Available online at: http://www.ijmtst.com/vol4issue4.html

International Journal for Modern Trends in Science and Technology

ISSN: 2455-3778 :: Volume: 04, Issue No: 04, April 2018

# An Outsourced Cloud Data with Safe and High Secured Ranked Keyword Search

K Aruna Kumari[1] | N Srikanth[1] | G Pavani[1] | G Bharath Kumar[1]

[1]Assistant Professor, Department of CSE, Chalapathi Institute of Engineering and Technology, Guntur, Andhra Pradesh

**To Cite this Article**

K Aruna Kumari, N Srikanth, G Pavani and G Bharath Kumar, "An Outsourced Cloud Data with Safe and High Secured Ranked Keyword Search", *International Journal for Modern Trends in Science and Technology*, Vol. 04, Issue 04, April 2018, pp.-99-106.

## ABSTRACT

*Cloud computing empowers the prototypical of data service outsourcing. To defend data concealment, penetrating cloud data has to be encrypted before outsourced to the salable public cloud. Traditional searchable encryption techniques provision only Boolean search and are not yet plenty to meet the operative data utilization need that is inherently demanded by large number of users and enormous amount of data files in cloud. In this project, we delineate and elucidate the tricky of sheltered ranked keyword search over encrypted cloud data. Ranked search momentously enhances system usability by enabling search result relevance ranking in its place of sending undifferentiated results, and auxiliary warrants the file retrieval precision. we explore the statistical measure methodology, i.e. relevance score, from information retrieval to physique a sheltered searchable index, and progress a one-to-many order-preserving mapping technique to appropriately shelter those penetrating score information.*

***Keywords:*** *keyword, Index, Trap door, search, Safe, High secured, confidential, data, Files, Cloud, sky drive, One drive*



## I. INTRODUCTION

### 1.1 History

Cloud Computing is a lately emerged prototypical which is attractive prevalent among practically all initiatives. It encompasses the thought of on mandate amenities which means using the cloud resources on demand and we can scale the resources as per demand. Cloud computing indubitably affords unending reimbursements and is a cost operative model. The major apprehension in this model is Retreat in cloud. This is the intention of many enterprises of not fancying the cloud computing. This paper provides the criticism of security research in the turf of cloud security.

The commencement of cloud computing canister be traced support to the workstation days of the 1960s when the idea of "utility computing"[1] was coined by computer scientist and Turing award winner John McCarthy. Utility computing ended up becoming something of a big business for companies IBM. The concept was : that computing power could be broken down as a service for businesses much like how the power and telephone companies operated for their customers. Indeed, it was an article "The Computers of Tomorrow" for the Atlantic Monthly in May of 1964 where author Martin Greenberger pointed out the concept that "advanced arithmetical machines of the future" were now being used not only institutionally for scientific calculation and research but for business functions such as accounting and inventory. The potential for huge profit to be made in this type of invested had for terminal machines that would cost less than $300. These

ideas were indeed profound, but they never really took off as consumers were looking for more complete personal computer solutions that had, for example, some storage capacity available.

### 1.2 Open source in cloud

The cloud contribution an open-source enclose called Cloud Foundry, a Platform-as-a-Service[2] that should wallopfright in the hearts of its competitors, in particular the likes of Salesforce.com, Microsoft and Rack space. The platform will recommend developers the tools to accumulate out applications[13] on public clouds, private clouds and anyplace else, whether the underlying server runs.

Developers and enterprise IT shops will rapidly have another option for platform as a service in the form of Cloud Swing, an upcoming offering from Open Logic that builds on its interior business of given thatprocedural support for open source software. Cloud Swing customers can use the platform to bring together software stacks of in cooperation open source and commercial products for use on cloud communications services such as Amazon Elastic Compute Cloud.

### 1.3 Cloud computing and its security

The sanctuary in cloud is to integrate seamlessly with the IT defense in your own statistics centre. nevertheless, the cloud service contributor implements its own IT security procedures.

