International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6, August, 2019

Efficient Multi-Keyword Search Through Ciphertext Data in the Cloud

Shridevi Soma, Vinaya S Kavalgi

Abstract:Cloud computing has become essential for storing sensitive data sets that are centralized in the cloud. The need for privacy and protection of files and documents has increased day by day. Data users typically dump the most powerful information in cloud storage to prevent third parties from accessing data in cloud storage. In legacy systems, end users safely retrieved encrypted information using keyword search. However, in existing systems, we recommend only individual keywords and Boolean keywords, which is not yet sufficient to ensure efficient data usage for a vast number of data users and the number of documents in the cloud repository. This work aims to develop a systematic approach to searching multi keywords in the cloud with ciphertext data. The cloud server carries out risk-free investigations with no clear information about keywords and trap doors. The client or user uses multiple keywords to perform data retrieval. When the client enters a question for many words, the server breaks the question into one word and retrieves the word from the index. In this task, the cipher text policy attribute-based encryption (CPABE) algorithm is used to perform encryption of files and documents. Experimental results show 95% accuracy with a data set size of 1000, for both single and multiple keyword searches. Because previous research were limited to single searches, this new work performs multiple keyword searches with unique security aspects to create a multi-keyword search system rather than the cryptographic data in the cloud.

Index terms-- Cloud storage, MKS (Multi-keyword search), Rank search, Collective data owners, SKS(Single keyword search), Security, Secret key generation.

I. INTRODUCTION

Cloud computing has created tremendous momentum in the IT industry that can be used to understand the kinds of computing, storage, and applications. Several IT companies dump data to cloud storage. Different users can access or send information stored in the cloud, regardless of their location. Research shows that mobile cloud provides a bandwidth and energy efficient encrypted search architecture. It is limited to a single keyword search system to efficiently retrieve encrypted data. They used the mobile cloud architecture to perform encrypted searches to reduce computation time and reduce energy consumption, but were limited to just one keyword search [1]. A special tree-based index structure that uses greedy depth-first search algorithms to introduce search through multiple keyword searches. There is not enough symmetry in the search to handle the problem [2].

Keyword-based searches have been introduced directly into encrypted data outsourced to cloud servers. There are errors related to spell checking to eliminate people who have tried local sensitive hashing techniques [4]. Multiple keyword searches have been calculated to provide an efficient search. When users searching for multiple keywords search for matching data or files on the cloud server, the words are identified and the keywords are organized. When merged together, all indexes are formed. Security k - Internal computation adapted from the nearest neighbours approach [5]. They collect the relevant data values from the search query and choose a similar match as possible so that the match measure [3] can be calculated. A statistical approach to creating a one-to-one retention mapping to protect sensitive sensitive data and develop a secure search index on file retrieval. This solution creates a privacy guarantee as robust as possible compared to previous searchable encryption methods in the range of ranked keyword searches [6]. Honest and uninteresting cloud servers have introduced the VSSE framework to provide provable search beyond file privacy and have been validated through rigorous security analysis [7]. Multidimensional (MD) algorithm adaptation methods are used for linear retrieval to support multi-dimensional searches introduced to solve multi-keyword text search (MTS) problems [13]. For ranked searches, the system improves through checks, returns similar files in order, and uses the OPSE algorithm [12]. In this task, several keywords search for ciphertext files encrypted with different keys for different information owners. Cloud servers can be risk-free without having clear information about keywords and trap doors. The client or user uses multiple keywords to retrieve data. When a client enters a question for many words, the server breaks the question into one word and searches for the word in the index. As a result, many information owners use different keys to encrypt files and keywords. Authorized users can execute questions without the important secret key of many information owners. This work uses the CPABE algorithm to convert plaintext to ciphertext (encrypted format).

II. PROPOSED SYSTEM

With the development of various stages for sharing the data's by means of cloud server, cloudlets and so on, the keywords search is necessary to search the files from the cloud storage.

Revised Manuscript Received on October 30, 2019.
* Correspondence Author
Dr. Shridevi Soma, Computer Science and Engineering, PDACE, Kalaburagi, Karnataka, India. Email: shridevisoma@gmail.com
Vinaya S Kavalgi, Computer Science and Engineering, PDACE, Kalaburagi, Karnataka, India. Email: vinayakavalgi@gmail.com

To maintain the information safety the security level should be enhanced. The proposed system is designed using three modules namely. Data Owners, Data Users, Cloud Servers as illustrated in the figure 1.

