International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 05 | May -2017 www.irjet.net p-ISSN: 2395-0072

A Review on Muilt-Keyword Search in Encrypted Data with Privacy Preservation

Swati S. Deshmukh
M.Tech Student, Department Of Computer Science & Engineering
J. D. College Of Engineering & Management
Prof. Shrikant V. Sonekar
Asst. Professor, Department Of Computer Science & Engineering
J. D. College Of Engineering & Management

***Abstract*— Cloud computing is an on-request computing. It is an Internet-based computing. Around there shared assets, information and data are given on request to PCs and different gadgets. It additionally gives the administrations over the web. In distributed computing, specialist organizations have capacity to give stockpiling at server as per clients require. They permit clients to store and recover the information in cloud server on request from anyplace and on a gadget. This control of information at cloud server offers ascend to such a variety of security issues since information is gotten to over web. For security reason information is store in encoded organized. In this, customer has no immediate control on information once it is transferred on cloud server. In this paper, we examine the thought behind single watchword hunt over encoded information and furthermore multi catchphrase positioning. Cloud information proprietors need their records in a scrambled frame with the end goal of protection saving. In this manner it is important to create productive and solid ciphertext seek systems. One test is that the connection between records will be ordinarily hidden during the time spent encryption, which will prompt critical inquiry precision execution debasement.**

***Keywords*— Cloud computing, Encryption, Inner product similarity, Single Keyword Search, Multi-keyword search, ranking**

I. INTRODUCTION

As we venture into the enormous information period, terabyte of information is created overall every day. In the late 1960's the possibility of "Utility computing" that was authored by MIT PC researcher and Turing grant champ John McCarthy was ideally known as the idea of distributed computing over a system. Enterprises were searching for some kind of significant arrangement, since utility computing wound up getting to be something of a major business for organizations, for example, IBM. To be sure, Martin Greenberger called attention to the idea that "cutting-edge arithmetical machines without bounds" were currently being utilized institutionally for logical figuring and research as well as for business capacities, for example, bookkeeping and stock. Advance, he foreseen his bit of work in which PCs would be widespread practically like the real power organizations running wires wherever in due time. Endeavours and clients who possess a lot of information more often than not outsource their valuable information to cloud office with a specific end goal to diminish information administration cost and storeroom spending. Therefore, information volume in distributed storage offices is encountering an emotional increment. Despite the fact that cloud server suppliers (CSPs) guarantee that their cloud administration is outfitted with solid safety efforts, security and protection are real hindrances keeping the more extensive acknowledgment of distributed computing administration [1]. A conventional approach to decrease data spillage is information encryption. Be that as it may, this will make server-side information use, for example, looking on encoded information, turn into an exceptionally difficult assignment. In the current years, analysts have proposed numerous ciphertext seek plans [35-38] [43] by joining the cryptography methods. These techniques have been demonstrated with provable security, yet their strategies require gigantic operations and have high time intricacy. Consequently, previous techniques are not appropriate for the enormous information situation where information volume is huge and applications require online information handling. Also, the connection between archives is covered in the above strategies. The connection between reports speaks to the properties of the archives and subsequently keeping up the relationship is essential to completely express a record. For instance, the relationship can be utilized to express its classification. On the off chance that a record is autonomous of some other reports aside from those archives that are identified with games, then it is simple for us to state this archive has a place with the classification of the games. Because of the visually impaired encryption, this vital property has been covered in the customary strategies. Thusly, proposing a strategy which can keep up and use this relationship to speed the pursuit stage is alluring. Then again, because of programming/equipment disappointment, and capacity debasement, information list items coming back to the clients may contain harmed information or have been twisted by the pernicious manager or interloper. In this manner, an unquestionable instrument ought to be given to clients to check the accuracy and culmination of the list items. Because of a progressive change in the field of enterprises over past decade, there has been increment

sought after of outsourcing of information over an extensive variety of system. With a specific end goal to control this enormous measure of information in financially savvy way venture has adjusted a predominant innovation called distributed computing that expel the weight of information administration. In this information driven condition endeavour tend to store their information onto cloud that involves profitable resource of client information like messages, individual wellbeing information and so on. Distributed computing is ending up being most basic worldview in the improvement of data innovation which offer adaptable get to, pervasive, on request get to and capital consumption sparing.

