A STUDY ON PRIVACY PRESERVING RANKED MULTI- KEYWORD SEARCH FOR MULTIPLE DATA OWNERS IN CLOUD COMPUTING.

Sri Ramakavacham Prudhvi Raj [1]

1. Working as Assistant Professor in the department of Computer Science & Engineering, Sri Indu Institute of Engineering and Technology, Sheriguda (Village), Ibrahimpatnam (Mandal), Rangareddy (Dist), Hyderabad, Telangana, 501510, India. Email id: **Abstract:**

With the rapid development of various multimedia technologies, more and more multimedia data are generated and transmitted in the medical, also the internet allows for wide distribution of digital media data. It becomes much easier to edit, modify and duplicate digital information .Besides that, digital documents are also easy to copy and distribute, therefore it will be faced by many threats. Privacy preserving Ranked Multi-keyword Search in a Multi-owner model (PRMSM). To enable cloud servers to perform secure search without knowing the actual data of both keywords and trapdoors, in this proposed system systematically construct a novel secure search protocol. To rank the search results and preserve the privacy of relevance scores between keywords and files, we propose a novel Additive Order and Privacy Preserving Function family. To prevent the attackers from eavesdropping secret keys and pretending to be legal data users submitting searches, we propose a novel dynamic secret key generation protocol and a new data user authentication protocol. Furthermore, PRMSM supports efficient data user revocation. Extensive experiments on real-world datasets confirm the efficacy and efficiency of PRMSM.

Many people are using the cloud storage for storing their large amount of data. Not only by individuals many companies, are industrialists also using the cloud storage. Day-by-day the amount of people using the cloud storage is increasing due to its easiness of use. The data that has been stored in cloud may contain some secret documents also. Thus a secured storage and secured data retrieval is necessary. Many searchable algorithms for cloud is existing. But less of them provide proper protection for the data that is stored. To increase confidentiality in the case of multiple data owners a tree-based ranked multi-keyword search scheme can be used. By considering a large amount of data in the cloud, the TF-IDF model is used to develop a multi-keyword search and return the top search results. The cloud server also uses a depth first search algorithm to find the corresponding file from the cloud.

Keyword- AES, DES, 3DES, Cloud Computing, Ranking etc

I INTRODUCTION

Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network (typically the Internet). The name comes from the common use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. Cloud computing entrusts remote services with a user's data, software and computation. Cloud computing consists of hardware and software resources made available on the Internet as managed third-party services. These services typically provide access to advanced software applications and high-end networks of server computers. In cloud storage the data is stored in logical pools as digital data. In multi-owner scenario, the same data will contain several owners.

A main server will be there to handle the entire data. The cloud may contain multiple servers may be reside in multiple locations. The main server or the cloud storage providers will be responsible for protection and handling of the stored data. The cloud users will buy or lease the storage capacity from these cloud storage providers. Cloud storage enables distributed and scalable network access to the digital data. A problem that has to be faced in cloud storage is the secured search over the encrypted data.

The most challenging task in cloud storage is secured search on encrypted cloud data. There are various search schemes exist. But they results either in system overhead or sometimes those methods will be really hard to implement over large data sets. To prevent the unauthenticated access the data will be stored in cloud as in the encrypted form.

To provide an efficient search, a tree based multi keyword search scheme is constructed [1]. The words that are seemed as keywords for a document are identified and an index is formed. All the indexes such formed are then merged into one. For each search requests a depth first search is used to identify the corresponding data file of the user. The TF-IDF model is used to return the top results. A depth first search is used to perform efficient search.

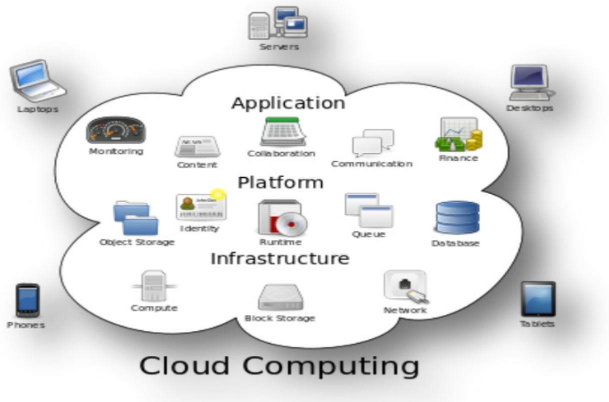


Fig 1 Cloud Computing Architecture

Computing resources are shared by many users. The benefits of cloud can be extended from individual users to organizations. The data storage in cloud is one among them. The virtualization of hardware and software resources in cloud nullifies the financial investment for owning the data warehouse and its maintenance. Many cloud platforms like Google Drive, cloud; Sky Drive, Amazon S3, Drop box and Microsoft Azure provide storage services. Security and privacy concerns have been the major challenges in cloud computing. The hardware and software security mechanisms like firewalls etc. have been used by cloud provider. Due to advantage of cloud services, more and more sensitive information are being centralized into the cloud servers, such as emails, personal health records, private videos and photos, company finance data, government documents, etc.

