is analogous to the proof and verification algorithms respectively. The SCSP receives a request that includes the file name. Upon receiving the request and file name, the SCSP determines whether the user is an authorized to download the file. If unsuccessful, the SCSP notifies the user of the unsuccessful download by returning an abort signal. The user employs the localized key to restore the original file upon receiving encrypted data from the SCSP. In this scenario, the SCSP produces the corresponding cipher text, CF.

D.Ciphertext ABE algorithm:

In the Cipher text policy attribute-based encryption system a user's private key is linked to a set of attributes that define the user permissions. A collection of characteristics is chosen when a cypher text is encrypted, and only people who have access to those attributes can decrypt the cypher text.

Setup (): A reputable organization is in charge of this algorithm. After a security parameter is supplied, the programme outputs master secret key and a public parameters.

Key Generation (): The trustworthy authority oversees this algorithm as well. It receives a set of user attributes the master secret key MK and the public parameters PK as inputs, The output of the process is the secret key for a user who possesses the attribute set The output comprises two components and can be utilized by proxy B to aid in decryption while is used directly by the user to recover a plaintext message from the partially decrypted ciphertext generated by proxy B.

Policy Creation (): The final ciphertext CT is produced by proxy A using this technique, The creation of the cryptographic access policy associated with the access tree involves utilizing the partial ciphertext and the proxy encryption secret key as inputs.

Policy Verification (): The final ciphertext CT is produced by proxy A using this technique, The proxy encryption secret key and the partial cipher text are inputs used to build the access trees associated cryptographic access policy.

Decryption(): The user executes the decryption algorithm. The partially decrypted cipher text as well as the user's secret key are inputs into the decryption process. The outcome of this stage is either the decrypted message in case of a successful execution, or an error message otherwise.

E. Equality Checking Algorithm

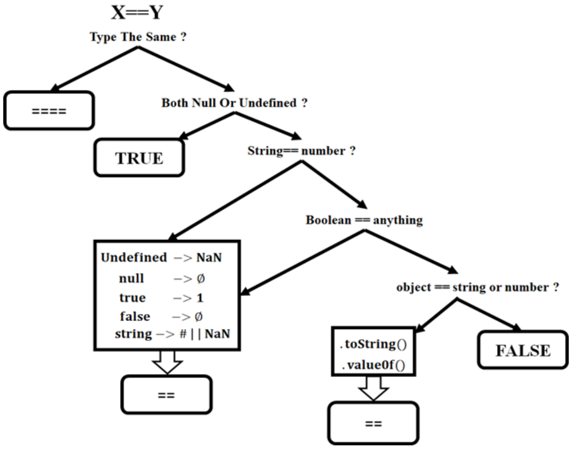


Fig 3.4 Equality Checking Algorithm

This encryption's most basic configuration involves creating a hash from an uploaded file. Next, encrypt the remaining portions of the file using this hash as the key. Finally, the hash key is encrypted and saved using the password. Only if the user has the original file can they obtain the password. The same hash and encrypted file must be calculated when there are two or more people using the same file. One copy is needed because the encrypted version is identical. The user must first use their password to decrypt the hash, then use the hash to decode the file in order to decrypt it.

STEP 1: Client! Server: The server responds with the tag of the current node is the (gri; gr0 h(mi)) when the client requests the deduplication of fresh data (m_). The current node is initially the trees root and its tag is (gr0; gr0 h(m0)).

STEP 2 : Client side : The client compute gri_h(m_) and verifiesgri_h(m_)?=gri_h(mi).

STEP 3: Client! Server : If the client notifies the server of a duplicate. If not it computes 1g and sends b to the server.

STEP 4 : Server side : The equality checking algorithm involves several steps. First the server adjusts the tree's current pointer based on the value of b If b is equal to 0 the server moves the pointer to the leftmost child and if not it moves it to the appropriate child. The algorithm then loops back to step 1. The algorithm terminates when the server comes across duplicate find or needs to transfer the reference to an empty node.

