Secure File Storage System on Cloud using Hybrid Cryptography

Submitted by :

SHIVAM MISHRA

(1613101687 / 16SCSE101248)

A Report for the Evaluation 3 of Project 2

In partial fulfillment for
the award of the degree of

Bachelor of Technology

IN

Computer Science and Engineering

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING

Under the Supervision of :

Dr. Ritu sindhu

Professor

April May 2020

Table of content

Secure File Storage System on Cloud using Hybrid Cryptography

UNDER THE GUIDANCE OF:
DR. RITU SINDHU
SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY, GREATER NOIDA
RITUSINDHU@GALGOTIASUNIVERSITY.EDU.IN

SHIVAM MISHRA
SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
GALGOTIAS UNIVERSITY,GREATER NOIDA
RB982684@GMAIL.COM

Abstract—

The Digital Revolution has brought with it an exponential growth in the usage of digital computation and along with it, the start of the Information Era. Across the globe, companies are expanding globally and opening offices at various locations . This has brought the need to make access to data from any location possible and feasible. This is where Cloud Computing and Storage comes into the picture**.** But security risks and data leak possibilities with cloud storage comes. Hence in cloud storage data security is most important. This paper presents a review of a system which stores data on the cloud after encrypting it. Hence even if a security breach were to take place, the attacker would get access to encrypted data, which would still ensure data confidentiality. In this system, an encrypted file is uploaded on the cloud storage system, which results in the decrypted (or original**)** file getting downloaded to their local computer. The system also uses two different hybrid approaches for encryption and decryption, namely AES and RSA algorithms, and AES and Blowfish algorithms, and shows a comparative study on the difference between the two approaches.

Keywords— Cloud Computing and Storage,- AES Algorithm,• RSA Algorithm,• Blowfish Algorithm.

I. INTRODUCTION

Traditional storage devices such as flash drives, hard disks and other kinds of physical storage

devices are slowly becoming obsolete. The reason for this is that, on the business front, global expansion of companies requires data to be shared amongst employees for collaborative working. On the user's personal usage front, many users nowadays have multiple devices, such as one or more mobile/cell phones, tabs, laptops, desktop PCs et cetera. Hence cloud storage provides a way to access one's personal data across all of one's personal devices. Hence more and more people are shifting towards the more convenient option of cloud for storing their data. The ability to access files from remote locations using just a stable internet connection gives cloud an edge over other storage options.

How cloud storage works is that it stores the users' confidential files on the storage servers, and users have the freedom of accessing their files from any location. All of a user's devices such as tablets, laptops, mobile phones, desktop PCs and other technology gadgets can be used to store and access files stored on the cloud. Businesses can also benefit from cloud storage by being able to improve productivity considerably with the help of cloud storage. Cloud storage thus eliminates the need for carrying physical storage devices.

Another advantage of cloud storage is that users can store all kinds of files, such as text documents, images, spreadsheets, videos, PDFs et cetera. Various types of features are provided by different cloud storage providers. Additionally, cloud storage provides a backup option as well. If data on one's local storage gets deleted accidentally, or if one loses the physical storage device such as a hard disk, then one's data can be permanently lost. Also, physical storage devices have a fixed storage capacity, and more the storage capacity, the more it costs. Compatibility or detection issues could possibly arise with physical storage devices. Another issue is that a virus that could inhabit one's computer can move to the flash drive and infect its digital data, or loss due to server failures, employee mistakes, natural disasters are also possible. From the infrastructure point of view, the cost of buying new servers, installing them, and maintaining them is also much higher than the alternative of cloud storage. Buying new servers, installing them, and maintaining them. Additionally, this helps in cutting back on one's energy bill and becoming eco-friendlier.

Cloud storage also helps in immediate data exchange, thus giving access to multiple people. This makes this service a perfect tool for both distant and in-house work. Thus, online cloud storage is beneficial for all types of businesses. Cloud storage is a more cost-efficient platform that does not require a huge investment and it can be actively used for connecting and collaborating with clients and employees. Hence more and more users are turning to cloud storage, making it a very popular alternative to traditional storage options.

II. RELATED WORK

Hybrid Cryptography concept is used for securing storage systems of clouds. Two different approaches are used to show the difference between less secure and more secure systems.