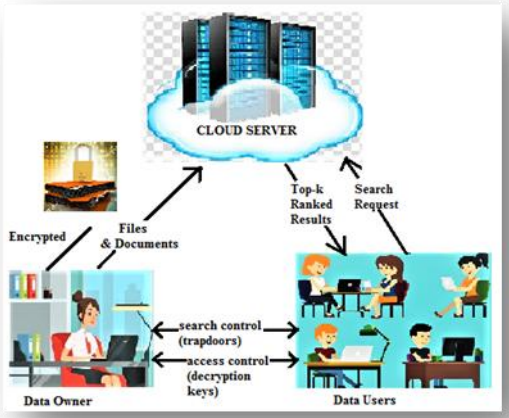

Figure 1 : System architecture of multi-keyword search through ciphertext data

A. *Data owners*

The data owners contains confidential information that attempts to upload a large number of files, documents, and encryption information in encrypted form to cloud storage. There are other things that a data user must provide in reply to a data access request sent by a data user so that the data user can gain access to the files on the cloud server. The data owner must generate a secret key for the user to proceed to the authentication process and the users must be able to download the file.

B. *Data Users*

The authorized data users only access the files of data owners with the keywords. In the data users the important part involves the multi-keywords searches the relatable keywords documents from the cloud servers. Users will find the files or documents using multi-keywords. When the questions for multi-keywords is entered, the cloud server will divide the questions into one word and searches that words from the database. Data users needs transfer the data access request so they can go through the files from the cloud servers. When they gets the secret key from the data owners authentication process involves then the authorized users can easily view and copy the files.

C. *Cloud servers*

The cloud server stores encrypted collections of records. The Cloud Server contains all the information for the data owners, data users, and uploaded file details. Cloud Server returns the highest rank files and document results.It will display top-k files and apply the security to that files then decrypt and download the files and documents into host system.

In this work, which searches for multiple keywords, users can use the keywords searches count on the cloud server to find documents. When a user enters multi keywords, the cloud server splits the word, which is searched from a set of similar or matching files that are in the database or uploaded to the cloud server. In this work, the first data owner and user are logged in with their name and password. The security authentication process is performed as shown in Figure 2. The data owners has uploaded the appropriate file to the cloud. The data access request must be answered by the data owner.

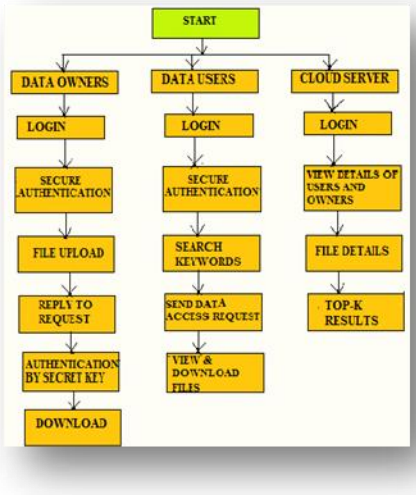

Figure 2: Block diagram of multi-keyword search through ciphertext data.

In data users to find the files from cloud they need to enter multi-keywords to search the similar data or files from the cloud. For authentication secret key is sent to data users for giving permission to access the data. Then data users able to copy the files. In cloud server it will return the top-k results of files.

III. ALGORITHM

The proposed work aims to perform multi-keywords search over cipher text data. An encryption algorithm Cipher text Policy Attribute-Based Encryption (CPABE) used to encrypt files and documents for privacy. It is extensively used in various cyber real systems and the Internet of Things for maintaining data security. CPABE contains four algorithms as shown in the figure 3.

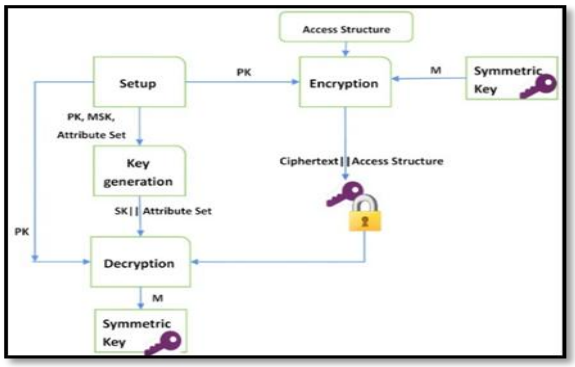

Figure 3: Diagram of Cipher text Policy Attribute-Based Encryption (CPABE)

A. ***Setup:*** This algorithm creates the public key PK and the master key MK.

B. ***Encryption:*** Data users uses the public key PK to execute the data owners to encrypt the plaintext M and output the ciphertext CT.

C. ***Key Generation:*** This algorithm creates the secret key SK according to the set of attributes S provided by data users. The user's attributes describe their private key SK and determine the decryption authority.

