Securing Data on the Cloud Server by the User Authentication and Data Security Techniques

A. Anusha Priya
Research scholar
Department of Computer Science,
Karpagam University,
Karpagam Academy of Higher Education
Coimbatore.

R. Gunasundari
Associate Professor &Head
Department of Information Technology
Karpagam University,
Karpagam Academy of Higher Education
Coimbatore.

ABSTRACT
Cloud computing is a model which offers a large number of applications under the different topologies. It is the technology of building a strong data security between the Cloud Service Provider (CSP) and User. Authentication is a direct need of an each and every organization and so it is becoming supreme necessary for an organization not because it copes only with the security threats but the reason to develop the policies, procedures and appliances that provide administrative, physical and logical security. Whenever a distinct request to access to a pool of resources, to use them or update them as desired, then to authenticate such an individual is denoted as an authentication. Securing data is always of vital importance because of the serious nature of the Cloud Computing and large amounts of complex data it carries, the need is also equally important. In this paper, the user authentication on the cloud server can be developed by implementing the different types of authentication methods. The three metric based, of bio metric authentication improve the authentication process on the cloud server which protects the data from the unauthorized users. The data in the cloud storage can be protected by implementing data security algorithms such as RSA and AES. The security algorithms can perform very well to protect the data in the cloud server.

Keywords
Cloud server, Authentication, Data security, Bio-metric, AES, RSA.

1. INTRODUCTION

1.1 Cloud Computing
Cloud Computing [1] is a model for enabling the 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. It is the delivery of the computing services over the Internet. Cloud services [2] allow the individuals and businesses to use the software and hardware that are managed by the third parties at remote locations. Examples of the cloud services include online file storage, social networking sites, webmail, and online business applications. The cloud computing model allows the access to give information and to do computer resources from anywhere that a network connection is available. The Cloud Computing [3] provides a shared pool of resources, including data storage space, networks, computer processing power, and specialized corporate and user applications. This cloud model promotes the availability and composed of five essential characteristics, three service models, and four deployment models.

2. OVERVIEW OF THE AUTHENTICATION AND ITS TECHNIQUES ON THE CLOUD SERVER

2.1 Authentication
Authentication [4] [12] means enabling the network to declare the authorized users to access to its resources. It provides the way where the claimed identifier is verified by the access control mechanisms through some means.

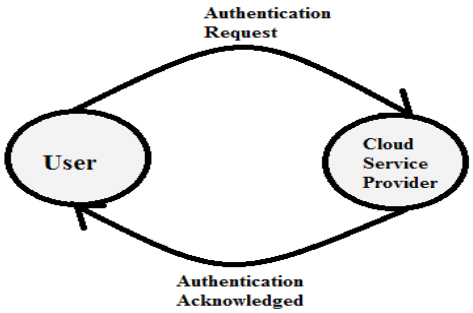

Figure 1. User authentication method

The figure 1 shows that the user requests the user authentication from the cloud service provider and the service provider acknowledge the user authentication. After receiving the authentication from the service provider the user can access the data from the cloud server.

2.2 Authentication Techniques on Cloud
The user authentication involves confirming with a certain degree of confidence that the electronic form of the user's identity represented in the IT System corresponds to the real life identity of the user. There are three factors of user authentication that is used in the combination to increase the level of confidence in the claimed identity of a user.

2.2.1. Single factor / knowledge-based authentication
This type of authentication technique consists of text base that uses the passwords or Personal Identification Numbers (PINs) and graphic based authentication that uses the graphics for the authentication. Knowledge based authentication uses secret information. When the user provides some information to authenticate himself as a legitimate user, the system processes this information and suggests whether the user is legitimate or not. Knowledge based authentication is based on " Something You Know" assumption, in which the user types a password to login to a computer or enters his Personal Identification Number (PIN) to access his/her bank account from an ATM.

