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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

Multi Keyword Ranked Search over Anonymous Encrypted Cloud Data

Shiba S. Kale, Shivaji R. Lahane

Student, Dept. of Computer, GES's R. H. Sapat College of Engineering, Management Studies & Research, Nashik, Savitribai Phule Pune University, India

Assistant Professor, Dept. of Computer, GES's R. H. Sapat College of Engineering, Management Studies & Research, Nashik, Savitribai Phule Pune University, India

ABSTRACT: The advancement in cloud computing has motivated the data owners to outsource their data management systems from local sites to commercial public cloud for great flexibility and economic savings. For real privacy, user identity should remain hidden from CSP (Cloud service provider) and to protect privacy of data, data which is sensitive is to be encrypted before outsourcing. Thus, enabling an encrypted cloud data search service is of great importance. By considering the large number of data users, documents in the cloud, it is important for the search service to allow multi-keyword query and provide result similarity ranking to meet the effective need of data retrieval, search, and not often differentiate the search results. In this system first time, we define and solve the challenging problem of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE), and establish a set of strict privacy requirements for such a secure cloud data utilization system to be implemented in real. We first propose a basic idea for the Multi-keyword Ranked Search over Encrypted cloud data (MRSE) based on secure inner product computation and efficient similarity measure of coordinate matching, i.e., as many matches as possible, in order to capture the relevance of data documents to the search query, then give two significantly improved MRSE schemes to achieve various stringent privacy requirements in two different threat models. Assignment of anonymous ID to the user to provide more security to the data on cloud server is done. To improve search experience of the data search service, further extension of these two schemes to support more search semantics is done.

KEYWORDS: Cloud Computing; Searchable Encryption; Privacy-Preserving; Keyword Search; Ranked Search Anonymization; MRSE.

I. INTRODUCTION

Cloud computing has been a long dreamed vision of computing as a utility, where the cloud customers can remotely store their data into cloud so they can enjoy the on-demand high-quality applications and services from a shared pool of configurable computing resources. Great flexibility and economic savings in cloud computing are motivating the individuals and enterprises to outsource their local complex data management system into the cloud. In order to protect data privacy and combat unsolicited accesses in the cloud and beyond, sensitive data, for instance, e-mails, personal health records, photo albums, tax documents, financial transactions, etc. may have to be encrypted by data owners before outsourcing them to the commercial public cloud; this, however, obsoletes the traditional data utilization service based on plaintext keyword search. The trivial solution of downloading all the data and decrypting locally is clearly not practical, due to the huge amount of bandwidth cost in cloud scale systems. Moreover, besides from eliminating the local storage management, storing data into cloud serves no purpose unless and until they can be easily searched and utilized by users.

By considering the potentially large number of on-demand data users and huge amount of outsourced data documents in the cloud, this problem is particularly challenging as it is very much difficult to meet the requirements of performance, system usability, and scalability. Ranked search can elegantly eliminate the unnecessary network traffic by sending back only the most relevant data, which is very highly desirable in the "pay-as-you-use" cloud paradigm.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

For protection of privacy, such ranking operation, however, should not leak any keyword related information. On the other hand, to improve the search result accuracy as well as to enhance the user searching experience, it is also essential for such ranking system to support multiple keywords search, as single keyword search often yields far too coarse results. As a common practice indicated by today's web search engines (like Google search), data users may tend to provide a set of keywords instead of only one as the indicator of their search interest to retrieve the most relevant data. Along with privacy of data and efficient searching schemes, real privacy is gained only if the user's identity remains hidden from the cloud service provider i.e. CSP as well as the third party user on the cloud server [1].

The remainder of this paper is as follows: In Section II, we describe the literature survey. Section III describes the proposed system, algorithms. Section IV describes mathematical model. Section V describes the implementation strategy. Section VI describes the experimental setup. Section VII describes the results. Finally we conclude the paper in Section VIII.

