International Journal of Engineering Research-Online
A Peer Reviewed International Journal
Articles available online http://www.ijoer.in

Vol.3., Issue.6., 2015
(Nov.-Dec.,)

REVIEW ARTICLE

ISSN: 2321-7758

A SURVEY ON MULTI-KEYWORD RANKED SEARCH SCHEME OVER ENCRYPTED DATA IN CLOUD COMPUTING

AKHIL KUMAR GOUR[1], AMRUTA DESHMUKH[2], PRITAM BANKAR[3], PRANITA GAWADE[4], Prof. RANJANA BADRE[5]

[1,2,3,4]UG Student, [5]Assistant Professor

Computer Department, MIT Academy of Engineering, Alandi, Pune (MAHARASHTRA), INDIA

ABSTRACT

Cloud computing represents today's most significant issue in information technology paradigm. Utilizing cloud computing, people can store, manage and process their data on remote servers. The primary barriers to its wide adoption are the security and privacy concerns. Sensitive data generally has to be scrambled before outsourcing for the shielding of data privacy. Accordingly, empowering a scrambled cloud information pursuit administration is of principle significance. This discards proper data utilization like keyword-based document retrieval. Thus, enabling an encrypted cloud data search service is of paramount importance. In consideration of the large number of cloud users and their data, it is critical for the search service to allow multi-keyword search and provide ranked result to achieve the successful recovery of data need. A multi-keyword ranked search scheme, in cloud computing, underpins dynamic operations like insertion and deletion of documents.

I. INTRODUCTION

The cloud computing regards registering as a utility and leases out the processing and stockpiling abilities to people in general. In such a framework, the individual can remotely store the data on the cloud server, specifically data outsourcing, and after that make the cloud data open for free, through the cloud server. This speaks to a more versatile, minimal effort and stable path for open information access on account of the adaptability and high efficiency of cloud servers.

Utilizing cloud computing, people can store their information on remote servers and permit information access to open clients through the cloud servers. As the outsourced information are prone to contain delicate protection data, they are commonly scrambled before transferred to the cloud. This significantly constrains the ease of use of outsourced information because of the difficulty of looking over the encoded information. To address this issue, some universally useful arrangements with completely homomorphic encryption [1] or negligent RAMs [2] have been outlined by analysts.

In any case, because of their high computational overhead for both the cloud server and client, these systems are not reasonable. Unexpectedly, searchable encryption schemes empower the client to store the scrambled information to the cloud and execute catchphrase search over ciphertext area. Searchable encryption

(SE) schemes are more practical and have made specific contributions in terms of efficiency, functionality and security. Among them, multi-keyword ranked search accomplishes more consideration for its practical applicability. In the proposed system, this issue of multi-keyword ranked search over encrypted cloud data, while preserving strict system-wise privacy in cloud computing paradigm, will be addressed by building up the fine-grained multi-keyword pursuit plans over encrypted cloud information.

III. Related Works

A. Order preserving symmetric encryption (OPSE)

To protect the privacy, clients need to encode their delicate information before outsourcing it to the cloud. Then again, the conventional encryption plans are lacking since they make the utilization of indexing and searching operations all the more difficult. Accordingly, searchable encryption systems are developed to conduct search operations over a set of encrypted data. Unfortunately, these systems just permit their customers to perform a precise search but not approximate search; a vital requirement for all the present data retrieval systems.

As of late, an expanded consideration has been paid to the approximate searchable encryption frameworks to find keywords that match the submitted queries approximately. It concentrates on building a flexible secure index that permits the cloud server to perform the surmised search operations without uncovering the content of the query trapdoor or the index content. Specifically, the most recently cryptographic primitive, order preserving symmetric encryption (OPSE), has been utilized to protect the keywords [3], [4], [5], [6]. It further divides the search operation into two steps. The first step, finds the candidate list in terms of secure pruning codes. In particular, two methods are developed to construct these pruning codes. The second step utilizes a semi honest third party to decide the best matching keyword relying upon secure comparability function. As little information as possible is revealed to third party. Developing such a system enhances the utilization of retrieval information systems and makes these systems more user-friendly.

