INTERNATIONAL JOURNAL FOR TRENDS IN ENGINEERING & TECHNOLOGY
VOLUME 6 ISSUE 1 – JUNE 2015 - ISSN: 2349 - 9303

Isolation Conserves Multi-Keyword Graded Up Encoded Cloud Data

C.Devipriya
PG Scholar, Department of CSE,
Pallavan College of Engineering, Kanchipuram,
cdevipriya91@gmail.com

S.Kerthy
Assistant Professor, Department of CSE,
Pallavan College of Engineering, Kanchipuram,
skerthy_it@yahoo.co

Abstract — The Data owner outsources their data to public cloud owing to safety. These data are encoded before they are transferred to the cloud. Cloud user need a different type of data's from cloud. These data's must be relevant to the query of the user. Multi keywords are allowed to the user to search the query in cloud. This frame work is called Multi keywords ranked search over encrypted data is MRSE. To capture the relevance data documents to the search query "coordinate matching principle" is used. "Inner product similarity" measure is used to evaluate the similar document and used for perfectly matching the query given by the user. Confidential Data stored in the cloud server is encrypted and decrypted by using blowfish algorithm. The Message Authentication Code (MAC) algorithm is used to check the integrity of the data stored in the cloud service provider. Message authentication code is created by the data owner before the data uploaded in the cloud server because of confidentiality. The user receives the data from the service provider and decrypts the data by using the decryption key received from the data owner .After that user verifies the message authentication code generated by him with the code generated by the data owner. Experiments on the real-world data set further show proposed schemes indeed introduce low overhead on computation and communication.

Index Terms - Multi keywords, MRSE, co-ordinate matching, blowfish, inner product similarity, message authentication code

1 INTRODUCTION

Cloud computing provides an environment for storing the data and it is used to provide services based on on-demand service. In cloud computing, the data owner outsources their complex database management systems into the cloud server because of safety. To protect data privacy against attacks from the cloud service provider, confidential data must be encoded before being uploaded to the cloud server. However it is crucial to achieve range queries on the encoded data in the cloud provider. This problem faces two challenges. First, the confidentiality needs hiding the documents and the relative content of the document from the cloud server, should not to be leaked to the hacker. Second, relevant to the query given by the user in the search result the cloud provider must return the matched document to the customer. The Data owner selects the data and creates the bit vectors for that data. Using that bit vectors of the data the binary data is generated. The binary data is the indexes for the user to search their relevant data and the indexes are outsourced by data owner to the cloud service provider. The user has different choices of data and they sent the query to the server or service provider. Before that the user have to get the access from the data owner. For that the user sends the details about him or her to the data owner. we propose an efficient indexing method to support faster query evaluation than the trivial linear scan manner. Until now, the query answer is computed by evaluating the condition for every encrypted data point e(pi) in the database. This linear scan is not acceptable for large databases.

We adopt the ball tree indexes to retrieve such points while pruning as many data points as possible. Briefly, a ball tree is a binary tree such that each non-leaf node represents a ball and has two child nodes. A data point belonging to a parent goes to the child ball whose center is closer to the data point. All data points are only stored at the leaf nodes. To build such a ball tree, we could keep separating the data point space into two partitions (left and right child) recursively until the number of data points in some partition is below a predefined threshold and we make this partition as a leaf node. We call this threshold as "max leaf size" and we will test it in our experiment chapter. The detailed algorithm to construct the ball tree (i.e., how to encrypted records into balls) could be found and we don't bother to bring the exactly same algorithm here..

2 BACKGROUNDS AND RELATED WORK

Encryption is a perfect way to protect data privacy which also gives rise to the difficulty in executing a regular query on such encoded data. To keep the advantage of computing ability on the server side, it is not acceptable to download the whole encoded data and execute the queries on the data after decryption; the cloud server has to have the ability to execute queries directly over the encrypted data and return the correct result to the user. There have been considerable interests in querying encrypted data and various queries are considered.

