Privacy Preserving And Multi Keyword Search Over Encrypted Outsource Data.

Ravindra Hyalij [1]

[1]P.G. Student, Department of Computer Engineering, SKN Sinhgad Institute of Technology & Science, Lonavala, Pune, Maharashtra, India

ABSTRACT

Cloud computing offers flexible computation and resources for storage purpose, but here user poses challenges on verifiability of computations and data privacy. Because of data privacy reason some data owner away from cloud facility. Proposed system resolve this problem by providing facility of verifiability for privacy preserving multi-keyword search over outsourced documents. In this system integration of two technique use first is homomorphic MAC technique and second one is privacy-preserving multi-keyword search. The proposed scheme enables the client to verify search results efficiently without storing a local copy of the outsourced data.

Keyword: *Decryption, Encryption, Multi keyword Search technique, Private Cloud, Public Cloud*.

1. INTRODUCTION

Cloud computing provide flexible computation to achieve computation in less time and storage space for storing huge amount of data. In this proposed system client firstly encrypt his data then outsource it and later access data from anywhere / any device by using single keyword and multi keyword search. The client can additionally request that cloud server to perform some computation over his data on his behalf. For access outsource data client have to create trapdoor. For creating trapdoor client register their identity token on data owner side then owner give him decryption key of data. But, here user faces some challenges like verifiability of computations and data privacy. In this study the focus is on verifiability for privacy-preserving multi-keyword search over outsourced documents.

In previous system, sometimes cloud server may not return correct result due to system fault or incentive to reduce computation cost, due to this behaviour of cloud it create critical problem. In previous system user should have a local copy of data on which he wants to perform computation. This technique is used to increase the efficiency of searching over the cloud data. This technique provide security for cloud data and also provide facility of searching by using multiple keyword. Data owner does not need to store local copy of data.

Privacy preserving multi keyword search technique:
Proposed system design by integrating ’Homomorphic MAC technique’ with ’Privacy Preserving Multi keyword search technique’. This proposed technique provide facility to verify search result without store the local copy of outsource data. It also display semantic result for multi keyword searching.

Supporting fine-grained access control:
Through constructing an access tree for each document in database as the access policy by using the cipher text policy attribute based encryption techniques, our scheme can achieve the search user authorization. That is, even users with different attributes have the same keyword, they will receive different search results.

We generate an authentication tag by the reciprocal of data owners secret value for each encrypted file to prevent cloud sever from returning false or inaccurate search results. In this way, our solution can realize both search results verification.

2. RELATED WORK

In this area, displaying the diverse technique to take care of the issue related the cloud security: Here they are built up an arrangement of troublesome protection necessities for such a safe cloud information usage framework. Among different multi-keyword semantics, they pick the efficient closeness measure of "organize coordinating", i.e., however many matches as would be prudent, to capture the significance of information reports to the inquiry question. Likewise further utilize "internal item likeness" to quantitatively assess such closeness measure. They first propose an essential thought for the MRSE in light of secure inward item calculation, and after that give two enhanced MRSE plans to accomplish different stringent protection prerequisites in two diverse danger models. Exhaustive examination recognizing security and efficiency certifications of proposed plans is given. [1].

In another examination proposed an efficient comparability searchable Symmetric encryption plot. To do as such, they used locality sensitive hashing which is utilized for quick similarity search in high dimensional spaces for plain information. They proposed LSH based secure file and a pursuit plan to begin fast similarity seek with regards to scrambled information. In such a context, it is extremely basic not to misfortune the confidentiality of the delicate information while giving usefulness. They have provided a thorough security definition and demonstrated the security of the proposed plan under the gave definition to ensure the confidentiality. To clear up the points of interest of the proposed scheme, we introduced a certifiable use of it, in particular the error recognizing watchword seek. This application empowers keyword search which is tolerant to the writing mistakes both in the inquiries and the information sources. [2].

This examination first endeavors the well known similitude measure, i.e., vector space display with cosine measure, to viably obtain the precise query item. They proposed two secure list plans to meet different protection prerequisites in the two danger models. In the end, the spillage of touchy recurrence data can be stayed away from.

To lift look effectiveness, they propose a tree-based file structure for the entire archive set. From the usage of the model of our safe inquiry framework, distinguish three fundamental productivity related elements, by which the effectiveness of the pursuit calculation upon our file tree can be essentially moved forward. What's more, entire pursuit prepare put forth evident in defense that clients need to guarantee the genuineness of the returned seek results.[3]