II Literature survey

Cloud computing, the long-held dream of computing as a utility, has the potential to transform a large part of the IT industry, making software even more attractive as a service and shaping the way IT hardware is designed and purchased. Developers with innovative ideas for new Internet services no longer require the large capital outlays in hardware to deploy their service or the human expense to operate it. Various methods are used for searching of data files in cloud for multi-owner scenario. Some of them are discussed below.

2.1 Practical Techniques for Searches on Encrypted Data

Dawn Xiaodong et al. [2] proposed a method for searching without any loss of data confidentiality. If a mobile user wants to retrieve the documents containing a particular keyword from the mail storage server with limited bandwidth. The problem is the server has to know about the content of the documents. So the problem is to support the search queries without revealing all the data. The servers must be trusted and must not reveal the data without proper authorization. The unfrosted server leads to undesirable security and privacy risks in applications.

The untreated server must not learn anything about the plaintext rather than the ciphertext. So that the untrusted server cannot search for a word without the user's authorization by using the techniques of controlled searching. The user can ask the untrusted server to search for a secret word without revealing the word to the server by supporting hidden queries. The untrusted server learns nothing more than the search result about the plaintext by supporting query isolation.

First the problem of searching on encrypted data is defined. Assume user A has a set of documents and stores them on an untrusted server S. For example, A could be a mobile user who stores her email messages on an untrusted mail server. Because S is untrusted, A wishes to encrypt her documents and only store the ciphertext on S. Each document can be divided up into 'words'. Each 'word' can be any token such as a word or a sentence. The user A may have only a low-bandwidth network connection to the server S, he/she wishes to only retrieve the documents which contain the word W. In order to achieve this goal, we need to design a

A. Secured Multi-keyword Ranked Search over Encrypted Cloud Data:

In cloud computing data possessor are goaded to farm out their complex data management systems from local sites to the commercial public cloud for greater flexibility and economic savings. To ensure safety of stored data, it is must to encrypt the data before storing. Necessary to invoke search with the encrypted data also.

B. Privacy Preserving Keyword Searches on Remote Encrypted Data:

Consider the problem: a user U wants to store his files in an encrypted form on a remote file server S. Later the user U wants to efficiently retrieve some of the encrypted files containing specific keywords, keeping the keywords themselves secret and not to endanger the security of the remotely stored files.

C. Search on Encrypted Cloud Data:

Today's pay-as-you-use cloud paradigm. This paper has defined and solved the problem of effective yet secure ranked keyword search over encrypted cloud data [4]. Ranked search greatly enhances system usability by returning the matching files in a ranked order regarding to certain relevance criteria (e.g., keyword frequency) thus making one step closer towards practical deployment of privacy-preserving data hosting services in Cloud Computing.

2.2 Secure Index for Resource-Constraint Mobile Devices in Cloud Computing

Hanbing Yao et al. [3] proposed a secure index based on counting Bloom filter (CBF) for ranked multiple keywords search. Nowadays more organizations and users are outsourcing their data into cloud server. In order to protect data privacy, the sensitive data have to be encrypted, which increases the heavy computational overhead and brings great challenges to resource-constraint devices. In this scheme, several algorithms are designed to maintain and lookup CBF, while a pruning algorithm is used to delete the repeated items for saving the space.

The problem of secure ranked search over encrypted data in the cloud server is discussed here. In the proposed scheme, counting Bloom filter is used to generate the secure index for ranked multiple keywords search. Moreover, several algorithms are designed to maintain and lookup CBF and a pruning algorithm is used to delete the repeat items for saving the space. The Paillier cryptosystem is employed to encrypt relevance scores. It ensures that even the same relevance scores will be encrypted into different bits, which can help to resist statistical analyses. The major computing work in rank is done by the cloud server on the encrypted relevance scores, which make the resource constraint mobile devices can easily search over encrypted data.

The Parlier cryptosystem is used to encrypt relevance scores. It will make sure that the same relevance scores are encrypted into different bits. So this can resist the statistical analyses on the cipher text of the relevance scores. Moreover, the Parlier cryptosystem supports the homomorphism addition of cipher text without the knowledge of the private key, the major computing work in ranking could be moved from user side to the cloud server side. Therefore, this scheme can effectively use in resource-constraint mobile devices such as 5G mobile terminals.

III RELATED WORK

OBJECTIVES

To enable efficient, secure and dynamic multi-keyword ranked search over outsourced encrypted cloud data under the aforementioned models, our system design should simultaneously achieve the following design goals.

1) **Dynamic Multi Owner Multi-Keyword Ranked Search**

2) **Search Efficiency**

3) **Privacy-preserving**

IV MOTIVATION

Cloud computing is the long dreamed vision of computing as a utility, where cloud customers can remotely store their data into the cloud Privacy concerns, secure searches over encrypted cloud data has motivated several research works under the single owner model. However, most cloud servers in practice do not just serve one owner; instead, they support multiple owners to share the benefits brought by cloud computing. In this paper, we propose schemes to deal with Privacy preserving Ranked Multi-keyword Search in a Multi-owner model (PRMSM).

V PROBLEM STATEMENT

Traditional authentication methods often follow three steps. First, data requester and data authenticator share a secret key, say, k0. Second, the requester encrypts his personally identifiable information d0 using k0 and sends the encrypted data (d0)k0 to the authenticator. Third, the authenticator decrypts the received data with k0 and authenticates the decrypted data.

