Phrase Search Scheme Based on Bloom Filter for Cloud Storage

Dr. R. Usha Rani[1], S. GnaamTayaru[2]
[1]*Associate Professor, Department of CSE, CVR College of Engineering, Hyderabad, India*
[2]*M.Tech Student, Department of CSE, CVR College of Engineering, Hyderabad, India*

Abstract- In the present generation, the data owners stored their data in the form of encryption. The encrypted document's keys are stored in the cloud instead of data owner. So, this scenario not secures for the data owners as well data users. Hence, we need to improve the security of the file or data in the cloud storage. In existing we used different searching schemes to get secure searched results such as conjunctive keyword search, fuzzy searching and multi keyword searching schemes. But, these previous schemes not providing time efficient and secure results to the users and to avoid the searching delay as well as security we required an efficient scheme for cloud storage. In this paper we proposed phrase searching scheme which is provide the fast and secure searched results to the data users.To improve the searched results response time, we are also implementing fuzzy search scheme. By using fuzzy searching, user satisfaction may improve along with time efficient results. In this proposed system we are using bloom filters to provide the secure searching to the users.

Keywords- Cloud Storage, Keyword Searching, Bloom filters

I. INTRODUCTION

Information retrieval, as experienced with the aid of maximum Web users, is an iterative and interactive manner which includes filing a query, seeing the ranked record summaries back in response to the query (which may additionally probable lead to download the related complete documents), and submitting a new question, until the sought data were determined or the quest has been deserted. Unfortunately, the unmanageably huge reaction sets of web search engines like Google coupled with their low precision and ranked listing presentation may additionally make precise perusal tough, time eating and expensive for the user. Research in data retrieval is therefore an increasing number of specializing in the lack of effectiveness of modern-day retrieval engine's interfaces.

With the speedy development of cloud computing, there's a developing fashion of cloud storage utilization. Since the data is encrypted, the looking of documents which incorporate specific keywords will become rather difficult. We may also give you a solution that the person downloads all the encrypted records and decrypts them along with his/her secrete key, after which he/she makes use of normal search techniques to look documents containing precise key phrases. This method is glaringly inefficient and calls for customers to have sturdy storage capability. Organization can buy simplest wished quantity of storage from cloud carrier provider to satisfy their data storage need rather than maintaining their own datastorage. The data owner is relieved from purchasing hardware and software program to manipulate data themselves. Instead of those top notch blessings, the privateness and protection concerns are stopping corporations to utilize these benefits. The records proprietor and the cloud server are not in identical relied on area. Security of remotely stored data is a huge concern because user lost his physical control on data. The control is in the service provider's hand. The data isn't secured as well as there are many attacks with the aid of inner outside attackers. The valuable data which includes social safety range, e-mail, and personal health data and group's financial data should be saved securely. Solution is encryption of data at customer aspect earlier than outsourcing. But if you encrypt data the searching over chipper text is difficult. The current search techniques are only applied on undeniable text information. The trivial solution of downloading all of the records and decrypting regionally is really impractical due huge amount of bandwidth value in cloud scale device. Searchable encryption allows storing data in encrypted layout and you may apply key-word search over chipper text records.

Cloud storage is an online storage model in which humans add their documents and their data might be stored on multiple digital servers, which might be generally hosted via 0.33 parties, in preference to on dedicated servers. Only legal customers, inclusive of the data owners, can access the stored data. To in addition shield their information, human beings commonly encrypt no longer only their report content, however additionally their record names, before uploading them to the cloud which makes it tough for the cloud storage company to search through the information. In current years, searchable encryption strategies have been developed to resolve those problems. However, those strategies are too gradual to be used on a large dataset, i.e.they're not scalable. Furthermore, customers frequently have spelling errors or use

morphological variations of a phrase. Hence, cloud storage seek service should help fuzzy searching.

Cloud computing has generated a whole lot hobby inside the research community in recent years to look over encrypted files saved on cloud many schemes has been proposed but much less attention have been cited on greater seek techniques. To overcome the storage and access of confidential files saved in cloud. We proposed a phrase search using bloom filters which is faster than present system. Our strategies make use of a conjunctive keyword that is downside of present device to get the stored record quicker and get entry to secure.

