RESEARCH ARTICLE OPEN ACCESS

MULTI-KEYWORD SEARCH FOR MULTIPLE DATA OWNERS USING ID-BASED ENCRYPTION TECHNIQUE

*Arumugam. P, **Pavithra. M, ***Shyamala. S, ****Ezhilarasan. M

*(Information Technology, Pondicherry Engineering College, Pondicherry)
** (Information Technology, Pondicherry Engineering College, Pondicherry)
*** (Information Technology, Pondicherry Engineering College, Pondicherry)
****(Professor,Information Technology, Pondicherry Engineering College, Pondicherry)

--************************----------------------------------

Abstract:

Cloud computing is a standout amongst the most famous innovation among the new rising advancements in the field of web and PCs. With the expanding appropriation of cloud computing, a developing number of clients re-appropriate their documents to the cloud. To protect the security, the datasets are normally scrambled before redistributing. In any case, the regular routine with regards to encryption makes the powerful use of the information troublesome. For instance, it is hard to look through the given watchwords in scrambled records. Customary symmetric encryption strategy won't fit into the security bend where there is an increase in the number of information proprietors and clients, even the information proprietors can't have their very own documents. Consequently, ID-based Encryption procedure has been proposed for productive recovery of information from the cloud.

Keywords **—**MRSE, IBE, PKG, PKI

--************************----------------------------------

I. INTRODUCTION

Cloud computing is a computing paradigm, where an extensive pool of frameworks are associated in private or open systems, to give powerfully versatile foundation to application, information, and record stockpiling. With the appearance of this innovation, the expense of calculation, application facilitating, content stockpiling and conveyance are decreased altogether. Cloud computing is a viable way to deal with experience direct money-saving advantages and it can possibly change a server farm from a capital-serious set up to a variable evaluated condition. Cloud registering depends on an essential vital of 'reusability' of IT industries. The distinction that cloud computing brings contrasted with conventional ideas of "lattice figuring", "disseminated registering", "utility processing", or "autonomic processing" is to widen skylines crosswise over authoritative limits. Cloud computing incorporates a gathering of PCs that are mutually used to give distinctive calculations and errands. Cloud computing is a standout amongst the most imperative IT standards over the most recent couple of years. One of the key advantages that are offered from this IT innovation for the organizations is decreased time and expenses available. Cloud computing is giving organizations and associations to utilize shared capacity and registering assets. It is superior to create and work with the claim foundation. Cloud computing additionally gives associations and organizations to have an adaptable, secure, and financially savvy IT foundation. There are diverse sorts and models in cloud computing with respect to the distinctive gave administrations. Along these lines, cloud computing includes open cloud, private cloud, half and half cloud, and network cloud. Administration conveyance models, then again, could be arranged as SaaS (Software as an administration), PaaS (Platform as a Service), and IaaS (Infrastructure as a Service).

The existing encryption techniques makes retrieval of documents difficult as data users need to manage multiple keys for different data owners. Data users need to generate multiple trapdoors for data owner's data even for the same query condition. In the paper, ID-based encryption has been proposed which results in efficient data retrieval from cloud without the need to manage multiple encryption keys by the data users for the same data.

II. LITERATURE SURVEY

Encryption plays a major role in retrieval of data from cloud securely. The two major types of encryption are symmetric and asymmetric encryption.

Symmetric Encryption [1] is the simplest kind of encryption that involves only one secret key to cipher and decipher information. It utilizes a secret key that can either be a number, a word or a string of random letters. It is combined with the plain text of a message to change the content in a particular way. The sender and the recipient should know the secret key that is used to encrypt and decrypt all the messages. DES, Triple-DES, Blowfish, RC5 and Advanced Encryption Standard are examples of symmetric encryption. The main limitation of the symmetric key encryption is that all parties involved have to exchange the key used to encrypt the data before they can decrypt it.

