International Conference on Engineering & Technology (ICET-2020)

Available at: *International Journal of Scientific Research and Engineering Development*

RESEARCH ARTICLE OPEN ACCESS

PRIVACY-PRESARVING KEYWORD SEARCHING OVER ENCRYPTED DATA IN CLOUD COMPUTING

Mr.P.Rajendran[1], M.E., Akhilesh kumar[2], P.Dhineshkumar[3], S.Parthiban[4]

[1] Assistant professor, [234] UG students

[1234] Department of CSE Selvam College of technology, Namakkal.

[1] rajendran.cse@selvamtech.edu.in [2] akhileshdev329@gmail.com

[3] dhikumar.cs4@gmail.com [4] parthasundaraj98@gmail.com

----------**********----------

ABSTRACT:

With the appearance of distributed computing, information proprietors are animated to contract out their composite information official frameworks from nearby locales to the exchange open cloud for extraordinary give and money related investment funds. Therefore, empowering a scrambled cloud information investigate administration is of predominant significance. Taking into account the huge number of insights clients and passage license in the cloud, it is required to permit different catchphrases in the pursuit request and return grant in the request for their essentialness to these watchwords. Right now, the first occasion when, we portray and tackle the requesting issue of protection saving multi-catchphrase positioned search over encoded information in distributed computing (MRSE). Among a decision of multi-watchword semantics, we choose the capable comparability decide of "arrange coordinating," i.e., as a ton of matches as could be expected under the circumstances, to proficiently catch the essentialness of re-appropriated accreditations to the inquiry catchphrases, and use "internal item similitude" to quantitatively gauge such correspondence measure.

Keywords — ***MRSE, keyword search, coordinate matching.***

----------**********----------

1. INTRODUCTION

Cloud Computing gets pervasive, increasingly more delicate data are being incorporated into the cloud, for example, messages, individual wellbeing records, government archives, and so on. By putting away their information into the cloud, the information proprietors can be calmed from the weight of information stockpiling and upkeep in order to appreciate the on-request great information stockpiling administration. Nonetheless, the way that information proprietors and cloud server are not in the equivalent believed space may put the redistributed information in danger, as the cloud server may never again be completely trusted. It follows that touchy information normally ought to be scrambled before redistributing for information security and fighting spontaneous gets to.

Data encryption makes compelling information use a difficult errand given that there could be a lot of redistributed information records. Besides, in Cloud Computing, information proprietors may impart their redistributed information to an enormous number of clients. The individual clients may need to just recover certain particular information documents they are keen on during a given meeting.

One of the most mainstream ways is to specifically recover documents through catchphrase based hunt as opposed to recovering all the scrambled records back which is totally illogical in distributed computing situations. Such watchword based hunt system permits clients to specifically recover documents of intrigue and has been generally applied in plaintext search situations, for example, Google search. Sadly, information encryption confines client's capacity to perform catchphrase search and consequently makes the conventional plaintext look strategies unsatisfactory for Cloud Computing. Other than this, information encryption additionally requests the security of watchword protection since catchphrases typically contain significant data identified with the information records.

Searchable encryption (SE) is a hot research field, particularly with the development of distributed computing. Right now, audit and investigate the current accessible encryption plans. SE can be isolated into open key accessible encryption and symmetric accessible encryption (SSE) as indicated by various cryptography natives. Right now, center around the symmetric accessible encryption since open key accessible encryption as a rule is keen on during a given meeting.

One of the most mainstream ways is to specifically recover documents through catchphrase based hunt as opposed to recovering all the scrambled records back which is totally illogical in distributed computing situations. Such watchword based hunt system permits clients to specifically recover documents of intrigue and has been generally applied in plaintext search situations, for example, Google search. Sadly, information encryption confines client's capacity to perform catchphrase search and consequently makes the conventional plaintext look strategies unsatisfactory for Cloud Computing. Other than this, information encryption additionally requests the security of watchword protection since catchphrases typically contain significant data identified with the information records. Searchable encryption (SE) is a hot research field, particularly with the development of distributed computing. Right now, audit and investigate the current accessible encryption plans. SE can be isolated into open key accessible encryption and symmetric accessible encryption (SSE) as indicated by various cryptography natives. Right now, center around the symmetric accessible encryption since open key accessible encryption as a rule is computationally costly. Rich works are proposed to manage symmetric accessible encryption. Melody et al. First characterized the issue of looking on scrambled information and proposed a symmetric accessible encryption conspire with direct intricacy.

We first propose the random traversal algorithm which makes the cloud server randomly traverse on index and returns different results for the same query, and in the meantime, it maintains the accuracy of queries unchanged for higher security. Based on the random traversal algorithm, we present one both efficient and secure searchable encryption scheme, which can support top-*k* similarity search over encrypted data. In this scheme, the data owner can control the level of query unlink ability without sacrificing accuracy.