The first approach uses RSA and AES algorithms; RSA is used for key encryption and AES is used for text or data encryption. In the second or we can say more secured approach, AES and Blowfish algorithms are used. In this approach, these two algorithms provide double encryption over data and key which provides high security compared to the first one.

[1]. To make the centralised cloud storage secure ECC(Elliptic Curve Cryptography) algorithm is implemented. This approach uses a single key for encryption and decryption and a complete process takes place at the client side. This methodology performs steps such as: a.Authentication, b.Key generation operation, c.Encryption, d.Decryption.

[2]. In this proposed system a three step procedure is used. Firstly, Diffie Hellman is used for exchanging keys. Thereafter authentication is performed using a digital signature scheme. Finally data is encrypted using AES and then uploaded to the required cloud system. A reverse procedure is implemented.

[3]. Combination of RSA algorithm and MD5 to assure various security measures such as confidentiality, data integrity, non- repudiation etc. It uses RSA key generation algorithm for generation of encrypted key for encryption and decryption process. MD5 digest is used for accepting an input of length up to 128 bit and processing it and generating an output of padded length for encryption and decryption process.

[4]. Implementation of Trusted Storage System using Encrypted File System (EFS) and NTFS file system drive with help of cache manager for securing data files. EFS encrypts stored files by automatically using cryptographic systems. The process takes place as follows, firstly application writes files to NTFS which in turn places it in cache and returns back to NTFS.
After this NTFS asks EFS to encrypt files and heads them towards the disk.

[5]. Cloud Storage Security Service is provided by using separate servers viz. User Input, Data Storage and User Output. Three different servers are used to ensure that failure of any of the servers doesn't harm the data. User Input server is used for storing user files and input data by providing user authentication and making sure the data is not accessed by any of the unauthorized means. Data storage server is the place where the encryption using AES is performed to secure user input and then the encrypted files are transferred to the User Output server. User Output Server is the place from where the user gets the output file or the decrypted file and uses it for further use.

III. PROBLEM FORMULATION AND DESIGN

The many advantages of using cloud storage include:

1. It eliminates the need for carrying physical storage devices.
2. Data in any format can be stored using cloud storage.
3. Cloud storage provides safe backup, as opposed to physical storage devices where loss of device, data corruption by a computer virus, natural disasters, amongst other causes, can lead to loss of data.
4. Cloud storage is more cost-effective as it eliminates the need to invest in hardware,

TABLE I. COMPARATIVE STUDY OF CLOUD STORAGE SYSTEMS

No.	Title	Methodology	Limitations
	Secure storage and access of data in Cloud computing.	ECC (Elliptic Curve Cryptography) algorithm. Performs authentication, key generation, encryption and decryption.	Uses a single key for encryption and decryption hence providing less security.
2	Use of Digital Signature with Diffie Hellman key exchange and AES encryption algorithm to enhance Data Security in Cloud Computing.	Use of Diffie Hellman for key exchange. Authentication provided by Digital Signature scheme and lastly files encrypted using AES.	Time consuming procedure as three different steps using different techniques are performed.
3	RSA Encryption and Digital Signature.	Use of RSA algorithm in combination with MD5 Digest to ensure data security on cloud	RSA algorithm only provides key encryption and along with MD5 it provides single text encryption

The system is designed such that it works in the following way:

1. The user signs in if already registered, or signs up to register themselves by providing their details such as name, email id, phone number, password for account et cetera.
2. The user then selects the file that is to be uploaded by browsing from local storage.
3. The user then selects the encryption algorithm that they want to use. The proposed system provides the choice between using a combination of AES and RSA or AES and Blowfish.
4. The selected file gets uploaded after getting encrypted using the selected encryption algorithm combination.
5. The user also has the option of viewing the files that they have uploaded or have access to and downloading them.
6. On selecting a file to download it, the user is sent the decryption key on their email id that was entered on registration or sign-up.
7. Using this key, the user can download the decrypted or original file.
8. The system also provides a comparison with respect to security between the two hybrids .