Briefly, the work includes equality test, range queries search, and aggregation query computation as well as keyword based query search or similarity based query search. Existing schemes allow a user to securely search the documents with conjunctive or disjunctive or single keyword search. These three approaches return the undifferentiated results and it does not provide ranked results. The query given by the user does not match the result provided by the cloud server. None of this approach provides multi keyword ranked search over encrypted cloud data. So due to this problem more traffic occur in the cloud server and the user cannot able to download the document needed by them. The main drawback is data

encryption only create on earliers system, so security less. The keyword isolation could not be secured in the public key setting since the server could encrypt any keyword progress. This system is built upon with the expensive calculation of the pairing process on elliptic curves. The adopted secure inside product calculation scheme is not well sufficient for our MRSE design. The trusted hardware, which is placed on the the server side and it is tamper-resistant, has limited capacity in storage and processing power. Since it is Symmetric key Technique and small key sizes-It is less secure.

2.1 Existing System

Earlier systems works for secure ranked keyword search which handles keyword density to rank results rather of retracing identical outcomes. To develop search performance, conjunctive keyword search over encoded data. The Boolean keyword searchable encryption processes backing multiple keywords rated search over encrypted cloud data while securing privacy. This scheme to support more search denotation which improve the search experience of the strategy. The basic challenge comes from the following dilemma a range query requires comparing the values of the attribute. The scheme protects the confidential values and does not preserve the relative order of such values in encoded records.

- Public cloud having plain text of local site data. i.e. outsourcing original data to public cloud from local site which may lead to replay attacks or brute force attacks.
- The Existing System has the problem to determine that, multi-keyword search includes retrieval of the undifferentiated search term as a result for many iterations.

Disadvantages

1. Since it is Symmetric key technique and also small key sizes it is less secure.
2. Execution time for encryption and decryption are high
3. Network traffic is very high

2.2 Proposed System

Multi keywords are allowed to the user queries when searching in cloud. This frame work is called Multi keywords ranked search over encrypted data is MRSE. The main purpose of the proposed system is to solve the problem of multi keyword ranked search over encrypted cloud data (MRSE). To capture the relevance data documents to the search queries "coordinate matching principle" is used. Dynamic updating of index and the document can also be applied in this system models. During the index structure, each document is combined with binary vectors as sub indices. The data owner receives the information from the user then sends the decoded key and then cloud service provider sends the decoded data to the user. This system is not only designed for support the entry pattern for the performance concerns.

During the index construction, each document is converted into binary The binary data is the index for the data in the data owner. The bit vector is the bytes form of the data in the data owner. The bit vector is converted into the binary data. These bit vector and the binary data are the index for the user to search the queries. However each bit vector represents whether the corresponding keyword is contained in the document. The search queries is also described as a binary vector where each bit means, whether the corresponding keyword appears in this search request, so the closeness could be exactly measured by the inner product of the query vector with the data vector. And both index and documents are uploaded from the data owner to the cloud server but before uploading the data both the index and data are ciphered using blowfish algorithm.Integrity of the data stored in the cloud server is verified using message authentication code.Access control only given for the authorised user because unknown person cannot access the data without knowing decryption key.After that Decryption key only sent by the data owner to the authorised user.

Advantages

1. Its great flexibility and industrial savings are prompting both individuals and enterprises to source their local complicated data administrated system into the cloud.
2. MRSE schemes to achieve different stringent privacy requisites in two various threat designs.
3. Service provider did not know about the original content in the data owner that is the encoded data most secure.
4. Efficiency can be achieved while processing the Multi Keyword search request.
5. Blowfish algorithm used here is faster than AES and DES.

