International Journal of Scientific Research and Engineering Development-– Volume X Issue X, Year

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Data Security And Duplication Checker

Riddhi Parmar*, Arpan Kanani**

*(Information Technology, Charusat University Of Science And Technology, Anand, India
Email: riddhi2000par@gmail.com)

** (Information Technology, Charusat University Of Science And Technology, Anand, India
Email: arpanpatel090@gmail.com)

----------**********----------

Abstract:

In today's world for data storage cloud servers are preferred by most of the clients. So, it is arduous to provide absolute security and with no similarity to the client data. Therefore, for client to share their confidential data to untrusted cloud service providers is dreadful. we have proposed a model that uses cross-bred encryption algorithms to ensure the security for clients and md5 checksum for measuring the similarity of uploaded data for an efficient use case.

Keywords — **Hybrid cryptography, Flask, SQL server, AES, DES, FERNET, MD5 checksum.**

----------**********----------

I. INTRODUCTION

Due to the rapid increase in technology gigantic amount of data is created every day. As we know that this vast amount is shared among various cloud servers. So, data confidentiality becomes crucial as it is necessary to maintain privacy to client's data. Secondly, another important service is integrity if the servers cannot handle assault which is intentional or not but data must be protected from such activities and lastly available. It means that the user or client could access it data from anywhere at any time. This paper of the given encryption algorithm model is to provide a base to all the security problems faced in the cloud services. In our archetype model, we have divided the stipulated input file into three separate part and stored them as individual and different encryption algorithm (i.e. AES, TRIPLE DES, FERNET) is used to encrypt different parts which makes it more arduous for the assaulter to manipulate the data making it much secure.

II. LITERATURE

A. Related Work

In the proposed model symmetric key cryptography with steganography is used. Data encryption is done with the help of four different algorithms named: AES, blowfish, RC6, and BRA, which provides block-level security. Multithreading options are attempted for selection of algorithms; however, all algorithm key size is 128 bits. Lastly, LSB techniques are implemented for key information security.[1]

Paper scripted by V. S. Mahalle and A. K. Shahade at International Conference on Power, Automation and Communication (INPAC), named "Enhancing the data security in Cloud by implementing hybrid (Rsa & Aes) encryption algorithm," uses two different algorithms and varied combination

of three keys is used in encryption and decryption. One key is made public and one is the private key kept with a user. The proposed model facilitates secure upload and downloads with two keys.[2]

The work by Padmavathi, B. and S. R. Kumari named "A Survey on Performance Analysis of DES, AES and RSA Algorithm along with LSB Substitution Technique." In which they presented a technique to share information over an unsecure channel. Their model is comprised of three encryption techniques and a steganographic algorithm.[3]

The paper named "Use of cryptography in cloud computing" written by Jaber, Aws Naser, and Mohamad Fadli Bin Zolkipli works on the advancement of cryptography in the cloud computing sector. In the aspect of discussing security issues with the help of cryptographic algorithms in cloud computing.[4]

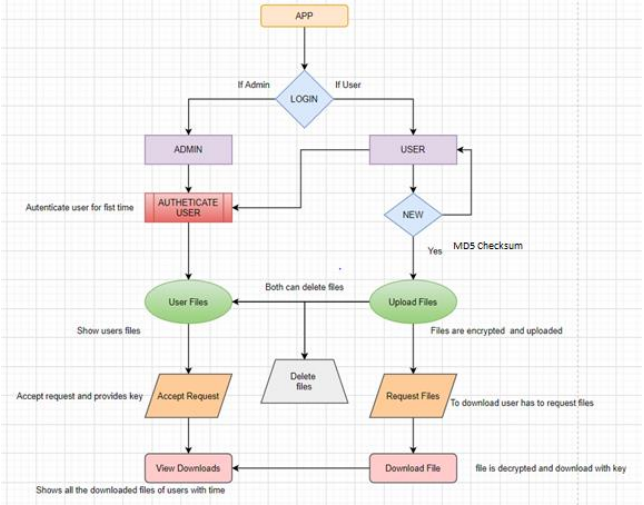
The merger of RSA algorithm and DSA algorithm for data integrity, signature validation, encryption as well as decryption. The functioning of RSA is for encryption and decryption of text however, DSA is also included in the public key cryptosystem and key authorization. Here only a single algorithm, RSA, is used to provide end-to-end security to the user's data.[5]

B. Background Of Invention

The reason behind it is simple yet complex, public-key encryption systems are generally reliable on mathematical terms and therefore prove to be inefficient when compared with symmetric-key encryption systems. Here comes hybrid encryption in the picture which is a bend of the convenience of public-key crypto schemes and efficiency of symmetric key crypto schemes.

