IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

# TOWARDS ACHIEVING KEYWORD SEARCH OVER DYNAMIC ENCRYPTED CLOUD DATA WITH SYMMETRIC-KEY BASED VERIFICATION

**Dr K. Vijaya Bhaskar[1] , U. Muni Sekhar[2], S. Nayaz[3], B VaraLakshmi[4]**

[1] Associate Professor, Department of Computer Applications , Chadalawada Ramanamma Engineering College Tirupati , Andhra Pradhesh, India.

[2,3,4] Student, Department of Computer Applications, Chadalawada Ramanamma Engineering College Renigunta Rd, Tirupati, Andhra Pradesh, India

**Abstract**-Cloud computing has become a popular approach to manage personal data for the economic savings and management flexibility in recent year. However, the sensitive data must be encrypted before outsourcing to cloud servers for the consideration of privacy, which makes some traditional data utilization functions, such as the plaintext keyword search, impossible. To solve this problem, we present multi-keyword ranked search scheme over encrypted cloud data supporting dynamic operations efficiently. Our scheme utilizes the vector space model combined with TF IDF rule and cosine similarity measure to achieve a multi-keyword ranked search. However, traditional solutions have to suffer high computational costs. In order to achieve the sub-linear search time, our scheme introduces Bloom filter to build a search index tree. What is more, our scheme can support dynamic operation properly and effectively on the account of the property of the Bloom filter, which means that the updating cost of our scheme is lower than other schemes. We present our basic scheme first, which is secure under the known cipher text model. Then, the enhanced scheme is presented later to guarantee security even under the known background model. The experiments on the real-world data set show that the performances of our proposed schemes are satisfactory.

**Keywords**— Bloom filter, Cloud computing, IDF rule, Scheme, Vector space.

## I. INTRODUCTION

Cloud computing has become a popular approach to manage personal data for the economic savings and management flexibility in recent year. However, the sensitive data must be encrypted before outsourcing to cloud servers for the consideration of privacy, which makes some traditional data utilization functions, such as the plaintext keyword search, impossible. To solve this problem, we present multi-keyword ranked search scheme over encrypted cloud data supporting dynamic operations efficiently. Our scheme utilizes the vector space model combined with TF IDF rule and cosine similarity measure to achieve a multi-keyword ranked search. However, traditional solutions have to suffer high computational costs. In order to achieve the sub-linear search time, our scheme introduces Bloom filter to build a search index tree. What is more, our scheme can support dynamic operation properly and effectively on the account of the property of the Bloom filter, which means that the updating cost of our scheme is lower than other schemes. We present our basic scheme first, which is secure under the known cipher text model. Then, the enhanced scheme is presented later to guarantee security even under the known background model.The experiments on the real-world data set show that the performances of our proposed schemes are satisfactory.

## II. RELATEDWORKS

The concept of identity-based encryption was introduced by Shamir , and conveniently instantiated by Boneh and Franklin . IBE eliminates the need for providing a public key infrastructure (PKI). Regardless of the setting of IBE or PKI, there must be an approach to revoke users from the system when necessary, e.g., the authority of some user is expired or the secret key of some user is disclosed. In the traditional PKI setting, the problem of revocation has been well studied and several techniques are widely approved, such as certificate revocation list or appending validity periods to certificates. However, there are only a few studies on revocation in the setting of IBE. Boneh and Franklin first proposed a natural revocation way for IBE. They appended the current time period to the ciphertext, and non-revoked users periodically received private keys for each time period from the key authority. Unfortunately, such a solution is not scalable, since it requires the key authority to perform linear work in the number of non-revoked users. In addition, a secure channel is essential for the key

authority and non-revoked users to transmit new keys. To conquer this problem, Boldyreva, Goyal and Kumar introduced a novel approach to achieve efficient revocation. They used a binary tree to manage identity such that their RIBE scheme reduces the complexity of key revocation to logarithmic (instead of linear) in the maximum number of system users. However, this scheme only achievesselective security. Subsequently, by using the aforementioned revocation technique, Libert and Vergnaud proposed an adaptively secure RIBE scheme based on a variant of Water's IBE scheme , Chen et al. constructed a RIBE scheme from lattices. Recently, Seo and Emura proposed an efficient RIBE scheme resistant to a realistic threat called decryption key exposure, which means that the disclosure of decryption key for current time period has no effect on the security of decryption keys for other time periods. Inspired by the above work and, Liang et al. introduced a cloud-based revocable identity- based proxy re-encryption that supports user revocation and ciphertext update. To reduce the complexity of revocation, they utilized a broadcast encryption scheme to encrypt the ciphertext of the update key, which is independent of users, such that only non-revoked users can decrypt the update key. However, this kind of revocation method cannot resist the collusion of revoked users and malicious non-revoked users as malicious nonrevoked users can share the update key with those revoked users. Furthermore, to update the ciphertext, the key authority in their scheme needs to maintain a table for each user to produce the re-encryption key for each time period, which significantly increases the key authority's workload.