Asymmetric encryption is also known as public key cryptography, which is a relatively new method, compared to symmetric encryption. It uses two keys to encrypt a plain text. A public key is made available to anyone who wants to send a message. The second private key is kept secret so that sender can only know. A message that is encrypted with the public key can only be decrypted using a private key, while also, a message encrypted using a private key can be decrypted using a public key. Security of the public key is not required because it is publicly available and can be passed over the internet.Diffie Hellman key exchange algorithm, RSA (Rivest–Shamir–Adleman), ID-Based Encryption, Elliptic Curve Cryptography (ECC), Digital Signature Algorithm (DSA), Attribute based encryption(ABE) are some examples of asymmetric encryption. Asymmetric key has a far better efficiency in ensuring the security of data transmitted during communication.

Cong Wang [2] proposed search which solves processing overhead, data and keyword privacy, minimum communication and computation aerial. The information proprietor constructs file alongside the keyword frequency based relevance scores for documents. The user request to Cloud Server using the private key. The Cloud Server searches the index with scores and sends encrypted file based on the ranked sequence. It does not perform multiple keyword searches. Minimal overhead in index building.

Jiadi [3] proposed search using two round searchable encryption (TRSE). In the first round, users present multiple keyword 'REQ' 'W' as the encrypted query for accomplishing data, keyword privacy and create trapdoor (REQ, PK) as Tw and send to Cloud Server. Then Cloud Server figure out the score from the encrypted index for files and returns the encrypted score result vector to the user. In the second round, user decodes N with the secret key and calculates the file ranking and then request files with Top k scores. The ranking of record is done on the client side and scoring is done on the server side. The contraction and confining are utilized to decrease cipher text size, still the key size is too large. The communication aerial will be very high if the encrypted trapdoor's size is too large. It does not make an effective searchable index update.

Ning [4] designed search for known cipher text model and background model over encrypted data providing low computation and communication overhead. The coordinate matching is chosen for multi-keyword search. They used inner product similarity to quantitatively evaluate similarity for ranking files. The limitation is that MRSE have small standard deviation σ which weakens keyword privacy.

Qin Liu [5] proposed search that provides keyword privacy, data privacy and semantic secure by public key encryption. Cloud Server is involved in partial decipherment by reducing the communication and computational aerial in the decryption process for end users. The user submits the keyword trapdoor encrypted by the user's private key to Cloud Server securely and retrieve the encrypted documents. The communication and computational expense for encryption and decoding are more.

A secure tree-based search scheme [6] over the encrypted cloud storage, which supports multi keyword ranked search along with dynamic operation on document collection available at server has also been designed. The vector space model and term frequency (TF) × inverse document frequency (IDF) model are combined and used in the construction of index and generation of query to provide multi keyword ranked search output.. The Searchable Symmetric Encryption has been applied to encrypt the index and query vectors, and till then ensure accurate relevance score calculation between encrypted index and query vectors.

Wenhai Sun [7] proposed Attribute-based Keyword search that facilitates conjunctive keyword search, keyword semantic security and Trapdoor unlinkability. The owners makes an index with all keywords and access list with policy attributes which specifies the user's list authorized for searching. Now owners encrypt the document, index with access list utilizing ciphertext policy attribute-based encryption technique. To have user membership management, they used proxy re-encryption and lazy re-encryption techniques to share the workload to Cloud Server. The user requests the Tw to Cloud Server using its private key. Now Cloud Server recovers Tw and searches the encrypted indexes and returns documents only if the user's attributes in Tw satisfies access policies in indexes which makes coarse-grained dataset search authorization. Trapdoor generation will need additional time with the increased number of attributes.

In the paper, the proposed technique Identity Based Encryption has characterized and solved the problem of multi- keyword search over encrypted cloud data without the need of acquisition of multiple encryption keys by the data users with less retrieval time compared with existing searchable symmetric encryption technique.

III. PROPOSED SYSTEM

3.1 Problem Definition

The retrieval of documents owned by multiple data owners becomes difficult for data users because of different encryption keys used by data owners causing the delay in retrieval time. With the help of the ID-based encryption technique, this issue can be eliminated.