- To defend customers from external coercion.
- To guarantee that beingconsumer environments are isolated from one another.
- For every type of cloud service, the contributor delivers a good deal of the IT defence.
- IT defense software and firewalls, intrusion detection systems, virtual private networks , and secure connections that the cloud provider has in place.
- Know how the cloud providers are defending the generally computing situation.

### 1.4 Basic Model of system

As Cloud Computing becomes widespread, progressivelysusceptiblein sequence are creaturefederal into the cloud, such as e-mails, guise health proceedings, company economicsfacts and government credentials, etc. The reality that information owners and cloud server are no longer in the identical trusted domain couldlay the outsourced unencrypted statistics. Cloud computing is the extensive dreamed hallucination of computing[16] as a efficacy, where cloud patrons can remotely hoard their data into the cloud so as to have the benefit of the on-demand elevatedsuperiority applications and military from a collectivepuddle of configurable computing possessions. The reimbursement brought by this new computing representationinclude but are not restricted to assistance of the burden for luggage compartment management, widespread data admittance with self-governingecological locations and avoidance of capital disbursement on hardware, software and workforce maintenances etc. It follows that perceptivefacts has to be encrypted earlier to outsourcing for data isolation and skirmishingspontaneous accesses. However, data encryption makes valuable data exploitation a very exigentundertaking given that there could be a large amount of outsourced data files[11].

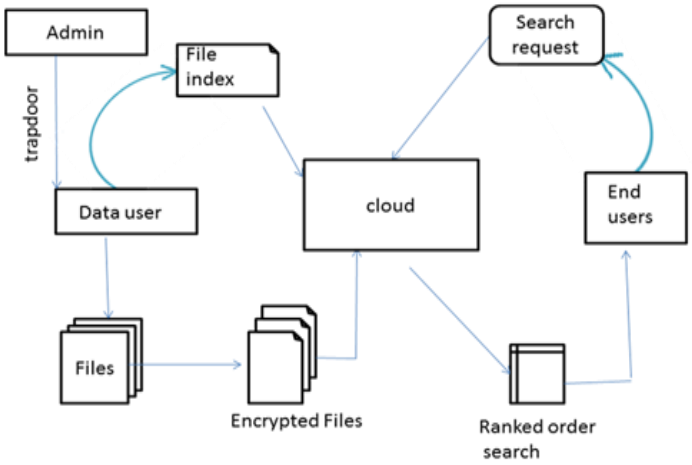


***Fig. 1: Basic model of system***

## II. LITERATURE SURVEY

Writing overview is the most imperative stride in programming improvement prepare. Before building up the device it is important to decide the time component, economy and organization quality. Once these things are fulfilled, then next stride is to figure out which working framework and dialect can be utilized for building up the apparatus. Once the software engineers begin assembling the instrument the developers require parcel of outside support. This support can be acquired from senior software engineers, from book or from sites. Before building the framework the above thought are considered for building up the proposed framework.

Boneh.D, Crescenzo G. D.,Ostrovsky.R and Persiano.G[1] depict the idea of open key encryption with catchphrase seek .Consider client Bob who sends email to client Alice scrambled under Alice's open key. An email passage needs to test whether the email contains the catchphrase "dire" with the goal that it could course the email as needs be. Alice, then again does not wish to give the passage

the capacity to unscramble every one of her messages. A component that empowers Alice to give a key to the entryway that empowers the passage to test whether "earnest" is a watchword in the email without learning whatever else about the email is appeared. This instrument is spoken to as Public Key Encryption with watchword Search. As another illustration, consider a mail server that stores different messages openly encoded for Alice by others. Utilizing this Alice can send the mail server a key that will empower the server to recognize all messages containing some particular catchphrase, however learn nothing else.