II. RELATED WORK

In [2], main focus is to find the solution of multi-keyword ranked search over encrypted cloud data (MRSE) while preserving strict system-wise privacy in the cloud computing paradigm. A variety of multi- keyword semantics are available, an efficient similarity measure of "coordinate matching" i.e. as many matches as possible, to capture the data documents' relevancy to the search query is used. Specifically "inner product similarity", i.e., the number of query keywords appearing in a document, to quantitatively evaluate such similarity measure of that document to the search query is used in MRSE algorithm. The important drawback of this paper is that the user's identity (ID) was not kept hidden. Hence, whoever puts the data on Cloud Service Provider (CSP) was known to the CSP. This might be risky in some situations. Hence, this limitation is overcome in proposed system. For instance, a user may want to store old e-mail messages encrypted on a server managed by Yahoo or another large vendor, and later retrieves certain messages while travelling with a mobile device.

In [3], the schemes are efficient as no public-key cryptosystem is involved. Indeed, this approach is independent of the encryption method chosen for remote files. The main theme taken from the paper is of storing data remotely on other server and retrieving that data from anywhere through mobile, laptop etc. In [4], an overall summary of the benefits of a cryptographic storage service, for instance, reducing the legal exposure of both customers and cloud providers, also achieving regulatory compliance is provided. Besides all this, cloud services that could be built on top of a cryptographic storage service such as secure back-ups, archival, health record systems, secure data exchange and e-discovery is stated.

The [5], the paper has defined and found solution to the problem of effective yet secured ranked keyword search over encrypted cloud data. For the first time, paper has defined and solved the most challenging problem of privacy-preserving multi-keyword ranked search over encrypted cloud data (MRSE), and established a set of strict privacy requirements for such a secure cloud data utilization system to become a reality. The proposed ranking method proves to be efficient to return highly relevant documents which correspond to the submitted search terms. Idea of proposed ranking method is used in the proposed system in order to enhance the security of data and the documents on Cloud Service Provider (CSP).

The [6] paper tells the significance of protecting individual's privacy in cloud computing, it also provides some privacy preserving technologies used in cloud computing services. Paper tells that it is quite important to take privacy into account while designing cloud services, if all this involves the collection, processing or sharing of personal data. From this paper, main idea taken is of privacy preserving of data. This paper importantly describes privacy of data but doesn't allow indexed search as well as the user's identity is not kept hidden. Thus, these two main limitations are overcome in the proposed system.

In [7], the paper mainly focuses on existing and new algorithms for assigning anonymous IDs and their examination with respect to trade-offs between communication and computational requirements. These algorithms are built on top

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

of a secure sum data mining operation using Newton's identities and Sturm's theorem. The main focus in this paper is of assigning anonymous ID to the user on the cloud.

In [9], paper's main idea is to formalize and provide solution to the problem of effective fuzzy keyword search over encrypted cloud data while maintaining keyword privacy. This basic theme is taken but it is mianly for multi-keyword ranked search (MRSE scheme) in the proposed system. In [8],design of secure cloud storage service which addresses the reliability issue with near-optimal overall performance is mainly proposed.

Achieving fine-grainedness, scalability, and data confidentiality of access control simultaneously is a problem for which no solution is found yet. The paper [10] addresses this challenging open issue on one hand, defining and enforcing access policies based on data attributes, and, on the other hand, allowing the data owner to refer most of the computation tasks involved in fine-grained data access control to untrusted cloud servers without disclosing the underlying data contents.

In [11], authors have proposed a privacy-preserving public auditing system for data storage security in Cloud Computing scheme. It utilizes homomorphic linear authenticator and random masking to guarantee that the TPA would not learn any knowledge about the data content stored on the cloud server during the efficient auditing process, which eliminates the burden of cloud user from the difficult, tedious and possibly expensive auditing task, it also alleviates user's fear of his/her outsourced data leakage. In [12], authors have defined and solved the problem of privacy-preserving query over encrypted graph-structured data in cloud computing (PPGQ), and established a set of strict privacy requirements for such a secure cloud data utilization system to become a reality. The basic idea taken from this paper is of privacy preserving over encrypted cloud data. The limitation in this is that the query is not indexed to provide fast searching. This drawback is overcome in the proposed system.

III. PROPOSED ALGORITHM

A. *Methodology:*

1. Methodology for Registration:

Step 1 : User open URL of registration.
Step 2 : User provide necessary details and register on KDC (Key Distribution Cloud).
Step 3 : KDC provide token to User (Via email).

