AN EFFICIENT RANKED MULTI-KEYWORD SEARCH, FUZZY KEYWORD SEARCH FOR MULTIPLE DATA OWNERS OVER ENCRYPTED CLOUD DATA

Ms. Shireen I. Kudle

Department of Computer Engineering

Zeal College of Engineering and Research, Pune

Dr. Swapnaja A. Ubale

Department of Information Technology,

Zeal College of Engineering and Research, Pune.

ABSTRACT: - With the coming of cloud computing, it has turned out to be providing security for information. In existing system, data user can access the without authentication Cipher text-Policy Attribute-based Encryption (CP-ABE) can be utilized to conduct fine-grained and owner-centric access control. In a cloud computing system we are developed the system providing security for information. In this system, data owner can upload different file using AES algorithm in the encrypted format for maintaining the security. For insurance concerns, secure endeavours over scrambled cloud data have propelled a couple of research works under the single proprietor model. In our system we developed this system for multiple owners' model with different functionality. In this system, implemented plans to tree based ranked multi-keyword search scheme for multiple data owners (TBMSM), o proficiently create novel inquiry convention dependent on bilinear blending, which empowers various information proprietors to utilize various keys to scramble their keywords and trapdoors.. In cloud server module, view all users, data owners and all encrypted file also. User also view attacker of the system. Data user can search over encrypted data using hash value md5 algorithm. Data Users can also fuzzy keyword algorithm search technique also used moreover; User can download file at particular place only as well as at particular times only. Also find out attacker of system if any user enters 3 time wrong key.

Keywords- Attacker, Cloud Server, Fuzzy Keyword Search, Hash Value, Ranked keyword Search.

I. INTRODUCTION

Encryption on touchy information before re-appropriating can save information protection. Be that as it may, information encryption makes the customary information usage administration dependent on plaintext keyword search a difficult issue. The category of search function, including secure ranked multi-keyword search, and similarity search. A different data owner can upload this any file in a encrypted format then encrypted index is generated. This encrypted index goes to administrator system. Different data owners can upload files on a cloud so for every file is stored in encrypted format. In this system, user can search that file with different searching techniques like fuzzy keyword search, Hash value search and multikeyword search. Data owners uploaded file store on a cloud server an answer for this issue is to download all the hidden information and make the first information utilizing the hidden key, yet this is not practical because it make additional overhead. In this system, Data owner can file upload in different file in encrypted format using AES 128 bit. When user can search any file then after checking authentication user get file. If user wants to download that file then data user request to data owner. After getting the request user can send the key for download the file. hence , propose when user search keywords that time give the security and demonstrate the bring about positioning structure to make simple cloud servers to perform safe excluding knowing the real value of both keywords and trapdoors, We proposed fuzzy keyword search, using this we can easily search the information. We also introduced any file can download from particular location only. Also find out attacker of system if any user enters 3 time wrong key.

II. OBJECTIVE AND SCOPE

1. File search using multikeyword search as well as search using hash value over encrypted data.
2. File upload in different format like in encrypted format.
3. User can search the encrypted data using fuzzy keyword.
4. User file download at particular place and particular time so system becomes more secure.
5. Find out attacker of the system.

III. PROBLEM STATEMENT

Encryption on sensitive data before outsourcing can preserve data privacy. However, data encryption makes the traditional data utilization service based on plaintext keyword search a very challenging problem. The category of search function contains secure ranked multi-keyword search and similarity search. However, all these schemes are limited to the single-owner model. Search over encrypted data using hash value and attacker problem occurred in existing system.

