IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

Cloud Based Secured Multi Keyword Ranked Search for Multi-Owned

Ms. Vanita Gadekar [1], Dr.Baisa Gunjal[1]

Computer Engineering Department, AVCOE, Sangamner, SPPU[1]

Abstract: For the purpose of economic saving as well as to achieve flexibility users outsourced their data on public cloud. Before storing data on cloud user wish to apply some security constraints to protect their data form malicious attacks. Therefore they used to refer encryption technique to encode their data and then they will outsource it. Existing systems which refers encryption technique are inefficient as they are having only single keyword search over large document. To the best solution for this inefficiency we proposed multikeyword search strategy for large number of documents. Our multikeyword search technique is efficiently word against the searching of multiple keywords simultaneously. It may be referred as lightweight process of searching. In our system we used two protocols to generate a secret key dynamically as well as for vali user authentication. Our system effectively works for large amount of databases and multiple keyword searching. In our system we provide higher privacy requirements to implement the data security for cloud stored data. Our main contribution in this project is to give backup and restore facility at the user end and mainly we focused on multikeyword searching with ranking approach.

Keywords: ranked keyword search, multiple owners, privacy preserving, dynamic secret key.

I. INTRODUCTION

Cloud computing provides multiple facility such as, to protect sensitive data, like emails, governments files, employee personal records in various sectors etc. In cloud computing, privacy is provided for accessing data, to outsourced the data and also to maintain flexibility of data[1]. By using the concept of virtualization and firewall CSP gives guarantee of data privacy to the data owner in cloud system. CSP has full control on cloud hardware, software as well as on data owner's [3]. Previously used technique that is data encryption suffered from many challenges such as inefficiency, costing when there is large amount of data is present [1]. Due to single keyword search strategy it is most time consuming task and hence it result into inefficiency [7]. J. Li, Q. Wang introduced fuzzy multi-keyword search technique over an encrypted data [9]. Whereas Boolean keyword search provides the solution for ranked keyword search for large dataset[12].

We proposed multikeyword search technique for ranked keywords where large number of encrypted data is available. From security perspective our system provides unique ID for users. This ID is unique or hidden from CSP and third party user. Therefore, better privacy and data confidentiality is achieved with our proposed system. Our system works effectively where there is need to protect data sensitivity. Some medical application in which patient details as well as some important diseases information needs to preserve privately. With implementation of such type of system we aim to provide a backup and restore facility at users end and ranked multikeyword search over large database in cloud. Our system gives the best solution against single keyword search over large amount of cloud data. And also it proves its efficiency as well as security. Furthermore, we are going to discussed about our system related work done, problem statement, proposed system and its architecture, algorithm and mathematical model in following section. Lastly we are giving our contribution in this system by conclusion our system.

II. RELATED WORK

There are two new protocols has been design for different data owners to use different keys for encryption of data as well as keywords. Also to protect data from attackers who monitors the secret keys and covering to be legal data. In this system two protocol namely, dynamic secret key generation protocol and a data user authentication protocol is implemented. To protect legal data from the attackers a secure search protocol is utilized [1]. M. Armbrust, [2], discussed about simplification of system by comparing cloud computing and conventional computing. In this system he examines functional and non-functional opportunities of cloud storage. Therefore, their system decreases confusion by depicting that how to rectify the comparisons between of cloud and conventional computing. Q. Wang, K. Ren,[3], suggested data security in cloud. Author introduced privacy-preserving in public auditing system. Proposed system manages multiple audit sessions of multiple users for their outsourcing of data. In auditing process TPA would not have any idea about the data content stored that are on the cloud server.

Cryptographic schemes are described by D.Song, D.Wagner, and A.Perrig [4], for the issues of searching of encrypted data over cloud. They give proofs of security for the resulting crypto systems. Their system is secure for remote searching on encrypted data over an untrusted server.R. Curtmola, [5], provides review on some previously existed notions of security. They suggest new stronger security definitions known as Searchable

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

symmetric encryption (SSE). This scheme allows outsourcing the data to other party. With implementation of such system they proved higher security levels. In this system "trapdoors" are generated by the owner of the private key. A protocol that allows conjunctive keyword queries over encrypted data is introduced by P. Golle, J. Staddon, and B. Waters [6]. It solves the problem of secure Boolean search. They considered documents that are stored on untrusted server. C.