3 SYSTEM ARCHITECTURE

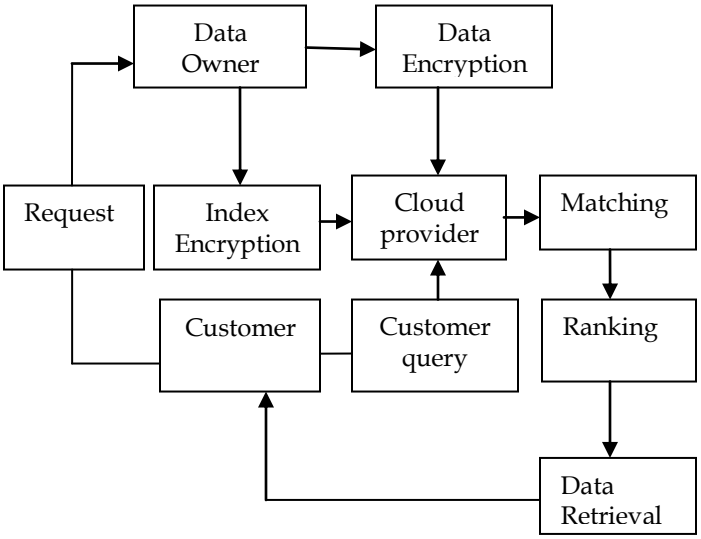

Fig 1 Architecture diagram for proposed system

Fig 1 Architecture diagram involves three entities and are classified as Data Owner, Cloud Provider and Customer (user).The Data owner has different document collection. The data owner will create the binary data for their document collection .This binary data act as the index for the customer to search their query. And then the data owner will ciphered both index and document and upload to the cloud provider. The query is processed by the customer to search the relevant document in the cloud. After that cloud provider matched With the relevant document to the query given by the user and

produce ranked result to the customer to retrieve their relevant data.

4 ALGORITHM USED

4.1 Blowfish Algorithm

For secure storing original data and index to the cloud server Blowfish algorithm is used. Blowfish algorithm is a symmetric block cipher with a variable-length key, 64-bit block cipher. The algorithm consists of two parts: a key-expansion part and a data-encryption part. Blowfish algorithm is faster than AES and DES. Mainly it is used to reduce the execution time for encryption and decryption. Security for the data stored in the cloud server is increased due to this blowfish algorithm. Encoded and decoded key are same in the blowfish algorithm.

Sub keys: Blowfish uses a large number of sub keys. These keys must be pre computed before any data encryption or decryption.

Encryption: Blowfish is a Feistel network consisting of 16 rounds. The input data to the encryption is plaintext .Encryption steps for blowfish algorithm:

Decryption: Decryption is exactly the same as encryption, except that P1, P2... P18 are used in the reverse order. The input data to the decryption is cipher text.

4.2 Message Authentication Code

MAC is applied here to check the integrity of the data stored in the cloud provider. The Data owner before uploading the data to the cloud provider generates a MAC code and sent the code to the authorized user. The user decrypts the data using decryption key sent by the data owner, after that user verifies the MAC code generated by him with the code generated by the data owner. The MAC algorithm mainly focuses on authentication and integrity, so that the security for the data is very high by applying this algorithm. Original data cannot be hacked by unauthorized user. The MAC algorithm, accept an input as secret key and an arbitrary-length message to be authenticated, and output a tag. The MAC values are generated and protected using the same secret key. This algorithm is useful for secret sharing of information between the data owner to the cloud server via cloud server. Steps for MAC Algorithm:

1. K ← K
2 .Run AMACK (·), VFK (·,·) where VFK(M, T) is 1 if MACK(M) = T and 0 otherwise if A made a VFK query (M, T) such that – The oracle returned 1, and – A did not, prior to making verification query (M, T), make tag-generation query M ,then return 1 else return. A MAC is a Cryptographic Checksum denoted by ,MAC = C K(M) where M is Variable-Length Message and K is a Secret Key to a Fixed Size Authenticator.

4.3 Coordinate Matching Principle

Coordinate matching principle generally used for identifying the relevant document. Ranking the user query can also effectively achieved by applying this coordinate matching technique. Considering large number of user and data in the cloud server it is necessary to estimate the document according to the query given by cl. This principle mainly eliminates traffic and reduces the burden for the user to retrieve their related data.