3.2 System Model

In our multi-owner and multi-user cloud computing model, four entities are involved, as illustrated in Fig.1, they are data owners, the cloud server, administration server, and data users. Data owners have a collection of files F. To enable efficient search operations on these files which will be encoded, data owners first build a secure searchable index I on the keyword set W extracted from F, then they submit I to the administration server. Finally, data owners encrypt their files F and re-appropriate the corresponding encrypted files C to the cloud server. After accepting I, the administration server re-encrypts I for the authenticated data owners and outsources the re-encrypted index to the cloud server. Once a data user wants to search t keywords over these encrypted files stored on the cloud server, he first computes the corresponding trapdoors and submits them to the administration server. Once the data user is confirmed by the administration server, the administration server will further re-encrypt the trapdoors and submit them to the cloud server. Upon receiving the trapdoor T, the cloud server searches the encrypted index I of each data owner and returns the corresponding set of encrypted files. To improve the file retrieval accuracy and save communication cost, a data user would tell the cloud server a parameter k and cloud server would restore the top-k relevant files to the data user. Once the data user receives the top-k encrypted files from the cloud server, he will decrypt these returned files.

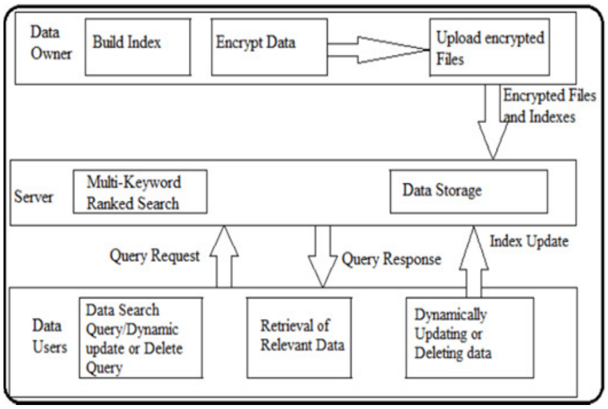
Data Owner
Build Index
Encrypt Data
Upload encrypted Files
Encrypted Files and Indexes
Server
Multi-Keyword Ranked Search
Data Storage
Query Request
Query Response
Index Update
Data Users
Data Search Query/Dynamic update or Delete Query
Retrieval of Relevant Data
Dynamically Updating or Deleting data

Figure. 1 Architecture Diagram for Data retrieval

3.3 ID-Based Encryption Technique

Identity-Based Encryption (IBE) is the best alternative to public key encryption, which is proposed to simplify key management in a certificate-based Public Key Infrastructure (PKI) by using human-intelligible identities (e.g., remarkable name, email address, IP address, etc) as public keys. Therefore, sender utilizing IBE does not need to look up public key and certificate, but directly encrypts the message with the receiver's identity. Accordingly, receiver acquiring the private key associated with the corresponding identity from Private Key Generator (PKG) is able to decrypt such ciphertext.

An IBE scheme which typically involves two entities, PKG and users (including sender and receiver) has consisted of the following four algorithms.

- Setup(λ) : The setup algorithm takes as input a security parameter λ and outputs the public key PK and the master key MK. Note that the master key is maintained secret at PKG.
- KeyGen(MK,ID) : The private key generation algorithm is run by PKG, which takes as input the master key MK and user's identity ID ∈ {0, 1}∗. It restores a private key SKID corresponding to the identity ID.
- Encrypt(M,ID) : The encryption algorithm is run by the sender, which takes as input the receiver's identity ID and a message M to be encrypted. It outputs the ciphertext CT.
- Decrypt(CT,SKID) : The decryption algorithm is run by the receiver, which takes as input the ciphertext CT and his/her private key SKID . It restores a message M or an error ⊥.