Truth be told, the major issues with past idea of security for SSE are watched, and demonstrate to outline developments which stay away from these pitfalls. Further, second arrangement additionally accomplishes what we call versatile SSE security, where inquiries to the server can be picked adaptively (by the enemy) amid the execution of the inquiry, this idea is both essential by and by and has not been already considered. Shockingly, in spite of being more secure and more effective, SSE plans are astoundingly straightforward.

As an extra commitment, multiuser SSE is considered. All earlier work on SSE concentrated the setting where just the proprietor of the information is equipped for submitting look inquiries. the regular expansion where a self-assertive gathering of gatherings other than the proprietor can submit seek questions is additionally taken into contemplations. SSE in the multi-client setting, and present an effective development that accomplishes preferred execution over essentially utilizing access control components is characterized.

Singhal.A[3] characterizes how to relegate a similitude measure to each report that demonstrates how intently it coordinates an inquiry. Boolean questions are not by any means the only strategy for hunting down data .If some correct subset of the report being looked for is known, then they are unquestionably suitable, which is the reason they have been so effective in territories, for example, business databases and bibliographic recovery frameworks .Often, notwithstanding, the data prerequisite is less definitely known. Therefore, it is once in a while helpful to have the capacity to determine a rundown of terms that give a decent sign of which reports are significant, however they won't really all be available in the archives looked for. The framework ought to rank the whole gathering as for the question, so that the main 100, say, positioned reports can be inspected for significance and those that constitute the appropriate response set extricated.

Song.D, wagner.D, and perrig.A [4] propose cryptographic plans for the issue of seeking on encoded information and give confirmations of security to the subsequent such crypto frameworks. It is attractive to store information on information stockpiling servers, for example, mail servers and record servers in scrambled frame to lessen security and protection dangers. Be that as it may, this for the most part suggests that one needs to give up usefulness for security. For instance, if a customer wishes to recover just records containing certain words, it was not beforehand known how to let the information stockpiling server play out the pursuit and answer the question without loss of information secrecy.. The strategies have various vital points of interest.

### 2.1 Disadvantages of Existing System

- When the data is modified by the owner the index is not updated and end users will not get updated data.
- Till the search results are not saved we need to search again all the keywords.
- It takes more time to recalculate the score for updated data without new index.

## III. PROPOSED APPROACH

**3.1 Problem Statement:** When user updates the data in cloud, it becomes critical again to calculate the relevance score for the data retrieval. Hence, it is necessary to find an approach for efficient data retrieval by managing the data updation so as to retrieve them efficiently.

### 3.2 Aim of this approach

- Achieving valuableexploitation of vaguely stored encrypted facts in cloud computing
- Effective use of ranked searchable symmetric encryption.
- endorsement of ranked search consequences.
- Rehashing of index list.

### 3.3 Objectives of Proposed System

- In Cloud Computing, outsourced dossieranthologyvalor not only be accessed but also restructuredrepeatedly for diverseappliance purposes.
- Score dynamics are designed as calculationonly just encrypted scores for

only justformed files, or modifying old encrypted scores for adjustment of presented files in the file gathering.

- The proposed system provides automatic updating of index and ranks based on owner modification and user access.
- It saves lot of re-computation visual projection on owners all throughindex keep posted.

### 3.4 System Architecture

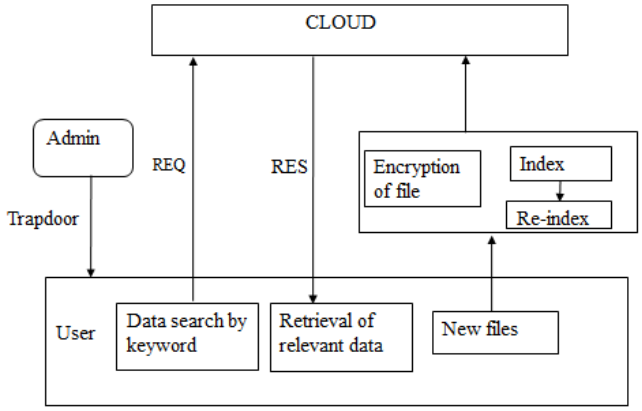


**Fig2: Architecture of Rank based index system**

The architecture of the rank based index system is briefed as follows:

1) Admin, who is the authenticholder of the catalog. Admin gives trapdoor to the authorized users.

