Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.017

IJCSMC, Vol. 6, Issue. 12, December 2017, pg.91 – 122

Implementation of a Secure and Dynamic Multi-Keyword Ranked Search Scheme over Encrypted Cloud Data

A.P.Jaware, N.R.Borkar

Computer Science and Engineering .Amravati University INDIA
anita.jaware@gmail.com; namrata.borkar@gmail.com

Abstract— Due to the increasing popularity of cloud computing, more and more data owners are motivated to outsource their data to cloud servers for great convenience and reduced cost in data management. However, sensitive data should be encrypting before outsourcing for privacy requirements, which obsoletes data utilization like keyword-based document retrieval. In this paper, we present a secure multi-keyword ranked search scheme over encrypted cloud data, which simultaneously supports dynamic update operations like deletion and insertion of documents.

Keywords— multi-keyword ranked search, encryption, decryption, dynamic update, cloud computing

I. INTRODUCTION

Cloud computing has been considered as another model of enterprise IT infrastructure, which can compose gigantic resource of computing, storage and applications, and empower users to appreciate pervasive, helpful and on-demand network access to a mutual pool of configurable computing resources with incredible efficiency and insignificant economic overhead [1]. Pulled in by these engaging features, both individuals and enterprise are used to outsource their data to the cloud, rather than buying software and hardware to deal with the data themselves. In spite of the different points of interest of cloud services, outsourcing delicate information, (for example, e-mail, individual health records, organization account information, government archives, and so forth.) to remote servers brings privacy concerns. The cloud service providers (CSPs) that keep the data for users may access users' sensitive information without authorization. A general way to deal with secure the data privacy is to encrypt the data before outsourcing [2].On the other hand, this will bring about a gigantic expense in terms of data ease of use. For example, the current techniques on keyword-based

© 2017, IJCSMC All Rights Reserved

information retrieval, which are broadly utilized on the plaintext data, can't be straightforwardly connected on the encrypted data. Downloading all the data from the cloud and decrypt locally is clearly unrealistic.

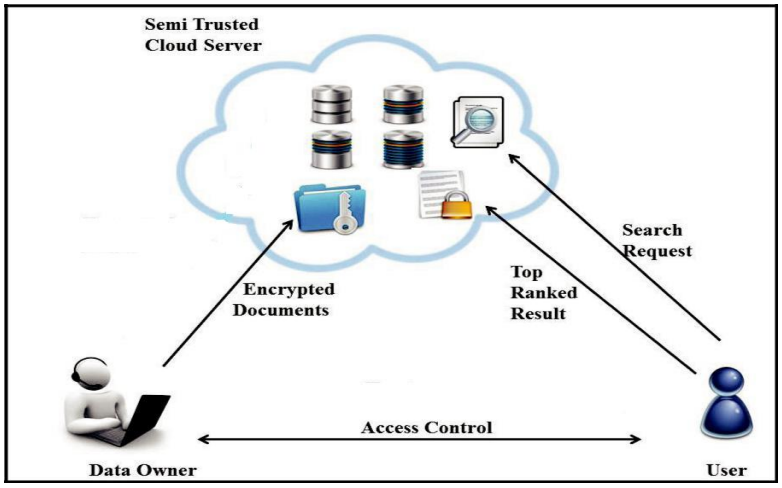


Figure1.1 Architecture of the search over encrypted cloud data

In the simplest terms, cloud computing means storing and accessing data and programs over the Internet instead of your computer's hard drive. The cloud is just a metaphor for the Internet. What cloud computing is not about is your hard drive. When you store data on or run programs from the hard drive, that's called local storage and computing. Everything you need is physically close to you, which means accessing your data is fast and easy, for that one computer, or others on the local network. Working off your hard drive is how the computer industry functioned for decades; some would argue it's still superior to cloud computing, for reasons I'll explain shortly.
For it to be considered "cloud computing," you need to access your data or your programs over the Internet, or at the very least, have that data synchronized with other information over the Web. In a big business, you may know all there is to know about what is on the other side of the connection; as an individual user, you may never have any idea what kind of massive data processing is happening on the other end. The result is the same with an online connection, cloud computing can be done anywhere, anytime.

With a particular final objective to address the above issue, analysts have illustrated some all around helpful arrangements with totally homomorphism encryption [3] or missing RAMs [4]. In any case, these schedules are not down to earth in light of their high computational overhead for both the cloud server and user. In spite of what may be normal, more useful unique reason arrangements, for instance, searchable encryption (SE) plan have made specific responsibilities to the extent productivity, value and security. Searchable encryption scheme engage the user to store the encrypted data to the cloud and execute unequivocal word look for over cipher text domain. As being what is indicated, abundant works have been proposed under assorted risk models to finish distinctive interest value, for instance, single keyword search, closeness look, multi-keyword Boolean search, ranked search, multi-keyword ranked search, etc. Among them, multi keyword positioned quest finishes more thought for its pragmatic propriety. Starting late, some component arrangements have been proposed to reinforce embedding and erasing operations on archive gathering. These are colossal goes about as it is exceptionally possible that the data owner need to overhaul their data on the cloud server.

Yet, few of the dynamic plan support successful multi keyword situated look Proposed cloud storage systems that offer privacy, reliability and authentication of data user against a UN trusted cloud provider. This data user firstly registers for login to see data in cloud but when client try to login then user got a message that

© 2017, IJCSMC All Rights Reserved

he is not an active client until owner not active data user. Data owner have all rights to active client to see cloud data. when data user search a file and want to upload or download the file, then data owner will decided whether send rights or not to data user for uploading and downloading data to utilize the file Presently in the existing system the cloud server hosts third-party data storage and get back services. As information may have sensitive information, the cloud servers cannot be fully hand over in protecting data. For this cause, outsourced files must be encrypted. Any type of data leakage that would involve data privacy considered as undesirable. To meet the effective data retrieval requirement, the huge amount of documents command the cloud server to achieve result relevance ranking, as an alternative of returning undifferentiated results. Such ranked search system allows data users to find the most appropriate information quickly, rather than burdensomely sorting through every match in the content collection. Ranked search can also gracefully eradicate avoidable network traffic by transferring back only the most appropriate data, which is highly attractive in the "pay-as-you-use" cloud concept. For privacy protection, such ranking process, yet, should not leak any keyword related information. On the other hand, to progress the search result correctness as well as to improve the user searching experience, it is also essential for such ranking system to maintain multiple keywords search, as single keyword search regularly yields far too common results.

1.2 ORGANIZATION OF REST OF THE CHAPTERS OF THE REPORT

This Project Consist of:

1. Algorithm for Encryption and decryption in the file, which are used for secured the files data.
2. After securing the files, authenticated user want to access the file so need to search a file by using multiple keywords.
3. In this User gets numbers of file that contain those multiple keywords then user need to ranked that particulars file.
4. By using ranked user gets access most relevant file with total time that required for the file.

II. LITERATURE SURVEY

2.1 Literature Survey

Cloud computing transforms the way information technology (IT) is expended and oversaw, promising Enhanced expense efficiencies, quickened development, speedier time-to-market, and the capacity to scale Applications on interest (Leighton, 2009). [1]. as per Gartner, while the buildup developed exponentially amid2008 and proceeded since, it is clear that there is a noteworthy movement towards the cloud computing model and that the advantages may be significant (Gartner Hype-Cycle, 2012). Be that as it may, as the cloud's state Processing is rising and growing quickly both theoretically and actually, the legitimate/contractual, monetary, administration quality, inter-operability, security and protection issues still posture critical difficulties. In this Part, we depict different services and organization models of distributed computing and recognize significant Difficulties.

© 2017, IJCSMC All Rights Reserved

2.2 Security challenges for the public cloud

Kui Ren, Cong Wang, and Qian Wang, In this paper, Cloud computing represents today's most exciting computing paradigm shift in information technology. However, security and privacy are perceived as primary obstacles to its wide adoption. Here, the authors outline several critical security challenges and motivate further investigation of security solutions for a trustworthy public cloud environment. cloud computing is the newest term for the long-dreamed vision of computing as a utility. The cloud provides convenient, on-demand network access to a centralized pool of configurable computing resources that can be rapidly deployed with great efficiency and minimal management overhead. With its un-precedence advantages, cloud computing enables a fundamental paradigm shift in how To deploy and deliver computing services that is, it makes possible computing outsourcing such that both individuals and enterprises can avoid committing large capital outlays when purchasing and managing software and hardware, as Toll as dealing with the operational overhead therein.[1]

2.3 Cryptographic cloud storage

S. Kamara and K. Lauter, In this paper, To consider the problem of building a secure cloud storage service on top of a public cloud infrastructure where the service provider is not completely trusted by the customer. To describe, at a high level, several architectures that combine recent and non-standard cryptographic primitives in order to achieve our goal. To survey the benefits such an architecture would provide to both customers and service providers and give an overview of recent advances in cryptography motivated specifically by cloud storage. [2] A fully homomorphism encryption scheme C. Gentry, In this paper, To propose the first fully homomorphism encryption scheme, solving a central open problem in cryptography. Such a scheme allows one to compute arbitrary functions over encrypted data without the decryption key { i.e., given encryptions E(m1); … ; E( ) of m1;… ; mt, one can efficiently compute a compact cipher text that encrypts f(m1;… ; m) for any efficiently computable function f. This problem posed by Rivest et al. in 1978. Fully homomorphic encryption has numerous applications. For example, it enables private queries to a search engine { the user submits an encrypted query and the search engine computes a succinct encrypted answer without ever looking at the query in the clear. It also enables searching on encrypted data { a user stores encrypted files on a remote file server and can later have the server retrieve only files that (when decrypted) satisfy some Boolean constraint, even though the server cannot decrypt the files on its own. More broadly, fully homomorphic encryption improves the efficiency of secure multiparty computation. Our construction begins with a somewhat homomorphism boot-strappable" encryption scheme that works when the function f is the scheme's own decryption function. To then show how, through recursive self-embedding, boot-strappable encryption gives fully homomorphic encryption. The construction makes use of hard problems on ideal lattices. [3]

2.4 Software protection and simulation on oblivious rams

O. Goldreich and R. Ostrovsky, [4] In this paper, To present a theoretical treatment of software protection. In particular, To distill and formulate the key problem of learning about a program from its execution, and reduce this problem to the problem of on-line simulation of an arbitrary program on an oblivious RAM. To then present our main result: an efficient simulation of an arbitrary (RAM) program on a probabilistic oblivious

© 2017, IJCSMC All Rights Reserved

RAM. Assuming that one-way functions exists, To show how one can make our software protection scheme robust against a polynomial-time adversary who is allotted to alter memory contents during execution in a dynamic fashion.

2.5 Public key encryption with keyword search

D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano, In this paper, To study the problem of searching on data that is encrypted using a public key system. Consider user Bob who sends email to user Alice encrypted under Alice's public key. An email gateway wants to test whether the email contains the keyword \urgent" so that it could route the email accordingly. Alice, on the other hand does not wish to give the gateway the ability to decrypt all her messages. To define and construct a mechanism that enables Alice to provide a key to the gateway that enables the gateway to test whether the word urgent" is a keyword in the email without learning anything else about the email. As another example, consider a mail server that stores various messages publicly encrypted for Alice by others. Using our mechanism Alice can send the mail server a key that will enable the server to identify all messages containing some specific keyword, but learn nothing else. To define the concept of public key encryption with keyword search and give several constructions [5].

2.6 Public key encryption that allows pir queries

D.Boneh, E. Kushilevitz, R. Ostrovsky, and W. E. Skeith III, In this paper, Consider the following problem: Alice wishes to maintain her email using a storage-provider Bob (such as a Yahoo! or hotmail email account). This storage-provider should provide for Alice the ability to collect, retrieve, search and delete emails but, at the same time, should learn neither the content of messages sent from the senders to Alice (with Bob as an intermediary), nor the search criteria used by Alice. A trivial solution is that messages will be sent to Bob in encrypted form and Alice, whenever she wants to search for some message, will ask Bob to send her a copy of the entire database of encrypted emails. This however is highly inefficient. To will be interested in solutions that are communication-efficient and, at the same time, respect the privacy of Alice. In this paper, To show how to create a public-key encryption scheme for Alice that allows PIR searching over encrypted documents. Our solution is the first to reveal no partial information regarding the user's search (including the access pattern) in the public-key setting and with nontrivially small communication complexity. This provides a theoretical solution to a problem posed by Boneh, DiCrescenzo, Ostrovsky and Persiano [6] on "Public-key Encryption with Keyword Search." The main technique of our solution also allows for Single-Database PIR writing with sub linear communication complexity, which to consider of independent interest.

2.7 Practical techniques for searches on encrypted data

D. X. Song, D. Wagner, and A. Perrig, In this paper, It is desirable to store data on data storage servers such as mail servers and file servers in encrypted form to reduce security and privacy risks. But this usually implies that one has to sacrifice functionality for security. For example, if a client wishes to retrieve only documents containing certain words, it was not previously known how to let the data storage server perform the search and answer the query without loss of data confidentiality. In this paper, To describe our cryptographic schemes for the problem of searching on encrypted data and provide proofs of security for the resulting crypto

© 2017, IJCSMC All Rights Reserved

systems. Our techniques have a number of crucial advantages. They are provably secure; they provide provable secrecy for encryption, in the sense that the un-trusted server cannot learn anything about the plaintext when only given the cipher text. they provide query isolation for search, meaning that the un-trusted server cannot learn anything more about the plaintext than the search result; they provide controlled searching. So, That the un-trusted server cannot search for an arbitrary word without the user's authorization. they also support hidden queries, so that the user may ask the un-trusted server to search for a secret word without revealing the word to the server. The algorithms To present are simple, fast (for a document of length n, the encryption and search algorithms only need stream cipher and block cipher operations), and introduce almost no space and communication overhead, and hence are practical to use today.[7]

2.8 Privacy preserving keyword searches on remote encrypted data

Y.-C. Chang and M. Mitzenmacher,[8] In this paper, To consider the following problem: a user U wants to store his files in an encrypted form on a remote file server S. Later the user U wants to efficiently retrieve some of the encrypted files containing (or indexed by) specific keywords, keeping the keywords themselves secret and not jeopardizing the security of the remotely stored files. For example, a user may want to store old e-mail messages encrypted on a server managed by Yahoo or another large vendor, and later retrieve certain messages while travelling with a mobile device. In this paper, To offer solutions for this problem under Toll-defined security requirements. Our schemes are efficient in the sense that no public-key cryptosystem is involved. Indeed, our approach is independent of the encryption method chosen for the remote files.