IV. REVIEW OF LITERATURE

Ravindra R. et.al [1] state that as cloud computing is exceptionally commanding innovation lately, whole delicate data is being put away onto the cloud. For keeping up information secrecy, touchy information are by and large scrambled, which makes viable information usage an extremely mind boggling task. The Existing accessible encryption plans gives a clear way to deal with secure hunt over encoded information utilizing catchphrases and recovering the vital documents of intrigue. While these methods bolster just careful fuzzy keyword search. That is, there is no acknowledgment of slight mistakes and organization irregularities which are run of the mill client looking through conduct. On account of this disadvantage, the current procedures gets inconsistent in distributed computing, influencing the framework convenience. This makes the client looking through encounters extremely baffling and brings about low framework effectiveness. This system incorporates the formalization and arrangement of the issue of viable fuzzy keyword search over encoded cloud information just as saving catchphrase protection. Conquering the downsides of customary inquiry strategies, the fuzzy keyword search supports the framework convenience by creating the coordinating and applicable records when clients' looking through sources of info precisely coordinate the predefined catchphrases or the nearest conceivable coordinating or important documents dependent on catchphrase closeness semantics, when definite match falls flat.

Sofiane Mounine Hemam et.al [2] proposed that in this system, research the heap adjusting between hubs in the volunteer distributed computing. We propose another methodology which depends on cloning a cloud administration on at least one hubs when the quantity of the client solicitations will be significant at a given time. Our answer permits a superior framework unwavering quality and lessens the reaction time of the clients by appropriating their solicitations between the volunteer hubs. Shockingly, the replication of cloud administrations limits the extra room limit. Therefore, we propose a second calculation that chooses and erases the imitations of a cloud administration without corruption of the heap adjusting, utilizing for this the Markov Chain Models. The trial results, in view of PeerSim test system, show that the proposed calculations can viably accomplish great execution (load adjusting) and improve the reaction time.

Hongwei Li et.al [3] introducing utilizing cloud computing, individuals can store their information on remote servers and permit information access to open clients through the clouds servers. As the outsourced data are likely to contain sensitive privacy information, they are commonly encoded before transferred to the cloud. This, in any case, significantly restrains the ease of use of redistributed information due to the difficulty of looking over the encoded information. In this paper, we address this issue by building up the fine-grained multi-keyword search plots over encoded cloud information. Our unique commitments are three-overlay. To begin with, we present the importance scores and inclination factors upon keywords which empower the exact catchphrase search and customized client experience. Second, we build up a useful and very efficient multi-catchphrase search conspire. The proposed plan can bolster convoluted rationale search the mixed"AND","OR"and"NO"operations of keywords. Third, we further utilize the classified sub-word references system to accomplish better efficiency on file building, trapdoor producing and question. In conclusion, we examine the security of the proposed plans as far as confidentiality of records, protection insurance of file and trapdoor, and unlink capacity of trapdoor. Through broad investigations utilizing this present reality dataset, we approve the exhibition of the proposed plans. Both the security examination and trial results exhibit that the proposed plans can accomplish a similar security level contrasting with the current ones and better execution as far as usefulness, inquiry unpredictability and efficiency.