SCOPE

I. In this paper, 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.

II. We propose a 3DES 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

III .Experimental results demonstrate the efficiency of our proposed scheme. There are still many challenge problems in symmetric SE schemes.

IV .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.

V .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

VI ALOGORITHM

The need for data encryption arose by the growing concern of the safety and security of the data. In this article, various data encryption algorithms under comparison are DES (Data Encryption Standard), 3DES (Triple DES), AES (Advanced Encryption Standard) and Blowfish (Best performance).

Various data encryption algorithms under comparison are DES (Data Encryption Standard), 3DES (Triple DES), AES (Advanced Encryption Standard) and Blowfish.

DES

Data Encryption Standard was the first encryption technique based on the Lucifer algorithm proposed by IBM. Being the first encryption standard it had many defects and several exploits were discovered which made it very unsafe.

3DES

Triple DES is an enhancement to DES, which provided triple security in comparison to DES. The algorithm is same, only the encryption technique is applied thrice in order to increase the level of security.

AES

Advanced Encryption Standard was proposed by National Institute of Standard and technology (NIST) in order to replace DES. The only known attack to AES is the brute force attack that allows an attacker to test combination of characters in order

To break the security. However, Brute Force is not an easy job even for a super computer if the number of combination is arbitrarily high.

Searchable Encryption Algorithm

An algorithm that consists of the polynomial time randomized algorithms.

They are:

Key Gen(s) - s is a security parameter taken and used to generate a key pair either public or private.

PEKS (Apub, w) - Apub is a public key and w is a word which are used to produce a searchable encryption. Trapdoor (Apriv, w) - Apriv is a private key and w is a word which are used to produce a trapdoor Tw.

Cipher text Security

It is a technique that is used to provide security for the encrypted data. A cipher text attacker could easily break semantic security by reordering the keywords and submitting the resulting cipher text for decryption. A standard technique is used to break this and this technique is called the cipher text security.

Private Key Searchable Encryption

A model called private key searchable encryption is used to search on a private key encrypted data. The user himself encrypts data, so as to organize in an arbitrary way.

Public Key Searchable Encryption

Public key searchable encryption is a model that allows user to encrypt data and send it to the server. The owner provides decryption key may be different.

VII MRSE FRAMEWORK

For easy presentation, operations on the data documents are not shown in the framework since the data owner could easily employ the traditional symmetric key cryptography to encrypt and then outsource data. With focus on the index and query, the MRSE system consists of four algorithms as follows

1. Setup (ℓ) Taking a security parameter ℓ as input, the data owner outputs a symmetric key as SK.

2. Build Index (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.

3. Trapdoor (fW) with t keywords of interest in fW as input, this algorithm generates a corresponding trapdoor Tf W.

4. Query (Tf W, k, I)When the cloud server receives a query request as (Tf W, k), it performs the ranked search on the index I with the help of trapdoor TfW, and finally returns FfW, the ranked id list of top-k documents sorted by their similarity with f.W

VIII CONTRIBUTION

1 We suggest two MRSE schemes based on the Similarity calculation of ―coordinate matching‖ at the time of assembling different privacy needs in two different threat models.

2 We examine some further improvements of our ranked search method to maintain more search semantics and dynamic data process.

3 We determine the problem of multi keyword ranked search over encrypted cloud data, and set up a set of privacy needs for such a secure cloud data operation system.

4. Detailed analysis investigating privacy and Efficiency assurance of the proposed schemes is known, and testing

On the real-world data set further show the proposed schemes certainly bring in low overhead on calculation and communication. In this paper we propose hybrid approach **AES, DES, 3DES** to maintain secure and search semantics.

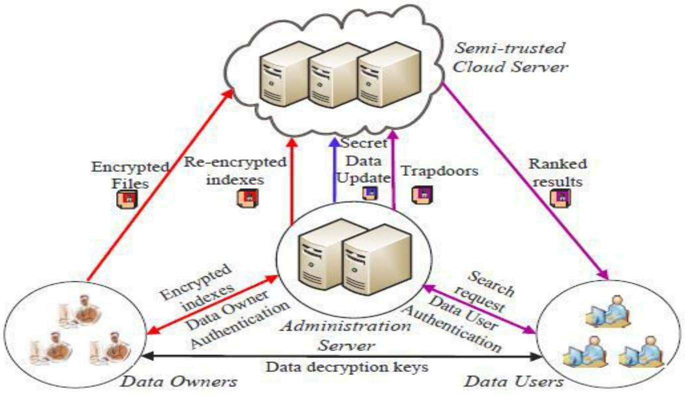


Figure 2: Architecture of multi-keyword synonym query over encrypted cloud data.

