Impact Factor (SJIF): 4.542

International Journal of Advance Research in Engineering, Science & Technology

e-ISSN: 2393-9877, p-ISSN: 2394-2444

Volume 4, Issue 5, May-2017

Secure Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing

Rahul S Badhekar[1],Devika K Shete[2], Reshma K Shinde[3], Prof.Shrikant N Dhage[4]

[1]*Final Year Student of Deptartment of Computer Engineering, Jaihind College of Engineering, University of Pune, India*

[2]*Final Year Student of Deptartment of Computer Engineering, Jaihind College of Engineering, University of Pune, India*

[3]*Final Year Student of Deptartment of Computer Engineering, Jaihind College of Engineering, University of Pune, India*

[4]*Assistant Professor of Deptartment of Computer Engineering, Jaihind College of Engineering, University of Pune, India*

ABSTRACT: The commencement of cloud computing, it has become with time trendy for data owners to outsource their data to public cloud servers while permit data users to recover this data. For privacy concerns, protected searches over encrypted cloud data have aggravated several research works under the single owner model. However, most cloud servers in practice do not just serve one owner; instead, they support multiple owners to share the benefits brought by cloud computing. In this paper, we propose schemes to deal with Security Maintaining Ranked Multi-keyword Search in a Multi-owner model (PRMSM). To allow cloud servers to perform safe search without knowing the real data of both keywords and trapdoors, we scientifically build a original secure search protocol. To rank the search results and preserve the privacy of relevance scores between keywords and files,The propose a novel preservative arrange and Privacy Preserving Function family. To protect the attackers from snooping secret keys and imagine to be official data users submitting searches, we propose a new dynamic secret key creation protocol and a new data user verification protocol. Also, PRMSM supports well-organized data user revocation. wide experiments on real-world datasets confirm the efficacy and efficacy of PRMSM.

KEYWORDS: PRSM,Cloud computing,Ranked keyword search, multiple owners, Privacy preserving, dynamic secret key, Cryptography, Authentication.

1. INTRODUCTION

Cloud computing is the wide visualized idea of computing as a service, where cloud customers can somewhat store their data into the cloud so as to like the on-demand high quality applications and services from a collective pool of configurable computing resources. Its huge flexibility and economic savings are motivating both individuals and enterprises to outsource their limited complex data management system into the cloud. To protect data privacy and conflict unsolicited accesses in the cloud and beyond, sensitive data, e.g., emails, personal health records, photo albums, tax documents, financial transactions, etc., may have to be encrypted by data owners before outsourcing to the commercial public cloud this, however, obsoletes the conventional data utilization service based on plaintext keyword search. The small solution of downloading all the data and decrypting locally is clearly impractical, due to the huge amount of bandwidth cost in cloud extent systems. Moreover, aside from eliminating the local storage administration,storing data into the cloud serves no.purpose unless they can be easily searched and utilized. Thus, exploring privacy-preserving and effective search service over encrypted cloud data is of supreme significance. Considering the potentially large number of on-demand data users and huge amount of outsourced data documents in the cloud, this problem is particularly difficult as it is extremely hard to meet also the requirements of performance, system usability and scalability. On the one hand, to meet the effective data recovery need, the large amount of documents insist the cloud server to perform result relevance grade, instead of returning undifferentiated results. Such ranked search system enables data users to find the most relevant information quickly, rather than burdensomely categorization through every match in the content collection. Ranked search can also gracefully eliminate unnecessary network traffic by sending back only the most relevant data, which is highly desirable in the "pay-as-youuse" cloud model. For privacy protection, such ranking operation, however, should not seep out any keyword related information. On the other hand, to get better the search result.