D. ***Decryption:*** This algorithm is executed by the data users to decrypt the ciphertext CT with the generated secret key.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The work processes of the system used Eclipse tools. First, run the Tomcat server in Eclipse connected to the SQL database. For cloud storage, DriveHQ is used as a cloud service provider. The data owner manages user information and generates a secret key, the owner uploads the file to the cloud with access control to the file, access control grants access to specific authorized data users, and the private key Through users. Data users must enter multiple keywords to retrieve data files from cloud servers and files, and files must be converted to cipher text. The data user downloads and uploads a valid private key. Encryption is performed using the CP-ABE algorithm. With multiple keyword-based searches, you can use your keywords in the cloud to efficiently locate and download files and documents as a decryption method. The SQL database contains all data items and upload file details by registered data owners and users.

The Table I describes the data collected to analyze the performance parameters where the parameters are collected for making the performance great for multi-keywords search. In existing systems it's limited to single keyword search and lack of problems while in searching the keywords.

Table I: Data of number of files and single/multi-keyword search.

No. of files	Total Keywords	Single keyword search	Multi-keyword search
50	100	25	75
120	170	50	120
250	550	100	450
500	780	150	630
750	900	250	650
1000	1200	400	800

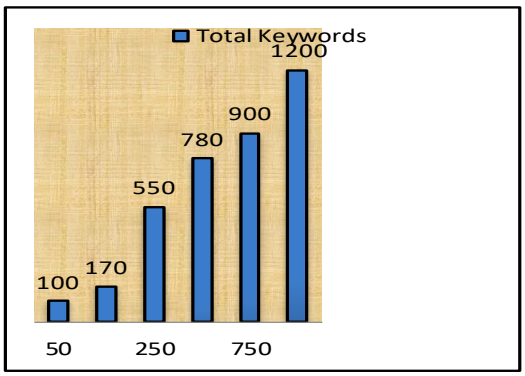

Figure 4: Generation of total keywords of the files

Figure 4 shows the generation of total keywords of the input files. For example:- First we checked 50 files containing 100 keywords, then these keywords were included in the MKS and SKS search process. The process is performed with a total of 1000 files and many related file keywords are used as the data set in the experimental process for MKS and SKS searches.

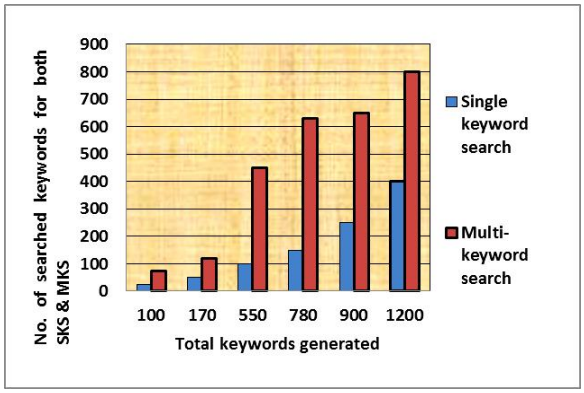

Figure 5: Chart of comparison between single keyword search and multi-keyword search

Figure 5 shows the comparative analysis of single keyword search and multiple keyword search. Searching for individual keywords creates the confusion of searching for similar keywords to find the files, since there are many keywords that are related to each other. This leads to a lack of time calculation, limited security guarantees and a lower speed searching process. Multi-keyword searches improve the fast search process by allowing you to easily search for relevant files on the cloud server without confusion and speeds up the search process. The security level has been increased by the created secret key, so that third parties do not misuse the process. In this work, a total of 1,000 files and many keywords from associated files are used as a set of data for experimentation purposes to search for multi-keywords search. The total accuracy of the system is estimated at 95% depending on the results.

The overall accuracy of the system is calculated using the formula,

$$\text{Accuracy (\%)} = \frac{\text{Number of Single/Multi keywords}}{\text{Total number of files}} * 100$$

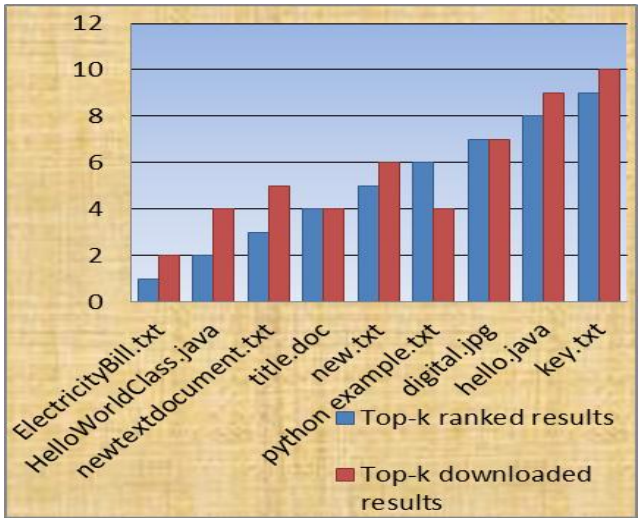

Figure 6: Chart of top-k downloaded results and top-k ranked results

Figure 6 shows the list of files that have been given the highest priority, i.e top-k ranked. If you use the values of the data users, the results with the highest k-rank that these files prefer to the users are displayed. The top-k downloaded results summarize number of times the files were downloaded by users from the cloud server. The download can be repeated or after the time ranking has been entered through the table.

