ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

An Efficient Keyword Hierarchical Search for Safeguarding Privacy of Encrypted Cloud

M.D.Surwase[1], A. V. Mophare[2],

Department of Computer Science and Engineering, N B Navale Sinhgad College of Engineering, Kegaon, Solapur, India

Assistant Professor, Department of Computer Science and Engineering, N B Navale Sinhgad College of Engineering, Kegaon, Solapur, India

ABSTRACT: The project solves and defines the difficulty of multi-keyword ranked search over encrypted cloud data (MRSE) while preserving firm system wise privacy in the cloud computing hypothesis. Hence to protect privacy of the data, before privacy data also outsourced to the cloud data that has delicate to be encrypted, which make the valuable data utilization service not a easy task. Even though searchable encryption method allows users to firmly search over encrypted data all the way through the keywords, they carry only search i,e Boolean. They are not yet enough to meet the utilization of the data successfully because there is instinctively demanded by large number of data files and users located in cloud. Hence it is required to allow multiple keywords in the search request and return documents in the order of their significance to the keywords require. Related takes a shot at searchable encryption concentrate on single catchphrase inquiry or Boolean watchword seek, and once in a while separate the indexed lists. In this paper, surprisingly, we characterize and tackle the testing issue of protection saving Multi-keyword Ranked Search over Encrypted Cloud Data [MRSE] and build up an arrangement of strict security necessities for such a safe cloud information use framework to wind up noticeably a reality. Then according to Top K Query scheme the sorted results are created.

KEYWORDS: Encryption Scheme, Secure Data Maintenance, Dynamic Searching Scheme, Cloud Computing.

I. INTRODUCTION

Cloud computing has been considered as a new model of enterprise IT infrastructure, which can organize huge resource of computing, storage and applications, and enable users to enjoy ubiquitous, convenient and on-demand network access to a shared pool of configurable computing resources with great efficiency and minimal economic overhead . Attracted by these appealing features, both individuals and enterprises are motivated to outsource their data to the cloud, instead of purchasing software and hardware to manage the data themselves.

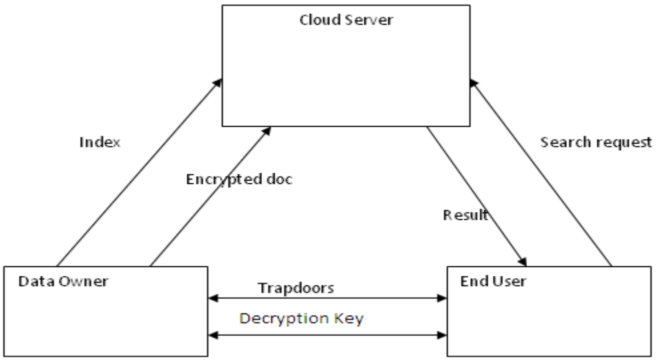


Fig.1 System Architecture Design

DOI: 10.15680/IJIRCCE.2017. 0504379

Despite of the various advantages of cloud services, outsourcing sensitive information (such as e-mails, personal health records, company finance data, government documents, etc.) to remote servers brings privacy concerns. The cloud service providers (CSPs) that keep the data for users may access users' sensitive information without authorization. A general approach to protect the data confidentiality is to encrypt the data before outsourcing. However, this will cause a huge cost in terms of data usability. For example, the existing techniques on keyword-based information retrieval, which are widely used on the plaintext data, cannot be directly applied on the encrypted data. Downloading all the data from the cloud and decrypt locally is obviously impractical.

II. EXISTING SYSTEM

A general way to deal with secure the information privacy is to scramble the information before outsourcing. Searchable encryption plans empower the customer to store the encoded information to the cloud and execute watchword seek over ciphertext space. As such, bottomless works have been proposed under various danger models to accomplish different inquiry usefulness, for example, single catchphrase pursuit, closeness seek, multi-watchword Boolean hunt, positioned look, multi-watchword positioned seek, and so forth. Among them, multi-watchword positioned look accomplishes increasingly consideration for its down to earth pertinence. As of late, some dynamic plans have been proposed to bolster embeddings and erasing operations on archive accumulation. These are critical acts as it is exceptionally conceivable that the information proprietors need to redesign their information on the cloud server.

