International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 29 Issue 3 – APRIL 2022.

Secure File Storage on Cloud Using Hybrid Cryptography

M.Akhshaya
Department of Computer Science and engineering
Kalasalingam Academy of Research & Education
Krishnankovil, TamilNadu

J.Jeyaranjani
Assistant professor
Department of Computer Science and engineering
Kalasalingam Academy of Research & Education
Krishnankovil, TamilNadu

Abstract— Cloud computing enables ease of use with simplified user interface. Cloud computing has become the tool of choice for big data processing and analytics due to its reduced cost, broad network access, elasticity, resource pooling and measured service. It also reduces the cost spent in infrastructure. With the cloud computing the budget spent on purchasing and leasing licenses for the software has come down. Cloud computing empowers the initiation of sustainability in terms of green computing, less carbon emission and energy efficiency. At the same time it poses a huge challenge in terms of security. The users can access the cloud with their computing devices – mobile/laptop/PC/Tab. There by the challenge is to provide highly secured cloud services. In this project, we propose secure file storage on cloud using Hybrid cryptography. The scope of the project is to give a secure access, to ensure data integrity and to provide confidentiality. In which we have implemented two algorithms such as SHA algorithm for key generation which will verify user credential and RSA algorithm for achieving the task of encryption and decryption which secures the data in cloud.

Keywords—SHA; RSA; Hybrid Cryptography.

I. INTRODUCTION

Cloud computing has become integral part many a corporate. It has made life easy because it can be accessible by common man by any computing device like mobile/laptop/PC etc. The user data has to stored and retrieved securely. There are many security mechanisms in place. The objective is to propose a secure file storage on cloud using hybrid approach. To give a secure access – the users are required to register themselves with login credentials. The credentials are stored with hash value. The Identity Access Management facilitates the roles for the users and user groups. To ensure data integrity - The moment an user gives value of the credentials, the system compares the hash. If in case an intruder changes the values of the credentials, then the SHA algorithm helps to find the same. To provide confidentiality – the user data stored and retrieved should not be accessed by unauthorized personal. RSA algorithm is used to provide confidentiality. It is important to bring together the inherent features of public clouds and private clouds to build a trustworthy big data processing platform.

II. RELATED WORK

Muhammad et al. proposed a novel framework for the secure and dynamic access of IoT services in a multicloud environment. To facilitate IoT multicloud collaboration, they designed and implemented a novel protocol on a cloud platform. The framework involves various stages, including service matchmaking, authentication, and SLA management. SLA management offers service execution in an external cloud according to the agreed SLA, and it monitors the operation to verify whether or not cloud providers comply with the agreed upon SLA. Fadi et al..

They proposed a seamless and secure key agreement-based framework in cloud-assisted IoT services. A seamless key approach satisfies security properties using a bilinear pairing method and elliptic curve cryptography techniques. In this work, a mobile sink strategy was introduced for user authentication over cloud-based environments. The proposed framework consists of seven phases, namely, initialization, system registration, system login, authentication, extraction of sensitive information, and secret key updating. In the system registration phase, IoT devices register their user ID and password to the TA. In the system login phase, a user needs to enter their valid smart card into the terminal to provide their credentials, such as identity and secret keys. In the authentication phase, the communication with the cluster head is authenticated.

III. LITERATURE SURVEY

The highlighted security issue of cloud is they are stored in a common server where all other data around the world is also stored, and it's not safe to upload the sensitive data in cloud which may cause any issues related to security. And there are also no backup services which most of the cloud service providers provide automatically but others makes as to take out the backup of data. The survey reveals a negative image regarding the security of cloud storage. The survey shows an estimated 61% of SMBs situated across the U.K. and France still believe that their organizational data is unsafe in the Cloud substantial. While 45% contend that migrating their data to the Cloud has compromised their security.

