IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN : 2278-0661, p-ISSN : 2278-8727
PP 28-32
www.iosrjournals.org

A Review Paper on Multi keyword Ranked Search on Encrypted Cloud Data

Purva Jain[1], Dr. Abhijit Banubakode[2]
(Computer, RSCOE, University of Pune, Pune, Maharashtra India)[1]
(*I.T., RSCOE, University of Pune, Pune, Maharashtra India*)[2]

Abstract: *Because of the expanding prominence of distributed computing, more information proprietors are inspired to outsource their information to cloud servers for extraordinary accommodation and diminished expense in information administration. Then again, delicate information ought to be scrambled before outsourcing for security prerequisites, which obsoletes information use like catchphrase based report recovery. In this paper, we show a safe multi-essential word positioned inquiry plan over encoded cloud information, which at the same time bolsters element overhaul operations like cancellation and insertion of archives. Specifically, the vector space model and the broadly utilized TF×IDF model are joined as a part of the record development and question era. We build a unique tree-based file structure and propose an "Avaricious Depth-first Search" calculation to give efficient multi-magic word positioned inquiry. The protected kNN calculation is used to scramble the file and question vectors, and in the interim guarantee exact pertinence score count between encoded list and inquiry vectors. With a specific end goal to oppose factual assaults, apparition terms are added to the list vector for blinding indexed lists. Because of the utilization of our exceptional tree-based file structure, the proposed plan can accomplish sub-direct inquiry time and manage the erasure and insertion of archives flexibly. Broad analyses are led to exhibit the efficiency of the proposed scheme.*

Keywords*: Searchable encryption, multi-keyword ranked search, dynamic update, cloud computing.*

I. Introduction

CLOUD computing has been considered as another model of enterprise IT infrastructure, which can compose gigantic resource of computing, storage and applications, and empower users to appreciate pervasive, helpful and on-demand network access to a mutual pool of configurable computing resources with incredible efficiency and insignificant economic overhead [1]. Pulled in by these engaging features, both individuals and enterprises are roused to outsource their data to the cloud, rather than buying software and hardware to deal with the data themselves.

In spite of the different points of interest of cloud services, outsourcing delicate information, (for example, e-mail, individual health records, organization account information, government archives, and so forth.) to remote servers brings privacy concerns. The cloud service providers (CSPs) that keep the data for users may access users' sensitive information without authorization.

A general way to deal with secure the data privacy is to encrypt the data before outsourcing [2].
On the other hand, this will bring about a gigantic expense in terms of data ease of use. For example, the current techniques on keyword-based information retrieval, which are broadly utilized on the plaintext data, can't be straightforwardly connected on the encrypted data. Downloading all the data from the cloud and decrypt locally is clearly unrealistic.

With a particular final objective to address the above issue, analysts have illustrated some all around helpful arrangements with totally homomorphic encryption [3] or missing RAMs [4]. In any case, these schedules are not down to earth in light of their high computational overhead for both the cloud server and user. In spite of what may be normal, more useful unique reason arrangements, for instance, searchable encryption (SE) plan have made specific responsibilities to the extent productivity, value and security. Searchable encryption scheme engage the user to store the encrypted data to the cloud and execute unequivocal word look for over ciphertext domain. As being what is indicated, abundant works have been proposed under assorted risk models to finish distinctive interest value, for instance, single keyword search, closeness look, multi-keyword boolean search, ranked search, multi-keyword ranked search, etc. Among them, multikeyword positioned quest finishes more thought for its pragmatic propriety. Starting late, some component arrangements have been proposed to reinforce embedding and erasing operations on archive gathering. These are colossal goes about as it is exceptionally possible that the data owner need to overhaul their data on the cloud server. Yet, few of the dynamic plan support successful multikeyword situated look. Inverse document recurrence (IDF)" model are

joined in the list development and inquiry era to give multikeyword positioned seek. Keeping in mind the end goal to get high search

Effectiveness, we develop a tree based list structure and based on this tree list we propose a "Greedy Depth –first Search" calculation. Because of the uncommon structure of our tree-based list, the proposed search scheme can flexibly accomplish sub-straight search time and manage the deletion and insertion of reports. The protected kNN algorithm is used to encrypt the index and query vectors, and in the interim guarantee relevance score calculation between encrypted index and query vectors. To oppose distinctive attacks in different threat models, we build two secure search schemes: the basic dynamic multi-keyword ranked search (BDMRS) scheme in the known ciphertext model, and the enhanced dynamic multi-keyword ranked search (EDMRS) scheme in the known background model. Our commitments are condensed as takes after:

1) We design a searchable encryption scheme that underpins both the precise multi-keyword ranked search and flexible dynamic operation on document collection.