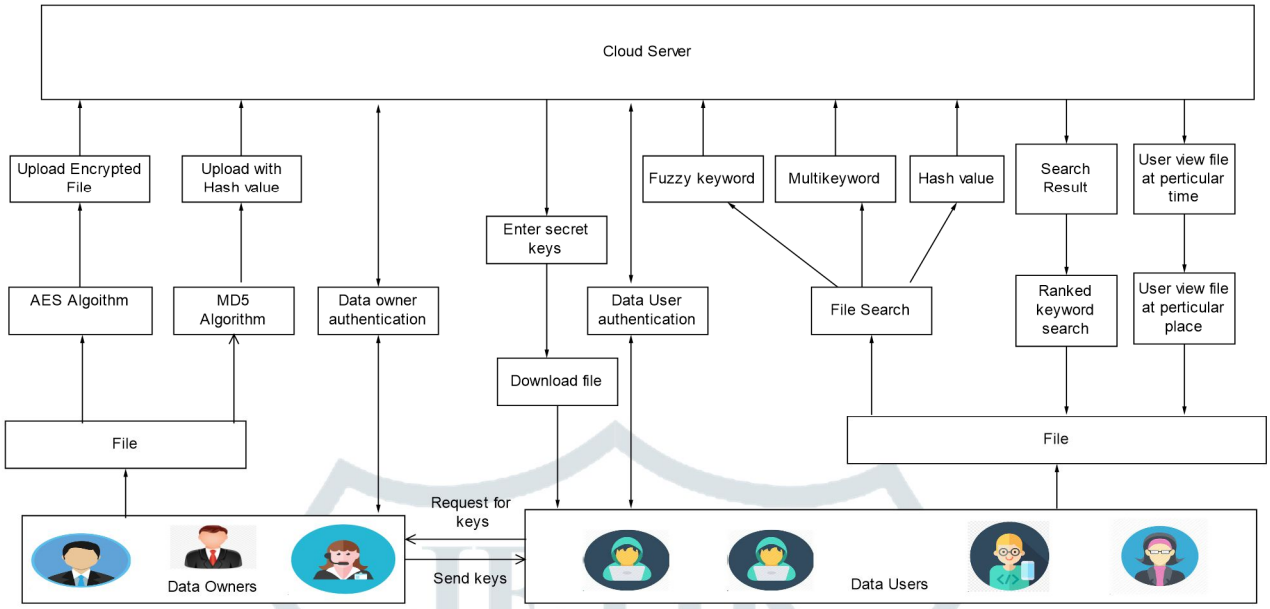
Wei Zhang et.al [4] state that cloud computing provides abundant benefits including simple access, diminished expenses and flexible asset the executives. For security concerns, touchy information must be scrambled before re-appropriating, which obsoletes conventional information usage dependent on plaintext catchphrase search. Hence, building up a protected pursuit administration over encoded cloud information is of principal significance. There are a few explores worried about this issue. Be that as it may, every one of these plans depend on a solitary cloud model which has the danger of single purpose of disappointment, misfortune and defilement of information, loss of accessibility and loss of security. In this system, we investigate the issue of secure appropriated keyword search in a multi-cloud worldview. We first define a conveyed search model. In light of this model, we propose two plans. In plot I, we propose to cross-store all encoded file cuts, keywords and keys. In plot II, we deliberately develop a catchphrase circulating technique and a file conveying methodology. Further, we expand the two plans with Shamir's mystery plans to accomplish better accessibility and heartiness. Broad investigations on genuine world datasets confirm the efficacy and efficiency of our plan

Zhiyong Xu et.al [5] proposed that cloud processing is getting progressively pervasive as of late. It acquaints an efficient route with accomplish the board flexibility and monetary investment funds for distributed applications. To exploit registering and capacity assets offered by cloud specialist organizations, information proprietors must re-appropriate their information onto open cloud servers which are not inside their confided in spaces. In this manner, the information security and protection become a major concern. To forestall data revelation, touchy information must be scrambled before transferring onto the cloud servers. This makes plain content keyword inquiries inconceivable. As the aggregate sum of information put away in broad daylight mists collects exponentially, it is trying to help efficient catchphrase based inquiries and rank the coordinating outcomes on encoded information. Most current works just consider single catchphrase inquiries without proper positioning plans. The multi-keyword inquiry issue was being viewed as of late. MRSE [1] is one of the first inquire about attempts to define and address the issue of compelling yet secure positioned multi-keyword search over encoded cloud information. Nonetheless, the catchphrase word reference utilized in MRSE is static and must be remade when the quantity of keywords in the lexicon increments. It additionally has extreme out-of-request issues in the coordinating outcomes and doesn't consider the catchphrase get to frequencies, which extraordinarily influences its convenience. In this paper, we propose a novel methodology, called MKQE, to address these issues. Just minor changes in the lexicon structure must be done when additional keywords are presented.

Qin Liu et.al[6] proposed that distributed computing as a developing innovation pattern is required to reshape the advances in data innovation. In this paper, we address two central issues in a cloud situation: protection and efficiency. We first survey a private catchphrase based file recovery conspire proposed by Ostrovsky et. al. At that point, in light of an aggregation and distribution layer (ADL), we present a plan, named efficient information retrieval for ranked query (EIRQ), to additionally decrease questioning expenses brought about in the cloud. Inquiries are classified into numerous positions, where a higher positioned inquiry can recover a higher level of coordinated files. Broad assessments have been led on an explanatory model to inspect the viability of our plan.