An IBE scheme must fulfill the definition of consistency. Specifically, when the private key SKID generated by algorithm KeyGen when it is given ID as the input, then Decrypt(CT,SKID) = M where CT = Encrypt(M,ID).The motivation of IBE is to simplify certificate management. For instance, when Alice sends a message to Bob at bob@company.com, she basically encode her message with Bob's email address "bob@company.com", yet does not need to acquire Bob's public key certificate. At this point, when Bob receives the encrypted email he confirms himself at PKG to obtain his private key, and read his email with such a private key.

3.4 Modules Description

3.4.1 Admin/Data Owner Login Module

This module is used for login purpose. In this module, the admin or data owners can enter their user-name and their password. If both username and password matches with the database then it goes to the home page.

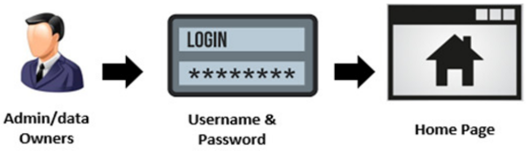

Figure.2 Admin/Data Owner Login Module

3.4.2 Video Upload Module

In this module, the data owners is going to upload the Video like 2D, 3D and other relevant videos in the server space.

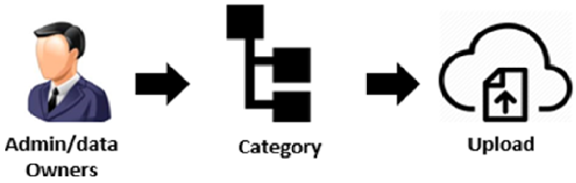

Figure.3 Video Upload Module

3.4.3 Admin Authentication Module

The admin has an important role in giving approval to the data owner in uploading the video. Only if admin gives approval, the data gets stored in database or else the data doesn't get stored.

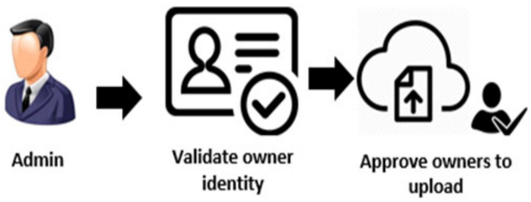

Figure. 4 Admin Authentication Module

3.4.4 Data Encryption Module

This module consists of adding the signature with the help of the ID which can be used by the recipient of a message to verify that the message has not been altered during transmission as well as to verify the originator's identity so that the integrity of the data and programs may be verified at any later time.

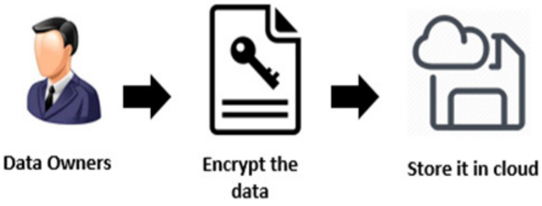

Figure. 5 Data Encryption Module

3.4.5 Multi-keyword Search Module

This module helps in retrieving the data from the database that correspond to multi- keywords or characters specified by the user.

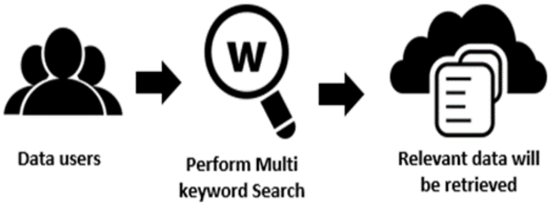

Figure. 6 Multi-keyword Search Module

3.4.6 .Data Request Module

The Data User needs to request for the complete data with user identity (uid) key. The Data Owner will verify the uid Key and approve the data to be accessed with the data owner. Once the data owner gave access, the data will be visible to the user requested.