Disadvantages

- Huge cost in terms of data usability. For example, the existing techniques on keyword-based information retrieval, which are widely used on the plaintext data, cannot be directly applied on the encrypted data. Downloading all the data from the cloud and decrypt locally is obviously impractical.
- Existing System methods not practical due to their high computational overhead for both the cloud sever and user.

III. RELATED STUDY

In the year of 2012, the authors K. Ren, C.Wang, Q.Wang et al. revealed a paper titled "Security challenges for the public cloud" and describe into the paper such as Distributed computing speaks to today's most energizing processing outlook change in data innovation. In any case, security and protection are seen as essential deterrents to its wide reception. Here, the creators plot a few basic security challenges and rouse promote examination of security answers for a dependable open cloud environment.

In the year of 2010, the authors S. Kamara and K. Lauter Cryptographic cloud storage in Financial Cryptography and Data Security" and describe into the paper such as We consider the issue of building a protected distributed storage benefit on top of an open cloud framework where the specialist organization is not totally trusted by the client. We portray, at an abnormal state, a few designs that join later and non-standard cryptographic primitives with a specific end goal to accomplish our objective. We review the advantages such engineering would give to both clients and specialist organizations and give an outline of late advances in cryptography inspired particularly by distributed storage.

In the year of 2009, the author C. Gentry revealed a paper titled "A fully homomorphic encryption scheme" and describe into the paper such as We propose the primary completely homomorphic encryption plot, taking care of an old open issue. Such a plan permits one to register subjective capacities over encoded information without the decoding key—i.e., given encryptions E(m1), ..., E(mt) of m1, ..., m t, one can effectively process a conservative ciphertext that scrambles f(m1, ..., m t) for any proficiently calculable capacity f. Completely homomorphic encryption has various applications.

For instance, it empowers encoded internet searcher questions—i.e., a web crawler can give you a compact scrambled solution for your (boolean) inquiry without recognizing what your inquiry was. It additionally empowers seeking on encoded information; you can store your scrambled information on a remote server, and later have the server recover just records that (when unscrambled) fulfill some boolean requirement, despite the fact that the server can't decode the documents all alone.

All the more comprehensively, it enhances the proficiency of secure multiparty calculation. In our answer, we start by planning a to some degree homomorphic "boostrappable" encryption plot that works when the capacity f is the plan's own unscrambling capacity. We then show how, through recursive self-implanting, bootstrappable encryption gives completely homomorphic encryption.

IV. PROPOSED SYSTEM

In this proposed framework, surprisingly, we characterize and tackle the issue of Multi-keyword Ranked Search over Encrypted Cloud Data [MRSE] while safeguarding strict framework shrewd protection in the cloud computing worldview.We enhance the of ranked search mechanism, including supporting more search semantics, i.e., TF _IDF, and dynamic data operations. Also performs the provision of maintaining the integrity of rank order in search result and the cloud server is untrusted.Beacouse of providing the integrity to rank order the quality of search is enhanced or improved .User save the time to get relevance document to their search query. In order to improve the document retrieval accuracy, the search result should be ranked by the cloud server according to ranking in order to make the data on cloud more secure. To reduce the cost of communication data user can provide N number along with the trapdoor so that cloud server return only top-N document which having are relevance to user query.

The hunt question is additionally depicted as a parallel vector where each piece implies whether relating catchphrase shows up in this pursuit ask for, so the comparability could be precisely measured by the inward result of the inquiry vector with the information vector. Be that as it may, specifically outsourcing the information vector or the inquiry vector will disregard the list protection or the pursuit security. To address the difficulty of supporting such multi-catchphrase semantic without protection ruptures, we propose an essential thought for the MRSE utilizing secure inward item calculation, which is adjusted from a safe k-Nearest Neighbor (kNN) method [23].

System Contributions

(a) For the first occasion when, we investigate the issue of multi-keyword positioned seek over encoded cloud information, and build up an arrangement of strict protection necessities for such a safe cloud information usage framework.

(b) We propose two MRSE plans in view of the comparability measure of 'Organize Coordinating' while at the same time meeting distinctive protection necessities in two diverse risk models.

