ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

A Survey on Multi-Keyword Ranked Search Scheme over Encrypted Cloud Storage

Shaikh Saniya, Sherkar Shilpa, Kulkarni Aditi, Kulkarni Medha, Prof.S.P.Vidhate

VACOEA, A.Nagar, Maharashtra, India

ABSTRACT: Because of the expanding prominence of distributed computing, an ever increasing number of information proprietors are persuaded to outsource their information to cloud servers for extraordinary accommodation and lessened cost in information administration. In any case, delicate information ought to be encoded some time recently outsourcing for security necessities, which obsoletes information usage like catchphrase based archive recovery. In this paper, we introduce a protected multi-watchword positioned seek conspire over scrambled cloud information, which at the same time underpins dynamic refresh operations like erasure and addition of records. In particular, the vector space show and the generally utilized TF IDF demonstrate are joined in the record development and inquiry age. We develop an extraordinary tree-based file structure and propose an "Eager Depth-first Search" calculation to give productive multi-catchphrase positioned look. The safe KNN calculation is used to encode the record and inquiry vectors, what's more, in the interim guarantee exact pertinence score estimation between scrambled list and inquiry vectors. So as to oppose measurable assaults, apparition terms are added to the record vector for blinding query items. Because of the utilization of our extraordinary tree-based record structure, the proposed plan can accomplish sub-direct pursuit time and manage the cancellation and inclusion of archives adaptable. Broad tests are led to show the effectiveness of the proposed conspires.

I. INTRODUCTION

Distributed computing has been considered as another model of big business IT foundation, which can sort out gigantic asset of registering, stockpiling and applications, and empower clients to appreciate omnipresent, advantageous and on demand organize access to a common pool of configurable figuring assets with extraordinary proficiency and negligible monetary overhead [1]. Pulled in by these engaging highlights, the two people and endeavors are spurred to outsource their information to the cloud, rather than obtaining programming and equipment to deal with the information them.

Regardless of the different focal points of cloud administrations, outsourcing delicate data, (for example, messages, individual wellbeing records, organization back information, government reports, and so on.) to remote servers brings security concerns. The cloud specialist co-ops (CSPs) that keep the information for clients may get to clients' delicate data without approval. A general way to deal with secure the information privacy is to scramble the information before outsourcing [2]. Be that as it may, this will cause a tremendous cost as far as information ease of use. For instance, the current systems on catchphrase based data recovery, which are generally utilized on the plaintext information, can't be specifically connected on the scrambled information. Downloading every one of the information from the cloud and unscramble locally is clearly unfeasible.

Keeping in mind the end goal to address the above issue, specialists have outlined some broadly useful arrangements with completely homomorphism encryption [3] or neglectful RAMs [4]. Be that as it may, these techniques are not down to earth because of their high computational overhead for both the cloud disjoin and client. On the opposite, more useful exceptional reason arrangements, for example, accessible encryption (SE) plans have made particular commitments as far as effectiveness, usefulness and security. Accessible encryption plans empower the customer to store the scrambled information to the cloud and execute watchword look over ciphertext space. Up until this point, bounteous works have been proposed under various danger models to accomplish different look usefulness, for example, single watchword seek, likeness look, multi-watchword boolean pursuit, positioned seek, multi-watchword positioned look, and so forth. Among them, multi-catchphrase positioned look accomplishes increasingly consideration

Copyright to IJIRCCE
DOI: 10.15680/IJIRCCE.2018. 0601067

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

for its pragmatic appropriateness. As of late, some unique plans have been proposed to help embeddings and erasing operations on report gathering. These are noteworthy works as it is exceptionally conceivable that the information proprietors need to refresh their information on the cloud server. However, few of the dynamic plans bolster proficient multi-catchphrase positioned look.