The study by I.S. Decisions exposes further that cloud storage services customers have reservations regarding the

CSPs' ability to detect unauthorized access of their data. The survey reveals that this concern arises from the fact that cloud storage is typically associated with a lack of thorough access control mechanisms that make it challenging to detect persons who misuse employee data to access customer data.

Hybrid cryptography was implemented by combining at least two varying cryptographic algorithms by Information systems security experts . RSA and AES algorithms was used in the first approach, whereas AES and Blowfish algorithms in the second approach. In the first approach, key encryption was done using RSA algorithms, and to encrypt text AES algorithm was used. ARS secret key and RSA public key are required for uploading the data. The files will be stored temporarily in a directory until it's being encrypted. And then RSA algorithm is applied for the encryption of the data. And then AES algorithm is applied to the file. To convert the file into an encoded for ARS key is used.

IV. METHODS

3.1 Existing Method:
Blowfish used for the encryption of file slices takes minimum time and has maximum throughput for encryption and decryption from other symmetric algorithms. The idea of splitting and merging adds on to meet the principle of data security. The disadvantages of Blowfish algorithm are it must get key to the person out of the band specifically not through the unsecured transmission channel. Each pair of users' needs a unique, so as number of the user's increase, KEY MANAGEMENET becomes complicated,
Blowfish algorithm can't provide authentication as well as non REPUDIATION as two people have the same key.
Disadvantages:
1. Computation Overhead.
2. No reliability of security..
3. Time consuming.

3.2 Proposed Method:
The hybrid approach when deployed in cloud environment makes the remote server more secure and thus, helps the cloud providers to fetch more trust of their users. It implements a double stage encryption algorithm that provides high security, scalability, confidentiality and the easy accessibility of multimedia content in the cloud.
SHA most secure public key cryptography algorithm the stored image file is completely secured, as the file is being encrypted not by just using one but two encryption algorithms. The system is very secure and robust in nature. Data is kept secured on cloud server which avoids unauthorized access.
Advantages:
• High Security.
• Less Computation Process.

V. REQUIREMENTS

4.1 Hardware Requirements:
H/W Configuration:
• Processor - Intel i3
• Hard Disk -160GB
• Key Board - Standard Windows Keyboard
• Mouse - Two or Three Button Mouse
• Monitor – SVGA
• RAM - 8GB

4.2 Software Requirements:
S/W Configuration:
• Operating System - Windows 10
• Language - Java 17
• IDE - Net Beans
• Database - My SQL
.

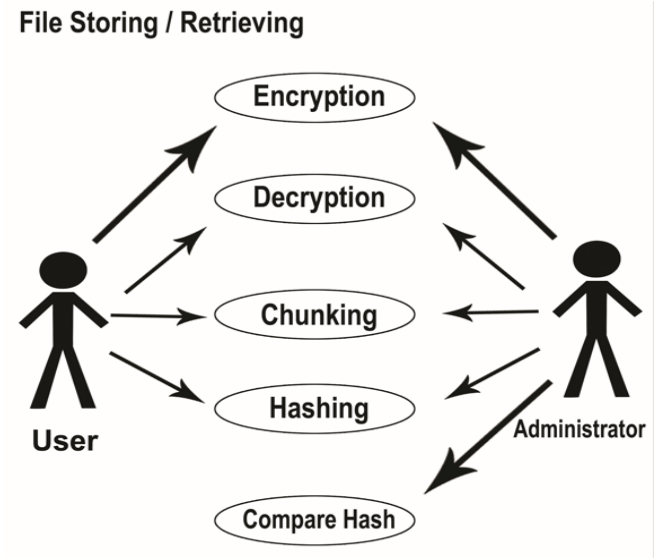
VI. SYSTEM DESIGN

5.1 UML Diagrams:
The condensing "Unified Modeling Language" (UML) means "Brought together Modeling Language." UML is a normalized universally useful demonstrating language utilized in the field of item arranged programming. UML will likely turn into a standard language for making object-situated programming models. UML has two essential parts in its present structure: a meta-model and documentation. Some sort of technique or interaction might be acquainted with, or related with, UML later on. The Unified Modeling Language (UML) is a norm for characterizing, picturing, assembling, and reporting programming framework curios, just as business displaying and other non-programming frameworks. The UML is a bunch of demonstrated designing practices for displaying enormous and complex frameworks. The Unified Modeling Language (UML) is a significant piece of article situated programming advancement and the product improvement process. To communicate the plan of programming projects, the UML basically utilize graphical documentations.