Wang, N. Cao, J. Li, K. Ren[7], represents the schemes that support searching of boolean keyword. In this system authors solve the problem of managing an efficient ranked keyword search.Because of this one can effectively manages encrypted data that is stored remotely.It provides usability by supplying matching files. This scheme assumes "honest-but-curious" server in the model where cloud server behaves as "honest" and "curious" analyze the message. N. Cao, C. Wang, M. Li, K. Ren[8], described the solution for the problem of multi-keyword ranked search over encrypted cloud data (MRSE). Authors mainly they also concern on preserving strict system-wise privacy in the cloud storage. MRSE schemes to achieve various stringent privacy requirements in two different threat models. MRSE scheme evaluates "coordinate matching" to achieve privacy requirements into various threat models. Many systems simplify and give solution for effective fuzzy keyword search over encrypted cloud data, J. Li, Q. Wang, C. Wang, et al.[9], and manages keyword privacy. A wildcard-based technique is used for searching of fuzzy keywords. Tg is system consisting cloud data system i.e. data owner, data user and cloud server for fuzzy keyword searching approach.

A probabilistic public key system i.e. PEKS is introduced by P. Xu, H. Jin, Q. Wu, and W. Wang [10]. This scheme is convenient to search ciphertexts for multiple users. With this system they aim to proved SS-CKA and IK-NCK-KGA securities.

W. Lou, and Y. T. Hou [11], described a novel multi keyword fuzzy search technique. This scheme is used to exploit locality-sensitive hashing technique. Rather than expanding the index file fuzzy matching is done through algorithmic design. This approach of leveraging LSH functions in the Bloom filter provides efficient solution to the Secure fuzzy search of multiple keywords.

W. Lou [12], build a new crypto primitive OPSE. It achieves efficient one-to-many order-preserving function for mapping. This primitive expands ranked search over encrypted data. OPSE achives an efficient one-to-many order-preserving mapping function, to allow the effective RSSE to be designed.

Yu, W. Lou, Y. T. Hou[13], examined the issues of outsourced dataset can be supplied from different owners to different users. It is enabling scalable file-level search authorization. G. Ateniese[14], implements ABE to developed a primitive known as attribute-based keyword search (ABKS). In this system keywords are encoded according to an access control policy. In this system data user generates the credentials tokens for encrypted data. A scheme known as VABKS allow data owner to control the search.

T. Jung, X. Y. Li, Z. Wan, and M. Wan [15], locates user privacy problem. In this attribute-based privilege control scheme Anony Control is proposed.According our review on previous system we analyse a problem statement given as follow:

III. PROBLEM DEFINITION

As per our objectives and from security point of view we have to propose a system that is having high level data privacy on cloud. For data privacy system must be armed with encryption techniques. Along with the encryption of data on cloud we have to propose a technique that helps to search such encrypted data. As we know cloud is following strategy like pay-as-you-use. Hence sometimes search technique is so costly as it consider huge amount of data on cloud so at time of result fetching system must consider the large amount of data users and documents in the cloud and must consider crucial nature of searching so system should contain multi-keyword query service. For effective bandwidth utilization proposed system must be armed with result similarity ranking service to meet effective retrieval. Also System should implement proper techniques for encryption. System should implement proper techniques for searching such as coordinate matching, inner product similarities. As a part of contribution system should contain module that hides user identity.

IV.PROPOSED SYSTEM

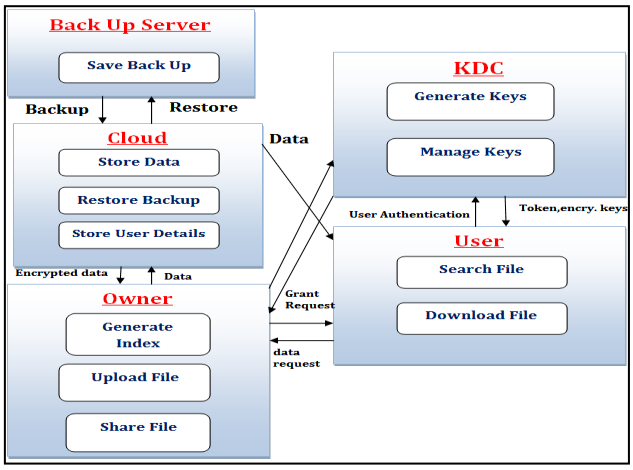

Fig.1 System Architecture

Above we represent system architecture

1. Register:

Step 1: User Registration

Step 2: During registration user provides some necessary information.