2. Methodology to Upload File On Cloud:
Step 1 : User Login.
Step 2 : User select files to upload.
Step 3 : User first get encryption key from KDC.
Step 4 : User get index of files.
Step 5 : User encrypt index.
Step 6 : SHA-1 algorithm to encrypt index.
Step 7 : Encrypt document using encryption key given by KDC.

3. Methodology to Share document with other User:
Step 1 : Add user access privileges to data structure present on KDC.
Methodology to Search (Search by other user).
Step 1 : User login and verified by KDC.
Step 2 : User get key from KDC.
Step 3 : Generate Trapdoor for search (Trapdoor includes query words and number of ranked document).
Step 4 : Get result set.

B. *Prerequisites:*

1. OCR for image extraction(Optical Character Recognition):

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

Optical Character Recognition, or OCR, is a technology that enables you to convert different types of documents, such as scanned paper documents, PDF files or images captured by a digital camera into editable and searchable data. It is a common method of digitizing printed texts so that it can be electronically edited, searched, stored more compactly, displayed on-line, and used in machine processes such as machine translation, text-to-speech key data and text mining. OCR is a field of research in pattern recognition, artificial intelligence and computer vision.

C. *Algorithms:*

1. Data Structure:
   a. Dictionary D:

id	word
1	java
2	C
3	php

   b. Index Generation:
      1. Create index Φ
      2. Identify TF-IDF of document keyword DK
      3. Define Threshold T
      4. for each DKi in DK

ifDKi ≥T &&DKi € D then
add T in index

2. System Algorithms:
   a. Sha1 Algorithm:
      Input: Message
      Output: 160 bit hash value
      Processing:
      Message m is divided into 512bit size parts as ML
      Initialize:
      h0 = 0x67452301
      h1 = 0xEFCDAB89
      h2 = 0x98BADCFE
      h3 = 0x10325476
      h4 = 0xC3D2E1F0
      For each mi in ML break
      mi into sixteen 32-bit big-endian words array W
      Extend W list upto eighty 32-bit words:

fori from 16 to 79
w[i] = (w[i-3] xor w[i-8] xor w[i-14] xor w[i-16]) left shift rottate 1
End For
a = h0
b = h1
c = h2
d = h3
e = h4
fori from 0 to 79
if 0 ≤ i ≤ 19 then
f = (b and c) or ((not b) and d)
k = 0x5A827999
else if 20 ≤ i ≤ 39
f = b xor c xor d
k = 0x6ED9EBA1
else if 40 ≤ i ≤ 59
f = (b and c) or (b and d) or (c and d)

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

```
        k = 0x8F1BBCDC
else if 60 ≤ i ≤ 79
        f = b xor c xor d
        k = 0xCA62C1D6

temp = (a leftrotate 5) + f + e + k + w[i]
      e = d
      d = c
      c = b leftrotate 30
      b = a
      a = temp
End For
   Add this chunk's hash to result so far:
   h0 = h0 + a
   h1 = h1 + b
   h2 = h2 + c
   h3 = h3 + d
   h4 = h4 + e
Produce the final hash value (big-endian) as a 160 bit number as
hh = (h0 leftshift 128) or (h1 leftshift 96) or (h2 leftshift 64) or (h3 leftshift 32) or h4
End for
```

b. AES Algorithm:

Encryption phase:

```
Constants: int nob = 4;
int R = 14;
```

Inputs: input plaintext :array in of 4*Nob bytes

Output: output ciphertext : array out of 4*Nob bytes

```
expanded key: array w of 4*Nob*(R+1) bytes
        Internal work array:
        state, 2-dim array of 4*Nob bytes, 4 rows and Nob cols
```

Start:

```
state : 2D byte array
out : 2D byte array (having actual cipher)

state = in; // actual component-wise copy

AddRoundKey(state, w, 0, Nob – 1)
 For each round to R
        Replace      in state matrix with SubByteS()
        In each row cyclically shifts the      by certain offset
        MixColumns(state)
        AddRoundKey(state, w, round*Nob, (round+1)*Nob – 1)
End For
        Replace      in state matrix with SubByteS(    )
        In each row cyclically shifts the      by certain offset
        AddRoundKey(state, w, R*Nob, (R+1)*Nob – 1)
out = state
End
```