Accuracy as well as to enhance the user searching experience, it is also necessary for such grade system to support multiple keywords search, as single keyword search frequently yields far too coarse results. As a common practice indicated by today's web search engines (e.g., Google search), data users may tend to provide a set of keywords instead of only one as the indicator of their search interest to retrieve the most related data. And each keyword in the search request is able to help fine down the search result further. "Coordinate matching", i.e., as many matches as possible, is an

well-organized similarity measure among such multi-keyword semantics to refine the result significance, and has been widely used in the plaintext information retrieval (IR) community. However, how to apply it in the encrypted cloud data search system remains a very difficult task because of intrinsic security and privacy obstacles, including various strict requirements like the data privacy, the index privacy, the keyword privacy, and many others. In the literature, searchable encryption is a helpful technique that treats encrypted data as documents and allows a user to strongly search through a single keyword and retrieve documents of attention. However, direct application of these approaches to the secure large scale cloud data consumption system would not be necessarily suitable, as they are residential as crypto primitives and cannot accommodate such high service-level necessities like system usability, user searching experience, and easy information discovery. Although some recent designs have been proposed to support Boolean keyword search as an effort to enrich the search flexibility, they are still not sufficient to provide users with satisfactory result position functionality. Our before time work has been aware of this problem, and provided a solution to the secure ranked search over encrypted data problem but only for queries consisting of a single keyword. How to design an efficient encrypted data search method that supports multi-keyword semantics without privacy breaches still remains a difficult open problem.

2. LITERATURE SURVEY

1) Single keyword search:

Author: Deepali D. Rane et.al, D. Song, D. Wagner et al,Y.-C. Chang et al,C. Wang et al

proposed implementation of the encryption and decryption, Secure index construction issuccessfully completed with desirable performance. After index construction it will get compressed and will be stored in.cfs file format. After firing single-keyword query, user will get all documents that contain the specified keyword. Theadvantages are protects data privacy by encrypting documents before outsourcing, rank based retrieval of the documents,To easily access the encrypted data by multi keyword rank search using keyword index. The Disadvantages of the proposedsystem are single-keyword search without ranking, Boolean keyword searching without ranking, single-keyword searchwith ranking, Rarely sorting of the results i.e. no index creation and ranking, Single User search.

2) Multi-keyword search:

Author: Zhihua Xia et.al,Bing Wang et.al,Yanzhi Ren et.al, Hongwei Li et.al

proposed a secure, efficient and dynamic search scheme, which supports not only the accurate multikeywordranked search but also the dynamic deletion and insertion of documents. They construct a special keywordbalanced binary tree as the index, and proposed a "Greedy Depth-first Search" algorithm to obtain better efficiency thanlinear search. In addition, the parallel search process can be carried out to further reduce the time cost. The security of thescheme is protected against two threat models by using the secure KNN algorithm. Experimental results demonstrate the efficiency of proposed scheme. The Advantages of the proposed system are searchable encryption schemes enable theclient to store the encrypted data to the cloud and execute keyword search over cipher text domain and a secure tree-basedsearch scheme over the encrypted cloud data, which supports multi-keyword ranked search and dynamic operation on thedocument collection. The disadvantages are the cloud service providers (CSPs) that keep the data for users may accessusers sensitive information without authorization. A general approach to protect the data confidentiality is to encrypt the data before outsourcing. However, this will cause a huge cost in terms of data usability.

3. Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing

Author: Wei Zhang , Yaping Lin , Sheng Xiao, Jie Wu ,Siwang Zhou:

With the advent of cloud computing, it has become increasingly popular for data owners to outsource their data to public cloud servers while allowing data users to retrieve this data. For privacy concerns, secure searches over encrypted cloud data has motivated several research works under the single owner model. However, most cloud servers in practice do not just serve one owner; instead, they support multiple owners to share the benefits brought by cloud computing. In this paper, we propose schemes to deal with privacy preserving ranked multi-keyword search in a multi-owner model (PRMSM). To enable cloud servers to perform secure search without knowing the actual data of both keywords and trapdoors, we systematically construct a novel secure search protocol. To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a novel additive order and privacy preserving function family. To prevent the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol. Furthermore, PRMSM supports efficient data user revocation. Extensive experiments on real-world datasets confirm the efficacy and efficiency of PRMSM.

