Dynamic Searchable over Encrypted Cloud Data for Multi keyword Ranked Search Scheme

P.SUKANYA

Abstract— **Due to the increasing recognition of cloud computing, more and more data owners are encouraged to outsource their data to cloud servers for convenience and reduced cost in information control. however, sensitive information have to be encrypted before outsourcing for privacy requirements, which obsoletes records utilization like keyword-primarily based document retrieval.**

In this paper, we introduced a comfortable multi-keyword ranked search scheme over encrypted cloud data, which concurrently supports dynamic operations like insertion and deletion of documents. Specifically, the vector area model and the widely-used TF×IDF model are blended in

the index creation. We construct a tree-based index structure and advise a "Greedy Depth-first Search" algorithm to offer efficient multi-keyword ranked search. Due to the use of our special tree- based index structure, the proposed scheme can achieve sub-linear search time and deal with the deletion and insertion of documents flexibly.

***Index Terms*— Multi-keyword ranked search over encrypted cloud data, OTP, Product resemblance, Cloud, Data owners**

1. INTRODUCTION

Cloud computing has been considered as a new version of company IT infrastructure, that may prepare large useful resource of computing, storage and applications, and allow users to experience ubiquitous, convenient and on-demand community get admission to a shared pool of configurable computing sources with incredible performance and minimum financial overhead . Attracted by these attractive features, each individuals and firms are inspired to outsource their data to the cloud, instead of buying software and hardware to manage the data themselves. Regardless of the various advantages of cloud services, outsourcing sensitive information (such as e-mails, personal health statistics, business enterprise finance records, authorities files , and many others) to remote servers brings privacy concerns. The cloud service providers (CSPs) that hold the information for users may additionally get admission to users' sensitive data with out authorization. A fashionable method to shield the records confidentiality is to encrypt the data earlier than outsourcing but, this could cause a massive fee in terms of information usability. For an example, the existing techniques on keyword-based data retrieval, which are broadly used on the plaintext records, can not be directly implemented on the encrypted data. Downloading all the data from the cloud and decrypt regionally is impractical. On the contrary, more practical special purpose solutions, such as searchable encryption (SE)schemes have made specific contributions in terms of performance, capability and safety. Searchable encryption schemes permit the client to store the encrypted data to the cloud and execute keyword search over ciphertext domain. To this point, abundant works had been proposed under different threat models to obtain various search capability, which include single key-word search, similarity search, multi-keyword boolean search, ranked search, multi-keyword ranked search, and so on. Among them, multi-keyword ranked search achieves increasingly more attention for its realistic applicability. These days, some dynamic schemes have been proposed to help inserting and deleting operations on document collection. These are significant works as it is highly possible that the data proprietors need to update their records at the cloud server. However few of the dynamic schemes support efficient multi-key-word ranked search. This paper proposes a secure tree-based scheme over the encrypted cloud data, which helps multi-keyword ranked search and dynamic operation on the document collection. Mainly, the vector space version

and the extensively-used ―time frequency (TF) × inverse document frequency (IDF)‖ model are blended within the index production and query technology to provide multi-keyword ranked search. In order to obtain high search efficiency, we assemble a tree-based index structure and propose a ―Greedy Depth-first Search‖ algorithm primarily based on this index tree. Due to the special structure of our tree-based index, the proposed search scheme can flexibly acquire sub-linear search time and deal with the deletion and insertion of files. we construct secure search schemes: the basic dynamic multi-keyword ranked search (BDMRS) scheme in the known ciphertext model, and the enhanced dynamic multi-key-word ranked search(EDMRS) scheme within the recognized background model. Our contributions are summarized as follows:

1) We layout a searchable encryption scheme that helps both the accurate multi-keyword ranked search and bendy dynamic operation on document collection.

2) The proposed scheme can achieve higher search performance through executing our ―Greedy Depth-first search‖ algorithm. Moreover, parallel search can be flexibly carried out to further reduce the time cost of search method.

2. LITERATURE SURVEY & RELATED WORK

1. Secure and privacy preserving keyword search:

Qin Liu [3] proposed in this paper the pursuit that gives catchphrase protection, information protection and semantic secure by open key encryption. CSP is included in fractional decipherment by lessening the correspondence and computational flying in decoding process for end clients. The client presents the watchword trapdoor encoded by user's private key to

CS (Cloud Server) safely and recovers the scrambled records.