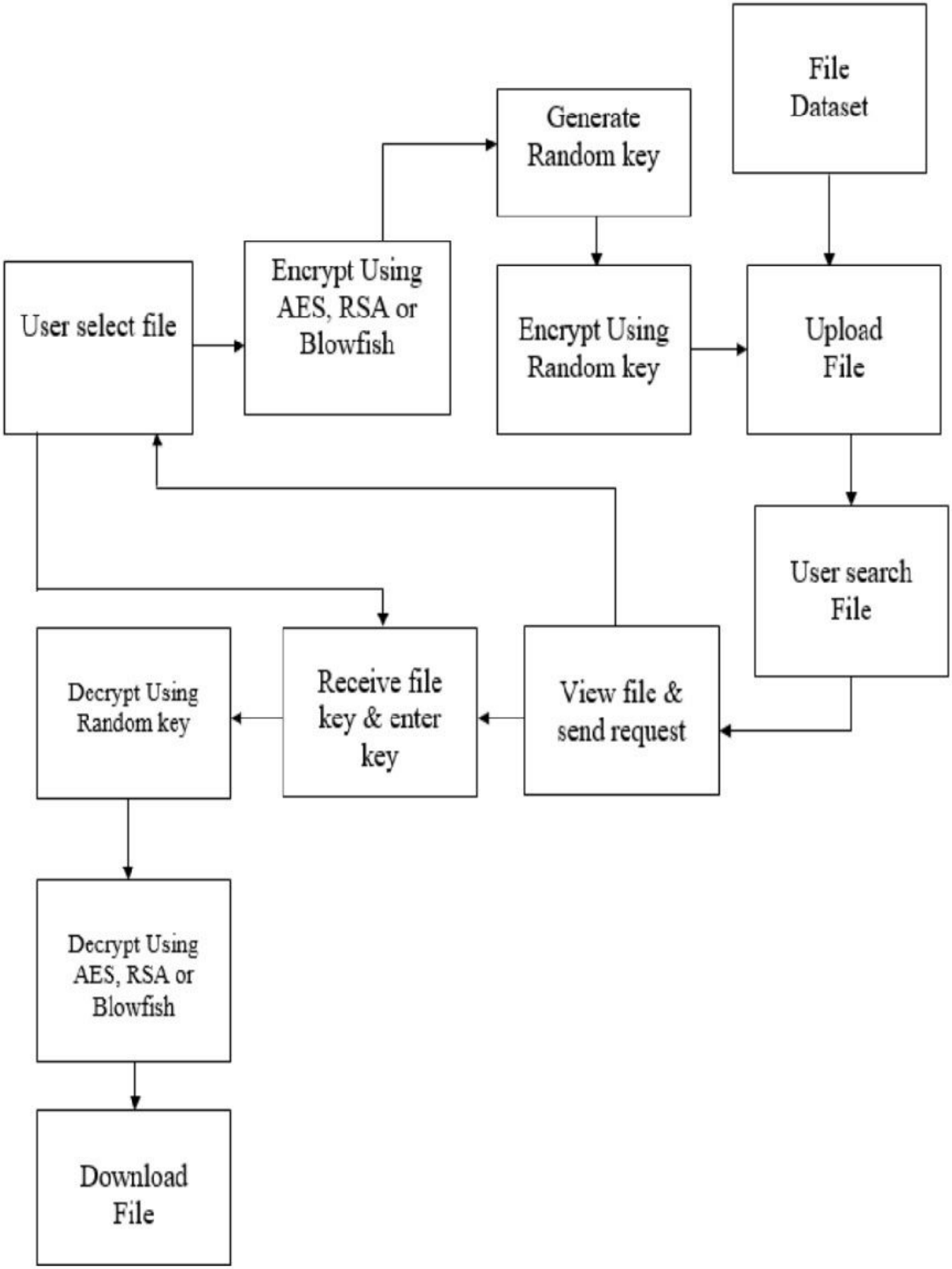


Fig. 1. Block Diagram showing working of system

IV. LEARNING METHODOLOGY

A. AES Algorithm

The Advanced Encryption Standard (AES) also known as 'Rijndael' is a symmetric-key block cipher algorithm having three fixed 128-bit block ciphers with cryptographic key sizes of 128, 192 and 256 bits respectively. The AES algorithm has a maximum block size of 256 bits whereas Key size is unlimited. The AES design is based on a substitution-permutation network (SPN) and does not use the Data Encryption Standard (DES) Feistel network, thus making it stronger and faster than Triple-DES. Step-wise description of the algorithm: Key Expansions: Round keys are derived from the cipher key using AES key schedule, it also requires a separate 128-bit round key block for each round plus one more. Initial Round: Add Round Key - using bitwise xor each byte of the state is combined with a block of the round key.