The classic form of single factor authentication is called as a User ID and Password. Where the user claims his/her identity by presenting a User ID to the IT access control system. The system then checks the password for the claimed identity against its secure list of known identities and passwords. If the User ID and Password pair entered by the user, match the User ID and password stored in the IT access control system, then the user is judged to be authentic and given access to the system.

2.2.2. Two Factor/Token Based Authentication

This scheme uses some physical items called tokens such as smart cards, passports and physical keys. Authentication token or simply a token may be a physical device that an authorized user of the computer is given to aid in authentication. Such a token may be physically connected or plugged into the client system. The term may refer to software token as well. Hardware tokens are typically small enough to be carried out in a pocket or purse and often are designed to attach to the user's keychain. Some may store the cryptographic keys such as a digital signatures or biometric data such as a fingerprint. Other may include small keypads to allow the entry of a PIN.

Token based authentication [6] is based on "Something You Have "assumption, in which the user carries a wallet full of credentials (a driver's license, credit card, a university ID card) to certify his/her identity (as a driver, as a credit worthy consumer, or as a student). This system uses both forms of authentication. I.e. it involves using "Something You Know" (i.e. a PIN) and "Something You Have" (i.e. a token). Most widely used forms of two factor authentication are.

(i) Automated Teller Machine (ATM) or Cash point Machine Card and PIN.

(ii) Access Control Token and PIN.

At an ATM, the user puts his/her Cash point/ATM card into the ATM and the ATM requests the user to enter his/her PIN. The information held on the magnetic stripe of the card together with the PIN, encrypted in a secure block of data, is sent to the Bank's Central Authentication System, where the PIN entered by the user, is compared with the PIN held on the file against the user's account number and details. However, in this scheme, personally designed unique information is used as a token. Each user is registered against the unique token which becomes his identifying label of the token. Stored information is presented to the system (e.g. ATM card) as well as PIN code to authenticate a user.

2.2.3. Three Factor / Biometric-based authentication

Three factor authentication or Biometric based authentication involves using an access control token such as smart card, a PIN to access the smart card and a biometric value held in the central database. The card is entered into a reader, the PIN is entered, and the biometric is read and encrypted under a cryptographic key held on the smart card. The User ID is read from the smart card together with the encrypted biometric are sent to the central database, where the biometric can be decrypted and compared with the value on the central access control system/database. It is to be noted that the user's PIN is not sent to the central access control system but is checked locally by the smart card. Biometrics is the technologies that analyse the human characteristics for the automated personal authentication. In this scheme, behavioural characters (i.e. voice signature, gait of a human) as well as psychological characters (i.e. fingerprint, hand, iris, retina, face) describing the human characteristics are used for the authentication. Biometric based authentication is used for both the authentications as well as for identification. In short, this system uses some physical or behavioural traits of a human for the authentication.

2.3 Authentication Attacks

Attacks regarding authentication are those which target a web site's method of validating the identity of a user, service or application. These are the following types of attacks.

2.3.1. Brute Force Attack

It is an automated process of trial and error used to guess a person's user name, password, credit card number or cryptographic key. A normal brute force attack uses a single user name against many passwords. A reverse brute force attack uses many user names against a password. When a guessed password allows access to the system, the brute force attack has been successful and the attacker is able to access the account.

Brute Force techniques are highly popular and often successful in the systems with millions of the user accounts.

Example:
Username = Michael
Passwords = Faraday, Jordan, Osterman, [pet names], [birthdays], [car names]
Usernames = Dad, Jon, Barbara, Ed, Sara
Password = 12345678

2.3.2. Insufficient Authentication

This type of attack occurs when a website permits an attacker to access the sensitive content or functionality without having to properly authenticate. Web based administration tools are a good example of web site providing access to sensitive functionality.

Example:
Many web applications are designed with the administrative functionality location directory and off the root directory (admin). This directory is usually never linked anywhere on the web site, but can still be accessed using a standard web browser.