5.1.1 Goals:
Coming up next are the essential points of the UML plan:
1. Give clients a prepared to-utilize visual demonstrating language that permits them to make and share significant models.
2. Give means to expanding and practicing the key ideas.
3. Be indifferent with regards to specific programming dialects or advancement processes.
4. Build up a conventional establishment for grasping the displaying language.
5. Empower the utilization of more elevated level advancement thoughts like joint efforts, structures, examples, and parts.
6. Accumulate a rundown of best practices.

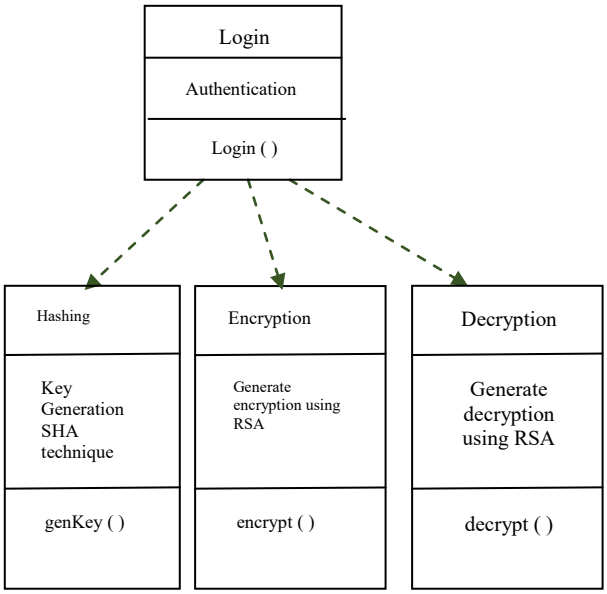
5.2 Use Case Diagram:
A use case diagram is a form of behavioral diagram specified by and derived from a Use-case analysis in the Unified Modeling Language (UML). Its motivation is to give a visual portrayal of a framework's working as far as

entertainers, objectives (addressed as use cases), and any connections between use cases. A use case diagram's principal aim is to indicate which system functions are executed for specific actor. The roles of the system's actors can be shown.



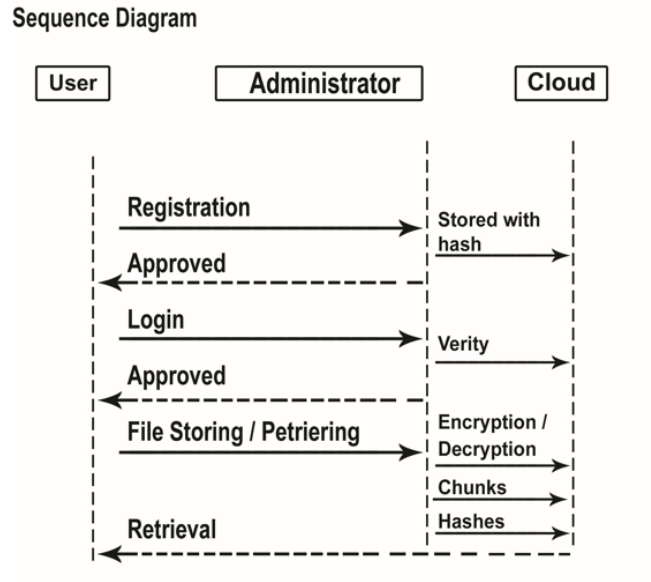
5.3 Class Diagram:

In software engineering, a class diagram in the Unified Modeling Language (UML) displays the structure of a system by presenting the system's classes, characteristics, actions (or methods), and links among the classes. It clarifies which class holds data.