(c) Thorough examination exploring security and proficiency certifications of the proposed plans is given, and analyses on this present reality dataset additionally demonstrate the proposed plots in fact present low overhead on calculation and correspondence.

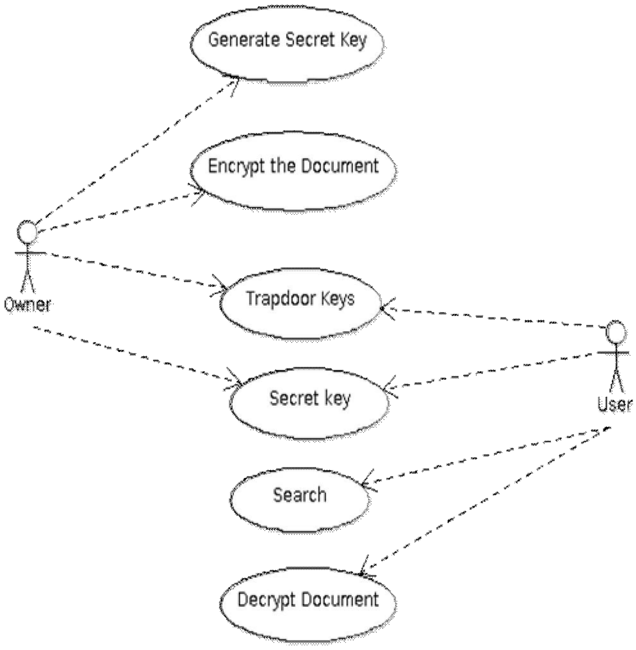


Fig.2. System UseCase Design

Cloud Server

Cloud server stores the encrypted document collection C and the encrypted searchable tree index I for data owner. Upon receiving the trapdoor TD from the data user, the cloud server executes search over the index tree I, and finally returns the corresponding collection of top-k ranked encrypted documents. Besides, upon receiving the update information from the data owner, the server needs to update the index I and document collection C according to the received information.

Data Owner

Data owner has a collection of documents that he wants to outsource to the cloud server in encrypted form while still keeping the capability to search on them for effective utilization. In our scheme, the data owner firstly builds a secure searchable tree index I from document collection F, and then generates an encrypted document collection C for F. Afterwards, the data owner outsources the encrypted collection C and the secure index I to the cloud server, and securely distributes the key information of trapdoor generation (including keyword TF values) and document decryption to the authorized data users.

Data User

Data users are authorized ones to access the documents of data owner. With t query keywords, the authorized user can generate a trapdoor TD according to search control mechanisms to fetch k encrypted documents from cloud server. Then, the data user can decrypt the documents with the shared secret key.

KNN Algorithm

This is the main driver of the code. To do the classification, we are essentially interested in finding the distance between the particular instance we are trying to classify to other instances. We then determine the classification of the instance we want from a “majority vote” of the other k closest instances. Each feature of an instance is a separate class that essentially just stores a continuous or discrete value depending on if you are using regression or not to classify your neighbors. The additional feature classes and file reader are left to the reader as an exercise. Note that it would be fairly easy to weight features using this model depending on if you want to give one feature more clout than another in determining the neighbors.

Multi-keyword Ranked Search over Encrypted Cloud Data

Setup (1ℓ) Taking a security parameter ℓ as input, the data owner outputs a symmetric key as SK. BuildIndex(F, SK):Based on the dataset F, the data owner builds a searchable index I which is encrypted by the symmetric key SK and then outsourced to the cloud server. After the index construction, the document collection can be independently encrypted and outsourced. Trapdoor($\tilde{w}$) : With t keywords of interest in $\tilde{w}$ as input, this algorithm generates a corresponding trapdoor T$\tilde{w}$. Query(T$\tilde{w}$, k, I):When the cloud server receives a query request as (T$\tilde{w}$, k), it performs the ranked search on the index I with the help of trapdoorT$\tilde{w}$, and finally returns F$\tilde{w}$, the ranked id list of top-k documents sorted by their similarity with $\tilde{w}$.

3DES Algorithm

Which consists of three different DES keys K1, K2 and K3. This means that the actual 3TDES key has length 3×56 = 168 bits.