3. RELATED WORK

In this paper, for the first time, we define and solve the problem of multi-keyword ranked search over encrypted cloud data (MRSE) while preserving demanding system-wise privacy in the cloud computing paradigm. Among a variety of multi keyword semantics, we choose the efficient match measure of "coordinate matching", i.e., as many matches as feasible, to capture the relevance of data documents to the search reservation. Specifically, we use "inner product similarity" i.e., the number of query keywords appearing in a document, to quantitatively estimate such similarity compute of that document to the search query. During the index construction, each document is associated with a binary vector as a sub-index where each bit represents whether corresponding keyword is contained in the document. The search query is also described as a binary vector where each bit means whether corresponding keyword appears in this search request, so the similarity could be exactly measured by the inner product of the query vector with the data vector. However, directly outsourcing the data vector or the query vector will violate the index privacy or the search privacy. To meet the challenge of supporting such multi-keyword semantic without privacy breaches, we propose a basic idea for the MRSE using secure inner product computation, which is adapted from a secure k-nearest neighbor (*kNN*) technique.

4. PROPOSED SYSTEM

We define a multi-owner model for privacy preserving keyword search over encrypted cloud data. We propose an capable data user authentication protocol, which not only prevents attackers from eavesdropping secret keys and pretending to be prohibited data users performing searches, but also enables data user authentication and revocation. We methodically construct a novel secure search protocol, which not simply enables the cloud server to perform secure ranked keyword search without knowing the real data of both keywords and trapdoors, but also allows data owners to encrypt keywords with self-chosen keys and allows legal data users to query without knowing these keys. We propose an Additive Order and Privacy Preserving Function family (AOPPF) which allows data owners to defend the privacy of relevance scores using special functions according to their preference, while still permitting the cloud server to rank the data files perfectly. We conduct extensive experiments on real-world datasets to verify the efficacy and efficiency of our proposed schemes. system of search as show in below Fig.1

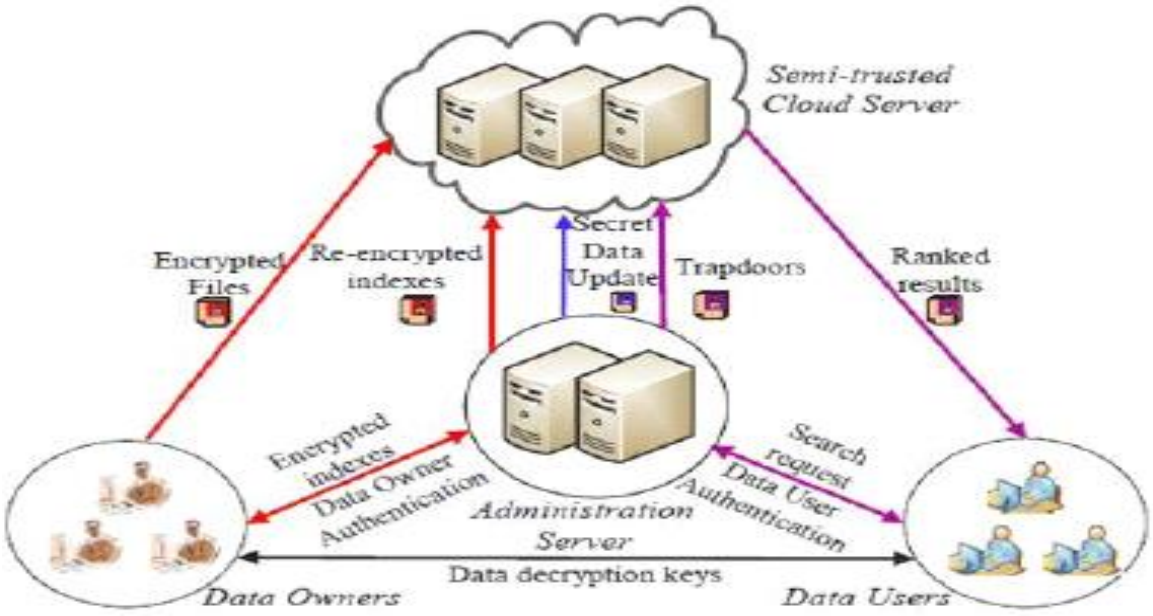


Fig.1.Proposed system architecture.