2.3.3. Weak password recovery validation

When a website permits to illegally obtain, change or recover another user's password. Conventional web site authentication methods require users to select and remember a password. The user should be the only person who knows the password and it must be remembered precisely. With the passage of time, a user's ability to remember a password fades. The complication increases further when the average user visits 20 or more sites requiring them to supply a password. A website is considered to have Password Recovery Validation when an attacker is able to foil the recovery mechanism being used. This happens when the information required to validate a user's identity for recovery which is either easily guessed or can be circumvented. Password recovery systems may be compromised through the use of brute force attacks, inherent system weaknesses or easily guessed secret questions.

Example of an automated password recovery processes include requiring the user to answer a "secret question" defined as a part of the user's registration process. The second mechanism in use is having the user to provide a "hint"during the registration that will help the user to remember his password.

Example: (Weak methods of Password Recovery)

3. DATA SECURITY

The data can secure by using the cryptography [9] [10] methods. In that an encryption algorithm is a statistical procedure used to encrypt the data. During the use of an algorithm and a key, in sequence is encoded into cipher text and requires the use of a key to transform the data back into its original form. Algorithms are an essential part of a technology to ensure the effective and secure authentication, as well as to provide reliability and encryption. The figure 2 shows the methodology of the data security in the cloud servers.

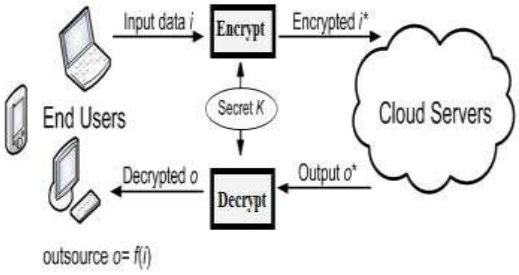

Figure 2. Data security methodology

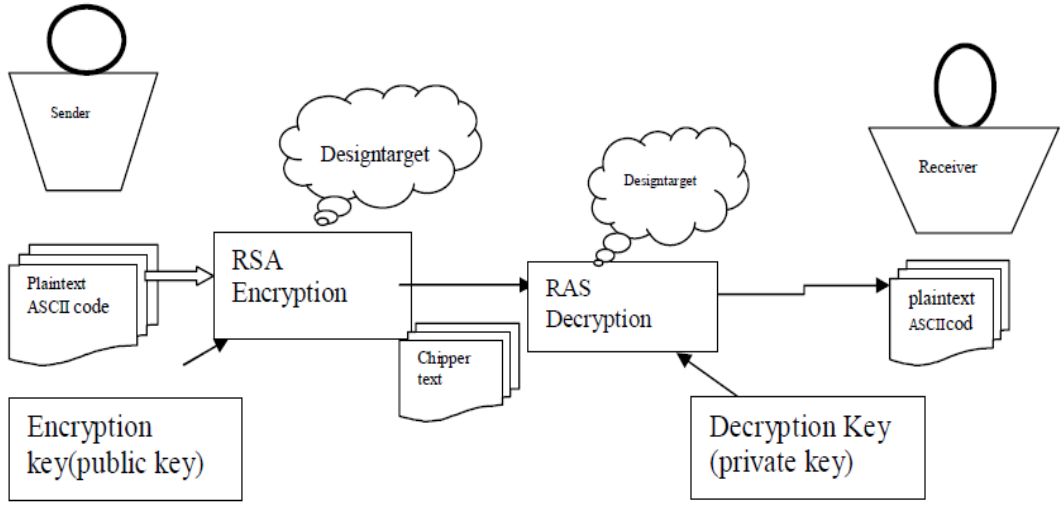

Figure 3. Flow diagram for RSA algorithm

3.1.1 Pseudo code for RSA Algorithm

Step 1. Choosing two very large prime numbers p and q.

Step 2. Compute their system modulus, n =p*q and the 'totient' function φ (n) = (p -1) (q -1). Note that the factors p and q remain secret and n is public.