III.PROBLEM DEFINITION

3.1 ANALYSIS OF PROBLEM

In previous system, the cloud service providers (CSPs) that keep the data for users may access users' sensitive information without authorization. A general approach to protect the data confidentiality is to encrypt the data before outsourcing [2]. However, this will cause a huge cost in terms of data usability. For example, the existing techniques on keyword-based information retrieval, which are widely used on the plaintext data,cannot be directly applied on the encrypted data. Downloading all the data from the cloud and decrypt locally is obviously impractical. Searchable encryption schemes enable the client to store the encrypted data to the cloud and execute keyword search over cipher text domain. Multi-keyword Boolean search allows the users to input multiple query keywords to request suitable documents. Among these works, conjunctive keyword search schemes [15], [16], [17] only return the documents that contain all of the query keywords. Disjunctive keyword search schemes [18], [19] return all of the documents that contain a subset of the query keywords. Predicate search schemes [20], [21], [22] are proposed to support both conjunctive and disjunctive search. All these multi-keyword search schemes retrieve search results based on the existence of keywords, which cannot provide acceptable result ranking functionality. Ranked search can enable quick search of the most relevant data. Sending back only the top-*k* most relevant documents can effectively decrease network traffic. Some early works [23], [24], [25] have realized the ranked search using order-preserving techniques, but they are designed only for single keyword search. In this dissertation, we can solve the above problem by using Xoring algorithm for encryption. We can also search multiple keywords with relevant ranking.

© 2017, IJCSMC All Rights Reserved

3.1.1 Problem 1 (regarding encryption)

Given a file 1.text user want to upload file on to the cloud. In the cloud number of user connected so any one can access my file information. My file 1.txt having private data then I need to secured my file 1.txt

- **Definition 1(Answer)**

  Given file 1.txt user need to secure file 1.txt on to the cloud so that other user can not access this private information. For this purpose user first need to study for cryptography then decided which encryption algorithm is suitable to secure 1.txt file. In this thesis we can provided simple and easy to implemented algorithm for encryption is Xoring. By using Xoring user can secure 1.txt file in this scenario when user upload 1.txt file then this plain text convert into cipher text this cipher text cannot understand by other user and the result is user file will be secured.

3.1.2 Problem 2(regarding searching)

When user uploads a file on a cloud then authenticated user want to search particular file. Now, user try to searching a file by entering file keyword then as a output user will gets numbers of file those contain this keyword but the problem is that user can not decided which is the best one that he needed.

- **Definition (Answering)**

  Suppose user trying to search documents. For example user want to search a document that information related to object oriented programming then user enter single keywords like object user gets number of file that contain keyword object now user should be confused which is the most important file by resolving this problem in this proposed system we can used multi keywords search. By using multi keywords, search user gets appropriate documents.

3.1.3 Problem3 (regarding ranking)

In searching process when user search a particular file by entering multiple keywords as input then he got number of file that contain these multi keywords but user confuse which one is the best.

- **Definition (Answering)**

  In this searching process when user gets number of file but which is the best one is decided by using ranking this ranking state that how many(times keyword) count occurs in the files.

  By using count user decided which the appropriate file that he want to search.

3.2 HARDWARE REQUIREMENT

- Hardware - Pentium Processor
- Speed – 1.1GHz
- RAM – 2 GB
- Hard Disk -20GB
- Floppy Drive – 1.44 MB

3.3 SOFTWARE REQUIREMENT

- Operating System –Windows
- Technology –Dot net
- IDE – Visual Studio

© 2017, IJCSMC All Rights Reserved

- Database –SQL Server 2014
- Visual Studio Version –Visual Studio 2015

IV.METHODLOGY

During the course of exploring this area of research, we came across certain that we have to know for research. These things are discussing below-

4.1 CLOUD COMPUTING

We are seeing technology moving to the cloud. It is not just a fad—the shift from traditional software models to the Internet has steadily gained momentum over the last 10 years. Looking ahead, the next decade of cloud computing promises new ways to collaborate everywhere cloud computing is a kind of outsourcing of computer programs. Using cloud computing, users are able to access software and applications from wherever they need, while it is being hosted by an outside party in "the cloud." This means that they do not have to worry about things such as storage and power, they can simply enjoy the end result Life before cloud computing.

Traditional business applications have always been very complicated and expensive. The amount and variety of hardware and software required to run them are daunting. You need a whole team of experts to install, configure, test, run, secure, and update them. When you multiply this effort across dozens or hundreds of apps, it is easy to see why the biggest companies with the best IT departments aren't getting the apps they need. Small and mid-sized businesses don't stand a chance. Cloud computing: a better way with cloud computing, you eliminate those headaches that come with storing your own data, because you're not managing hardware and software — that becomes the responsibility of an experienced vendor like salesforce.com. The shared infrastructure means it works like a utility: you only pay for what you need, upgrades are automatic, and scaling up or down is easy.

Cloud-based apps can be up and running in days or weeks, and they cost less. With a cloud app, you just open a browser, log in, customize the app, and start using it.

Businesses are running all kinds of apps in the cloud, like customer relationship management (CRM), HR, accounting, and much more. Some of the world's largest companies moved their applications to the cloud with salesforce.com after rigorously testing the security and reliability of our infrastructure. As a cloud computing grows in popularity, thousands of companies are simply rebranding their non-cloud products and services as "cloud computing." Always dig deeper when evaluating cloud offerings and keep in mind that if you have to buy and manage hardware and software, what you're looking at isn't really cloud computing but a false cloud.

There are three types of cloud computing:

- **Infrastructure as a Service (IaaS)-** A third party hosts elements of infrastructure, such as hardware, software, servers, and storage, also providing backup, security, and maintenance.
- **Software as a Service (SaaS)-** Using the cloud, software such as an internet browser or application is able to become a usable tool.
- **Platform as a Service (PaaS)-** The branch of cloud computing that allows users to develop, run, and manage applications, without having to get caught up in code, storage, infrastructure and so on. There

© 2017, IJCSMC All Rights Reserved

are several types of PaaS. Every PaaS option is public, private, or a hybrid mix of the two. Public PaaS is hosted in the cloud and its infrastructure managed by the provider. Private PaaS, on the other hand, is housed in on-site servers or private networks, and is maintained by the user. Hybrid PaaS uses elements from both public and private, and is capable of executing applications from multiple cloud infrastructures. PaaS can be further category depending on whether it is open or closed source. When choosing a PaaS solution, the most important considerations beyond how it is hosted are how well it integrates with existing information systems, which programming languages it supports, what application-building tools it offers, how customizable or configurable it is, and how effectively it is supported by the provider.

As digital technologies grow ever more powerful and available, apps and cloud-based platforms are becoming almost universally widespread. Businesses are taking advantage of new PaaS capabilities to further outsource tasks that would have otherwise relied on local solutions. This is all made possible through advances in cloud computing. Traditional business applications have always been very complicated and expensive. The amount and variety of hardware and software required to run them are daunting. You need a whole team of experts to install, configure, test, run, secure, and update them.

When you multiply this effort across dozens or hundreds of apps, it's easy to see why the biggest companies with the best IT departments aren't getting the apps they need. Small and mid-sized businesses don't stand a chance. The affordability of cloud-hosted data makes it an essential tool for these types of situations.

Here are some other benefits of cloud computing.

- **Adaptable**: Cloud computing allows for adaptable programs and applications, that are customizable, while allowing the owners control over the core code. Multi-tenancy Cloud software provides the opportunity to provide personalized applications and portals to a number of customers or tenants.
- **Reliable:** Because it is hosted by a third party, businesses and other users have greater assurance of reliability, and when there are problems, easy access to customer support.
- **Scalability**: With the Internet of Things, it is essential that software functions across every device and integrates with other applications. Cloud applications can provide this. Secure Cloud computing can also guarantee a more secure environment, thanks to increased resources for security and centralization of data.

4.2 Security

Cloud computing is the delivery of computing services over the Internet. Cloud services allow individuals and businesses to use software and hardware that are managed by third parties at remote locations. Examples of cloud services include online file storage, social networking sites, webmail, and online business applications. The cloud computing model allows access to information and computer resources from anywhere that a network connection is available. Cloud computing provides a shared pool of resources, including data storage space, networks, computer processing power, and specialized corporate and user applications. Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. Because of these benefits each and every organizations are moving their data to the cloud. So there is a need to protect that data against

© 2017, IJCSMC All Rights Reserved

unauthorized access, modification or denial of services etc. To secure the Cloud means secure the treatments (calculations) and storage (databases hosted by the Cloud provider). Security goals of data include three points namely: Availability Confidentiality, and Integrity. Confidentiality of data in the cloud is accomplished by cryptography.

4.2.1 Xoring Algorithm

The XOR operator is extremely common as a component in more complex ciphers. By itself, using a constant repeating key, a simple XOR cipher can trivially be broken using frequency analysis. If the content of any message can be guess or otherwise known then the key can be reveal. Its primary merit is that it is simple to implement, and that the XOR operation is computationally inexpensive. A simple repeating XOR cipher is therefore sometimes use for hiding information in cases where no particular security is required. If the key is random and is at least as long as the message, the XOR cipher is much more secure than when there is key repetition within a message. There will be basic formula for XOR encryption algorithm that count number of character.

Plain Text= number of character in file

Encrypted text = 4* Plain text

XOR Algorithm

1. Convert string TO Hex value
2. Convert hex value to binary string
3. Convert Key into Hex Value
4. Convert Hex to binary string
5. Perform XOR operation
6. X1 =X Xor Key
7. Add length of key
8. X2 =X1+Length of a key
9. Convert X2 into 16 bit
10. Split the binary string X2 to 4 byte chucks and assign each the proper b64 value
11. m<--0
12. if i=0 and i<binary data. Length then
13. i+=4
14. v<--Bin to Dec (i,4)
15. Cipher += GetB64Form (v*4+m)
16. If ++m>3 then
17. m=0
18. Return cipher
19. Decode
20. m = 0
21. Convert the b64 string to binary string
22. v<--(GetNFromB64(c) - m) / 4
23. binary data += Dec To Binary (v,4)
24. If ++m>3 then
25. m=0
26. Chop the 8-bit binaries and mix back the key into it
27. if i = 0 and i < binary data. Length then
28. if i + 16 > binary data. Length
29. break
30. c<-- Bin To Dec(i, 16)
31. dc = (c - key. Length) ^ (int) key [keypos]
32. ++keypos >= key. Length

© 2017, IJCSMC All Rights Reserved

33. keypos = 0;
34. Decoded += new string ((char)dc, 1);
35. Return Decode

4.2 KEYWORDS SEARCH

Keyword search has been extensively used for extracting information from structured data such as relational databases, which can modelled as multi keywords. Users of such databases usually do not have sufficient knowledge about the structure of data, and are often not familiar with query languages such as SQL. Thus, they need a simple system that receives some keywords and returns a set of files that together cover all or part of the input keywords.

4.3 PROPOSED SYSTEM

In our proposed system, we used the one algorithm for encryption i.e. Xoring algorithm our system works by following way:

4.3.1 Register

User need to register first by clicking register page and fill information inside register page get user id and Password by using this id and Password user gets login.

4.3.2 Access Permission

By using valid Id and password user try to login but never get login until owner get access permission to hat user when owner gave permission then and only then user gets login. After login user need access permission to upload and download the files so user again need to send permission for access to the owner and owner itself decided which access permission allotted to that user.

4.3.3 Upload The File

For our project after access permission user want to upload file over cloud when user upload file into the cloud firstly user store the files into database of various extension like txt,doc,pdf etc.

4.3.4 Encrypt File

When user upload file into cloud this file not stored in plaintext format. These files should store in cipher text format so that an unauthenticated user should not read this file. For this purpose we need to used encryption algorithm we used Xoring algorithm for encryption.

4.3.5 Search files

When authenticated user want a particular file then user enter some keywords and search that file in the result of output user gets numbers of files that contains these keywords. By using ranking user gets most appropriate file.

V. SYSTEM MODULE

In this chapter, we have given a graphical representation of the system, which is useful to understand the systematic flow of the data from which it is easy to reach at required goal. This chapter includes the data flow diagrams.

© 2017, IJCSMC All Rights Reserved

5.1 DATA FLOW DIAGRAM

A Data Flow Diagram (DFD) is a graphical tool that allows system analyst to depict the flow of data in an information system. The DFD is one of the method that system analyst used to collect information necessary to determine information system requirements. Data flow diagram (DFD) is used to show how data flows through the system and the processes that transform the input data into output. Data Flow diagrams is a way of expressing system requirements in a graphical manner.DFD represents one of the most ingenious tools used for structured analysis.