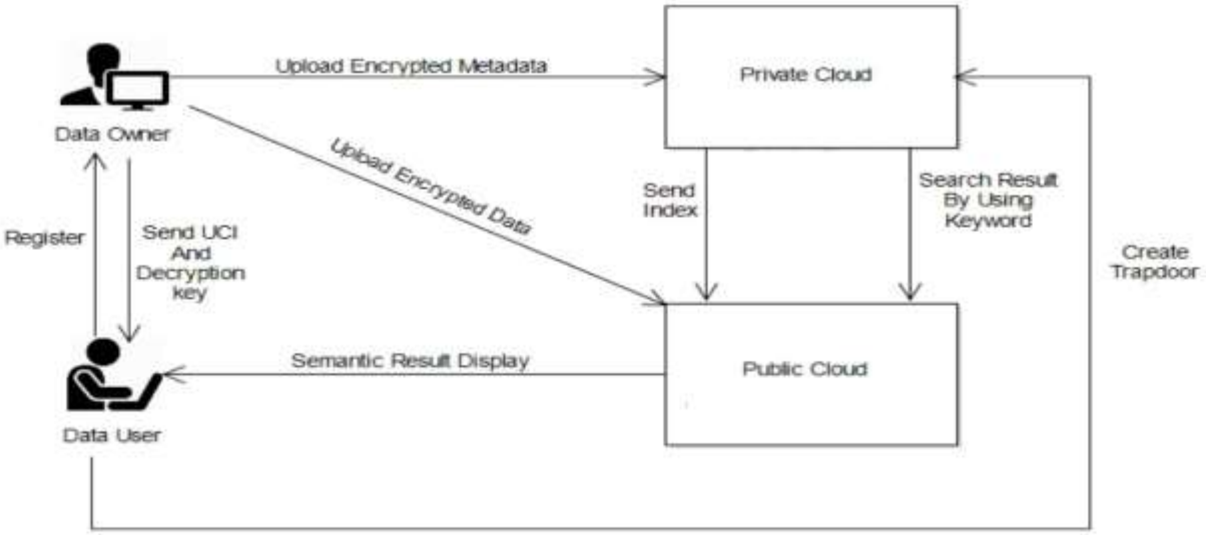
This approach guarantees that lone the most related things are recovered by the client, counteracting superfluous correspondence and calculation trouble on the client. Framework executes the entire framework and shows the adequacy and efficiency of our answer through tests utilizing the openly accessible Enron dataset. Our investigation portrayed that the proposed plan is turned out to be secure, privacy-saving, efficient and powerful.[4]

This exploration handled the testing multi-catchphrase fluffy hunt issue over the encoded information. In that proposed and incorporated a few new plans to tackle the different catchphrases seek and the fluffy pursuit issues at the same time with high efficiency. Our approach of utilizing LSH works in the Bloom filter to build the file list is novel and gives an efficient answer for the protected fluffy pursuit of different watchwords.[5]

3. PROBLEM STATEMENT

Although cloud computing offers elastic computation and storage resources, it poses challenges on verifiability of computations and data privacy. In this work we investigate verifiability for privacy-preserving multi-keyword search over outsourced documents by using homomorphic MAC technique. As the cloud server may return incorrect results due to system faults or incentive to reduce computation cost, it is critical to offer verifiability of search results and privacy protection for outsourced data at the same time. To fulfill these requirements, we design a verifiable Privacy-preserving keyword Search scheme, called VP Search, by integrating an adapted homomorphic MAC technique with a privacy-preserving multi-keyword search scheme.

4. ARCHITECTURE AND DESIGN



Data Owner: Data owner is who creates his data and upload on cloud.
Data User: Data user is who use data owner data by using decryption key.
Private Cloud: also known as an internal or enterprise cloud, resides on company's intranet or hosted data center where all of your data is protected behind a firewall.
Public Cloud: Your data is stored in the provider's data center and the provider is responsible for the management and maintenance of the data center.

Operations Details:
1] Data owner encrypt data and upload it on public cloud. Then he creates metadata for encrypted data.
2] He encrypts created metadata and uploads it on private cloud.
3] Encrypted metadata index uploaded on public cloud.
4] When data user wants to search data owner data, first he needs to create trapdoor.
5] After successfully creation of trapdoor data user search data over public cloud using Keywords.
6] This system use homomorphic MAC technique and multi keyword search technique to show semantic result.
7] Data user registers him on data owner side and request for data decryption key.
8] By using decryption key data user decrypt required data and use it.

5. METHODS AND SCHEME

Set Theory:
N : A universal attribute set {1. . . n} for some nature number n.
G : Access structure space**.**
W : Keyword space comprised of keywords w.
I : An attribute set used for an access structure GT Ɛ G on an encrypted index and I
S : An attribute set for a user secret key SK and S
i : An attribute in N either refers to a positive attribute I or its negation :i.
D : An encrypted index for a file.
Q : A trapdoor for an intended keyword w 2 W.
rk :A proxy re-encryption key set.
PSK : A user's partial secret key.
F : An attribute set containing the attributes to be updated.
D : An attribute set including all the attributes in D's access structure with the re-encryption keys not being 1 in rk.
V : An attribute set containing all the attributes in PSK with the re-encryption keys not being 1 in rk.

5.1 Secure Key Generation:

Secure index generation. Before outsourcing a file to the CS, the data owner calls EncIndex algorithm to generate a secure index D for this file.

5.2 Search:

Upon receipt of a trapdoor Q and the user identity IDf,
1) The CS finds out if IDf exists on the user list of the target dataset. If not, the user is not allowed to search over the dataset
2) otherwise, the CS continues the Search algorithm with the input of trapdoor Q, encrypted index D and
Df from the user list. We call this process dataset search authorization.

