Privacy Preserving Multi-Keyword Ranked Search Over Secured Cloud

Abhish Chalke[1], Aniket Bagwe[2], Sameer Chorge[3], Yogesh Gite [4]

[1]*Student, Computer Engineering, Dilkap Research institute of engineering & Management, Maharashtra, India*
[2]*Student, Computer Engineering, Dilkap Research institute of engineering & Management, Maharashtra, India*
[3]*Student, Computer Engineering, Dilkap Research institute of engineering & Management, Maharashtra, India*
[4]*Professor, Computer Engineering, Dilkap Research institute of engineering & Management, Maharashtra, India*

ABSTRACT

As the cloud technology is getting more and more popular day by day users are inspired to store their data on cloud. The clouds like amazon azure are getting more and more popular day by day. Clouds provide resilience and reasonable in price. But for privacy preserved it is necessary to encrypt that file before upload on cloud. Since there so many files are uploaded by various users on cloud so it is required to provide facility of searching not only searching but the multi keyword search with the proper ranking. Currently the search facilities which are provided are mostly a single keyword search or a Boolean search. So we think it is necessary to provide multi keyword ranked search to enhance the experience of user. From many of the algorithm, we used coordinate matching for getting more and more matches as possible and to give more flexibility towards users we also provide backup facility.

1. INTRODUCTION:

Cloud Currently we are in an information-explosion era where constantly purchasing new hardware, software and training IT professional is becoming a nightmare for almost every IT person. Coincidentally, we are witnessing an enterprise IT architecture which shifted to a centralized, more powerful computing paradigm known as Cloud Computing, in which enterprise's or personage's databases and applications are moved to the servers in the large data centres (i.e. the cloud) managed by the third-party cloud service providers (CSPs) in the Internet. Cloud computing has been recognized as the most momentous turning point in the development of information technology during the past decade. People are attracted by the benefits it offers, such as personal and flexible access, on-demand computing resources configuration, considerable capital expenditure savings, etc. Therefore, many companies, organizations, and individual users have adopted the cloud platform to improve their business operations, research, or everyday needs. Various domains where searching is performed on outsourced Cloud data are: Search Engine, where a document collection is outsourced to cloud storage and client can retrieve documents which contain the query keywords .

2. LITERATURE REVIEW:

In previous system i.e Secure and privacy preserving keyword search that provides keyword privacy, data privacy and semantic secure by public key encryption. CSP is involved in partial decipherment by reducing the communication and computational aerial in decryption process for end users. The trapdoor request is submitted by

user that is encrypted by DES algorithm. Which provide some security but that is not much.
The data owner build index but not provide keyword frequency based relevance scores for files.. The cloud server searches the index with scores and sends encrypted file based.

Limitation:

- It does not perform multiple keyword ranked searches.
- The communication and computational cost for encryption and decryption is more
- It's less secured.
- Does not take the backup for the data.

3. PROPOSED SYSTEM:

This proposed method has defined and solved the problem of effective but safe and sound rank keyword search over Encrypted cloud data. Ranked search greatly enhances system usability by returning the matching files in a ranked order regarding to certain important criteria. Preserve privacy and provide multi keyword search are one of the biggest challenges that are solved in this paper. We provide efficient ranking method. By each download the count is updated and by that count it gives the rank to each File. For privacy preserving we encrypt the file with AES algorithm of 256 bits

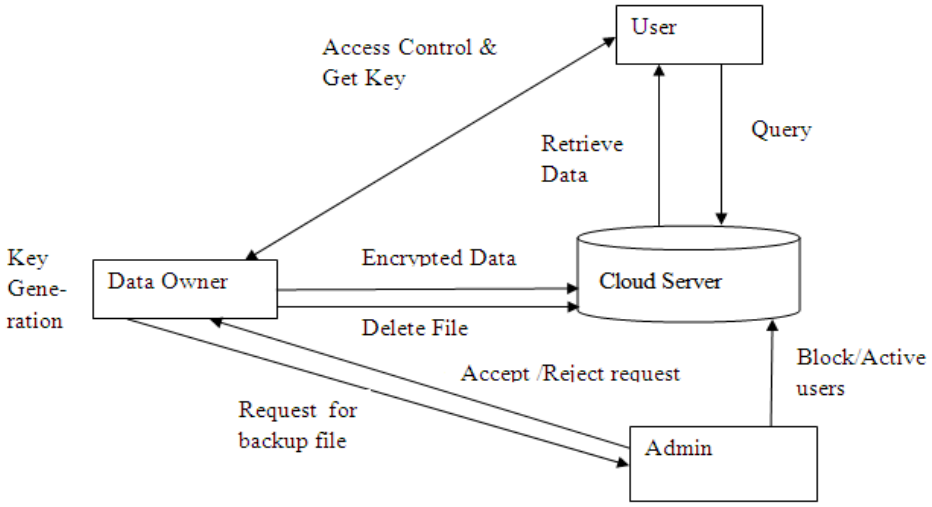

Fig.1 Proposed system

Data owner module:

In this module, the data owners should be able to upload the files. The files are encrypted before the files are uploaded to the cloud. These files when once available on the cloud, the data users should be able search using the keywords. The data owners will also be provided with a request approval screen so they are able to approve or reject the request that are received by the data users.

Data user module:

Data users are users on this system, who will be able to download files from the cloud that are uploaded by the data owners. Since the files stored on the cloud server could be in huge numbers, there is a search facility provided to the user. The user should be able to do a multi-keyword search on the cloud server. Once, the result appears for the specific search, these users should be able to send a request to the respective data owners of the file through the system for downloading these files. The data users will also be provided a request approval screen, where it will

notify if the data owner has accepted or rejected the request. If the request has been approved, the users should be able to download the decrypted file.

File Upload and Encryption Module:
In this module, the data owners should be able to upload the files. The files are encrypted before the files are uploaded to the cloud. These files when once available on the cloud, the data users should be able to search using keywords. The data owners will also be provided with a request approval screen so that they are able to approve or reject the request that are received by the data users. The file before upload will have to be encrypted with a key so that the data users cannot just download it without this key. The encryption of these files uses AES algorithm so that unauthorized users will not be able to download these files.

File Download and Decrypt Module:
Data users are users on this system, who will be able to download files from the cloud that are uploaded by the data owners. Since the files stored on the cloud server could be in huge numbers, there is a search facility provided to the user. The user should be able to do a multi-keyword search on the cloud server. Once, the result appears for the specific search, the users should be able to send a request to the respective data owners of the file through the system for downloading these files. The data users will also be provided a request approval screen, where it will notify if the data owner has accepted or rejected the request. If the request has been approved, the users should be able to download the decrypted file. The file before download will have to be decrypted with a key. Once the key is provided during the download, the data users will be able to download the file and use them.

Admin Module:
In this module, we provide some special functions to the admin. If data owner by mistake delete some data but then if owner want to restore that file then owner sends request to admin. Then admin has special authority to either accept that request or reject it. Admin can manage all the users, If admin thinks that some user is fake then admin can block it and admin also has authority to active that user again.

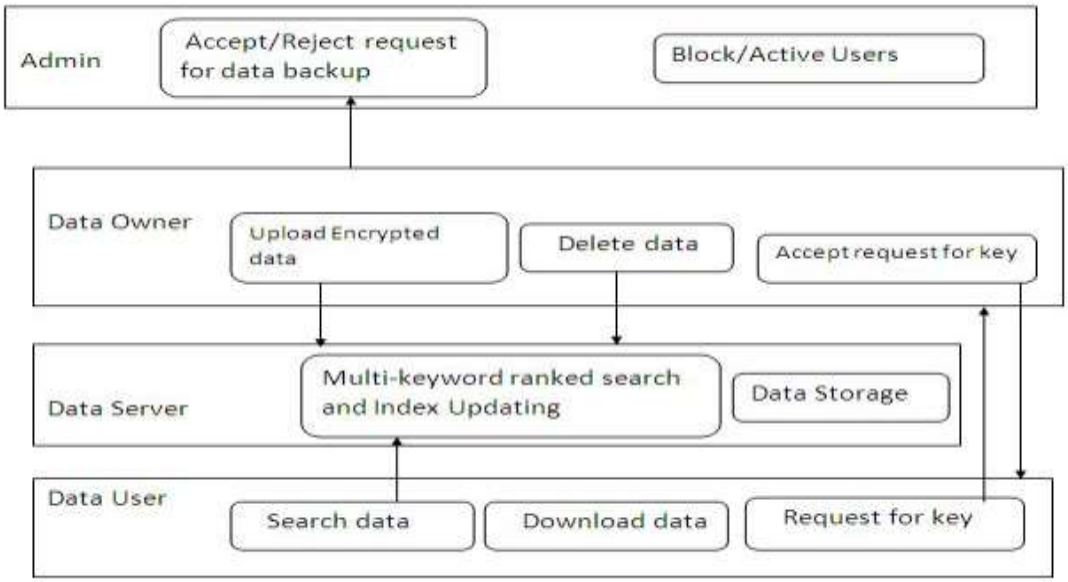

Fig. 2 System Architecture

4.Algorithm Used:

4.1 AES algorithm:

a) **Key Expansions**: By using rijndael's key schedule round keys are derived from ciphers . key block of 128 – bit

rounds are required for AES
b) **Initial Round**: Add Round Key: Each byte of the state is combined with a block of the round key using bitwise xor.
c) **Rounds:**

- Sub Bytes: A non-linear substitution step where each byte is replaced with another according to a lookup table.
- Shift Rows: A transposition step where the last three rows of the state are shifted cyclically a certain number of steps.
- Mix Columns: A mixing operation which operates on the columns of the state, combining the four bytes in each column.
- Add Round Key

d) **Final Round (no Mix Columns)**

- Sub Bytes
- Shift Rows
- Add Round Key

4.2 Seed Block Algorithm:

Initialization: Main cloud Mc; Remote Server Rs; Seed Block Sb; Random Number Rn;
Client Ci; Client ID: C_ID; File a1,a2;
Input: Ci create a1, r generated at Mc;
Output: Recover file a1;
Step 1: Generate Random number r;
Step 2: Generate Seed block for each Client Ci and store at Rs;
Sb = r Ex-OR C_ID ;
Step 3:If Ci creates/modifies a1 a2 stores at Mc, then a2 create as
a2= a1 Ex-OR Si;
Step 4: Store a2 T Rs;
Step 5: If server get crashed a1 deleted from Mc then, we do Ex-OR to retrieve the original file as: a1=a2 Ex- OR Si
Step 6: Return a1 to Si;
Step 7:END.

5. APPLICATION

- **Personalized Medication**: where patient's medical record is outsourced to hospital's server and an authorized doctor can perform secure searching on patient's medical record for diagnosis.
- **Email Server**: where a collection of private emails is outsourced to email server and client can retrieve pertinent emails based on the content of the mail/sender names/receiver names or email IDs.
- **Crime Investigation:** where the Interpol's criminal database acts as the server and clients are the authenticated crime investigation agencies like police departments.

6. CONCLUSION:

In this survey paper we have summarized one of the prominent search technique i.e. multi keyword ranked search and we apply this technique with the cloud server that enhance the user experience towards the use of cloud. But when we outsourced any data on cloud then the security is necessary for it. So for "preserved the privacy" of user we use AES algorithm encryption. We discuss the detailed steps of implementing the AES algorithm in this paper. AES is one of the reliable algorithm for providing encryption. When users are tends to use the cloud facility then they may do some mistakes as they are not the professionals users they might delete their data which is important for them that is why we include a admin module in this paper who can restore their data after the request of that data owner and admin can block the user as well as active the users. Hence this system is one of the most efficient model that enhance the user experience also gives the flexibility towards the data privacy and reduce the redundancy and provide scalability

7. REFERENCE:

[1] L. M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A break in the clouds: towards a cloud definition," ACM SIGCOMM Comput. Commun. Rev., vol. 39, no. 1, pp. 50–55, 2009.

[2] I. H. Witten, A. Moffat, and T. C. Bell, "Managing gigabytes: Compressing and indexing documents and images," Morgan Kaufmann Publishing, San Francisco, May 1999.
[3] D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. of S&P, 2000.
[4]E.-J. Goh, "Secure indexes," Cryptology ePrint Archive, 2003, http://eprint.iacr.org/2003/216.
[5] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in Proc. of ACNS, 2005.
[6] R. Curtmola, J. A. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in Proc. of ACM CCS, 2006.
[7] D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. of EUROCRYPT, 2004.
[8] M. Bellare, A. Boldyreva, and A. ONeill, "Deterministic and efficiently searchable encryption," in Proc. of CRYPTO, 2007.
[9] M. Abdalla, M. Bellare, D. Catalano, E. Kiltz, T. Kohno, T. Lange,J. Malone-Lee, G. Neven, P. Paillier, and H. Shi, "Searchable encryption revisited: Consistency properties, relation to sanonymous ibe, and extensions," J. Cryptol., vol. 21, no. 3, pp. 350–391, 2008.