II. LITERATURE REVIEW

1. **Enabling secure and efficient ranked keyword search over outsourced cloud data**
   Cloud computing economically enables the paradigm of data service outsourcing. However, to protect data privacy, sensitive cloud data has to be encrypted before outsourced to the commercial public cloud, which makes effective data utilization service a very challenging task. Although traditional searchable encryption techniques allow users to securely search over encrypted data through keywords, they support only Boolean search and are not yet sufficient to meet the effective data utilization need that is inherently demanded by large number of users and huge amount of data files in cloud. In this paper, we define and solve the problem of secure ranked keyword search over encrypted cloud data. Ranked search greatly enhances system usability by enabling search result relevance ranking instead of sending undifferentiated results, and further ensures the file retrieval accuracy. Specifically, we explore the statistical measure approach, i.e. relevance score, from information retrieval to build a secure searchable index, and develop a one-to-many order-preserving mapping technique to properly protect those sensitive score information. The resulting design is able to facilitate efficient server-side ranking without losing keyword privacy. Thorough analysis shows that our proposed solution enjoys "as-strong-as possible" security guarantee compared to previous searchable encryption schemes, while correctly realizing the goal of ranked keyword search. Extensive experimental results demonstrate the efficiency of the proposed solution. [1]

2. **An efficient public key encryption with conjunctive-subset keywords search**
   We consider the advancement of conjunctive catchphrase accessible plan which empowers one to look encoded archives by utilizing more than one watchword. The thought of conjunctive watchword seeking was displayed by Golle et al. in 2004. Be that as it may, their security display was built in a symmetric-key setting which isn't material for the general applications in the truth. So Park et al. broadened Golle et al's. security display into a publickey setting which calls the Public Key Encryption with Conjunctive Field Keyword Search (PECKS) plot. In this paper, we look at six security models by finishing up the mystery key setting and open key setting, and aggregate up six security prerequisites that must fulfill to develop a safe conjunctive catchphrase accessible plan. At that point we think about and examine the security and the execution of the security models. At long last, we show a few issues that need to additionally examine later on. [2]

3. **Privacy-preserving multi-keyword fuzzy search over encrypted data in the cloud**
   The appearance of distributed computing, information proprietors are roused to outsource their mind boggling information administration frameworks from neighborhood locales to business open cloud for incredible adaptability and financial funds. In any case, for securing information protection, touchy information must be scrambled before outsourcing, which obsoletes conventional information usage in view of plaintext watchword seek. In this manner, empowering an encoded cloud information look benefit is of central significance. Considering the vast number of information clients and reports in cloud, it is urgent for the hunt administration to permit multi-watchword inquiry and give result likeness positioning to meet the viable information recovery require. Related chips away at accessible encryption concentrate on single catchphrase look or Boolean watchword seek, and once in a while separate the query items. In this paper, out of the blue, we characterize and take care of the testing issue of protection saving multi-catchphrase positioned seek over encoded cloud information (MRSE), and build up an arrangement of strict protection prerequisites for such a safe cloud information use framework to end up plainly a reality. Among different multi-watchword semantics, we pick the effective standard of "facilitate coordinating", i.e., whatever number matches as could

Copyright to IJIRCCE
DOI: 10.15680/IJIRCCE.2018. 0601067

be expected under the circumstances, to catch the similitude between look question and information reports, and further utilize "internal item closeness" to quantitatively formalize such rule for comparability estimation. We initially propose an essential MRSE plot utilizing secure inward item calculation, and afterward altogether enhance it to meet distinctive protection necessities in two levels of danger models. Intensive examination researching security and proficiency certifications of proposed plans is given, and investigations on this present reality dataset additionally indicate proposed plots in fact present low overhead on calculation and correspondence. [3]

4. **Achieving usable and privacy-assured similarity search over outsourced cloud data**
Distributed computing is imagined as the cutting edge design of IT ventures, giving advantageous remote access to information stockpiling and application administrations. While this outsourced stockpiling model can possibly bring extraordinary practical reserve funds for information proprietors and clients, however because of wide worries of information proprietors that their private information might be automatically uncovered or dealt with by cloud suppliers. Despite the fact that conclusion to-end encryption methods have been proposed as promising answers for secure cloud information stockpiling. In this article, we distinguish the framework prerequisites and difficulties towards accomplishing security guaranteed accessible outsourced cloud information administrations. This paper display a general system for this, utilizing accessible encryption methods, which permits encoded information to be sought by clients without spilling data about the information itself and clients inquiries. The factual measure approach, i.e., significance score, from data recovery to construct a safe accessible file, and build up a one-to-many request saving mapping system to legitimately ensure those delicate score data. The subsequent plan can encourage productive server side positioning without losing watchword protection. [4]