Step 3. Select the encryption key e at random, so that gcd (e, φ (n)) =1, where 1<e < φ (n).

Step 4. Solve the following equation to find the decryption key d: =1mod φ (n), where $0 \leq d \leq n$.

Step 5. Publish the public encryption key: PU = {e, n}, which is known to everyone.

Step 6. Keep the secret or private decryption key:PR = {d, n}, which is known only to the person who has to decrypt or sign the message.

Types of security algorithms

1. Rivest, Shamir &Adleman (RSA)
2. Advanced Encryption Standard (AES)

3.1 Rivest, Shamir & Adleman (RSA)

RSA [5] security is so much depends on the difficulty of the factoring large prime number. At this time the calculation required to find the factor value to break the key is slow, many cryptanalysts consider that linear attack and differential attack are difficult to applied in RSA, only mathematical attack and brute force are feasible, but the time needed to use these attacks are quite long, therefore they are considered as impractical. As the computer speed increased, this encryption algorithm may consider to be insecure and unused, but do not forget that a person can just simply increase the size of the key to prevent it from the attacker. The common size for the key length now is 1024 bits for P and Q, therefore N is 2048 bits, if the implementation of the RSA is fast enough, the key size can be doubled.

3.2 Advanced Encryption Standard (AES)

Advanced Encryption Standard (AES) algorithm is not only used for the security but also for the great speed. Both hardware and software implementation are faster still. New encryption standard recommended by the NIST to replace the RSA. Encrypts data blocks of 128 bits in 10, 12 and 14 round depending on the key size. It can be implemented on the various platforms especially in the small devices. It is carefully tested for many security applications.

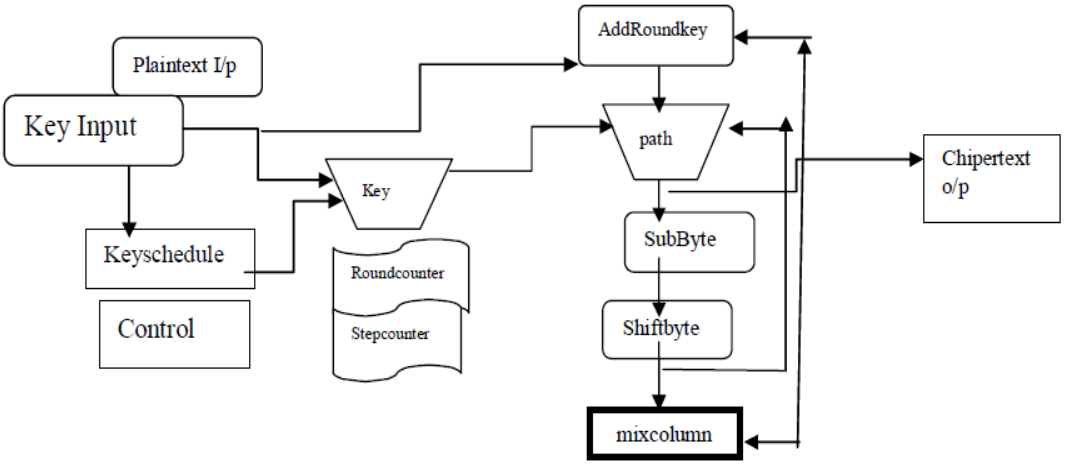

Figure 4. Flow diagram for AES algorithm

3.2.1 Implementation of AES Algorithm

AES is a block cipher with a block length of 128 bits. It allows three different key lengths: 128, 192, or 256 bits. This paper proposes AES with 128 bit key length. The encryption process consists of 10 rounds of processing for 128-bit keys. Except for the last round in each case, all the other rounds are identical. 16 byte encryption key, in the form of 4-byte words is expanded into a key schedule consisting of 44/4-byte words. The 4 x 4 matrix of bytes made from 128-bit input block is referred to as the state array. Before any round-based processing for encryption can begin, input state is XOR with the first four words of the schedule.