5.3 Index Generation:

Our proposed method utilizes the idea of bucketization which is a data partitioning technique widely used in literature. Here, each object is distributed into several buckets via min hash functions introduced in III-A and the bucket-id is used as an identifier for each object in that bucket. This method maps objects such that the number of buckets, in which two objects collide, increases as the similarity between those objects increases. In other words, while two identical objects collide in all of the buckets, number of common buckets decreases as dissimilarity between objects increases. The proposed secure index is generated by the data owner utilizing the following phases, namely: feature extraction, bucket index construction and bucket index encryption.

6. ANALYSIS AND DISCUSSION

In proposed system use multi-keyword by using this if data user search any keyword in outsource data the files which contain related keywords also display, the files which contain most of the words will also be rank forward. e.g. if you search for keyword Protocol then system also return the files which contain internet, network, authentication. Here design an efficient, verifiable and privacy-preserving multi-keyword search for outsourced cloud data under the partially honest cloud server model. It is realized by integrating an adapted homomorphic technique with a privacy preserving multi-keyword search scheme. The proposed scheme is exceptionally effective as it depends on just a single route work for security. In this system provide detailed analysis on security, privacy, verifiability and efficiency of VP Search. Specially, the underlying homomorphic MAC scheme used in VP Search can be proved to be secure. Here implement VP Search using java for implementation and evaluate its performance over three UCI (Unique Client Identification) data sets. VP Search is very efficient on authentication tag generation and keyword search operations.

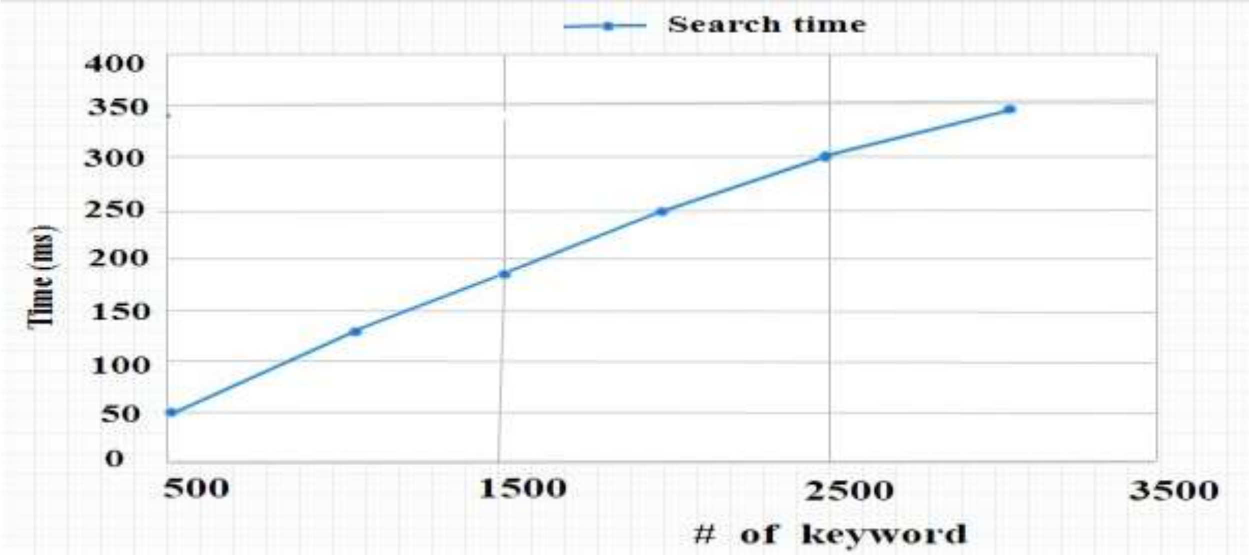


Chart -1: Inner Product Computation Related To The File Length.