The results of our tests confirm that the ciphertext retrieval system performs as expected and Despite the inclusion of access and integration middleware across the system and providing operators with transparent access and cross database queries to encrypted heterogeneous databases the system still operates efficiently in comparison to the existing solution. The heterogeneous multidatabase searches take longer than single-database inquiries with the same amounts of accessed data and querying times because interactions between various database types are required. Most crucially, employing simply the encrypted cloud database system will not allow for cross-database searches.

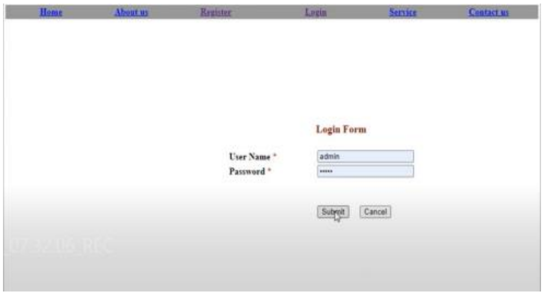
Encryption Process

The plaintext and key are the two inputs used by the encryption algorithm. has only one output, which is the ciphertext. A key is a secret that controls the encryption algorithm's output; various keys will result in a variety of ciphertexts. This component lets you control who has access to the message because only those with the key can decrypt it. An algorithm for key generation can produce keys. These are the instructions that were used to make the key. The keys used in current encryption are apparently random strings of numbers and letters, but it is still difficult to create keys that are truly random and cannot be guessed. In many cases, keys are pseudorandom, which implies that upon analysis, they don't actually appear to be random.It's possible that this is just the opposite of the instructions that were used to encrypt the plaintext.

Decryption Process

Decrypted data can be restored to their original format after encryption, which typically involves a reverse process. Encrypting and decrypting data is often justified by the need for privacy. This is particularly important when information is transmitted over the Internet, as unauthorized access from third parties can compromise its confidentiality. By encrypting data, it is protected against potential theft and loss. Encrypted data can include various types of files such as text files, images, emails, user data, and directories. When attempting to access the encrypted data, the recipient must enter a password or provide some form of authentication to initiate the decryption process. Data can be manually or automatically decrypted. A set of keys or a password might also be used to carry it out. One of the most significant and popular conventional cryptography techniques is the Hill cipher Encryption and Decryption procedure, which involves creating a random Matrix and is fundamentally important for security. In the Hill cipher, decryption is dependent on the inverse of the matrix. A challenge arises during decryption in the Hill cipher due to the non-existence of the inverse of the matrix in some cases.The encrypted content cannot be decoded if the matrix is not invertible. The improved Hill cipher algorithm totally eliminates this flaw.

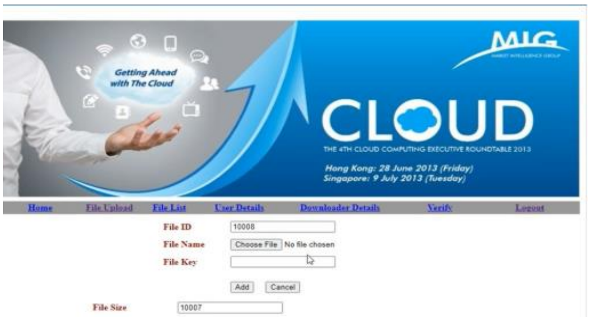
IV. RESULT AND DISCUSSION



Login Form

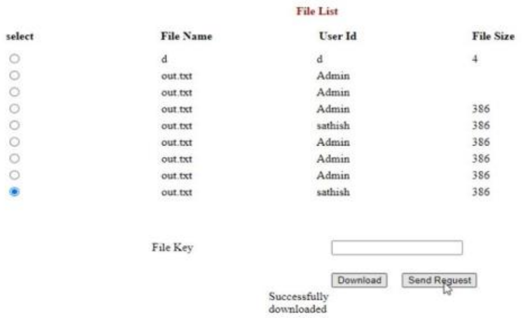
User Name * admin
Password * •••••

Figure 4.1 Log in Form


Getting Ahead with The Cloud
MIG
CLOUD
THE 4TH CLOUD COMPUTING EXECUTIVE ROUNDTABLE 2013
Hong Kong: 28 June 2013 (Friday)
Singapore: 9 July 2013 (Tuesday)
Home
File Upload
File List
User Details
Downloader Details
Verify
Logout
Choose File
Add
Cancel

