International Journal of Research in Advent Technology, Vol.4, No.5, May 2016
E-ISSN: 2321-9637
Available online at www.ijrat.org

Implementing a Multikeyword Ranked Retrieval Search using an Efficient Procure Technique over Cipher Data

Mr. Vilas D. Ghonge[1], Prof Dr.V.S.Gulhane[2],
Student , Computer Science and Engineering Department, Sipna college of Engineering and Technology[1,]
Associate Professor , Information Technology Department, Sipna college of Engineering and Technology
[2]*,Amravati,Maharashtra,India.*
vilasghonge77@gmail.com[1] *, v_gulhane@rediffmail.com*[2]

Abstract- Cloud Computing provides unlimited data storage services and high performance Computing and reduces the cost of managing large amount of data. This paper solves and defines the problem of multikeyword ranked search over encrypted Cloud Data while preserving the security. Therefore it is necessary to outsource the data on the Cloud in order to maintain Data security. Existing Keyword search support only Boolean queries i.e either a file matches or does not match a query. Boolean Search technique only produces the unsorted result. Here we have proposed an efficient, secure and fast data searching technique which will help us to handle data efficiently in Cloud Storage or Server. Multikeyword ranked retrieval search technique helps user to retrieve relevant data or files in which they are interested in and also the search operation performed over encrypted data , information leakage can be eliminated and data can be searched and retrieved efficiently.

Index Terms- :- Cloud Computing, Upload Data, Data Retrieval, Multikeyword Search, Boolean Queries.

1. INTRODUCTION

In Cloud Computing, Data Owner may share their outsourced data with a number of users, who only wants to retrieve the data files they are interested in. Multikeyword Ranked retrieval Search is the best option to implement this concept. It is preferred to get the retrieval result with the most relevant files that match users interest instead of all the files, which indicates that the files should be ranked in the order of relevance by users interest and only the files with the highest relevance are sent back to the users. A series of Searchable Symmetric encryption Schemes have been Proposed to enable search on encrypted data, whether a keyword exists in a file or not without considering the difference of relevance with the queried keyword of these files in the result.

2. LITERATURE REVIEW

In the work [1], describes the cryptographic schemes for the problem of searching on encrypted data and provide confidentiality preserving ranked ordered search. A searchable encryption techniques [3], [4] are able to provide secure search over encrypted data for users. They build a searchable inverted index that stores a list of mapping from keywords to the corresponding set of files which contain this keyword. When data users input a keyword, a trapdoor is generated for this keyword and then submitted to the cloud server. Some researchers study the problem on secure and ranked search over outsourced cloud data. In the work [7], introduces a new framework for confidentiality preserving top k-retrieval using multikeyword . In the work [10], the advent of cloud computing, data owners are motivated to outsource their complex data management systems from local sites to the commercial public cloud for great flexibility and economic savings. But for protecting data privacy, sensitive data has to be encrypted before outsourcing, which obsoletes traditional data utilization based on plaintext keyword search. It shows increasing importance as cloud computing drives more businesses to outsource their data and querying services. However, most existing studies, including those on data outsourcing, address the data privacy and query privacy separately and cannot be applied to this problem. In the work [6], the main idea is to formalize and provide solution to the problem of effective fuzzy keyword search over encrypted cloud data while maintaining keyword privacy. The basic theme in this proposed work is multi-keyword ranked search (MRSE scheme) over cipher data. We believe this work steps towards practical applications of privacy homomorphism to secure query processing on large-scale, structured datasets.