III. SYSTEM DESIGN

The system is a combination of four services they are Flask, Encryption mechanism, MD5 checksum, and SQL server. The flask framework is used to make the services user-friendly so that the model looks attractive and convenient. The model is based on the cloud server so all the files and user data and accounts are stored and handled by cloud storage. Now for the security of the files, they are encrypted through a cross-breed algorithm so it becomes arduous to assault files and data remains safe. And to check whether duplicate data or duplicate file is already there or not for that we used MD5 checksum. Now basically lets us try to understand the working of the model with a flowchart given below.



(Figure 1: Activity Diagram For System)

IV. FLASK FRAMEWORK

FLASK, A THIRD-PARTY LIBRARY THAT IS MAINLY USED FOR DEVELOPING WEB APPLICATIONS IN PYTHON, WHICH IS A WELL- KNOWN WEB FRAMEWORK. IT IS THE MOST SUCCESSFUL PYTHON MICRO-FRAMEWORK. AS IT IS A WEB FRAMEWORK IT HELPS US TO BUILD WEB APPLICATIONS. SECONDLY, CONSIDERING IN TERMS OF SPEED AND PERFORMANCE, AND THE FLASK IS PREFERRED OVER DJANGO AS IT DID PROVE MUCH FLEXIBILITY. FINALLY, IT IS USED AT THE BACKEND, HOWEVER,

TEMPLATES OF HTML AND XML ARE ALSO DEVELOPED USING A DIFFERENT TEMPLATING LANGUAGE WHICH IS KNOWN AS JINJA2 WHICH WORKS ON HTTP REQUESTS.

V. ENCRYPTION MECHANISM

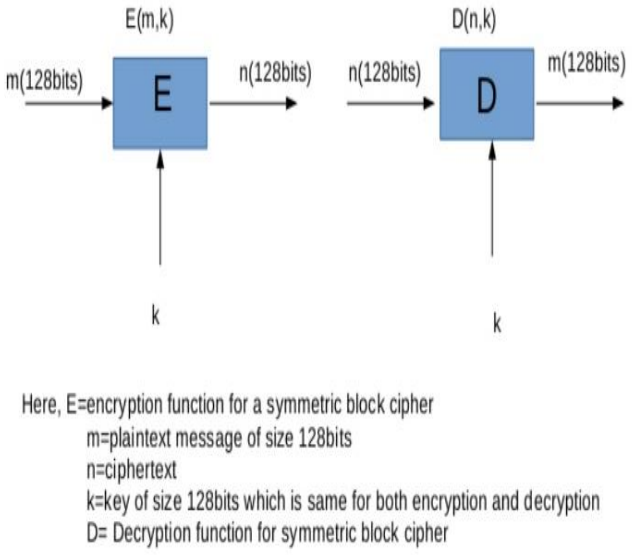
THE MAIN MOTIVE OF OUR PROPOSED MODEL IS TO PROVIDE CERTIFIED SECURITY TO CLOUD STORAGE FOR THAT A HYBRID CRYPTOSYSTEM IS USED. MAINLY ALL THE DATA ARE ENCRYPTED WITH THE HELP OF ENCRYPTION ALGORITHMS THE ENCRYPTION MECHANISM USES THREE ALGORITHMS WITH FILE SPLITTING AND MERGING TECHNIQUES. THE ALGORITHMS ARE:

- AES (Advanced Encryption Standard)
- Triple DEA (Data Encryption Algorithm)
- Fernet