Limitations: The correspondence and computational expense for encryption and decoding is more

2. Secure and Efficient Ranked Keyword Search:

Cong Wang [4] proposed seek which unravels preparing overhead, information and catchphrase security, least correspondence and calculation aeronautical. The information proprietor assemble list alongside the catchphrase recurrence based importance scores for documents. Client ask for „w– to cloud server with discretionary k– as Tw utilizing the private key. The cloud server looks the list with scores and sends scrambled document in light of positioned grouping.

Limitations: It doesn't play out numerous catch phrase seeks. Minimal overhead in record building

3. Single Keyword Search over Encrypted data on cloud:

Reachable searchable encryption plan agree to a client to solidly search for over scrambled information through watchwords without first applying decoding on it, the proposed systems bolster just traditional Boolean catchphrase look, without catching any relevance of the records in the query item. At the point when straightforwardly connected in substantial joint information outsourcing cloud environment, they experience next deficiency.

Limitations: Single-watchword hunt without positioning. Boolean-catchphrase look without positioning. Try not to get applicable information.

Privacy-preservingMulti-keyword Text Search:

Wenhai Sun [6] proposed this inquiry that gives comparability based item positioning, catchphrase security, Index and Query classification and Query Unlink capacity. The scrambled record is worked by vector space model supporting solidified and particular document look. The searchable file is fabricated utilizing Multidimensional B tree. Proprietor makes scrambled question vector ® for document catch phrase set. Client gets the individual encoded inquiry vector of W from proprietor which is given to CS. Presently CS looks list by Merkle–Damgård development calculation and thinks about cosine measure of document and question vector and returns top k scrambled records to client.

Limitations: The likeness rank score of the record vector completely relies on upon the kind of the report

Secure Multi-catchphrase Top-k Retrieval Search:

Jiadi [7] proposed this pursuit utilizing Two round searchable encryption (TRSE). In first round, clients presents various catchphrase "REQ" 'W– as encoded question for finishing information, watchword protection and make trapdoor (REQ, PK) as Tw and sends to cloud server. At that point cloud server figure the score from encoded file for documents and returns the scrambled score result vector to client. In second round, client decode N with mystery key and ascertains the document positioning and afterward ask for records with Top k scores. The positioning of record is done on customer side and scoring is done on server side.

Limitations: The withdrawal and restricting is utilized to diminish figure content size, still the key size is too extensive. The correspondence elevated will be high, if the scrambled trapdoor's size is too vast. It doesn't make powerful searchable file redesign.

Privacy Preserving Multi-Keyword Ranked Search (MRSE):

Ning [8] proposed this quest for known figure content model and foundation model over encoded information giving low calculation and correspondence overhead. The direction coordinating is decided for multi- watchword seek. They utilized internal item likeness to quantitatively assess similitude for positioning documents. The downside is that MRSE have little standard deviation σ which debilitates catchphrase protection.

Limitations: Multi-watchword positioned look (MRSE) for known figure content model may deliver two diverse trapdoor which dubious the protection spillage issue of trapdoor unlink capacity which may debilitate the catchphrase security. MRSE has little standard deviation σ which thusly debilitates the watchword protection. The honesty of the rank request is not checked in MRSE.

Attribute-based Keyword Search:

Wenhai Sun [9] proposed Attribute-based Keyword Search that gives conjunctive catchphrase look; watchword semantic security and Trapdoor unlink capacity. The proprietors makes file with all catchphrases and access list with strategy characteristics which determines the clients list approved for seeking. Presently proprietors scramble the archive, record with access list utilizing ciphertext strategy characteristic based encryption procedure. To have client enrollment administration, they utilized intermediary re-encryption and languid re-encryption strategies to share the workload to CS. The client asks for the Tw to CS utilizing its private key. Presently CS recovers Tw and ventures the encoded files and return records just if the user's properties in Tw fulfills access approaches in files which makes coarse-grained dataset look approval.

Limitations: Trapdoor era will require additional time with the expanded number of traits.

Efficient and Secure Multi-Keyword Search on Encrypted Cloud Data:

This proposed technique has characterized and tackled the issue of successful yet protected and sound rank watchword seek over Encrypted cloud information [10]. Positioned look enormously upgrades framework ease of use by giving back the coordinating records in a positioned request with respect to certain vital criteria (e.g. watchword recurrence) accordingly making one stage nearer towards sensible utilization of secure information facilitating administrations in Cloud Computing. These papers has characterized and tackled the testing issue of security saving and productive multi watchword positioned look over scrambled cloud information stockpiling (MRSE), and set up an

arrangement of strict protection prerequisites for such an ensured cloud information use framework to wind up a reality. The proposed positioning strategy ends up being effective to do a reversal to a great degree applicable reports comparing to submitted seek terms. Proposed positioning strategy is utilized as a part of our future framework keeping in mind the end goal to improve the security of data on Cloud Service Provider.

