Enabling Secure and Efficient Multi Keyword Ranked Search over Encrypted Cloud Data

Revathy B.D[#1], Anbumani .A[#2] , Rohith .V[#3]

[#1]*ME, CSE, Mahendra Institute of Technology,Mallasamuram,Namakkal-637503*

[#2]*AssistantProfessor, CSE, Mahendra Institute of Technology,Mallasamuram,Namakkal-637503*

[#3]*AssistantProfessor, CSE, Coorg Institute of Technology, Ponnampet-571216*

Abstract-Cloud computing is a new paradigm for new generation, organization and companies are started switching from local server to cloud server because of scalability and large storage capacity. Data owners are motivated to store their valuable data in the cloud, now the problem arises how to protect valuable data against attacks and threats. In this paper we are introducing secure Attribute Based Encryption for protecting valuable data and further faster searching and accessing of encrypted data based on automatic annotation based searching technique. We also use k-NN technique for displaying top k ranked search result. Further it show effectively our scheme provides security and how effectively improves system performance and reduces communication overhead and also eliminates unnecessary traffic.

1. INTRODUCTION

Cloud computing is famous for storage, scalability and remote access. It allows the data owners to store and retrieve the data remotely, once the data is deployed it on the cloud, the complete control of the data comes under cloud server, cloud servers are the untrusted entity, to protect our valuable data(password, Account number), sensitive data's must be encrypted before outsourcing. Cloud storage services allow the users to outsource their data in the cloud storage servers and retrieve them whenever and wherever required. To provide faster access of data, it is necessary to provide searching capabilities to the data user (like Google) and the searching ability can be achieved based on automatic Annotation, and effective retrieval of data requires most frequently displayed result and that is done by using top-K ranked results. The main aim of this paper is,

1. To provide a data privacy and data security
2. Decrease the computational over head
3. Provide accurate ranked search result
4. Increase the communication capacity
5. Increase the performance by decreasing network traffic.
6. To improve the system usability.

2. BACKGROUND AND RELATED WORK

Due to remote access and storage of cloud, large companies and organization started switching from local server to cloud server to store more and more valuable data, it is necessary to provide security and faster data retrieval whenever and wherever require for the users.

A. Existing System

Existing system is based on the Searchable Encryption (SE). SE allows the user to search the data based on keywords, this technique supports, single keyword search, Boolean keyword search and pain text search. This technique has some disadvantages, they are

1. It cannot provide high level system requirement like usability
2. It is not adequate to provide search result based on ranking
3. Sharing of data is not secured under this technique
4. It support only single and Boolean keyword searching so it is not flexible and efficient
5. Single keyword search often yields far too coarse result

B. Proposed System

To meet the effective data retrieval, the result must be returned based on ranking. The ranked search result improves system usability and resource, also eliminates un-necessary traffic by returning relevant and accurate result. It also improves system performance. The enabling secure MRSE contains four different entities.

Data Owner: Data owner has a collection of documents, that documents are deployed on cloud server, data owners must specify the accessing permission to the data user.

Data User: Data users are authorized customer and they must have to account in order to communicate with data owners. Cloud is based on pay-as-you-use rule. Data users have a permission to search for the file by specifying file name and used ID. Only the authorized data users are capable to download and view the file. By using secrete key they can decrypt the file.

ISSN: 2278 – 7798

Cloud server: cloud servers are semi trusted server, they provides large storage and remote access, and data owners are able to deploy their documents along with key.

Cloud Manager: Cloud manager are trusted and authorized manager, they are responsible to provide and manage user and data owners account. It is also responsible for managing and handover keys to authorized users.