5. **Efficient similarity search over encrypted data**
In the present time, because of appealing highlights of distributed computing, the monstrous measure of information has been put away in the cloud. Despite the fact that cloud-based administrations offer many advantages yet protection and security of the touchy information is a major issue. These issues are settled by putting away delicate information in scrambled frame. Encoded stockpiling secures the information against unapproved get to, yet it debilitates some essential and critical usefulness like inquiry operation on the information, i.e. looking through the required information by the client on the encoded information expects information to be unscrambled first and afterward seek, so this in the long run, backs off the way toward seeking. To accomplish this numerous encryption plans have been proposed, be that as it may, the majority of the plans handle correct Query coordinating yet not Similarity coordinating. While client transfers the record, highlights are removed from each archive. At the point when the client fires a question, trapdoor of that inquiry is created and look is performed by finding the connection among archives put away on cloud and question catchphrase, utilizing Locality Sensitive Hashing. [5]

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

Sr. No.	Paper	Advantages	Disadvantages	Method
1	Enabling secure and efficient ranked keyword search over outsourced cloud data.	The relevant search result which greatly improve the efficiency of search.	The effectiveness of ranked keyword search is increased by concepts public-key systems that support comparison queries on encrypted data as well as more general queries such as subset queries and has to support arbitrary conjunctive queries (P1, P2...... Pn) without leaking information on individual conjuncts	preserving symmetric algorithm
2	An efficient public key encryption with conjunctive-subset keywords search	the advantages of security and convenience since it will reveal the least amount of information to the server	Fixed keyword fields and variable keyword fields have advantage and drawback.	PECKS algorithm.
3	Privacy-preserving multi-keyword fuzzy search over encrypted data in the cloud	Improve it to meet different privacy requirements in two levels of threat models.	Boolean Symmetric search technique can be included to support multiple keyword search without making any changes to the existing architecture	KeyGen algorithm, SVD algorithm
4	Achieving usable and privacy-assured similarity search over outsourced cloud data	efficiency could be improved due to the adopted index structure, care needs to be taken to prevent leakage of private information to the cloud server	Plan to optimize the index construction algorithm and continue to research on usable and secure mechanisms for the effective utilization over outsourced cloud data.	index construction algorithm
5	Efficient similarity search over encrypted data	The advantages of multi-probe LSH over entropy based LSH.	The disadvantage of this structure is that it is a probabilistic data structure.	Locality Sensitive Hashing (LSH) and KNN algorithm

Copyright to IJIRCCE
DOI: 10.15680/IJIRCCE.2018. 0601067

III. SYSTEM DESIGN

Propose the first expressive SE scheme in the public-key setting from bilinear pairings in prime order groups. As such, our scheme is not only capable of expressive multi-keyword search, but also significantly more efficient than existing schemes built in composite-order groups. Using a randomness splitting technique, our scheme achieves security against offline keyword dictionary guessing attacks to the cipher texts. Moreover, to preserve the privacy of keywords against offline keyword dictionary guessing attacks to trapdoors, we divide each keyword into keyword name and keyword value and assign a designated cloud server to conduct search operations in our construction.

Formalize the security definition of expressive SE, and formally prove that our proposed expressive SE scheme is selectively secure in the standard model. Implement our scheme using a rapidly prototyping tool called Charm, and conduct extensive experiments to evaluate its performance. Our results confirm that the proposed scheme is sufficiently efficient to be applied in practice.

A trusted trapdoor generation center who publishes the system parameter and holds a master private key and is responsible for trapdoor generation for the system, data owners who outsource encrypted data to a public cloud, data users who are privileged to search and access encrypted data, and a designated cloud server who executes the keyword search operations for data users.

Attribute-based encryption storage system supporting secure de-duplication of encrypted data in the cloud, in which the cloud will not store a file more than once even though it may receive multiple copies of the same file encrypted under different access policies. A data provider intends to upload a file M to the cloud, and share M with users having certain credentials. In order to encrypts M under an access policy A over a set of attributes and uploads the corresponding ciphertext to the cloud, such that only users whose sets of attributes satisfying the access policy can decrypt the ciphertext.

