International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 2 Issue 3, March - 2013

Implementation Of Secure Ranked Keyword Search By Using "RSSE"

P. Naresh
M.Tech(CSE)
KVSR Engg College
Kurnool

K. Pavan kumar
Asst.prof & HOD in IT
KVSR Engg College
Kurnool

D. K. Shareef
Asst.prof in CSE
INTELL Engg College
Anantapur

Abstract

Cloud computing provides the services for on demand users. Before exchanges, it must encrypt the data files and stored in cloud server. Present traditional applications are using keyword search to find out the results with encrypted format. It can show the results as huge amount of files. These search results are not utilized efficiently in user's side. In this paper we are going to present a new way to solve the problems by introducing Secure Keyword Search mechanism. Secure keyword search displays the top list of files with fewer results as output. Using fewer results as output we increase the file retrieval accuracy and reduce the communication overhead. Secure Keyword Search starts in ranked list files only. Ranked list files can identify with relevance score and statistical measure. To achieve our design goals on both system security and usability, we propose to bring together the advance of both crypto and IR community to design the ranked searchable symmetric encryption (RSSE) scheme.

Keywords: cloud computing, ranked keyword search, Relevance score, and accuracy.

1. Introduction

Cloud computing is the use of computing resources that are delivered as a service over a network. The name comes from the use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. Cloud computing entrusts remote services with a user's data, software and computation.

Cloud computing is an emerging technology which helps as an utility, through which clients are going to store their data in the cloud server and using applications from a set of computing resources[1]. Here sensitive data is going to be centralized in the server. In some times the cloud server may leaks the data to hackers [2]. The data is going to encrypted before outsourced to achieve privacy. The encryption techniques increase the data utilization from a large amount of data. To retrieve data files we introduced keyword search mechanism. By this mechanism the users are going to retrieve the data files of their interest. In traditional search, encryption techniques the users are going to search data by using keywords without decrypting it, they support only Boolean keyword search only [2][10]. In cloud computing graded keyword search enhances the system usability by displaying the matching files by the help of relevance score. To achieve security and usability we introduce advanced cryptographic and information retrieval techniques, and using one-to-many order preserving symmetric encryption [3].

Basically there are three types of public cloud Services:

- Infrastructure as a service (IaaS): In this most basic cloud service model, IaaS providers offer computers, as physical or more often as virtual machines, and other resources.
- Platform as a service (PaaS): In the PaaS model, cloud providers deliver a computing platform typically including operating system, programming language execution environment, database, and web server. Application developers can develop and run their software solutions on a cloud platform without the cost and complexity of buying and managing the underlying hardware and software layers.
- Software as a service (SaaS): In the SaaS model, cloud providers install and operate application software in the cloud and cloud users access the software from cloud clients. The cloud users do not manage the cloud infrastructure and platform on which the application is running. This eliminates the need to install and run the application on the cloud user's own computers simplifying maintenance and support. What makes a cloud application different from other applications is its scalability.

Cloud computing is a type of computing that relies on sharing computing resources rather than having local servers or personal devices to handle applications. In cloud computing, the word cloud is used as a metaphor for "the Internet," so the phrase cloud computing means "a type of Internet-based computing," where different services such as servers, storage and applications are

delivered to an organization's computers and devices through the Internet. Cloud computing is comparable to grid computing, a type of computing where unused processing cycles of all computers in a network are harnesses to solve problems too intensive for any stand-alone machine**.** This shared IT infrastructure contains large pools of systems that are linked together. Often, virtualization techniques are used to maximize the power of cloud computing.

2. Background And Related Work

Now-a-days cloud servers get to store large amount of files. Here select and processing the files is the main problem. Whenever large numbers of files are available in cloud server under encryption some problems are generated. Totally all files are not encrypted. Although traditional searchable encryption schemes allow a user to securely search over encrypted data through keywords without first decrypting it, these techniques support only conventional Boolean keyword search[2], without capturing any relevance of the files in the search result. That's here there is no sufficient privacy and security in outsourcing. Some unauthorized users are entering and corrupt the content of information. Privacy-preserving multi-keyword ranked search over encrypted cloud data will also enables us to encrypt data in a privacy provided manner [9][10], and search through multiple keywords within large amount of data in cloud server.