Limitations: Dynamic redesigning and erasure of the record from the cloud is unrealistic.

A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data:

This proposed technique [11] recommend a protected tree-based hunt plan over the encoded distributed storage, which bolsters multi catchphrase positioned seek alongside element operation on record accumulation accessible at server. The vector space model and term recurrence (TF) × backwards report recurrence (IDF) model are comfily utilized as a part of the development of record and era of inquiry to give multi catchphrase positioned look yield. To get high pursuit proficiency results, creators build a tree-based list structure and proposed a Greedy Depth-first Search calculation in light of this list tree. As a result of this uncommon structure of tree-based list, the proposed seek plan can adaptably accomplish sub straight hunt time and can successfully manage the erasure and insertion of records. The kNN calculation is connected to scramble the file and inquiry vectors, and till then guarantee precise pertinence score estimation between encoded file and question vectors.

3 IMPLEMENTATION DETAILS

3.8 MRES System

For our living being, we pick the disposition of fit coordinating, to recognize the correspondence in the

midst of hunt request and information certifications. Especially, we utilize inner information correspondence, i.e., the figure of question watchwords showing up in an archive, to assess the comparability of that report to the hunt inquiry in direction coordinating rule. Every archive is associated with a twofold vector as a sub record where every piece speak to whether practically equivalent to catchphrase is contained in the

report [6] The hunt reservation is likewise portray as a double vector where every piece implies whether comparing watchword shows up in this inquiry demand, so the likeness could be precisely measured by inward result of question vector with data vector. Be that as it may, straightforwardly outsourcing information vector or question vector will encroach file protection or pursuit security.

To enhance report recovery precision, query output ought to be positioned by cloud server as indicated by some positioning criteria. Cloud server just sends back top-k archives that are most important to the inquiry question.

In the longed for creature the stream begins from the client. The client needs to enlist in CSP to get the enhancements. When client information is put away in CSS it has no unswerving control above it. Client needs to procure any evaluator called TPA who will consistently check the client information in CSS. The TPA ought to be conceded by the client to check the respectability for a particular information and for an unambiguous time without getting to the careful information. Underneath a calculation is give which portray how the TPA does the review.

AES calculation is utilized to store the information in scrambled structure in cloud server. So when TPA does the review it just gets the bogus impression of unique documents. The qualities on which TPA figures or check the respectability is really the hash estimation of encoded document computed agreeably. The TPA is just permitted to check the uprightness nothing else. It checks the trustworthiness interestingly then checks whether uprightness pruned and in conclusion check whether respectability remains or lost. Client can whenever concede or renounce the concession from TPA. Client has the benefit to transfer, download and alter information. Client's alter solicitation is likewise

served for a particular part of the record rather than recover the entire file [6].

1) **AES Algorithm**:

Notation and Definitions

AES(K, W) Encrypt W using the AES codebook with key K

AES-1(K, W) Decrypt W using the AES codebook with key

K MSB(j, W) Return the most significant j bits of W

LSB(j, W) Return the least significant j bits of W

B1 | B2 Concatenate B1 and B2 K The key-encryption key K

s The number of steps in the wrapping process, = 6n

P[i] The ith plaintext key data block

C[i] The ith cipher text data block

A The 64-bit integrity check register

R[i] An array of 64-bit registers where i = 0, 1, 2,..., n

Algorithm

1. The set of round keys from the cipher key.

2. Initialize state array and add the initial round key to the starting state array.

3. Perform round = 1 to 9: Execute Usual Round.

4. Execute Final Round.

5. Corresponding cipher text chunk output of Final Round Step

iii. Usual Round Execute the following operations which are described above.

1. Sub Bytes

2. Shift Rows

3. MixColumns

4. Add Round Key, using K(round)

iv. Final Round:
Execute the following operations which are described above.

1. Sub Bytes

2. Shift Rows

3. Add Round Key, using K(10)

In the key wrap algorithm, the concatenation occupation will be used to concatenate 64-bit quantity to form the 128-bit input to the AES codebook. The pulling out functions will be used to split the 128-bit amount produced from the AES codebook into two 64- bit quantities.

The arrangement of the key enfold algorithm requires the use of the AES codebook [AES]. The next sections will describe the key envelop algorithm, the key unwrap algorithm, and the inherent data integrity check.