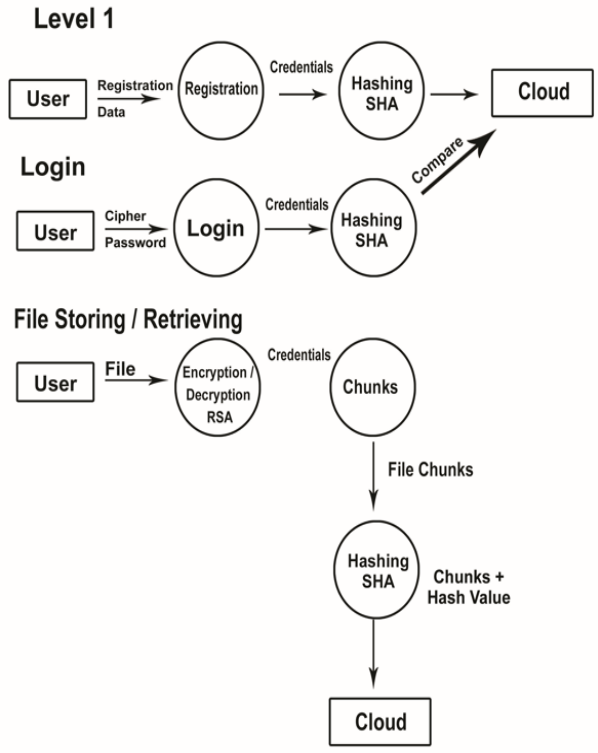
5.4 Sequence Diagram:

A grouping outline is a type of communication graph in the Unified Modeling Language (UML) that shows how cycles interface and in what request. It's a Message Sequence Chart construct.



Arrangement outlines are otherwise called occasion charts, occasion circumstances, and timing graphs.

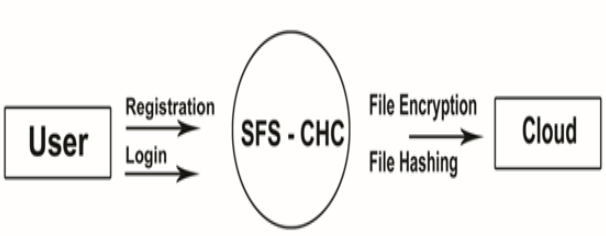
VII. ARCHITECHTURE



VIII. Design Alternatives

6.1 Collaboration Diagram:

The method call sequence is specified in the cooperation diagram via a numbering technique, as illustrated below. The number demonstrates the request where the strategies are called. To depict the cooperation diagram, we used the same order management system. The sequence diagram which indicates the method calls. The cooperation diagram, on the other hand, depicts the object organization, whereas the sequence diagram does not.



IX. Modules

1) Training DataSet

It collects the lot of images with specified angles and dimensions to train the system to understand about the objects and their types. Large scale annotated image dataset ImageNet which contained high resolution images, making it possible to train deep models with large scale training data. It also implements pre-process technique for implementation of image learning for each segmentation. It also verifies the alignment of every image; then only they are eligible for training phase.

2) Analyze

It is used for detection of faces and to reduce the noises present in the complicated movie scenes. It defines various graph editing operations as per the noise analysis and then designs the edit cost function to improve the performance.

3) Testing Dataset

After training and analyzing phase, go to testing phase to test the detection algorithm whether the given image is tested based on previous training working properly or not. we first highlight the importance of learning strategy of detection due to the difficulty of training detectors, and then introduce the optimization techniques for both training and testing stages in detail. Finally, we review some real-world object detection based applications including face detection, pedestrian detection, logo detection and video analysis. During testing, images are resized to different scales followed by multiple detectors and the detection results are merged.