5 IMPLEMENTATION

5.1 Binary Data Generation

The Data owner select the data and create the bit vector for that data. Using that bit vector of the data the binary data is generated. The binary data is the index for the data in the data owner. The bit vector is the bytes form of the data in the data owner. The bit vector is converted into the binary data. These bit vector and the binary data are ready for the data ciphering.

5.2 Data Ciphering

In this process, both the encrypted document and searchable index (binary data) are outsourced from the data owner to the cloud service provider. By using blowfish algorithm, the data owner has to encrypt the original data and sent it to server. And then encrypt the binary data or index by using that same algorithm . Service provider did not know about the original content in the data . This index are used to refer the data in the service provider. It gives more security in the server side, so that the attackers can't use the data.

5.3 Data User Access Control

The user needs data from the server. The user has different choices and the user sent the query to the server or service provider. Before that the user gets the access from the data owner. For that the user sends the details about him or her to the data owner. Then only the data owner receives the information from the user and ready to sends the decryption key. The access control mechanism is employed to manage decryption capabilities given to user .In this process ,Only the authorised user can access the data from the cloud server.The unauthorised user cannot access the data because they do not know the key. Secure access control of the outsourced data in the can be confidentiality achieved by using blowfish and message authentication code algorithm (MAC). This approach provides user access and control over their data. Lastly, this process achieved strong authentication and identity management for both cloud service providers and the user.

5.4 Data User Query

The data user query is processed by the service provider. The service provider generates the bit vector for the query from the user. Then the service provider converts the bit vector into binary data. Service provider finds the similar data from the indexes. And send the encrypted data to the data owner. Then the user decrypts the received data by the key from the data owner**.** multi-keyword are used for the user to search the query.Multi-keyword query is processed by Coornidate matching Technique to capture the relevance data to the search query. A query from User will go though the trusted proxy, which encrypts the query and submits the query to the server.The server computes and returns the answer to the proxy.

The proxy then decrypts the answer, and returns the answer to User. For example, Owner is a hospital, who outsources patient records to the cloud, and Users are various medical research labs, who post queries to retrieve patient records of interests.

5.5 Retrieving Data

User receive the encoded data from the cloud service provider.When other person attacking the data ,original data cannot be retrieved by them because they only know the encrypted data. Decryption key only known to the data owner and user.After that user retrieve their

data by using decryption key sent by the data owner.Decryption key is created by using blowfish algorithm. After decryption the user can gets the original data.Data owner and the user only know about the original data.Integrity of the retrieved data can be checked by applying message authentication code(MAC).

6 RESULTS

6.1 Data encryption and decryption result

Blowfish algorithm generally used for encrypting and decrypting the data while storing and retrieving the data from the cloud .The user can access the data after downloading and decrypting file. For encryption and decryption keys are provided.

6.2 Ranking result

When any User request for the data soon after ranking is done based on request given by the user. For ranking co-ordinate matching principle is used. Ranking is used for getting the expected results from the server depend on the user query.

6.3 Integrity result

Both the Data Owner and user check the authentication of the message by applying a MAC algorithm while storing and retrieving the data from the cloud server. The MAC algorithm is mainly used for checking the authentication of the secret sharing of information between the client and the data owner.