2) Users are the members in a group who are unrestricted to admittance the in sequence of the database. Users can upload files to cloud and repossess them according to ranking

3) Fig 2 shows the ranked based index architecture.

### 3.4Module description

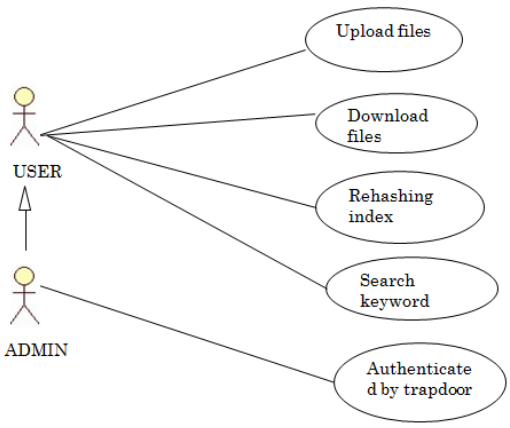


**Fig3: Process**

**Phase1:**Providing secured data transfer between owner and cloud

**Phase2:**Maintenance of index files with relevant keywords

Index structure

**Phase3:**Updating the index file at every updated data in the cloud

**Phase4:**Searching and Retrieval of files

**Phase5:** Ranking of search results

***Providing secured data transfer between owner and cloud***

As Cloud Computing gets to be distinctly predominant, more delicate data are being brought together into the cloud, for example, messages, individual wellbeing records, organization fund information, and government reports, and so on. The way that information proprietors and cloud server are no longer in the same trusted area may put the outsourced decoded information at risk[16]: the cloud server may spill information data to unapproved elements or even be hacked. It takes after that delicate information must be encoded before outsourcing for information protection and battling spontaneous gets to.This component is used to facilitate the user to upload data in a secured way using encrypt the document by DES Algorithm and to convert the encrypted document with some keys and then keys are send to the user for to retrieve the results. This module provides secured data transfer into cloud with series of encryption methods.