II. LITERATURE REVIEW

Qin Liu et al. proposed Secure and protection saving catchphrase seek in [1]. It gives watchword protection, information protection and semantic secure by open key encryption. The principle issue of this pursuit is that the correspondence and computational cost of encryption and decoding is more.

Ming Li et al. proposed Authorized Private catchphrase Search (APKS) in [2]. It gives watchword protection, Index and Query Privacy, Fine-grained Search Authorization and Revocation, Multi-dimensional Keyword Search, Scalability and Efficiency. This hunt technique builds the inquiry effectiveness utilizing trait chain of importance yet by and by every one of the characteristics are not progressive.

Cong Wang et al in [3] proposed Secure and Efficient Ranked Keyword Search which fathoms preparing overhead, information and catchphrase protection, least correspondence and calculation overhead. It is not valuable for different catchphrase seeks, Also there is a tad bit of overhead in record building.

Kui Ren et al. [4] proposed Secured fluffy catchphrase seek with symmetric searchable encryption (SSE). It doesn't bolster fluffy pursuit with open key based searchable encryption; likewise it can't play out various watchwords semantic hunt. The updates for fluffy searchable list are not effectively performed.

Ming Li et al. [5] proposed Privacy guaranteed searchable distributed storage technique. It is executed utilizing SSE, Scalar-Product-Preserving Encryption and Order-Preserving Symmetric Encryption. It bolsters the security and useful necessities. This plan does not bolster open key based searchable encryption.

Wei Zhou et al. [6] proposed K-gram based fluffy watchword Ranked Search. In this proprietor make k-gram fluffy watchword list for records D and tuple <I, D> is transferred to inquiry server (SS) which is embedded to sprout channel for size controlling. The encoded record D is transferred to capacity server. However, the issue is that, the measure of the k-gram construct fluffy catchphrase set depends with respect to the jacquard coefficient esteem.

J. Baek et al. in [7] proposed Secure Channel Free Public Key Encryption with Keyword Search (SCF-PEKS) strategy. In these strategy bunch servers makes its own open and private key match yet this technique experiences outside aggressor by KGA.

H. S. Rhee et al. [8] proposed Trapdoor in noticeability Public-Key Encryption with Keyword Search (IND-PEKS). In this outsourcing is done as SCF-PEKS. It experiences outside assailant utilizing KGA and breaking down the recurrence of event of watchword trapdoor.

Peng Xu et al. [9] proposed Public-Key Encryption PEFKS with Fuzzy Keyword Search, in this client makes fluffy catchphrase trapdoor Tw and correct watchword trapdoor Kw for W. Client asks for Tw to CS. At that point CS checks Tw with fluffy watchword list and sends superset of coordinating figure messages by Fuzz Test calculation that is executed by CS. The client procedure Exact Test calculation for confirming ciphertexts with Kw and recover the scrambled records. The way toward making fluffy watchword list and correct catchphrase list is troublesome for huge size database.

Ning et al. [10] proposed Privacy Preserving Multi Keyword Ranked Search (MRSE). It is valuable for known figure content model and foundation display over encoded information. It gives low calculation and correspondence overhead. The facilitate coordinating is chosen for multi-catchphrase seek. The disadvantage is that MRSE have little standard deviation which decreases the watchword security.

III. PROPOSED WORK

Cloud data proprietors need to outsource reports in an encoded outline with the true objective of security sparing. Thusly it is essential to make beneficial and tried and true ciphertext look techniques. One test is that the association between documents will be ordinarily masked amid the time spent encryption, which will provoke gigantic request precision execution degradation. In like manner the volume of data in server ranches has experienced an electrifying improvement. This will make it significantly all the more hard to arrange ciphertext look for contrives that can give gainful and tried and true online information recuperation on extensive volume of mixed data.