B.Attribute-based keyword search scheme with efficient user revocation (ABKS-UR)

This system concentrates on an alternate yet all the more difficult situation where the outsourced dataset is contributed from different proprietors and is searchable by numerous clients, i.e. multi-user multi-contributor case. Inspired by attribute-based encryption (ABE), the first attribute-based keyword search scheme with efficient user revocation (ABKS-UR) is presented that enables scalable fine-grained (i.e. file-level) search authorization [7]. The system permits different proprietors to scramble and outsource their information to the cloud server independently. Clients can produce their own particular search capacities without depending on an online trusted authority. Fine-grained search approval is likewise enforced by the proprietor authorized access scheme on the index of every file. Further, by consolidating intermediary re-encryption and lazy re-encryption strategies, it can designate substantial system update workload amid client repudiation to the resourceful semi-trusted cloud server. It formalizes the security definition and proves the ABKS-UR scheme selectively secure against chosen-keyword attack.

C. Privacy-aware bedtree based approach

Traditional searchable encryption schemes ordinarily only support accurate keyword matches. However, clients infrequently utilize somewhat distinctive formats e.g. “data-mining” versus “data mining”. In this way, fuzzy keyword search is a helpful feature to have. As of late, a few analysts proposed utilizing wildcard based approach to deal with fuzzy keyword search. They likewise proposed an answer for multi-catchphrase look.

Their methodologies have some limitations, namely (a) fuzzy keyword search solution devours expansive capacity size since it embeds each fuzzy keyword as a leaf node in the index tree, (b) fuzzy single-keyword search solution does not bolster multi-keyword search,(c) the current multi-keyword search scheme does not does not give proficient incremental updates. A privacy-aware bedtree based way to deal with multi-keyword feature is

designed in [8]. Incremental updates can be easily done byutilizing this solution. It is cost-effective in terms of storage size and development time. Additionally, the search time is typically superior to the wildcard approach for multi-keyword queries where numerous encrypted files are returned using single-word queries for approaches that do not support multi-keyword queries.

D. Verifiable Keyword-based Semantic Search

Despite the fact that the current searchable encryption plans empower clients to look over encoded information, these plans don't bolster unquestionable status of query item.So as to spare computation cost or download bandwidth, cloud server just directs a small amount of search operation or return a part of result, which is seen as selfish and semi-genuine however inquisitive. To upgrade adaptability of scrambled cloud information while supporting verifiability of query item is a major challenge. To handle the challenge, a keen semantic hunt plan is utilized which returns not just the result of keyword based accurate match, but also the result of keyword based semantic match [9]. In the meantime, this plan underpins the irrefutability of output.

E. Public key encryption with keyword search

Public key encryption with keyword search (PKES) empowers senders to send encoded information to a beneficiary like conventional public key encryption (PKE) plans. The disparity between PKES and PKE is that the recipient in PKES can seek on the encoded information which is stored on the third party server (like a cloud storage server). The greater part of the existed PKES plans depend on bilinear map, so they are expensive in calculation and difficult to be utilized as a part of practice. A PKES plan is built taking into account factoring, which is computationally proficient and secure [10]. A public modules and an irregular component of the arrangement of numbers is required.

F. TEES (Traffic and Energy saving Encrypted Search)

Data encryption is an over whelming overhead for the mobile devices, and information retrieval process causes a complicated correspondence between the information user and cloud. Generally, with restricted data transmission limit and constrained battery life, these issues acquaint heavy overhead with computation and communication and also a higher power utilization for cell phone users, which makes the encoded search over mobile cloud exceptionally difficult.

TEES (Traffic and Energy saving Encrypted Search) is data transmission and energy proficient scrambled search architecture over mobile cloud. This building design offloads the calculation from cell phones to the cloud, and further enhances the correspondence between the mobile customers and the cloud. It is shown that the information privacy does not corrupt when the performance improvement systems are applied. TEES lessens the calculation time by 23% to 46% and save the energy utilization by 35% to 55% for every document retrieval, in the meantime the system traffics amid the file retrievals are likewise fundamentally reduced.