Architecture:

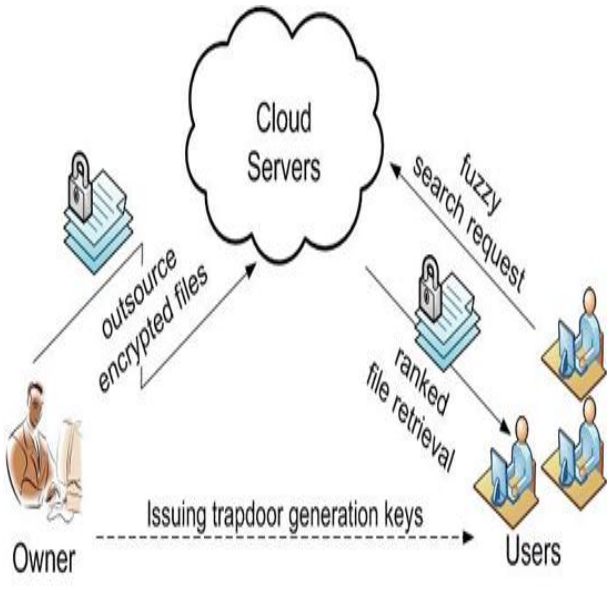

Fig.1: Architecture of cloud server

Advantages:

1. It can retrieve the results with less communication overhead.
2. It can provide the results with effective retrieval accuracy.
3. It can provide effective privacy and security application.

In the architecture we have three entities

1. Data owner: Data owner having collection of data files that he wants to outsource into the cloud server in encrypted format, this will increase effective data utilization.

2. Data user: When the data user wants to search the required files he enters a keyword in a secret form.

3. Cloud server: It is the place where a pool of data files and different applications can store.

Previously user can selects the files in the form of a plain text files. This is ailing under access the files. There is no perfect decryption technique to access the files of representation process. Here we introduce encryption based secure keyword searching mechanism. It can provide efficient solution for accessing the data. It is a good usability to display the effective matching details files [4]. These matching files are extracted with relevance score. This kind of matching files are retrieved with efficient mechanism. It can provide the results with guaranteed mechanism. All the files are collected with encryption format. All encrypted files are given weight in implementation process. These kinds of approaches show the better result in implementation.

The search result is displayed according to relevance score which improves file retrieval accuracy .In information retrieval process we maintain an inverted index to represent file ID's and relevance scores.

Inverted index maintains a list of mappings which corresponds to set of files that contain the keywords .*Ranking functions* are used to calculate relevance score of matching files for a given keyword[3] .To calculate the relevance score we use a statistical measurement i.e. TF*IDF ,where TF-term frequency and IDF –inverse document frequency.TF is number of times a given keyword find within a file and IDF is calculated by dividing the total number of files by the number of files containing the keyword[5].

$$score(Q, Fd) = \sum_{t \epsilon Q} \frac{1}{Fd} . (1 + \ln fd, t) . \ln (1 + \frac{N}{ft})$$

Here Q denotes the searched keywords, $f_{d,t}$ denotes the TF of term t in file F_d, f_t denotes the number of files that contain term t, N denotes the total number of files in the collection, and |Fd| is the length of file F_d, obtained by counting the number of indexed terms, functioning as the normalization factor.

3. Graded Keyword Search

Here we are going to implement graded searchable symmetric encryption based on existing searchable symmetric encryption schemes. Searchable symmetric encryption [6] allows data owner to outsource data in encrypted manner. In this data outsourcing, there may be loss of data by less security [7]. To overcome this we have ranked searchable symmetric encryption mechanisms (RSSE). This algorithm is based on four individual algorithms,they are "KeyGen, BuildIndex, TrapdoorGen and SearchIndex" which contributed the two step process.

Setup-Here data owner initialized public and private parameters by KeyGen algorithm and pre-process the data file C, to generate inverted index based on random keywords. The owner encrypts data file C and makes an index with relevance scores.

Retrieval-The user enters a keyword and makes a trapdoor by using TrapdoorGen, in the server. Then the cloud server will find the matching files, and displays file IDs based on Search Index algorithm.

Algorithm 1: One-to-many Order-preserving Mapping-OPM

Procedure OPMK (D, R, m,id(F))
While |D|! = 1 **do**
{D, R} ← BinarySearch(K,D,R,m);
End while
Coin $\xleftarrow{R}$ TapeGen(K,(D,R,1||m,id(F)));
c $\xleftarrow{coin}$ R;
Return c;
End procedure
Procedure BinarySearch(K,D,R,m);
M←D; N←R;
d←min (D)-1; r←min (R)-1;
y←r + [N/2];
Coin $\xleftarrow{R}$ TapeGen (K, (D, R, 0||y));
X $\xleftarrow{R}$ d + HYGEINV (coin,M,N,y-r);
If m≤x **then**
D ← {d +1,……., x};
R ←{r+ 1,……., y};
Else
D ←{x+ 1,……., d +M};
R ←{y+ 1,……., r + N};
end if
Return {D, R};
end procedure

In preserving symmetric encryption schema, numerical ordering of plaintext is preserved. In this, random order preserving will be considered. Thus high rate of information leakage is present. To avoid this we go for one-to-many Order preserving mapping, which preserves the data security and plaintext order. It incorporates random plaintext-to-bucket mapping of OPSE, and implemented in Algorithm 1 [3]. TapeGen( ) is a random coin generator, HYGEINV( ) is used instance of HGD( ). By the use of the OPM algorithm, server can rank the files efficiently.