The front line web called the Semantic Web will help the customer to recuperate the important data that is secured on the cloud as transcendentalism and make the data clear to the customer which is stayed behind the cloud. The purpose of the proposed situating count is to give customers the result set of correlated data.

A different levelled clustering technique is proposed to reinforce more interest semantics and moreover to deal with the interest for speedy figure content chase inside a noteworthy data condition. The proposed dynamic approach amasses the reports in light of the base centrality restrict, and after that bundles the ensuing packs into sub-groups until the basic on the most outrageous size of cluster is come to. In the interest arrange, this approach can accomplish a straight computational multifaceted design against an exponential size augmentation of chronicle social event. In order to check the authenticity of rundown things,

a structure called minimum hash sub-tree is arranged. We extend this considered semantic resemblance to consider natural associations between thoughts using ontologies. We propose situating figuring with multi catchphrase and cosmology.

Cloud data proprietors need to outsource reports in an encoded outline with the true objective of security sparing. Thusly it is major to make gainful and reliable ciphertext look strategies. One test is that the association between files will be regularly masked amid the time spent encryption, which will provoke tremendous request precision execution corruption. In like manner the volume of data in server ranches has experienced a hair-raising advancement. This will make it extensively all the more hard to arrange ciphertext look for contrives that can give beneficial and tried and true online information recuperation on sweeping volume of mixed data.

The forefront web called the Semantic Web will help the customer to recoup the profitable data that is secured on the cloud as mysticism and make the data evident to the customer which is stayed behind the cloud. The purpose of the proposed situating computation is to give customers the result set of relevant data.

A different levelled grouping system is proposed to reinforce more interest semantics and besides to deal with the interest for brisk figure content chase inside a noteworthy data condition. The proposed dynamic approach aggregates the reports in light of the base criticalness constrain, and after that packages the ensuing bundles into sub-groups until the basic on the most outrageous size of cluster is come to. In the interest organize; this approach can accomplish a straight computational multifaceted nature against an exponential size augmentation of document social event. To check the authenticity of rundown things, a structure called minimum hash sub-tree is arranged. We grow this considered semantic resemblance to consider characteristic associations between thoughts using ontologies. We propose situating computation with multi catchphrase and cosmology.

IV. RELATED WORK

A. Secure and privacy preserving keyword search

Qin Liu [16] proposed in this paper the inquiry 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 elevated in decoding process for end clients. The client's presents the watchword trapdoor encoded by users‟ private key to CS (Cloud Server) safely and recover the scrambled reports.

B. Secure and Efficient Ranked Keyword Search

Cong Wang [17] proposed look which illuminates handling overhead, information and watchword protection, least correspondence and calculation elevated. The information proprietor assemble file alongside the catchphrase recurrence based importance scores for records. Client asks for „w‟ to cloud server with discretionary "k‟ as Tw utilizing the private key. The cloud server seeks the record with scores and sends encoded document in light of positioned succession.

C. Single Keyword Search Over Encrypted data on cloud

Realistic searchable encryption conspire agree to a client to solidly search for over scrambled information through catchphrases without first applying decoding on it, the proposed procedures bolster just ordinary Boolean watchword look, without catching any materialness of the documents in the item. At the point when straightforwardly connected in huge joint information outsourcing cloud condition, they experience next deficiency.

D. Privacy-preserving Multi-keyword Text Search

Wenhai Sun [19] proposed this inquiry that gives comparability based item positioning, catchphrase protection, Index and Query privacy and Query Unlink capacity. The encoded document is worked by vector space show supporting solidified and particular record look. The searchable record is assembled utilizing Multidimensional B tree. Proprietor makes scrambled inquiry vector Q̄ for record watchword set. Client gets the individual scrambled question 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 inquiry vector and returns best k encoded records to client.

E. Secure Multi-keyword Top-k Retrieval Search

Jiadi [20] proposed this pursuit utilizing Two round searchable encryption (TRSE). In first round, clients presents various catchphrase "REQ" 'W‟ as scrambled inquiry for fulfilling information, watchword security and make trapdoor (REQ, PK) as Tw and sends to cloud server. At that point cloud server ascertains the score from encoded file for records and returns the scrambled score result vector to client. In second round, client decode N with mystery key and figures the document positioning and after that demand records with Top k scores. The positioning of document is done on customer side and scoring is done on server side.