The encryption-decryption process is as follows:

- Encrypt the plaintext blocks using single DES with key K1.
- Now decrypt the output of step 1 using single DES with key K2.
- Finally, encrypt the output of step 2 using single DES with key K3. The output of step 3 is the ciphertext.
- Decryption of a ciphertext is a reverse process. User first decrypt using K3, then encrypt with K2, and finally decrypt with K1. Due to this design of Triple DES as an encrypt–decrypt–encrypt process.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

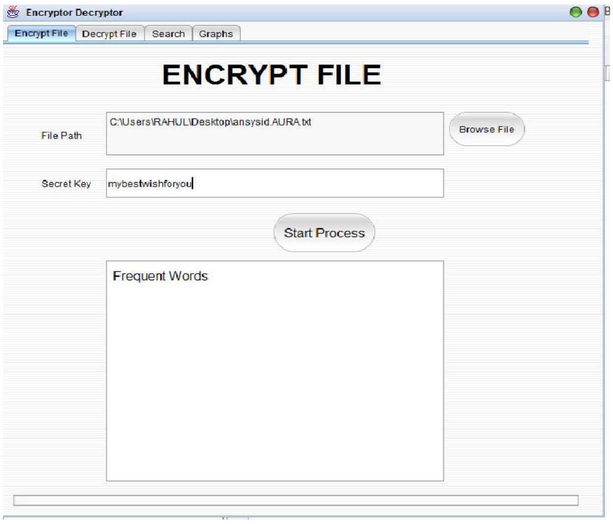
International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 4, April 2017

V. EXPERIMENTAL RESULTS

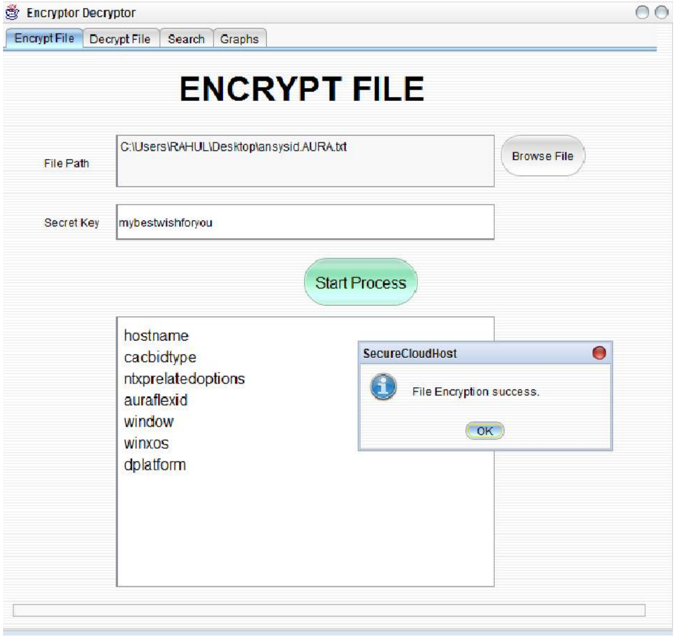


File Path: C:\Users\RAHUL\Desktop\ansysid.AURA.txt

Secret Key: mybestwishforyou

Frequent Words

Fig.3. Encrypt Input File



File Path: C:\Users\RAHUL\Desktop\ansysid.AURA.txt

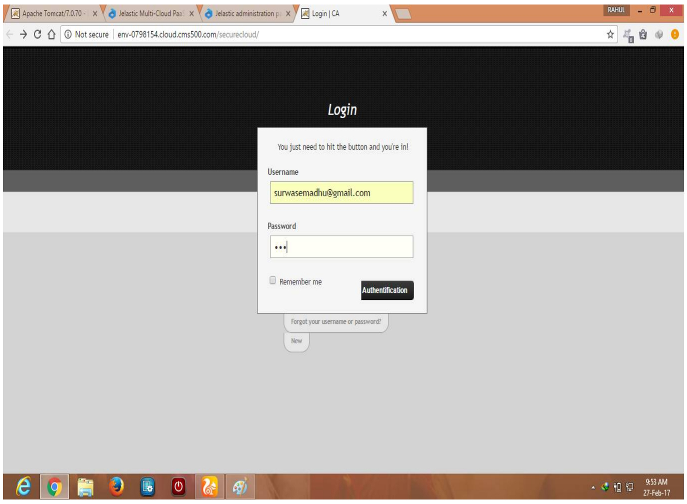
Secret Key: mybestwishforyou

hostname
cacbidtype
ntxprelatedoptions
auraflexid
window
winxos
dplatform

File Encryption success.

