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

Gauri Bodkhe[1],Prof. Gurudev B. Sawarkar[2]
Department of Computer Science & Engineering
[1]M.Tech Students,V. M. Institute Of Engineering & Technology, Nagpur
[2]Assistant Professor,V. M. Institute Of Engineering & Technology, Nagpur

Abstract-*Distributed computing is an on-demand computing. It is an Internet-based computing. Around there shared resources, data and information are given on demand to PCs and diverse contraptions. It also gives the organizations over the web. In circulated computing, expert associations have ability to give stockpiling at server according to customers require. They allow customers to store and recuperate the data in cloud server on demand from wherever and on a contraption. This control of data at cloud server offers climb to such an assortment of security issues since data is gotten to over web. For security reason data is store in encoded sorted out. In this, client has no prompt control on data once it is exchanged on cloud server. In this paper, we inspect the idea behind single watchword chase over encoded data and besides multi catchphrase situating. Cloud data proprietors require their records in a mixed edge with the ultimate objective of insurance sparing. In this way it is critical to make profitable and strong ciphertext look for frameworks. One test is that the association between records will be normally covered up amid the time spent encryption, which will incite basic request exactness execution degradation.*

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

I. INTRODUCTION

As we wander into the tremendous data time frame, terabyte of data is made in general each day. In the late 1960's the likelihood of "Utility computing" that was composed by MIT PC scientist and Turing stipend champ John McCarthy was in a perfect world known as disseminated computing over a framework. Undertakings were hunting down some sort of noteworthy course of action, since utility computing ended up getting the chance to be something of a noteworthy business for associations, for instance, IBM. Undoubtedly, Martin Greenberger pointed out the possibility that "front line arithmetical machines without limits" were right now being used institutionally for consistent figuring and research and additionally for business limits, for instance, accounting and stock. Propel, he predicted his bit of work in which PCs would be broad for all intents and purposes like the genuine power associations running wires wherever in due time. Attempts and customers who have a great deal of data as a rule outsource their profitable data to cloud office with a particular ultimate objective to reduce data organization cost and storeroom spending. In this way, data volume in disseminated stockpiling workplaces is experiencing an enthusiastic addition. Regardless of the way that cloud server providers (CSPs) ensure that their cloud organization is furnished with strong wellbeing endeavours, security and insurance are genuine deterrents keeping the more broad affirmation of disseminated computing organization [1]. A customary way to deal with reduction information spillage is data encryption. In any case, this will make server-side data use, for instance, looking on encoded data, transform into an especially troublesome task. In the present years, experts have proposed various ciphertext look for arrangements [35-38] [43] by joining the cryptography techniques. These systems have been exhibited with provable security, yet their techniques require huge operations and have high time unpredictability. Thusly, past strategies are not proper for the huge data circumstance where data volume is colossal and applications require online data dealing with. Likewise, the association between files is canvassed in the above procedures. The association between reports addresses the properties of the documents and consequently keeping up the relationship is basic to totally express a record. For example, the relationship can be used to express its arrangement. In case a record is self-ruling of some different reports beside those documents that are related to recreations, then it is straightforward for us to express this chronicle has a place with the characterization of the diversions. As a result of the outwardly disabled encryption, this imperative property has been shrouded in the standard methodologies. In this manner, proposing a methodology which can keep up and utilize this relationship to speed the interest stage is appealing. Of course, in light of programming/gear disillusionment, and limit corruption, data list things returning to the customers may contain hurt data or have been contorted by the malicious supervisor or gate crasher. In this way, a certain instrument should be given to customers to check the exactness and summit of the rundown things. In light of a dynamic change in the field of undertakings over past decade, there has been addition looked for after of outsourcing of data over a broad assortment of

framework. With a particular true objective to control this colossal measure of data in monetarily insightful way wander has balanced an overwhelming advancement called conveyed computing that oust the heaviness of data organization. In this data driven condition try tend to store their data onto cloud that includes productive asset of customer data like messages, individual prosperity data et cetera. Disseminated computing is winding up being most essential perspective in the change of information development which offer versatile get to, inescapable, on demand get to and capital utilization saving.