IV. LITERATURE SURVEY

At the point when sensitive information is outsourced to the cloud, information proprietors normally get to be worried with the security of their information in the cloud and beyond. Encryption-before-outsourcing has been viewed as a principal method for ensuring client information privacy against the cloud server [11]. However, how the encoded information can be viably used then gets to be another challenge.

In the previous systems, successful keyword searching schemes have been created which uses bilinear maps and which depend on public key encryption strategy. This plan works just for single client and all the more significantly, queries in this plan got created in an extremely dynamic way, and henceforth, not able to hide the search pattern [12]. A few schemes are created in which client must have knowledge about all the substantial keywords and their separate positions as compulsory data in order to produce a query [11]. Likewise, past plans have concentrated on multi-client searchable plan [13]. Drawbacks of these frameworks can be,

1. Single-keyword search without ranking.
2. Boolean keyword search without ranking.
3. Single-watchword search without ranking.
4. Once in a while sorting of the results i.e. no index

creation and ranking.

5. Single User pursuit.

Because of diverse cryptography primitives, searchable encryption plans can be built utilizing public key based cryptography [10],or symmetric key based cryptography [14].Song et al. [14] proposed the first symmetric searchable encryption (SSE) scheme, and the search time of their plan is linear to the measure of the information gathering.

These early works are single catchphrase boolean search schemes, which are exceptionally straightforward as far as usefulness. A while later, plenteous works have been proposed under distinctive threat models to accomplish different search functionality, for example, single keyword search, similarity search [15], multi-keyword boolean search[16], [17], [18], ranked search[19], and multi-keyword ranked search[20], [21] and so on.

Multi-keyword boolean search permits the clients to enter different query keywords to ask for suitable results. Among these works, conjunctive keyword search plans [16], just give back the documents that contain the greater part of the query keywords. Disjunctive keyword search schemes [17] give back the greater part of the documents that contain a subset of the query keywords. Predicate search plans [18] are proposed to bolster both conjunctive and disjunctive search. All these multi-keyword search plans retrieve indexed results taking into account the presence of keywords, which can't give adequate result ranking functionality.

Ranked search can empower speedy search of the most relevant information. Sending back just the top-k most significant documents can successfully lessen network traffic. Some early works [19] have understood the ranked search utilizing order preserving methods, yet they are designed just for single keyword search. Cao et al. [20] understood the first privacy safeguarding multi-keyword ranked search scheme, in which documents and queries are described as vectors of dictionary size. With the "coordinate matching", the records are ranked by number of coordinated query keywords.

Be that as it may, Cao et al's. plan does not consider the significance of the distinctive keywords, and accordingly is not sufficiently exact. Moreover, the search efficiency of the plan is linear with the cardinality of record collection. Sun et al. [21] proposed a protected multi-keyword search scheme that supports similarity based ranking. The authors developed a searchable index tree taking into account vector space model and adopted cosine measure together with TF×IDF to give ranked results. Sun etal's. search method accomplishes superior to anything linear query efficiency yet brings about accuracy loss.

Orencik et al. [22] proposed a secure multi-keyword search strategy which used local sensitive hash (LSH) methods to group the identical records. The LSH calculation is suitable for comparative query yet can't give definite ranking. In [23], Zhang et al. proposed a strategy to manage secure multi-keyword ranked search in a multi-proprietor model. In this scheme, distinctive information proprietors use different secret keys to scramble their files and keywords while authorized information users can query without knowing keys of these diverse data proprietors. The authors presented an "Additive Order Preserving Function" to retrieve the most pertinent search items.

On the other hand, these works don't bolster dynamic operations. Essentially, the information owner might need to update the record collection after he transfers the accumulation to the cloud server. Accordingly, the SE schemes are relied upon to support the deletion and insertion of the records. There are additionally a few dynamic searchable encryption strategies. In the work of Song et al. [14], the every document is considered as a succession of fixed length words, and is independently listed. This scheme underpins straight-forward update operations yet with low efficiency.

