Fetching Encrypted Cloud Data Using Searchable Index Scheme

Selvi. A
II[nd] Year – M.E. CSE
Srinivasan Engineering College
Peramabalur
Tamil Nadu, India

Suresh Kumar. N
II[nd] Year – M.E. CSE
Srinivasan Engineering College
Peramabalur
Tamil Nadu, India

Amritha. S
Asst. Professor / CSE
Srinivasan Engineering College
Peramabalur
Tamil Nadu, India

Abstract

To ensure security, encryption techniques play a major role when data are outsourced to the cloud. Retrieving the data from the cloud servers is measured. Many searching techniques are used for retrieving the data storage. This information focused on a set of keyword based search algorithms. This Methodology provides secure data retrieval with high effectiveness. We observe that server-side ranking based on order-preserving encryption (OPE) certainly leaks data privacy. To reduce the leakage, we propose a two-round searchable encryption (TRSE) scheme that supports top-k multikeyword retrieval. All the files stored in clouds which are trust servers, are in the encrypted form and are secret to other users. Using Attribute Based encryption (ABE) to encrypt the files. In this scheme, users are categorizing into personal and specialized domains which greatly reduce the key management complexity. A structured way to access the files for personal and professional purposes. Users are able to dynamically modify the access policy and attributes.

Keywords

Searchable Symmetric Encryption, Order Preserving Encryption, Two Round Searchable Encryption, Attribute Based Encryption

1. Introduction

Cloud Computing, the trendiest computing in information technology where everything is based on-demand service and pay-for-use service. It is bringing of compute services such as SaaS, PaaS and IaaS over the internet that are supervised by arbitrator at outback locations. There are many applications such as emails, business data file storage, etc. are outsourced to server (cloud). Only authorized user can access the data from the cloud server. Outsourcing unencrypted data to cloud by the owner is not much secure because server may leak information to unauthorized. Hence encryption plays a major position before outsourcing the data into the cloud server. In spite of encrypting, retrieval of data becomes an intriguing task when searching has to be made on vast data. The best way is to use keyword based search on encrypted data for data concealing. Many searchable techniques have been proposed on the basis of keyword search. Discussion is made on the existing techniques that are been intend by many authors. This study analyses the algorithms for searching the encrypted contented. In study is made on these algorithms based on the working principle, merits and demerits. It also compares the complexity, efficiency in the clouds of various algorithms and shows which technique is better to handle while retrieving the encrypted content.

2. Techniques For Searching Over Encrypted Data

2.1 TRSE (Two-Round Searchable Encryption)

The framework of TRSE includes four algorithms: Setup, IndexBuild, TrapdoorGen; Score Calculate, and Rank.

2.1.1 Setup(ƛ)

The data owner generates the secret key and public keys for the homomorphic encryption scheme. The securityparameter ƛ is taken as the input, the output is a secret key (SK), and a public key set PK.

2.1.2 IndexBuild (C, PK)

The data owner builds the secure searchable index scheme from the file collection C.Technologies from IR community like steaming are employed to build searchable index I from C,and then I is encrypted to I' with PK,output the secure searchable index I'.

2.1.3 TrapdoorGen(REQ, PK)

Thedata user generates secure trapdoor from his request REQ.VectorTϕ is built from user's multikeyword request REQ and then encrypted into secure trapdoor T with public key from PK,output the secure trapdoor Tϕ.

2.1.4 ScoreCalculate(Tϕ,I')

When receives secure trapdoor Tϕ ,the cloud server computes the scores of each files in I' with Tϕ and returns the encrypted result vector N back to the data user.

2.1.5 Rank(N,SK,K)

The data user decrypts the vector N with secret key SK' and then requests and gets the files with top-k scores.

2.2 Attribute Based Encryption (ABE)

Attribute based encryption (ABE) uses access policy while searching on encrypted data with its Boolean expressions.It works on the basis of nine algorithms. The first is the Setup algorithm used to compute secret key and masterkey by trusted authority. Second, KeyGenalgorithm is used to generate the public/private key pair. Fourth,fifth and sixth algorithm such as PseudoGen(), Encrypt(), are used for outsourcing the data usingcryptographic primitives such as access structure and attribute scramble procedure. Seventh,eighth and ninth algorithm used as query, retrieve and decrypt is mainly for the retrieval of data. In which queryalgorithm works as the retriever take on pseudonym list from the cloud service provider and receiver sends thescrambled index to the Cloud Service Provider (CSP).

Then the CSP checks whether the request made by theretriever and the encrypted index stored are same by using Retrieve algorithm. If it matches decrypt algorithmworks where the encrypted data are decrypted and sent to the retriever. It provides best quality for searchingover encrypted data and faster in accessing.

3. Related Work: Plain Text Fuzzy Keyword Search

Recently, the importance of fuzzy search has received attention in the context of plaintext searching in information retrieval community. They addressed this problem in the traditional information-access paradigm by allowing user to search without using try-and-see approach for finding relevant information based on approximate string matching algorithm. At the first glance, it seems possible for one to directly apply these string matching algorithms to the context of searchable encryption by computing the trapdoors on a character. However, this trivial construction suffers from the dictionary and statistics attacks and fails to achieve the search privacy.

3.1 Disadvantages

The secure searchable encryption scheme does notperform any function when new updates in files or whenany modifications are performed.

The relevance score algorithm is not updated frequentlywhen there are some modifications in the owner files.