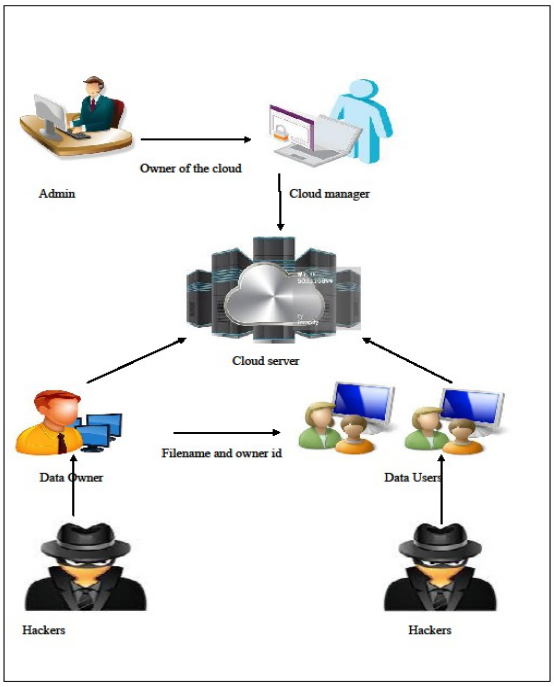


Fig (1): System Architecture

Admin are the owner of cloud and they hold and manage cloud manager. Cloud manger is the authorized and secured entity to manage cloud server and also responsible for managing and distributing key to the authorized user. Whereas cloud servers are semi trusted server they provides large storage and remote access. Cloud is based on pay-as-you-use. So data owner must want to buy a account in the cloud now the cloud manager creates authorized user. Data owner can able to outsource their sensitive information after encrypting the documents by using ABE. Then manually hand over the keys to cloud manager and keyword, filenames and owner ids are shared between the users who belong to same group. Data users are able to search the file based on keywords and owner IDs, download the file and decrypt the file using keys. User can able to view the file many times but downloading of particular file is only one time. If an authorized user try to access or hack the data, then generate a alert massage and that is sent to data owner as well as data users in the form if mail and also blocks the unauthorized access.

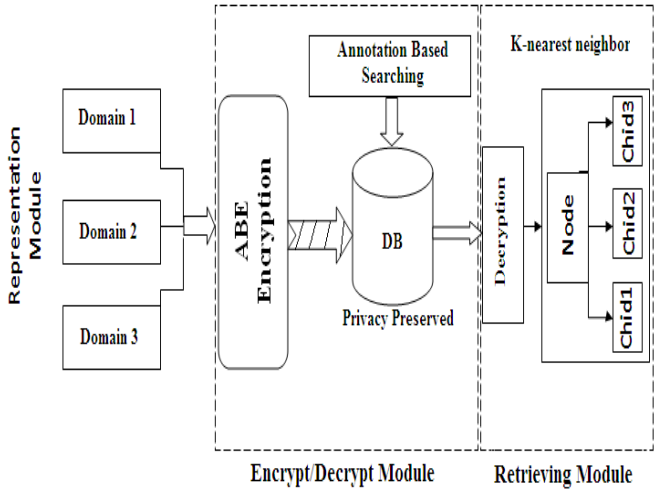
C. Design goals:

An effective ranked keyword search must provide usability, performances and security.

Secured Multi Keyword search: Our technique allows the authorized user to specify sing and multiple keywords in the search query (i.e. filename, date of publish, owner id). MRSE provides accurate and relevant search result.

Privacy Preserving: Providing data privacy and search privacy is very difficult to provide but MRSE provides data privacy, keyword privacy and search privacy to the user.

Efficiency: by providing accurate and relevant searched ranked result to the user when and wherever required, it improves the searching efficiency and data efficiency.



Fig(2) Block Diagram of secure MRSE

D. Modules Description:

System Setup Module:

In this module, first we develop the system module, which consists of Data Owner, which is controlled by domain authority, next data consumer and the Cloud Service Provider.

New User Grant Module:

When a new user wants to join the system, with the aid issues an attribute private key to him/her based on his/her attributes. Based on the system model provided we attempt to define an underlying primitive namely OABE with outsourced key-issuing and decryption for realizing our access control system.

File Upload Module

In this module, we develop the file upload module process, where, when a data owner wants to outsource and share a file with some users, he/she encrypts the file to be uploaded under a specified attribute set (resp. access policy). Whenever a data

ISSN: 2278 – 7798

owner wants to create and upload a file he/she firstly defines an attribute set (resp. access structure).

File Access Module:

In this module, we create the file access module, when a user wants to access an outsourced file; he/she downloads cipher text from S-CSP and decrypts it with the help of D-CSP.

User Revocation Module:

When there is a user to be revoked, updates \affected" users' private keys with the help of CSP, while the \affected" cipher texts having been stored on S-CSP will be updated as well.

3. ALGORITHM USED

Secure ABE: ABE is one type of public key encryption, here secretes keys of user & the cipher text is depends on attributes. Deception of the cipher text is possible only if the attributes of user key is matched with the attributes of cipher text. A key-policy attribute-based encryption system for message space M and access structure space G1 is a tuple of the following algorithms:

Setup (λ1, U1): it takes input as security parameter λ1 and a universe description U1 and returns public parameters PK1 and the master secret key MK1 as output.

Encrypt(PK1, M1, S1): It takes input as public parameters PK1,a message M1 and a set of attributes S1 and returns cipher text CT1 associated with the attribute set as output .

KeyGen(MK1, A1): It takes input as the master secret key MK1 and an access structure A1 and return private key SK1 associated with the attributes as output.

Decrypt(SK1, CT1): it takes input as private key SK1 associated with access structure A1 and a cipher text CT1 associated with attribute set S1 and return a message M1 if S1 satisfies A1 or it returns error message as output.