File ID 10008
File Name No file chosen
File Key
File Size 10007

Figure 4.2 Encryption



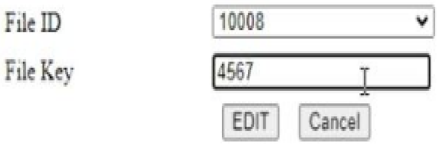
File List

select	File Name	User Id	File Size
○	d	d	4
○	out.txt	Admin	
○	out.txt	Admin	
○	out.txt	Admin	386
○	out.txt	sathish	386
○	out.txt	Admin	386
○	out.txt	Admin	386
○	out.txt	Admin	386
●	out.txt	sathish	386

File Key

Successfully downloaded

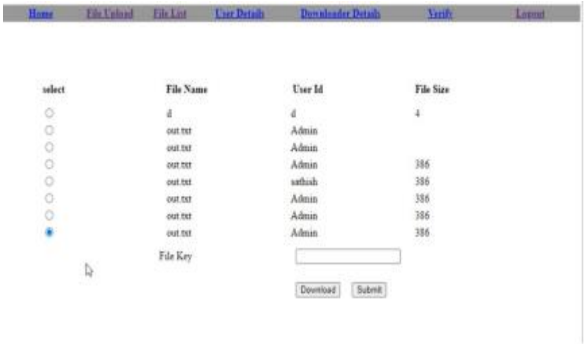
Figure 4.3 Decryption



File ID: 10008

File Key: 4567

Figure 4.4 File Key



select	File Name	User Id	File Size
○	d	d	4
○	out.txt	Admin	
○	out.txt	Admin	
○	out.txt	Admin	386
○	out.txt	sathish	386
○	out.txt	Admin	386
○	out.txt	Admin	386
●	out.txt	Admin	386

File Key

Figure 4.5 File

The Transparent Ciphertext Retrieval System which supports the integration of encrypted heterogeneous databases is a feasible approach for protecting confidential data in cloud environments. The system is very efficient and successful at protecting data confidentiality and privacy according to the studies evaluation of its performance in extracting data from encrypted heterogeneous databases. The system is a suitable alternative for protecting data in various cloud based applications due to its low overhead and high retrieval accuracy. Furthermore the transparency function of the system guarantees that users can access data without jeopardizing its security. The results are discussed highlighting the system's potential to improve data security in cloud environments, especially in situations where multiple entities must access sensitive data. In conclusion the results of the research indicate that the Transparent Ciphertext Retrieval System Supporting Integration of Encrypted Heterogeneous Database is an effective approach for securing confidential information in cloud environments.

V. CONCLUSION

Security is now one of the most crucial factors in every industry. Every piece of information needs to be protected because any alterations cause very serious issues. The protection of data from harmful assaults and illegal access is necessary. The cloud-assisted model uses ciphertext retrieval under heterogeneous datasets to secure data. Although the current approaches can somewhat lessen workload and relieve security concerns, they are unable to manage encrypted varied datasets. To solve this problem, a transparent ciphertext retrieval approach is proposed in this study. Programming language, accessible platforms, and dataset heterogeneity are not factors in the method. Also, this technology safeguards data security by fending off hostile hackers. It is effective and practicable. The proposed method takes less time for authentication purposes as well as for encrypting and decrypting data while exchanging files in a distributed database system.

VI. REFERENCES

1. Li Yong, Liu Hefei, Shen Xiujuan, Yuan Bin, Wang Kun, 2021, "Keyword Semantic Extended Top-k Cipher text Retrieval Scheme Over Hybrid Government Cloud Environment", IEEE Access, vol. 06669, pp. 155249 – 155259.