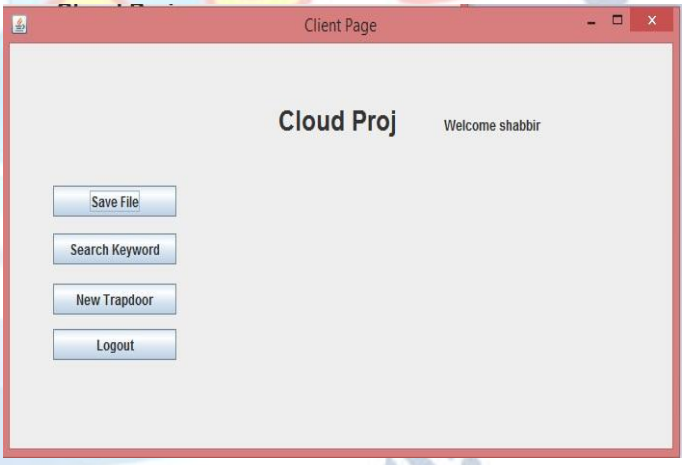


**Fig4: Client Page**

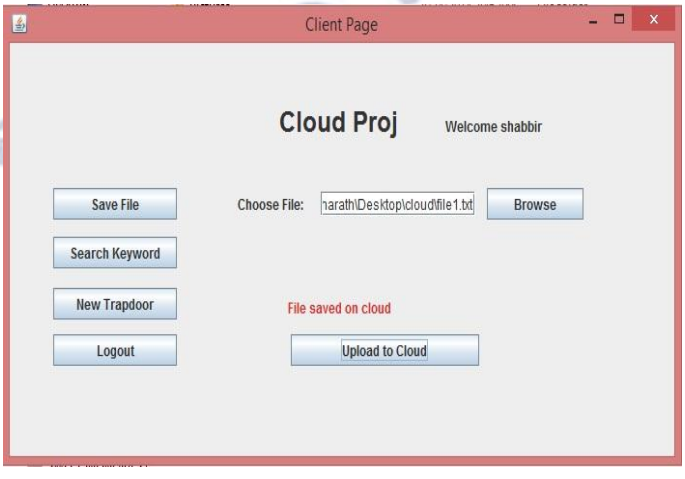

**Fig5: File Save**

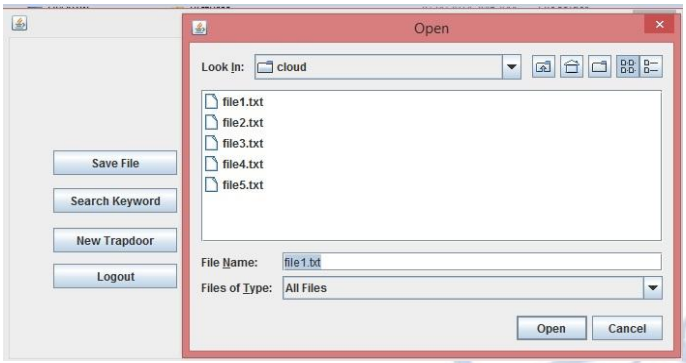


**Fig6: Multiple File Save**

***Maintenance of index files with relevant keywords***

***Index structure***

In information retrieval, indexing structure is used to stores a rundown of mappings from catchphrases to the relating set of documents that contain this watchword, permitting full pursuit. For positioned seek purposes, the errand of figuring out which documents are most applicable is normally done by allocating a numerical positions, which can be pre processed, to each record in view of some positioning strategy presented underneath.Ranking methodCollect the words which are greater than 4 letters, then search those words in file which is uploaded by user. Count those repetitions and save keyword in index along with file name, then again search the same word in other files and make count how many times it is repeated. Now rank one is given to the highest counted file for that particular keyword like that rank is calculated for remaining keywords. For example take keyword soap two files contains it so ranking is done like this. Soap – ravi.txt(3,1001) sha.txt(2,1001) here, Ravi file soap is repeated 3 times and in sha file it is repeated 2 times so rank 1 is given to Ravi and rank 2 is given to sha.