3. SYSTEM ARCHITECTURE

Cloud data storage service involves three parts 1.Data Owner Module 2.Data User Module and 3.File Uploaded Module (Server Module) as shown in Fig. Data Owner stores a set of document D on to the cloud server

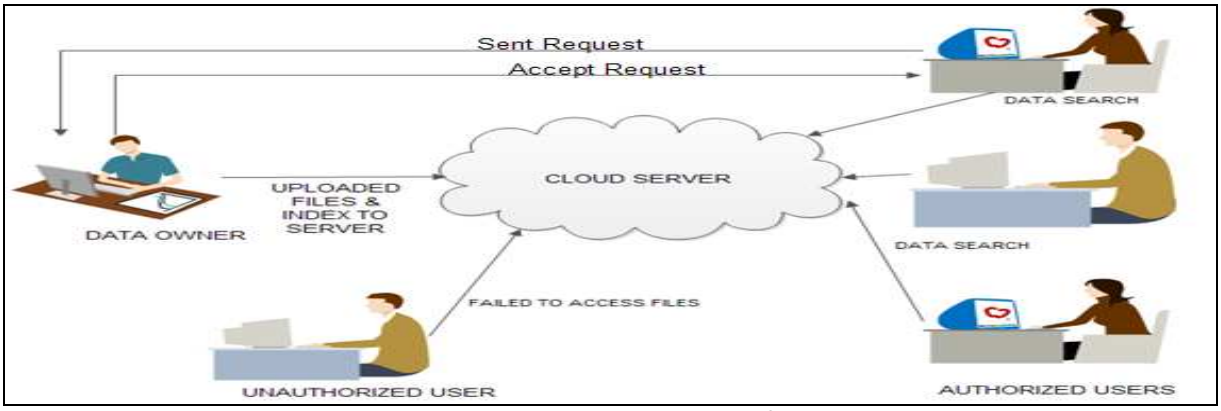


Fig.1. Components and Architecture of System

to avoid the security threats. To make a search, a set of keywords is given by an authorized user. The results are ranked using the ranking algorithm by the cloud server. The various components present in the architecture of proposed system are actual user stores the data on cloud. Cloud server stores the data and searching multikeyword. Data user retrieves the file from the cloud server using multikeyword ranked search.

4. MODULE DESCRIPTION

There are three Modules in this Proposed Work

1) Data Owner Module
2) Cloud Server Module
3) Data User Module

First of all Data Owner Signup in the system and provide his username ,password and some other information .Then Data Owner upload multiple number of data files providing with three to four keyword of each file in the database .The files and Encrypted multiple keyword are stored in the database . When Data user search particular file in search text Window with multiple keyword, he gets links of that particular file. In order to access that particular file he has to register his username and password on the site .After registering a request is to transferred to the Data Owner. Data Owner either accept or reject a request from a particular Data user through his Account. After getting Authentication from Data Owner Data user can access a file from his Registered Account. Data Owner can also check how many number of Users are there and who has accessed which file.

ALGORITHM:

Algorithm Multikeyword search
{
For all Document Ni do
{
Compare (level index of Ni, Query word) j= 1
While match do
{
Increament j
Compare (level indices of Ni , Query word)
End while
}
Rank of Ni highest level that match with Query Word
End for
}
}

Suppose we have uploaded N no. of Documents in the database and provided query word or multikeyword to each document in the database. Then we have searched a document by firing a query (Multikeyword) to the database. Database retrieve most relevant files according to the query and provide it to the user.

FLOW OF PROJECT

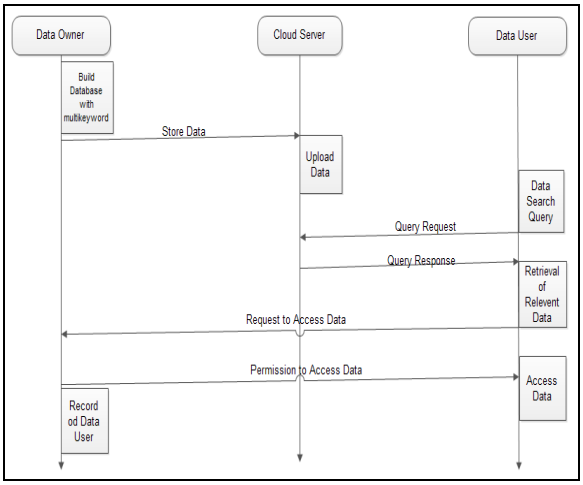


Fig.2. Flow of Project

5. EXPERIMENTAL SETUP

A. Hardware Requirement:

Processor	:	Pentium Dual core with 1.2 GHz
RAM	:	1GB
Hard Disk	:	20 GB
Key Board	:	Standard Windows Keyboard

B. Software Requirement:

Operating System	:	Windows OS
Technology	:	PHP
Web Server	:	WampServer
Database	:	Mysql

C. Dataset: For dummy records we have uploaded 50 files, data user selects random files from this dataset to download data from this clouds .