Algorithm steps:

These steps used to encrypt 128-bit block

Step 1. The set of round keys from the cipher key.

Step 2. Initialize state array and add the initial round key to the starting state array.

Step 3. Perform round = 1 to 9: Execute Usual Round.

Step 4. Execute Final Round.

Step 5. Corresponding cipher text chunk output of the Final Round Step

Step 6.Usual Round

Execute the following operations which are described above.

1. Sub Bytes
2. Shift Rows
3. Mix Columns
4. Add Round Key, using K(round)

Step 7. Final Round:

Execute the following operations which are described above.

1. Sub Bytes
2. Shift Rows
3. Add Round Key, using K(10)

Step 8.Encryption:

Each round consists of the following four steps:

i. Sub Bytes: The first transformation, Sub Bytes, issued at the encryption site. To substitute a byte, interpret the byte as two hexadecimal digits.

ii. Shift Rows: In the encryption, the transformation is called as Shift Rows.

iii. Mix Columns : The Mix Columns transformation operates at the column level; it transforms each column of the state to a new column.

iv. Add Round Key : Add Round Key proceeds one column at a time. Add Round Key adds a round key word with the each state column matrix. The operation in the Add Round Key is the matrix addition.

The last step consists of XOR the output of the previous three steps with the four words from the key schedule. And the last round for encryption does not involve the "Mix columns" step.

v. Decryption: Decryption involves reversing all the steps taken in the encryption using the inverse functions like,

a) Inverse shift rows, b) Inverse substitute bytes, c) Add round key, and d) Inverse mix columns.

The third step consists of XOR the output of the previous two steps with four the words from the key schedule. And the last round for the decryption does not involve the "Inverse mix columns" step.

3.3 Data security issues in the cloud

3.3.1 Data Availability

Customer data is normally stored in chunk on the different servers often residing in the different locations or in the different Clouds. In this case, data availability becomes a major legitimate issue as the availability of the uninterruptible and seamless provision becomes relatively difficult [11].

3.3.2 Privacy and Confidentiality

Once the client host data to the cloud there should be some guarantee that access to that data will only be limited to the authorized access. Inappropriate access to the customer's sensitive data by the cloud personnel is a risk that can pose potential threat to the cloud data. Assurances should be provided to the clients and proper practices and privacy policies and procedures should be in the place to assure the cloud users of the data safety. The cloud seeker should be assured that the data hosted on the cloud will be confidential.

3.3.3 Data location and Relocation

Cloud Computing offers a high degree of data mobility. Consumers do not always know the location of their data. However, when an enterprise has some sensitive data that is kept on a storage device in the Cloud, they may want to know the location of it. They may also wish to specify a preferred location (e.g. data to be kept in India). This, then, requires a contractual agreement, between the Cloud provider and the consumer that data should stay in a particular location or reside on a given known server. Also, cloud providers should take responsibility to ensure the security of the systems (including data) and provide robust authentication to safe guard the customers' information. Another issue is the movement of the data from one location to another. Data is initially stored at an appropriate location decide by the Cloud provider. However, it is often moved from one place to another. Cloud providers have contracts with each other and they use each other's resources.

3.3.4 Data integrity

By providing the security of data, cloud service providers should implement mechanisms to ensure the data integrity and be able to tell what happened to a certain dataset and at what point. The cloud provider should make the client aware of what particular data is hosted on the cloud, the origin and the integrity mechanisms put in place for the compliance purposes, it may be necessary to have exact records as to what data is placed in a public cloud, when it occurred, what will be the Virtual Memories (VMs).