II. RELATED WORK

Utkarsh Joshi, NeerajVishwakarma and A.Murugan proposed model of using AES Encryption for security of the facts is surprisingly efficient. The search over the encrypted base-64 file has additionally been effectively carried out. And the hunt optimization using Wildcard primarily based method in conjunction with Jaccard calculation for acquiring the similarity between the generated n-grams and previously encrypted n-grams to obtain the desired files. And the n-grams are broken into small pieces of 3 letters every. Rather than storing all the n-grams in single table we're storing the n-grams in exclusive desk as a result developing Indexes.

Min-Shiang Hwang, Shih-Ting Hsu and Cheng-Chi Lee supplied an efficient SCF-PECKS scheme that may stand in opposition to the off-line keyword guessing attack. Their scheme is constructed in bilinear pairing based totally on ElGamal machine and the security is below decisional Diffie-Hellman assumption without random oracle. Their scheme is extra efficient than other conjunctive keyword searchable schemes and is extra suitable for the susceptible gadgets. In addition, this scheme can be extended into the multi-person conjunctive keyword search scheme in the future.

P. Golle, J. Staddon and B. Waters offered two protocols for conjunctive search for which it's far provably hard for the server to differentiate between the encrypted keywords of documents of its personal selecting. Their protocols permit relaxed conjunctive search with small skills. Their work most effective in part solves the trouble of relaxed Boolean search on encrypted information. In particular, a whole solution calls for the capability to do disjunctive keyword seek securely, each throughout and within keyword fields.

D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persianoproposed an encrypted keyword search scheme based totally on public key encryption changed into most of the maximum noted in the area. The author considered a situation where a consumer wishes to have an e mail server verify messages related to certain keywords without revealing the content material of the emails. As sample software, the scheme would allow an urgent encrypted e-mail to be flagged to the attention of a consumer even as others dispatched to suitable folders. The proposed answer makes use of identity based encryption and a version the use of bilinear mapping.

III. FRAMEWORK

A. Proposed System Overview

In this paper, we are presenting a phrase search scheme along with fuzzy search scheme which achieves a much quicker response time than current searched answers. The scheme is also scalable, in which documents can without difficulty be eliminated and delivered to the corpus. We also described changes to the scheme to lower garage price at a small value in reaction time and to shield in opposition to cloud providers with statistical knowledge on saved statistics. Although word searches are processed independently using our technique, they're commonly a specialized function in a keyword search scheme, wherein the number one feature is to provide conjunctive keyword searches.

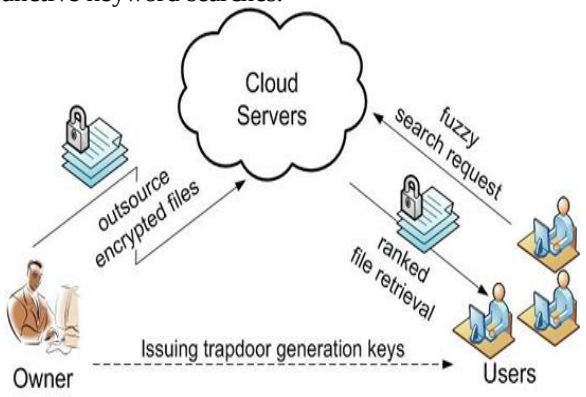

Fig.1: Proposed Framework

Our algorithms can effortlessly be tailored to the state of affairs of an enterprise wishing to setup a cloud server for its employees through enforcing a proxy server in vicinity of the information owner and having the personnel/users authenticate to the proxy server. In the framework, all through setup, the statistics proprietor generates the specified encryption keys for hashing and encryption operations. Then, all files inside the database are parsed for keywords. Bloom filters tied to hashed keywords and n-grams are attached. The files are then symmetrically encrypted and uploaded to the cloud server. To upload documents to the database, the records owner parses the files as in setup and uploads them with Bloom filters connected to the cloud server. To dispose of a report from the information, the records owner absolutely sends the request to the cloud server, who eliminates the document together with the attached Bloom filters. To carry out a search, the data owner computes and sends a trapdoor encryption of the queried key phrases to the cloud to provoke a protocol to search for the asked key phrases inside the corpus. Finally, the cloud responds to the records proprietor with the identifiers to the requested documents.

B. Bloom filters

Bloom filter out is a data structure that helps set membership queries. The fundamental operations involve including factors to the set and querying for element club inside the probabilistic set representation. The primary bloom filter does now not aid the removal of elements. However, these days the improved bloom filter helps deletion operation. The accuracy of a bloom filter relies upon on the dimensions of the clear out, the quantity of hash functions used in the filter out, and the wide variety of factors brought to the clear out. The more the factors added to a bloom filter, the higher the possibility that the query operation reviews false positives.

