JCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-5, May 2018 E-ISSN: 2347-2693

A Novel Approach for Searchable Keyword in Cloud Computing Using Efficient Algorithm

P.Sai Sirisha[1*], K.Praveen Kumar[2], M.Harsha Vardhan[3] , J.Karthik[4]

[1*]Computer Science & Engineering, St.Ann's College of Engineering & Technology, JNTUK, Chirala, India
[2] Computer Science & Engineering, St.Ann's College of Engineering & Technology, JNTUK, Chirala, India
[3] Computer Science & Engineering, St.Ann's College of Engineering & Technology, JNTUK, Chirala, India
[4] Computer Science & Engineering, St.Ann's College of Engineering & Technology, JNTUK, Chirala, India

**Corresponding Author: harshamaddu@gmail.com, Tel.: +91-9676811364*

Available online at: www.ijcseonline.org

Accepted: 18/May/2018, Published: 31/May/2018

***Abstract*—**In ancient, dynamic secret writing systems it provides solely information confidentiality, however, it's incompatible with dynamic multi-keyword over encrypted information by causing secret keys to the user through the mail. In traditional we'll be going to be operative on plain information however in our project we tend to encrypt the information. The info by generating secret keys and activity search on encrypted data by user request. During this paper, we tend to propose a dynamic multi-keyword hierarchical search theme over encrypted cloud information, that at the same time supports dynamic update operations like deletion and insertion of documents. Intensive experiments area unit conducted to demonstrate the potency of the projected theme.

***Keywords*—** MultiKeyword,Ranked,Efficient,Searchable.

I. INTRODUCTION

Cloud computing has been thought-about as a brand new model of enterprise IT infrastructure, which may organize immense resource of computing, storage, and applications, and alter users to get pleasure from the present, convenient and on-demand network access to a shared pool of configurable computing resources with nice efficiency and tokenism economic overhead. Attracted by these appealing options, each people and enterprises area unit actuated to source their information to the cloud, rather than getting computer code and hardware to manage the info themselves. Despite the assorted benefits of cloud services, outsourcing sensitive data (such as e-mails, personal health records, company financial information, government documents, etc.) to remote servers brings privacy considerations[1]. The cloud service providers (CSPs) that keep the data for users could access users' sensitive information while not authorization. A general approach to safeguard the info confidentiality is to write the info before outsourcing. However, this can cause a large price in terms of knowledge usability. For instance, the prevailing techniques on keyword-based data retrieval, that area unit wide used on the plaintext information, cannot be directly applied to the encrypted information[2]. So far, rife works are projected underneath completely different threat models to realize varied search practicality, like single keyword search, similarity search, multi-keyword Boolean search, hierarchical search, multi-keyword hierarchical search, etc[3]. Inverse document frequency (IDF) model area unit combined with the index construction and query generation to supply multi-keyword hierarchical search, so as to get high search efficiency[4]. We tend to construct a tree-based index structure and propose a "Greedy Depth-first Search" algorithmic program supported this index tree attributable to the special structure of our tree-based index, the projected search theme will flexibly reach sublinear search time and handle the deletion and insertion of documents.

Rest of the paper is organized as follows, Section I contains the introduction of Cloud computing, Section II contains the related work of different keyword approaches, Section III contains some measures of methodology and algorithms, Section IV contain the architecture and essential steps of the searching efficiently, section V explains the methodology with flow chart, Section VI describes results and discussion, Section VII contains the recommendation of and Section VIII concludes research work with future directions.

II. RELATED WORK

Searchable encryption schemes enable the clients to store the encrypted data to the cloud and execute keyword search over ciphertext domain [5]. Due to different cryptography primitives, searchable encryption schemes can be constructed using public key based cryptography or symmetric key based cryptography Song et al. proposed the first searchable encryption scheme, and the search time of their scheme is linear to the size of the data collection [6]. Goh proposed formal security definitions for SSE and designed a scheme based on Bloom filter. The search time of Goh's scheme is O(n), where n is the cardinality of the document collection. Curtmolaet al. proposed two schemes (SSE-1andSSE-2) which achieve the optimal search time. Their SSE-1 scheme is secure against chosen-keyword attacks (CKA1) and SSE-2 is secureagainst adaptive chosen-keyword attacks (CKA2) [7]. These early works are single keyword Boolean search schemes, which are very simple in terms of functionality Afterward, abundant works have been proposed under different threat models to achieve various search functionality, such as single keyword search, similarity search multi-keyword Boolean search ranked search

Multi-keyword Boolean search allows the users to input multiple query keywords to request suitable documents[8]. Among these works, conjunctive keyword search schemes only return the documents that contain all of the query keywords[9]. Disjunctive keyword search schemes. return all of the documents that contain a subset of the query keywords. Predicate search schemes are proposed to support both conjunctive and disjunctive search[10]. All these multi-keyword search schemes retrieve search results based on the existence of keywords, which cannot provide acceptable result ranking functionality[11].