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

Decryption phase:
Constants: int Nob = 4;
int R = 14;
Inputs:
in : array of 4*Nob bytes
out : array of 4*Nob bytes
w : array w of 4*Nb*(Nr+1) bytes
Output:
out : array of 4*Nob bytes
Steps:
voidInvCipher(byte[] in, byte[] out, byte[] w)
Start:
State : 2D byte array
state = in
AddRoundKey(state, w, R*Nob, (R+1)*Nob – 1)

For each round from (R-1) to 1
InvShiftRows(state)
InvSubBytes(state)
AddRoundKey(state, w, round*Nob, (round+1)*Nob-1)
MixColumns(state)
End For

InvShiftRows(state)
InvSubBytes(state)
AddRoundKey(state, w, 0, Nob – 1)
out = state
End

IV. MATHEMATICAL MODEL

S = { U , C , KDS } WHERE
U = { I , O , F , SoU} Here
I = { I1 , I2 , I3 , I4 , I5 }
O = {O1 , O2, O3, O4, O5}
F = {RegReq ,LoginReq , IndexGen , F6 , F7 , F8 , F9}
SoU = U1 , U2 HERE
U1=>Owner ; U2=> Third Party User ;
I1=> Registration Details ;
I2 => Login Details
I3 => File
I4 => Custom Keywords
I5 => Search Keyword
O1 => Verification Token
O2 => Encryption Key Generation
O3 => Cipher Text
O4 => Trapdoor Key
O5 => Tags Of a file
F6 => Encryption
F7 => Trapdoor Generation Request
F8 => Decryption
F9 => File Transfer

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

C = { Ip , Op , Fp}WHERE
Ip = I01 , I02 , I03 , I04 Here
I01 => Encrypted Input File
I02 => Index File
I03 => Keyword
I04 => User Token
Op = { Op1 , Op2 } Here
Op1 => Search result ; Op2 => User file to download
Fp = F1 , F2 , F3 , F4 Here
F1 => Save Index
F2 => Save File
F3 => Search Key
F4 => Top k result
KDS = { KI , KO , KF } WHERE
KI = UI HERE UI => User Information
KO = Key , Token
KF = KeyGen , Store Key , LoginReg , AccessPri , TrapdoorGen

V. IMPLEMENTATION STRATEGY

We have created a system in java and apache environment. We have created a desktop application for admin and user that communicate with CSP and KDC. An apache tomcat is installed on KDC and CSP. A servelet and jsp files provides web services to user and admin application. Data communication is carried out using http protocol. Desktop application uses HTTP client facility to communicate with KDS and cloud. JSON format is used to transfer the data. The proposed system is implemented using Divide and Conquer strategy.

VI. EXPERIMENTAL SETUP

A. *Dataset:* For dummy records we have used Enron Email Data Set, user selects random files from this dataset to upload data on clouds [13]. Keyword set: We have selected top 4000 keywords for index generation in the given dataset.
B. *System Setup:* We have tested the system on single node machine. 4 different applications i.e. user application, Admin application, KDS, Cloud is hosted on same windows -7 system. database used is mysql. We have hosted our system on cloud Windows 2008 R2 server.

VII. RESULTS

A. We individually tested each module and /or algorithm for correctness.
1. Document Preprocessing: TF and IDF of the document: In this, we calculate the Term Frequency and Inverse Document Frequency of the document keywords after using Lemmatization and Stop words removal algorithms. Finally we get the Frequency Count of probable keyword in the documents.
2. Generation of Index: Index of file/document is generated while uploading the file/document and manipulate the index generation time of file/document. Encryption of index is done after the index of particular file/document is generated.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

TABLE I

Document size (KB)	No. of index terms	Index generation time (ms)	Encryption time(ms)	Upload time (ms)	Download time (ms)
1	6	27817	1	59	2
2	20	1135	253	88	154
3	42	16238	146	103	2
4	73	24036	1	772	4
5	87	20556	3	98	3

The table shows the readings of document size vs time which is calculated from various operations (index generation, encryption,upload, download,etc) done.

