

Review of Secure File Storage on Cloud using Hybrid Cryptography

Shruti Kanatt
Department of Computer
Engineering
FCRIT, Vashi
Navi Mumbai, India

Amey Jadhav
Department of Computer
Engineering
FCRIT, Vashi
Navi Mumbai, India

Prachi Talwar
Department of Computer
Engineering
FCRIT, Vashi
Navi Mumbai, India

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. Furthermore, companies are expanding globally and opening offices at various locations across the globe. 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 with cloud storage comes security risks and data leak possibilities. Hence data security is a very important component of cloud storage. 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, the user uploads a file to the portal, it gets encrypted and then uploaded onto the cloud. The user can then download their files from the cloud through the portal, 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 require 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 help 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 and 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 system of cloud. 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 single key for encryption and decryption and 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 three step procedure is used. Firstly, Diffie Hellman is used for exchanging keys. Thereafter authentication is performed using digital signature scheme. Finally data is encrypted using AES and then uploaded to the required cloud system. For decryption 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 in cache and return backs 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 User Output server. User Output Server is the place from where user gets the output file or the decrypted file and use it for further use.

4	Survey Paper on Cloud Storage Security	Using EFS, NTFS with cache for securing data files by using automatic cryptographic systems inbuilt in EFS.	As cryptographic systems are inbuilt in EFS, modifications for providing better security measures is difficult to implement.
5	Secure File Storage and File Sharing.	Separate servers are used for input, storage and output functions. Providing better security by keeping separate modules.	As three different servers are used there can be connectivity issues as well as synchronization problems.

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,
5. Cloud storage also helps developers collaborate and share their work in a more efficient and speedy manner.

Another advantage of cloud storage could be additional security. The proposed system aims to make the cloud storage system secure using data encryption. Thus, the aim of the proposed system is to increase security of data uploaded onto the cloud by using encryption algorithms to make the system more secure.

TABLE I. COMPARITIVE STUDY OF CLOUD STORAGE SYSTEMS

No.	Title	Methodology	Limitations
1	Secure storage and access of data in Cloud computing.	ECC (Elliptic Curve Cryptography) algorithm. Performs authentication, key generation, encryption and decryption.	Uses 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 and not multiple 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 hybrid encryption algorithm combinations i.e. AES and RSA hybrid combination and AES and Blowfish combination.

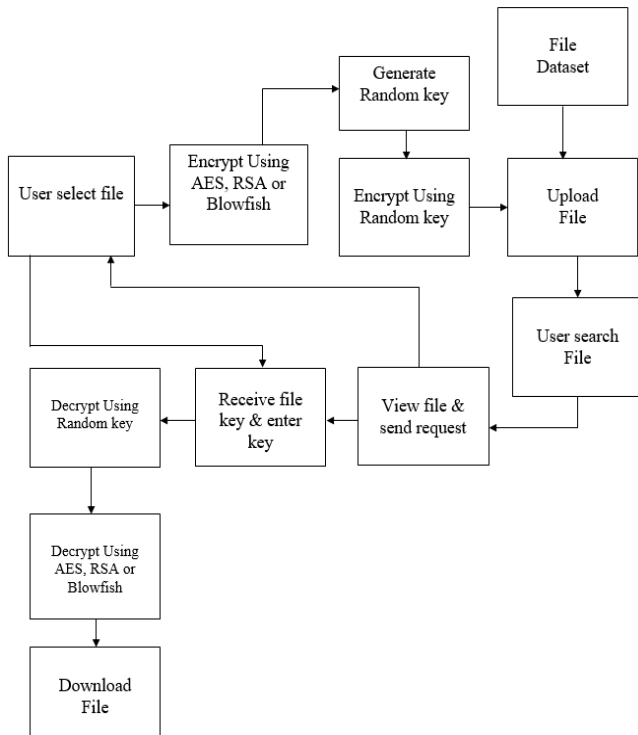


Fig. 1. Block Diagram showing working of system

The system is thus secure, as it provides a double layer of security. Confidential user login credentials are the first layer of security. The second layer is the encrypted file. Since the file is encrypted and then stored on the cloud, even if an attacker gains access to the cloud, they would only have access to the encrypted files. The file can be decrypted using only the decryption key, which is only sent to the user's email id which was entered during registration/sign-up time.