Later, another data provider uploads a ciphertext for the same underlying file M but ascribed to a different access policy A. Since the file is uploaded in an encrypted form, the cloud is not able to discern that the plaintext corresponding to ciphertext is the same as that this will store M twice. Obviously, such duplicated storage wastes storage space and communication bandwidth.

Distributed computing is used to refer to any large collaboration in which many individual personal computer owners allow some of their computer's processing time to be put at the service of a large problem. In our system the each cloud admin consist of data blocks. The cloud users upload the data into multi cloud. Cloud computing environment is constructed based on open architectures and interfaces; it has the capability to incorporate multiple internal and/or external cloud services together to provide high interoperability. We call such a distributed cloud environment as a multi-Cloud .A multi-cloud allows clients to easily access his/her resources remotely through interfaces.

Division of Data in the Cloud for Optimal Performance and Security that collectively approaches the security and performance issues. In the division methodology, we divide a file into fragments data over the cloud nodes. Each of the nodes stores only a single fragment of a particular data file that ensures that even in case of a successful attack, no meaningful information is revealed to the attacker. Our main aim is to build a system where Optimization of Performance and Security of data in the cloud is enhanced. We implemented division of data in the cloud for excellent performance and security; we correlated the particular data division of a file into fragments.

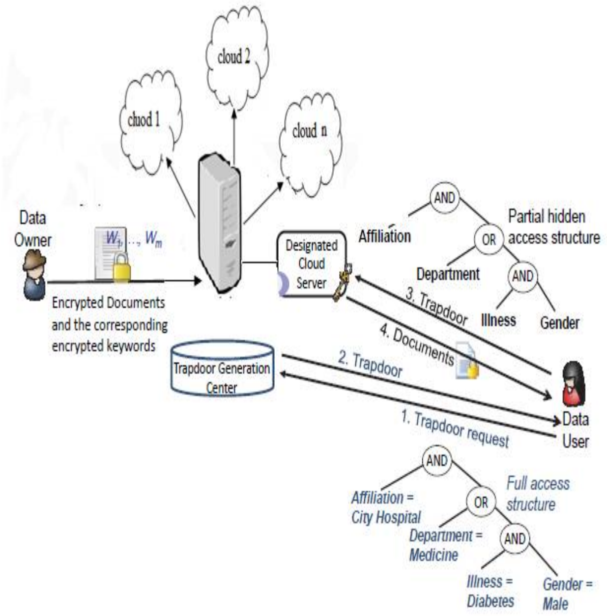

Fig: System Architecture

Algorithms :

Algorithm: Ranked Serial Binary Search (RSBS) algorithm

Input: Noised trapdoors (one per search keyword): T1,….., T e

Encrypted document indexes: A = I1----IN

The number of documents to return: k

Output:

Top-k documents that best match the search request:D = {D1.D2…..,Dk}

```
1: Scores = zeros(0;N) // create an array of N zeros
2: for i := 1 to N do
3: for n := 1 to e do
4: Score[i]   Score[i] + bsearch(Tn, Ii, 1 ,Si)// search if
the keyword appears in any of the s slices of the document
5: end for
6: end for
7: sorted, indices = sort(Scores) // sort the score array and
get the indices or old element in the sorted array.
```

```
8: D  <-- indice [0 : k - 1] // get the top-k documents
9: return D
```

IV. SYSTEM REQUIREMENT

Hardware resources required

- Processor – Pentium –III
- Speed – 1.1Ghz
- RAM – 256 MB(min)
- Hard Disk – 20 GB
- Key Board – Standard Windows Keyboard
- Mouse – Two or Three Button Mouse
- Monitor – SVGA
- Android Mobile

Software resources required

- Operating System: Windows 7
- Database : MYSQL
- Android Studio

V. APPLICATION

1. System also used for the personal use.
2. System can useful for storage the files secure manner
3. System also useful for organization.
4. Also use of college.

VI. CONCLUSION

In this paper, a safe, effective and dynamic pursuit conspire is proposed, which underpins not just the precise multi-keyword positioned look yet in addition the dynamic cancellation and inclusion of archives. We develop a unique watchword adjusted paired tree as the record, and propose a "Ravenous Profundity initially Search" calculation to acquire better proficiency than direct hunt. Also, the parallel pursuit process can be done to additionally lessen the time cost. The security of the plan is ensured against two risk models by utilizing the protected KNN calculation. Trial comes about illustrate the effectiveness of our proposed conspire.