4) Detection

In this module we are going to detect the face of the movie characters. In this module we are using cv library and facenet library. After installing the cv2 and facenet libraries in this project which refers to read the images from specified path, splits the regions and compare them with every trained data which was already implemented in pre-process mechanism. Facenet library is used to load the model from the given directory, get the RGB colors for the divided regions to match them. When the completion of this process it detects the objects which are available the given image or given video.

5) Recognition and result

In this module we are going to recognize the face of the movie characters which is we previously stored on the face database. We just found that the give the real name of it. This is going to be done here. Here we are using deep learning object detection algorithm to implement this process accurately and time cost effective. In fact, instance segmentation can be viewed as a special setting of object detection, where instead of localizing an object by a bounding box, pixel-level localization is desired.

X. Algorithm

9.1 SHA Algorithm:

Hashing is the process of scrambling raw information to the extent that it cannot reproduce it back to its original form. It takes a piece of information and passes it through a function that performs mathematical operations on the plaintext. This function is called the hash function, and the output is called the hash value/digest.

SHA 256 is a part of the SHA 2 family of algorithms, where SHA stands for Secure Hash Algorithm. Published in 2001, it was a joint effort between the NSA and NIST to introduce a successor to the SHA 1 family, which was slowly losing strength against brute force attacks.

Message Length: The length of the clear text should be less than 264 bits. The size needs to be in the comparison area to keep the digest as random as possible.

Digest Length: The length of the hash digest should be 256 bits in SHA 256 algorithm, 512 bits in SHA-512, and so on. Bigger digests usually suggest significantly more calculations at the cost of speed and space.

The RSA algorithm (Rivest-Shamir-Adleman) is the basis of a cryptosystem -- a suite of cryptographic algorithms that are used for specific security services or purposes -- which enables public key encryption and is widely used to secure sensitive data, particularly when it is being sent over an insecure network such as the internet.

RSA derives its security from the difficulty of factoring large integers that are the product of two large prime numbers. Multiplying these two numbers is easy, but determining the original prime numbers from the total -- or factoring -- is considered infeasible due to the time it would take using even today's supercomputers.

The public and private key generation algorithm is the most complex part of RSA cryptography. Two large prime numbers, p and q, are generated using the Rabin-Miller primality test algorithm. A modulus, n, is calculated by multiplying p and q. This number is used by both the public and private keys and provides the link between them. Its length, usually expressed in bits, is called the key length.

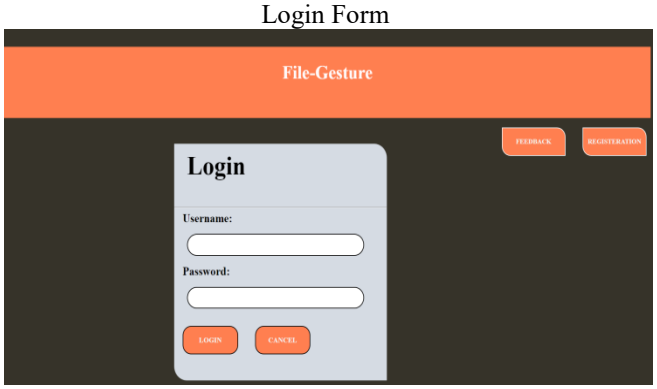
The public key consists of the modulus n and a public exponent, e, which is normally set at 65537, as it's a prime number that is not too large. The e figure doesn't have to be a secretly selected prime number, as the public key is shared with everyone.

9.2 METHODOLOGY:

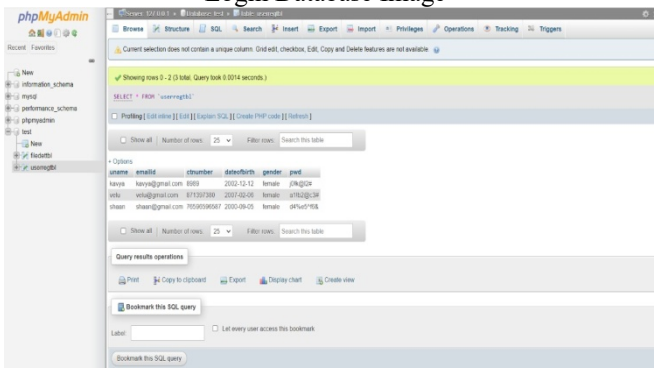
1. Install the required packages
2. Register user details
3. During login, it generates a key.
4. After key is verified, user can access the cloud.
5. Upload file into the cloud using RSA algorithm.
6. Decrypt and download the file using RSA algorithm.