3.3.5 Storage, Backup and Recovery

When you decide to move your data to the cloud the cloud provider should ensure the adequate data resilience storage systems. At the minimum they should be able to provide RAID (Redundant Array of Independent Disks) storage systems although most of the cloud providers will store the data in the multiple copies across many independent servers [6]. In addition to that, most cloud providers should be able to provide options on the backup services which are certainly important for those businesses that run cloud based applications so that in the event of a serious hardware failure they can roll back to an earlier state and the storage it resided on, and where it was processed. When such data integrity requirements exists, thus the origin and custody of data or information must be maintained in order to prevent tampering or to prevent the exposure of data beyond the agreed territories

3.4 Importance of security

Because of increasing threats to the networked computer systems, there is great need for security [8] innovations. Security practitioners and researchers have made strides in protecting the systems and, correspondingly, individual users' digital asset. However, the problem arises that, until in the recent times, security is treated wholly as a technical problem. The system user is not factored into the equation. Users interact with the security technology either passively or actively. For a passive use beneath the stand ability may be sufficient for the users. For an active use people require much more from their security solution: ease of use, memo ability, efficiency, effectiveness and satisfaction. Today there is an increasing acknowledgment that security issues are also fundamentally human computer interaction issues. Authentication is the process of determining whether a user should be allowed access to a particular system or resource. It is a critical area of the security research and practice. Alphanumeric passwords are used widely for the authentication. Yet traditional alphanumeric passwords have drawbacks from a usability viewpoint, and these usability problems tend to translate directly into the security troubles. That is, users who fail to decide and handle passwords securely open holes so that the attackers can exploit.

4. EXPERIMENTAL RESULTS AND ANALYSIS

The cloud 3.0 simulation tool is used for the cloud computing process. In that simulation tool the different types of document files are used to conduct the experimentation with comparison of the two algorithms namely RSA and AES which are performed very well. The Performance of the encryption algorithms are evaluated considering the following parameters.

- Time taken to encrypt the file
- Time taken to decrypt the file

4.1 Performance analysis of RSA and AES by Encryption process

The encryption time is considered the time taken by the encryption algorithms to produces a cipher text from a plain text. Encryption time is used to calculate the throughput of an encryption scheme, is calculated as the total plaintext in bytes and the encrypted is divided by the encryption time. Comparisons analyses of the results of the selected different encryption scheme are performed. From the experiment results the two algorithms are executed and compared with the different types of files which are pdf, text, doc and jpeg. The algorithms are performed very well and the observed results are depicted in the below Table 1.

Table 1. Encryption time comparison analysis for RSA and AES algorithms

S. No	File type	ENCRYPTION TIME (in SEC)	
		RSA	AES
1	DOC	7.0	4.2
2	TEXT	9.3	3.6
3	PDF	8.0	5.7
4	JPEG	12.1	9.4
5	BMP	4.8	5.3

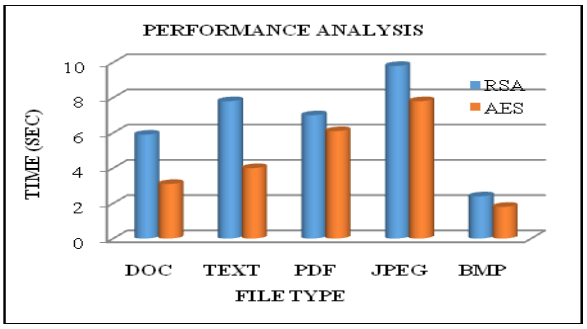

Figure 5. Encryption time comparison of RSA and AES algorithms

The Figure5, shows the encryption time taken for the encryption on the different types of files by the two algorithms i.e. AES and RSA, it is noticed that the AES algorithm takes much less time when compared to the time taken by RSA algorithm. Hence the AES algorithms produce better performance and protect the data from the unauthorized users very well.

4.2 Performance analysis of RSA and AES algorithms by Decryption process

The decryption time is considered the time taken by the decryption algorithms to produces a plain text from a cipher text. Decryption time is used to calculate the throughput of a decryption scheme, is calculated as the total cipher text in bytes and the decrypted is divided by the decryption time. The two algorithms are executed and compared with different types of files like pdf, text, doc and jpeg. The algorithms are performed very well and the observed results are listed in the Table 2 below.