Our experimental results show that our methods are more efficient than the state-of-the-art methods and can better protect data privacy. Especially, our proposed method has good scalability performance when dealing with large data sets.

2 .RELATED WORKS

Cloud computing has developed as a problematic pattern in both IT enterprises and research networks as of late, its remarkable attributes like high versatility and pay-asyou-go style have empowered cloud shoppers to buy the amazing processing assets as administrations as per their genuine prerequisites, with the end goal that cloud clients have never again need to stress over the squandering on figuring assets and the multifaceted nature on equipment stage the board. Accessible encryption (SE) is a hot research field, particularly with the development of distributed computing. Right now, audit and break down the current accessible encryption plans. SE can be separated into open key accessible encryption and symmetric accessible encryption (SSE) as indicated by various cryptography natives. Right now, centers around the symmetric accessible encryption since open key accessible encryption for the most part are computationally costly. Rich works are proposed to manage symmetric accessible encryption. Melody et al. [6] first characterized the issue of looking on encoded information and proposed a symmetric accessible encryption conspire with direct intricacy. From that point forward, Goh et al. detailed a security definition for SSE and proposed a safe record which depends on pseudo-irregular capacities and Bloom channels, however the time cost of Goh's plan is O(n). Curtmola et al. presented two conventional meanings of SSE and proposed a strategy which depends on modified rundown to improve the question execution, their technique is end up being more proficient than different works. In any case, a large portion of these works can just help single catchphrase boolean hunt, which isn't propelled enough to help complex functionalities. As of late, numerous works have been proposed to accomplish various types of complex inquiries like likeness search, multi-catchphrase search, and so forth. When all is said in done, the written works utilized special case based systems, in view of Bed-tree and applied the area touchy hashing (LSH) to manage similitude search. Works support multi-watchword boolean hunt, yet boolean inquiry is wasteful in light of the fact that it restores all the records that fulfill the question criteria. Thus, some ongoing works are proposed to manage the transmission capacity sparing multi-catchphrase positioned search. Cao et al. proposed the multi-watchword positioned scan over scrambled information just because and manufactured an accessible list dependent on the vector space model, and picked "arrange coordinating" to gauge the comparability among inquiries and reports. Nonetheless, in their plans, the time multifaceted nature of search is O(nm) (n is the quantity of catchphrases in lexicon, m is the size of the reports that put away in the cloud server), and the time intricacy of trapdoor development is additionally extremely high. Sun et al. proposed a tree-based record structure which depends on the vector space model and the TF IDF model. This structure accomplishes sub-straight time multifaceted nature, yet it is helpless in ensuring information security. Above and beyond, Xia et al. proposed a Greedy Depth-first Search tree-based accessible encryption plot EDMRS, which accomplished more effectiveness than

early works, however the expense of search stays high and the time intricacy of making trapdoor is high O(n2).

3. METHODOLOGIES

Key Authority (KA)

It is a semi-trusted entity in cloud system. Namely, KA is honest-but-curious, which can honestly perform

The Assigned takes and return correct results. However it will collect as many sensitive contents as possible. In cloud system, the entity is responsible for the users' enrollment. Meanwhile, it not only generates most part of system parameter, but also creates most part of secret key for each user.

Cloud Service Provider (CSP)

It is the manager of cloud servers and also a semi-trusted entity which provides many services such as data storage, computation and transmission. To solve the key escrow problem, it generates both parts of system parameter and secret key for each user. The cloud server is an intermediate entity which stores the encrypted documents and corresponding indexes received from the data owner, and then provides data access and search services to authenticated search users.

Data Owners (DO)

They are owners of files to be stored in cloud system. They are in charge of defining access structure and executing data encryption operation. The data owner outsources her data to the cloud server for convenient and reliable data access to the corresponding search users. To protect the data privacy, the data owner encrypts the original data using symmetric encryption algorithm. To improve the search efficiency the data owner generates some keywords for each outsourced document.

Users

They want to access cipher text stored in cloud system. They download the cipher text and execute the corresponding decryption operation.