a. Document size vs number of index terms

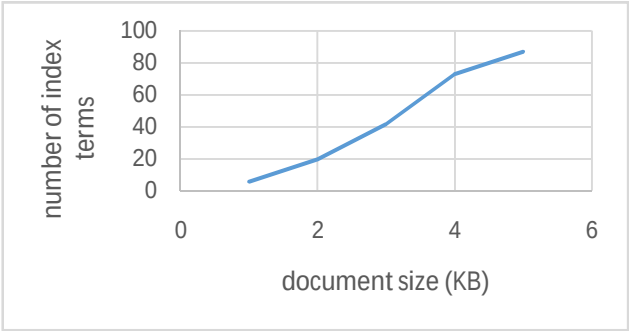


Fig.2: document size vs number of index terms

Fig.2 shows that as the document size increases, the number of index terms which are extracted also increase.

b. Document size vs index generation time

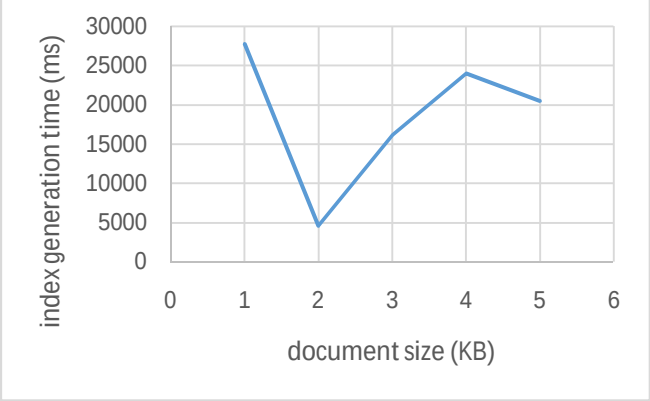


Fig.3: document size vs index generation time

Fig.3 shows the graph of document size vs time required for index generation. This time varies depending upon the background proceeses running along with the project.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

c. Document size vs encryption time

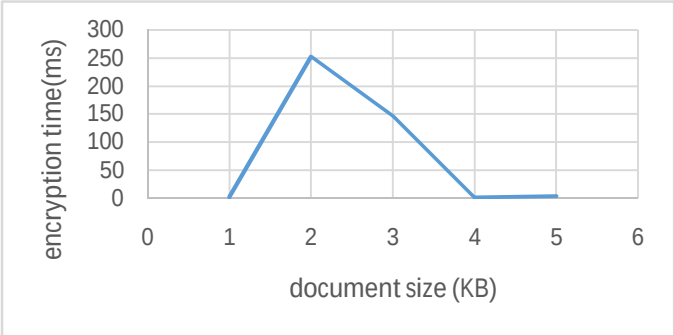


Fig.4: document size vs encryption time

Fig.4 shows the graph of document size vs time required for encryption of file/document. This time varies depending upon the background proceeses running along with the project.

d. Document size vs upload and download time

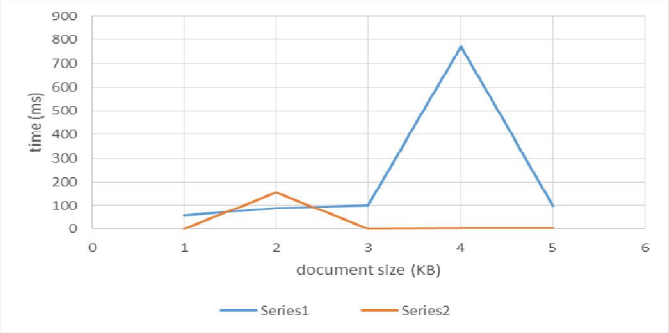
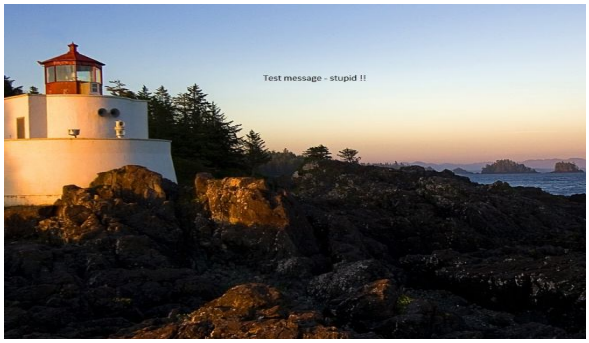


Fig.5: document size vs upload and download time

In the Fig.5,
Series 1 – Upload time
Series 2 – Download time

3. Image Text Extraction: Here, OCR (Optical Character Recognition) technique is used to extract the text from image in image containing some text. The pixels of the text are mapped with inbuilt patterns and a pattern is generated. This pattern is machine readable and text extraction is done.