4. Performance Analysis

Performance of our proposed system is calculated based on the effectiveness and efficiency.

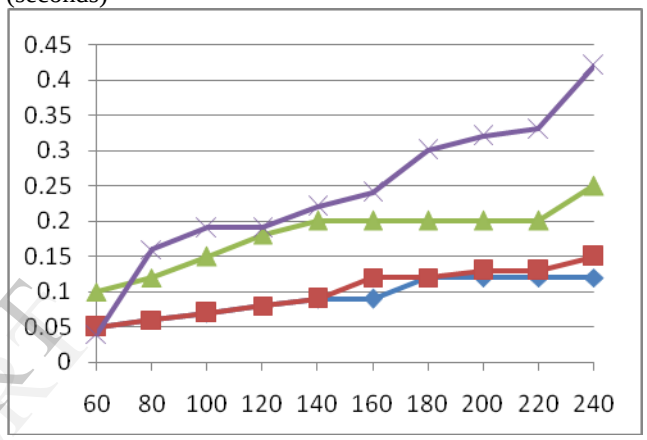

Fig.2: Time cost graph of OPM

This includes time and cost for searching the matching's. Effectiveness of one-to-many order preserving mapping is determined by relevance score R. Efficiency is measured by both size and range R. The result showing in Fig.2 is about different range values. Here the range values are as going changed and larger R, but the time cost for mapping took less than 220 milliseconds. When the range is 2^{40} then time cost is less than 70 milliseconds. This shows good performance compared to existing one.

5. Conclusion

In this paper we solve the security problems which may occur during outsourcing the data. In previously, there is no security for the data because for leakage. Any third party may hack the leaked data and also there is a burden in accessing the data items from different vast retrieved searches. To avoid this by implementing an OPSE (ordered preserving symmetric encryption) with RSSE mechanism. Present solution is best for effective data utilization and provides security for outsourced cloud

data compared to previous one. Hence accuracy is also increased.

6. References

1. 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.
2. D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. of IEEE Symposium on Security and Privacy'00, 2000.
3. A.Boldyreva, N. Chenette, Y. Lee, and A. O'Neill, "Order preserving symmetric encryption," in Proc. of Eurocrypt'09, volume 5479 of LNCS. Springer, 2009.
4. N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," in Proc.of INFOCOM'11, 2011.
5. C.Wang, Q.Wang, K.Ren, and W.Lou, "Toward Secure and Dependable Storage Services in Cloud Computing",IEEE Transactions on service Computing.
6. M. Bellare, A. Boldyreva, and A. O'Neill, "Deterministic and efficiently searchable encryption," in Proc. Of Crypto'07, volume 4622 of LNCS. Springer, 2007.
7. C. Wang, S. Chow, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Secure Cloudud Storage," IEEE Transactions on Computers (TC), to appear.
8. L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword searches over encrypted data," in Proc. Of ICICS'05, 2005.
9. N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data" in Proc.of INFOCOM'11, 2011.
10. D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. of EUROCRYP'04, volume 3027 of LNCS. Springer, 2004.

About The Authors

P.Naresh, received his B.Tech degree in Computer Science and Engineering from Jawaharlal Nehru Technological University, Anantapur, India, in 2011. Currently pursuing M.Tech in computer science and engineering at KVSR Institute of Technology, Kurnool, India. His interesting research area is Cloud Computing, Network Security and Data Mining.

K.Pavan kumar, received his B.Tech degree in Computer Science and Engineering from Jawaharlal Nehru Technological University, Hyderabad, India. 2006; M.Tech in Computer Science from Jawaharlal Nehru Technological University, Hyderabad, India.in 2009 . He is an Asst.Professor at DR.K.V.S.R.I.T, Kurnool, India. His current research is on Mobile Ad-hoc Networks, computer networks and cloud computing

D.K.Shareef received his M. Tech degree in Computer Science from S.K.T.R.M College of Engineering KURNOOL, India in 2012, B.Tech form JNTU Anantapur 2010 respectively. He is currently working as Assistant Professor at INTELL Engineering College of JNTUA. Published FOUR International Conferences and TWO National Conference.