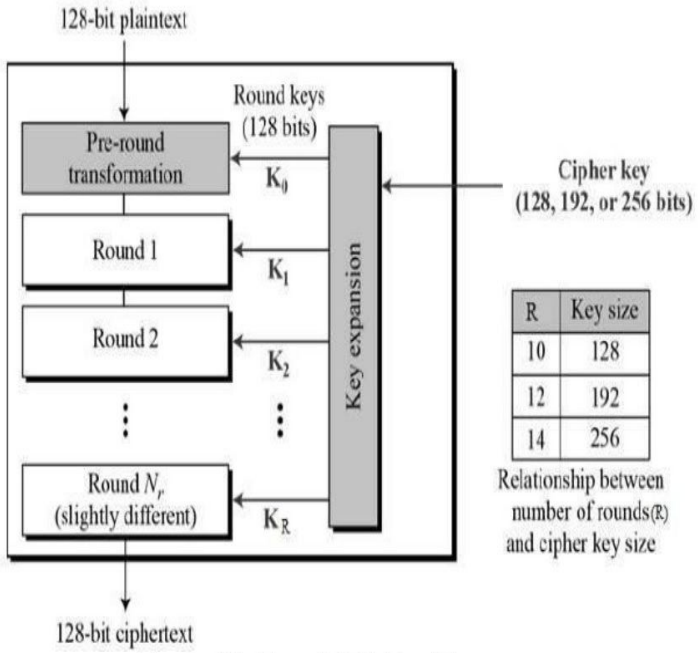
Rounds:

(a) Sub Bytes - according to a lookup table each byte is replaced with another in a non-linear substitution step

.(b) Shift Rows - a transposition step where the last 3 rows of the state are shifted cyclically a certain number of steps.

(c) Mix Columns - a mixing operation which operates on the columns of the state, combining the 4 bytes in each column.

(d) Add Round Key Final Round (no Mix Columns). (a) Sub Bytes (b) Shift Rows (c) Add Round Key



R	Key size
10	128
12	192
14	256

Relationship between number of rounds(R) and cipher key size

Fig. 2. Working of AES Algorithm

The key size used for an AES cipher specifies the number of transformation rounds that convert the input, called the plaintext, into the final output, called the ciphertext. The number of rounds are as follows:

- 10 rounds for 128-bit keys.
- 12 rounds for 192-bit keys.
- 14 rounds for 256-bit keys.

Each round consists of several processing steps, including one that depends on the encryption key itself. A set of reverse rounds are applied to transform the ciphertext back into the original plaintext using the same encryption key.

B. Blowfish Algorithm

Blowfish is a symmetric block encryption algorithm designed which is fast, compact, simple and secure to use as:Blowfish has a 64-bit block size and a variable key length from 32 bits up to 448 bits. It is a 16-round Feistel cipher and uses large key-dependent S-boxes. In structure it resembles CAST-128, which uses fixed S-boxes.

Description of Algorithm:

Blowfish symmetric block cipher algorithm encrypts block data of 64-bits at a time. It will follow the 16 round Feistel network and this algorithm is divided into two parts.