LEVEL 0

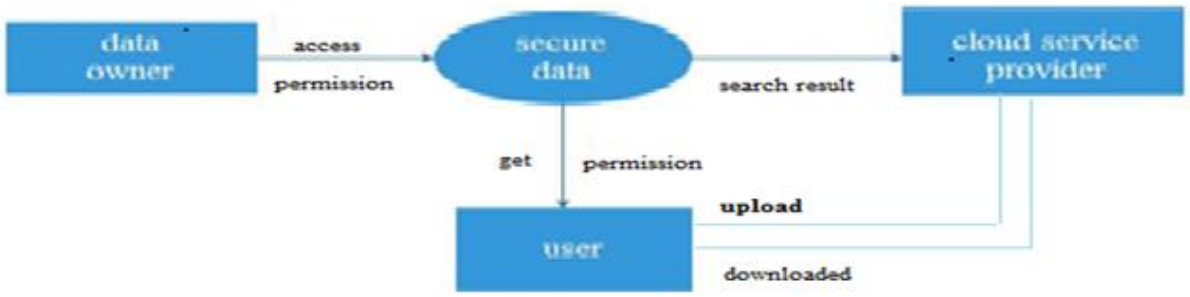


Figure 5.1: Level 0 DFD for Proposed System

The above figure 5.1 depicts Level 0 DFD for secured multiply keywords search. it shortly seems that how our system will flow and user will interact with system for uploading files, searching files and downloading files, so system containing Encryption Algorithm for securing data.

Level 1

In level 1 Data Flow Diagram is intended to serve as a communication tool among

- System analysts
- End users
- Database designers
- System programmers
- Other member of the project team

In this DFD user representing by external entity is the origin of destination of data entities are external to the system. In this DFD shows numbers of processes doing in our system, with their sequence flow, process performs some action on data, such as extract the keywords, retrieve files, ranking the sequence of files etc.

© 2017, IJCSMC All Rights Reserved

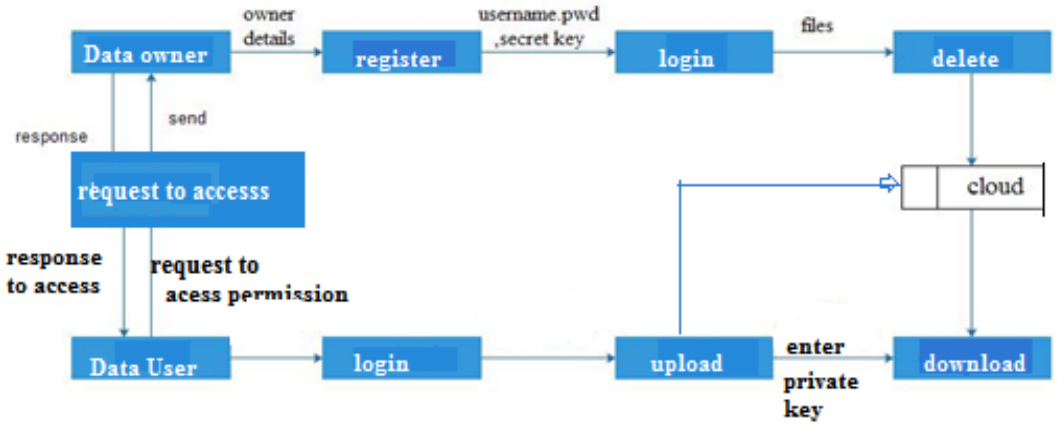


Figure 5.2: Level 1 DFD for Proposed System

Figure 5.2 shows the data flow diagram of our system. In the above figure 5.2 when user get the query it passes to next level i.e. getting input keywords that keywords are process and system search for that keywords. Now, System matches the keywords into database through the files and ranking to the file according to the count. Finally, display the actual files that we want.

5.2 UNIFIED MODELING LANGUAGE

Unified Modelling Language (UML) is a graphical language for visualizing, specifying, constructing and documenting the artifacts of software. UML is expressive language, addressing all the views needed to develop and then deploy such system.

5.2.1 Use case Diagram

Use case diagram is one of the UML diagrams which specify actors and their roles in the system. Following figure shows actors as user and system and their roles are also specified. Roles of the actor are entering input keyword, Pick relevant files. Roles of the system are giving extract the relevant keyword, match the keywords with files in database and store in database, ranking the files base on input keyword and displaying relevant result.

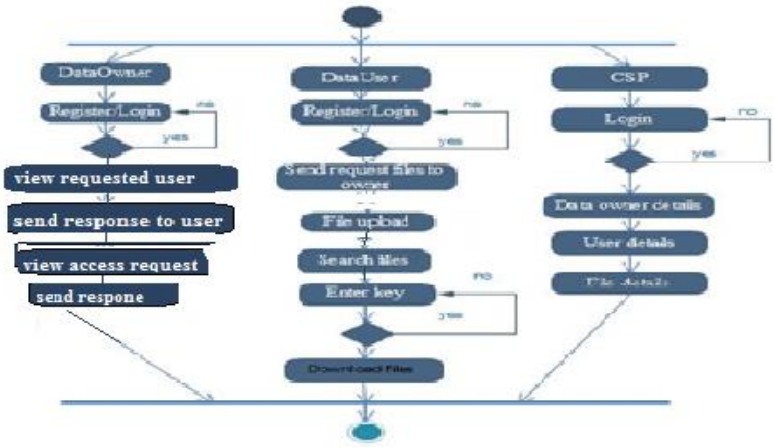


Figure 5.3: Use Case Diagram

© 2017, IJCSMC All Rights Reserved

VI. SYSTEM IMPLEMENTATION

6.1 INTRODUCTION

This chapter includes the system implementation details for secure and dynamic multi keywords ranked search scheme over encrypted cloud data. Microsoft Sql server 2012 for storing data, Microsoft Visual Studio 2015 is used for implementation of system. It integrates computation, visualization, and programming in easy to use environment. As discussed in literature review, there are various approaches are available for secured and dynamic multi keywords ranked search scheme over encrypted cloud data, each with its own merits and demerits. In our implementation, we have used Xoring encryption algorithm for security and words by words search by using search button and cloud is used to stored files.

6.2 DATA OWNER LOGIN FOR PROVIDING ACCESS TO THE US ER

In our Project, there are three main fields

1. Data Owner
2. Data User
3. Cloud

The figure 6.1 shows the Data Owner Login form that providing access to the authenticated data user. By using this window Data owner login after that data owner, check how many access request come from data user and decided to which user gets this access.

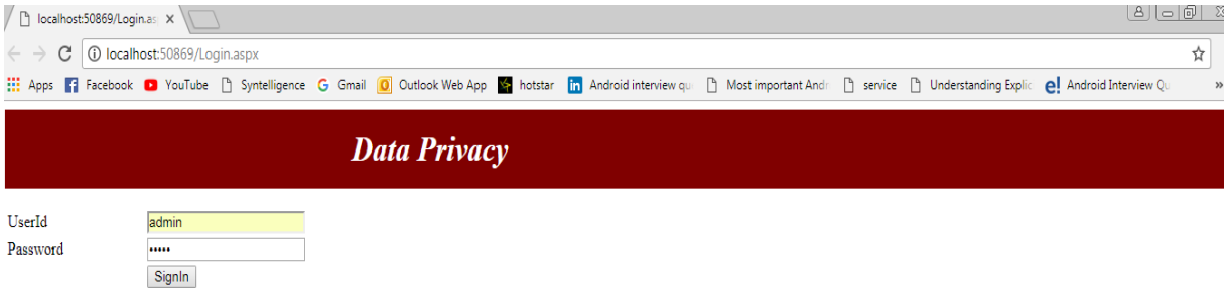


Figure 6.1: Login Module for Data Owner

6.3 REGISTER DATA USER FOR LOGIN

In the previous window, we saw login form of data owner. Now in figure 6.2 we see the register window for data user. In this figure 6.2 shows user registration page that require user information like full name, mobile number, email address, passwords and confirm password. When user enter email address, password and

© 2017, IJCSMC All Rights Reserved

click to register button then these email id and passwords will be using in Data user login form so we need to remember email address and password.

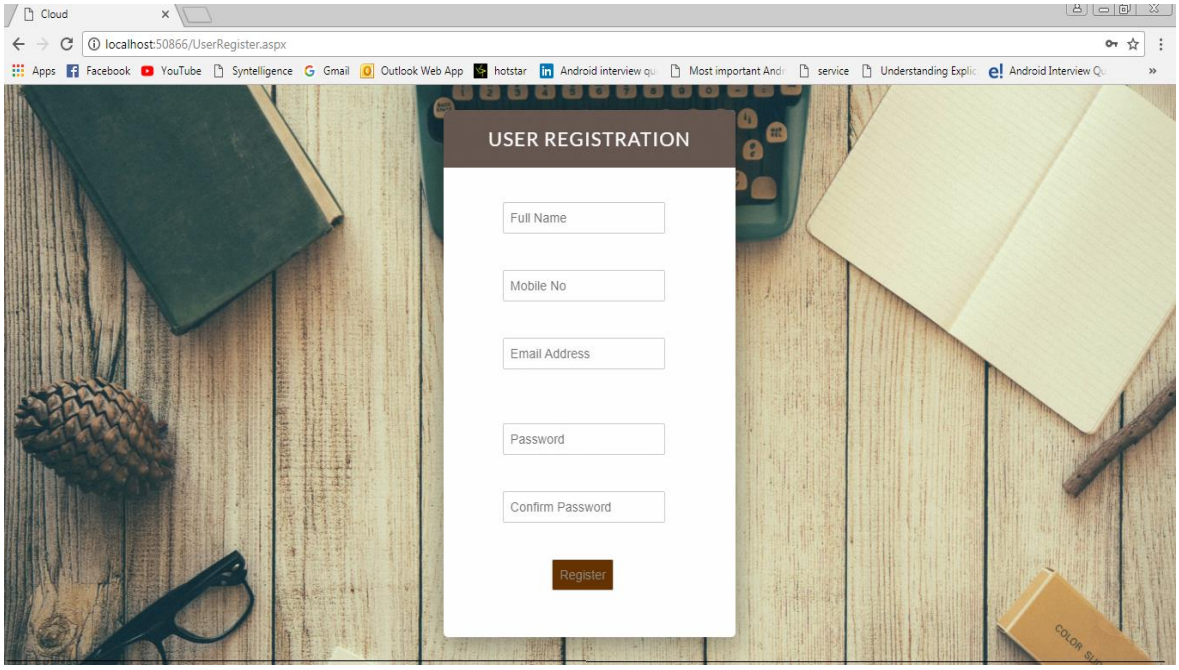


Figure 6.2: Registration Page for Data User Login

6.4 DATA USER LOGIN FOR A SECURE MULTI KEYWORDS RANKED SEARCH.

The figure 6.3 show login page for Data User in this figure user will try to login with the help of UserId and Password that created after register. So user enter valid user id and password.

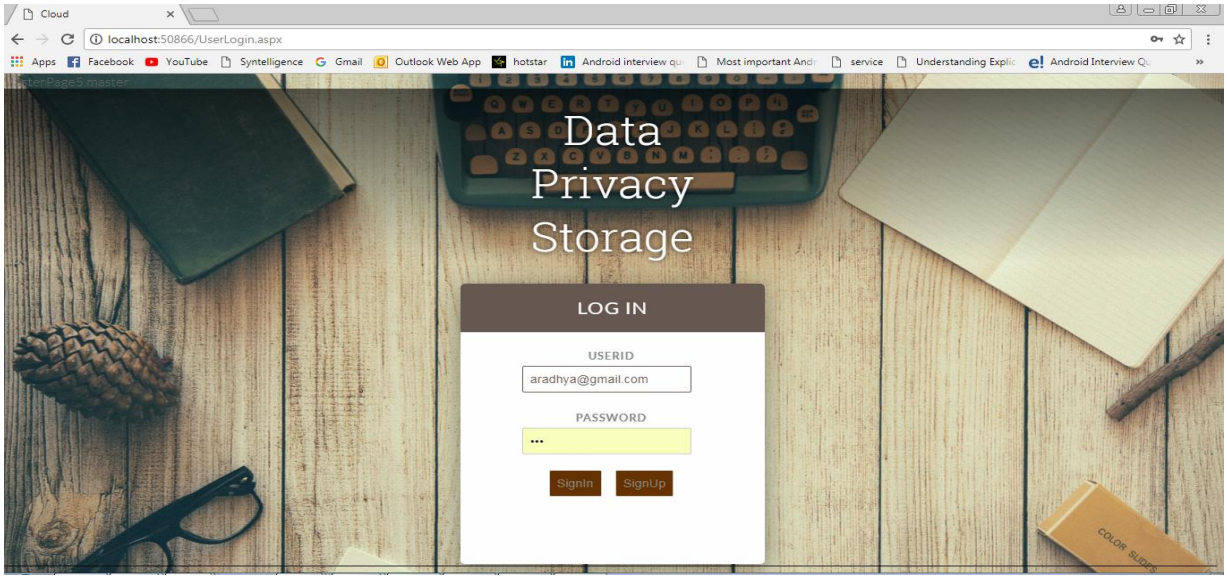


Figure 6.3: Login Page for Data User

By using valid user id and password user will not gets login. In this case, user needs to activate their account from data owner. We can see the message in figure 6.4 like sorry your account not active.