Step 3: Token is given to end user by administrative server provide.

2. Upload File:

Step 1: Login.

Step 2: Upload file.

Step 3: Administrative Server provides an encryption key to the user.

DOI 10.17148/IJARCCE.2016.55127

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

Step 4: Index of files for user.
Step 5: Encrypt user index.
Step 6: To encrypt index SHA-1 algorithm is used.
Step 7: Encrypt document AES algorithm
3. Share document with other User:
Step 1: Access privileges to data structure present on Administrative Server.
4. Search:
Step 1: User login that verified by Administrative Server.
Step 2: Get key from Administrative Server.
Step 3: Generate Trapdoor for search.
Step 4: Achieved result set.
5. Backup:
Step 1: User login that verified by Administrative Server.
Step 2: User get key from Administrative Server.
Step 3: Show own files.
Step 4: Select file and use backup facility where last modified copy is saved.
6. Restore:
Step 1: User login and verified by Administrative Server.
Step 2: User get key from Administrative Server.
Step 3: Show own deleted files.
Step 4: Select file for restore.

IV. ALGORITHMS

1. AES Algorithm
Advance Encryption Standard is a symmetric key block cipher. AES is a non-Feistel cipher that encrypt and decrypts 128-bits block of the data. The size of the key can be 128,129 or 256-bits.It depends on the number of rounds.
The number of the rounds:
10 rounds for 128-bits,
12 rounds for 192-bits,
14 rounds for 256-bits.
Input: secret key k, Message M
Output: Encrypted Message EM
Process:
I. Key Expansions Round:
Keys are derived from the cipher key using Rijndael's key record. Sseparate 128-bit round key block for each round plus one more needed for AES algorithm.
II. Initial Round:
AddRoundKey each byte of the state is combined with a block of the round key using bitwise XOR.

III. Rounds:
-SubBytesa non-linear substitution step where each byte is replaced with another according to a lookup table.
-ShiftRowsa transposition step where the last three rows of the state are shifted cyclically a certain number of steps.
-MixColumnsa mixing operation which operates on the columns of the state, combining the four bytes in each column.
-Add Round Key
4. Final Round (no Mix Columns)
-S-Sub Bytes.
-Shift Rows.
-Add Round Key.
2. SHA-1:
Input: Message
Output: 160 bit hash value
Process:
Message m is divided into 512-bit size parts as ML
Initialize:
h0, h1, h2,h3,h4 For each mi in ML.
Break mi into sixteen 32-bit big-endian words array W
Extend W list up to eighty 32-bit words:
for i 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
for i from 0 to 79
if $0 \leq i \leq 19$ then
f = (b and c) or ((not b) and d)
k = 0x5A827999
else if $20 \leq i \leq 39$
f = b xor c xor d
k = 0x6ED9EBA1
else if $40 \leq i \leq 59$
f = (b and c) or (b and d) or (c and d)
k = 0x8F1BBCDC
else if $60 \leq i \leq 79$
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

V. 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, F10 , F11}
SoU = {U1, U2} here,
U1 = Owner
U2= Third Party User

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

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
F10 = Backup
F11 = Restore
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

VI. EXPERIMENTAL RESULTS

To develop a system we have used java platform. We have created 3 different entities Cloud, KDC and client. Cloud and KDC server provides web services to the client system. Client can be owner or data user. Client system is build using swing components HTTP client facility is used for communication. Mysql database is used to store data. Apache tomcat server is used to at server end to serve multiple client requests at a time.