For meeting the challenge of supporting multi-keyword ranked search over scrambled cloud data, while safeguarding strict system-wise privacy in cloud computing, a fundamental idea of MRSE utilizing secure inner product computation is proposed.

V. CONCLUSION

The issue of multi keyword search scheme over encrypted data in cloud computing will be taken care of. These arch operation will be more comfortable yet secure by allowing the user to search the encrypted cloud. It will viably perform result relevance ranking instead of returning undifferentiated results. In addition, the issue of preserving the privacy of sensitive information will be tackled.

VI. REFERENCES

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

[2]. 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.

[3]. http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6468245

[4]. https://www.deepdyve.com/lp/institute-of-electrical-and-electronics-engineers/approximate-keyword-based-search-over-encrypted-cloud-data-UE47GL7Mpt

[5]. https://www.researchgate.net/publication/262325511_Approximate_Keyword-based_Search_over_Encrypted_Cloud_Data

[6]. Ayad Ibrahim, Hai Jin, Ali A. Yassin, Deqing Zou, "Approximate Keyword-based Search over Encrypted Cloud Data," in 2012 Ninth IEEE International Conference on e-Business Engineering.

[7]. Wenhai Sun, Shucheng Yu, Wenjing Lou, Y. Thomas Hou, Hui Li, "Protecting Your Right: Verifiable Attribute-based Keyword Search with Fine-grained Owner-enforced Search Authorization in the Cloud," IEEE Transactions on Parallel and Distributed Systems, 2014.

[8]. M. Chuah, W. Hu, "Privacy-aware BedTree Based Solution for Fuzzy Multi-keyword Search over Encrypted Data," 2011 31st International Conference on Distributed Computing Systems Workshops.

[9]. Zhangjie Fu, Jiangang Shu, Xingming Sun, and Nigel Linge, "Smart Cloud Search Services: Verifiable Keyword-based Semantic Search over Encrypted Cloud Data," IEEE Transactions on Consumer Electronics, Vol. 60, No. 4, November 2014.

[10]. 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.

[11]. Ayad Ibrahim, Hai Jin, Ali A.Yassin, DeqingZou, "Secure Rank Ordered Search of Multi-Keyword Trapdoor over Encrypted Cloud Data" IEEE Asia-Pacific Services Computing Conference 2012.

[12]. Ning Cao, Cong Wang, Ming Li, KuiRen, Wenjing Lou , "Privacy Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data" IEEE Transactions on parallel and Distributed Systems, Vol. 25,January 2014.

[13]. Yanjiang Yang, "Towards Multi-User Private Keyword Search for Cloud Computing" IEEE 4th International Conference on Cloud Computing. 2011.

[14]. D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in IEEE Proceedings of S&P, 2000.

[15]. 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.

[16]. Y. H. Hwang and P. J. Lee, "Public key encryption with conjunctive keyword search and its extension to a multi-user system," in Proceedings of the First international conference on Pairing-Based Cryptography. Springer-Verlag, 2007, pp. 2–22.

[17]. B. Zhang and F. Zhang, "An efficient public key encryption with conjunctive-subset keywords search," Journal of Network and Computer Applications, vol. 34, no. 1, pp. 262–267, 2011.

[18]. J. Katz, A. Sahai, and B. Waters, "Predicate encryption supporting disjunctions, polynomial equations, and inner products," in Advances in Cryptology–EUROCRYPT 2008. Springer, 2008, pp. 146–162.

[19]. C. Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data," Parallel and Distributed Systems, IEEE Transactions on, vol. 23, no. 8, pp. 1467–1479, 2012.

[20]. N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," in IEEE INFOCOM, April 2011, pp. 829–837.

[21]. 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, computer and communications security. ACM, 2013, pp. 71–82.

[22]. C.Orencik, M.Kantarcioglu, and E.Savas,"A practical and secure multi-keyword search method over encrypted cloud data," in Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013, pp. 390–397.

[23]. 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 Systems and Networks (DSN), 2014 44th Annual IEEE/IFIP International Conference on. IEEE, 2014, pp. 276–286.