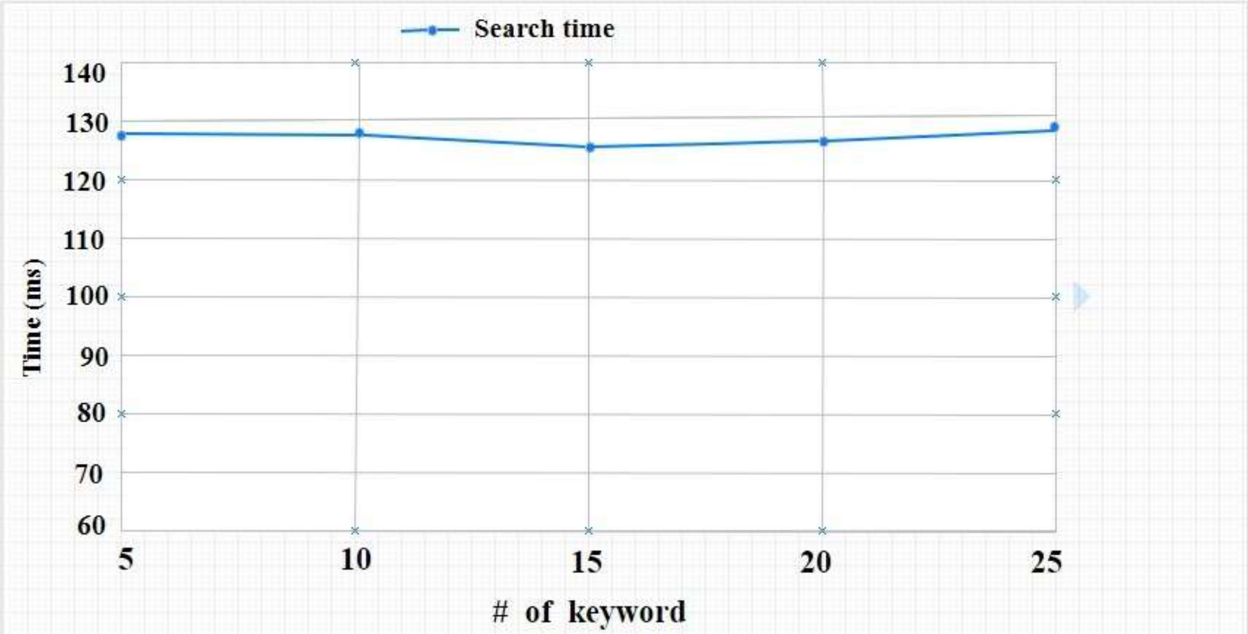
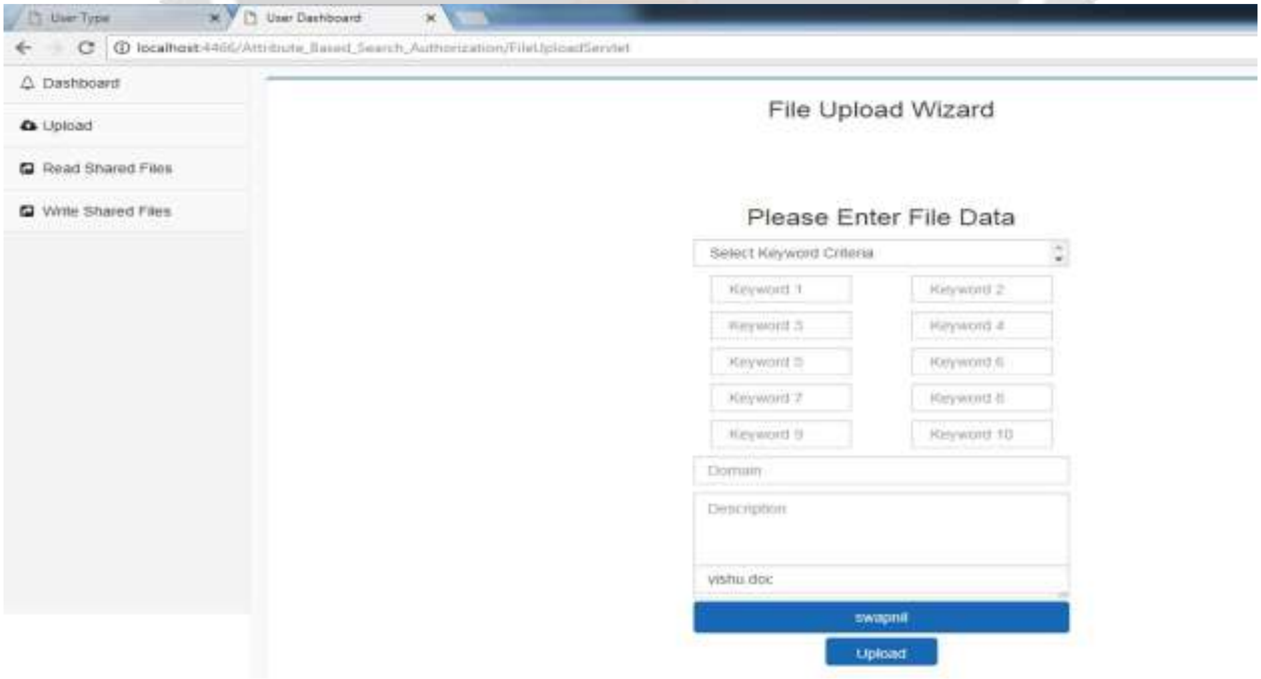


Chart -2: Inner Product Computation Not The Number Of Keywords.

7. RESULTS

The application will be able to Encryption of documents at clients machine. A secure system having , efficient and dynamic searching and storing system using different algorithms and models. Obtains better search efficiency and linear search, greater time efficiency. Data owner will responsible for generating index tree is a updation in system.