© 2017, IJCSMC All Rights Reserved

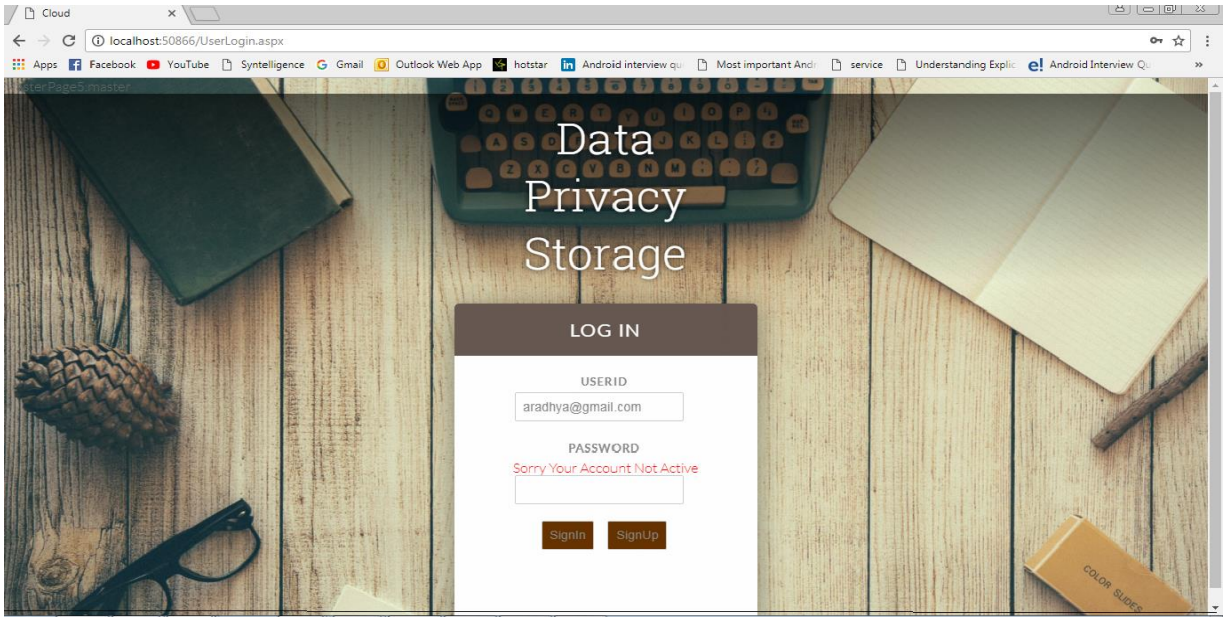
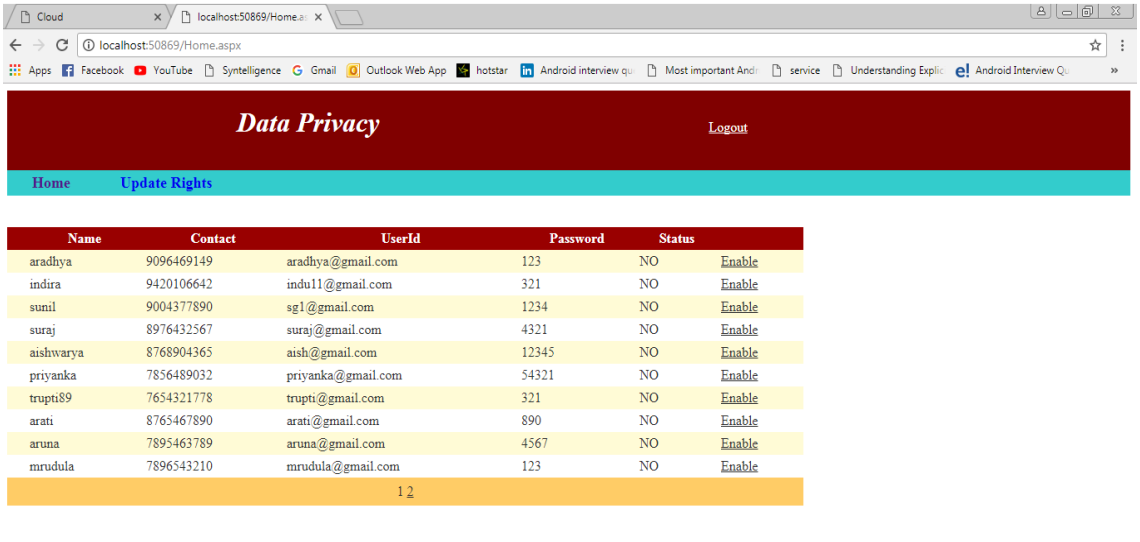


Figure 6.4: Data User Account Not Active

6.5 DATA OWNER ACTIVE AUTHENTICATED DATA USER

In above window we saw that data user would not get login. When user enter correct user id and password and click to sign up button then user request will automatically send to the data owner. Data owner decided to enable user or not. If data owner will not enable user then data user will not access these account. The figure 6.5 shows the list of data users that are waiting to enable from data owner.



Name	Contact	UserId	Password	Status	
aradhya	9096469149	aradhya@gmail.com	123	NO	Enable
indira	9420106642	indu11@gmail.com	321	NO	Enable
sunil	9004377890	sg1@gmail.com	1234	NO	Enable
suraj	8976432567	suraj@gmail.com	4321	NO	Enable
aishwarya	8768904365	aish@gmail.com	12345	NO	Enable
priyanka	7856489032	priyanka@gmail.com	54321	NO	Enable
trupti89	7654321778	trupti@gmail.com	321	NO	Enable
arati	8765467890	arati@gmail.com	890	NO	Enable
aruna	7895463789	aruna@gmail.com	4567	NO	Enable
mrudula	7896543210	mrudula@gmail.com	123	NO	Enable

1 2

Figure 6.5: List of Data Users that are not active

The figure 6.6 show Activation window is used to providing rights to the data user. In this owner select user name and decided which rights allotted to which user as shown in figure.

© 2017, IJCSMC All Rights Reserved

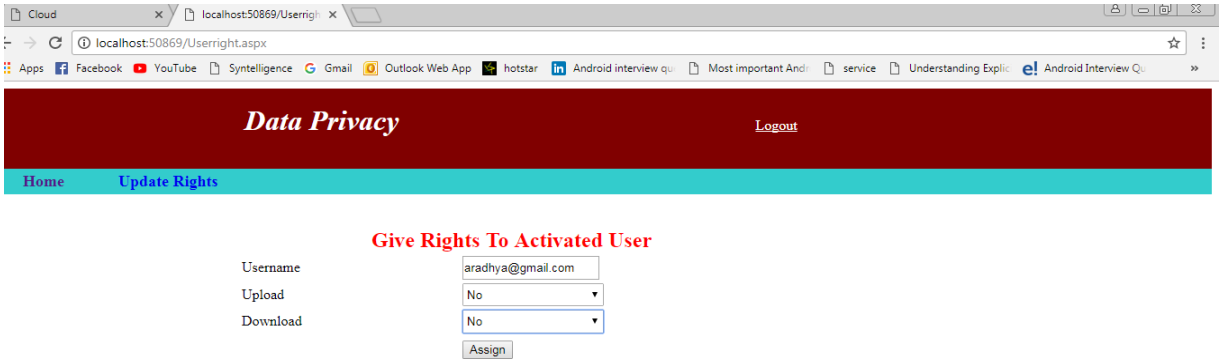


Figure 6.6 Activation Window

In the figure 6.7 shows Username waiting to enable. In this module could not display User Name like aradhya,indira,sunil,suraj as shown in figure 6.5 so that all Data user were enable. User name like aishwarya,priyanka and so on that all were not enable as shown in below figure.

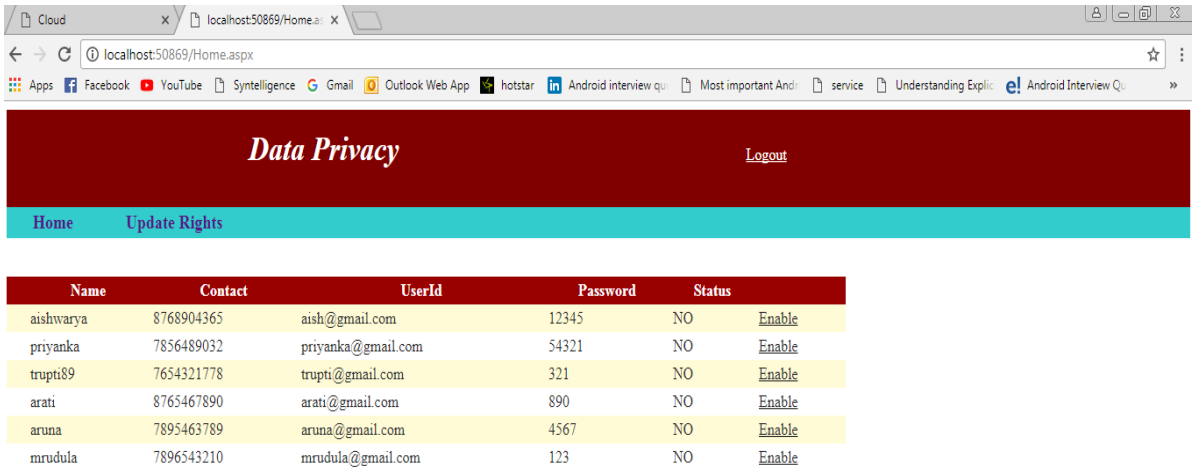


Figure 6.7: list of Inactivated Data User

Owner Provided Rights to the data user but currently Data Owner cannot provide any rights like upload and download to the data user. As shown in figure 6.6 data owner could not provides any rights to the user means user only gets login but not doing any rights if user want any rights then user need to send request to the Data owner. Data owner is only owner that decided which right allotted to which Data User.

6.6 DATA USER LOGIN AFTER ACTIVATED FROM DATA OWNER.

In the figure 6.8 shows, the data user gets login after activated from Data User.

© 2017, IJCSMC All Rights Reserved

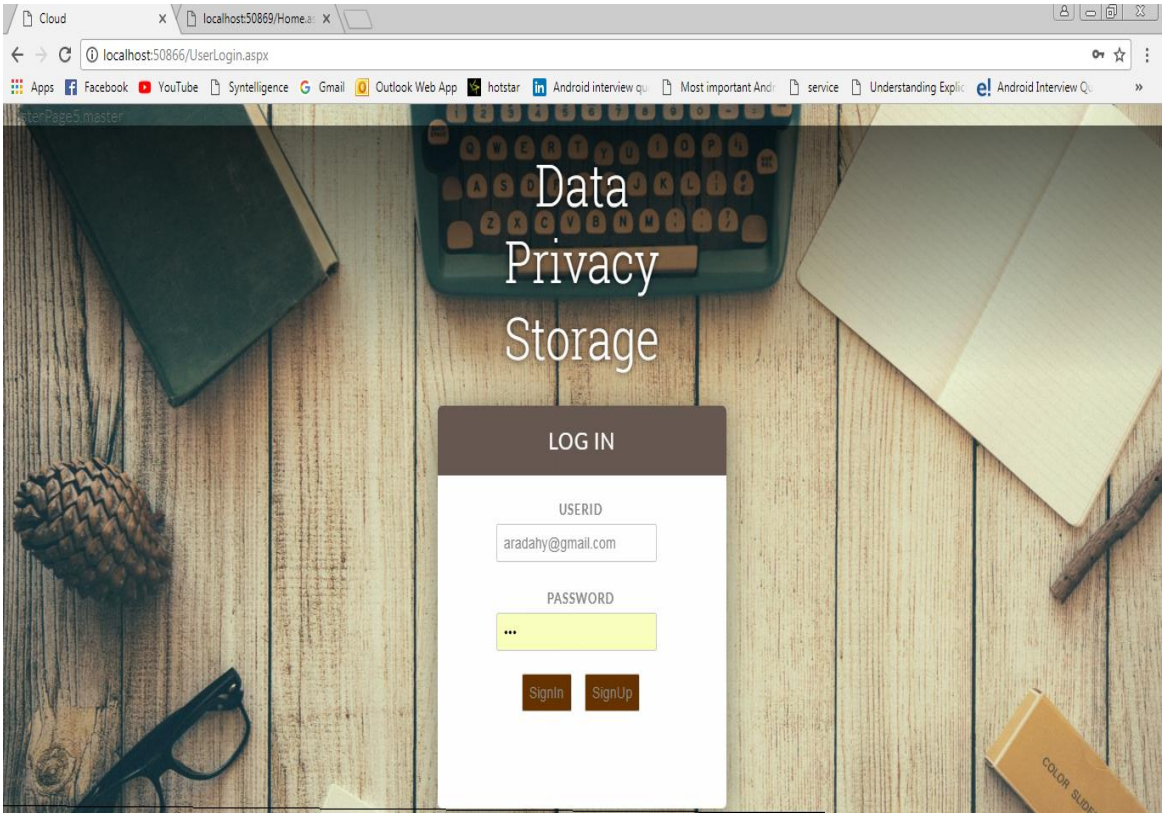


Figure 6.8: Login Window for Data User

In the figure 6.8 user get login by using valid Userid and Password. When new User wants to get login them user not only required registering but also need data owner activation permission. After getting Data owner Permission then user will get to login in this module.

6.7 ARCHITECTURE OF A SECURE AND DYNAMIC MULTI KEYWORDS RANKED SEARCH OVER ENCRYPTED CLOUD DATA

The figure 6.9 shows the GUI of main windows which shows the principle steps in a secure and dynamic multi keywords ranked search over encrypted cloud data.