II. LITERATURE REVIEW

Qin Liu et al. proposed Secure and insurance sparing catchphrase look for in [1]. It gives watchword assurance, data insurance and semantic secure by open key encryption. The rule issue of this interest is that the correspondence and computational cost of encryption and disentangling is more.

Ming Li et al. proposed Authorized Private catchphrase Search (APKS) in [2]. It gives watchword assurance, Index and Query Privacy, Fine-grained Search Authorization and Revocation, Multi-dimensional Keyword Search, Scalability and Efficiency. This chase strategy assembles the request viability using attribute chain of significance yet eventually every one of the qualities is not dynamic.

Cong Wang et al in [3] proposed Secure and Efficient Ranked Keyword Search which comprehends planning overhead, data and catchphrase assurance, slightest correspondence and count overhead. It is not significant for various catchphrase looks for, Also there is a slight bit of overhead in record building.

Kui Ren et al. [4] proposed Secured fleecy catchphrase look for with symmetric searchable encryption (SSE). It doesn't reinforce feathery interest with open key based searchable encryption; in like manner it can't play out different watchwords semantic chase. The updates for cushy searchable rundown are not adequately performed.

Ming Li et al. [5] proposed Privacy ensured searchable conveyed stockpiling method. It is executed using SSE, Scalar-Product-Preserving Encryption and Order-Preserving Symmetric Encryption. It supports the security and valuable necessities. This arrangement does not support open key based searchable encryption.

Wei Zhou et al. [6] proposed K-gram based cushioned watchword Ranked Search. In this proprietor make k-gram soft watchword list for records D and tuple <I, D> is exchanged to request server (SS) which is installed to grow channel for size controlling. The encoded record D is exchanged to limit server. In any case, the issue is that, the measure of the k-gram build cushioned catchphrase set depends regarding the jacquard coefficient regard.

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

H. S. Rhee et al. [8] proposed Trapdoor in perceptibility Public-Key Encryption with Keyword Search (IND-PEKS). In this outsourcing is done as SCF-PEKS. It encounters outside attacker using KGA and separating the repeat of occasion of watchword trapdoor. Peng Xu et al. [9] proposed Public-Key Encryption PEFKS with Fuzzy Keyword Search, in this customer makes cushioned catchphrase trapdoor Tw and right watchword trapdoor Kw for W. Customer requests Tw to CS. By then CS checks Tw with feathery watchword list and sends superset of organizing figure messages by Fuzz Test computation that is executed by CS. The customer strategy ExactTest count for affirming ciphertexts with Kw and recuperate the mixed records. The path toward making cushioned watchword list and right catchphrase rundown is troublesome for immense size database.

Ning et al. [10] proposed Privacy Preserving Multi Keyword Ranked Search (MRSE). It is profitable for known figure content model and establishment show over encoded data. It gives low computation and correspondence overhead. The encourage planning is decided for multi-catchphrase look for. The hindrance is that MRSE have minimal standard deviation which diminishes the watchword security.

III. RELATED WORK

A. Secure and privacy preserving keyword search

Qin Liu [16] proposed in this paper the request that gives catchphrase assurance, data insurance and semantic secure by open key encryption. CSP is incorporated into partial decipherment by reducing the correspondence and computational lifted in unraveling process for end customers. The customer shows the watchword trapdoors encoded by users‟ private key to CS (Cloud Server) securely and recuperate the mixed reports.

B. Secure and Efficient Ranked Keyword Search

Cong Wang [17] proposed look which lights up taking care of overhead, data and watchword security, minimum correspondence and count lifted. The data proprietor amass document nearby the catchphrase repeat based significance scores for records. Customer request "w" to cloud server with optional 'k' as Tw using the private key. The cloud server looks for the record with scores and sends encoded archive in light of situated progression.