The repeated ones will take the highest rank as users will download these repeated files and display them.

V. CONCLUSION AND FUTURE WORK

In this research work, the efficient multi-keyword search through ciphertext data present in the cloud is implemented using the CPABE algorithm. In this approach, users search for the appropriate matching keywords of documents in the cloud server. Users can easily find the files or documents with multi-keywords. Cloud servers achieved risk-free investigation without significant authentic information on the keywords and trapdoors. It reduces the computation time, increases security and performs a quick search.

This work can be further enhanced by efficiently storing the huge data in the cloud storage. In today's world, it would be a challenging and efficient way to store big data in the cloud. Hence this proposed method is developed to meet the challenges. The future scope of this work can be enhancement in the performance accuracy.

REFERENCES

1. J.Li, R.Ma, and H.Guan, "TEES: An Efficient Search Scheme overEncrypted Data on Mobile Cloud" IEEE Transactions on Cloud Computing, vol.3, pp.2168-7161 2015.
2. Z.Xia, X.Wang, X.Sun, and Q.Wang "A Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data" IEEE Transactions on parallel and distributed systems vol:pp no:99 2015.
3. W. Zhang, Y. Lin, S. Xiao, J. Wu, S. Zhou, "Privacy preserving ranked multi-keyword search for multiple data owners in cloud computing," IEEE Trans Comput., vol. 65, no. 5, pp. 1566 – 1577, 2016.
4. B.Wang, S.Yu, W.Lou, Y.T.Hou, "Privacy-Preserving Multi-Keyword Fuzzy Search over Encrypted Data in the Cloud " IEEE INFOCOM 2014-IEEE Conference on computer communications.
5. N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacy preserving multi-keyword ranked search over encrypted cloud data," Parallel and Distributed Systems, IEEE Transactions on, vol. 25, no. 1, pp. 222-233, 2014.
6. Cong Wang, Ning Cao, Ming Li, Kui Ren, Wenjing Lou "Enabling Secure and Efficient Ranked Keyword Search over Outsourced Cloud Data" IEEE Transactions on parallel and distributed systems, vol.23,August 2012.
7. Qi Chai, Guang Gong "Verifiable Symmetric Searchable Encryption For Semi-honest-but-curiousCloud Servers" IEEE ICC 2012-Communication and information systems security symposium.
8. W. Sun, B. Wang, N. Cao, H. Li, W. Lou, Y. Hou, H. Li, "Privacy preserving multi-keyword text search in the cloud supporting similarity based ranking," IEEE T Parall Distr., vol. 25, no. 11, pp. 3025 – 3035, 2014.
9. R. Li, Z. Xu, W. Kang, K. Yow, C. Xu, "Efficient multi-keyword ranked query over encrypted data in cloud computing," FUTURE GENER COMP SY., vol. 30, pp. 179 – 190, 2014.
10. C. Wang, N. Cao, J. Li, K. Ren, W. Lou, "Secure ranked keyword search over encrypted cloud data," in: ICDCS'10, Genoa, Italy, 2010.
11. N. Cao, C. Wang, M. Li, K. Ren, W. Lou, "Verifiable Privacy-preserving multikeyword text search in the cloud supporting similarity based ranking" IEEE transactions on parallel and distributed systems Vol 25, nov 2014

AUTHORS PROFILE

Dr. Shridevi Soma M.Tech, Ph.D.

Author working presently as Professor in Department of Computer Science and Engineering, Poojya Doddappa Appa College of Engineering Kalaburagi, Karnataka, India. She has 18 years of Teaching and 10 years of Research Experience, and completed her B.E, M.Tech. and Ph.D. in Computer Science and Engineering. Her Research Area includes Digital Image Processing and Pattern Recognition, Cloud Computing, Internet of Things, Big Data Analytics. She published more than 30 Research papers in above mentioned areas, also Guiding Research Students.

Vinaya S Kavalgi is currently pursuing M.Tech. in Computer Science and Engineering from the Department of Computer Science and Engineering at Poojya Doddappa Appa College of Engineering, Kalaburagi, Karnataka, India.