2) The proposed scheme is in a general sense kept to logarithmic for search complexity in the uncommon structure of tree-based index. What's more, practically speaking, the proposed scheme can accomplish higher search proficiency by executing our "Greedy Depth-first Search" algorithm. Additionally to reduce the time cost of search process parallel search can performed.

II. Literature Survey

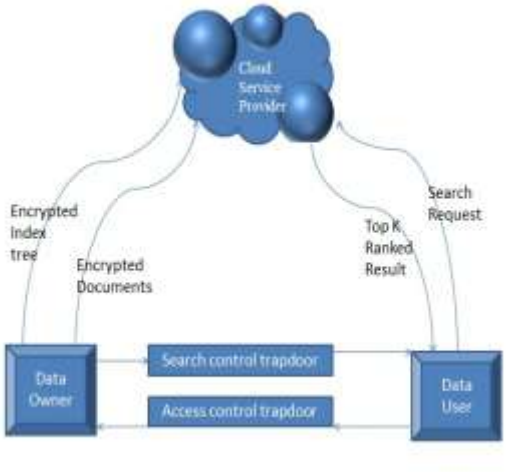
cloud computing transforms the way information technology(IT) is expended and oversaw, promising enhanced expense efficiencies, quickened development, speedier time-to-market, and the capacity to scale applications on interest (Leighton, 2009).[1] As per Gartner, while the buildup developed exponentially amid 2008 and proceeded since, it is clear that there is a noteworthy movement towards the cloud computing model and that the advantages may be significant (Gartner Hype-Cycle, 2012). Be that as it may, as the cloud's state processing is rising and growing quickly both theoretically and actually, the legitimate/contractual, monetary, administration quality, inter-operability, security and protection issues still posture critical difficulties. In this part, we depict different services and organization models of distributed computing and recognize significant difficulties.

[2]We consider the issue of building a safe cloud storage services on top of an open cloud foundation where the service provider is not totally trusted by the user. We depict, at an abnormal state, a few architectures that consolidate late and non-standard cryptographic primitives with a specific end goal to accomplish our objective. We review the benefits such a construction modeling would give to both customers and service providers and give an outline of late advances in cryptography roused specifically by cloud storage.

We propose the first completely homomorphic encryption scheme, taking care of a focal open issue in cryptography. Such a plan permits one to figure subjective capacities over encypted data without the decoding key – i.e., given encryptions E(m1),...,E(mt) of m1,...,mt, one can efficiently process a smaller ciphertext that encrypts f(m1,...,mt) for any efficiently calculable capacity f. This issue was postured by Rivest et al. in 1978. [3]Completely homomorphic encryption has various applications. For instance, it empowers private queries to a search engine– the user presents an encypted query and the search engine processes a brief encrypted answer while never taking a gander at the query in the clear. It likewise empowers looking on encrypted data – a user stores encypted files on a remote file server and can later have the server recover just files that (when decoded) fulfill some boolean limitation, despite the fact that the server can't unscramble the files all alone. All the more comprehensively, completely homomorphic encryption enhances the efficiency of secure m.

We concentrate on the issue of looking on data that is encrypted using a public key system. [5] Consider user Bob who sends email to user Alice encrypted under Alice's public key. An email portal needs to test whether the email contains the keyword "urgent" so that it could course the email in like manner. Alice, then again does not wish to give the entryway the capacity to decrypt every one of her messages. We build a component that empowers Alice to give a key to the passage that empowers the entryway to test whether "urgent" is a keyword in the email without learning whatever else about the email. We allude to this component as Public Key Encryption with keyword Search.

III. System Architecture:



The system model in this paper incorporates three unmistakable substances: data owner, data user and cloud server, as illustrated in Fig. 1.

Data owner has a gathering of records F = {f1; f2; :::; fn} that he needs to outsource to the cloud server in encoded structure while up 'til now keeping the ability to check on them for convincing utilization. data owner firstly manufactures a secure searchable tree index I from archive accumulation F, and a short time later makes a encrypted document gathering C for F. A brief span later, the data owner outsources the encoded accumulation C and the secure index I to the cloud server, and safely disseminates the key data of trapdoor era and document decryption to the approved data users. Additionally, the data owner observe his documents those are stored on cloud server. When updating, the data owner creates the upgradable data locally and sends it to the server.
Data users are approved ones to get to the archives of data owner. With t query keywords, the approved user can create a trapdoor TD as indicated by search control mechanisms to get k encrypted documents from cloud server. By then, documents are decrypt using shared secret key.

Cloud server stores the encrypted document accumulation C and the encrypted searchable tree index I for data owner. In the wake of tolerating the trapdoor TD from the data user, look over the index tree I, in conclusion gives back the relating gathering of top-k situated encoded reports. Also, in the wake of tolerating the update information from the data owner, the server needs to update the index I and document gathering C as per the received information.