**Fig7:Keyword Search Example**

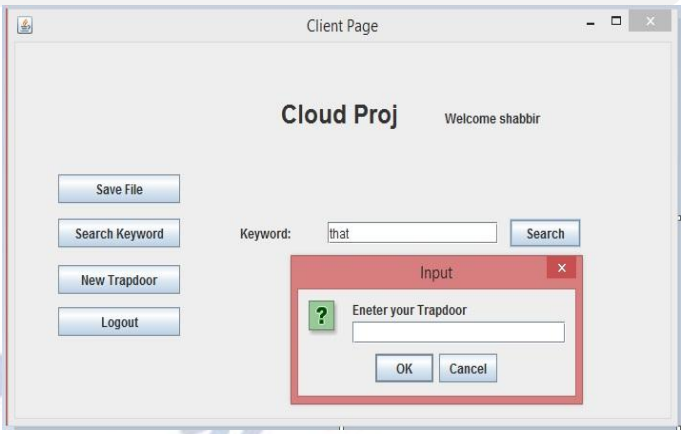


**Fig8: Trapdoor Window**

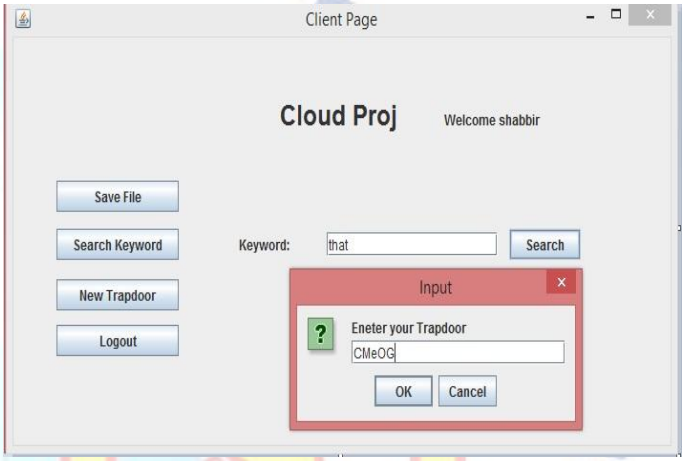


**Fig9: Trapdoor Window**

***Updating the index file at every updated data in the cloud***

For every new file updation, the index has to be modified and it should be updated with new relevancy scores and with new ranking. Two types of updation are present. They are as follows:

- Index is updated based on the previous indexing results by taking the old scores we will calculate the new scores and add to the index
- Rehashing of index is done based on users criteria .When the users choose the file based on their interest then ranking order will change.

Depending upon the two results the re-index is calculated

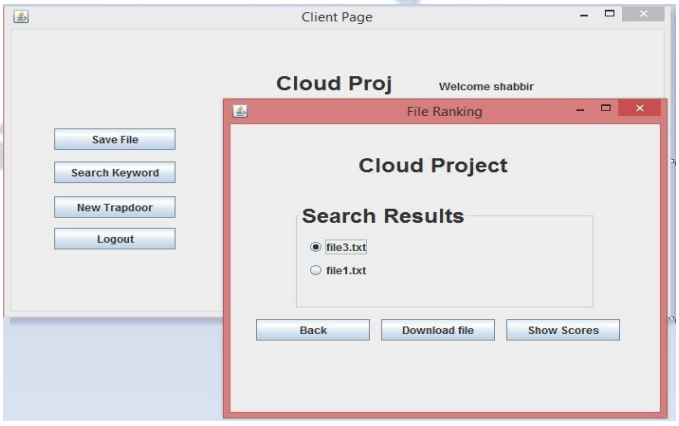


**Fig10: Search Result Window**

***Searching and Retrieval of files***

Searchable and Retrieval of files through encryption has been broadly considered with regards to cryptography. Among those works, most are centered around productivity upgrades and security definition formalizations. The main development of searchable encryption was proposed by Song et al in which each word in the archive is encoded autonomously. To accomplish more proficient pursuit comparative "record" methodologies is utilized, where a solitary scrambled hash table list is worked for the whole document accumulation. In the record table, every passage comprises of the trapdoor of a catchphrase and an encoded set of document identifiers whose comparing information documents contain the watchword. As an integral approach displayed an open key based searchable encryption plot, with an undifferentiated from situation. [10]shows the search and retrieval process, where the users can send the search keyword request to the cloud but the cloud will give the ranking order of the files retrieval response to only authorized users.

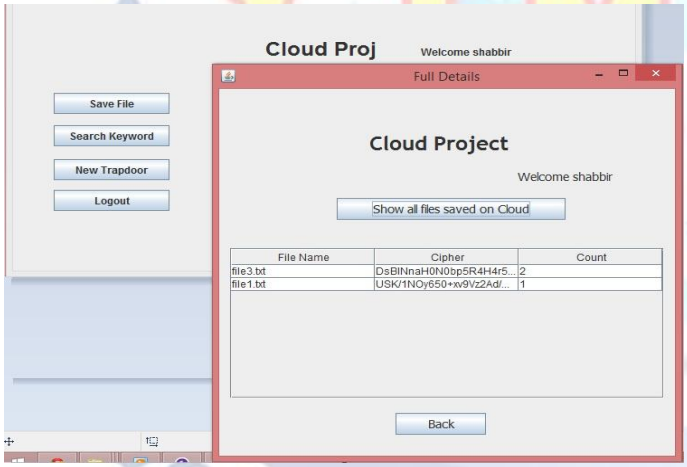


**Fig11: Show all files saved on cloud**


**Fig12: Pending Requests**


**Fig13:Trapdoor Assign to User**

***Ranking of search results***

At the point when the client scan for any magic word, the cloud server will send the top –n documents which are put away in the cloud focused around positioning in the list record. The client can choose any of the document in the top –n records focused around their advantage. The positioning request of the records will be changed focused around the picked document. Case when he chose the main 3 record in the rundown then the rank of the third document turns into one, next time when client look the decisive word the positioning of the document request is changed. This positioning of documents is spared and further utilized for the planning re-index.