III.METHODOLOGY

The methodology we tend to area unit exploitation to implement this project area unit are:-

1)Greedy Depth First Search:

A greedy Depth First Search algorithmic program is the associate degree recursive paradigm that follows the matter determination heuristic of constructing the domestically optimum alternative at every stage with the hope of finding a world optimum. The Greedy DFS algorithmic program may be an algorithmic program that uses the thought of backtracking. It involves complete searches of all the nodes by going ahead, if attainable, else by backtracking. Here, the word go back implies that after you area unit moving forward and there aren't any a lot of nodes on the present path, you progress backward on a similar path to seek out nodes to traverse. All the nodes are visited on the present path until all the unvisited nodes are traversed once that following path is selected.

Algorithm:

```
if the node u isn't a leaf node then
    if RScore(Du ,Q) >k th score then
        GDFS(u.hchild);
        GDFS(u.lchild);
    else
        return
    end if
else
    if RScore(Du,Q) >kth score then
        Delete the element with the smallest relevance score from RList.
        Insert a brand new component (RScore(Du,Q),u.FID) and type all the weather of RList;
    end if
    return
end ifWhere u-Leaf node
```

Where,

u-Leaf node
RScore (Du,Q)—The perform to calculate the connectedness score for question vector Q and index vector Du keep in node u, that is defined in Formula .
kthscore—The smallest connectedness score in current RList, that is initialized as zero.
hchild—The child node of a tree node with higher relevance score.
lchild—The child node of a tree node with lower relevance score.

2) KNN algorithm:

K-nearest neighbor search identifies the highest k nearestneighbors to the question. This system is usually utilized inprophetical analytics to estimate or classify a degreesupported the accord of its neighbors. K-nearest neighbour graphs area unit graphs during which each purpose isconnected to its k nearest neighbors.

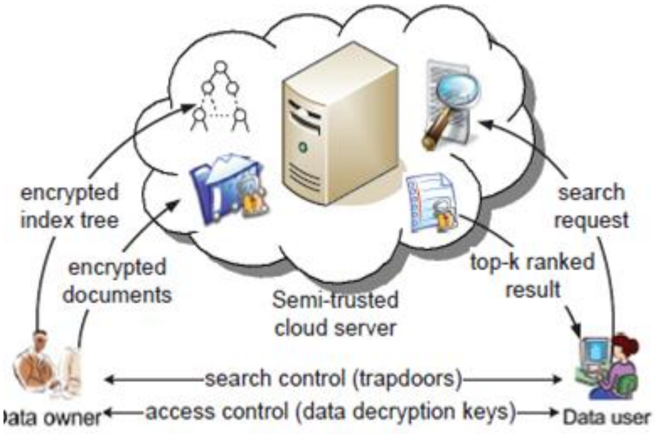


Figure1: Architecture of Multi-keyword Search

The basic plan of our new algorithm:

The worth of dmax is weakened keeping step with the continued actual analysis of the thing similarity distance forthe candidates. At the top of the step by step refinement,dmax reaches the optimum question vary disfunction andprevents the tactic from manufacturing a lot of candidatesthan necessary therefore fulfilling the r-optimality criterion.

Nearest Neighbor Search (q, k) // optimum algorithmic

Program:-

```
1. Initialize ranking = index.increm-ranking (F(q), df)
2. Initialize result = new sorted-]list (key, object)
3. Initialize dmax = w
4.Whereas o = ranking.getnext and d,(o, q) I d,,, do
5. If do, s> s dmax then result.insert (d,(o, q) , o)
6. If result.length two k then dmax = result[k].key
7. take away all entries from result wherever key >dmax
8. End while
Report all entries from result wherever key 1 dmax.
```

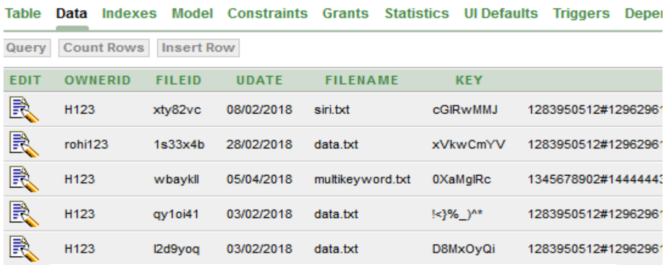
3) RSA Algorithmic Program

This algorithmic program is employed to write n rewrite filecontents. It's associate degree uneven algorithmic program.The RSA algorithmic program involves 3 steps: keygeneration, secret writing and coding. Key generation RSAinvolves a public key and a non-public key. The generalpublic key are often notable to everybody and is employedfor encrypting messages. Messages encrypted with thegeneral public key will solely be decrypted exploitation thenon-public key. The keys for the RSA algorithmic programarea unit generated the subsequent way:

1. Select 2 distinct prime numbers a and b.
2. Reason n = ab.n is employed because the modulus for each the general public and personal keys
3. Reason φ(n) = (a –1)(b –1), wherever φ is Euler's totient perform.
4. Select associate degree number e specified one < e < φ(n) and greatest common factor of (e, φ(n))1; i.e., e and φ(n) area unit co-prime. e is free because the public key exponent having a brief bit-length

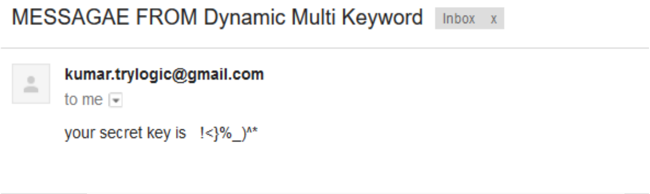
IV. RESULTS AND DISCUSSION

Results for the dynamic multi-keyword search are shown below.



EDIT	OWNERID	FILEID	UDATE	FILENAME	KEY	
	H123	xty82vc	08/02/2018	siri.txt	cGlRwMMJ	1283950512#1296296
	rohi123	1s33x4b	28/02/2018	data.txt	xVkwCmYV	1283950512#1296296
	H123	wbaykll	05/04/2018	multikeyword.txt	0XaMglRc	1345678902#1444444
	H123	qy1oi41	03/02/2018	data.txt	!<}%_)^*	1283950512#1296296
	H123	l2d9yoq	03/02/2018	data.txt	D8MxOyQi	1283950512#1296296

Figure 2: Encrypting the files whereas uploading the documents into the cloud

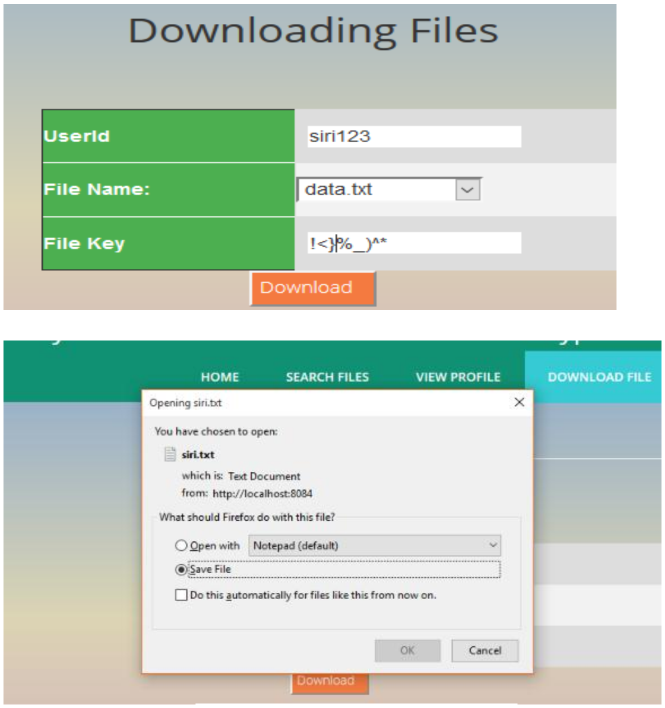


MESSAGAE FROM Dynamic Multi Keyword

kumar.trylogic@gmail.com
to me

your secret key is !<}%_)^*

Figure 3: Secret Key obtained to the Users mail Id



Downloading Files

UserId	siri123
File Name:	data.txt
File Key	!<}%_)^*

Opening siri.txt

You have chosen to open:

siri.txt

which is: Text Document

from: http://localhost:8084

What should Firefox do with this file?

○ Open with Notepad (default)

● Save File

□ Do this automatically for files like this from now on.

Figure 4: File Download Page

Data Owner Module:
This module helps the owner to register those details and additionally embrace login details. This module helps the owner to transfer his file with secret writing exploitation KNN algorithmic program. This ensures the files to be shielded from unauthorized user information owner encompasses an assortment of documents F = that he needs to source to the cloud server in encrypted kind whereas still keeping the potential to look on them for effective utilization.