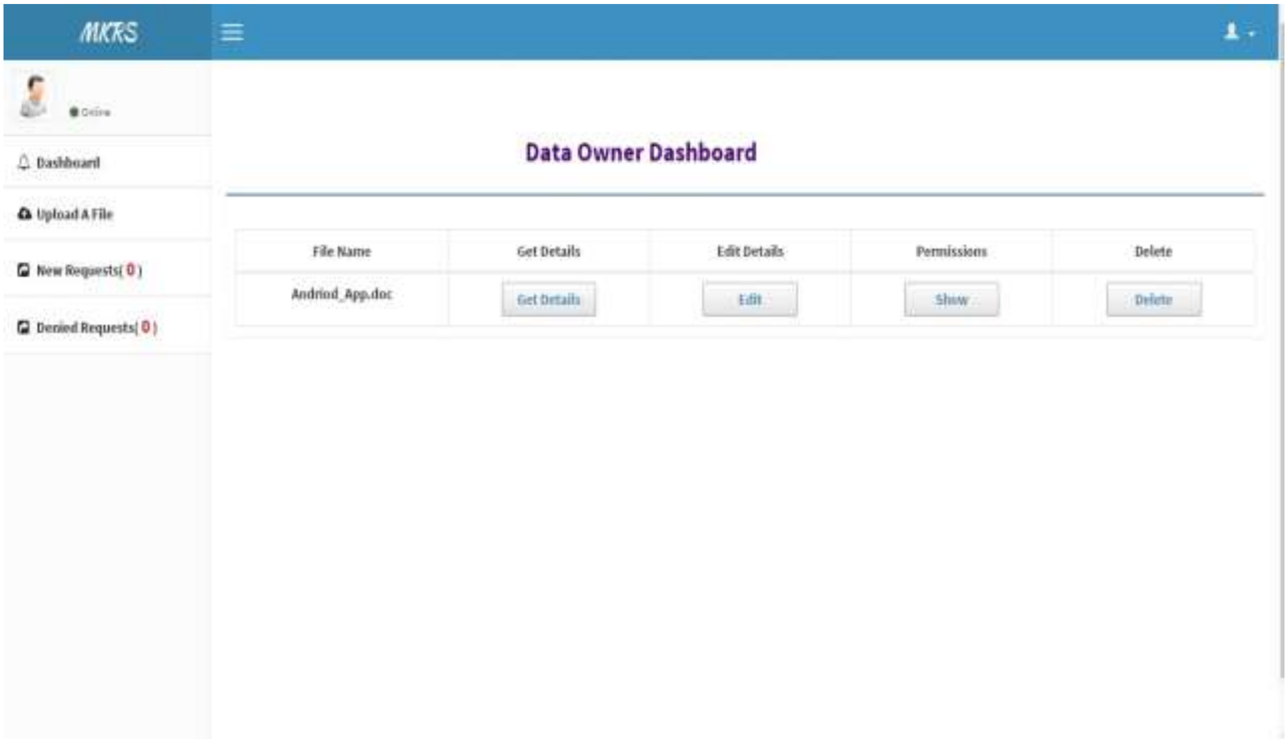


localhost:4466/Attribute_Based_Search_Authorization/FileUploadServlet

vishu.doc

swapnil

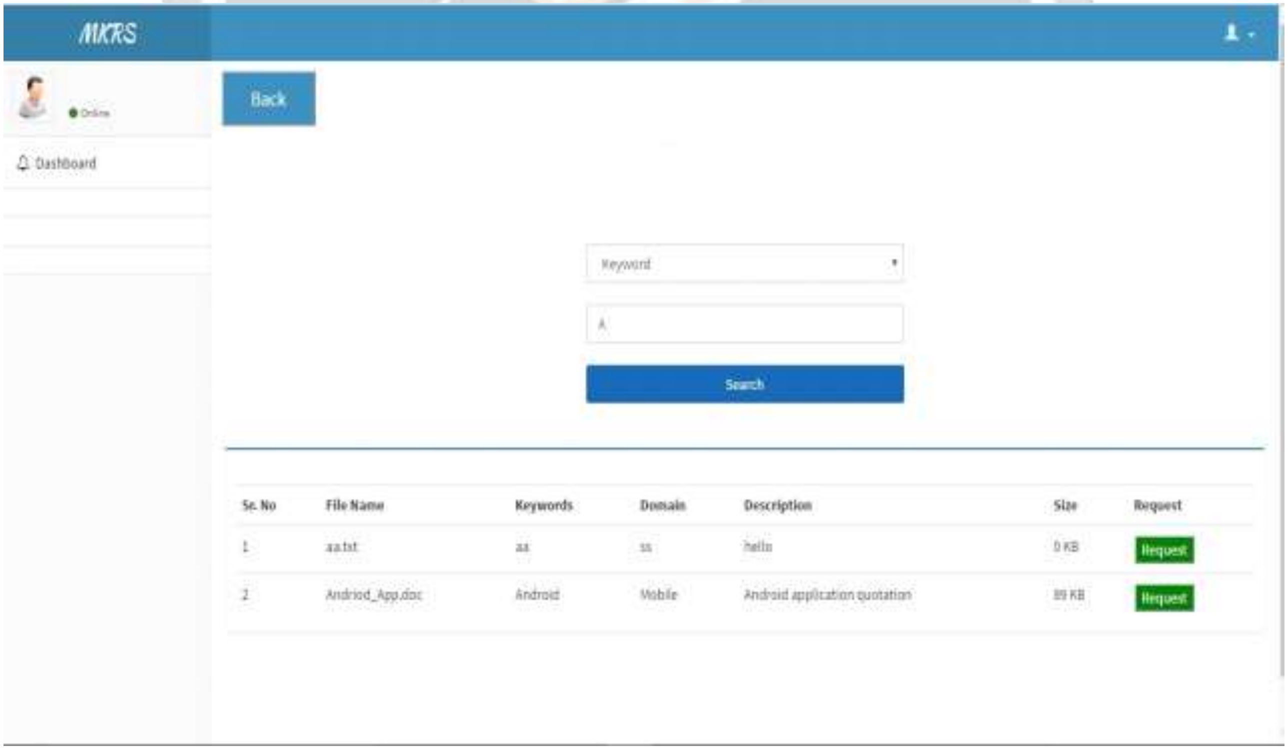
Fig -1: Multi-Keyword with file Upload



Data Owner Dashboard

File Name	Get Details	Edit Details	Permissions	Delete