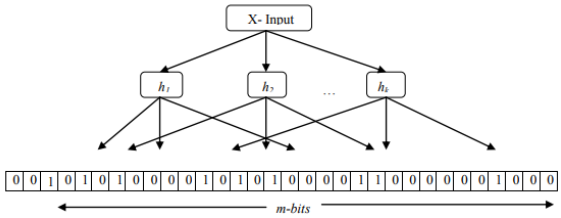

Fig.2: Bloom filter Structure

In the fig2, we passed the input X and that input processed into the bloom filter as hash functions. The bloom filter will check the input through the bits.

C. Phrase Search Scheme Using Bloom Filter

The proposed phrase searching scheme worked by using bloom filters to improve the time efficiency of the searched results in the cloud storage; in a keyword search scheme, Bloom filters can be used to test whether a keyword is associated with a document.

Advantages of Proposed Phrase Search Scheme

1. The proposed system achieves a much faster response time than existing solutions.
2. The proposed algorithms can easily be adapted to the scenario of an organization wishing to setup a cloud server for its employees by implementing a proxy server in place of the data owner and having the employees/users authenticate to the proxy server.

IV. EXPERIMENTAL RESULTS

In this experiment, we implemented three modules such as data owner, data user and cloud. Here, the data owner and data user can register and login into the cloud. The cloud will authenticate the users. Who are authorized users, they can login into the cloud.

Data owner can upload the files and while uploading the file, data owner must enter the file name along with keys. When file will be uploaded then the index will be generated to the uploaded file. The cloud can login and can view the authenticated users list and their uploaded files with content. But, here the content cannot display to the cloud because of encryption.

Next, data user can login into the application and he can search the files using one secrete key with file name. The bloom filter works here to search. The bloom filter gives the time efficient and secure searched results to the data user.

V. CONCLUSION

The conclusion of this paper is that we proposed a novel and time efficient phrase searching scheme to search the keywords from cloud with secure. And we proposed another searching scheme named as fuzzy searching scheme which included in the phrase searching to get time efficient results.By using fuzzy searching, even though we entered keyword query with wrong spelling, we can get the exact searched results. In the proposed scheme we implemented bloom filters to search the phrases from the cloud documents. Mainly we protected here our files from the cloud by storing only encrypted file in the cloud.

VI. REFERENCES

[1]. Hoi Ting Poon and Ali Miri,"Fast Phrase Search for Encrypted Cloud Storage", DOI 10.1109/TCC.2017.2709316, IEEE Transactions on Cloud Computing.

[2]. Y. Tang, D. Gu, N. Ding, and H. Lu, "Phrase search over encrypted data with symmetric encryption scheme," in International Conference on Distributed Computing Systems Workshops, 2012, pp. 471–480.

[3]. Utkarsh Joshi, NeerajVishwakarma and A.Murugan, "Fuzzy Keyword Search over Encrypted Data", Vol. 6, Issue 4, April 2017

[4]. Min-Shiang Hwang, "A New Public Key Encryption with Conjunctive Field Keyword Search Scheme", ISSN 1392–124X (print), ISSN 2335–884X (online) INFORMATION TECHNOLOGY AND CONTROL, 2014, T. 43, Nr. 3

[5]. P. Golle, J. Staddon and B. Waters, "Secure Conjunctive Keyword Search over Encrypted Data," Lecture Notes in Computer Science, Applied Cryptography and Network Security, Vol. 3089, 2004, pp. 31-45.

[6]. D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in In proceedings of Eurocrypt, 2004, pp. 506–522.

[7]. H. Poon and A. Miri, "An efficient conjunctive keyword and phrase search scheme for encrypted cloud storage systems," in IEEE International Conference on Cloud Computing, 2015.
[8]. "A low storage phrase search scheme based on bloom filters for encrypted cloud services," to appear in IEEE International Conference on Cyber Security and Cloud Computing, 2015.
[9]. H. S. Rhee, I. R. Jeong, J. W. Byun, and D. H. Lee, "Difference set attacks on conjunctive keyword search schemes," in Proceedings of the Third VLDB International Conference on Secure Data Management, 2006, pp. 64–74.
[10]. K. Cai, C. Hong, M. Zhang, D. Feng, and Z. Lv, "A secure conjunctive keywords search over encrypted cloud data against inclusion-relation attack," in IEEE International Conference on Cloud Computing Technology and Science, 2013, pp. 339–346.