C. Single Keyword Search over Encrypted data on cloud

Practical searchable encryption contrive consent to a customer to firmly scan for over mixed data through catchphrases without first applying interpreting on it, the proposed methodology reinforce simply normal Boolean watchword look, without getting any genuine nature of the records in the thing. Right when clearly associated in colossal joint data outsourcing cloud condition, they encounter next lack.

D. Privacy-preserving Multi-keyword Text Search

Wenhai Sun [19] proposed this request gives equivalence based thing situating, catchphrase assurance, Index and Query protection and Query Unlink limit. The encoded archive is worked by vector space indicate supporting set and specific record look. The searchable record is amassed using Multidimensional B tree. Proprietor makes mixed request vector $\bar{Q}$ for record watchword set. Customer gets the individual mixed question vector of W from proprietor which is given to CS. By and by CS looks list by Merkle–Damgård advancement estimation and contemplates cosine measure of archive and request vector and returns best k encoded records to customer.

E. Secure Multi-keyword Top-k Retrieval Search

Jiadi [20] proposed this interest using two round searchable encryption (TRSE). In first round, customers presents different catchphrase "REQ" 'W‟ as mixed request for satisfying data, watchword security and make trapdoor (REQ, PK) as Tw and sends to cloud server. By then cloud server determine the score from encoded petition for records and returns the mixed score result vector to customer. In second round, customer interpret N with riddle key and figures the archive situating and after that request records with Top k scores. The situating of record is done on client side and scoring is done on server side.

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

Ning [21] proposed this output for known figure content model and establishment show over mixed data giving low estimation and correspondence overhead. The orchestrate planning is chosen for multi-watchword look for. They used internal thing similarity to quantitatively survey equivalence for situating records. The drawback is that MRSE have minimal standard deviation σ which cripples catchphrase security.

G. Attribute-based Keyword Search

Wenhai Sun [22] proposed Attribute-based Keyword Search that gives conjunctive watchword look for; catchphrase semantic security and Trapdoor unlink limit. The proprietors makes list with all watchwords and get the chance to list with course of action qualities which demonstrates the customers list affirmed for looking for. Directly proprietors scramble the report, list with get to rundown using ciphertext approach property based encryption technique. To have customer cooperation organization, they used go-between re-encryption and drowsy re-encryption systems to share the workload to CS. The customer requests the Tw to CS using its private key. By and by CS recoups Tw and chases the encoded documents and return reports just if the user's qualities in Tw satisfy get to approaches in records which make coarse-grained dataset look for endorsement.

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

This proposed method has described and handled the issue of convincing however shielded and sound rank watchword look for over Encrypted cloud data [23]. Situated look massively redesigns system convenience by giving back the organizing records in a situated mastermind as for certain basic criteria (e.g. watchword repeat) thusly making one phase closer towards sensible use of secure data encouraging organizations in Cloud Computing. These papers has described and handled the testing issue of security sparing and gainful multi watchword situated look for over mixed cloud data stockpiling (MRSE), and set up a plan of strict security necessities for such a guaranteed cloud data utilization system to twist up unmistakably a reality. The proposed situating system ends up being gainful to retreat to an extraordinary degree vital files contrasting with submitted look for terms. Proposed situating strategy is used as a piece of our future system with a particular ultimate objective to enhance the security of information on Cloud Service Provider.