System accomplishment consist of various parts described as follows: We are implementing our project by using Java Technology and MySQL database. Various components of our system are:

1. Data Owner.
2. Data user.
3. Application server.
4. Cloud server.

1. Data Owner: Data owner have the set of files, they create the index file ad send that file to the application server. Finally Data owner encrypt that file and send encrypted file to the cloud server as well as send the encryption key to the data user.

2. Application server: Application server re-encrypt the index file of legitimate user and send that re-encrypted file to the cloud server.
3. Data user: Data user send keywords to search to words the application server, application server send that ask for to the cloud server if the data user is the legal user by creating the trapdoor.
4. Cloud server: Upon receiving the trapdoor, the cloud server searches the encrypted index of each data owner and returns the matching set of encrypted files.

5. ALGORITHM

Advance Encryption Standard :

AES is an iterative rather than Feistel cipher. It is based on 'substitution–permutation network'. It comprises of a series of linked operations, some of which involve replacing inputs by specific outputs (substitutions) and others involve shuffling bits around (permutations).
Interestingly, AES performs all its computations on bytes rather than bits. Hence, AES treats the 128 bits of a plaintext block as 16 bytes. These 16 bytes are arranged in four columns and four rows for processing as a matrix
Unlike DES, the number of rounds in AES is variable and depends on the length of the key. AES uses 10 rounds for 128-bit keys, 12 rounds for 192-bit keys and 14 rounds for 256-bit keys. Each of these rounds uses a different 128-bit round key, which is calculated from the original AES key.

- The schematic of AES structure is given in the following illustration –
- AES Structure
- Encryption Process

Here, we restrict to description of a typical round of AES encryption. Each round comprise of four sub-processes. The first round process is depicted below –

- First Round Process
- Byte Substitution (SubBytes)

The 16 input bytes are substituted by looking up a fixed table (S-box) given in design. The result is in a matrix of four rows and four columns.

- Shiftrows

Each of the four rows of the matrix is shifted to the left. Any entries that 'fall off' are re-inserted on the right side of row. Shift is carried out as follows –

I. First row is not shifted.
II. Second row is shifted one (byte) position to the left.
III. Third row is shifted two positions to the left.
IV. Fourth row is shifted three positions to the left.

The result is a new matrix consisting of the same 16 bytes but shifted with respect to each other.

- MixColumns

Each column of four bytes is now transformed using a special mathematical function. This function takes as input the four bytes of one column and outputs four completely new bytes, which replace the original column. The result is another new matrix consisting of 16 new bytes. It should be noted that this step is not performed in the last round.

- Addroundkey

The 16 bytes of the matrix are now considered as 128 bits and are XORed to the 128 bits of the round key. If this is the last round then the output is the ciphertext. Otherwise, the resulting 128 bits are interpreted as 16 bytes and we begin another similar round.

- Decryption Process

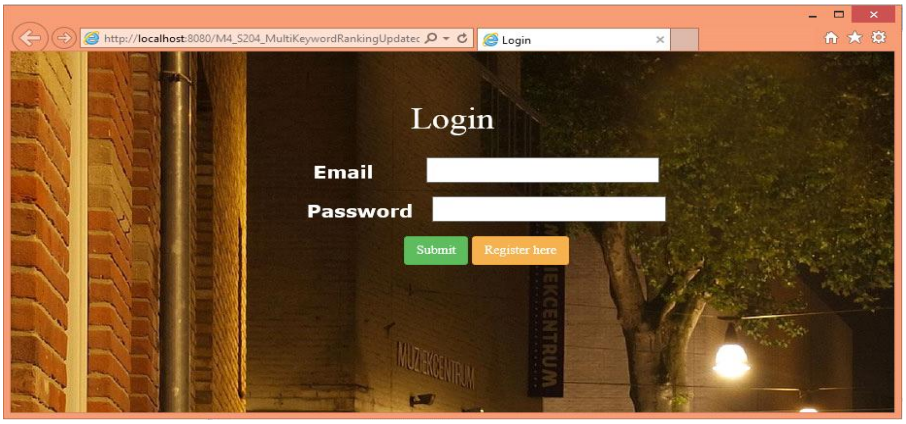
The process of decryption of an AES ciphertext is similar to the encryption process in the reverse order. Each round consists of the four processes conducted in the reverse order –