There are as yet many test issues in symmetric SE plans. In the proposed plot, the information proprietor is dependable for producing refreshing data and sending them to the cloud server. Along these lines, the information proprietor needs to store the decoded list tree and the data that are important to recalculate the IDF esteems. Such a dynamic information proprietor may not be extremely reasonable for the distributed computing model. It could be an important however troublesome future work to outline a dynamic accessible encryption plot whose refreshing operation can be finished by cloud server just, in the interim saving the capacity to help multi-watchword positioned look. Furthermore, as the a large portion of works about accessible encryption, our plan for the most part considers the test from the cloud server. In reality, there are many secure difficulties in a multi-client conspire. To start with, every one of the clients for the most part keep the same secure key for trapdoor age in a symmetric SE plot. For this situation, the repudiation of the client is huge test. On the off chance that it is expected to repudiate a client in this plan, we require to remake the record and disseminate the new secure keys to all the approved clients. Second, symmetric SE plots as a rule accept that every one of the information clients are dependable. It isn't reasonable what's more, an untrustworthy information client will prompt many secure issues. For instance, an untrustworthy information client may look through

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 1, January 2018

the records and convey the unscrambled archives to the unapproved ones. Considerably more, an unscrupulous information client may circulate his/her safe keys to the unapproved ones. In the future works, we will endeavor to enhance the SE plan to handle these test issues.

REFERENCES

1. C. Wang, N. Cao, K. Ren, and W. Lou, “Enabling secure and efficient ranked keyword search over outsourced cloud data,” IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 8, pp. 1467–1479, Aug. 2012.
2. B. Zhang and F. Zhang, “An efficient public key encryption with conjunctive-subset keywords search,” J. Netw. Comput. Appl., vol. 34, no. 1, pp. 262–267, 2011.
3. B. Wang, S. Yu, W. Lou, and Y. T. Hou, “Privacy-preserving multi-keyword fuzzy search over encrypted data in the cloud,” in Proc.IEEE INFOCOM, 2014, pp. 2112–2120.
4. C. Wang, K. Ren, S. Yu, and K. M. R. Urs, “Achieving usable and privacy-assured similarity search over outsourced cloud data,” in Proc. IEEE INFOCOM, 2012, pp. 451–459.
5. M. Kuzu, M. S. Islam, and M. Kantarcioglu, “Efficient similarity search over encrypted data,” in Proc. IEEE 28th Int. Conf. Data Eng., 2012, pp. 1156–1167.
6. N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, “Privacy-preserving multi-keyword ranked search over encrypted cloud data,” in Proc. IEEE INFOCOM, Apr. 2011, pp. 829–837.
7. 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 Proc. 8th ACM SIGSAC Symp. Inf., Comput. Commun. secur., 2013, pp. 71–82.
8. C. Orencik, M. Kantarcioglu, and E. Savas, “A practical and secure multi-keyword search method over encrypted cloud data,” in Proc. IEEE 6th Int. Conf. Cloud Comput., 2013, pp. 390–397.
9. W. Zhang, S. Xiao, Y. Lin, T. Zhou, and S. Zhou, “Secure ranked multi-keyword search for multiple data owners in cloud computing,” in Dependable Syst. Networks (DSN), IEEE 44th Annu. IEEE/IFIP Int. Conf., 2014, pp. 276–286.
10. S. Kamara, C. Papamanthou, and T. Roeder, “Dynamic searchable symmetric encryption,” in Proc. ACM Conf. Comput. Commun. Secur., 2012, pp. 965–976.
11. S. Kamara and C. Papamanthou, “Parallel and dynamic searchable symmetric encryption,” in Proc. Financ. Cryptography Data Secur., 2013, pp. 258–274.
12. D. Cash, S. Jarecki, C. Jutla, H. Krawczyk, M.-C. Ro¸su, and M. Steiner, “Highly-scalable searchable symmetric encryption with support for boolean queries,” in Proc. Adv. Cryptol, 2013, pp. 353–373.

Copyright to IJIRCCE DOI: 10.15680/IJIRCCE.2018. 0601067