INTERNATIONAL JOURNAL OF RESEARCH
In Advanced Engineering Technologies

Volume 6, Issue 1 JAN 2017

# Efficient and Dynamic Multi-Keyword Query Search over Secure Encrypted Cloud Environments

**[1]K.TEJASWI, [2]MD.AHMED**

[1]M.Tech Student, Department of CSE, Sri Vasavi Institute of Engineering & Technology, Village Nandamuru,Mandal Pedana, District Krishna, Andhra Pradesh, India.

[2]Assistant Professor, Department of CSE, Sri Vasavi Institute of Engineering & Technology, Village Nandamuru, Mandal Pedana, District Krishna, Andhra Pradesh, India.

***Abstract— Due to the growing popularity of cloud computing, more and more data owners are motivated to outsource their data to cloud servers for great convenience and reduced price in data management. However, sensitive data should be encrypted before outsourcing for privacy requirements that obsoletes data utilization like keyword-based document retrieval. During this paper, we present a secure multi-keyword ranked search scheme over encrypted cloud data, which at the same time supports dynamic update operations like deletion and insertion of documents. Specifically, the vector space model and also the widely-used TF_IDF model are combined within the index construction and query generation. We construct a special tree-based index structure and recommend a "Greedy Depth-first Search" rule to supply efficient multi-keyword ranked search. The secure KNN rule is utilized to encrypt the index and query vectors, and meanwhile ensure accurate relevance score calculation between encoded index and query vectors. So as to resist statistical attacks, phantom terms are added to the index vector for blinding search outcome. Because of the utilization of our special tree-based index structure, the projected scheme can achieve sub linear search time and deal with the deletion and insertion of documents flexibly. Intensive experiments are conducted to demonstrate the efficiency of the projected scheme.***

***Index Terms*** **-- Secure Multi-Keyword Search, Cloud Service Provider (CSP), Cloud Storage;**

## I. INTRODUCTION

Cloud Computing, a vital pattern for advanced information services, must outsource a necessary feasibility for data users data. Controversy on privacy, however were presented as outsourcing of sensitive data, as well as e-mail, medical records and private photos endlessly expands explosively. Reports of data loss and data breaches in cloud computing systems from time to time appear. The most vital threat to privacy roots once users source their personal data to the cloud within the cloud itself. The figure one shows the cloud service suppliers capable of the data and thus the communication between the users and therefore the cloud will, lawful or unlawful to manage and monitor. To make sure privacy, cipher users usually the data before it brings to cloud out-sourcing, the main challenges for effective data use one in all the most common ways that to do this could be through keyword-based retrieval. Keyword-based retrieval is also a typical data service and wide utilized in the text scenarios applied, where the users supported keywords retrieve relevant files during a file record. However, it turns out to be a tough task in cipher text scenario, attributable to the restricted operations on encoded data besides to enhance feasibility and save on costs at intervals the cloud paradigm, it is desirable to the question result to obtained with the foremost necessary files in place of all the files that got to purpose the interests of users that the files are elect thus as of relevance by users' corresponding interest and alone the files with the most effective relevancy are returned for users. To date, economical multi-keyword search on encoded data remains a tough drawback. It suggests that efforts embrace the search on encoded data not only data retrieval techniques like advanced data structures wont to represent the searchable index, and economical search algorithms, that lead through the corresponding system, but additionally the proper design of security protocols to create positive the security and privacy of the complete system. The blurring the keyword is detected by an innovative system and recursive vogue, without increasing the index so a high efficiency in terms the calculation and storage.