I. Add round key
II. Mix columns
III. Shift rows
IV. Byte substitution

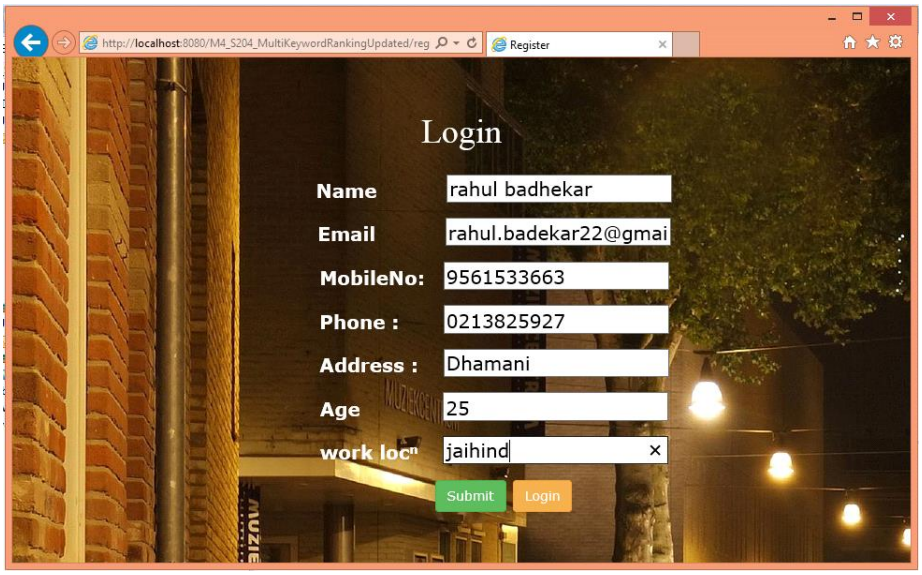
Since sub-processes in each round are in reverse manner, unlike for a Feistel Cipher, the encryption and decryption algorithms needs to be separately implemented, although they are very closely related.

