International Journal of Contemporary Research in Computer Science and Technology (IJCRCST) ISSN: 2395-5325
Volume 4, Special Issue 3 (March '2018)
Proceeding of - National Conference on Recent Trends in Information Technology and Communication,
Conference held at Velammal Engineering College,Chennai,India.

EFFICIENT PRIVACY-PRESERVING RANKED KEYWORD SEARCH METHOD

Vijayan.A,
Assistant Professor,
Department of Information Technology,
Velammal Engineering College,
Chennai,Tamilnadu,India.

Priyadarshini.S,
Student,
Department of Information Technology,
Velammal Engineering College,
Chennai,Tamilnadu,India.

Rubitha.S,
Student,
Department of Information Technology,
Velammal Engineering College,
Chennai,Tamilnadu,India.

Abstract: Cloud information proprietors like to outsource archives in a scrambled frame with the end goal of protection saving. Thusly it is basic to create productive and solid ciphertext look methods. One test is that the connection between archives will be regularly disguised during the time spent encryption, which will prompt noteworthy inquiry exactness execution debasement. Likewise, the volume of information in server farms has encountered an emotional development. This will make it much additionally difficult to plan ciphertext seek conspires that can give effective and solid online data recovery on expansive volume of encoded information. In this paper, a progressive grouping strategy is proposed to help more hunt semantics and furthermore to take care of the demand for quick ciphertext seek inside a major information condition. The proposed various levelled approach bunches the records in light of the base importance limit, and after that segments the subsequent groups into sub-bunches until the point when the limitation on the most extreme size of bunch is come to. In the hunt stage, this approach can achieve a straight computational intricacy against an exponential size increment of report gathering. With a specific end goal to check the genuineness of indexed lists, a structure called least hash sub-tree is outlined in this paper. Investigations have been led utilizing the accumulation set worked from the IEEE Xplore. The outcomes demonstrate that with a sharp increment of reports in the dataset the pursuit time of the proposed technique increments straight though the hunt time of the conventional strategy increments exponentially. Moreover, the proposed technique has leeway over the customary strategy in the rank protection and significance of recovered reports.

Keywords: Cloud computing, ciphertext search, ranked search, multi-keyword search, hierarchical clustering, security.

I.INTRODUCTION

As we venture into the enormous information time, terabyte of information are created overall every day. Endeavours and clients who claim a lot of information as a rule outsource their valuable information to cloud office keeping in mind the end goal to decrease 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 benefit is outfitted with solid safety efforts, security and protection are significant impediments keeping the more extensive acknowledgment of distributed computing administration. A customary method to decrease data spillage is information encryption. In any case, this will make server -side information usage, for example, seeking on scrambled information, turn into an extremely keep up and use this relationship to speed the hunt stage is attractive.

Then again, because of programming/equipment disappointment, and capacity defilement, information list items coming back to the clients may contain harmed information or have been mutilated by the malignant chairman or interloper. Subsequently, an evident system ought to be given to clients to confirm the accuracy and culmination of the query items.

In this paper, a vector space display is utilised and each record is spoken to by a vector, which implies each report can be viewed as a point in a high dimensional space. Because of the connection between various reports, every one of the archives can be partitioned into a few classes. At the end of the day, the focuses whose separation are short in the high dimensional space can be ordered into a particular class. The pursuit time can be to a great extent decreased by choosing the coveted classification and deserting the superfluous classifications. Contrasting and every one of the records in the dataset, the quantity of archives which client goes for is little. Because of the modest number of the coveted archives, a particular class can be additionally separated into a few sub-classifications. Rather than utilizing the conventional succession look technique, a backtracking calculation is delivered to look through the objective archives. Cloud server will first hunt the classifications and

get the base wanted sub-class. At that point the cloud server will choose the coveted k reports from the base wanted sub-classification. The estimation of k is already chosen by the client and sent to the cloud server. In the event that present sub-class cannot fulfil the k archives, cloud server will follow back to its parent and select the coveted records from its sibling classifications.

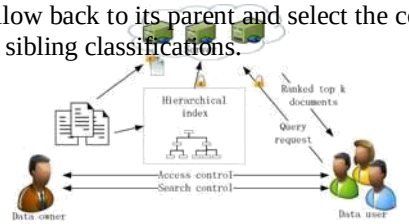

Figure1: Architecture of ciphertext search.

This procedure will be executed recursively until the point when the coveted k reports are fulfilled, or the root is come to. To confirm the respectability of the query item, an undeniable structure in view of hash work is developed. Each report will be hashed and the hash result will be utilized to speak to the record. The hashed consequences of archives will be hashed again with the classification data that these records have a place with and the outcome will be utilized to speak to the present class. Likewise, every class will be spoken to by the hash consequence of the mix of current classification data and sub-classes data. A virtual root is built to speak to every one of the information and classifications. The virtual root is indicated by the hash aftereffect of the link of the considerable number of classes situated in the primary level. The virtual root will be marked so it is unquestionable. To check the query item, client just needs to confirm the virtual root, rather than confirming each record.