- The keyword-based search is such one widely used data operator in many database and information retrieval applications, and its traditional processing methods cannot be directly applied to encrypted data. Therefore, how to process such queries over encrypted data and at the same time guarantee data privacy becomes a hot research topic.
- Song et al. first defined the problem of searching on encrypted data and proposed a symmetric searchable encryption scheme with linear complexity.
- Most of these methods cannot meet the high search efficiency and the strong data security simultaneously, especially when applying them to big data encryption poses great scalability and efficiency challenges.

## III. PROPOSED SYSTEM ARCHITECTURE

In this paper, we focus on a special type of multi-keyword ranked search, namely the multikeyword top-*k* search, which has been a very popular database operator in many important applications, and only needs to return the *k* documents with the highest relevance scores.

we propose a group multi-keyword top-*k* search scheme (GMTS), which is based on partition and supports top-*k* similarity search over encrypted data.

we propose a random traversal algorithm (RTRA) to strengthen the data security, where the data owner builds a binary tree as searchable index and assigns a random switch to each node, so the data user can assign a random key to each query.

- Improving the efficiency and the security of multi-keyword top-*k* similarity search over encrypted data.
- Data user receives different results but with the same high level of query accuracies in the mean time.
- Experimental results show that our methods are more efficient and more secure than the state-of-the-art methods.

**MODULES:**

- *Data Owner*
- *Data User*
- Authority
- Cloud Server
- Enclave

**MODULES DESCRIPTION:**

***Data owner:***

Data owner holds the data and wants to outsource his data to the cloud. In particular, data owners only want to share their data with those who satisfy certain conditions (e.g., student, professors or principal). They will be offline once their data have been uploaded to the cloud.

***Data User:***

Data user wants to download and decrypt the encrypted data shared in the cloud. Those who are authorized can download the encrypted file and further decrypt it to access the plaintext.

***Authority:***

Authority is responsible for initializing system parameters and data user registration. Also, it handles the call request from the cloud in the first proposed construction.

***Cloud Server:***

Cloud provides convenient storage service for data owners and data users. Specifically, it stores the outsourced data from data users and handles the download requests sent by data users.

***Enclave:***

Enclave handles the call request from the cloud (used in the second system).

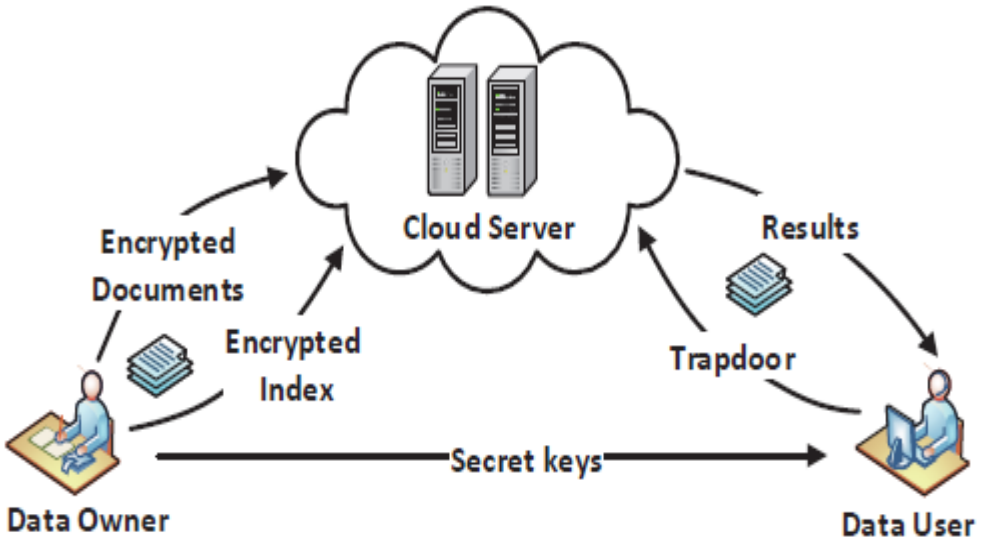


Fig.1 proposed system architecture

## IV. RESULTS AND DISCUSSION

The output screens obtained after running and executing the system are shown from Fig.2 to Fig. 10