D. Keyword set: We have selected approximately top 200 keywords for index generation in the given dataset.

E. System Setup: We have tested the system on single node machine for different applications i.e. Data user application, Data Owner, KDC Cloud is hosted on same Windows-7 system. Mysql Database is used.

Comparative Analysis & Discussions

1.Existing System retrieve irrevalent amount of data e.g suppose we search for a particular file on browser unnecessary files are retrieved from that search it causes loss of time and data Wastage. In case of Proposed System relevant amount of data is retrieved it leads to the save of time and data wastage is avoided.

2. Existing System does not maintain auditing details or record of Data owner hence information maintenance is not to be carried out properly. In case of Proposed System Auditing details and record is maintained.

3. Privacy shield is provided as only authorized persons are able to access important data files.

4. Number of Documents Vs Searching Time (ms)

The following graph shows that Searching Time (ms) required to Retrieve Documents is less in Implemented System Compared to that of Existing System.

Table 1. Existing System

Existing System	
Sr. No. of Document	Searching Time (ms)
1	0.65
2	0.55
3	0.42
4	0.65
5	0.45

Table 2. Proposed System

Proposed System	
Number of Keyword	Searching Time (ms)
1	0.15
2	0.06
3	0.21
4	0.087
5	0.028

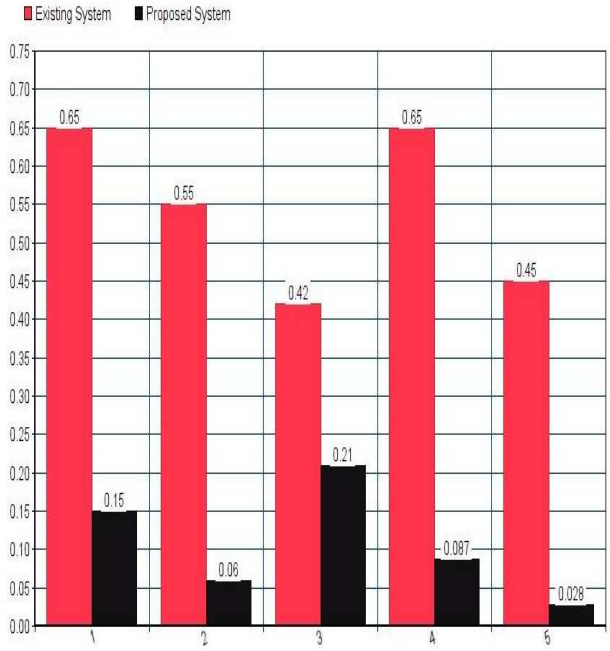
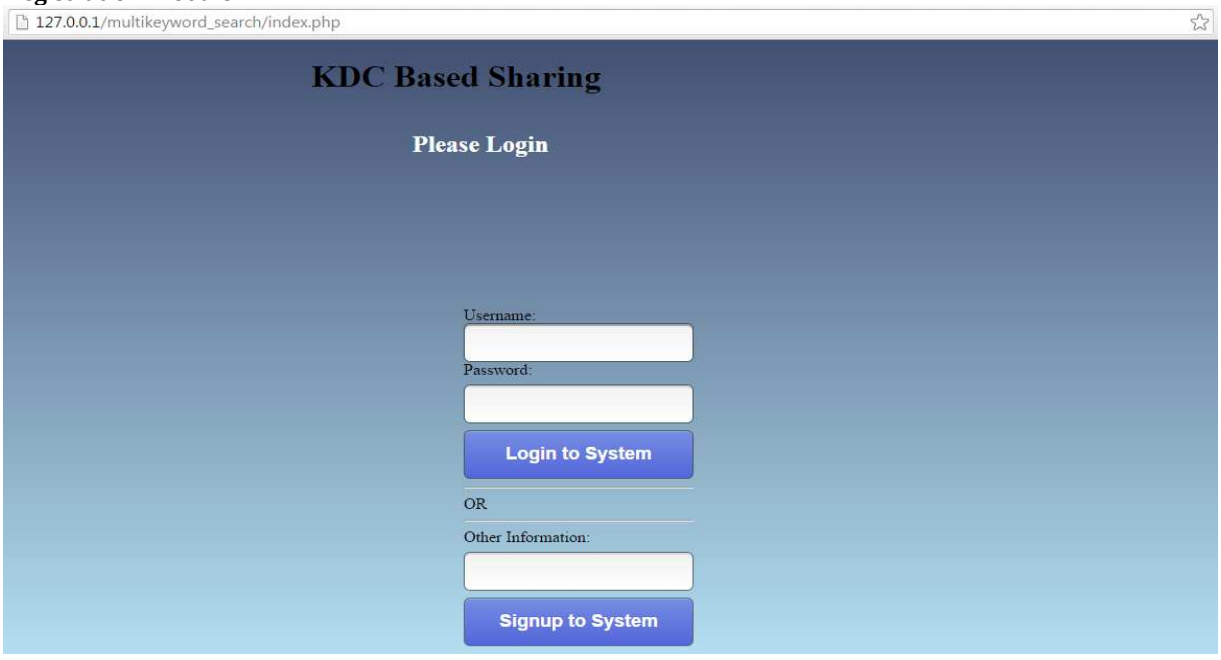