V.PROPOSED SYSTEM APPROACH

Fig.1 Block Diagram of Proposed System



In this proposed system consist of mainly 3 modules data owners, data users and cloud server.In our proposed system first data owner registration with login with proper authentication. Data owner upload files using AES algorithm in encrypted format, this file is store on the cloud and also upload file with hash value using MD5 algorithm .Data User registration and login with proper authentication, After login user search different file with multi-keyword search, Fuzzy keyword search and Search using hash value also. After Searching user view the file and send request to particular data owner. Data owner accept request and send secret keys to user. Data user enters secret keys and download file at particular time and particular place. If user entered 3 times wrong key user become attacker also cloud server view the attackers. User can view ranked multi-keyword search also.

In proposed system consist following functions like

User Interface for user Registration, Login,

Home Page: -User, Data Owner and Cloud Server

User: - User login with proper authentication, view file, file search using multi-keyword search, fuzzy keyword search, send request, display messages And for download any file from particular place and particular time only.

Data Owner: - Data owner upload file in encrypted format Send secret keys and token to authenticate users only.

Cloud server: - Cloud view info of user and data owner info. Also view file in encrypted format.

VI.MATHAMATICAL MODEL

Mathematical Model in Equation format

Notation

TFU=Total number file upload

FU1=Number of file upload 1

FU2=Number of file upload 2

FU3=Number of file upload 3

TDF=Total number file download

DF1=Number of file download 1

DF2=Number of file download 2

DF3=Number of file download 3

FSK=Total file search by keyword

FSK1=Total file search by keyword 1

FSK2=Total file search by keyword 2

FSK3=Total file search by keyword 3

For calculate total number of file upload by following equation 1

Total number of file upload= Number of file upload 1+ Number of file upload 2+…..+ Number of file upload N

∑ TFU = ∑ FU1+ ∑ FU2+…+∑ FUN
……equation 1

For calculate total number of file download by following equation 2

Total number of file download= Number of file download 1+ Number of file download 2+…..+Number of file download N

∑ TDF= ∑ DF1+ ∑ DF2+…+∑ DFN
……equation 2

For calculate total number of file search by keyword following equation 3

Total number of file search by keyword= Number of file search by keyword 1+ Number of file search by keyword 2+…..+Number of file search by keyword N

∑ FSK= ∑ FSK1+ ∑ FSK2+…+∑ FSKN
……equation 3

VII.ALGORITHMS IN PSEUDO CODE

1. AES Algorithm for Encryption.

AES is an iterative instead of Feistel cipher. It is based on two common techniques to encrypt and decrypt data knows as substitution and permutation network (SPN). SPN is a number of mathematical operations that are carried out in block cipher algorithms. AES has the ability to deal with 128 bits (16 bytes) as a fixed plaintext block size. These 16 bytes are represented in 4x4 matrixes and AES operates on a matrix of bytes. In addition, another crucial feature in AES is number of rounds. The number of rounds is relied on the length of key. There are three different key sizes are used by AES algorithm to encrypt and decrypt data such as (128, 192 or 256 bits). The key sizes decide to the number of rounds such as AES uses 10 rounds for 128-bit keys, 12 rounds for 192-bit keys and 14 rounds for 256-bit keys.

Input:

128_bit /192 bit/256 bit input (0, 1)

Secret key (128_bit) +plain text(128_bit).

Process:

10/12/14-rounds for-128_bit /192 bit/256 bit input

Xor state block (i/p)

Final round:10,12,14

Each round consists: sub byte, shift byte, mix columns, add round key.