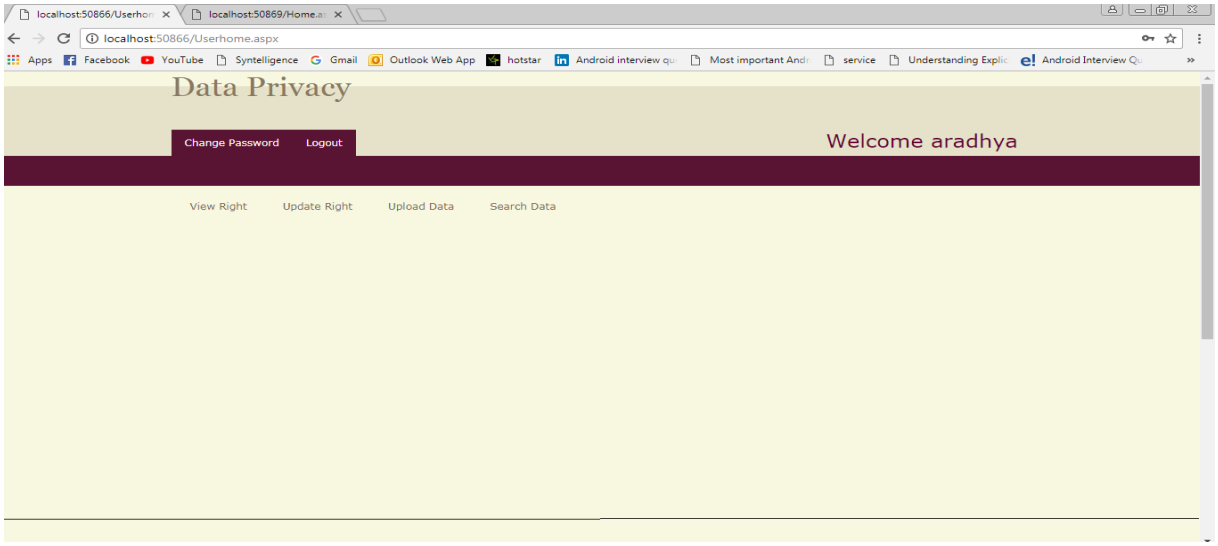


Figure 6.9: Main GUI Module of Project

© 2017, IJCSMC All Rights Reserved

The figure 6.9 shows main GUI of system module of a secure and dynamic multi keywords ranked search over encrypted cloud data. In the main GUI of our propose system there are four main links i.e.

1. View Right
2. Update Right
3. Upload Data
4. Search Data

6.8 VIEW RIGHT

This is the first step of secure and dynamic multi keywords ranked search over encrypted cloud data. In this user can view the rights that allotted from Data owner.

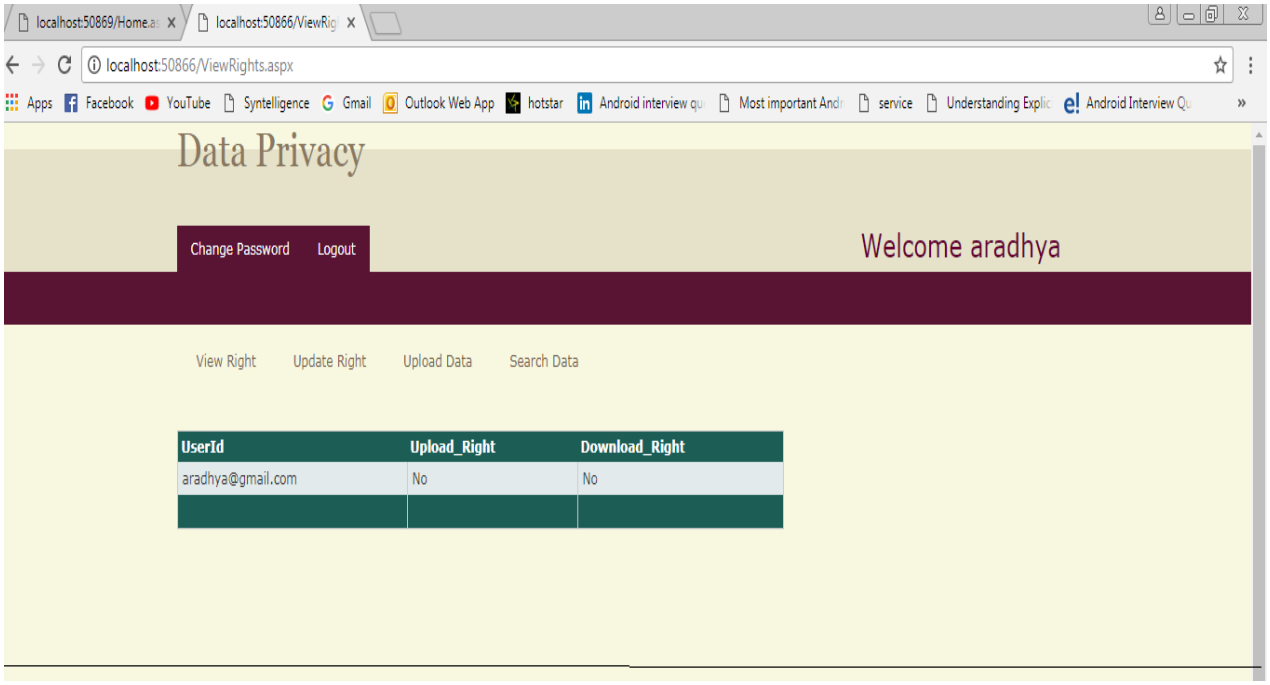


Figure 6.10: GUI of View Right

The figure 6.10 shows GUI of View Right i.e. in this we can view rights like Upload and Download that allotted from data owner in this figure Data owner cannot provide any rights to the data user that is why user can view upload and download rights allocates to No.

6.8 UPDATE RIGHT

This is the second step of secure and dynamic multi keywords ranked search over encrypted cloud data. In this user can update the rights that allotted from Data owner by sending request to the owner for update rights.

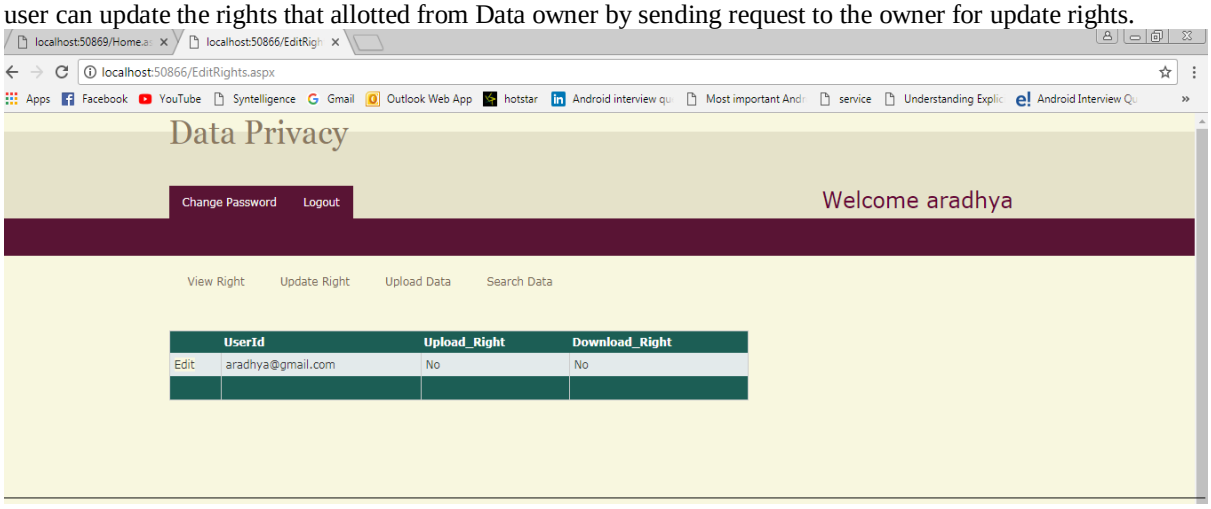


Figure 6.11: Update Window

© 2017, IJCSMC All Rights Reserved

In the figure 6.11 update window there is a Edit button when user click to the Edit button to update rights. In this figure 6.12 User edit right like upload and download to yes and send request to the data owner and wait for Administrator reply for update rights

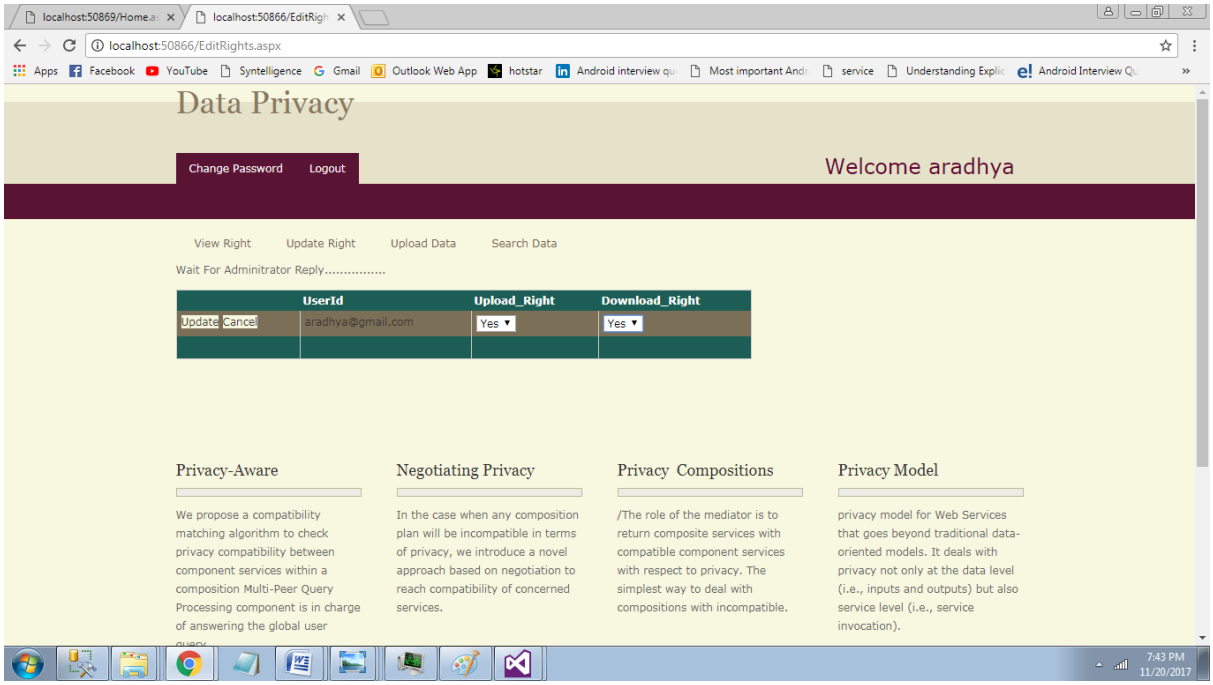


Figure 6.12 Update Window send request to Data Owner

When data user is send request to the owner for update rights then data owner received request and process that request to update right by selecting Userid to that particular user the figure 6.13 shows the windows for requested right and original rights to the data owner.

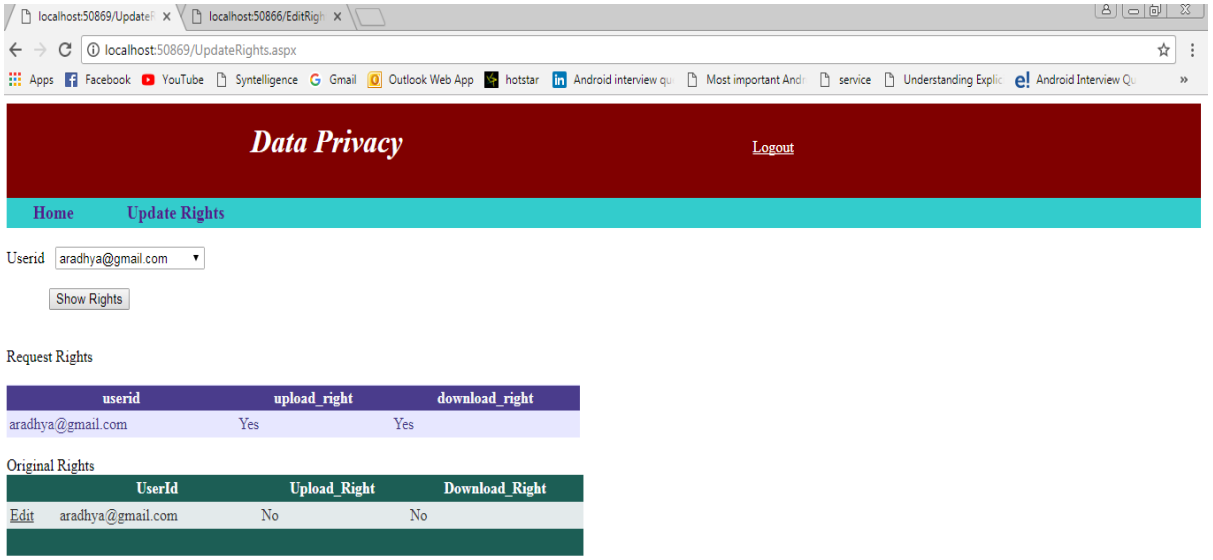


Figure 6.13 Update Rights window of Data owner

When owner want to update rights then owner click to the edit button of original right and select upload right and download rights to yes then click update button it will display message like Right Update success

© 2017, IJCSMC All Rights Reserved

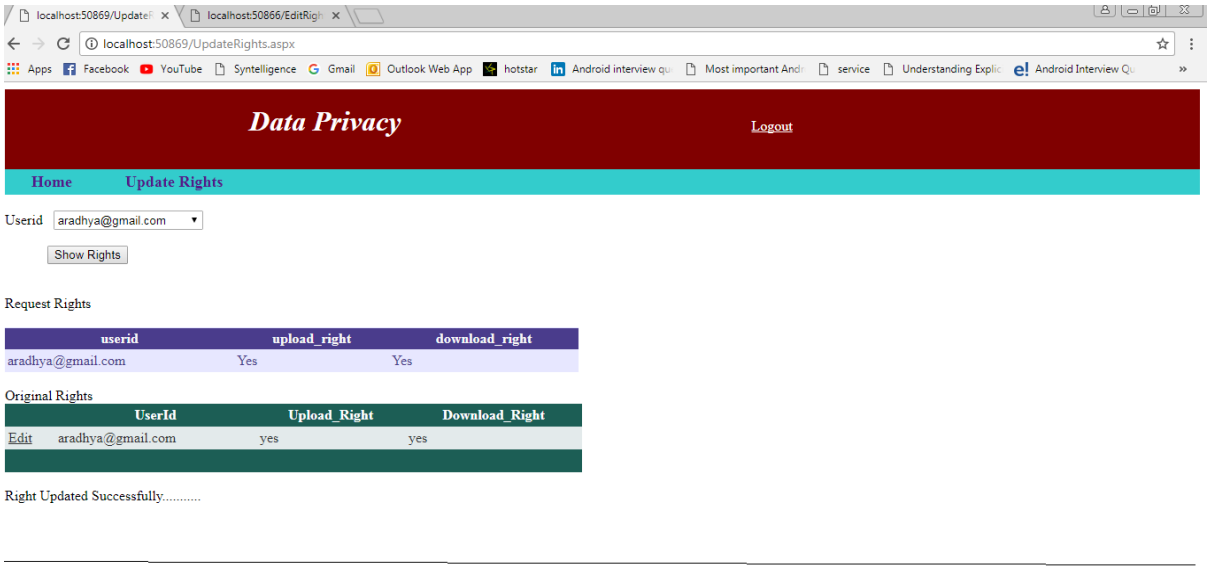


Figure 6.14: Completing Update process window of Data Owner

In figure 6.15 shows the updated rights window to the data user. In this window, display update right update from no to yes this allotted from data owner.