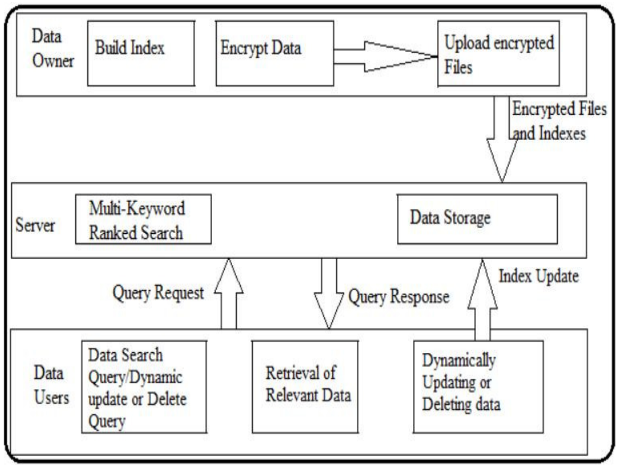

Data Owner
Build Index
Encrypt Data
Upload encrypted Files
Encrypted Files and Indexes
Server
Multi-Keyword Ranked Search
Data Storage
Query Request
Query Response
Index Update
Data Users
Data Search Query/Dynamic update or Delete Query
Retrieval of Relevant Data
Dynamically Updating or Deleting data

Fig3 Internal Flow Diagram

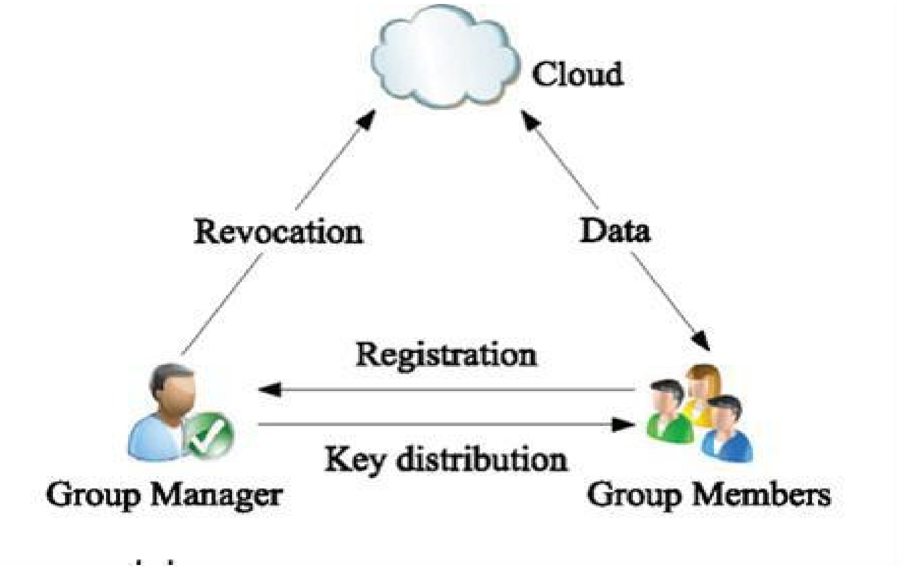


Fig 4 Administration & Key distribution

IX PROPOSED IMPLEMENTATION

A. Data User

This includes the user registration login details.

B. Data Owner

This helps the owner to register those details and also include login details

C. File Upload

This helps the owner to upload his file with encryption using AES,3DES algorithm. This ensures the files to be protected from unauthorized user.

D. Rank Search

This ensures the user to search the files that are searched frequently using rank search.

E. File Download

This allows the user to download the file using his secret key to decrypt the downloaded data.

F. View Uploaded and Downloaded File

This allows the Owner to view the uploaded files and downloaded files

PERFORMANCE AND EVALUATION

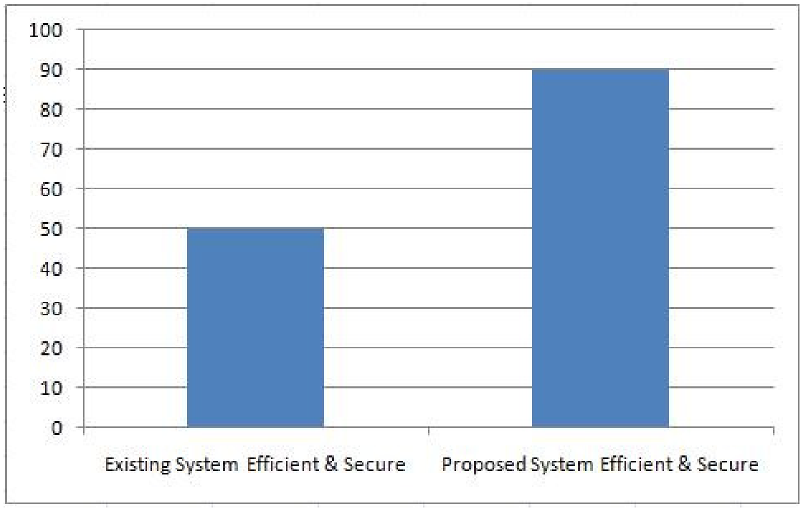
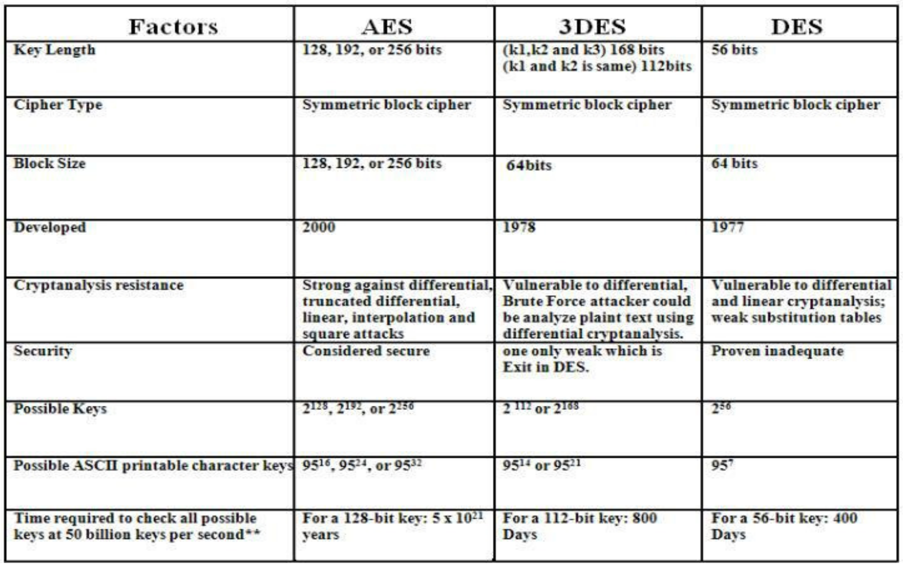