Fig.4. Encryption Process Succeeded

DOI: 10.15680/IJIRCCE.2017. 0504379


Apache Tomcat/7.0.70 -
Jelastic Multi-Cloud PaaS
Jelastic administration p
Login | CA
RAHUL
Not secure
env-0798154.cloud.cms500.com/securecloud/
Remember me
Authentification
Forgot your username or password?
New
9:53 AM
27-Feb-17

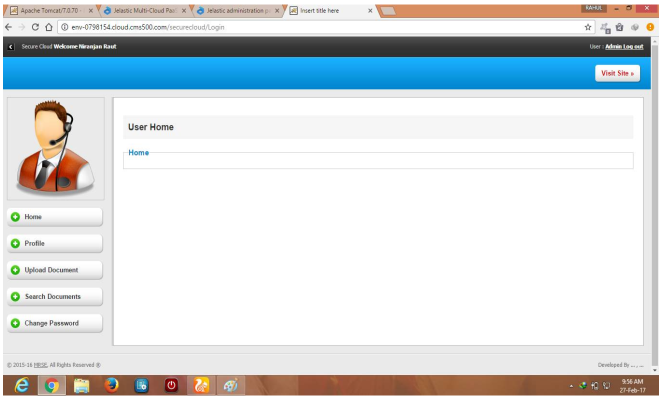
Login

You just need to hit the button and you're in!

Username: surwasemadhu@gmail.com

Password: •••

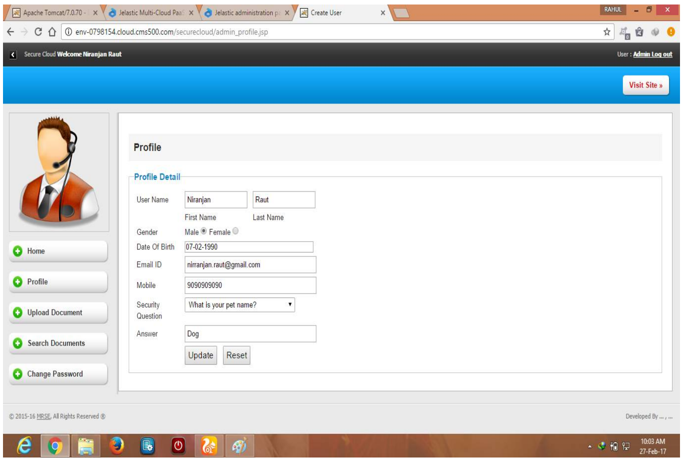
Fig.5.Web Login Page



User Home

Home

Fig.6.User Home Page


Apache Tomcat/7.0.70 -
Jelastic Multi-Cloud PaaS
Jelastic administration p
Create User
RAHUL
env-0798154.cloud.cms500.com/securecloud/admin_profile.jsp
Secure Cloud Welcome Niranjan Raut
User : Admin Log out
Visit Site »
Home
Profile
Upload Document
Search Documents
Change Password
Update
Reset
© 2015-16 MRSE, All Rights Reserved ®
Developed By ... , ...
10:03 AM
27-Feb-17

Profile

Profile Detail

User Name: Niranjan (First Name) Raut (Last Name)