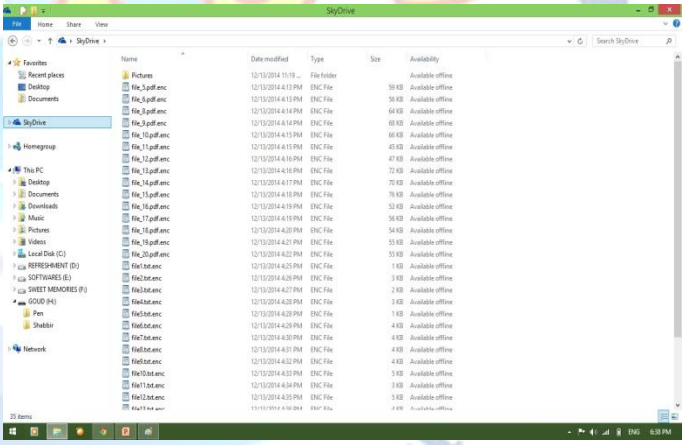

**Fig14: Skydrive folder files**

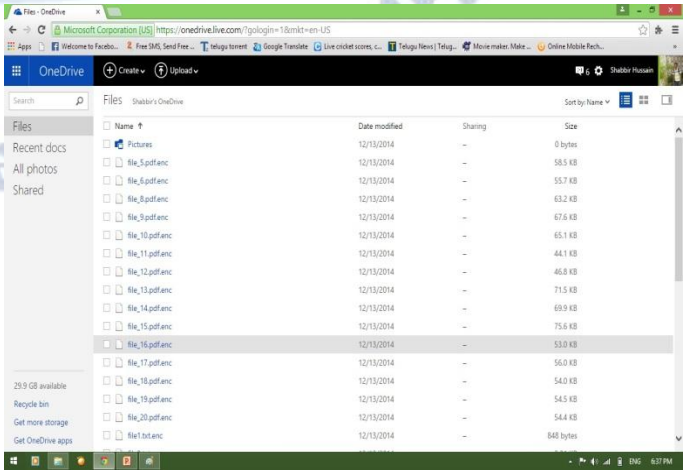


**Fig15: OneDriveCloud Files**

### 3.5 Algorithm

#### 3.5.1 Algorithm for Search

```
Algorithm cloud_search(){
search_word,index_file; while(index_file.hasNext()){
if(search_word==index_file.next()){
 display("search_word found");
  Choose file for ranking;
update index_file with new rank;
 } else {
display(Search_word not found");
  }
}
```

**Description**

step 1: Enter keyword to search
step 2: verify whether keyword is present in index file or not
step 3: if not present , display error message " word not found" . Else display the order of filenames , as placed in the index file(ranked order)
step 4: If updated by user, re arrange the order, placing the selected file 1st , and add to the index file in place of previous order
step 5: save index file,
step 6: close

#### 3.5.2 Algorithm for Ranking

```
Algorithm cloud_Save()
{
File actual_file,index_file;
while(word:=actual_file.next()){
if(word.length<4)
  ignore;
else{
  for( count=0;word:==word in actual_file;count++){
 actual_file.next();
 }
  for((iword:=index_file.next())!=null){
  if(word==iword){

temp_index.append(iword+actual_file(count,highest
_rank of iword+1));
}
 }
  if(word not found in index){
temp_index.append(word                    +"
---"+actual_file(count,1000));
 }
 }
for(index_file.hasNext(){
 if(index.next() in temp_index.next()){
 ignore ;
  } else {
temp_index.append(line from index);
   }
 }
```