Example:



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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

The above image contains text "Text message stupid!". This message is extracted using the OCR technique and the time required for extraction is displayed on the console as follows:
Result: Test message 7 stupid 11
Final Result : Test message 7 stupid 11
BUILD SUCCESSFUL (total time: 2 seconds)

4. Data Encryption and Data Decryption: Data encryption and decryption of file/document is done and corresponding time is calculated during upload and download of file/document.
5. File/Document Search: In this, keywords in query are searched from the documents/files uploaded and shared. Relevant results are returned to the user based on the number of matching keywords with file/document. Ranking is also provided in percentage.

a. Keywords in query vs time

TABLE II

Keywords	Time(ms)
2	103
6	160
10	203
14	432
18	445

Above table shows number of keywords in query for search and the time required to search those keywords in the file/document.

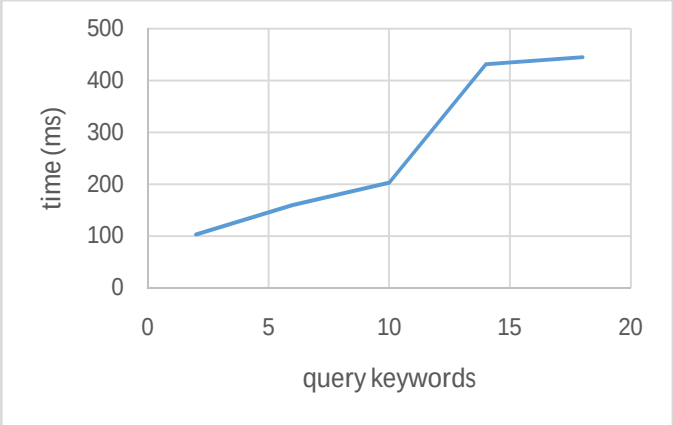


Fig.6: query keywords vs time

In the Fig.6, we observe that the time and query keywords vary proportionally i.e. time required to serch the quey keyords goes on increasing as the number of query keywords to search increaes.

b. Documents in data set vs time

TABLE III

Documents in dataset	Time(ms)
4	188
8	245
12	325
16	351

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

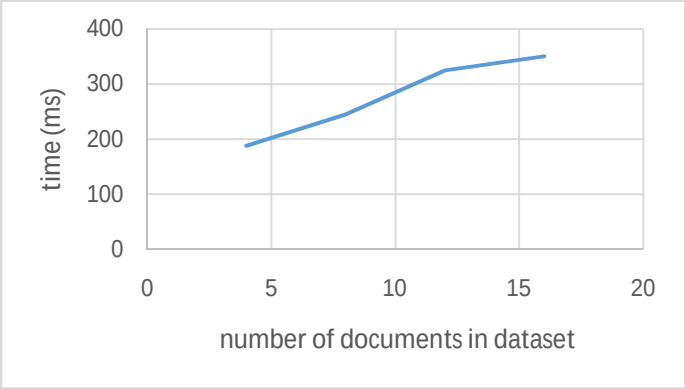


Fig.7: number of documents in dataset vs time

In the Fig.7, we observe that as the number of documents in a dataset increases, the time also increases accordingly.

B. Comparison of index of basic MRSE scheme with our proposed scheme index is done. In both cases, time for index creation is same as we calculate TF id IDF for document words. And match it with dictionary of our basic index term array. Existing system creates sub index in the form of 1 – 0. 1 represents existence of word while 0 represents absence of keyword.The proposed structure represents index with only matched keywords and hence size of index is reduced, therefore storage overhead is reduced. Total keyword dictionary size: 4000.

TABLE IV

FREQUENCY COUNT

Document	***Keyword Matched***	***Index size (MRSE Scheme)***	***Index size (P-MRSE Scheme)***
D1	10	4000	10
D2	15	4000	15
D3	13	4000	13
D4	12	4000	12

Above table shows compaison of the index size of various documents for the existing and proposed system.