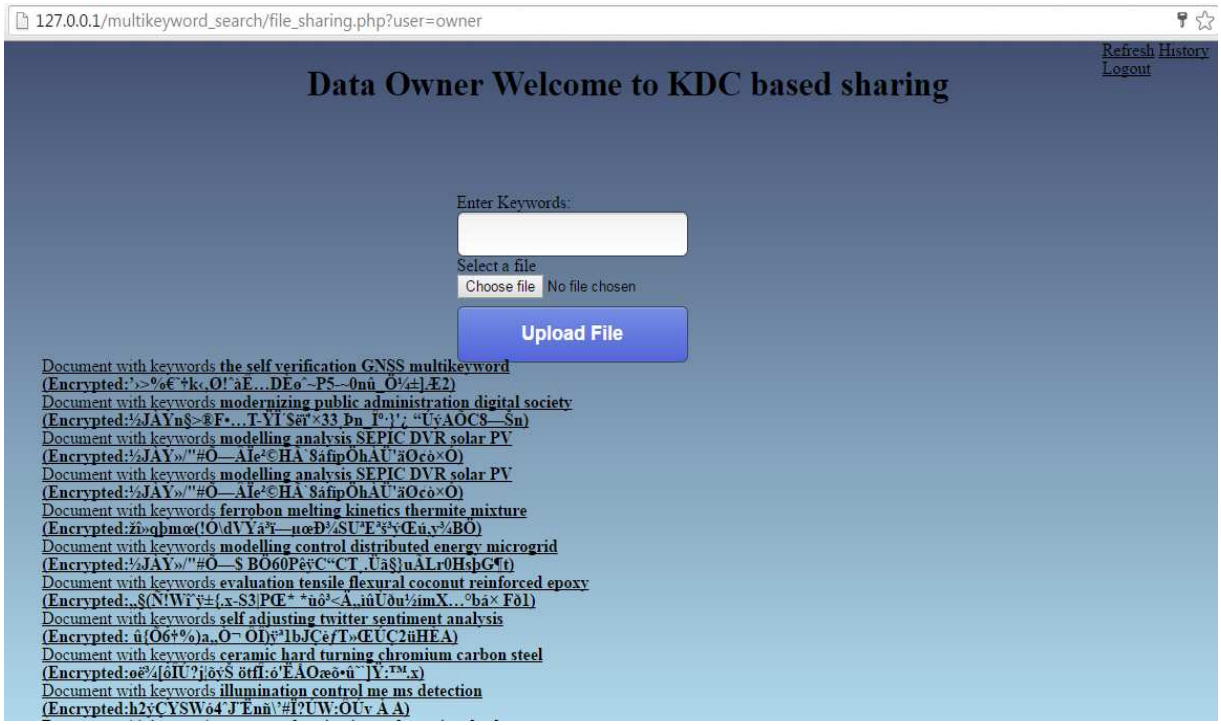
Fig.1.3. Number of Keyword Vs Searching Time