**Description**

step 1: Read word by word from the file which is to be uploaded
step 2: count number of time the word repeated in that file
step 3: check for that word whether present in index file or not
step 4: *if* found, re arrange the order of previously saved file according to their word counts.
*else*
append the word with its file name and count at the end of index file
step 5: upload the file to cloud
step 6: close

## IV. CONCLUSION

Ranked keyword search on remotely stored data is done by saving files in cloud and retrieve the files by searching through the keywords. Retrieved files are presented in ranked order which is done by using ranking algorithm in the index page. Security for data stored in cloud is done trough saving encrypted files and privacy of data is maintained by providing different trapdoors to different users. Ranked analysis is done by score dynamics i.e. taking the user choices into consideration and giving highest rank to user chosen file so that user can get more efficient results.

## REFERENCES

[1] P. Mell and T. Grance, “Draft nist working definition of cloud computing,” Referenced on Jan. 23rd, 2010 Online at http://csrc. nist.gov/groups/SNS/cloud-computing/index.html, 2010.
[2] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, “Above the clouds: A berkeley view of Utiliy computing,” University of California, Berkeley, Tech. Rep. UCBEECS- 2009-28, Feb 2009.
[3] Cloud Security Alliance, “Security guidance for critical areas of focus in cloud computing,” 2009, http://www. cloudsecurityalliance.org.
[4] Z. Slocum, “Your google docs: Soon in search results?” http:// news.cnet.com/8301-17939 109-10357137-2.html, 2009.
[5] B. Krebs, “Payment Processor Breach May Be Largest Ever,” http://voices.washingtonpost.com/securityfix/2009 /01/ payment processor breach may b.html, Jan. 2009.
[6] I. H. Witten, A. Moffat, and T. C. Bell, “Managing gigabytes: Compressing and indexing documents and

images," Morgan Kaufmann Publishing, San Francisco, May 1999.

[7] D. Song, D. Wagner, and A. Perrig, "Practical techniques for searches on encrypted data," in Proc. of IEEE Symposium on Security and Privacy'00, 2000.

[8] E.-J. Goh, "Secure indexes," Cryptology ePrint Archive, Report 2003/216, 2003, http://eprint.iacr.org/.

[9] D. Boneh, G. D. Crescenzo, R. Ostrovsky, and G. Persiano, "Public key encryption with keyword search," in Proc. of EUROCRYP'04, volume 3027 of LNCS. Springer, 2004.

[10] Y.-C. Chang and M. Mitzenmacher, "Privacy preserving keyword searches on remote encrypted data," in Proc. of ACNS'05, 2005.

[11] Shabbir Hussain Shaik, Pranathi.K and Kranthi.S "Outsourced Cloud Data with Ranked Keyword Search" " in IJARCSMS ISSN: 2321 7782 (Online).

[12] R. Curtmola, J. A. Garay, S. Kamara, and R. Ostrovsky, "Searchable symmetric encryption: improved definitions and efficient constructions," in Proc. of ACM CCS'06, 2006.

[13] G Bharath Kumar, E Sai Kumar "Prism: Portion of Resources in Phase-Level Using MapReduce In Hadoop" in International Journal of Research Volume 03 Issue 10 June 2016.

[14] R. Curtmola, J. A. Garay, S. Kamara, and R. Ostrovsky, "Searchablesymmetric encryption: improved definitions and efficientconstructions," in Proc. of ACM CCS'06, 2006.

[15] A. Singhal, "Modern information retrieval: A brief overview," IEEE Data Engineering Bulletin, vol. 24, no. 4, pp. 35–43, 2001.

[16] Poojitha Koneru, Dr. S.Prabakaran "A Secured and High Octane Rank Based Analysis in Cloud Computing Environment" in International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 1973-1976