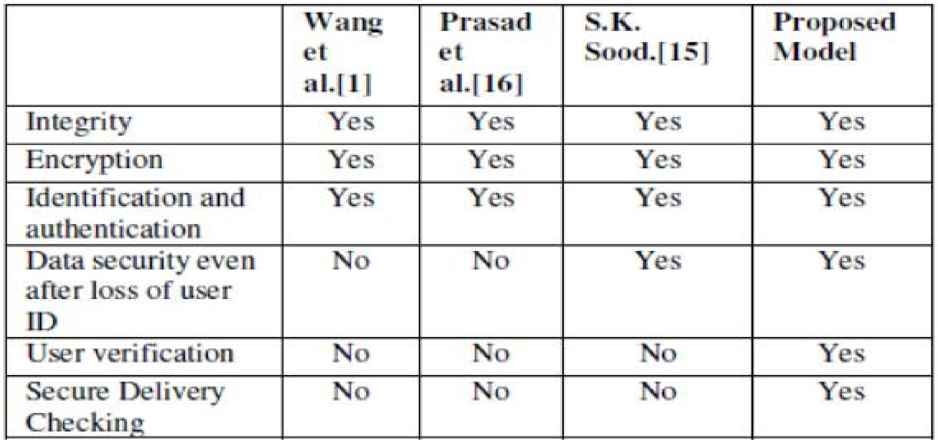
Fig 5 Shows Existing Vs Proposed

X COMPARISION AES DES 3DES



Factors	AES	3DES	DES
Key Length	128, 192, or 256 bits	(k1,k2 and k3) 168 bits (k1 and k2 is same) 112bits	56 bits
Cipher Type	Symmetric block cipher	Symmetric block cipher	Symmetric block cipher
Block Size	128, 192, or 256 bits	64bits	64 bits
Developed	2000	1978	1977
Cryptanalysis resistance	Strong against differential, truncated differential, linear, interpolation and square attacks	Vulnerable to differential, Brute Force attacker could be analyze plaint text using differential cryptanalysis.	Vulnerable to differential and linear cryptanalysis; weak substitution tables
Security	Considered secure	one only weak which is Exit in DES.	Proven inadequate
Possible Keys	2^{128}, 2^{192}, or 2^{256}	2^{112} or 2^{168}	2^{56}
Possible ASCII printable character keys	95^{16}, 95^{24}, or 95^{32}	95^{14} or 95^{21}	95^{7}
Time required to check all possible keys at 50 billion keys per second**	For a 128-bit key: 5 x 10^{21} years	For a 112-bit key: 800 Days	For a 56-bit key: 400 Days

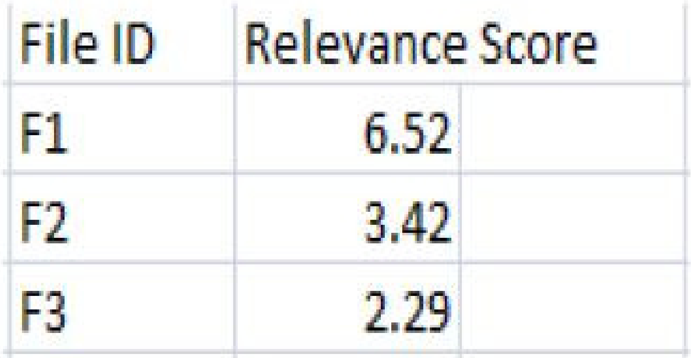
Fig 6 Compare AES DES 3DES



	Wang et al.[1]	Prasad et al.[16]	S.K. Sood.[15]	Proposed Model
Integrity	Yes	Yes	Yes	Yes
Encryption	Yes	Yes	Yes	Yes
Identification and authentication	Yes	Yes	Yes	Yes
Data security even after loss of user ID	No	No	Yes	Yes
User verification	No	No	No	Yes
Secure Delivery Checking	No	No	No	Yes

Fig 7 EVALUATION OF THE MODEL

RESULTSET



File ID	Relevance Score
F1	6.52
F2	3.42
F3	2.29

Figure 8 : Table for relevance score.