Gender: Male ● Female ○

Date Of Birth: 07-02-1990

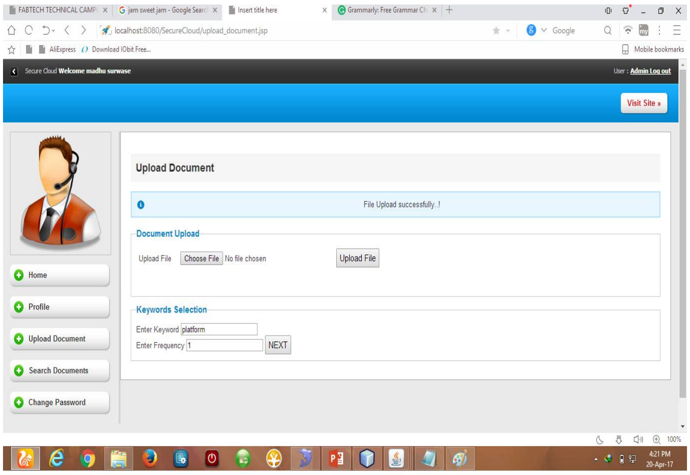
Email ID: nirranjan.raut@gmail.com

Mobile: 9090909090

Security Question: What is your pet name?

Answer: Dog

Fig.7.User Profile



Upload Document

File Upload successfully..!

Document Upload

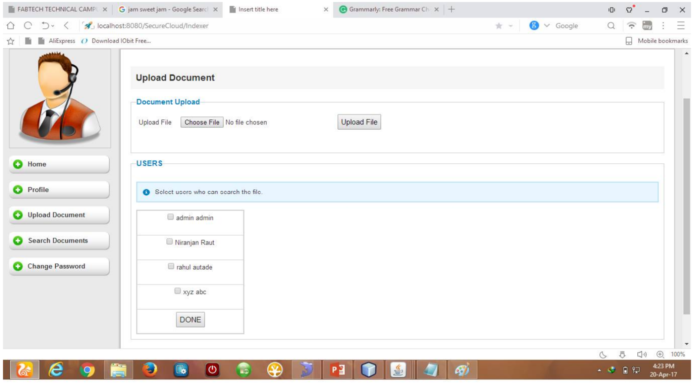
Upload File [Choose File] No file chosen [Upload File]

Keywords Selection

Enter Keyword: platform

Enter Frequency: 1 [NEXT]

Fig.8.Upload Document



Upload Document

Document Upload

Upload File [Choose File] No file chosen [Upload File]

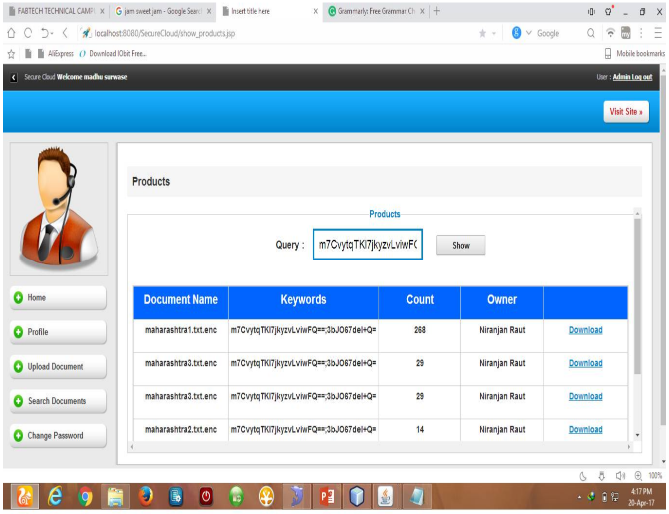
USERS

Select users who can search the file.

□ admin admin
□ Niranjan Raut
□ rahul autade
□ xyz abc

[DONE]

Fig.9.Document Uploaded Details



Products

Products

Query : m7CvytqTKI7jkyzvLviwF([Show]

Document Name	Keywords	Count	Owner	
maharashtra1.txt.enc	m7CvytqTKI7jkyzvLviwFQ==;3bJO67del+Q=	268	Niranjan Raut	Download
maharashtra3.txt.enc	m7CvytqTKI7jkyzvLviwFQ==;3bJO67del+Q=	29	Niranjan Raut	Download
maharashtra3.txt.enc	m7CvytqTKI7jkyzvLviwFQ==;3bJO67del+Q=	29	Niranjan Raut	Download
maharashtra2.txt.enc	m7CvytqTKI7jkyzvLviwFQ==;3bJO67del+Q=	14	Niranjan Raut	Download

Fig.10.User Query and Result

Result Analysis:

To measure the practicality and usability of our mechanism, owner has uploaded encrypted file and create index and generate trapdoor, using this trapdoor user got the result and this result is display on some ranking order, using secret key user got decrypt file.Ranking is depends on following way:

TF(t) = (Number of times term t appears in a document) / (Total number of terms in the document).