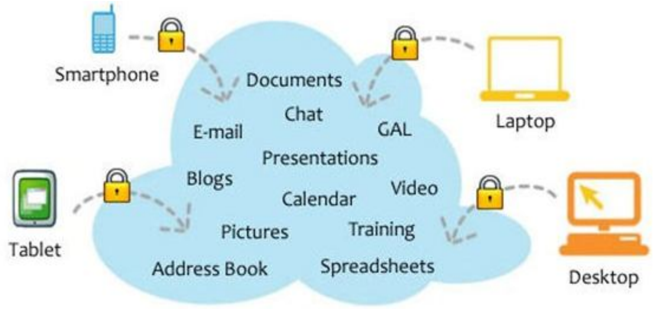


**Figure 1: Architecture of Cloud Storage**

A general approach to protect the confidentiality of data is to inscribe the data before outsourcing. However, this is often a massive value in terms of data, the user expertise lead. As an example, these techniques for keyword-based data retrieval that are sometimes used on the plaintext data cannot directly access the encoded applied data. Transfer all data among the cloud and regionally to decipher, is clearly impractical. to resolve the higher than drawbacks, researchers have some general purpose solutions with completely homomorphism cryptography or blind Rams created .These ways in which are impractical due to their high procedure issue for each the cloud Sever and users. Versatile search sub linearly accomplish by planned theme Search time and touch upon the deleting and inserting of documents. The secure KNN rule is employed to encipher the index and query vector, and within the meanwhile precise which means score calculation between encoded to create certain, index and query vectors.

## II. RELATED WORK

Multi-keyword Boolean search permits the users to input Multiple query keywords to request suitable documents among these works, conjunctive keyword search schemes are only return the documents that contain all of the query keywords. Divisional keyword search schemes return all of the documents that contain a collection of the question keywords. Predicate search schemes are projected to support every conjunctive and disjunctive search of these multi keyword search schemes retrieve search results based totally on the existence of keywords that cannot offer acceptable result ranking practicality. Ranked search can modify fast search of the foremost relevant data. Causation back solely the top-k most relevant documents can effectively decrease network traffic. Some near the beginning works have completed the ranked search using order-preserving techniques; but they are designed just for single keyword search. Cao et al. complete the first privacy-preserving multi-keyword ranked search approach, within which documents and queries are painted as vectors of dictionary size. With the “coordinate-matching”, the documents are ranked according to the number of coordinated query keywords. However, Cao et al.'s theme does not contemplate the importance of the varied keywords, so isn't correct enough. To boot, the search potency of the theme is linear with the amount of document collection. Sun et al. given a secure multi-keyword search approach that supports similarity based ranking. The authors construct a searchable index tree supported vector house model and adopted cosine live along side TF×IDF to supply ranking results. Sun et. al.'s search rule achieves better-than-linear search efficiency but ends up in precision loss. O¨ rencik et al. planned a secure multi keyword search technique that used native sensitive hash (LSH) functions to cluster the similar documents. The LSH rule is appropriate for similar search but cannot offer precise ranking. In, Zhang et al. projected an approach to deal with secure multi keyword hierarchal search throughout a multi-owner model. Throughout this approach, totally different data owners use whole different secret keys to code their documents and keywords whereas authorized data users can query whereas not knowing keys of those dissimilar data owners. The authors projected an "Additive Order preserving Function" to retrieve the most important search results. However, these works don't support dynamic operations.

## III. FRAME WORK

This paper proposes a secure tree based search process over the encoded cloud data that supports multi keyword ranked search and dynamic procedure on the document collection. Specifically, the vector house model and therefore the widely-used "TF ×IDF" model area unit collective inside the index-construction and query-generation to produce multi keyword ranked search. Therefore on get high search efficiency, we have a tendency to create a tree based index approach and propose a Greedy Depth initial Search technique supported this index tree. The secure KNN rule is used to write in code the index and query vectors, and meantime guarantee correct relevancy score calculation between encoded index and query vectors. To resist completely different attacks in varied threat models, we create a two secure search schemes: the essential dynamic multi-keyword ranked search scheme among the best-known cipher-text model, and additionally the improved dynamic multi keyword ranked search scheme among the notable background model. By using this model the following benefits area unit thanks to the special construction of our tree based mostly index, the projected search scheme can flexibly reach sub linear search time and trot out the insertion and of deletion file. We design a searchable cryptography scheme that supports each the right multi keyword ranked search and flexible dynamic operation on file collection. Because of the special technique of our tree based index, the search complexity of the projected approach is basically reserved to logarithmic. And in practice, the projected scheme can achieve higher search efficiency by execution our Greedy Depth initial Search technique. Moreover, similar search can be flexibly performed to any reduce the time and price of search technique.