SCREENSHOTS

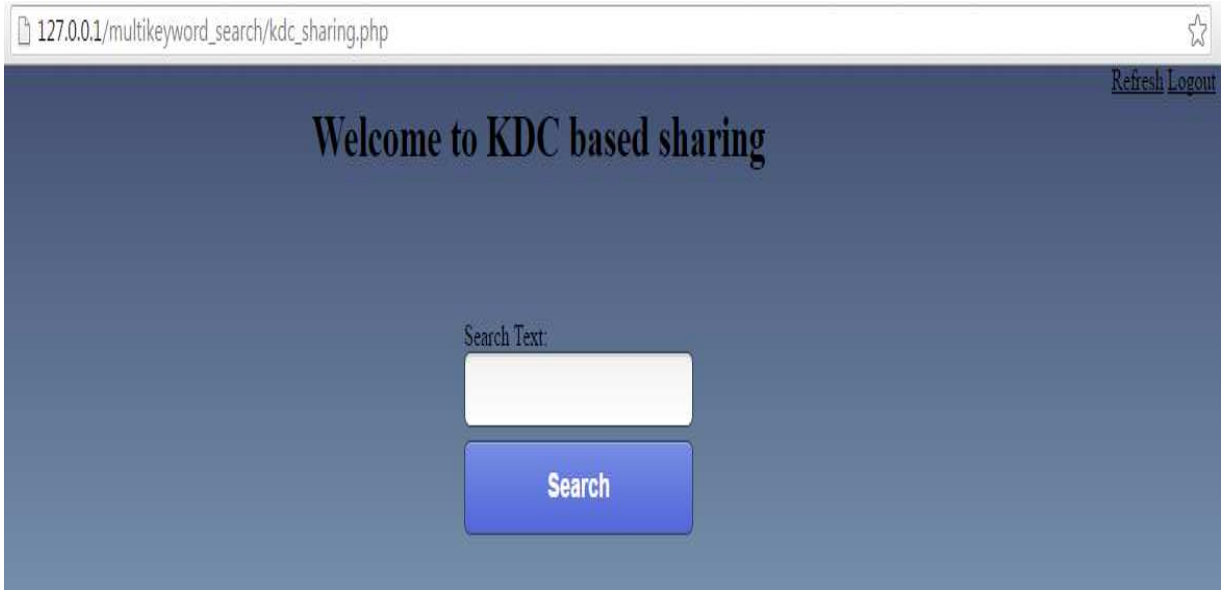
Registration Module



Data Owner Module



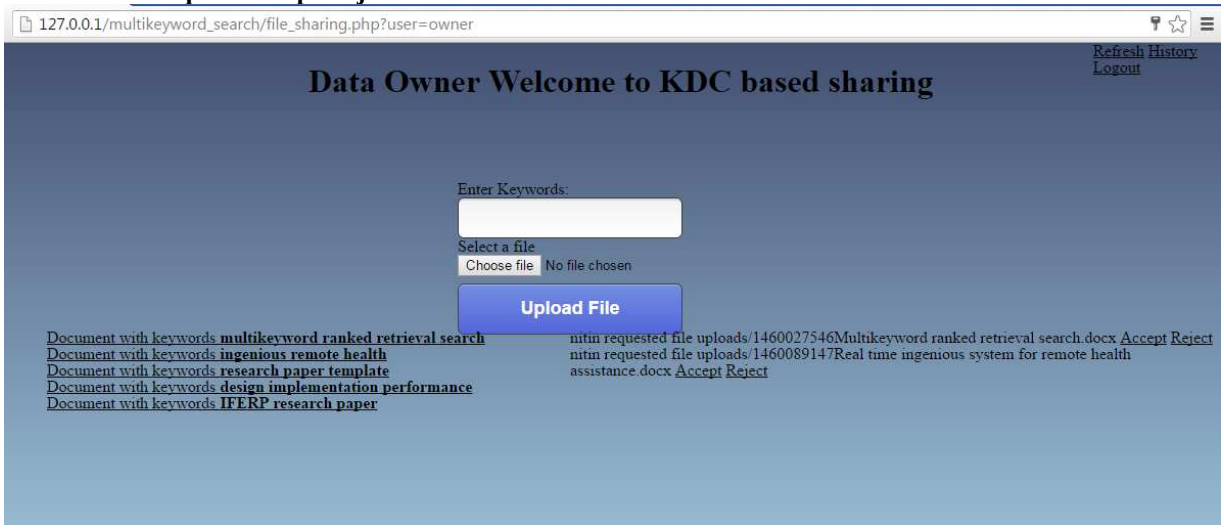
Search Text Module



Result Showing Search Text Module

Data User Module

Data Owner Request Accept/Reject Module



File Download Module

Data owner Record Module

6. CONCLUSION

In this paper we are trying to propose a system which is able to provide multikeyword ranked retrieval search over encrypted cloud data which improves the basic privacy requirement of data .It also provides efficiency of a system by low communication and computation overhead. Hence it will enable fast similarity search over Cloud data without compromising the security. The multikeyword ranked retrieval search technique will provide more efficiency and data can be searched and retrieved efficiently.

REFERENCES

[1] A. Swaminathan, Y. Mao, G.-M. Su, H. Gou, A.L. Varna, S. He, M.Wu, and D.W. Oard, "Confidentiality-Preserving Rank-Ordered Search," Proceeding Workshop Storage Security and Survivability, 2007.

[2] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure Ranked Keyword Search over Encrypted Cloud Data," Proceeding IEEE 30th International Conference Distributed Computing Systems (ICDCS), 2010.

[3] D. Boneh, G. Crescenzo, R. Ostrovsky, and G. Persiano, "Public- Key Encryption with Keyword Search," Proceeding International Conference Theory and Applications of Cryptographic Techniques (Eurocrypt), 2004.