2. Jung-Shian Li, I-Hsien Liu, Chin-Jui Tsai, Zhi-Yuan Su, Chu-Fen Li, Chuan-Gang Liu, 2020, "Secure Content-Based Image Retrieval in the Cloud with Key Confidentiality", IEEE Access, vol. 08, pp. 114940 – 114952.

3. Ion-Dorinel Filip, Andrei Vlad Postoaca, Radu-Dumitru Stochitoiu1, Darius-Florentin Neatu1, Catalin Negru1, Florin Pop, 2019, "data capsule : representation of heterogeneous data in cloud – edge computing" IEEE Access, vol. 07, pp. 49558 – 49567.

4. Anat Goldstein, Lior Fink, Gilad Ravid, 2022, "A Cloud-Based Framework for Agricultural Data Integration: A Top-Down–Bottom-Up Approach", IEEE Access, vol. 10, pp. 88527 – 88537.

5. Heng He, Renju Chen, Chengyu Liu, Ke Feng, Xiaohu Zhou, 2021, "An Efficient Cipher text Retrieval Scheme Based on Homomorphic Encryption for Multiple Data Owners in Hybrid Cloud", IEEE Access, vol. 09, pp. 168547 – 165887.

6. Gai Keke, Qiu Meikang, Sun Xiaotong, Zhao Hui, 2016, "Smart data deduplication for telehealth systems in heterogeneous cloud computing", IEEE Access, vol. 01, pp. 93 – 104.

7. Jidong Xiao, Lei Lu, Hai Huang, Haining Wang, 2018, "Virtual Machine Extrospection: A Reverse Information Retrieval in Clouds", IEEE Access, vol. 09, pp. 401 – 413.

8. Hanli Wang, Bo Xiao, Lei Wang, Fengkuangtian Zhu, Yu-Gang Jiang, Jun Wu, 2015, "CHCF: A Cloud-based Heterogeneous Computing Framework for Large-Scale Image Retrieval", IEEE Access, vol.09, pp. 401 – 413.

9. Die Wang, Catherine Prigent, Filipe Aires, Carlos Jimenez, 2017, "A Statistical Retrieval of Cloud Parameters for the Millimetre Wave Ice Cloud Imager on Board MetOp-SG", IEEE Access, vol. 05, pp. 4057 – 4076.

10. Heng He, Renju Chen, Chengyu Liu, Ke Feng, Xiaohu Zhou, 2021, "An Efficient Ciphertext Retrieval Scheme Based on Homomorphic Encryption for Multiple Data Owners in Hybrid Cloud", IEEE Access, vol. 09, pp. 168547 – 168557.

11. Shridevi Soma, Vinaya S Kavalgi, 2019, "Efficient Multi-Keyword Search Through Ciphertext Data in the Cloud", International Journal of Engineering and Advanced Technology, vol. 08, DOI: 10.35940/ijeat.F8534.088619.

12. Siti Dhalila Mohd Satar, Mohamad Afendee Mohamed, Masnida Hussin, Zurina Mohd Hanapi, Siti Dhalila Mohd Satar, 2021, "Cloud-based Secure Healthcare Framework by using Enhanced Ciphertext Policy Attribute-Based Encryption Scheme", International Journal of Advanced Computer Science and Applications, vol. 12, no. 06.

13. Niloufer Rafath, Wahaj Ghouri, Syed Raziuddin, 2015, "Security in Cloud using Ciphertext Policy Attribute-Based Encryption with Check ability", International Journal of Innovative Research in Computer and Communication Engineering, vol. 03, DOI: 10.15680/ijircce.2015.0305062.

14. Mohd Muhibuddin, Dhanunjayudu. K, 2021, "Design of Secure and Efficient Product Information Retrieval System Model in Cloud Computing", Journal of Emerging Technologies and Innovative Research, vol. 08, pp. 460 – 467.

15. Shobhanjaly P Nair, Dr. A. Kathirvel, 2018, "Cipher text attribute-based encryption algorithm to confiscate deduplication in hybrid cloud storage", International Journal of Advance Research, Ideas and Innovations in Technology, vol. 04, pp. 1678 – 1686.