1. Key-expansion
2. Data Encryption

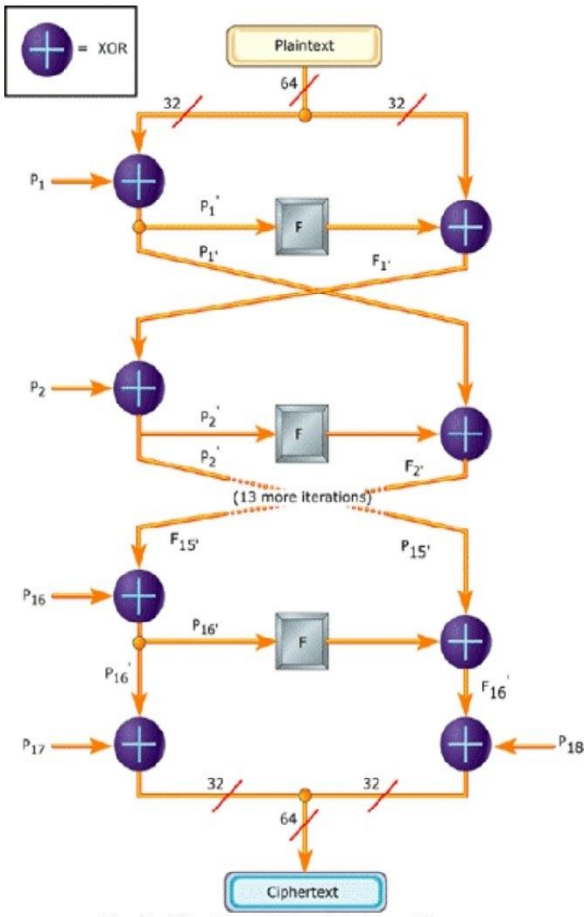


Fig. 3. Working of Blowfish Algorithm

C. RSA Algorithm

RSA (Rivest–Shamir–Adleman) is an algorithm used by modern computers to encrypt and decrypt messages. It is an asymmetric cryptographic algorithm. Asymmetric means that there are two different keys. This is also called public key cryptography, because one of the keys can be given to anyone.

Following is the algorithm using an encryption key as (e,n):

1 Message is represented as an integer between 0 and (n1). Large messages are broken-up into a number of blocks which are then represented by an integer in the same range.

2. Encrypt the message by raising it to the eth power modulo n resulting in a ciphertext message C.

3. To decrypt that message, raise it to another power d modulo n.

The encryption key (e,n) is made public while the decryption key (d,n) is kept private by the user

The Appropriate Values for e, d, and n are determined as follows:

1. Choose two very large (100+ digit) prime numbers represented as p and q.

2. Set n equal to p * q.

3. Choose any large integer d, such that GCD(d, ((p-1) * (q1))) = 1 4. Find e such that e * d = 1 (mod ((p-1) * (q-1)))

BLOWFISH ALGORITHM PSEUDOCODE :

```
uint32_t P[18];
uint32_t S[4][256];

uint32_t f (uint32_t x) {
  uint32_t h = S[0][x >> 24] + S[1][x >> 16 &
0xff];
  return ( h ^ S[2][x >> 8 & 0xff] ) + S[3][x &
0xff];
}

void encrypt (uint32_t & L, uint32_t & R) {
  for (int i=0 ; i<16 ; i += 2) {
    L ^= P[i];
    R ^= f(L);
    R ^= P[i+1];
    L ^= f(R);
  }
  L ^= P[16];
  R ^= P[17];
  swap (L, R);
}

void decrypt (uint32_t & L, uint32_t & R) {
  for (int i=16 ; i > 0 ; i -= 2) {
    L ^= P[i+1];
    R ^= f(L);
    R ^= P[i];
    L ^= f(R);
  }
  L ^= P[1];
  R ^= P[0];
  swap (L, R);
}

{
  for (int i=0 ; i<18 ; ++i)
    P[i] ^= key[i % keylen];
  uint32_t L = 0, R = 0;
  for (int i=0 ; i<18 ; i+=2) {
    encrypt (L, R);
    P[i] = L; P[i+1] = R;
  }
  for (int i=0 ; i<4 ; ++i)
    for (int j=0 ; j<256; j+=2) {
      encrypt (L, R);
      S[i][j] = L; S[i][j+1] = R;
    }
}
```

III. RESULT AND ANALYSIS

OWNCLOUD

Owncloud is a cloud storage system similar to Dropbox. It is free and a Open source software that provides unlimited storage capacity of the disk. Owncloud is hosted in our data center on physical and personal cloud servers.
Owncloud security features ensures that data stored on the Owncloud server remains secure and accessible to the user.

V. **Comparison of OneDrive Google Drive DropBox Mega and OwnCloud[9]**

Service Name	Free storage	OS Supported	Security
OneDrive	15 GB	Windows, Mac, Android, iOS, Blackberry	SSL only
Google Drive	15 GB	Windows, Mac, Android, iOS	SSL/TLS only
Dropbox	2 GB	Windows, Mac, Linux, Android, iOS, Blackberry, Kindle Fire	Secure Sockets Layer (SSL) and AES-256 bit
MEGA	50 GB	Windows Android, BlackBerry OS and iOS	2048-bit private/public
ownCloud	no limits on storage space (except for disk capacity)	Windows, Mac ,OS X , Linux, Android and iOS	TLS,HTTPS,CSR, 2048-bit RSA key, owner should be root and the permissions 640