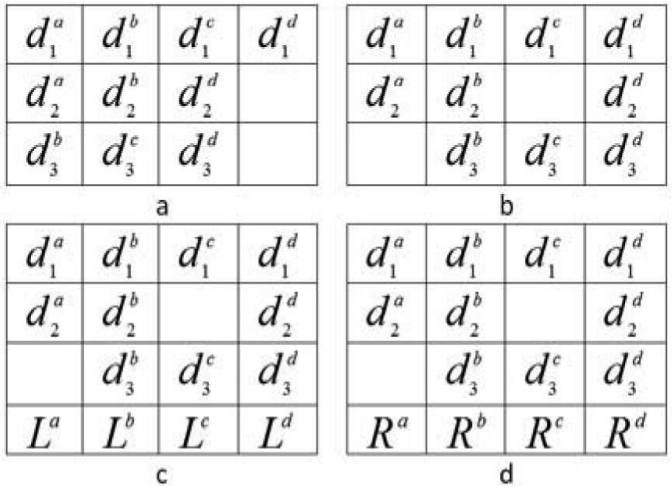
Automatic Annotation:

These techniques first align all the data units on a result page then align the result into different groups i.e. according to same data semantics. Let $data_j^k$ denote the data unit belonging to the $_j$th SRR of concept k(fig 3a). The Search Result Records (SRR) on a result page can be represented in a table format with each row representing an SRR. Automatic annotation search techniques are divided into 3 different phases.

Alignment: Identify all data units in the SRR and then orange them into different groups with each group corresponding to a different concept (e.g., all titles are grouped together). Grouping of same data units of the same semantic help to identify the common patterns and features among these data units (fig 3b) and this is the input for annotators.

Annotation: Each basic annotator is used to produce a label for the units within their group holistically, and a probability model is adopted to identify the most appropriate label for each group (fig 3c).

Annotation wrapper generation: For each identified concept, generate an annotation rule R that describes how to extract the data units from the concept present in the result page and what is the appropriate semantic label should be. The rules for all aligned groups, collectively, form the annotation wrapper for the corresponding Web Database (WDB), which can be used to directly annotate the data retrieved from the same WDB in response to new queries without the need to perform the alignment and annotation phases again (fig 3d).



Fig(3): Phases of automatic annotation

K-Nearest Neighbour

The kNN[3] query is an important analysis operation applied for database and it is used as a standalone query or core module for data mining. A kNN query searches is applied for K points in the database and that are the nearest to a given query point Q. Distance Preserving Transformation (DPT) supports kNN technique.

A DPT holds distances in the transformed space. if E_1 is a DPT, then $D(E_1(P_1;K);E_1(P_2;K)) = D(P_1; P_2)$. Then F is the Euclidean distance.

Database (DB) contains column vector and DPT transforms the space by translation and rotations for a point P and the encrypted value E1(PP;K) of P.

A DPT E1 can be expressed as np + m, where n is a D×D orthogonal matrix and m is a d-dimensional column vector. Distance between points $D(P_1; P_2) = D(E_1(P_1;K);E_1(P_2;K))$.

Here n and m together form the encryption key K.

4. EXPECTED RESULTS

1. Data Encryption and decryption Result

The ABE algorithm is applied on data for encryption and decryption. Then deploy the encrypted data on the cloud.

ISSN: 2278 – 7798

User can access the data after downloading then decrypt the file. For encryption and decryption keys are provided.

2. Ranking Result

When any User need the data for quicker access, the searched accurate result based on Ranking using k-nearest neighbor algorithm

3. Alert System Results

If any unauthorized User tries to access or updating the data on cloud, then alert will be generated in the form of mail and messages. The alert intimates the authorized user and the data owner too.

.

5. CONCLUSION AND FUTURE SCOPE

Cloud computing is one of the current most important and promising technologies. A user Can able to store their personal files in a cloud and retrieves them wherever and whenever they want. Here we presented a new efficient method for privacy preserving encrypted keyword search. It enables the service provider to determine whether a document contains a specified keyword without getting any information about the document or keyword. And also return accurate and effective ranking result based on top k retrieval. It provides low overhead on computation & communication. It eliminates network traffic and also provides data privacy and search privacy. This technique provides stronger security and improves system usability and it also support for multi user systems. This system is currently work on single cloud, In future is will extend up to sky computing & Provide better security in multi-user systems.

6. REFERENCES

[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] S. Kamara and K. Lauter, "Cryptographic cloud storage," in *RLCPS,January 2010, LNCS. Springer, Heidelberg*.

[3] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in *Proc. of ACNS*, 2005.

[4] R. Curtmola, J. A. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions,"
in *Proc. of ACM CCS*, 2006.

[5] D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in *Proc. of EUROCRYPT*, 2004.

[6] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. S. III, "Public key encryption that allows pir queries," in *Proc. of CRYPTO*, 2007.

[7] R. Brinkman, "Searching in encrypted data," in *University of Twente, PhD thesis*, 2007.

[8] A. Lewko, T. Okamoto, A. Sahai, K. Takashima, and B. Waters, "Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption," in *Proc. of EUROCRYPT*, 2010.

[9] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in *Proc. of ICDCS'10*, 2010.

[10] W. K. Wong, D. W. Cheung, B. Kao, and N. Mamoulis, "Secure knn computation on encrypted databases," in Proceedings of the 35th SIGMOD international conference on Management of data, 2009, pp. 139–152.

ISSN: 2278 – 7798