Algorithm Key Wrap:

1 Sub Bytes: The first transformation, Sub Bytes, is used at the encryption site. To substitute a byte, we interpret the byte as two hexadecimal digits.

2. Shift Rows: In the encryption, the transformation is called Shift Rows.

3. Mix Columns: The Mix Columns transformation operates at the column level; it transforms each column of the state to a new column.

4. Add Round Key: Add Round Key proceeds one column at a time.

The inputs to the key wrapping procedure are the KEY and the plaintext to be wrapped. The plaintext consists of n 64-bit blocks, contain the key data life form wrapped. The key covering process is described beneath.

A substitute explanation of the key wrap algorithm involves indexing quite than shifting. This approach allows one to calculate the wrap key in place, avoiding the rotation in the previous account. This produces the same results and is more easily implement in software.

CONCLUSION

In this paper, efficient and dynamic search scheme is proposed, which supports not only the accurate multi-key-word ranked search but also the dynamic deletion and insertion of files. We construct a special keyword balanced binary tree as the index, and suggest a
–Greedy Depth-first search‖ algorithm to obtain better efficiency than linear search. In addition, the parallel search manner can be performed to further reduce the time cost. The data proprietor is responsible for producing updating records and sending them to the cloud server. Accordingly, the data owner desires to store the unencrypted index tree. For the effectiveness viewpoint, we propose a tree based file structure. The tree-based searchable list composed in our plan can bolster dynamic redesign well, just getting to a phase of the report tree. Our deliberate plan can give insertion and erasure remodel. We endorse a secure plan to fulfill safety conditions inside the risk model. we will accomplish more research later on.

REFERENCES

[1] L. M. Vaquero, L. Rodero-Merino, J. Caceres and M. Lindner, –A break in the clouds: towards a cloud definition‖, ACM SIGCOMM Computer Communication Review, vol. 39, no. 1, (2009), pp. 50– 55.

[2] S. Kamara and K. Lauter, –Cryptographic cloud storage‖, Financial Cryptography and Data Security, Springer Berlin Heidelberg Publishing, (2010).

[3] C. Wang, N. Cao, K. Ren and W. Lou, –Enabling secure and efficient ranked keyword search over outsourced cloud data‖, IEEE Transactions on Parallel and Distributed Systems, vol. 23, issue 8, (2012)

August, pp. 1467–1479.

[4] D. Song, D. Wagner and A. Perrig, ―Practical techniques for searches on encrypted data‖, In Proceedings of S&P, (2000) May 14-17, Berkeley, CA.

[5] E. J. Goh, ―Secure indexes‖, Cryptology ePrint Archive, (2003), http://eprint.iacr.org/2003/216.

[6] R. Curtmola, J. A. 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, (2006), pp. 79-88.

[7] D. Boneh, G. D. Crescenzo, R. Ostrovsky and G. Persiano, ―Public key encryption with keyword search‖, In Proceedings of EUROCRYPT, (2004) May 2-6, Interlaken, Switzerland.

[8] 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‖, In Proceedings of the 2007 ACM Workshop on Storage Security and Survivability, (2007), pp. 7–12.

[9]S. Zerr, D. Olmedilla, W. Nejdl and W. Siberski, ―Zerber+r: Top-k retrieval from a confidential index‖, In Proceedings of EDBT, (2009), pp. 439–449.

[10]N. Cao, C. Wang, M. Li, K. Ren and W. Lou, ―Privacy-preserving multi-keyword ranked search over encrypted cloud data‖, In Proceedings of IEEE INFOCOM, (2011) April 10-15, Shanghai, China.

[11]W. Sun, B. Wang, N. Cao, M. Li, W. Lou, YT. Hou and H. L, ―Privacy-preserving multi-keyword text search in the cloud supporting similarity-based ranking‖, In Proceedings of the 8th ACM SIGSAC symposium on Information, computer and communications security, (2013), pp. 71-82.

[12]S. Kamara and C. Papamanthou, ―Parallel and Dynamic Searchable Symmetric Encryption‖ Cryptology ePrint Archive,(2013), http://eprint.iacr.org/2013/335.

[13]I. H. Witten, A. Moffat and T. C. Bell, Managing gigabytes: Compressing and indexing documents and images, Morgan Kaufmann Publishing, San Francisco, (1999) May.

Author's Profile :

Ms. P. Sukanya , is received her M.Tech Degree from Shri Vishnu Engineering College For Women (SVECW), Bhimavaram, Affiliated to JNTU Kakinada in the year 2016.