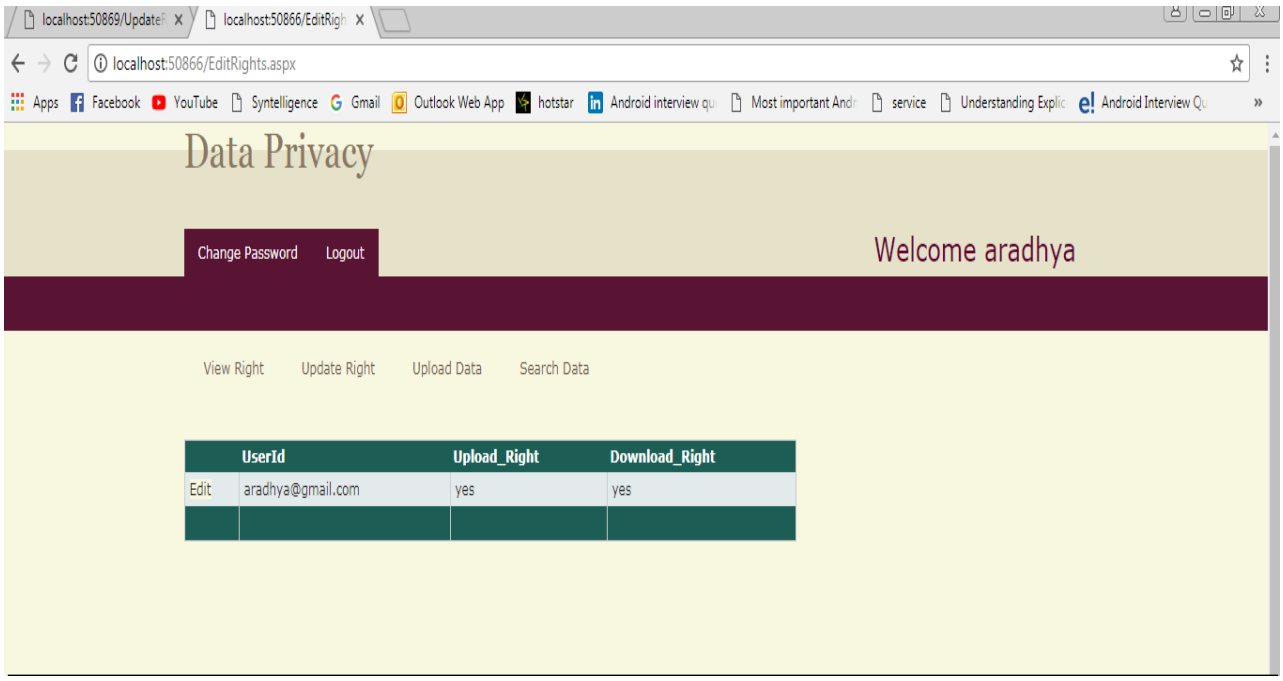


figure 6.15: updated right window

These are the entire step that follows by Update rights from data user to data owner from figure 6.11to 6.15.

6.10 UPLOAD RIGHT

Upload right in this module of data user upload file those he want to secure. In this we can upload numbers of file.

© 2017, IJCSMC All Rights Reserved

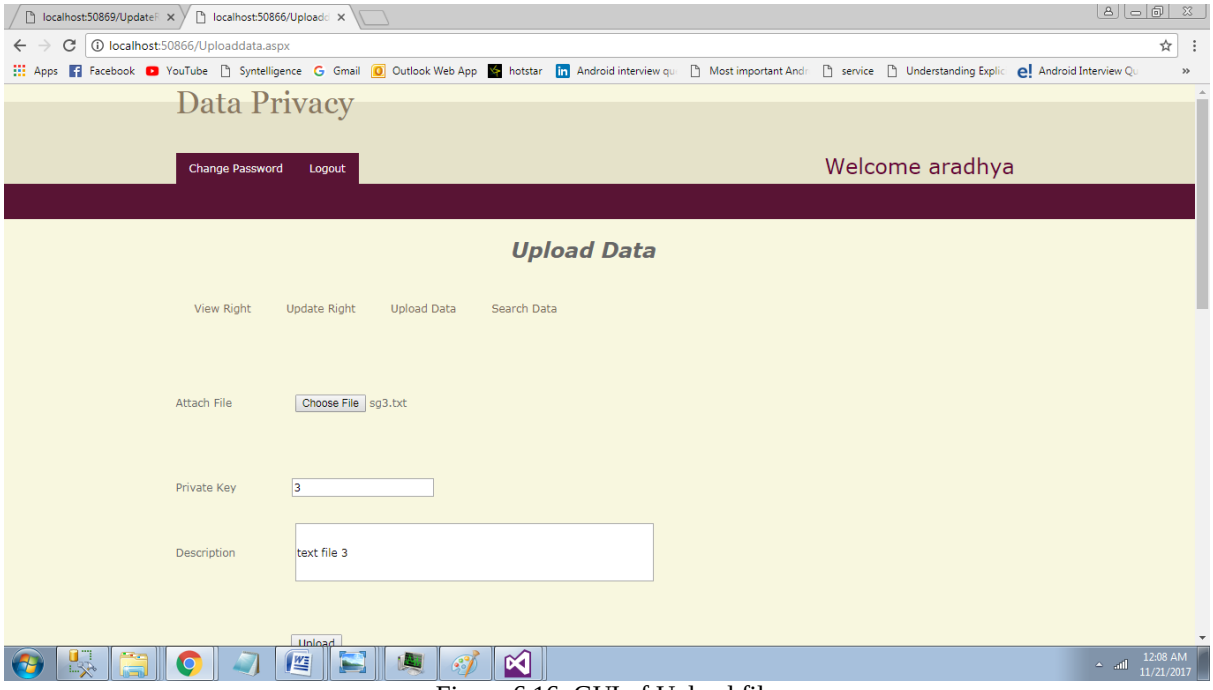


Figure 6.16: GUI of Upload file

In the figure 6.16 shows, GUI of upload file attachment i.e. in this we can upload a file with primary key and description. Uploaded file first stored in the database as well as cloud in the encrypted format for more security. In figure 6.17 shows time for encryption and text in the file that are in the encrypted form.

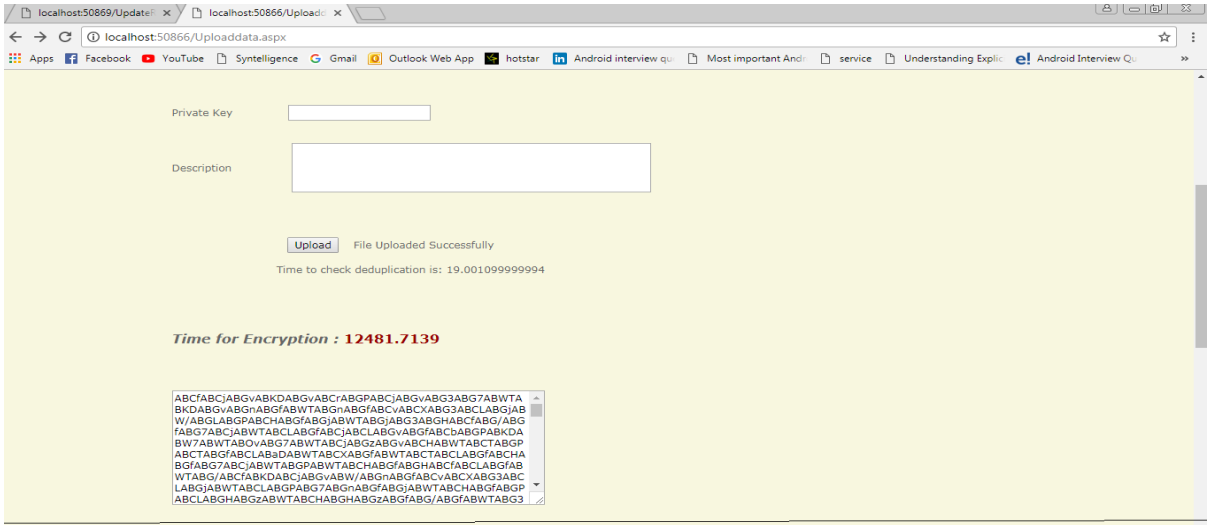


Figure 6.17 GUI of uploaded files in encrypted format

6.11 SEARCH DATA

In the search data when user want to search a particular file that contain those multi keywords. With the help of multiple keywords user search a file.

© 2017, IJCSMC All Rights Reserved

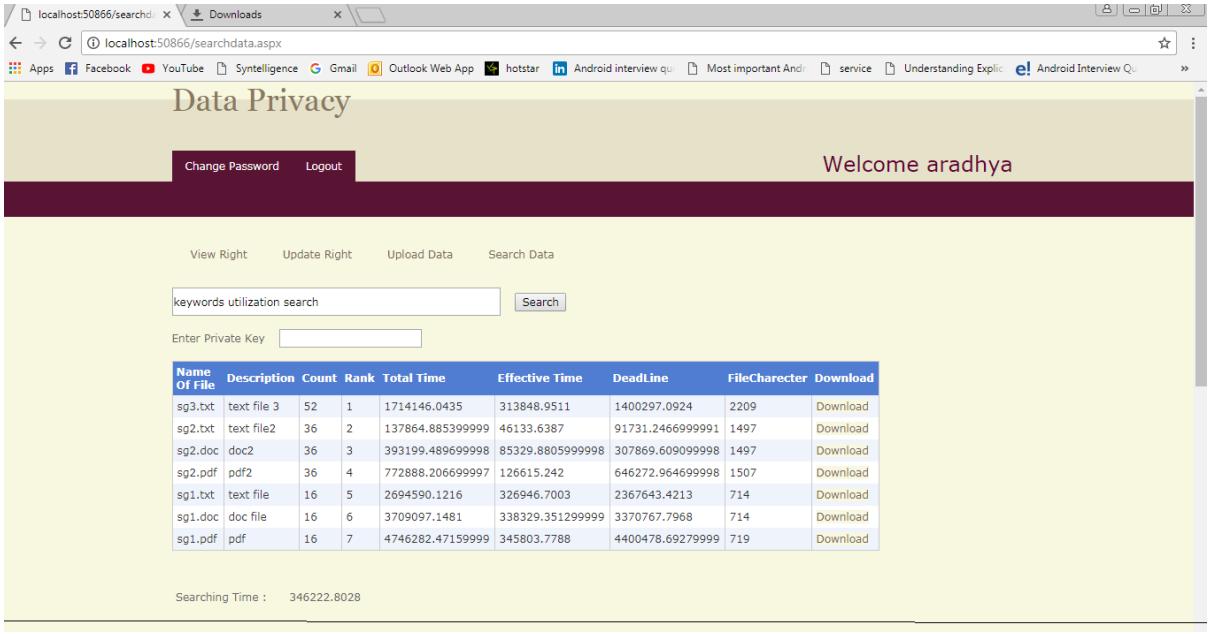


Figure 6.18: GUI of Search Module

In the figure 6.18 shows the GUI of search module, this is use to search by typing multiple keywords. In this module user, type keywords by using this search button search how many times this keywords used in the file according to this it will display count with rank. By using rank, it will consider which one the most appropriate file for the search. in this module include total time required for particular file, effective time also display, deadline and file character also display. It will display searching time that will decided how many time required to search keywords in the numbers of files.

In the Figure 6.19 shows download window when user click into download button. If the primary key is set to that file then it will require primary key for downloading the file. Otherwise, file will not be downloading.