Single Keyword Searchable Encryption

They propose to encode each word in the record freely. This technique has a high looking expense because of the filtering of the entire information gathering word by word. Goh formally characterized a protected list structure and detail a security display for list known as semantic security against versatile picked catchphrase assault (ind-cka). They likewise built up a productive ind-cka secure record development called z-idx by utilizing pseudo-arbitrary capacities and sprout channels. Because of the absence of rank component, clients need to set aside a long opportunity to choose what they need when monstrous archives contain the inquiry catchphrase. In this way, the request protecting systems are used to understand the rank instrument, In the pursuit stage, the cloud server processes the importance score amongst records and the question. Along these lines, pertinent archives are positioned by their significance score and clients can get the best k comes about. In people in general key setting, composed the principal accessible encryption development, where anybody can utilize open key to keep in touch with the information put away on server however just approved clients owning private key can seek. In any case, all the previously mentioned procedures just help single catchphrase seek.

Multiple Keyword Searchable Encryption

To advance inquiry predicates, an assortment of conjunctive watchword looked strategies have been proposed. These techniques demonstrate extensive overhead, for example, correspondence cost by sharing mystery, or computational cost by bilinear guide. String et al propose a safe hunt plot in view of vector space show. Because of the absence of the security examination for recurrence data and pragmatic pursuit execution, it is hazy whether their plan is secure and productive or not. Cao et al introduced a novel design to take care of the issue of multi-watchword positioned seek over encoded cloud information. Be that as it may, the pursuit time of this strategy becomes exponentially going with the exponentially expanding size of the archive accumulations. Be that as it may, at the phase of record building process, the importance between archives is disregarded. Therefore, the significance of plaintexts is hidden by the encryption, clients desire can't be satisfied well. For instance: given an inquiry containing Mobile and Phone, just the records containing both of the watchwords will be recovered by conventional strategies. Be that as it may, if taking the semantic connection between the archives into thought, the reports containing Cell and Phone ought to likewise be recovered. Clearly, the second outcome is better at meeting the clients desire.

Verifiable Search Based on Authenticated Index

The possibility of information check has been all around considered in the region of databases. In a plaintext database situation, an assortment of techniques have been created. The majority of these works depend on the first work by Merkle, and refinements by Naor and Nissm for declaration denial. Merkle hash tree and cryptographic mark procedures are utilized to develop confirmed tree structure whereupon end clients can check the accuracy and culmination of the question comes about. Throb and Mouratidis apply the Merkle hash tree in view of confirmed structure to content web crawlers. Notwithstanding, they just spotlight on the confirmation particular issues overlooking the pursuit security saving abilities that will be tended to in this paper. The hash fasten is utilized to develop a solitary watchword item confirmation plot by Wang et al. Be that as it may, their work can't be straightforwardly utilized as a part of our design which is situated for security saving various watchword seek. Along these lines, an appropriate instrument that can be utilized to check the list items inside huge information situation is basic to both the CSPs and end clients.

II.RELATED WORK

A number of well -known authentication protocols are considered in the context of next-generation mobile and CE network services. The potential weaknesses of current protocols can be overcome using Zero Knowledge Proof (ZKP) techniques to protect user passwords so an alternative ZKP protocol, SeDiCi 2.0, is described. This offers mutual and also two-factor authentication that is considered more secure against various phishing attempts than existing trusted third party protocols. The suitability of such a ZKP protocol for various CE-based cloud computing applications is demonstrated. It is desirable to store data on data storage

servers such as mail servers and file servers in encrypted form to reduce security and privacy risks. But this usually implies that one has to sacrifice functionality for security. For example, if a client wishes to retrieve only documents containing certain words, it was not previously known how to let the data storage server perform the search and answer the query, without loss of data confidentiality. We describe our cryptographic schemes for the problem of searching on encrypted data and provide proofs of security for the resulting crypto systems. Our techniques have a number of crucial advantages. They are provably secure: they provide provable secrecy for encryption, in the sense that the untrusted server cannot learn anything about the plaintext when only given the ciphertext; they provide query isolation for searches, meaning that the untrusted server cannot learn anything more about the plaintext than the search result; they provide controlled searching, so that the untrusted server cannot search for an arbitrary word without the user's authorization; they also support hidden queries, so that the user may ask the untrusted server to search for a secret word without revealing the word to the server. The algorithms presented are simple, fast (for a document of length n, the encryption and search algorithms only need O(n) stream cipher and block cipher operations), and introduce almost no space and communication overhead, and hence are practical to use today. The above paper has the following limitations:

- If you are stored data in the cloud after you want any file it takes more time because we are not assign index.
- If you kept any file in the cloud not provides security.

We proposed the MRSE-HCI architecture to speed up server-side searching phase. Accompanying with the exponential growth of document collection, the search time is reduced to a linear time instead of exponential time.