Therefore, the proposed system is designed to provide cloud storage features to users of the portal such as uploading and downloading files to the cloud, wherein the selected files are first encrypted and then uploaded to the file, and can be downloaded using only secret decryption key.

An additional feature is the comparative study between the two hybrid algorithm approaches, namely AES and RSA combination and AES and Blowfish combination.

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 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

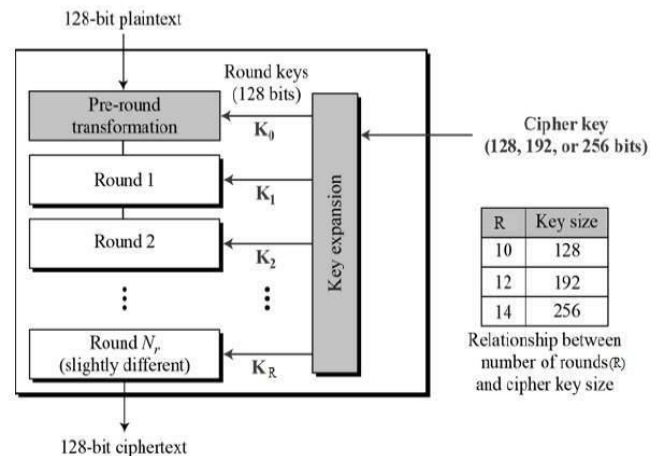


Fig. 2. Working of AES Algorithm

B. Blowfish Algorithm

Blowfish is a symmetric block encryption algorithm designed which is fast, compact, simple and secure to use as:

It encrypts data on large 32-bit microprocessors at a rate of 26 clock cycles per byte and can run in less than 5K of memory. It uses addition, XOR, lookup table with 32-bit operands. Also the key length is variable, it can be in the range of 32-448 bits: default 128 bits key length. It is suitable for applications where the key does not change often, like communication link or an automatic file encryptor. It is unpatented and royalty-free.

Description of Algorithm:

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

1. Key-expansion
2. Data Encryption

Key-expansion:

It will convert a key into several sub key arrays totalling 4168 bytes consisting at most 448 bits. Blowfish uses five subkey-arrays:

- One 18-entry P-array consisting of 32-bit sub keys: P1,P2,.....,P18 and four 256-entry S-boxes of 32-bit each:
 - S1,0, S1,1,..... S1,255
 - S2,0, S2,1,..... S2,255
 - S3,0, S3,1,..... S3,255
 - S4,0, S4,1,.....S4,255

These keys are generated earlier to any data encryption or decryption.

Data Encryption:

It is having a function to iterate 16 times of network. Each round consists of key-dependent permutation and a key and data-dependent substitution. All operations are XORs and additions on 32-bit words. The only additional operations are four indexed array data lookup tables for each round.

Algorithm: Blowfish Encryption

```

Divide x into two 32-bit halves: xL, xR
For i = 1 to 16:
xL = XL XOR Pi
xR = F(xL) XOR xR
Swap XL and xR
Swap XL and xR (Undo the last swap.)
xR = xR XOR P17
xL = xL XOR P18
Recombine xL and xR
    
```