I. A Secure and Dynamic Multi-keyword Ranked Search Scheme over Encrypted Cloud DataThis proposed technique [24] suggest an ensured tree-based chase plot over

the encoded conveyed capacity, which reinforces multi watchword situated look for close by component operation on report gathering open at server. The vector space model and term repeat (TF) × inverse file repeat (IDF) model are combinly used as a piece of the advancement of document and period of request to give multi watchword situated look for yield. To get high interest capability comes to fruition, maker build up a tree-based record structure and proposed a Greedy Depth-first Search count in perspective of this rundown tree. In light of this remarkable structure of tree-based record, the proposed look plan can adaptably fulfill sub straight request time and can effectively deal with the eradication and expansion of documents. The kNN estimation is associated with scramble the document and request vectors, and till then assurance exact relevance score number between encoded rundown and question vectors.

IV. PROPOSED WORK

Cloud information proprietors want to outsource reports in an encoded frame with the end goal of security saving. Along these lines it is fundamental to create productive and dependable ciphertext look methods. One test is that the connection between archives will be typically disguised during the time spent encryption, which will prompt huge inquiry exactness execution debasement. Likewise the volume of information in server farms has encountered a sensational development. This will make it considerably all the more difficult to plan ciphertext seek conspires that can give productive and dependable online data recovery on expansive volume of scrambled information.

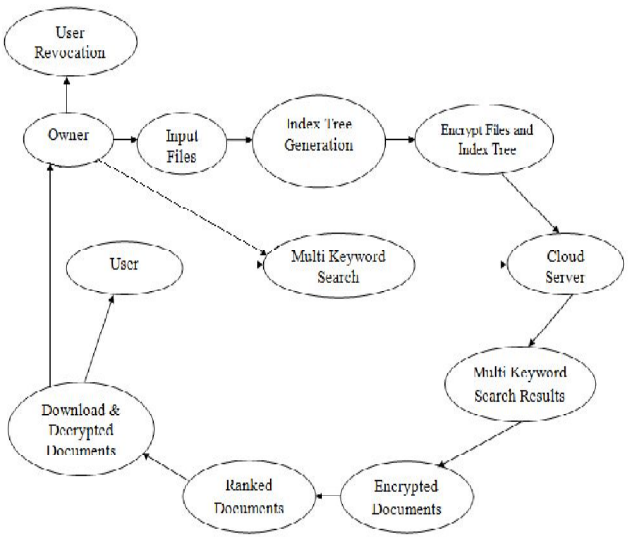

Figure 1: System Architecture

The cutting edge web called the Semantic Web will help the client to recover the valuable information that is put away on the cloud as metaphysics and make the information obvious to the client which is holed up behind the cloud. The point of the proposed positioning calculation is to give clients the outcome set of pertinent information.

A various levelled bunching strategy is proposed to bolster more pursuit semantics and furthermore to take care of the demand for quick figure content hunt inside a major information condition. The proposed progressive approach groups the reports in light of the base significance limit, and after that parcels the subsequent bunches into sub-bunches until the imperative on the most extreme size of bunch is come to. In the pursuit stage, this approach can achieve a straight computational intricacy against an exponential size increment of archive gathering. So as to check the legitimacy of list items, a structure called least hash sub-tree is planned. We expand this thought of semantic likeness to consider intrinsic connections between ideas utilizing ontologies. We propose positioning calculation with multi catchphrase and cosmology.

V. CONCLUSIONS

This paper concentrates different strategies of looking in the encoded cloud information stockpiling. We have methodically presents the security and information usage issues in the distributed storage identified with all accessible seeking procedures. Hence distinguished the primary issues that are to be fulfilled for secured information use are catchphrase protection, Data security, Index security, Query Privacy, Fine-grained Search, Scalability, Efficiency, Result positioning, Index secrecy, Query classification, Query Unlinkability, semantic security and Trapdoor Unlinkability. The vast majority of the looking strategies mostly concentrate on security and some on information use. The impediments of all the seeking procedures are additionally talked about. By the above review, security can be given by Public-Key Encryption and powerful information use by fluffy catchphrase look. We trust that this overview will make the specialists to shape their issue in the range of information use in distributed storage.

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,China, May 2013.

[20] 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.

[21] 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

[22] 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

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

[24] 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