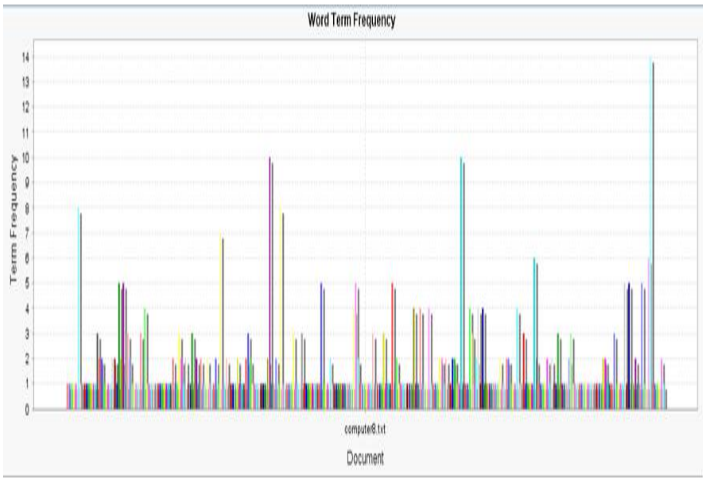


Fig.11.Word Term Frequency

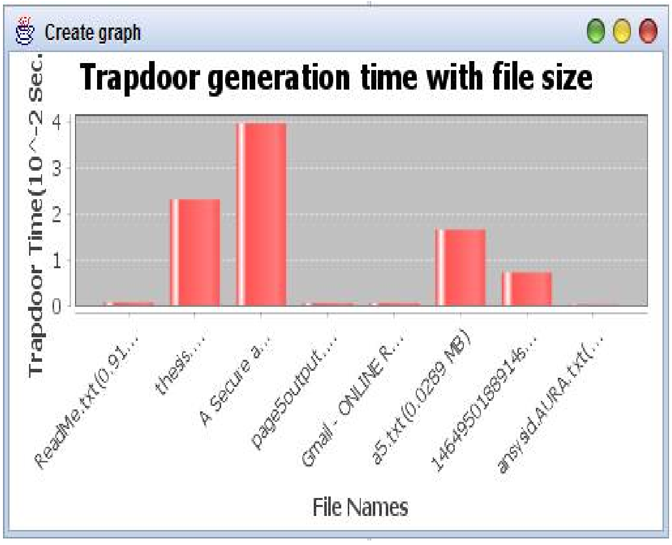


Fig.12.Trapdoor Generation Time with File Size

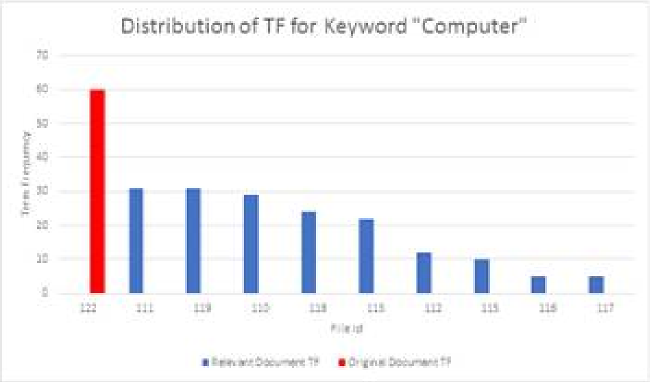


Fig.13. Distribution of TF for Keyword"Computer"

Fig.13 indicate that TF for original document in keyword "Computer" and TF for Relevent document in Keyword "Computer" .relevent documents means which documents present the keyword Computer and given the file ID.

VI. CONCLUSION AND FUTURE WORK

In this system, a secure, efficient and dynamic search scheme is proposed, which supports not only the accurate multi-keyword ranked search but also the dynamic deletion and insertion of documents.We have proposed a secured way of accessing files from cloud. We proposed a secured and efficient scheme for data owner to provide better services to the users. The owner side encryption scheme and index file generation helps the data user to get secure and protected data with better QOS. To improve the QOS a client side ranking process has been adopted. Searching the query in the index file rather than the file system cloud server can give quick response very quickly. Our proposed scheme fulfills the security requirements of multi keyword top-S retrieval over the encrypted cloud data.