F. Privacy Preserving Multi-Keyword Ranked Search (MRSE)

Ning [21] proposed this scan for known figure content model and foundation display over scrambled information giving low calculation and correspondence overhead. Then arrange coordinating is decided for multi-watchword seeks. They utilized inward item likeness to quantitatively assess comparability for positioning records. The downside is that MRSE have little standard deviation σ which debilitates catchphrase security.

G. Attribute-based Keyword Search

Wenhai Sun [22] proposed Attribute-based Keyword Search that gives conjunctive watchword seek; catchphrase semantic security and Trapdoor unlink capacity. The proprietors makes list with all watchwords and get to list with arrangement an trait which indicates the clients list approved for seeking. Presently proprietors scramble the report, list with get to rundown utilizing ciphertext

approach property based encryption method. To have client participation 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 hunts the encoded files and return documents just if the user's qualities in Tw fulfills get to approaches in records which makes coarse-grained dataset seek approval.

H. Efficient and Secure Multi-Keyword Search on Encrypted Cloud Data

This proposed technique has characterized and tackled the issue of compelling however sheltered and sound rank watchword seek over Encrypted cloud information [23]. Positioned look enormously upgrades framework ease of use by giving back the coordinating documents in a positioned arrange with respect to certain essential criteria (e.g. watchword recurrence) in this way 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 seek over scrambled cloud information stockpiling (MRSE), and set up an arrangement of strict security necessities for such an ensured cloud information usage framework to wind up distinctly a reality. The proposed positioning technique turns out to be productive to backpedal to a great degree important archives comparing to submit seek terms. Proposed positioning technique is utilized as a part of our future framework with a specific end goal to improve the security of data on Cloud Service Provider.

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

This proposed strategy [24] recommend a protected tree-based hunt conspire over the encoded distributed storage, which bolsters multi watchword positioned seek alongside element operation on report gathering accessible at server. The vector space model and term recurrence (TF) × opposite archive recurrence (IDF) model are combinly utilized as a part of the development of file and era of inquiry to give multi watchword positioned seek yield. To get high pursuit proficiency comes about, creator develop a tree-based record structure and proposed a Greedy Depth-first Search calculation in view of this list tree. In light of this extraordinary structure of tree-based record, the proposed look plan can adaptably accomplish sub straight inquiry time and can successfully manage the erasure and addition of archives. The kNN calculation is connected to scramble the file and inquiry vectors, and till then guarantee precise pertinence score count between encoded list and question vectors.

V. ARCHITECTURAL VIEW

A various leveled bunching technique is proposed to bolster more inquiry semantics and furthermore to take care of the demand for quick figure content pursuit inside a major information condition. The proposed progressive approach groups the archives in light of the base significance limit, and after that segments the subsequent bunches into sub-groups until the requirement on the most extreme size of bunch is come to. In the hunt stage, this approach can achieve a direct computational many-sided quality against an exponential size increment of archive accumulation.

With a specific end goal to confirm the legitimacy of indexed lists, a structure called least hash sub-tree is composed. We expand this idea of semantic closeness to consider innate connections between ideas utilizing ontology's. We proposed positioning calculation with multi watchword and philosophy.

VI. CONCLUSIONS

This paper concentrates diverse procedures of looking in the encoded cloud data stockpiling. We have methodically presents the security and data use issues in the appropriated stockpiling related to all open looking for methodology. Thus recognized the essential issues that are to be satisfied for secured data utilize are catchphrase assurance, Data security, Index security, Query Privacy, Fine-grained Search, Scalability, Efficiency, Result situating, Index mystery, Query grouping, Query Unlinkability, semantic security and Trapdoor Unlinkability. By far most of the looking systems for the most part focus on security and some on data utilize. The obstacles of all the looking for strategies are also discussed. By the above survey, security can be given by Public-Key Encryption and capable data use by soft catchphrase look. We assume that this review will make the pros to shape their issue in the scope of data use in conveyed stockpiling.