Eclipse and netbean-8.1 IDE is used to develop a system. Dataset: To test system performance we have used Enron dataset [16]. This is email dataset containing multiple files with different sizes. A dictionary of 4000 keyword is present to tag the file data and to generate an index. Following figure represents the index generation:

TABLE I INDEX GENERATION

File Size	Index gen time
1	5
2	11
3	16
4	19

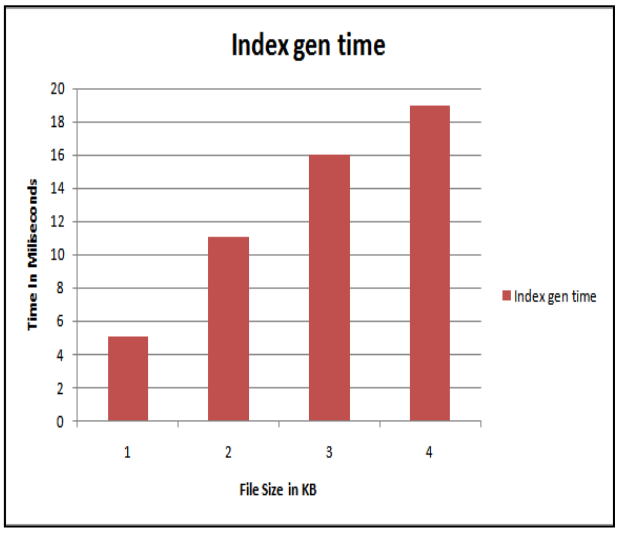

Fig.2 Time Required for index Generation

TABLE III FILE ENCRYPTION AND DECRYPTION

File Size	Encryption	Decryption
1	0.1	0.2
2	0.28	0.38
3	0.78	0.84
4	0.855	0.98

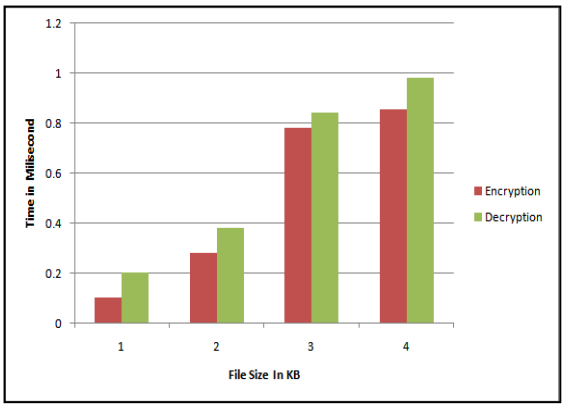

Fig.3 Time required for file Encryption & Decryption

TABLE IIIII FILE UPLOAD, DOWNLOAD AND RESTORE

File Size	Upload	Download	Restore Time
1	0.15	2.058	1.048
2	0.23	2.345	1.34
3	0.295	2.467	1.4
4	0.32	2.743	1.72

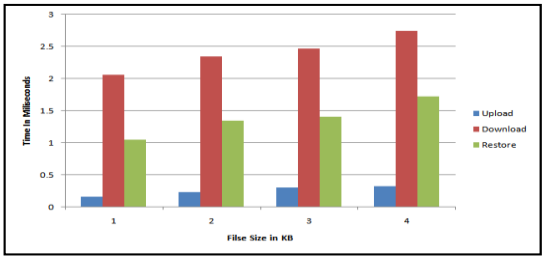

Fig.4 Time required for file upload, download and restore

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

TABLE IVV INDEX TERM

Index term	Index Size (Existing)	Index Size (Proposed)
10	4000	10
20	4000	20
30	4000	30
40	4000	40

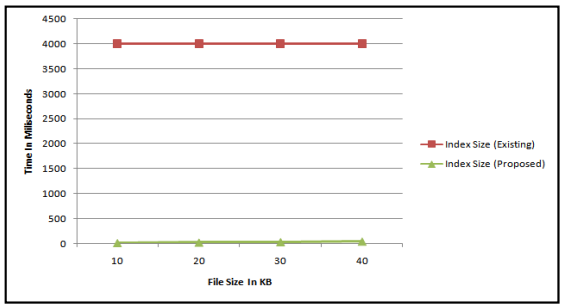

Fig.5:Time required for index generation

VII. CONCLUSIONS

Our multi-keyword ranked search system provides encrypted cloud data that utilizes efficient similarity measure of co-ordinate matching. Previous systems mainly focused on providing privacy to the data on cloud. We provide more real privacy system. In our system, stringent privacy is provided by allocating a unique ID to cloud user. Our system hides this user ID from the cloud service provider as well as the third party user, to protect the user's data on cloud from the CSP and the third party user. To manitain data confidentiality, to hide users identity may work efficiently. Also multiple owners for data are concept newly introduced in our system. While uploading the data index terms are generated that helped for indexed search which is efficient one. In contribution we provide an efficient backup and restore facility to the end user.