Data User Module:
This module includes the user registration login details. This module is employed to assist the consumer to look the file exploitation the multiple key words thought and find the correct result list supported the user question. The user goes to pick the specified file and register the user details and find an activation code in mail email before entering the activation code. Once user will transfer the file information user area unit licensed ones to access the documents of knowledge owner.

Cloud Service Provider (CSP):
A cloud supplier may be an element of cloud computing – generally Infrastructure as a Service (IaaS), computer code as a Service (SaaS) or Platform as a Service (PaaS) – to alternative businesses or people. Cloud suppliers area unit generally remarked as cloud service suppliers or CSPs. Cloud server stores the encrypted document assortment C and therefore the encrypted searchable tree index I for information owner. It won't read the files on the cloud server and he could delete the owner and user files additionally.

V.CONCLUSION and Future Scope

Finally,in this paper we presented a secure, efficient and dynamic search scheme is proposed, which supports the accurate multi-keyword ranked search. We constructed a special keyword balanced binary tree as the index, and propose a “Greedy Depth-first Search” algorithm to obtain better efficiency than linear search.In this scheme, the data owner is not responsible for generating updating information and sending them to the cloud server. The data owner needs to store the unencrypted index tree and the information that are necessary to recalculate the IDF values.

REFERENCES

[1] K.Ren, C.Wang, Q.Wangetal., *“Security challenges for the public cloud”*, IEEE Internet Computing, vol. 16, No. 1, pp. 69–73, 2012.

[2] S. Kamara and K. Lauter, *“Cryptographic cloud storage,”* in Financial Cryptography and Data Security. Springer, 2010, pp. 136– 149.

[3] C. Gentry, *“A fully homomorphic encryption scheme”*, Ph.D. dissertation, Stanford University, 2009.

[4] O. Goldreich and R. Ostrovsky, *“Software protection and simulation on oblivious rams”*, Journal of the ACM (JACM), vol. 43, No. 3, pp. 431–473, 1996.

[5] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, *“Public key encryption with keyword search”*, in Advances in Cryptology Eurocrypt 2004. Springer, 2004, pp. 506–522.

[6] D. Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, *“Public key encryption that allows pir queries”*, in Advances in Cryptology-CRYPTO 2007. Springer, 2007, pp. 50–67.

[7] Y. H. Hwang and P. J. Lee, “Public key encryption with conjunctive keyword search and its extension to a multi-user system,” in Proc. 1st Int. Conf. Pairing-Based Cryptography, 2007, pp. 2–22

[8] D. X. Song, D. Wagner, and A. Perrig, *“Practical techniques for searches on encrypted data”*, in Proc. IEEE Symp. Secur. Privacy, 2000, pp. 44–55.

[9] E.-J. Goh, *“Secure indexes”*, IACR Cryptol. ePrint Archive, vol. 2003, pp. 216.

[10] Y.-C. Chang and M. Mitzenmacher, *“Privacy preserving keyword searches on remote encrypted data,”* in Proc. 3rd Int. Conf. Appl. Cryptography Netw. Secur., 2005, pp. 442–455.

[11] J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, *“Fuzzy keyword search over encrypted data in cloud computing,”* in IEEE Proc. INFOCOM, 2010, pp. 1–5.

Authors Profiles

Ms.P.SaiSirisha completed her B.Tech(CSE) from J.N.T.University, Kakinada in 2018.She is currently working as an intern in Gemini Consulting and Services in the platform of Android(Mobile App Development) since from February 2018. Her main research work focuses in Cloud Computing, Mobile App Development.

Mr.K.Praveen Kumar completed his B.tech(CSE) from J.N.T.University, Kakinada in 2018. He is currently working as an intern in Gemini Consulting and Services in the platform of Web Development since from November 2017. His main research work focuses on Cloud Computing, Web Development.

.

Mr.M.Harsha Vardhan completed his B.tech(CSE) from J.N.T.University, Kakinada in 2018. He is a freelancer in writing professional blogs and making videos. He has many blogs and a youtube channel named as Top 3 in Telugu. His main research work focuses on Digital marketing, Data Privacy, and data mining. He has 2 years of experience in writing blogs.

Mr.J.Karthik completed his B.Tech(CSE) from J.N.T.University,Kakinada in 2010 and M.Tech(CSE) fromJ.N.T.University,Kakinada in2012. He is currently pursuing Ph.D. and working as Assistant Professor in Department of Computer Science & Engineering,St.Ann's College of Engineering &Technology, chirala since 2013. He has published 3 research papers in reputed international and 1 inInternational conference and it's also available online. His main research work focuses on DataMining, Information Security, Cloud Computing. He has 6 years of teaching experience.