III.PROPOSED SYSTEM

In this paper, we propose a multi-keyword ranked search over encrypted data based on hierarchical clustering index (MRSE-HCI) to maintain the close relationship between different plain documents over the encrypted domain in order to enhance the search efficiency. In the proposed architecture, the search time has a linear growth accompanying with an exponential growing size of data collection. We propose a hierarchical method in order to get a better clustering result within a large amount of data collection. The size of each cluster is controlled as a trade-off between clustering accuracy and query efficiency. According to the proposed method, the number of clusters and the minimum relevance score increase with the increase of the levels whereas the maximum size of a cluster reduces. Depending on the needs of the grain level, the maximum size of a cluster is set at each level. Every cluster needs to satisfy the constraints. If there is a cluster whose size exceeds the limitation, this cluster will be divided into several sub-clusters.proposed engineering not just legitimately understands the multi-watchword positioned seek issue, yet additionally acquires a change look effectiveness, rank security, and the importance between recovered reports.

IV. REFERENCES

[1]. S. Grzonkowski, P. M. Corcoran, and T. Coughlin, "Security analysis of authentication protocols for next-generation mobile and CE cloud services," in Proc. IEEE Int. Conf. Consumer Electron., 2011, Berlin, Germany, 2011, pp. 83–87.

[2]. D. X. D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. IEEE Symp. Security Priv.,BERKELEY, CA, 2000, pp. 44–55.

[3]. D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. EUROCRYPT, Interlaken, SWITZERLAND, 2004, pp. 506– 522.

[4]. Y. C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in Proc. 3rd Int. Conf. Applied Cryptography Netw. Security, New York, NY, 2005, pp. 442–455.

[5]. R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in Proc. 13th ACM Conf. Comput. Commun. Security, Alexandria, Virginia, 2006, pp. 79–88.

[6]. M. Bellare, A. Boldyreva, and A. O'Neill, "Deterministic and efficiently searchable encryption," in Proc. 27th Annu. Int. Cryptol. Conf. Adv. Cryptol., Santa Barbara, CA, 2007, pp. 535–552.

[7]. D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in Proc. 4th Conf. Theory Cryptography, Amsterdam, NETHERLANDS, 2007, pp. 535–554.

[8]. E.-J. Goh, Secure Indexes, IACR Cryptology ePrint Archive, vol. 2003, pp. 216. 2003.

[9]. 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.

[10]. A. Swaminathan, Y. Mao, G. M. Su, H. Gou, A. Varna, S. He, M. Wu, and D. Oard, "Confidentiality-preserving rank-ordered search," in Proc. ACM ACM Workshop Storage Security Survivability, Alexandria, VA, 2007, pp. 7–12.

[11]. S. Zerr, D. Olmedilla, W. Nejdl, and W. Siberski, "Zerber+R: Topk retrieval from a confidential index," in Proc. 12th Int. Conf. Extending Database Technol.: Adv. Database Technol., Saint Petersburg, Russia, 2009, pp. 439–449.

[12]. C. Wang, N. Cao, J. Li, K. Ren, and W. J. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. IEEE 30th Int. Conf. Distrib. Comput. Syst., Genova, ITALY, 2010, pp. 253–262.

[13]. P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in Proc. Proc. 2nd Int. Conf. Appl. Cryptography Netw. Security, Yellow Mt, China, 2004, pp. 31–45.

[14]. L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword

searches over encrypted data," in Proc. 7th Int. Conf. Inform. Commun. Security, Beijing, China, 2005, pp. 414–426.

[15]. R. Brinkman, "Searching in encrypted data" in University of Twente, PhD thesis, 2007

[16]. Y. H. Hwang and P. J. Lee, "Public key encryption with conjunctive keyword search and its extension to a multi -user system," in Proc. 1st Int. Conf. Pairing-Based Cryptography, Tokyo, JAPAN, 2007, pp.2–22.

[17]. H. Pang, J. Shen, and R. Krishnan, "Privacy-preserving similaritybased text retrieval," ACM Trans. Internet Technol., vol. 10, no. 1, pp. 39, Feb. 2010.

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

[19]. 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. Inform., Comput. Commun. Security, Hangzhou, China, 2013, pp. 71–82.

[20]. F. Li, M. Hadjieleftheriou, G. Kollios, and L. Reyzin, "Dynamic authenticated index structures for outsourced databases," in Proc. ACM SIGMOD, Chicago, IL, 2006, pp. 121– 132.

[21]. S. Kamara, C. Papamanthou, and T. Roeder, "Dynamic searchable symmetric encryption," in Proc. Conf. Comput.Commun. Secur., 2012, pp. 965–976.

[22]. R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: Improved definitions and efficient construc-tions," M. Chase and S. Kamara, "Structured encryption and controlled disclosure," in Proc. Adv. Cryptol., 2010, pp. 577–594.

[23]. D. Cash, J. Jaeger, S. Jarecki, C. Jutla, H. Krawczyk, M. C. Rosu, and M.