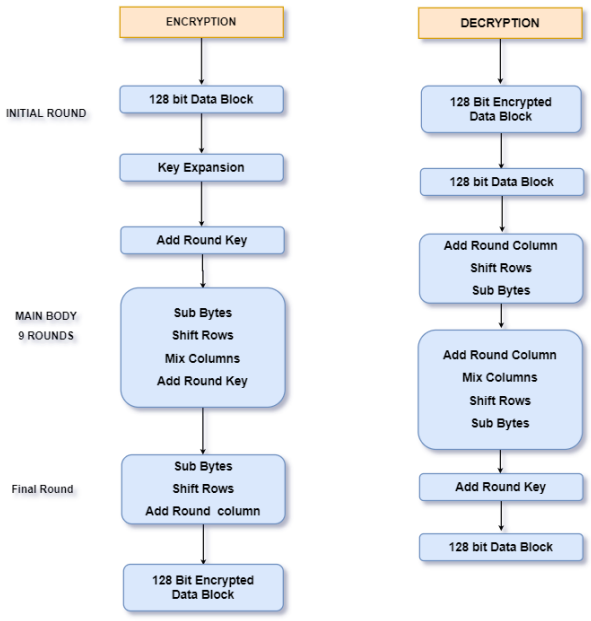
a. Advanced Encryption Standard

AES, which is a member of Rijndael block ciphers, basically this is a symmetric key algorithm. The block size used if 128-bit, but the variations in the length of the key are such as 128-bit key,196 bit, and 256-bit key. The conversion of plain text to cipher text is done with the transformation rounds. Moreover, it depends on the size of the key, 128-bit key consist of 10 rounds whereas 196-bit key has 12 rounds of transformation and 14 for 256-bit key length. There are mainly 4 rounds that are done in the repetitive mode for conversion of text in the AES algorithm. It takes the 128-bit block of data and a key to converting plain text to ciphertext. the functions are:

1) SubBytes
2) ShiftBytes
3) Mixcolumns
4) AddRoundkey



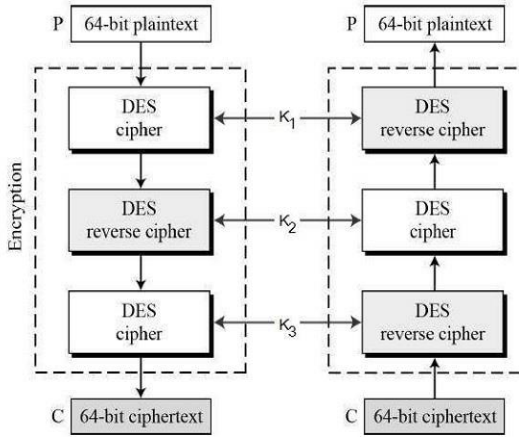
(Figure 2: AES algorithm)



(Figure 3: Steps for Encryption And Decryption)

b. Triple DES

Triples DES (3DES) is termed as Triple Data Encryption Algorithm (TDEA). Mainly in this block cipher in which DES algorithm is applied thrice to each data block uses symmetric keys. The traditional DES algorithm uses the 56-bit key but in 3DES size of the key is increased to three times which increases the security. There are three options available to define the key, we have applied the option in which all the keys are the same and use 56 bits.



(Figure 4: DES Encryption & Decryption)

c. Fernet

Fernet safeguards the message encrypted from it cannot be manipulated or read without the key. Using fernet we can fresh key but we need to keep it with safety anyone who has can easily read and manipulate the data. Similarly, in addition, it generates a URL-safe base64-encoded 32-byte key.

VI. MD5 Checksum

MD5(Message-Digest Algorithm) is a cryptographic hash function. Used as an encryption function for files. MD5 allows you to make 128 bits (32 carriers) "hash" from any thread taken as input, regardless of length (up to 2 ^ 64 bits). The idea behind the algorithm is to take up the random data as an input and generates it into a fixed-sized hash value. The input is in any size or any length while output size is always fixed.

VII. SQL Server

Standard Query Language (SQL) server is under decency software or and Relational Database Management System (RDBMS) tool that carries out SQL query statements. We have created the database with the help of a SQL server in which varied tables have been created such as user, admin, file system, request, etc. And relevant information of the client data is stored in such tables respectively.

VIII. Proposed Cloud Computing Security Architecture

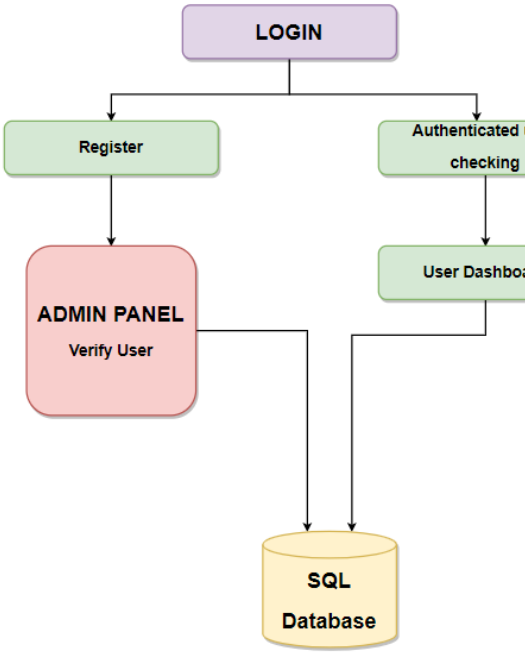
Having a particular goal to ensure endless security to the cloud servers and its data storing systems, the above hybrid encryption system is used. Knowing that the cloud servers are safe from assaults the data which is stored in the cloud storage are in utf-8 encoded format. There are mainly three phases in the cloud:

- Registration Phase
- Uploading Phase
- Downloading Phase

We have utilized MySQL Database for data storage. Cloud computing fulfils the responsibility of secure virtual service with the client's data, software, and computation. Cloud computing is composed of hardware and software resources made available on ethernet as well-handled third-party services. With the MySQL database backend storage of the user and its file is provided.

a. Registration Phase

Firstly, the customer as a new user needs to register themselves into the database to use the services of the proposed model. Secondly, after registration the user needs to be verified by the admin then only, they will able to login into the model. The user details are safe in the database with utf-8 encoding and the pass is encrypted with sha-256.



(Figure 5: Registration Process)

b. Uploading Phase

The process of uploading file is as follows:

1) The user uploads the text file to the SQL server.
2) Checksum key is generated and validated with existing files checksum for data duplication and if no likeness if found further processing is conducted.
3) Then the file is split into 3 different parts.
4) Each part is encrypted using the different algorithm (i.e. AES, TDES, Fernet)
5) After that, the original file parts are deleted and only encrypted parts are stored for the user's security.
6) Keys are stored in encrypting and encoded format into the database.

c. Downloading Phase

The process of downloading file is as follows:

1) The user needs to send a request to the admin to start the downloading process.
2) After the acceptance of request, the user is provided a key through the mail.
3) The user needs to enter the correct key to download the file.
4) After that, the decryption algorithms will decode the text.
5) Lastly, all the parts are merged and the file is downloaded by user attachment.

IX. Conclusion

Understanding that the requirements of the protection of information are necessary for all terms, is achieved through the end goal of the proposed model. One of the benefits of the conveyance model is that each level of authentication is indulged. Hence making it more secure. More acknowledgment is needs to be provided to the version of the proposed model. However, there is needs for security when the file is transmitted to the server from the user's side, which can be achieved by joining public-key cryptography in the upcoming period time.

Table-1

Reference No.	Background of Invention	Technology/ Algorithm	Limitations
1.	Secure file storage in cloud computing using hybrid cryptography.	Uses four different algorithms namely: AES, Blowfish, RC6, and BRA.	It is not designed for general-purpose encryption, and so

			has a very tight limit on the amount of data that can be encrypted.
2.	Enhancing the security of the cloud using hybrid algorithms.	Implementation of AES & RSA	RSA algorithm could be very slow when dealing with an enormous amount of data.
3.	Performance analysis of cryptographic algorithms.	The algorithms used are AES, DES & RSA using LSB substitution.	Message is hard to recover if the image is subject to attacks such as translation and rotation.
4.	The merger of an algorithm for data integrity and security.	Algorithm for encryption of text RSA & for key DSA	Here single algorithm is used and is not feasible for large data.

REFERENCES

[1] [1] P. V. Maitri and A. Verma, "Secure file storage in cloud computing using hybrid cryptography algorithm," 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 2016, pp. 1635-1638, doi: 10.1109/WiSPNET.2016.7566416.

[2] [2] V. S. Mahalle and A. K. Shahade, "Enhancing the data security in Cloud by implementing hybrid (Rsa & Aes) encryption algorithm," 2014 International Conference on Power, Automation and Communication (INPAC), Amravati, India, 2014, pp. 146-149, doi: 10.1109/INPAC.2014.6981152.

[3] [3] B. Padmavathi, S. Ranjitha Kumari, "A Survey on Performance Analysis of DES, AES and RSA Algorithm along with LSB Substitution Technique", International Journal of Science and Research (IJSR), PP.170-174, 2013

[4] [4] Jaber, Aws Naser, and Mohamad Fadli Bin Zolkipli. "Use of cryptography in cloud computing." *2013 IEEE International conference on control system, computing, and Engineering*. IEEE, 2013.

[5] [5] Aufa, F.J., Endroyono, & Affandi, A. (2018). Security System Analysis in Combination Method: RSA Encryption and Digital Signature Algorithm. *2018 4th International Conference on Science and Technology (ICST)*, 1-5.

[6] J. Padhye, V. Firoiu, and D. Towsley, "A stochastic model of TCP Reno congestion avoidance and control," Univ. of Massachusetts, Amherst, MA, CMPSCI Tech. Rep. 99-02, 1999.

[7] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std. 802.11, 1997.