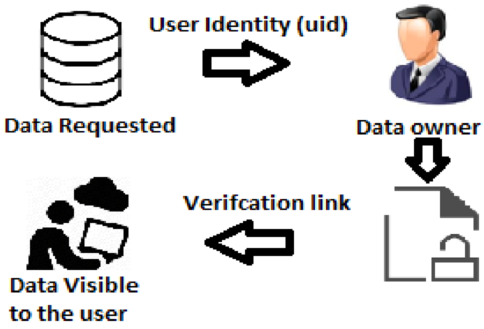

Fig.7. Data Request Module

3.5 Complexity involved in the proposal

The complexity involved in the proposal is to reduce the acquisition of multiple keys by the data users which helps in reducing the computational time for searching the files using ID-Based Encryption technique.

IV. RESULT ANALYSIS

The implementation has been done by using PHP language with SQL database as backend in Windows 10 environment. The response time of video retrieval based on the designation of data users using Identity Based Encryption has been compared with existing Searchable Symmetric encryption.

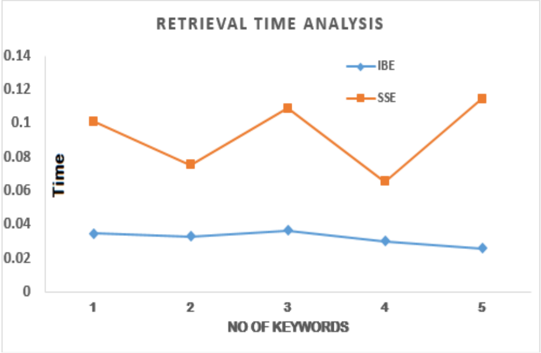

Fig 8. Response Time Graph for Data retrieval

V. CONCLUSION

In our proposed work, the problem of multi-keyword search for multiple data owners and multiple data users in the cloud computing environment has been analysed. Different from prior works, the proposed scheme enable authenticated data users to achieve secure, convenient, and efficient searches over multiple data owner's data. To efficiently authenticate data users and detect attackers who steal the secret key and perform illegal searches, we use ID-based Encryption, a public key cryptography algorithm. To empower the cloud server to perform secure search among multiple owner's data encrypted with different secret keys along with fast retrieval time, the ID-based Encryption plays a key role by limiting the use of more number of private keys for the data users in accessing the data with the help of Private Key Generator and reduces retrieval time of data.

REFERENCES

[1] S.Chandra , S. Paira , SkSafikulAlam ," A comparative survey of Symmetric and Asymmetric Key Cryptography", International Conference on Electronics, Communication and Computational Engineering, Pg no 83-93, 16 April 2015.

[2] Cong Wang et al., "Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data", IEEE Transactions on parallel and distributed systems, Vol. 23, no. 8, August 2012.

[3] Jiadi Yu, Peng Lu, Yanmin Zhu, GuangtaoXue, IEEE Computer Society, and MingluLi, "Toward Secure Multi-keyword Top k Retrieval over Encrypted Cloud Data", IEEE Journal of Theoretical

and Applied Information Technology , Vol. 66 No.1, 10th August 2014.

[4] Ning Cao et al., “Privacy-Preserving Multi Keyword Ranked Search over Encrypted Cloud Data”, IEEE Transactions on parallel and distributed systems, vol. 25, no. 1, Jan 2014.

[5] Qin Liuy, GuojunWangyz, and JieWuz, “Secure and privacy preserving keyword searching for cloud storage services”, ELSEVIER Journal of Network and computer Applications, March 2011.

[6] TianyuePeng ,Yaping Lin ,XinYao, Wei Zhang, “An Efficient Ranked Multi-Keyword Search for Multiple Data Owners OverEncrypted Cloud Data”, IEEE-Translations and content mining, Volume 6, Pg no. 21924- 21932, May 9, 2018.

[7] Wenhai Sun et al., “Privacy-Preserving Multi keyword Text Search in the Cloud Supporting Similarity-based Ranking”, the 8th ACM Symposium on Information, Computer and Communications Security, Hangzhou, China, May 2013.

[8] Zhihua Xia, “A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data”, IEEE Transactions on Parallel and Distributed Systems, Vol. Pg No: 99 Year 2015.