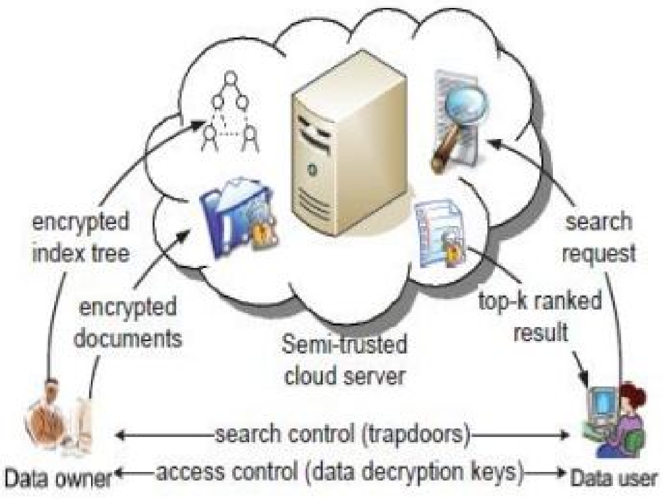


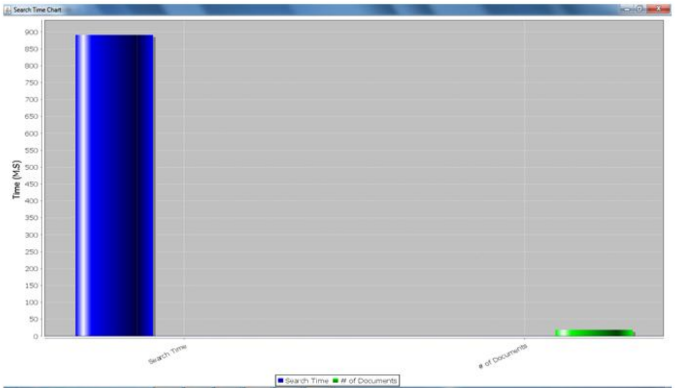
**Figure 2: Architecture of *Dynamic Multi Keyword Ranking Search***

Data Owner has a collection of documents that he needs to outsource to the cloud server in encrypted form whereas still keeping the power to look on them for effective utilization. Afterwards, the data owner outsources the encrypted assortment C and additionally the secure index I to the cloud server, and securely distributes the key data of trapdoor generation (including keyword inverse document frequency values) and document coding to the authorized data users. Besides, the data owner is liable for the update operation of his documents keep among the cloud server. Whereas change, the data owner generates the update data locally and sends it to the server data users are authorized ones to access the file of data owner with t query keywords; the authorized user will generate a trapdoor in line with search management mechanisms to fetch k encoded documents from cloud server. Then, the data user can decrypt the documents with the shared secret key. Cloud server stores the encoded document collection C and therefore the encoded searchable tree index I for data owner. Upon receiving the trapdoor from the info user, the cloud server executes search over the index tree I, and finally returns the corresponding collection of high k hierarchal encoded documents. Besides, upon receiving the update data from the data owner, the server wants to update the index I and document assortment C in line with the received data.

## IV. EXPERIMENTAL RESULTS

In our experiments, any number of data owners and data users can registered and login into the system. Who are authorized data owners they can upload the files into the cloud. Those uploaded files are stored in encryption format in cloud and the trapdoor file will be created (indexing) in index.txt file for those uploaded files trapdoor will be generated for data owners' uploaded files and those files are stored in encryption format in cloud after that who are authorized data users login into the system and search the keyword for data owners uploaded files after searching the keyword the searching particular keyword files will be generate those files the data users can download the file will be download in decryption format in your system after that authorized data owner is can also delete the uploaded files in cloud .