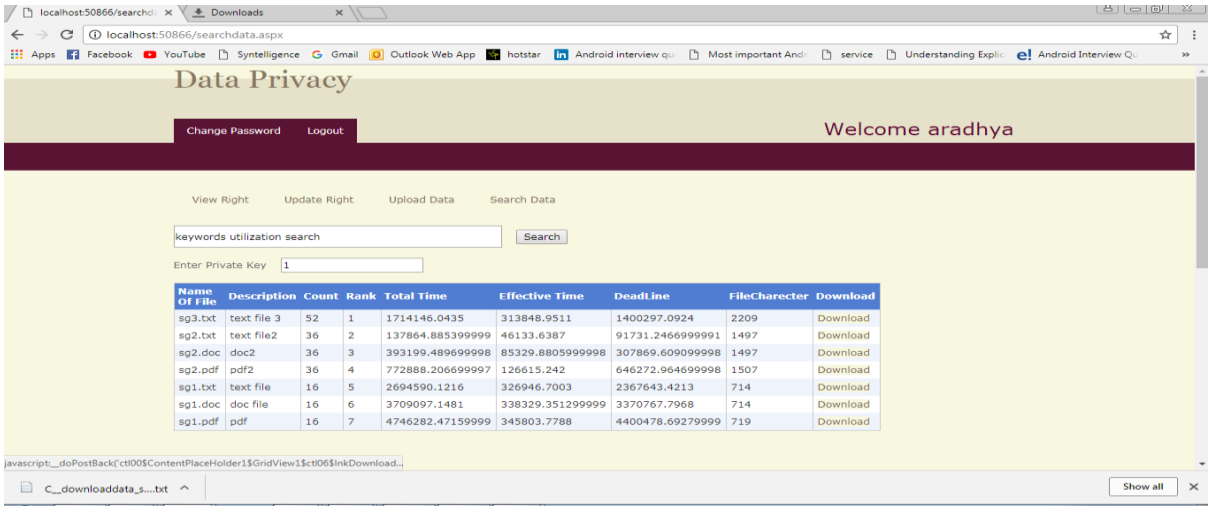


Figure 6.19 GUI for Download module

In the figure 6.19 GUI for Downloading Module shows, downloading modules that is using to download a file that user required. User decided by using rank which one is the most appropriate for these search.

© 2017, IJCSMC All Rights Reserved

VII. EXPERIMENTAL RESULT AND ANALYSIS

7.1 INTRODUCTION

Cloud Computing used for security purpose for that security of data is more important. For more security of data there are some algorithm used for encryption and decryption. We have sent all files to cloud for more security. In our system user upload files to the cloud and by using multi keywords user search a file after that user can download a file.

7.2 RESULT

When users upload a file then a file will be stored in the cloud. when a file stored in the cloud then a file is not stored as a plaintext for security purpose because user should not display inside data of a file so user need to try secure its file. So user need some algorithm that does not recognized its original file or file characters to unauthenticated user. In our project, we are using Xoring algorithm for encryption. In this algorithm plaintext converted into cipher text. so our first result analysis is based on character comparison between plaintext into cipher text with different extension of a files. There are some experimental results related to the following parameters:

7.2.1 RESULT -1 CHARACTER COMPARISON FOR TEXT FILES

Following table 7.1 gives the comparison for Text files between plaintext and cipher text by using Xoring algorithm. In this when user upload a file and a file stored in the cloud in the form of cipher text at that time size of the file will be change. Therefore, we are comparing them based on the character between plaintext and cipher text. In this we are first compare text files with each other for character comparison, also we can display encryption time that require for encrypting files.

Sr. no	File Names	File size in Bytes	Encryption Time in Ms	Plain Text Character (Pre-Encryption)	Cipher Text Character (Post Encryption)
1	1254.txt	4656	1115.06	4646	18624
2	text.txt	9856	4650.26	9845	39436
3	IJAIR.txt	14517	12589.20	14497	58068
4	4567.txt	19711	19100.092	19692	78844
5	Security.txt	24671	34790.53	24650	98684

Table 7.1: Comparison between Plaintext Character and Cipher text Character

In the table 7.1 shows the Plain text Character i.e. Pre-Encryption in this we can used Xoring encryption algorithm. when user upload a file then it will convert file character plaintext character into cipher text character. So, in this table 7.1 shows that numbers of character converted from plaintext into cipher text.

© 2017, IJCSMC All Rights Reserved

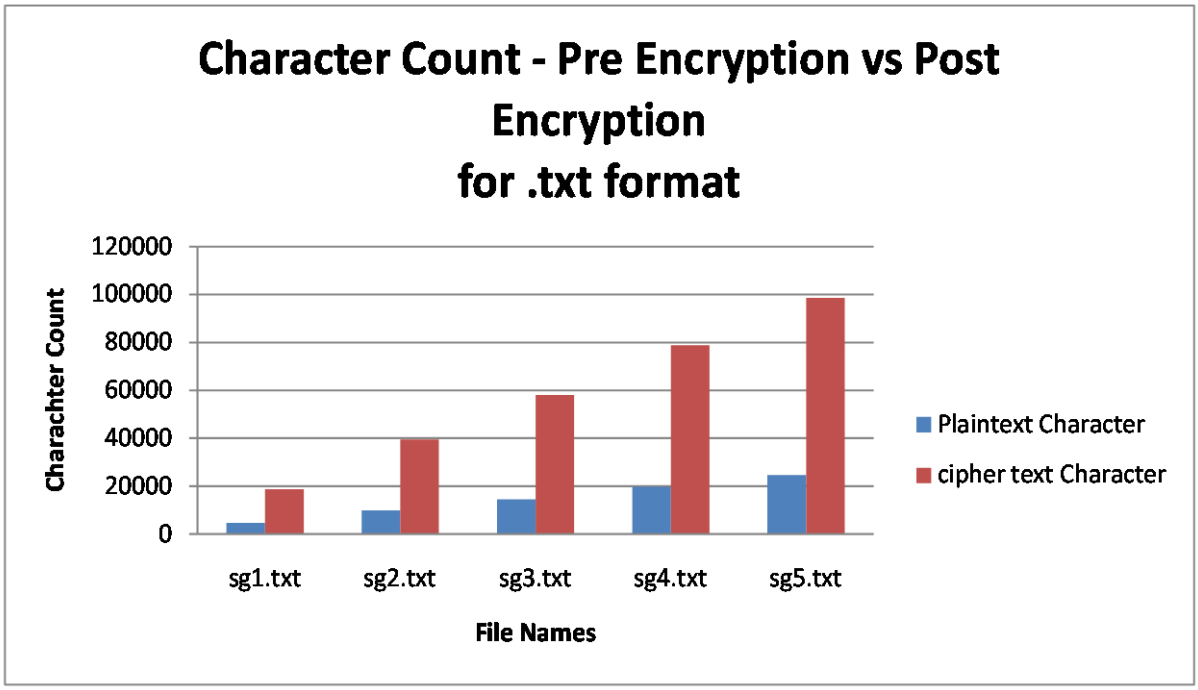


Figure7.1: Comparison between Plain text Character and cipher text Character for text file

Figure 7.1 shows the comparison between Plaintext Character and cipher text character for text files. The graph shows the character count between plaintext and cipher text. In this we can choose comparison between text file with different sizes. We first used plaintext means actual file character and cipher text character means encrypted file by using Xoring algorithm. In the graph blue bar shows Plain text character and red bar shows cipher text character. Cipher text character is more than plaintext character it depends on the algorithm. Y-axis shows character count and X-axis shows name of a files.

7.2.2 RESULT -2 CHARACTER COMPARISONS FOR DOC FILES

Following table 7.2 we can compare for Docs files only this is the comparison between Plain text and cipher text that gives total numbers of plain text character converted into cipher text by using Xoring Algorithm. We can display encryption time that require for encrypting files and also display size of files in Bytes.

Sr. No.	Doc file names	File Size in Bytes	Encryption Time in Ms	Plaintext Character (pre-Encryption)	Cipher Text Character (Post Encryption)
1	Text6.doc	10543	1072	488	1952
2	Management.doc	29696	1707.5	2601	4844
3	14566.doc	33280	2912.3	6486	25908
4	document.doc	36352	6415.3	9813	39240
5	Document47.doc	40448	11818.6	14108	56420

Table 7.2: Comparison between Plaintext Character and Cipher text Character

In the table 7.2 shows the Plain text Characters i.e. Pre-Encryption for doc files. When user upload a file then it will convert a file character into a cipher text character by using encryption algorithm. In this table 7.2 shows that numbers of character converted from plaintext character into cipher text character.

© 2017, IJCSMC All Rights Reserved

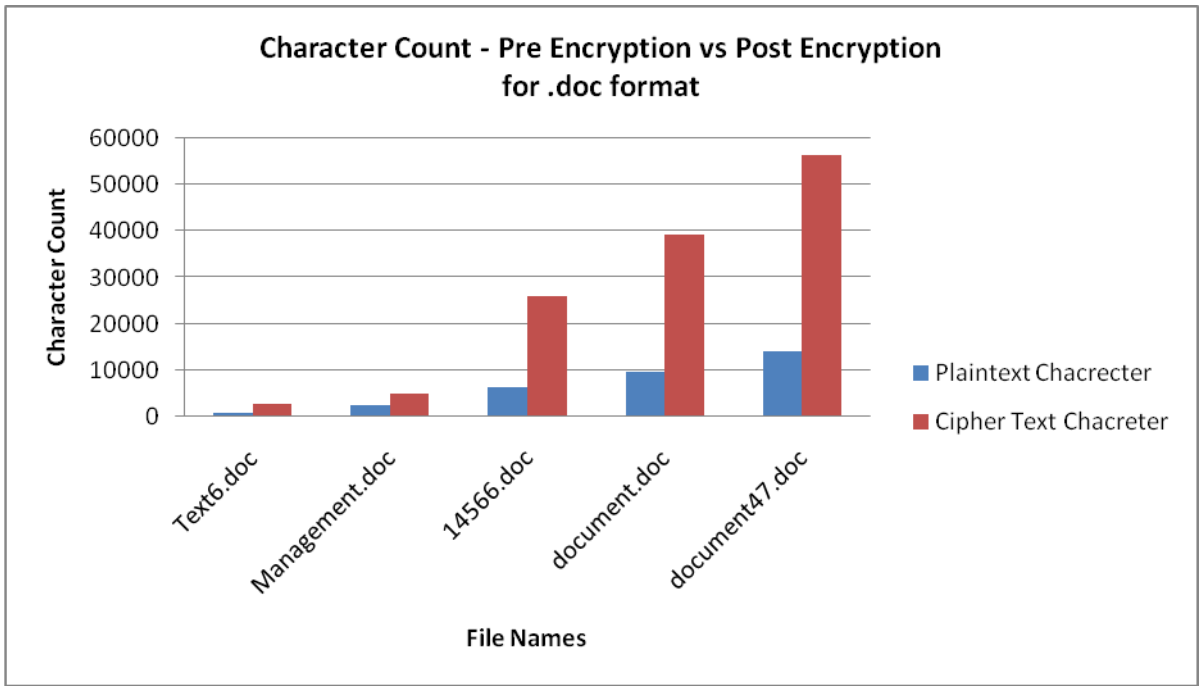


Figure7.2: Comparison between Plain text Character and cipher text Character for Doc Files

Figure 7.2 shows the comparison between Plaintext Character and cipher text character for Docs files. This graph shows the character comparison between plaintext and cipher text. We can used plaintext a files means file that cannot be encrypted means original file and we can choose a files that encrypted means file that should be encrypted by using encryption algorithm. In this graph, we can show that Post-encrypted file character is more than the Pre-encrypted files. In this graph Y-axis shows character count and X-axis shows name of a files.

7.2.3 RESULT -2 CHARACTER COMPARISONS FOR PDF FILES

Following table 7.3 in this we can compare for PDF files only this is the comparison gives the difference between plain text character and cipher text by using Xoring Algorithm.

Sr. No.	PDF file names	File Size in Bytes	Encryption Time in Ms	Plaintext Character (Pre-Encryption)	Cipher Text Character (Post Encryption)
1	document.pdf	118272	2022.11	93	376
2	cloud 2567.pdf	23320	51.002	510	2048
3	87564.pdf	32028	235.003	547	2196
4	pdf4567.pdf	44067	107.50	995	3984
5	document1.pdf	67775	1340.02	4645	18608

Table 7.3: Comparison between Plaintext Character and Cipher text Character for PDF files

In the table 7.3 shows converted character from plaintext into cipher text for Pdf files. In this we can also shows the file size in byte and encryption time that required encrypting a file in a millisecond.

© 2017, IJCSMC All Rights Reserved

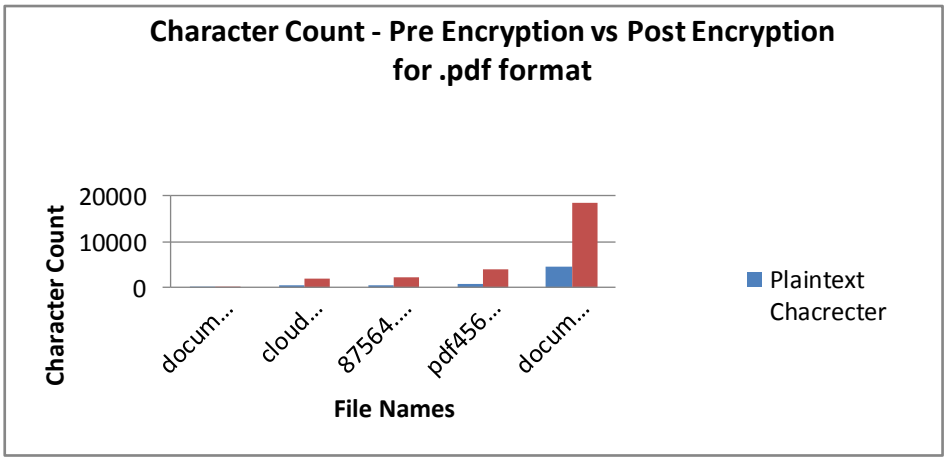


Figure7.3: Comparison between Plain text Character and cipher text Character for Pdf Files

Figure 7.3 shows the comparison between Plaintext Character and cipher text character for pdf files. This graph shows the character comparison between plaintext and cipher text. In this graph, we can show that Post-encrypted file character is more than the Pre-encrypted files. In this graph, Y-axis shows character count and X-axis shows name of a file

7.3 CHARACTER COMPARISONS FOR ALL FILES

Following table 7.4 in this we compare character count for Pre-encryption character and Post encryption character with different types of files.