XI. RESULT

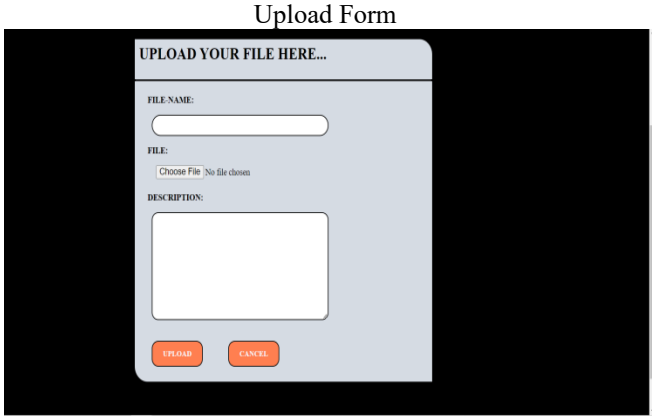
Login Form



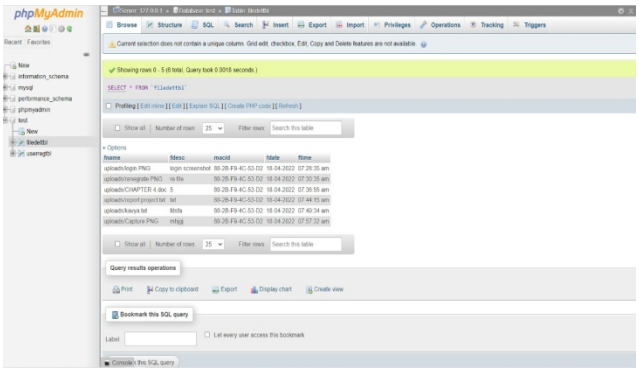
Login Database Image



Upload Form



Upload Database Image



XII. CONCLUSION

Data security is achieved by using SHA algorithm for generating key and RSA algorithm for encrypting and decrypting the data in cloud. Less time is used for encryption, decryption and key generation technique process using hybrid cryptography technique. With help of the proposed mechanism, we have accomplished data integrity, high security, reliability, low delay prompt authentication and confidentiality. The analysis results demonstrate that the proposed mechanism suitable for storing cloud tenants data securely with reduced computation cost and also is effective in data sharing using hybrid cryptography.

XIII. REFERENCES

[1] Oluwasanmi Stephen Alabi , , “Securing File Storage on the Cloud using Hybrid Cryptography”, IEEE, security, 2019

[2] Mazrekaj, A., Shabani, I. and Sejdiu, B., 2016. Pricing schemes in cloud computing: an overview. International Journal of Advanced Computer Science and Applications, 7), pp.80-86

[3] S. Carlin and K. Curran, “Cloud Computing Security”, Pervasive and Ubiquitous Technology Innovations for Ambient Intelligence Environments, vol. 1, no. 2, pp. 12-17, 2013. Available: https://www.igi-global.com/chapter/cloud-computing-security/68920. [Accessed 8 December 2020].

[4] Jyoti, T. and Pandi, G., 2017. Achieving Cloud Security Using Hybrid Cryptography Algorithm. International Journal of Advance Research and Innovative Ideas in Education, 3.

[5] "Cloud Storage Security Issues | A Research Report", IS Decisions, 2020. [Online]. Available: https://www.isdecisions.com/cloud-storage-security-issues/. [Accessed: 08- Dec- 2020].