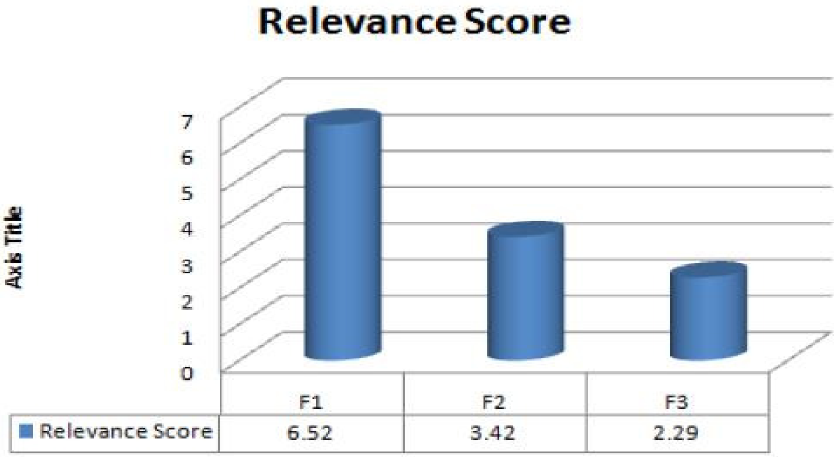


Figure 9: Graphical representation of relevance scores.

RESULT ANALYSIS

The experimental results can be explained with the set of some snapshots.

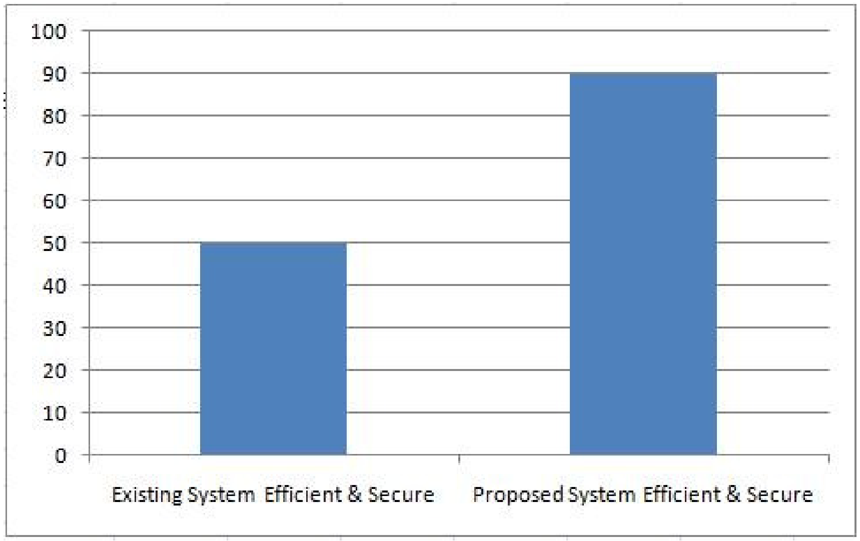
1. First of all a data owner uploads some data file, which is encrypted using AES algorithm.

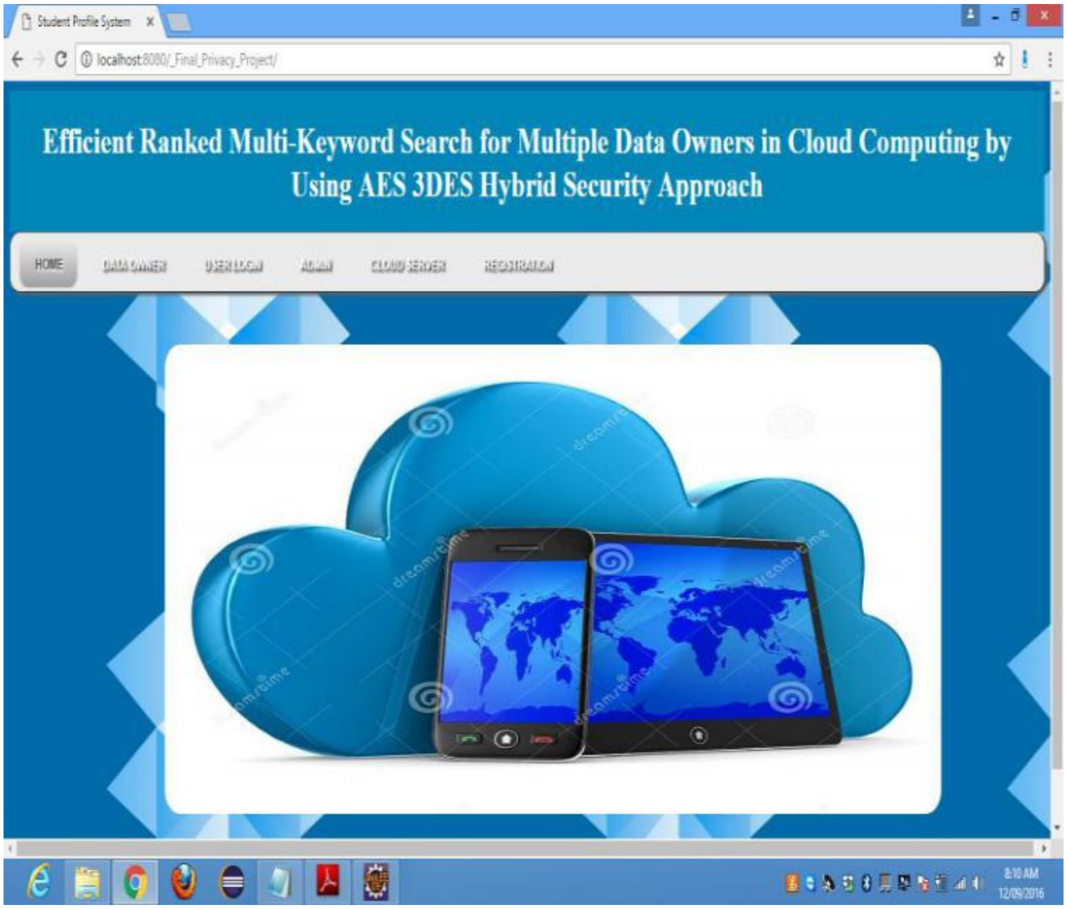
2. The uploaded on cloud server can be searched in a secure way using SSE (secure symmetric encryption) algorithm. The search results are displayed in ranked form using OPSE (order preserving symmetric encryption