SYSTEM ARCHITECTURE

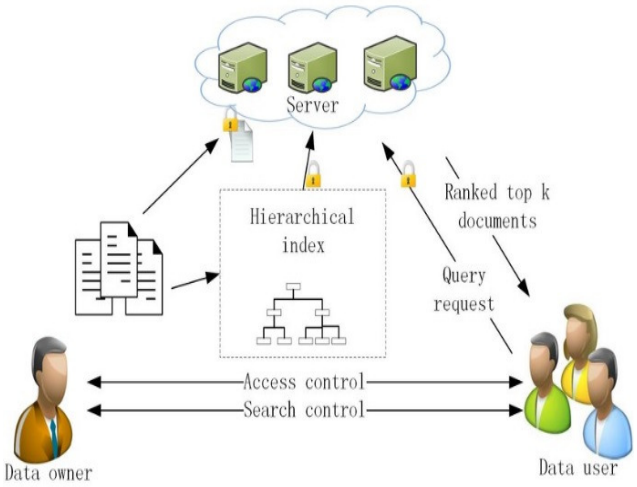

Figure 1 System Architecture

SYSTEM IMPLEMENTATION

A trump card based strategy, for the development of fluffy watchword sets. This system disposes of the requirement for specifying all the fluffy catchphrases and the came about size of the fluffy watchword sets is altogether decreased. In view of the developed fluffy watchword sets, we propose an effective fluffy catchphrase search plot. Through thorough security investigation, we show that the proposed arrangement is secure and protection safeguarding, while accurately understanding the objective of fluffy watchword search. This module is utilized to assist the client with getting the precise outcome dependent on the different catchphrase ideas. The clients can enter the numerous words question, the server is going to part that inquiry into a solitary word after hunt that word document in our database. At long last, show the coordinated word list from the database and the client gets the record from that rundown. The client will choose the necessary record and register the client subtleties and get initiation code in mail from the "customerservice404" email before enter the actuation code. After client can download the Zip record and concentrate that document. The administrator can change the secret word after the login and view the client downloading subtleties and the tallying of record demand subtleties on flowchart. The administrator can transfer the document after the transformation of the Zip record group.

4. CONCLUSION

The issue of multi-keyword ranked search over encrypted cloud data, and start an assortment of security prerequisites. Among various multi-catchphrase semantics, we pick the proficient guideline of "arrange coordinating", whatever number matches as could be expected under the circumstances, to viably catch likeness between question watchwords and redistributed reports, and use "internal item closeness" to quantitatively formalize such a standard for comparability estimation.

For meeting the trial of supporting multi-watchword semantic without assurance bursts, the propose a major MRSE plot using secure inner thing computation, and basically improve it to achieve insurance essentials in two degrees of hazard models.

Cautious assessment inquiring about security and viability affirmations of proposed plans is given, and examinations on this present reality dataset show our proposed plans present low overhead on both estimation and correspondence.

5 .FUTURE WORKS

With a structure like Nehalem close by, there are an assortment of open research issues, which it intend to address for future work. Specifically, it are keen on improving Nephele's capacity to adjust to asset over-burden or underutilization during the activity execution consequently. Our present profiling approach manufactures a significant premise right now the framework despite everything requires a sensible measure of client explanations

6 REFERENCES

[1] 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.

[2] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in *Security and Privacy, 2000. SP 2000. Proceedings. 2000 IEEE Symposium on*, 2000, pp. 44–55.

[3] D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in *Theory of cryptography*. Springer, 2007, pp. 535–554.

[4] E. Shen, E. Shi, and B. Waters, "Predicate privacy in encryption systems," in *Theory of Cryptography*. Springer, 2009, pp. 457–473.

[5] E.-J. Goh *et al.*, "Secure indexes." *IACR Cryptology ePrint Archive*, vol. 2003, p. 216, 2003.

[6] J. Tang, Y. Cui, Q. Li, K. Ren, J. Liu, and R. Buyya, "Ensuring security and privacy preservation for cloud data services," *ACM Computing Surveys*, 2016.

[7] L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword searches over encrypted data," in *Information and Communications Security*. Springer, 2005, pp. 414–426.

[8] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.

[9] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 1, pp. 222–233, 2014.

[10] P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in *Applied Cryptography and Network Security*. Springer, 2004, pp. 31–45.

[11] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: Improved definitions and efficient constructions," in *Proceedings of the 13th ACM Conference on Computer and Communications Security*. ACM, 2006, pp. 79–88.

[12] 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*, ser. ASIA CCS '13. ACM, 2013, pp. 71–82.

[13] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in *Applied Cryptography and Network Security*. Springer, 2005, pp. 442–455.

[14] Y. H. Hwang and P. J. Lee, "Public key encryption with conjunctive keyword search and its extension to a multi-user system," in *Pairing-Based Cryptography–Pairing*. Springer, 2007, pp. 2–22.

[15] Z. Ying, H. Li, J. Ma, J. Zhang, and J. Cui, "Adaptively secure ciphertext-policy attribute-based encryption with dynamic policy updating," *Sci China Inf Sci*, vol. 59, no. 4, pp. 042 701:1–16, 2016.