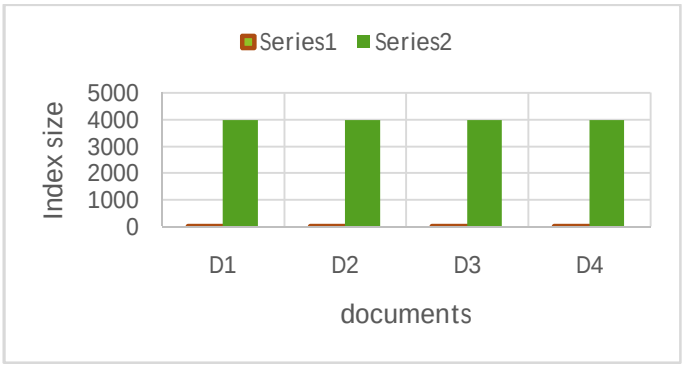


Fig.8: documents vs index size of existing and proposed system

In the Fig.8, comparison of the index size of documents of existing system and proposed system is done

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

Series 1 – Proposed system,Series 2 – Existing system

C. Security Analysis

We analyze our scheme according to the predefined privacy requirements

1) Index Confidentiality: the cloud server cannot infer the original data vector and the query vector as it is encrypted using SHA-1 algorithm.

2) Data Confidentiality: Data is secured as we are uploading encrypted data to the cloud server using AES algorithm.

3) User Identity Confidentiality: User identity is hidden from the cloud server.

VIII. CONCLUSION

The earlier work focused mainly on providing privacy to data on cloud in which using multi-keyword ranked search was provided over encrypted cloud data with the use of efficient similarity measure of co-ordinate matching. Also the previous work proposed a basic idea of MRSE using secure inner product computation. It became essential to provide more real privacy which this paper mainly focusses on. In this system, stringent privacy is provided by assigning the cloud user a unique ID. This user ID is kept hidden from the cloud service provider as well as the third party user in order to protect the user's data on cloud from the CSP and the third party user. Thus, by hiding the user's identity, the confidentiality of user's data is maintained.

REFERENCES

1. Ning Cao, Cong Wang, Ming Li, Kui Ren and Wenjing Lou ," Privacy Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data",2014.
2. Ankatha Samuyelu Raja Vasanthi ," Secured Multi-keyword Ranked Search over Encrypted Cloud Data",2012.
3. Y.-C. Chang and M. Mitzenmacher, "Privacy Preserving Keyword Searches on Remote Encrypted Data," Proc. Third Int'l Conf. Applied Cryptography and Network Security, 2005.
4. S. Kamara and K. Lauter, "Cryptographic Cloud Storage," Proc. 14th Int'l Conf. Financial Cryptograpy and Data Security, Jan. 2010.
5. Y. Prasanna, Ramesh . "Efficient and Secure Multi-Keyword Search on Encrypted Cloud Data", 2012.
6. Jain Wang, Yan Zhao , Shuo Jaing, and Jaijin Le, "Providing Privacy Preserving in Cloud Computing",2010.
7. Larry A. Dunning, Ray Kresman ," Privacy Preserving Data Sharing With Anonymous ID Assignment",2013.
8. 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.
9. 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.
10. S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving Secure, Scalable, and Fine-Grained Data Access Control in Cloud Computing," Proc. IEEE INFOCOM, 2010.
11. C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing," Proc. IEEE INFOCOM, 2010.
12. N. Cao, Z. Yang, C. Wang, K. Ren, and W. Lou, "Privacy preserving Query over Encrypted Graph-Structured Data in Cloud Computing," Proc. Distributed Computing Systems (ICDCS), pp. 393-402, June, 2011.
13. W.W. Cohen, "Enron Email Data Set," http://www.cs.cmu.edu/~enron/,2013.

BIOGRAPHY

Shiba S. Kaleis a Student in the Computer Department, GES's R. H. Sapat College of Engineering, Management Studies & Research, Maharashtra, Nashik. She received her Bachelor of Engineering Degree from KKWIEER ,Maharashtra,Nashik. Her research interests are Cloud Computing, Security.

Shivaji R. Lahaneis a Assistant Professor in the Computer Department, GES's R. H. Sapat College of Engineering, Management Studies & Research, Maharashtra, Nashik.Her research interests are Cloud Computing, Security, Networking, etc.