3. Algorithm generate the relevance score of files based on term frequency (TF) and inverse domain frequency (IDF), using the equation TF×IDF. Where TF can be defined as the number of times given keyword or term exists in a given file. IDF can be calculated by dividing the number of files in whole collection by number of files containing that keyword (Figures 8 and 9).

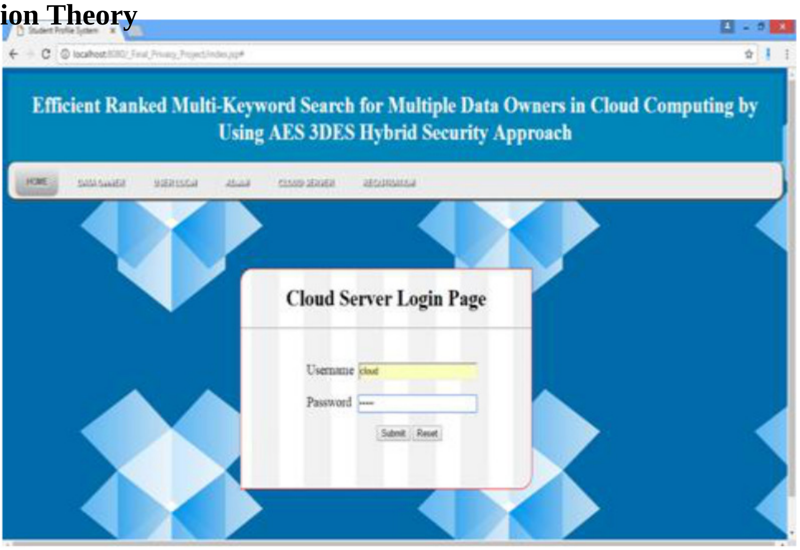
4. Hence based on relevance score files can be ranked for more symmetry.

OUTPUT SCREEN





Efficient Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing by Using AES 3DES Hybrid Security Approach

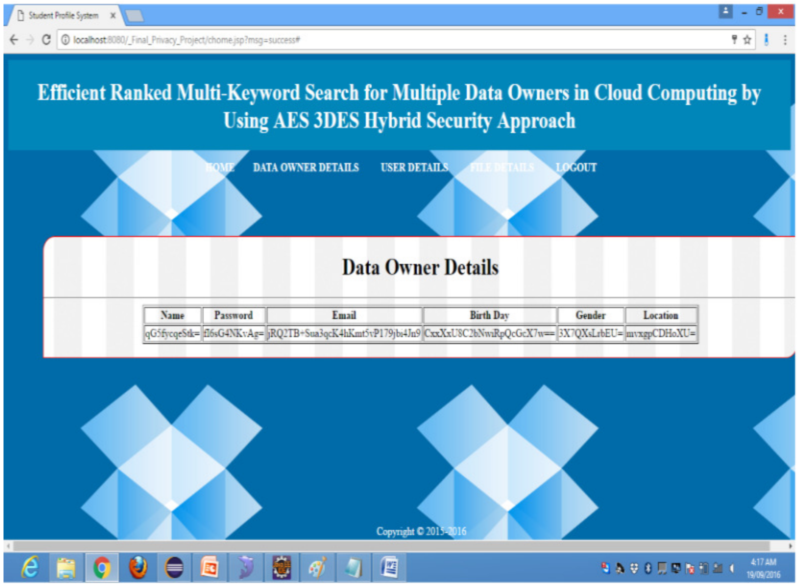


Efficient Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing by Using AES 3DES Hybrid Security Approach