In the below chart we can observe that difference between the length of both search time and number of documents.


Search Time Chart

We can observe that search time length is higher than documents length. The difference will be shown in the sense of time in milliseconds (M.S). So we can consider that the advantage dynamic multi keyword ranked search scheme.

Through our implementation the authorized data owner is upload files and those upload files stored in encryption format in cloud after that authorized data users can search the keyword query the particular searching keyword query files will be generate those generate keyword query files download the files in decryption format. The authorized data owner is upload the files as well as deleted the files in cloud based on that we can search the keyword in encrypted file in secure way and also reduced the searching time and cost.

## V.CONCLUSION

In this paper we describe and solve the drawback of multi key word ranked search over encrypted cloud data, and started a range of privacy requirements. Among various multi-keyword semantics, we choose the efficient similarity measure of "coordinate matching," i.e., as several equivalent as potential, to effectively capture the connation of outsourced documents to the question Keywords, and utilize "inner product similarity" to quantitatively calculate such comparison measure. So as to acquire the check of supporting multi-keyword semantic while not privacy violation, we provide a basic plan of MRSE using secure inner product calculation. Then, we provide two improved MRSE schemes to realize numerous severe privacy desires in 2 different threat models. The any enhancements of our ranked search technique, also as supporting additional search semantics, i.e., TF x IDF, and dynamic data technique elaborate analyses in investigating privacy and efficiency assurance of projected schemes are mentioned, and testing on the real-world data set demonstrate our projected schemes that introduces low transparency on both calculation and communication.

## REFERENCES

[1] 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.

[2] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in INFOCOM, 2010 Proceedings IEEE. IEEE, 2010, pp. 1–5.

[3] M. Kuzu, M. S. Islam, and M. Kantarcioglu, "Efficient similarity search over encrypted data," in Data Engineering (ICDE), 2012 IEEE 28th International Conference on. IEEE, 2012, pp. 1156–1167.

[4] Wenhai Sun et al., "Protecting Your Right: Attributebased Keyword Search with Finegrained Ownerenforced Search Authorization in the Cloud", IEEE INFOCOM 2014, Toronto, Canada, April 27 - May 2, 2014.

[5] Secure Ranked Keyword Search over Encrypted Cloud Data, IEEE PAPER, 2010.

[6] Zhihua Xia, "A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data", and IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL: PP NO: 99 YEAR 2015.

[7] Cong Wang et al.,"Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data", IEEE Transactions on parallel and distributed systems, vol. 23, no. 8, August 2012.

[8] K. Ren, C.Wang, Q.Wang et al., "Security challenges for the public cloud," IEEE Internet Computing, vol. 16, no. 1, pp. 69–73, 2012.

[9] S. Kamara and K. Lauter, "Cryptographic cloud storage," in Financial Cryptography and Data Security. Springer, 2010, pp. 136–149.

[10] C. Gentry, "A fully homomorphic encryption scheme," Ph.D. dissertation, Stanford University, 2009.

[11] O. Goldreich and R. Ostrovsky, "Software protection and simulation on oblivious rams," Journal of the ACM (JACM), vol. 43, no. 3, pp. 431–473, 1996.

[12] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Advances in CryptologyEurocrypt 2004. Springer, 2004, pp. 506–522.

[13] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, "Public key encryption that allows pir queries," in Advances in Cryptology-CRYPTO 2007. Springer, 2007, pp. 50–67.

[14] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on. IEEE, 2000, pp. 44– 55.

[15] E.-J. Goh et al., "Secure indexes." IACR Cryptology e Print Archive, vol. 2003, p. 216, 2003.