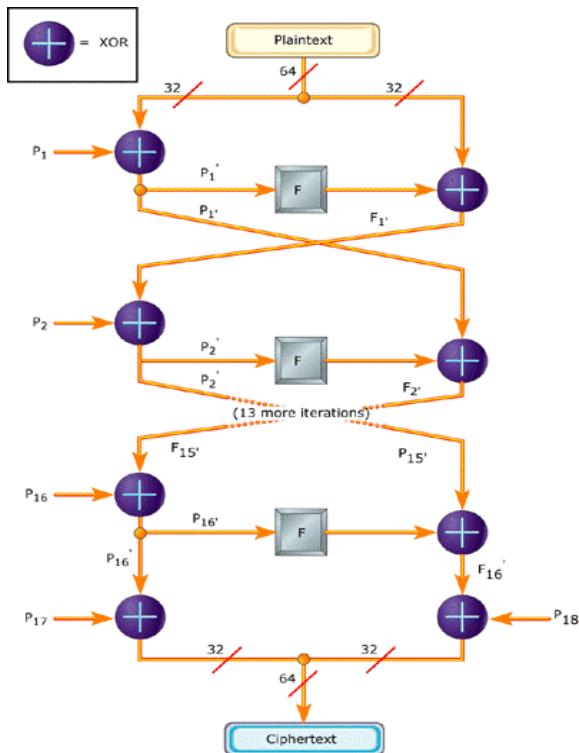


Fig. 3. Working of Blowfish Algorithm

The Rivest-Shamir-Adleman (RSA) algorithm is one of the most popular and secure public-key (asymmetric) cryptographic methods.

Since there is no efficient way to factor very large (100-200 digit) numbers, the algorithm capitalizes on the fact.

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

1. Message is represented as an integer between 0 and (n-1). 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) * (q-1))) = 1
4. Find e such that e * d = 1 (mod ((p-1) * (q-1)))

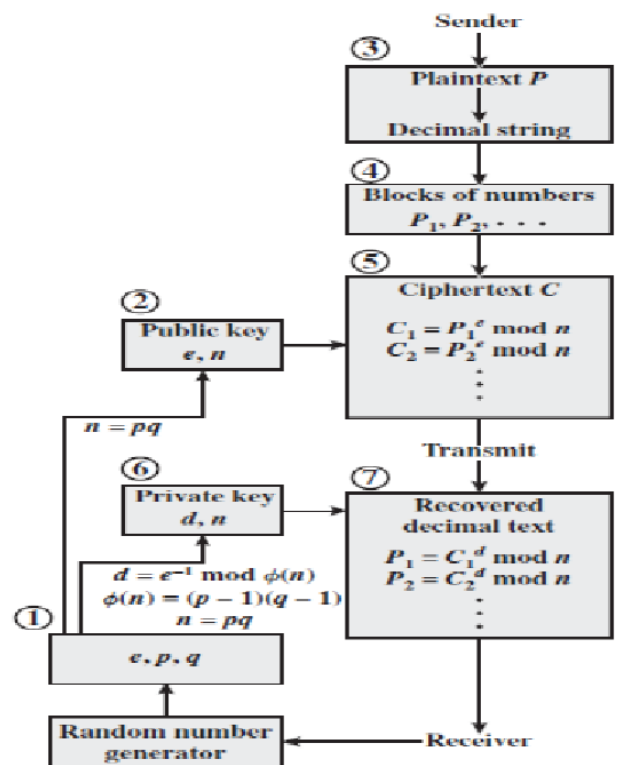


Fig. 4. Working of RSA Algorithm

V. 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 key. 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.

REFERENCES

- [1] Kumar, A., Lee, B. G., Lee, H., & Kumari, A. (2012). Secure storage and access of data in cloud computing. 2012 International Conference on ICT Convergence (ICTC).
- [2] Rewagad, P., & Pawar, Y. (2013). Use of Digital Signature with Diffie Hellman Key Exchange and AES Encryption Algorithm to Enhance Data Security in Cloud Computing. 2013 International Conference on Communication Systems and Network Technologies.
- [3] Ping, Z. L., Liang, S. Q., & Liang, L. X. (2011). RSA Encryption and Digital Signature. 2011 International Conference on Computational and Information Sciences.
- [4] Sunita Sharma, Amit Chugh: 'Suvey Paper on Cloud Storage Security'.
- [5] Rawal, B. S., & Vivek, S. S. (2017). Secure Cloud Storage and File Sharing. 2017 IEEE International Conference on Smart Cloud (SmartCloud).