6. RESULT AND DISCUSSION

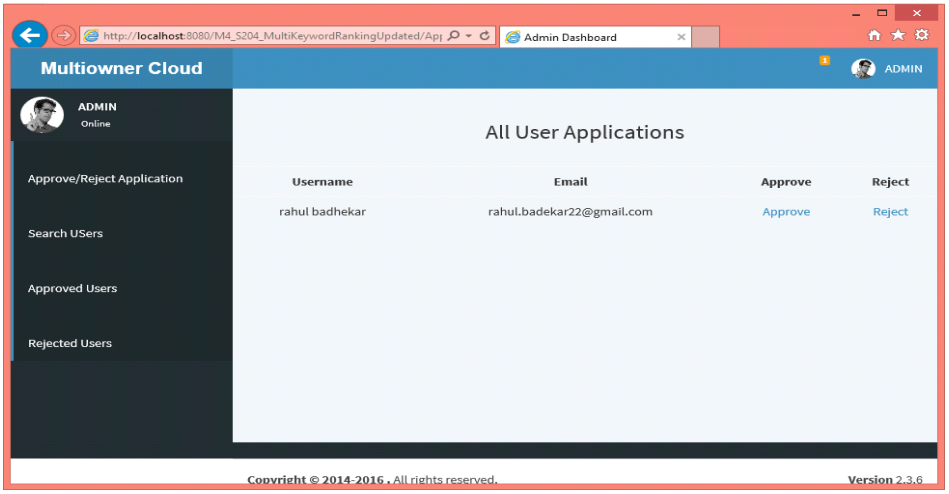
1.Home Page



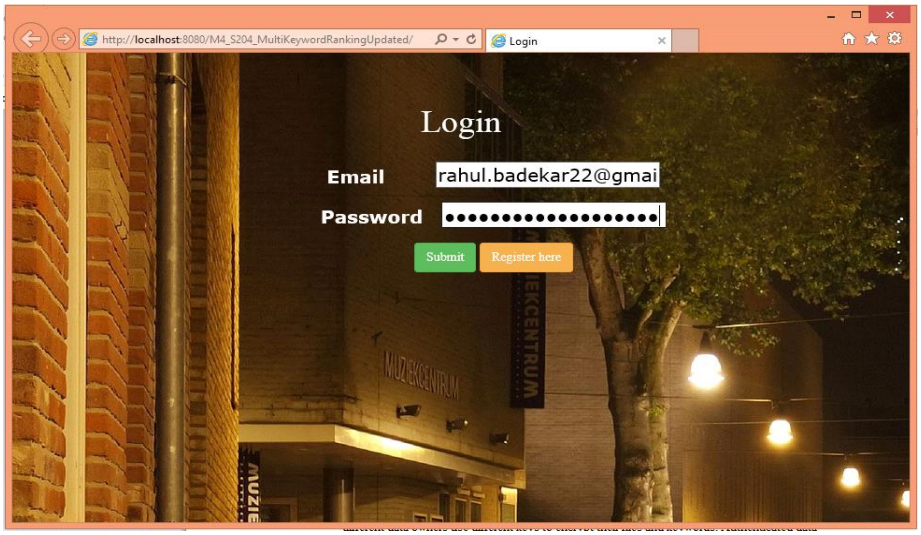
2 .User Register Page



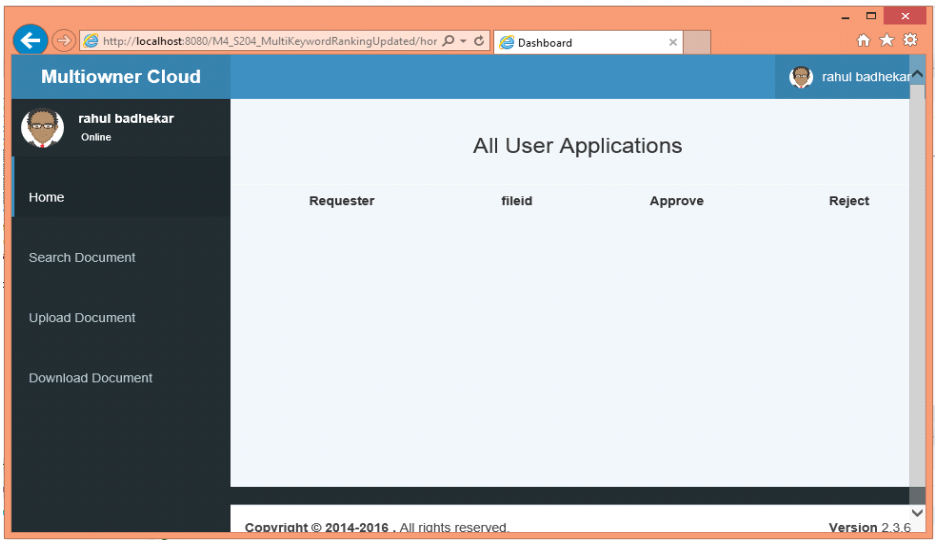
3. Admin Approve / Reject List



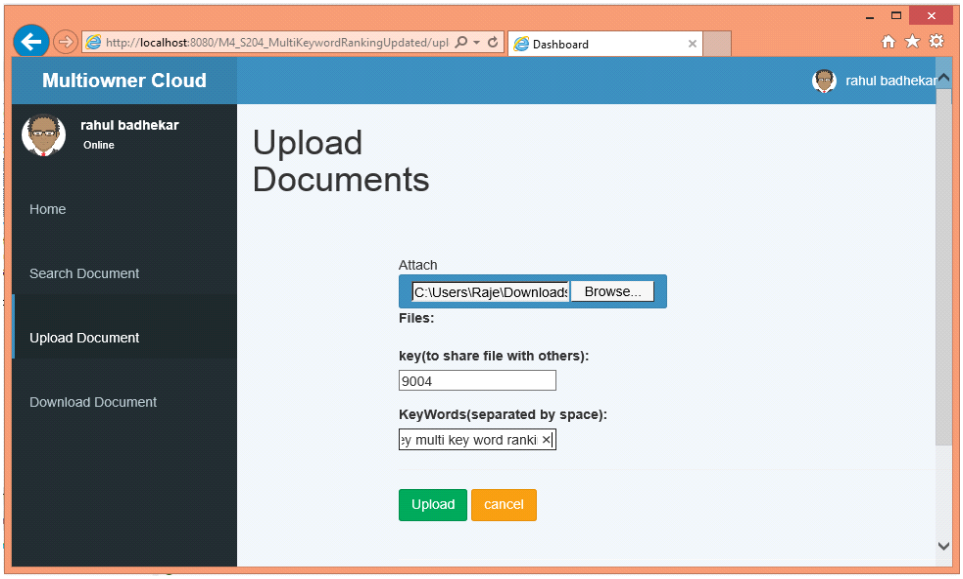
3.User Login Page



5.User Home Page



6. User File Upload Page.



7. ADVANTAGES OF PROPOSED SYSTEM:

I. The projected scheme allows multi-keyword search over encrypted files which would be encrypted with different keys for dissimilar data owners.
II. The proposed scheme allows new data owners to enter this system without disturbing other data owners or data users, i.e. the scheme supports data owner scalability in a plug-and-play model.
III. The proposed scheme ensures that only valid data users can perform correct searches. Moreover, once a data user is revoked, he can no longer perform correct searches over the encrypted cloud data.
IV. To enable cloud servers to perform secure search without knowing the actual value of both keywords and trapdoors, we systematically build a novel secure search protocol. As a result, different data owners use different keys to encrypt their files and keywords. valid data users can subject a query without significant secret keys of these different data owners.
V. To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a new additive order and privacy preserving function family, which helps the cloud server return the most related search results to data users without revealing any responsive information.
VI. To avert the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol.

8. CONCLUSION

In this paper, for the main time we define and solve the problem of multi-keyword ranked search over encrypted cloud data, and start a variety of privacy necessities. Among various multi-keyword semantics, we choose the efficient similarity measure of "coordinate matching", i.e., as many matches as feasible, to successfully capture the relevance of outsourced documents to the query keywords, and use "inner product similarity" to quantitatively evaluate such similarity