Andriod_App.doc	Get Details	Edit	Show	Delete

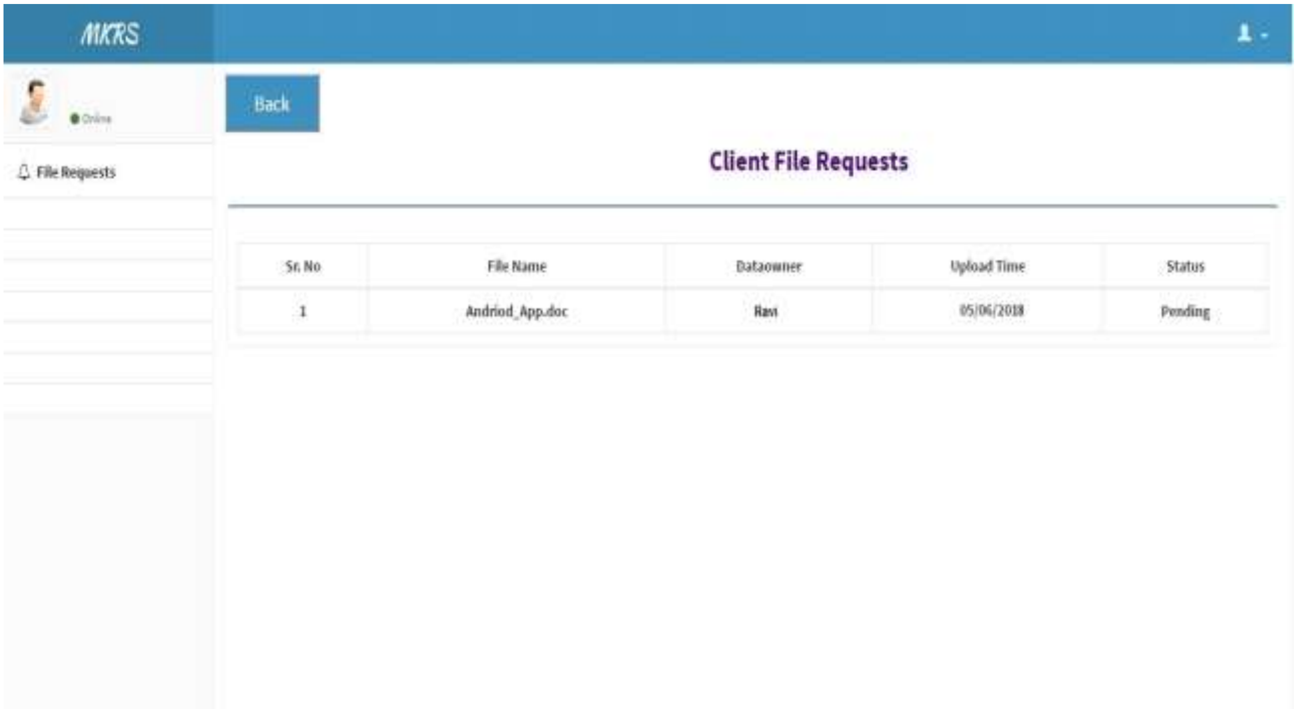
Fig -2: Uploaded File Details



A

Sr. No	File Name	Keywords	Domain	Description	Size	Request
1	aa.txt	aa	ss	hello	0 KB	Request
2	Andriod_App.doc	Android	Mobile	Android application quotation	89 KB	Request