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

[5] H. Hu, J. Xu, C.Ren, and B. Choi, "Processing Private Queries over Untrusted Data Cloud through Privacy Homomorphism," Proceeding IEEE 27th International Conference Data Engineering (ICDE), 2011.

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

[7] Jiadi Yu,"Toward Secure Multikeyword Top-k Retrieval over Encrypted Cloud Data" IEEE transactions on dependable and secure computing, vol. 10, no. 4, July/August ,2013.
[8] M. Perc, "Evolution of the Most Common English Words and Phrases over the Centuries," J. Royal Societies Interface, 2012.
[9] M.van Dijk, C. Gentry, S. Halevi, and V.Vaikuntanathan, "Fully Homomorphic Encryption over the Integers," Proceeding 29th Annual International Conference Theory and Applications of Cryptographic Techniques, 2010.
[10] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-Preserving Multikeyword Ranked Search over Encrypted Cloud Data," Proceeding IEEE INFOCOM, 2011.
[11] N. Howgrave-Graham, "Approximate Integer Common Divisors," Proceeding Revised Papers from International Conference Cryptography and Lattices (CaLC‟ 01), pp. 51-66, 2001.
[12] O. Regev, "New Lattice-Based Cryptographic Constructions," J. ACM, vol. 51, no. 6, pp. 899-942, 2004.
[13] S. Gries, "Useful Statistics for Corpus Linguistics," A Mosaic of Corpus Linguistics: Selected Approaches, Aquilino Sanchez Moises Almela, eds., pp. 269-291, Peter Lang, 2010.