4. Proposed System

The problem of supportingefficient ranked keyword search for achieving effectiveutilization of remotely stored encrypted data in CloudComputing. This is done by developing an efficient algorithm to group the 'related' keywordstogether.

4.1 Contributions

In Cloud Computing, an outsourced file collectionmight not only be accessed but also updated frequently forvarious application purposes. Hence, supporting the scoredynamics in the searchable index for a secure storage engine which is reflected from the corresponding filecollection updates, is thus of practical importance. In oursystem, we consider score dynamics as adding newlyencrypted scores for recently created files, or modify oldencrypted scores for modification of existing files in the filecollection.

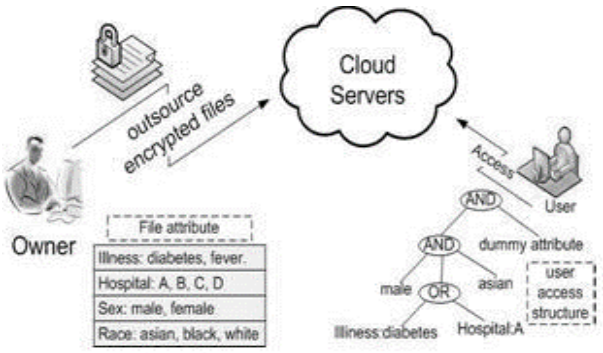

Fig . 1

Attribute Based encryption (ABE) to encrypt the files. Users are categorized into personal and professional domains which greatly reduce the key management complexity. There is a controlled way to access the files for personal and professional purposes. User is able to dynamically modify the access policy and attributes.This schemecontains four algorithms: Setup(), KeyGen(), Encrypt(),and Decrypt(), and they will be introduced as follows.

5. Attribute Encryption Scheme

According to these schemes, a summary of the criterion, that ideal attribute-based encryption schemes, are listed as follows.

5.1DataConfidentiality

Before uploading data to the cloud, the data was encrypted by the data owner. Therefore, unauthorized parties including the cloud cannot know the information about the encrypted data.

5.2 Fine-grained access control

In the same group, the system granted the different access right to individual user. Users classified on the same group, but each user can be granted the different access right to access data. Even same group of users have different access rights.

5.3Scalability

When the authorized users increase, the system can work efficiently. So the number of authorized users cannot affect the performance of the system.

5.4 User accountability

If the authorized user is dishonest, he would share his attribute private key with the other unauthorized user. It causes the problem that the illegal key would share among unauthorized users.

5.5User revocation

If the user quits the system, the scheme can revoke his access right from the system directly. The revocable user cannot access any stored data, because his access right was revoked.

5.6 Collusion resistant

Users cannot combine their attributes to decipher the encrypted data. Since each attribute is related to the polynomial or the random number, different users cannot collude each other.

6. Conclusion

In this study rigorous analysis is made on encryption techniques which relate to search based retrievalof files from the outsourced encrypted data. A lot of searchable encryption schemes have been analyzed based onsingle keyword and multi-keyword search. Many disadvantages have been focused on these techniques given thatthey rely on Boolean expressions. Hence rank based retrieval of data has been talk aboutwhich proves thedata security, search access and does not leak information to untrusted authorities. This study concludesrank based retrieval is most efficient for searching on encrypted data.Thus, based on the discussion above, these existingattribute-based encryption schemes have properties: (1)These schemes are encrypted with attributes, so a data owner just needs to predefine these attributes that hewould use, he doesn't need to care about the number ofusers in the system; (2) Each attribute has public key,secret key, and a random polynomial, so different userscannot combine their attributes to recover the data, anddifferent users cannot carry out collusion attacks; (3) Onlythe user who possesses the authorized attributes can satisfy the access policy to decrypt data; (4) The access policy contains a Boolean formula such as AND, OR, which can let the access structure be flexible to controlusers' access. However, almost all scheme exist that theauthority is used to generate keys.

6.1 References

[1] M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A.Konwinski, G. Lee, D. Patterson, A. Rabkin, and M. Zaharia, "AView of Cloud Computing," Comm. ACM, vol. 53, no. 4, pp. 50-58,2010.

[2] M. Arrington, "Gmail Disaster: Reports of Mass EmailDeletions,"http://www.techcrunch.com/2006/12/28/gmail-disasterreportsof-mass-email-deletions/, Dec. 2006.

[3] Amazon.com, "Amazon s3 Availability Event: July 20, 2008,"http://status.aws.amazon.com/s3-20080720.html, 2008.

[4] RAWA News, "Massive Information Leak Shakes Washingtonover Afghan War," http://www.rawa.org/temp/runews/2010/08/20/massive-information-leak-shakes-washington-overafghan-war.html, 2010.

[5] AHN, "Romney Hits Obama for Security InformationLeakage. "http://gantdaily.com/2012/07/25/romney-hits-obama-forsecurity-information-leakage/, 2012.

[6] Cloud Security Alliance, "Top Threats to Cloud Computing,"http://www.cloudsecurity alliance.org, 2010.

[7] C. Leslie, "NSA Has Massive Database of Americans' PhoneCalls," http://usatoday30.usatoday.com/news/washington/2006-05-10/, 2013.

[8] R. Curtmola, J.A. Garay, S. Kamara, and R. Ostrovsky, "SearchableSymmetric Encryption: Improved Definitions and Efficient Constructions," Proc. ACM 13th Conf. Computer and Comm. Security (CCS), 2006.