3.1) MODULES:

3.1.1) Index Construction of UDMRS Scheme

Amid the procedure of index development, we to begin with make a tree node for each document in the accumulation. These nodes are the leaf nodes of the index tree. By then, the internal tree nodes are made in view of these leaf nodes.

3.1..2)Search Process of UDMRS Scheme

The search procedure of the UDMRS scheme is a recursive methodology upon the tree, named as "Greedy Depth first Search (GDFS)" algorithm. We add to an outcome list meant as RList, whose components is described as ⟨RScore; FID⟩. Here, the RScore is the significance score of the archive fFID to the question. The RList stores the k got to reports with the biggest pertinence scores to the inquiry. The rundown's components are positioned in sliding request as indicated by the RScore, and will be upgraded opportune amid the search process.

3.1.3)BDMRS Scheme

In view of the UDMRS scheme, we build the essential element multi-keyword ranked search (BDMRS) scheme by utilizing the secure kNN algorithm [5]. The BDMRS scheme is intended to accomplish the objective of privacy preserving in the known ciphertext model. BDMRS scheme can secure the Index Confidentiality and Query Confidentiality in the known ciphertext model [6], [7], [8].

3.1.4) DMRS Scheme

Cloud server has the capacity interface the same search requests by following way of visited nodes. The Cloud server recognize a keyword as the standardized TF distribution of the keyword can be precisely acquired from the last computed relevance scores. A heuristic strategy to further enhance the security is to break such correct quality. Hence, we can acquaint some tunable haphazardness with exasperate the significance score estimation. Likewise, to suit diverse users' inclinations for higher exact positioned results or better protected keyword privacy, the arbitrariness are set movable.

3.1.5)Dynamic Update Operation of DMRS

After insertion or deletion of a record, we require to update synchronously the index. Since the index of DMRS scheme is planned as a balanced binary tree, the dynamic operation is done by redesigning hubs in the list tree. The report on record is just in view of archive recognizes, and no entrance to the substance of records is required.

IV. Conclusion

In this paper, a safe, effective and dynamic search scheme is proposed, which underpins the exact multi-keyword ranked search as well as the dynamic deletion and insertion of documents. We assemble a special keyword balanced binary tree as the index, and intend a "Greedy Depth-first Search" algorithm to acquire preferable effectiveness over linear search. Likewise, the parallel search procedure can be completed to further lessen the time cost. The plan's security is ensured against two risk models by utilizing the safe kNN algorithm. Trial results display the efficiency of our proposed scheme. In the proposed scheme, the information proprietor is in charge of producing overhauling data and sending them to the cloud server. Accordingly, the data owner needs to store the un-encrypted index tree and information that is required to recalculate the IDF values.

Such a active data owner may not be astoundingly suitable for the appropriated distributed computing model. It could be an important yet troublesome future work to design a dynamic searchable encryption scheme whose updating operation are performed by cloud server. Furthermore, as the large portion of works about searchable encryption, our scheme chiefly considers the test from the cloud server.

References

[1] K. Ren, C. Wang, Q. Wang et al., "Security challenges for the public cloud," IEEE Internet Computing, vol. 16, no. 1, pp. 69–73, 2012.
[2] S. Kamara and K. Lauter, "Cryptographic cloud storage," in Financial Cryptography and Data Security. Springer, 2010, pp. 136–149.
[3] C. Gentry, "A fully homomorphic encryption scheme," Ph.D. dissertation, Stanford University, 2009.
[4] 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.
[5] D. Boneh, G. Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Advance in Cryptology Eurocrypt 2004.Springer,2004 pp. 506–522.
[6] 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.
[7] 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.
[8] 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.

About Author

Dr. Abhijit Banubakode received Ph.D. degree in Computer Studies from Symbiosis Institute of Research and Innovation (SIRI), a constituent of Symbiosis International University (SIU), Pune, India in April 2014 and ME degree in Computer Engineering from Pune Institute of Computer Technology(PICT), University of Pune ,India in 2005 and BE degree in Computer Science and Engineering from Amravati University, India, in 1997. His current research area is Query Optimization in Compressed Object Oriented Database Management Systems (OODBMS). Currently he is working as Professor and Head of Department (HOD) in Department of Information Technology, Rajarshi Shahu College of Engineering,Pune, India. He is having 16 years of teaching experience. He is a member of International Association of Computer Science and Information Technology (IACSIT), ISTE, CSI and presented 12 papers in International journal and conference**.**

Purva Jain received B.TECH degree in Information Technology and Engineering from J.N.I.T. College of Engineering Jaipur, India in 2011 and pursuing ME degree in Computer Science and Engineering from Rajarshi Shahu College of Engineering, and Pune, India.