Cloud Server Login Page

Username: cloud

Password: *****



Efficient Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing by Using AES 3DES Hybrid Security Approach

Data Owner Details

Name	Password	Email	Birth Day	Gender	Location
qG5fycqeStk=	fl6sG4NKvAg=	jRQ2TB+Sua3qcK4hKmt5vP179jbi4Jn9	CxxXxU8C2bNwiRpQcGcX7w==	3X7QXsLrbEU=	mvxgpCDHoXU=

Copyright © 2015-2016

CONCLUSION

This project establishes the importance of faster retrieval of data from Cloud. This has become more important in the current scenario where usage of cloud infrastructure is on a rise. As more users move towards cloud for storing their information, it is essential for cloud providers to use newer algorithms which give speedy retrieval without compromising security of user data.

Various methods are used to make index and do searching in the encrypted text etc. But in a multiple data owner model which is considered for analyzing about the data sharing in cloud computing an efficient ranked multi-keyword search scheme over encrypted data is done. The index trees for each data files are merged into one. The searching is done using a DFS algorithm. That is a secure search protocol that allows different data owners to encrypt the files and indexes with different keys are used. Then, a tree-based index structure for each data owner allows the cloud server to merge encrypted indexes without knowing any information. This tree based search scheme is more efficient in keyword mapping that other existing methods.

FUTURE CHALLENGES

Future work will consider the problem of secure fuzzy keyword search in a multi-owner paradigm. On the other hand, plan to implement extend our work on images, video in commercial clouds. This study addresses these issues by proposing Visual Cryptography Scheme (VCS) technique for securing the files. Then the files are encrypted using Advanced Encryption Standard (AES). Then the encrypted files are securely sent to the cloud. This research work can be extended to implement image storage and retrieval.

V REFERENCES

[1] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, ―Fuzzy Keyword Search Over Encrypted Data in Cloud Computing,‖ Proc. IEEE INFOCOM, Mar. 2010.

[2] N. Cao, S. Yu, Z. Yang, W. Lou, and Y. Hou, ―LT Codes-Based Secure and Reliable Cloud Storage Service,‖ Proc. IEEE INFOCOM, pp. 693-701, 2012.

[3] S. Yu, C. Wang, K. Ren, and W. Lou, ―Achieving Se-cure, Scalable, and Fine-Grained Data Access Control in Cloud Computing,‖ Proc.IEEE INFOCOM, 2010.

[4] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. KatzA. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, ―A view of cloud computing,‖ Communication of the ACM, Vol. 53, No. 4, pp. 50–58, 2010.

[5] M. Abdalla, M. Bellare, D. Catalano, E. Kiltz, T. Kohno, T. Lange,J. Malone-Lee, G. Neven, P. Paillier, and H. Shi, ―Searchable encryption revisited: Consistency properties, relation to anonymous ibe, and extensions,‖ J. Cryptol., vol. 21, no. 3, pp. 350–391, 2008.

[6] M.Chuah and W. Hu,"Privacy-aware bedtree based solution for fuzzy multi-keyword search over encrypted data",Distributed Computing Systems Workshops, 2011 31st International Conference,IEEE, (2011).

[7] S. Deshpande, "Fuzzy keyword search over encrypted data in cloud computing",World Journal of Science and Technology, vol. 2, no. 10, (2013).

[8] P. Golle, J. Staddon, and B. Waters, ―Secure conjunctive keyword search over encrypted data,‖ in Proc. of ACNS, 2004, pp. 31–45.
L. Ballard, S. Kamara, and F. Monrose, ―Achieving efficient conjunctive keyword searches over encrypted data,‖ in Proc. of ICICS, 2005.

[9] D. Boneh and B. Waters, ―Conjunctive, subset, and range queries on encrypted data,‖ in Proc. of TCC, 2007, pp. 535–554.

[10]R. Brinkman, ―Searching in encrypted data,‖ in University of Twente, PhD thesis, 2007.

[11] Y. Hwang and P. Lee, ―Public key encryption with conjunctive keyword search and its extension to a multi-user system,‖ in Pairing, 2007.

[13]J. Katz, A. Sahai, and B. Waters, ―Predicate encryption supporting disjunctions, polynomial equations, and inner products,‖ in Proc. Of EUROCRYPT, 2008.

[14] C. Rong, S. T. Nguyen, M. G. Jaatun, ―Beyond Lightning: A survey on security challenges in cloud computing,‖ Computers and Electrical Engineering, vol. 39, no. 1, pp. 47-54, Jan. 2013.

[15] S. K. Sood, ―A combined approach to ensure data security in cloud computing,‖ Journal of Network and Computer Applications, vol. 35, no. 6, pp. 1831-1838, 2012.

[16] P. Prasad, B. Ojha, R.R. Shahi, R. Lal, ―3-dimentional security in cloud computing,‖ in 3rdInternational Conference on Computer Research and Development (ICCRD, 2011), pp. 198-208, 2011.

[17] Cong Wang, Qian Wang, and Kui Ren, Wenjing Lou, Ensuring Data Storage Security in Cloud Computing, IWQoS. IEEE International Workshop on Quality of Service, 2009. pp. 1-9.