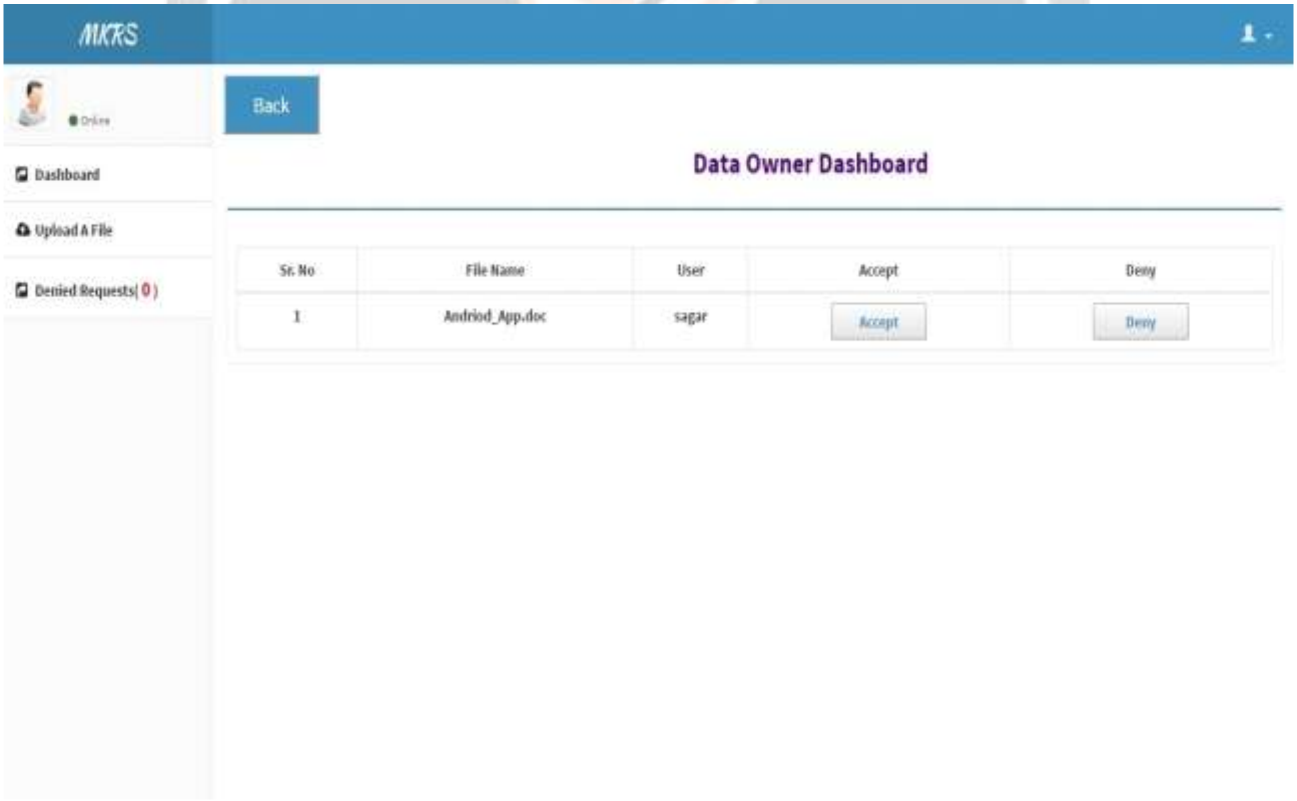
Fig -3: Data User Search File using keyword



Client File Requests

Sr. No	File Name	Dataowner	Upload Time	Status
1	Andriod_App.doc	Ravi	05/06/2018	Pending