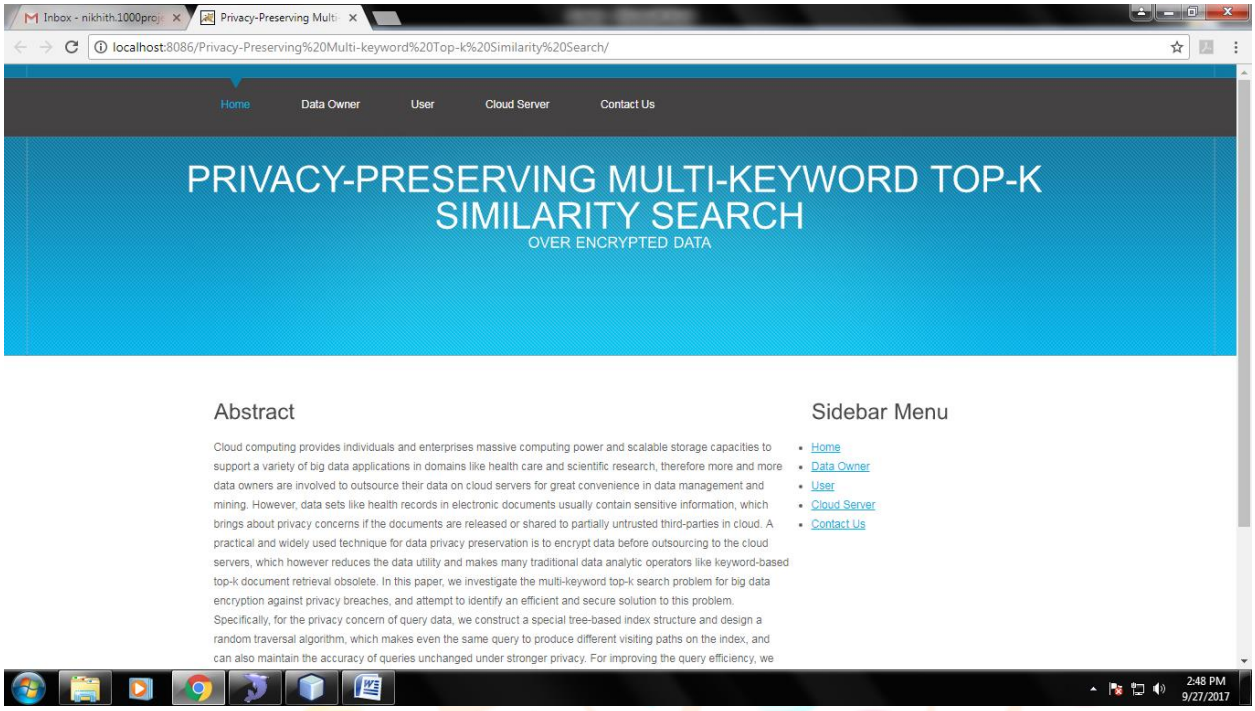


Fig.2 Home page

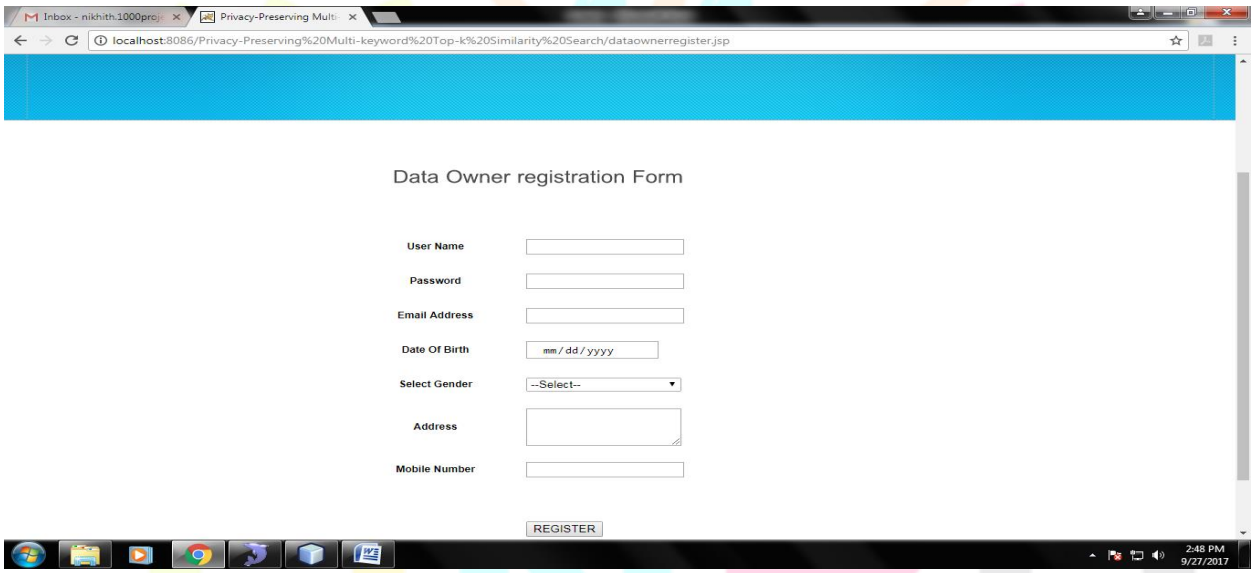


Fig:3 Data Owner Registration

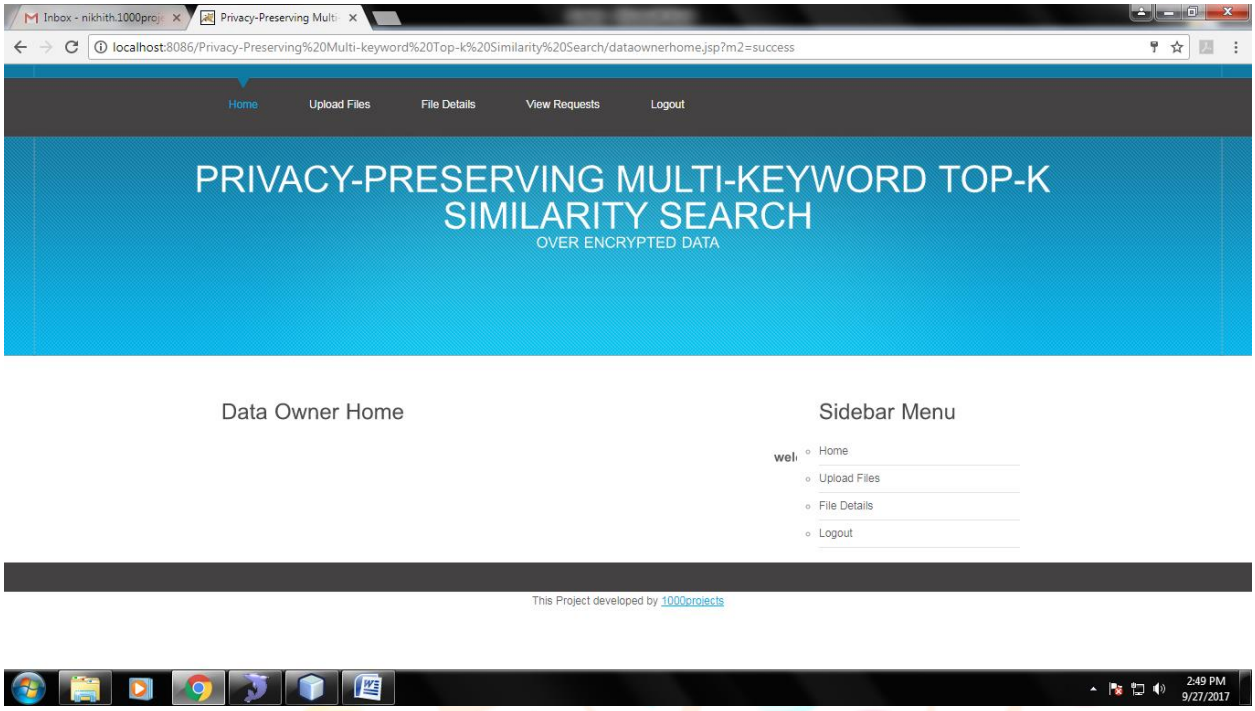


Fig: 4 Data Owner Home

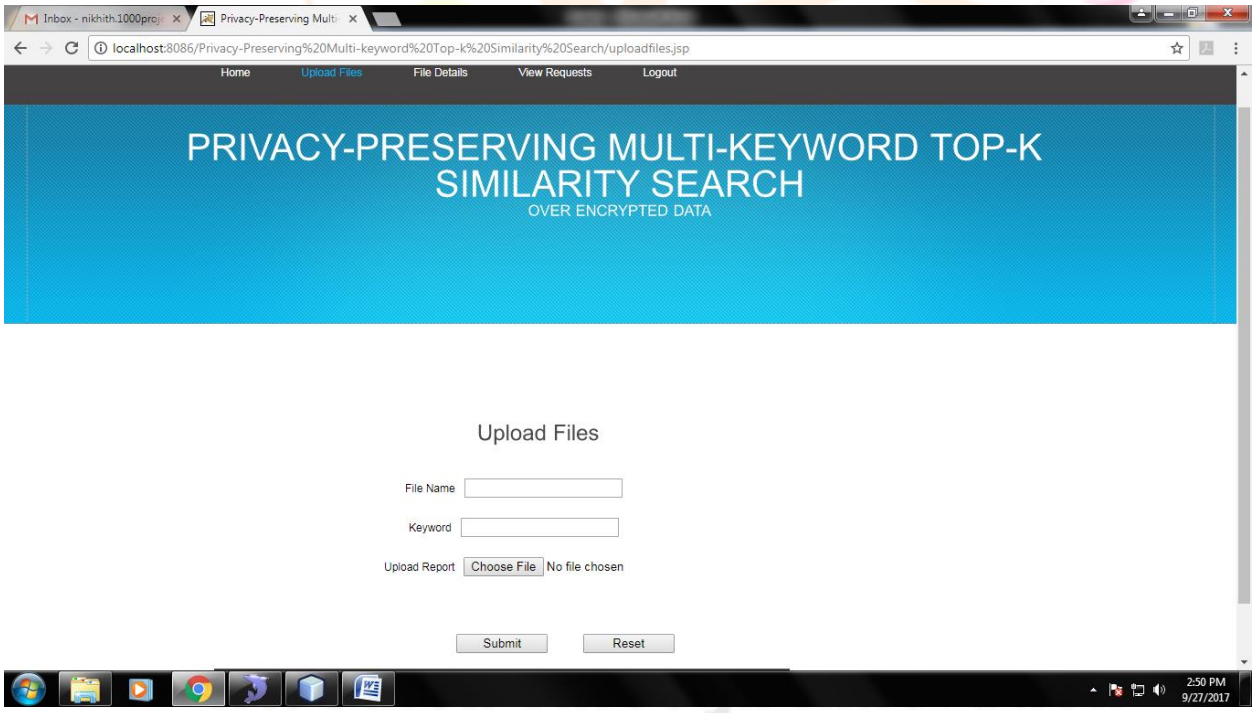


Fig:5 upload Files

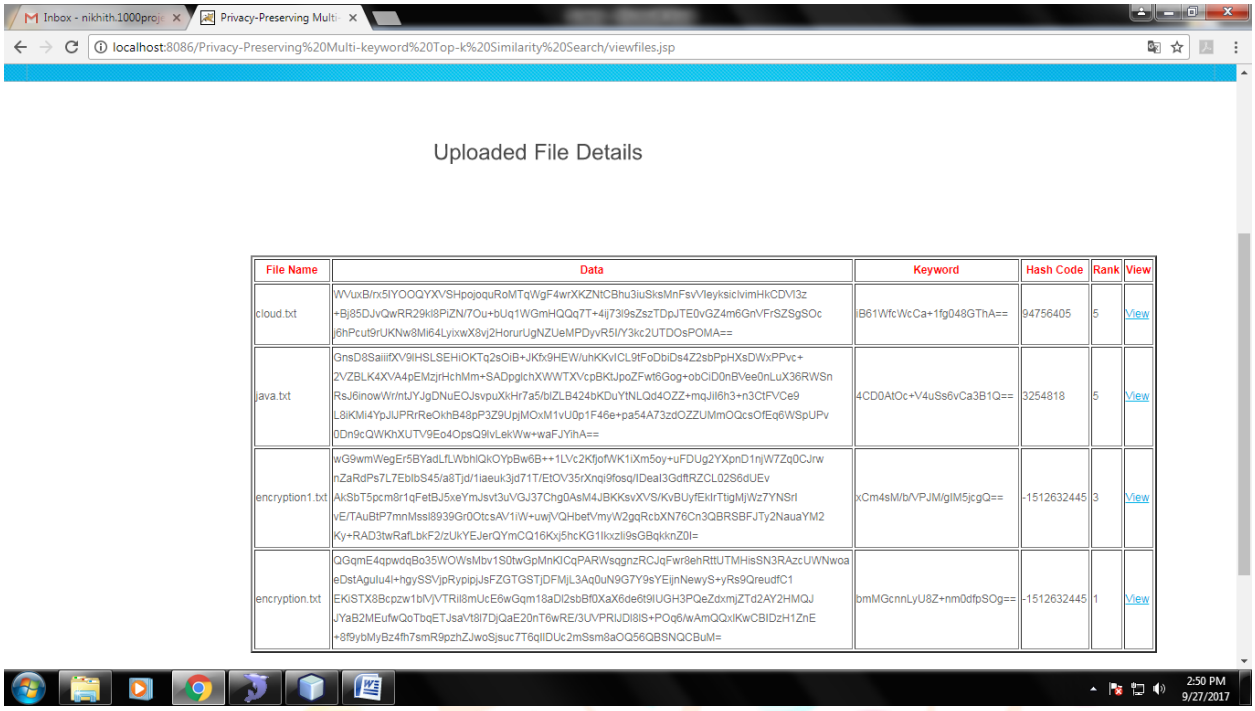


Fig: 6 View Uploaded File Details

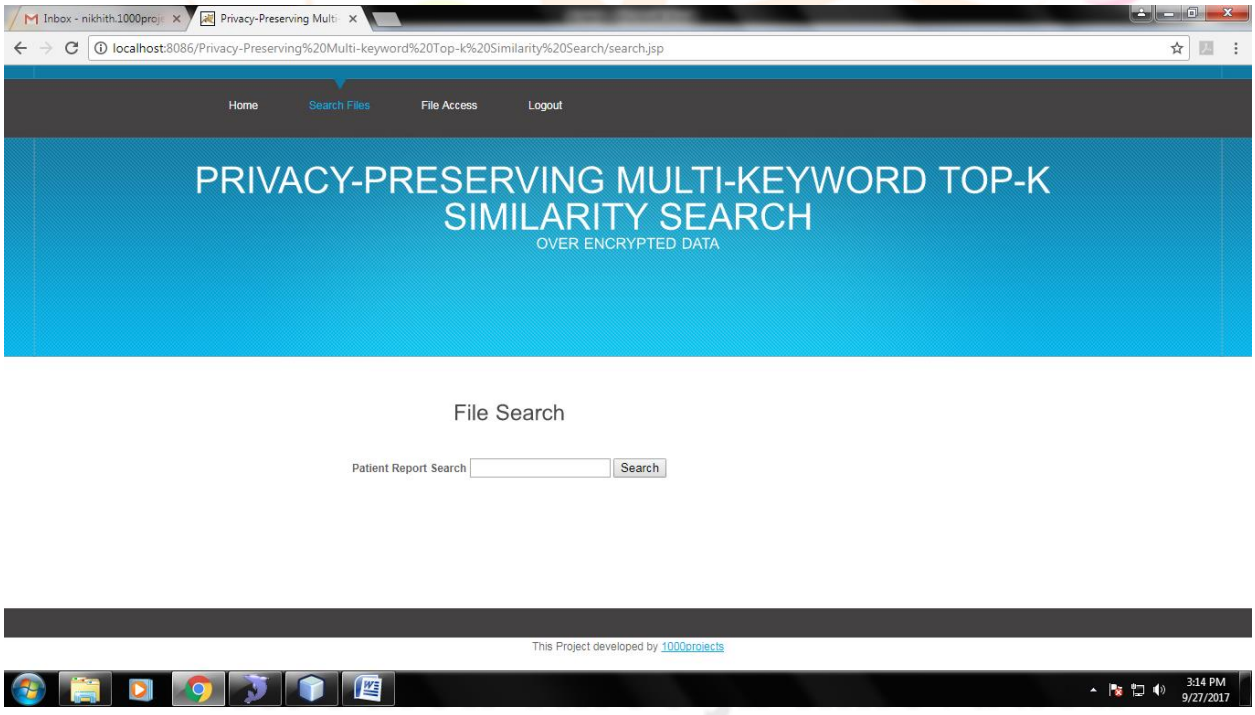


Fig:7 Search File

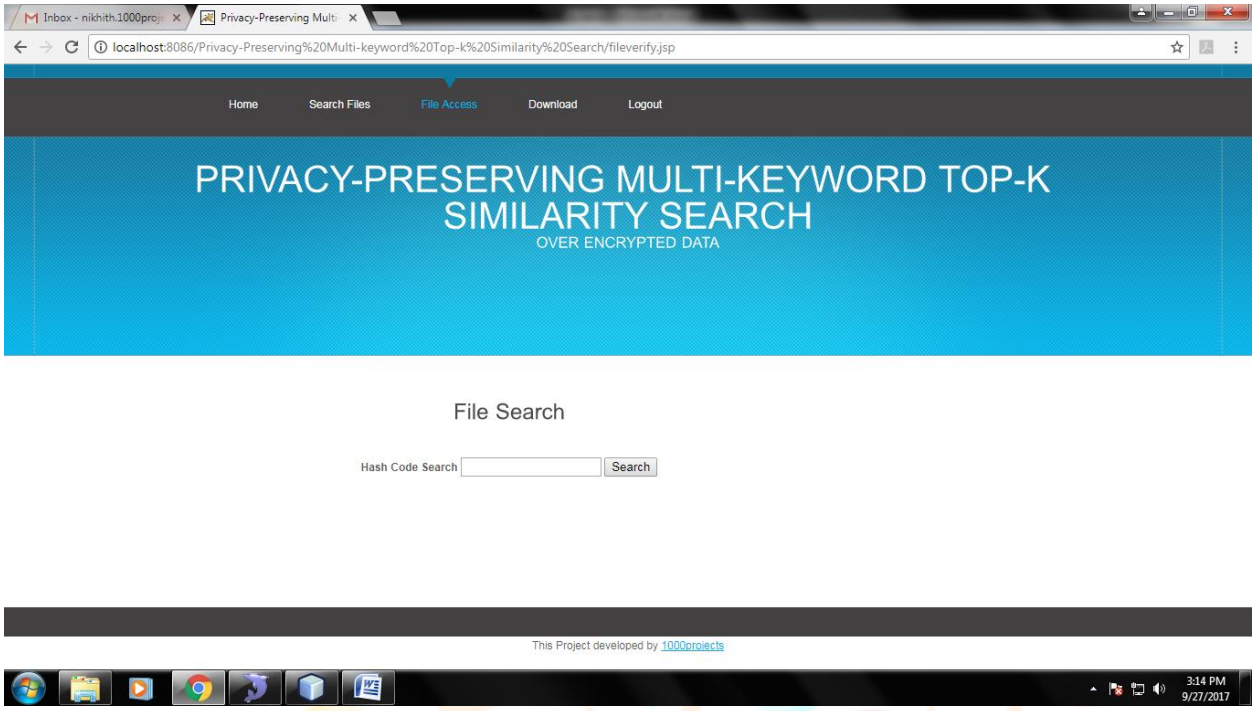


Fig:8 Hash Code Search

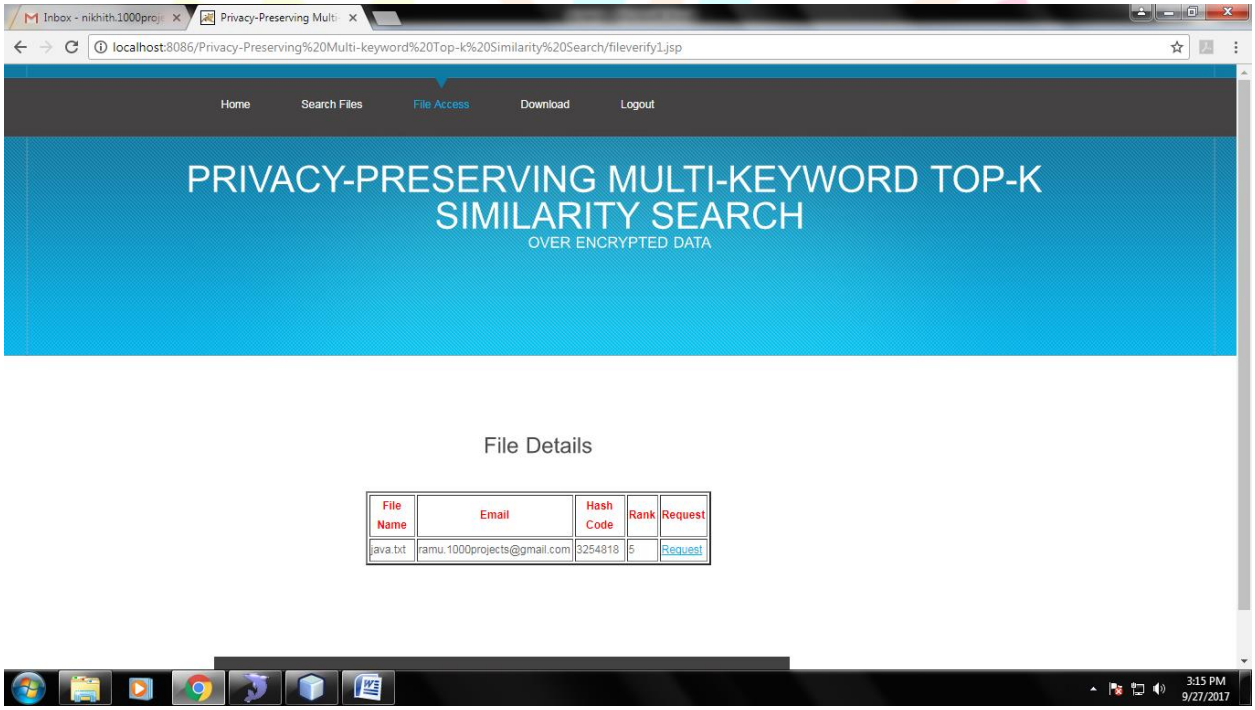


Fig: 9 View Hash Code Matched Files

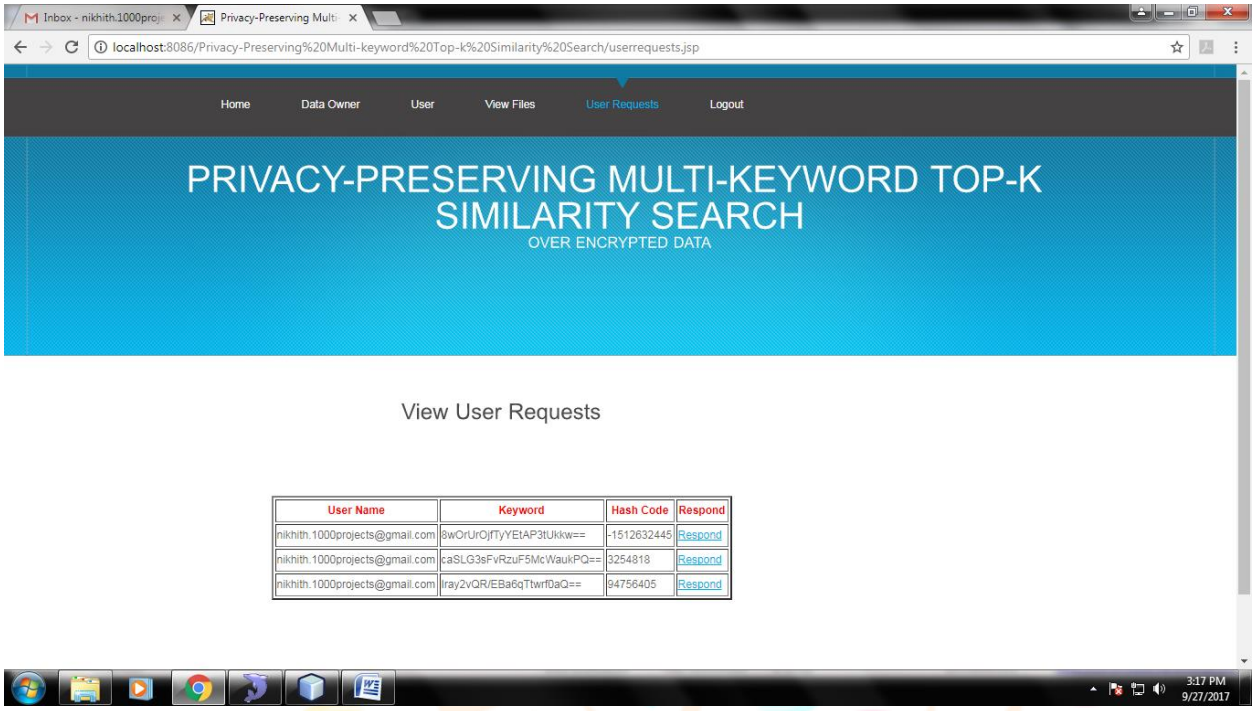


Fig:10 View User Search Requests

## V. FUTURE SCOPE AND CONCLUSION

In this paper, we focus on improving the efficiency and the security of multi-keyword top-*k* similarity search over encrypted data. At first, we