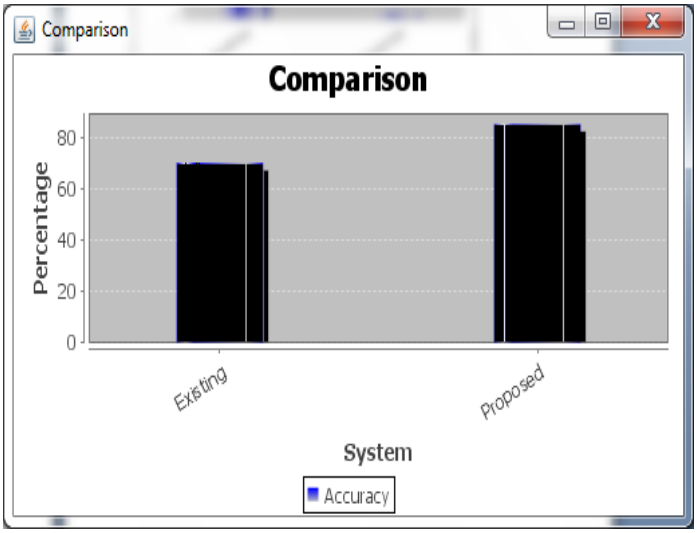

Fig 2 Comparison Graph between existing and proposed system

7 CONCLUSION

Query electrocution in the cloud server consists of computing and ranking similarity scores for every document in the data set. It defines multi-keyword graded search over encoded cloud data, and establish privacy requirements. A major challenge or concern in the cloud computing paradigm acquires data privacy. In this paper, for the first time defined and solved the problem of multi-keyword ranked search over encrypted cloud data, and establish a variety of privacyrequirements.To capture the relevance of data documents to t he search query "coordinate matching principle" is accomplished. It also interrogate some further improvement of our ranked search tool, contains supporting more search definition. There is an appeal to makes advantage the powerful sources of the cloud server to produce services to the user. Data privacy and integrity is achieved by employing a blowfish and the MAC algorithm. Finally, we analyze the performance of our scheme in detail by the experiment on real-world dataset. . But, there still exist some problems, such as dynamic update for searchable indices. We will do more research in the future.

REFERENCES

[1] N. Cao, S. Yu, Z. Yang, W. Lou, and Y. Hou, "LT Codes-Based Secure and Reliable Cloud Storage Service," Proc. IEEaE INFOCOM,pp. 693-701, 2012.

[2] M. Bellare, A. Boldyreva, and A. ONeill, "Deterministic and Efficiently Searchable Encryption," Proc. 27th Ann. Int'l Cryptology Conf. Advances in Cryptology (CRYPTO '07), 2007.

[3] 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 Anonymous Ibe, and Extensions," J. Cryptology, vol. 21, no. 3, pp. 350-391, 2008.

[4] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "FuzzyKeyword Search Over Encrypted Data in Cloud Computing,"Proc. IEEE INFOCOM, Mar. 2010.

[5] C. Wang, N. Cao, K. Ren, and W. Lou, "Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data," IEEE Trans. Parallel and Distributed Systems, vol. 23, no. 8, pp. 1467- 1479, Aug. 2012

[6] S. Kamara and K. Lauter, "Cryptographic Cloud Storage," Proc. 14th Int'l Conf. Financial Cryptograpy and Data Security, Jan. 2010.

[7] I.H. Witten, A. Moffat, and T.C. Bell, Managing Gigabytes: Compressing and Indexing Documents and Images. Morgan Kaufmann Publishing, May 1999.

[8] D. Song, D. Wagner, and A. Perrig, "Practical Techniques for Searches on Encrypted Data," Proc. IEEE Symp. Security and Privacy, 2000.

[9] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure Ranked Keyword Search over Encrypted Cloud Data," Proc. IEEE 30th Int'l Conf. Distributed Computing Systems (ICDCS '10), 2010.

[10] C. Wang, Q. Wang, K. Ren, N. Cao, and W. Lou, "Toward Secure and Dependable Storage Services in Cloud Computing," IEEE Trans. Services Computing, vol. 5, no. 2, pp. 220-232, Apr.-June 2012.

[11] A. Singhal, "Modern Information Retrieval: A Brief Overview," IEEE Data Eng. Bull., vol. 24, no. 4, pp. 35-43, Mar. 2001.

[12] E.-J. Goh, "Secure Indexes," Cryptology ePrint Archive, http:// eprint.iacr.org/2003/216. 2003.