Output:

cipher text(128 bit)

2. MD5(Message-Digest Algorithm)

MD5 or ”message digest 5” algorithm was designed by professor Ronald Rivest. Rivest is a professor in MIT who also invented RSA, RC5 and the MD-message digest hashing functions. MD5 is a one way hashing function. So by definition it should fulfill two properties. One, it is one way which means one can create a hash value from a message but cannot recreate the message from the hash value. Two, it should be collision free that is two distinct messages cannot have the same hash value.

Steps 1:A message digest algorithm is a hash function that takes a bit sequence of any length and produces a bit sequence of a fixed small length.

Steps 2: The output of a message digest is considered as a digital signature of the input data.

Steps 3:MD5 is a message digest algorithm producing 128 bits of data.

Steps 4: It uses constants derived to trigonometric Sine function.

Steps 5: It loops through the original message in blocks of 512 bits, with 4 rounds of operations for each block, and 16 operations in each round.

Steps 6: Most modern programming languages provides MD5 algorithm as built-in functions.

3. Fuzzy Keyword Search :-

Fuzzy keyword search greatly enhances system usability by returning the matching files when users searching inputs exactly match the predefined keywords or the closest possible matching files based on keyword similarity semantics, when exact match fails.

Inputs:-

1.$C=(F_1,F_2,\ldots,F_n)$

2.$W=\{W_1,W_2,\ldots,W_n\}$

3.Edit distance *d*

4.A searching input (w, k) (k<=d)

For Normal Search Set Up

$\Pi=(\text{Setup}(1^{\lambda}), \text{Enc}(sk, \cdot), \text{Dec}(sk, \cdot))$

$T_{wi} = f(sk, w_i)$

For Fuzzy Keyword

The wildcard-based fuzzy set of wi with edit distance d is denoted as Swi, d={Swi,0, Swi,1, · · · , Swi,d}.

d=1 (2L+1)*26+1

d=2 $C^1L+1+C^1L*C^1L+2C^2L+2$

For Searching Input:-

$\Pi=(\text{Setup}(1^{\lambda}), \text{Enc}(sk, \cdot), \text{Dec}(sk, \cdot))$

$Twi = f(sk,wi)$ $Tw'i= f(sk,w'i)$ for each $w'i \in Swi,d$

Step 1 FID_{wi} $\text{Enc}(sk,FID_{wi}||w_i)\{(\{T_{w'i}\}w'_{i\in Swi,d}, \text{Enc}(sk,FID_{wi}||w_i))\}_{wi\in W}$

Step 2 $\{Tw'\}w' \in S_{w,k}$

Step 3 $\text{Enc}(sk,FID_{wi}||w_i)$

Output:-

Get Expected result which is search by the user.

VIII.COMPARATIVE RESULTS

In our experimental setup, in table no.8.1, shows number of file upload and file download. In our system 70 total number of files. In that 39 were number file upload and 21 were downloading of files.

Table1: Number of File Upload and download

Sr. No.	Number of File Upload	Number of File Download
1	39	21

In our experimental setup, in table no.8.2, User can search different file with different keywords so get information about how to user can search any files. In that 46 users search by 1st keyword, 31 users search by 2nd keyword and 39 users search by 3rd keyword.

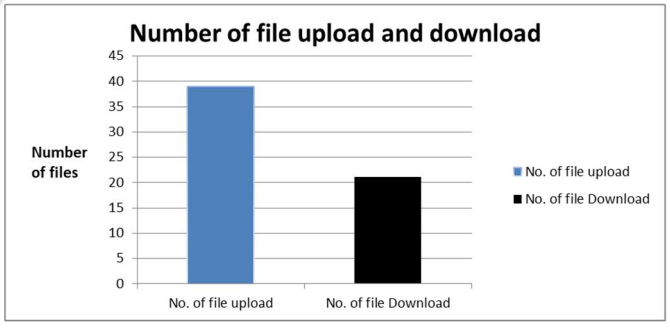
Table2: Number of file search by keyword

Sr. No.	No. of File Search by Keyword1	No. of File Search by keyword2	No. of File Search by keyword3
1	46	31	39

IX.RESULTS

From above data, as shown in graph 9.1, the total numbers of files were 70. The numbers of files found to be uploading were 39 and downloading files were 21.