propose the random traversal algorithm which can achieve that for two identical queries with different keys, the cloud server traverses different paths on the index, and the data user receives different results but with the same high level of query accuracies in the meantime. Then, in order to improve the search efficiency, we design the group multi-keyword top-*k* search scheme, which divides the dictionary into multiple groups and only needs to store the top-*ck* documents of each word group when building index. Next, to protect the query unlikability, we apply the random traversal algorithm to get the RGMTS, which can increase the difficulty of cloud servers to conduct linkage attacks on two identical queries, and we can also tune the value of *E* to make the level of query unlikability flexible for data owners. Finally, the experimental results show that our methods are more efficient and more secure than the state-of-the-art methods.

## REFERENCES

[1] J. Tang, Y. Cui, Q. Li, K. Ren, J. Liu, and R. Buyya, "Ensuring ecurity and privacy preservation for cloud data services," *ACMComputing Surveys*, 2016.
[2] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski,G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia,"A view of cloud computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.
[3] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: Improved definitions and efficient constructions," in *Proceedings of the 13th ACM Conference on Computer and Communications Security*. ACM, 2006, pp. 79–88.
[4] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in *Advances in Cryptology-Eurocrypt 2004*. Springer, 2004, pp. 506–522.
[5] Z. Ying, H. Li, J. Ma, J. Zhang, and J. Cui, "Adaptively secure ciphertext-policy attribute-based encryption with dynamic policyupdating," *Sci China Inf Sci*, vol. 59, no. 4, pp. 042 701:1–16, 2016.
[6] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Security and Privacy, 2000. SP 2000. Proceedings. 2000 IEEE Symposium on*, 2000, pp. 44–55.
[7] E.-J. Goh *et al.*, "Secure indexes." *IACR Cryptology ePrint Archive*, vol. 2003, p. 216, 2003.
[8] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in *Applied Cryptography and Network Security*. Springer, 2005, pp. 442–455.
[9] Y. H. Hwang and P. J. Lee, "Public key encryption with conjunctive keyword search and its extension to a multi-user system," in*Pairing-Based Cryptography–Pairing*. Springer, 2007, pp. 2–22.