DESIGN AND ANALYSIS

3.1 DESIGN CONCEPTS

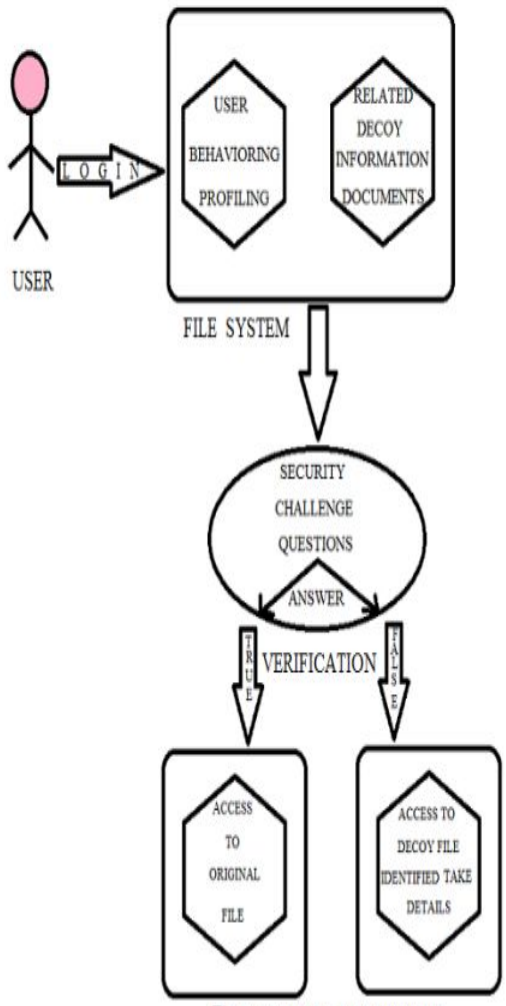
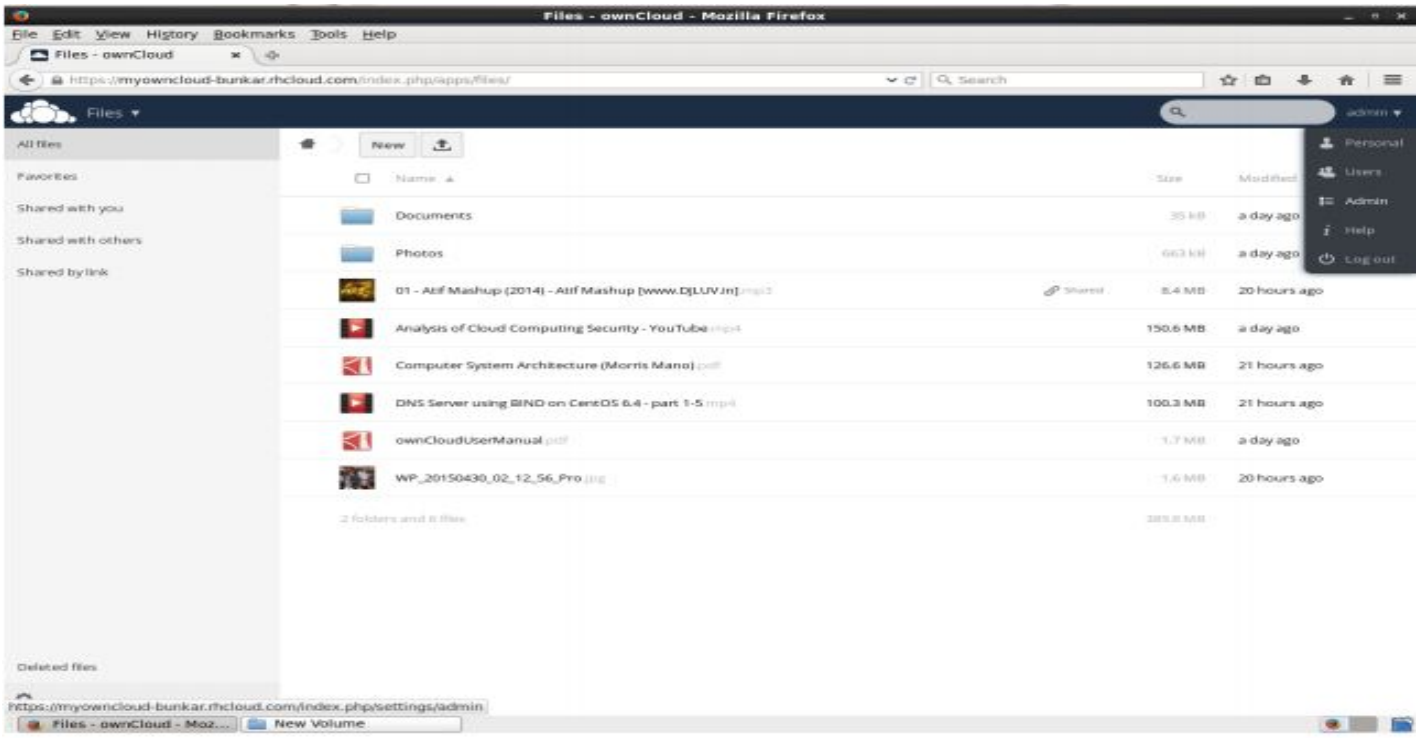