File Name	**Pre-Encryption** character count	**Post-Encryption** character Count		
	Plain Text	.txt	.doc	.pdf
sg1	4656	18624	18584	186.08
sg2	9845	39436	39340	39388
sg3	14499	58068	57924	57996
sg4	19692	78844	78740	78792

Table 7.4: Comparison between Plaintext Character and Cipher text Character count for all files

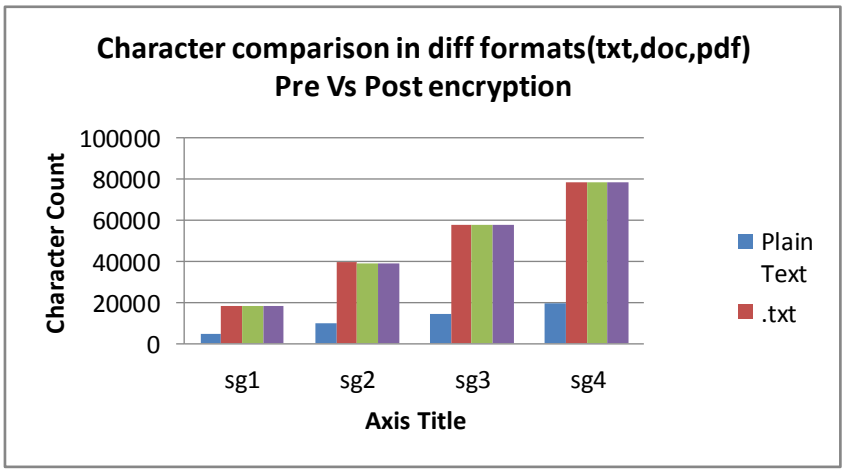


Figure7.4: Character Count Comparison

© 2017, IJCSMC All Rights Reserved

Figure 7.4 we shows the character count between Pre encryption and Post encryption for different extension of files. In these graph we can take fixed size of character files with different extension we can then encrypted that file. Based on the character count in the post encryption file we can compare both the files i.e. Pre and Post encrypted file. By the observation of the above figure 7.4, we can conclude that when we encrypt file with same character with different extension there would not contain so much difference between that Post encrypted files i.e. almost same. We can said that xoring provide any difference for different extension in this we see that character count almost same for different extension

7.4 KEYWORDS SEARCH FOR ALL FILES

In this table 7.5 we can show that time calculation for searching single to multiple keywords. In this we can take fixed character files size with different extension like .txt,.pdf,.rtf,.docx etc. when we upload that file after that we can search a file by typing keywords so in this we can first enter one keywords for .pdf files then calculate time for searching one keywords. After that, we can enter two keywords for the same extension calculates time for searching two keywords and so on.

Data Files	File Size in Bytes	Time for **One** Keyword search(ms)	Time for **Two** Keyword search(ms)	Time for **Three** Keyword search(ms)	Time for **Four** Keyword search(ms)	Time for **Five** Keyword search(ms)
1211.pdf	45660	225.012	223.012	170.009	175.009	281.016
1211.txt	1225	320.018	257.014	181.01	269.15	279.016
1211.docx	10930	1170.104	870.05	1139	1188.00	1203
1211.rtf	45660	2000.5	1949.008	2189	2400..9	2536.99

Table 7.5: Time calculation for searching one to multiple keywords

In the above table 7.5 shows time calculation for searching single to multiple keywords. We can calculate time for searching one to five keywords. In these firstly we can calculate searching time for .pdf file, secondly .txt. thirdly .docx file and finally we can calculate searching time for .rtf file.

Based on table 7.5 we can generate graph as shown in figure 7.5 calculate time for searching of multiple keywords. In this figure we can see the comparison for time with different extension. Searching time will be calculating by using count occurrence means how many times keywords occurs in a document according to that by using query searching time will be calculating.

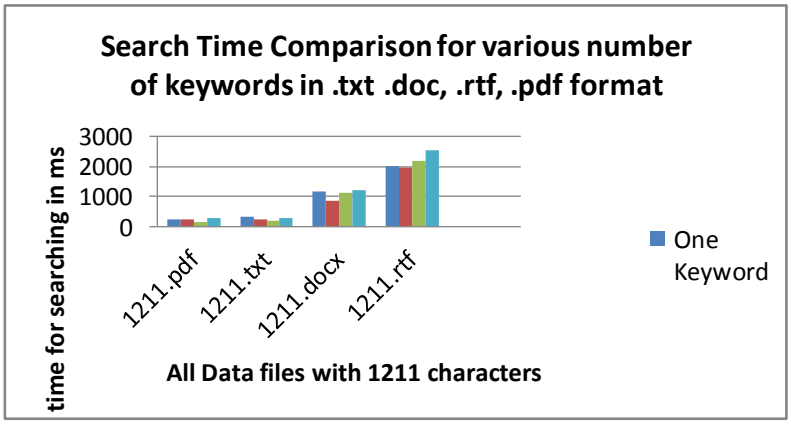


Figure7.5: Calculate time for Searching Multiple Keywords

© 2017, IJCSMC All Rights Reserved

In these, we can show that .pdf file require less time as compare to other file extension and .rtf file required higher searching time as compare to other extension. In searching time will be calculating in ms and all data files contains fixed character. In this blue line display one keywords searching time, red line display two keywords searching time, green line for three and blue line for five keywords.

7.5 Comparison with Existing System

In this, we will discuss about comparison between existing system and proposed system. New method that we proposed in our system is more efficient than existing system. Following are the few points of comparison between existing and proposed system.

Sr No.	Contents	Existing System	Proposed system
1	Algorithm used	Greedy Algorithm	Top-K count occurrence method.
2	Time Constraint	Only total time calculate	Effective time, total time & deadline calculate
3	Time Required for Searching	Greater	Less
4	For Probability computation	Separate Stack is use	Only database used
5	Answer Produce	Produced diversified answer	Produce unique answer
6	Generating Graph	More time required for graph generation	Less time required for graph generation
7	Searching efficiency	Medium	High

Table 7.6: Comparison for different parameter

Table shows the comparison between existing system and proposed system. In the existing system Greedy algorithm is used for searching and in our proposed system we used Top-K algorithm. In the existing system only total time will be calculated that's why time required for searching will be more as compared to our proposed system, because in the proposed system we calculate both total and throughput time means the actual searching time. For calculating the probability of searching, existing system used separate stack and in our system directly database is used. In the existing system diversified i.e. duplicate answers are produce and in our proposed system unique answers are produced. In the existing system more time required for graph generation and in the proposed system instantly graph generated. And efficiency of searching keywords in existing system will be medium and keyword searching efficiency of proposed system is more.

Now we compare the time required for keyword searching for in that greedy algorithm [1] and proposed system in that Top-K algorithm. For this, we have some timing for one keyword, two keyword upto the five keyword, and mention in below table.

© 2017, IJCSMC All Rights Reserved

Number of keywords	Time for Existing system(ms)	Time for Proposed system(ms)
1	1000.1	920.4
2	3924	2163.602
3	5009.23	3800.34
4	6830.4	4213.202
5	7092.36	5136.12

Table 7.7 Comparison between total times with existing system

Table shows the time searching for different input keyword in existing and proposed system. In existing system, time required will be greater than proposed system. Therefore, our system is more efficient than existing system. Graphical representation of this comparison is a given below. The result for greedy algorithm [1] taken from the existing work of the system and next result for proposed system in this we used top-K algorithm.

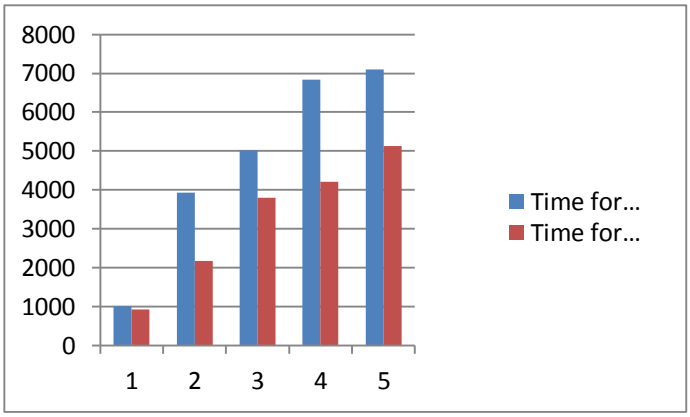


Figure 7.6: comparison graph for time to searching keywords with existing system

Figure 7.6 shows the comparisons graph for searching timing of keywords with existing system and proposed system. Blue bar indicates the time required for existing system and red bar indicates the time required for our proposed system. In the graph x-axis indicates the number of keywords and y-axis shows the time required in seconds. Existing system required more time than proposed system.

7.6 Comparison of Actual Keywords Conversion

We will discuss about actual comparison between two encryption algorithm like Triple DES and XOR algorithm. Now we have taken the original size in byte we can compare this with encrypted files.

File Name	Original Size in Bytes	Post -Triple DES Algorithm in bytes	Post-XOR algorithm in bytes
1675889.docx	9969	20930	10199
Document1.docx	11081	23810	13535.00
IJAIR.docx	11555	24766	12535.00
document45.docx	12435	26882	15735
security7543.docx	13697	29986	18637

Table 7.8: Comparison between Triple Des and XOR algorithm

© 2017, IJCSMC All Rights Reserved

Table 7.8 shows comparison between sizes of files in this we can compare triple DES and Xoring algorithm. Graphical representation of this comparison has given below.

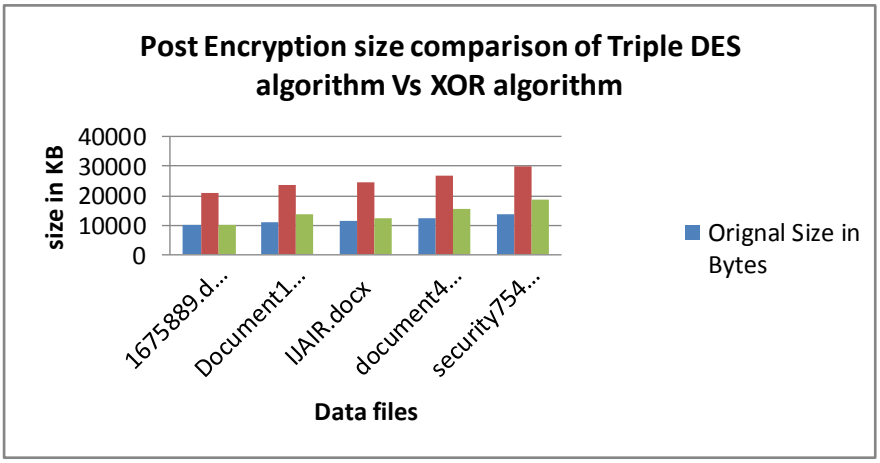


Figure 7.7: Comparison graph for Post-triple DES and Post-XOR algorithm

Figure 7.7 shows the comparisons graph for Post-triple DES and Post-XOR algorithm. Blue bar indicates the original size of file and red bar indicates encrypted size of a file by using Triple DES and green bar indicated encrypted size of file by using xoring algorithm. In the graph x-axis indicates the number of files and y-axis shows the size of a file. Triple DES required more time than XOR algorithm.

VIII. APPLICATIONS

Application of our proposed system is useful in many fields, mention as follow

1. Our system is useful for college administration for storing their data and finding the data easily. Our system stores the data securely because only system admin can access that data.

2. This system is also useful in library management system, for storing and finding the each book and its information easily

3. This system is also useful in bank also for finding customer records easily.

IX. CONCLUSION

This system proposed secured and dynamic multi keywords ranked search scheme over encrypted cloud data. In this system, we can use cloud for storing a file. When we will storing a file then we need to secure our file for this purpose we used XOR algorithm for security. Our system also finds the top-K file with their ranking for input keywords. We also get the effective time, total time and deadline. We have different graph depending on file Pre-character as well as Post-character with the different extension of a file. After that, we can compare search time for different inputs of keywords. We can also compare XOR and triple DES algorithm. Our approach is faster than previous because previous approach first search for unique answers the actual searching will be start that's they required more time. Finally, we show that the a secure data should be stored on the cloud and we can also show the Multi keywords ranked search scheme over encrypted cloud data.

© 2017, IJCSMC All Rights Reserved

X. FUTURE SCOPE

We have implemented our system thus there are many areas where we can apply keywords searching and we can also modified and extended our system. Following are possible areas for future work

KEYWORDS SEARCH BY USING VECTOR SPACE MODEL

We will try to improve the SE scheme to handle these challenges. There are still many challenges in symmetric SE schemes. In the proposed scheme, the data owner is responsible to update information and transfer them to the cloud server. Thus, the data owner needs to store the unencrypted index tree and the information that is necessary to recalculate the IDF values. Such an active data owner may not be very suitable for the cloud computing model. It is a meaningful yet difficult task to design a dynamic searchable encryption scheme where server at the cloud can perform the update operation only. As well, our scheme mainly considers the challenge from the cloud server. Actually, there are many secure challenges in a multi-user scheme.

© 2017, IJCSMC All Rights Reserved