.

In the proposed scheme, the data owner is responsible for generating updating information and sending them to the cloud server. Thus, the data owner needs to store the encrypted index and the information that are necessary to recalculate the TF values. Such an active data owner may not be very suitable for the cloud computing model.

It could be a meaningful but difficult future work to design a dynamic searchable encryption scheme whose updating operation can be completed by cloud server only, meanwhile reserving the ability to support multi-keyword ranked search. In addition, as the most of works about searchable encryption, our scheme mainly considers the challenge from the cloud server. Actually, there are many secure challenges in a multi-user scheme. Firstly, all the users usually keep the same secure key for trapdoor generation in a symmetric SE scheme. In this case, the revocation of the user is big challenge. If it is needed to revoke a user in this scheme, we need to rebuild the index and distribute the new secure keys to all the authorized users. Secondly, symmetric SE schemes usually assume that all the data users are trustworthy.

It is not practical and a dishonest data user will lead to many secure problems. For example, a dishonest data user may search the documents and distribute the decrypted documents to the unauthorized ones. Even more, a dishonest data user may distribute his/her secure keys to the unauthorized ones. In the future works, we will try to improve the SE scheme to handle these challenge problems.

REFERENCES

[1] K. Ren, C.Wang, Q.Wang et al., "Security challenges for the public cloud," IEEE Internet Computing, vol. 16, no. 1, pp. 69–73, 2012.
[2] S. Kamara and K. Lauter, "Cryptographic cloud storage," in Financial Cryptography and Data Security. Springer, 2010, pp. 136–149.
[3] C. Gentry, "A fully homomorphic encryption scheme," Ph.D. dissertation, Stanford University, 2009.
[4] O. Goldreich and R. Ostrovsky, "Software protection and simulation on oblivious rams," Journal of the ACM (JACM), vol. 43, no. 3, pp. 431–473, 1996.
[5] 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.
[6] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, "Public key encryption that allows pir queries," in Advances in Cryptology-CRYPTO 2007. Springer, 2007, pp. 50–67.
[7] D. X. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on. IEEE, 2000, pp. 44–55.
[8] E.-J. Goh et al., "Secure indexes." IACR Cryptology ePrint Archive, vol. 2003, p. 216, 2003.
[9] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in Proceedings of the Third international conference on Applied Cryptography and Network Security. Springer-Verlag, 2005, pp. 442–455.
[10] 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.
[11] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in INFOCOM, 2010 Proceedings IEEE. IEEE, 2010, pp. 1–5.
[12] M. Kuzu, M. S. Islam, and M. Kantarcioglu, "Efficient similarity search over encrypted data," in Data Engineering (ICDE), 2012 IEEE 28th International Conference on. IEEE, 2012, pp. 1156–1167.
[13] C. Wang, K. Ren, S. Yu, and K. M. R. Urs, "Achieving usable and privacy-assured similarity search over outsourced cloud data," in INFOCOM, 2012 Proceedings IEEE. IEEE, 2012, pp. 451–459.
[14] B. Wang, S. Yu, W. Lou, and Y. T. Hou, "Privacy-preserving multikeyword fuzzy search over encrypted data in the cloud," in IEEE INFOCOM, 2014.
[15] 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.
[16] Y. H. Hwang and P. J. Lee, "Public key encryption with conjunctive keyword search and its extension to a multi-user system," in Proceedings of the First international conference on Pairing-Based Cryptography. Springer-Verlag, 2007, pp. 2–22.
[17] L. Ballard, S. Kamara, and F. Monrose, "Achieving efficient conjunctive keyword searches over encrypted data," in Proceedings of the 7th international conference on Information and Communications Security. Springer-Verlag, 2005, pp. 414–426.
[18] D. Boneh and B. Waters, "Conjunctive, subset, and range queries on encrypted data," in Proceedings of the 4th conference on Theory of cryptography. Springer-Verlag, 2007, pp. 535–554.