Fig: SECURITY MECHANISM

Cloud Storage			OneDrive	Google Drive	DropBox	Mega	OwnCloud (OpenShift)
File Name	File Type	File Size(MB)	Uploading Time (Seconds)	Uploading Time (Seconds)	Uploading Time (Seconds)	Uploading Time (Seconds)	Uploading Time (Seconds)
WP_20150430_02_12_56_Pro	.jpg	1.6	73	13	10	8	4
Atif Mashup	.mp3	8.4	19	21	36	14	4
Cloud Computing	.mp4	81.7	40	35	73	70	31
DNS Server using BIND on Centos	.mp4	100.3	192	175	75	63	58
Computer System Architecture	.pdf	126.6	121	62	59	58	55

RESULTS DISCUSSION



UPLOADED FILE ON OWNCLOUD

III. CONCLUSION

This project implements a double stage encryption algorithm that provides high security, scalability, confidentiality and the easy accessibility of multimedia content in the cloud. The proposed algorithm is crucial in the second stage, the randomly generated key provides more security than the conventional encryption system. The ciphertext is stored in the cloud instead of original multimedia content. The cipher text is undoubtedly hard to recover the original content for random asymmetric keys. Wide application of the proposed algorithm protects the information from the side channel attacker to grab the multimedia data from the cloud. Thus, the multimedia content is safe in the cloud.Cloud computing is one of the most talked about IT trends today. This is because of the fact that cloud computing has helped several enterprises to save money while adding to the convenience of the users. The word 'Cloud' refers to the widespread internet, which means Cloud Computing is an internet based computing where services are delivered to the users via internet.

IV. REFERENCES

1) Cloud Security Alliance, "Top Threat to Cloud Computing V1.0," March 2010. [Online]. Available: https://cloudsecurityalliance.org/topthreats/csathreats.v1.0.pdf

2) M. Arrington, "In our inbox: Hundreds of confidential twitter documents," July 2009. [Online].Available:https://techcrunch.com/2009/07/14/in-our-inbox-hundreds-of confidential-twitter-documents/

3) D. Takahashi, "French hacker who leaked Twitter documents to TechCrunch is busted," March 2010. [Online].Available: http://venturebeat.com/2010/03/24/frenchhacker-wholeaked-twitter-documents-to-techcrunch-is-busted

4) D. Danchev, "ZDNET: french hacker gains access to twitter's admin panel," April 2009. [Online].Available:http://www.zdnet.com/blog/security/french-hacker-gainsaccess-totwitters-admin-panel/3292

5) P. Allen, "Obama's Twitter password revealed after french hacker arrested for breaking into U.S. president's account," March 2010. [Online].Available: http://www.dailymail.co.uk/news/article-1260488/Barack-Obamas-Twitter-passwordrevealed-French-hacker-arrested.html

6) F. Rocha and M. Correia, "Lucy in the sky without diamonds: Stealing confidential data in the cloud," in Proceedings of the First International Workshop on Dependability of Clouds, Data Centers and Virtual Computing Environments, Hong Kong, ser. DCDV '11, June 2011.

7) M. Van Dijk and A. Juels, "On the impossibility of cryptography alone for privacy preserving cloud computing," in Proceedings of the 5th USENIX conference on Hot topics in security, ser. HotSec'10. Berkeley, CA, USA: USENIX Association, 2010, pp.

1–8. [Online]. Available: http://dl.acm.org/citation.cfm?id=1924931.1924934