REFERENCES

[1] Qin Liuy, Guojun Wangyz, and Jie Wuz,"Secure and privacy preserving keyword searching for cloud storage services", ELSEVIER Journal of Network and computer Applications, March 2011

[2] Ming Li et al.," Authorized Private Keyword Search over Encrypted Data in Cloud Computing,IEEE proc. International conference on distributed computing systems,June 2011,pages 383-392

[3] 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

[4] Kui Ren et al., "Towards Secure and Effective Data utilization in Public Cloud", IEEE Transactions on Network, volume 26, Issue 6, November / December 2012

[5] Ming Li et al.,"Toward Privacy-Assured and Searchable Cloud Data Storage Services", IEEE Transactions on Network, volume 27, Issue 4, July/August 2013

[6] Wei Zhou et al., "K-Gram Based Fuzzy Keyword Search over Encrypted Cloud Computing "Journal of Software Engineering and Applications, Scientific Research , Issue 6, Volume 29-32,January2013

[7] Baek et al., "Public key encryption with keyword search revisited", in ICCSA 2008, vol. 5072 of Lecture Notes in Computer Science, pp. 1249 - 1259, Perugia, Italy, 2008. Springer Berlin/Heidelberg.

[8] H. S. Rhee et al., "Trapdoor security in a searchable public-key encryption scheme with a designated tester," The Journal of Systems and Software, vol. 83, no. 5, pp. 763-771, 2010.

[9] Peng Xu et al., Public-Key Encryption with Fuzzy Keyword Search: A Provably Secure Scheme under Keyword Guessing Attack",IEEE Transactions on computers, vol. 62, no. 11, November 2013

[10] Ning Cao et al.," Privacy-Preserving Multi- Keyword Ranked Search over Encrypted Cloud Data", IEEE Transactions on parallel and distributed systems, vol. 25, no. 1, jan 2014
[11] D. X. D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. S & P, BERKELEY, CA, 2000, pp. 44.
[12] C. Wang, N. Cao, K. Ren, and W. J. 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.
[13] 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. ASIACCS, Hangzhou, China, 2013, pp. 71-82.
[14] R. X. Li, Z. Y. Xu, W. S. Kang, K. C. Yow, and C. Z. Xu, Efficient multi-keyword ranked query over encrypted data in cloud computing, Futur. Gener. Comp. Syst., vol. 30, pp. 179-190, Jan. 2014.
[15] Gurdeep Kaur, Poonam Nandal, "Ranking Algorithm of Web Documents using Ontology", IOSR Journal of Computer Engineering (IOSR-JCE) eISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 3, Ver. VIII (May-Jun. 2014), PP 52-55
[16] Qin Liuy, Guojun Wangyz, and Jie Wuz, "Secure and privacy preserving keyword searching for cloud storage services", ELSEVIER Journal of Network and computer Applications, March 2011
[17] 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
[18] International Journal of Computer Applications (0975 –887) Volume 126 – No.14, September 2015
[19] Wenhai Sun et al., "Privacy-Preserving Multikeyword Text Search in the Cloud Supporting Similarity-based Ranking'', the 8th ACM Symposium on Information, Computer and Communications Security, Hangzhou,
[20] China, May 2013.
[21] Jiadi Yu, Peng Lu, Yanmin Zhu, Guangtao Xue,Member, IEEE Computer Society, and Minglu Li,"Toward Secure Multikeyword Top k Retrieval over Encrypted Cloud Data", IEEE Journal of Theoretical and Applied Information Technology 10th August 2014.
[22] Ning Cao et al.," Privacy-Preserving MultiKeyword Ranked Search over Encrypted Cloud Data", IEEE Transactions on parallel and distributed systems, vol. 25, no. 1, jan 2014
[23] 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
[24] Secure Ranked Keyword Search over Encrypted Cloud Data, IEEE PAPER, 2010.
[25] Zhihua Xia, "A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud Data", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL: PP NO: 99 YEAR 2015