ACKNOWLEDGMENT

I am Vanita Gadekar heartly thankful of my project guide Dr. Baisa Gunjal for their guidance and support for during project development. Also I am thankful to all AVCOE principal and staff for their well support.

REFERENCES

[1] Wei Zhang, Student Member, IEEE, Yaping Lin, and Siwang Zhou,"Privacy Preserving Ranked Multi-Keyword Search for Multiple Data Owners in Cloud Computing "

[2] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing,"

[3] C. Wang, S. S. Chow, Q. Wang, K. Ren, and W. Lou, "Privacypreserving public auditing for secure cloud storage," Computers, IEEE Transactions on, vol. 62, no. 2, pp. 362–375, 2013.

[4] D.Song, D.Wagner, and A.Perrig, "Practical techniques for searches on encrypted data," in Proc. IEEE International Symposium on Security and Privacy (S&P'00), Nagoya, Japan, Jan. 2000, pp. 44–55.

[5] R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in Proc. ACM CCS'06, VA, USA, Oct. 2006, pp. 79–88.

[6] P. Golle, J. Staddon, and B. Waters, "Secure conjunctive keyword search over encrypted data," in Proc. Applied Cryptography and Network Security (ACNS'04), Yellow Mountain, China, Jun. 2004, pp. 31–45.

[7] C. Wang, N. Cao, J. Li, K. Ren, and W. Lou, "Secure ranked keyword search over encrypted cloud data," in Proc. IEEE Distributed Computing Systems (ICDCS'10), Genoa, Italy, Jun. 2010, pp. 253–262.

[8] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, "Privacypreserving multi-keyword ranked search over encrypted cloud data," in Proc. IEEE INFOCOM'11, Shanghai, China, Apr. 2011, pp. 829–837

[9] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in Proc. IEEE INFOCOM'10, San Diego, CA, Mar. 2010, pp. 1–5.

[10] P. Xu, H. Jin, Q. Wu, and W. Wang, "Public-key encryption with fuzzy keyword search: A provably secure scheme under keyword guessing attack," Computers, IEEE Transactions on, vol. 62, no. 11, pp. 2266–2277, 2013.

[11] B. Wang, S. Yu, W. Lou, and Y. T. Hou, "Privacy-preserving multi-keyword fuzzy search over encrypted data in the cloud," in IEEE INFOCOM, Toronto, Canada, May 2014, pp. 2112–2120.

[12] C. Wang, N. Cao, K. Ren, and W. Lou, "Enabling secure and efficient ranked keyword search over outsourced cloud data," Parallel and Distributed Systems, IEEE Transactions on, vol. 23, no. 8, pp. 1467–1479, 2012.

[13] W. Sun, S. Yu, W. Lou, Y. T. Hou, and H. Li, "Protecting your right: Attribute-based keyword search with fine-grained owner-enforced search authorization in the cloud," in Proc.IEEE INFOCOM'14, Toronto, Canada, May 2014, pp. 226–234.

[14] Q. Zheng, S. Xu, and G. Ateniese, "Vabks: Verifiable attributebased keyword search over outsourced encrypted data," in Proc. IEEE INFOCOM'14, Toronto, Canada, May 2014, pp. 522– 530.

[15] T. Jung, X. Y. Li, Z. Wan, and M. Wan, "Privacy preserving cloud data access with multi-authorities," in Proc. IEEE INFOCOM'13, Turin, Italy, Apr. 2013, pp. 2625–2633.

BIOGRAPHIES

Ms..Vanita Arjun Gadekar Perusing masters in Computer Engineering at AVCOE, Sangamner. She has received Bachelors Degree in Computer Engineering from AVCOE, Sangamner. She is a member of Association of Computer Machinery (ACM).

Dr. Baisa Gunjal is Associate Professor and HOD in department of Information Technology at AVCOE, Sangamner. She is a member of Association of Computer Machinery (ACM). Dr. Baisa Gunjal received the PHD Degree in Computer Engg, From Pune University.