[10] P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in *Applied Cryptography and Network Security*. Springer, 2004, pp. 31–45.
[11] L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword searches over encrypted data," in *Informationand Communications Security*. Springer, 2005, pp. 414–426.

[12] D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in *Theory of cryptography*. Springer, 2007, pp. 535–554.
[13] E. Shen, E. Shi, and B. Waters, "Predicate privacy in encryption systems," in *Theory of Cryptography*. Springer, 2009, pp. 457–473.
[14] W. Sun, B. Wang, N. Cao, M. Li, W. Lou, Y. T. Hou, and H. Li, "Privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking," in *Proceedings of the 8th ACM SIGSAC Symposium on Information*, ser. ASIA CCS '13. ACM,2013, pp. 71–82.
[15] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 1, pp. 222–233, 2014.
[16] Z. Xia, X. Wang, X. Sun, and Q. Wang, "A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 2, pp. 340–352, 2016.
[17] C. D. Manning, P. Raghavan, H. Sch ¨ utze*et al.*, *Introduction to information retrieval*. Cambridge university press Cambridge,2008, vol. 1, no. 1.
[18] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer Networks and ISDN Systems*, vol. 30, no. 17, 1998.
[19] J. Baek, R. Safavi-Naini, and W. Susilo, "Public key encryption with keyword search revisited," in *Computational Science and ItsApplications*. Springer, 2008, pp. 1249–1259.
[20] D. J. Park, K. Kim, and P. J. Lee, "Public key encryption with conjunctive field keyword search," in *Information security applications*. Springer, 2004, pp. 73–86.