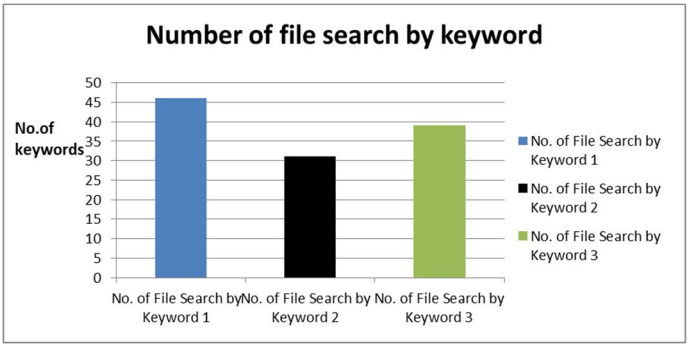
Graph 1: Number of file upload and download



In our experimental setup, as shown in graph 9.2, From above table data, In graph, we can see the no. of file search keyword by keyword 1, no of file keyword 2 and no of file keyword 3 in the graph ; we see 46 files search by keyword 1, 31 files search by keyword 2 and

39 files search by keyword 3 by different users are shown in the graph.

Graph 2: Number of file search by keyword



CONCLUTION

In this system, we implemented consider a multiple data owners model in cloud computing and propose an efficient ranked multikeyword search scheme over encrypted data. In existing system, data user can access the without authentication Cipher text-Policy Attribute-based Encryption (CP-ABE) can be utilized to conduct fine-grained and owner-centric access control. In this system, user can search using different searching techniques like multi-keyword search,Fuzzzy keyword search and Hash Value search. Upload a file in encrypted format for maintain the security. User can download any file in particular place and particular time only.

FUTURE WORK

In future, we can upload data with images and videos also.

ACKNOWLEDGMENT

This work is supported in multi-keyword search, fuzzy keyword search for multiple data owners over encrypted cloud data field in India. Authors are thankful to Faculty of Engineering and Technology (FET), Savitribai Phule Pune University, Pune for providing the facility to carry out the research work.

REFERENCES

[1] Ravindra R. Ghugare,Pranjuli Yavatkar, Nikita Patil, Sneha Kale **"Fuzzy Keyword Search over Encrypted Data in Cloud Computing"** International Journal of Recent Trends in Engineering & Research (IJRTER) Volume 04, Issue 01; January - 2018 [ISSN: 2455-1457]

[2] Sofiane Mounine Hemam, Ouided Hioual, Abbes Laghrour **"Load Balancing Between Nodes in a Volunteer Cloud Computing by Taking Into Consideration the Number of Cloud Services Replicas"** 2017 3rd International Conference of Cloud Computing Technologies and Application(CloudTech)

[3] Hongwei Li, Yi Yang, Tom H. Luan, Xiaohui Liang, Liang Zhou and Xuemin (Sherman) Shen **"Enabling Fine-grained Multi-keyword Search Supporting Classified Sub-dictionaries over Encrypted Cloud Data"** in IEEE Transaction on dependable and secure computing, vol 13, no. 3, May/June 2016.

[4] Wei Zhang Sheng Xiao Yaping Lin,Ting Zhou Siwang Zhou **"Secure Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing"** 2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks.

[5] Zhiyong Xu, Wansheng Kang, Ruixuan Li, KinChoong Yow, and Cheng-Zhong Xu **"Efficient Multi-Keyword Ranked Query on Encrypted Data in the Cloud"** 2016 IEEE 18th International Conference on Parallel and Distributed Systems

[6] Qin Liu, Chiu C. Tan, Jie Wu, and Guojun Wang, Hunan Province, P. R. China, **"Efficient Information Retrieval for Ranked Queries in Cost-Effective Cloud Environments"** 2016 The 31st Annual IEEE International Conference on Computer Communication