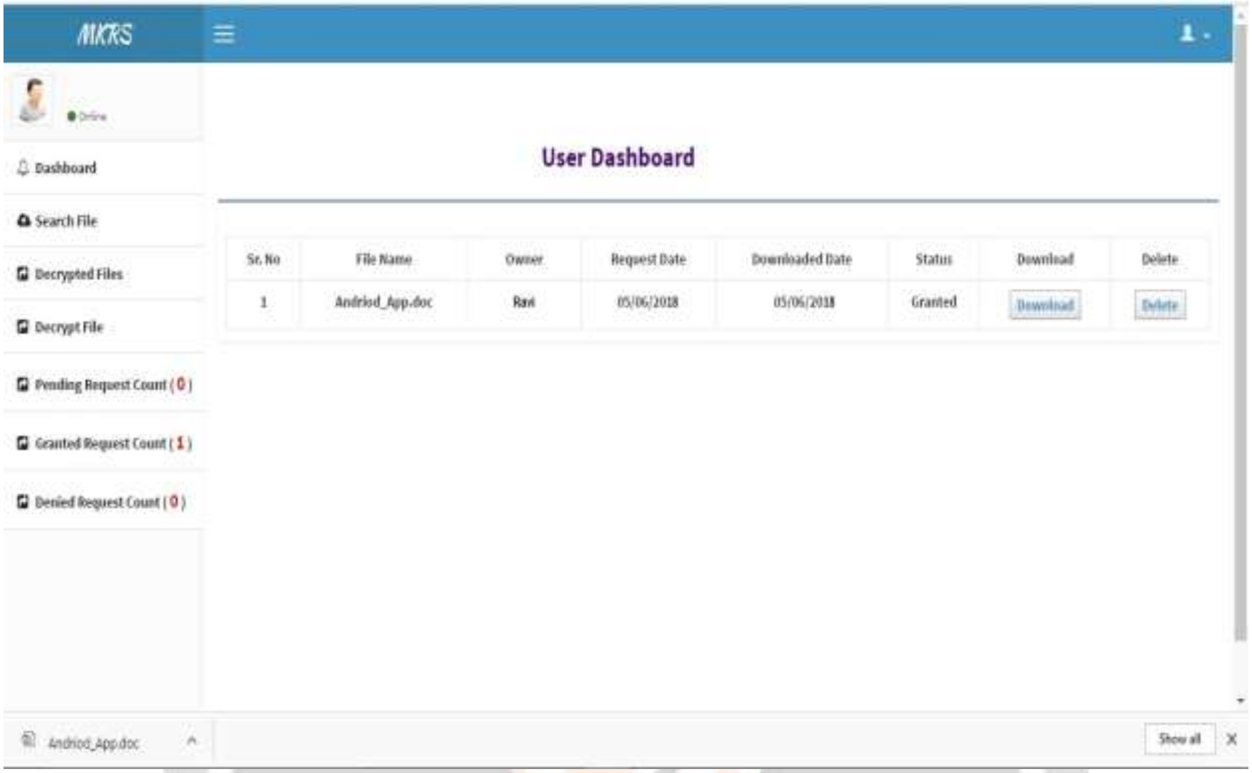
Fig -4: File access Status



Data Owner Dashboard

Sr. No	File Name	User	Accept	Deny
1	Andriod_App.doc	sagar	Accept	Deny

Fig -5: File Access Request


MKRS
Online
Dashboard
Search File
Decrypted Files
Decrypt File
Pending Request Count (0)
Granted Request Count (1)
Denied Request Count (0)
Andriod_App.doc
Show all

User Dashboard

Sr. No	File Name	Owner	Request Date	Downloaded Date	Status	Download	Delete
1	Andriod_App.doc	Ravi	05/06/2018	05/06/2018	Granted	Download	Delete

Fig -6: File Download

8. CONCLUSION

In this study, 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 construct a special keyword balanced binary tree as the index, and propose a Greedy Depth-rst Search algorithm to obtain better efficiency than linear search. In addition, the parallel search process can be carried out to further reduce the time cost. The security of the scheme is protected against two threat models by using the secure AES algorithm. Experimental results demonstrate the efficiency of our proposed scheme. There are still many challenge problems in symmetric SE schemes. 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 unencrypted index tree and the information that are necessary to recalculate the IDF values. Such an active data owner may not be very suitable for the cloud computing model. Mean while reserving the ability to support multi-keyword ranked search. Finally both the approaches like secure and accurate searching as well as dynamic operations are been performed.

9. FUTURE SCOPE

In large-scale databases (Outsource data) VP-Search method is not efficient since the cloud needs to search through the whole database, which is very inefficient. In future we have some work in this line that will be enhancements for efficient verification for large-scale outsourced data. This system works on semi trusted cloud but in future it will be extended up to all types of cloud environment and can provide better security. Furthermore in future we can extend our search scheme to use external storage more carefully while maintaining privacy.

REFERENCES

[1] Jianfeng Wang, Xiaofeng Chen (2016) Efficient and Secure Storage for Outsourced Data: A Survey. In: Springer Data Sci. Eng.1(3):178-188

[2] Ali Gholami and Erwin Laure (2015) Security and privacy of sensitive data in cloud computing: A survey of recent developments. HPCViz Dept., KTH- Royal Institute of Technology, Stockholm, Sweden, 10.5121/csit.2015.51611

[3] Syam Kumar Pasupuleti, Subramanian Ramalingam, Rajkumar Buyya (2016) An efficient and secure privacy-preserving approach for outsource data of resource constrained mobile devices in cloud computing. Journal of Network and Computer Applications 12-22

[4] Raghavendra S, Chitra S Reddy, Geeta C M, Rajkumar Buyya, Venugopal K R, S S Iyengar, L M Patnaik (2016) Survey on Data Storage and Retrieval Techniques over Encrypted Cloud Data. In: International Journal of Computer Science and Information Security (IJCSIS), Vol. 14, No. 9

[5] Anna Monreale, Wendy Hui Wang (2016) Privacy-Preserving Outsourcing of Data Mining. In: IEEE 40th Annual Computer Software and Applications Conference

[6] Dario Catalano (2014) Homomorphic Signatures and Message Authentication Codes. Universita di Catania, Italy

[7] Rosario Gennaro, Craig Gentry, and Bryan Parno. Non-Interactive Verifiable Computing: Outsourcing COmputation to Untrusted Workers.In: IBM T.J. Watson Research Center and CyLab, Carnegie Mellon University

[8] Dawn Xiaodong Song, David Wagner, Adrian Perrig. Practical Techniques for Searches on Encrypted Data. In: University of California, Berkeley

[9]Ms. Archana D. Narudkar, Mrs. Aparna A. Junnarkar (2015) A Survey on Searching Techniques over Encrypted Data. In: International Journal of Computer Science and Information TEchnologies, Vol. 6(2), 1007-1010

[10] Vishal R. Pancholi and Dr. Bhadresh P. Patel (2016) Enhancement of Cloud Computing Security with Secure Data Storage using AES. In IJIRST - International Journal for Innovative Research in Science & Technology| Volume 2| Issue 09