compute. For meeting the challenge of supporting multi-keyword semantic without privacy breaches, we propose a basic plan of MRSE using secure inner product working out. Then we give two significantly enhanced MRSE schemes to achieve various severe privacy requirements in two different risky models. Thorough analysis investigating privacy and efficiency guarantees of proposed schemes is given, and experiments on the real-world dataset show our proposed schemes establish low overhead on both computation and communication. As our future work, we will explore sustaining other multi keyword semantics (e.g., weighted query) over encrypted data, integrity confirm of rank order in search result and privacy guarantees in the stronger threat model.

9. REFERENCES

- 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.
- S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving secure, scalable, and fine-grained data access control in cloud computing," in *Proc. OfINFOCOM*, 2010.
- I. H. Witten, A. Moffat, and T. C. Bell, "Managing gigabytes: Compressing and indexing documents and images," Morgan Kaufmann Publishing, San Francisco, May 1999.
- R. Curtmola, J. A. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions,"in*Proc. of ACM CCS*, 2006
- A. Singhal, "Modern information retrieval: A brief overview," *IEEE Data Engineering Bulletin*, vol. 24, no. 4, pp. 35–43, 2001.
- L. M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner, "A break in the clouds: towards a cloud definition," *ACM* D.Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Proc. of S&P*, 2000.

- S. Kamara and K. Lauter, "Cryptographic cloud storage," in *RLCPS, January 2010, LNCS. Springer, Heidelberg*.
- C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for data storage security in cloud computing," in *Proc. of NFOCOM*, 2010
- R. Brinkman, "Searching in encrypted data," in *University of Twente, PhD thesis*, 2007.
- D.Shama and A.kush,'GPS Enabled E Energy Efficient Routing for Manet', International Journal of Computer Networks (IJCN), Vol.3, Issue 3, pp. 159-166, 2011.