Table 2. Encryption time comparison analysis for RSA and AES algorithms

S. No	File type	DECRYPTION ALGORITHMS	
		RSA	AES
1	DOC	5.9	3.1
2	TEXT	7.8	4.0
3	PDF	7.0	6.1
4	JPEG	9.8	7.8
5	BMP	2.4	1.8

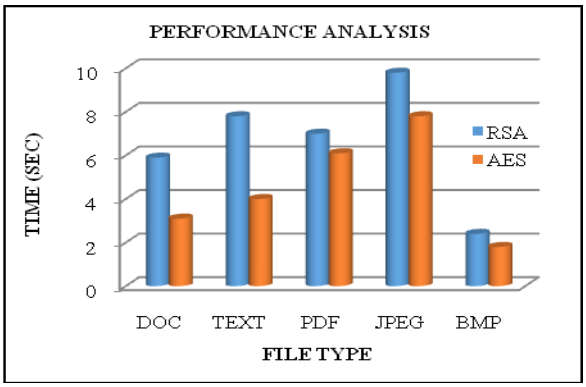

Figure 6. Decryption time comparison of RSA and AES algorithms

From the Figure6, it shows that the decryption times taken for the decryption process on the different types of files by the two algorithms (AES and RSA) are determined. It is observed that the AES algorithm obtain less time when compared to the time taken by RSA algorithm. The AES algorithm is most suitable for protecting the data from the unauthorized users than RSA algorithm.

4.3 Performance analysis of RSA and AES algorithms by different block size

The RSA and AES algorithms are executed with the different block size and observed the time taken to encrypt and decrypt the given file. The obtained results are listed in the Table 3 below.

Table3. Decryption time comparison analysis for RSA and AES algorithms

S. No	Block size	Encryption/Decryption time	
		RSK	AES
1	512	3	6
2	1024	5	8
3	1536	9	11
4	2048	15	12
5	2560	19	13
6	3072	29	18
7	3584	41	30
8	4096	46	33
9	4608	60	39
10	5120	72	41

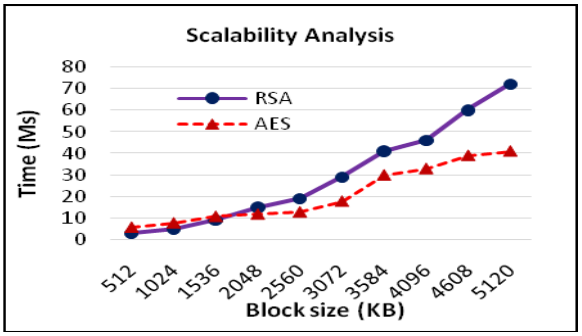

Figure 7. Scalability analysis of RSA and AES algorithms

The figure7 shows that the scalability of the RSA is decreased when increasing the block size from 2048 kb onwards but the AES algorithm shows that the scalability of the AES algorithm is increased when increasing the block size from 2048kb onwards. Hence the AES algorithm proves that it has more scalability than the RSA algorithm for the encryption and decryption process.

5. CONCLUSION

In this paper, it is overviewed the authentication process and its techniques. The authentication process on the cloud has been studied well and implemented using the simulation tool. In this study the three metric based bio-metric authentication types protects the unauthorized users so that they can access the data from the cloud storage and concludes that the Three Factor/Biometric-based authentication technique is convenient, safe and reliable. The data in the cloud storage can be protected by the two security algorithms which are RSA and AES. Five types of files (type i.e. PDF, JPEG, DOC, BMP and TEXT) of the different block size are used for the experimentation with the two encryption / decryption algorithms are AES and RSA.

The AES and RSA algorithms are tested by the two parameters scalability and encryption / decryption time for the same file. By analyzing the table1 and table2, it is noticed that the RSA has more memory usage as compared to AES algorithm. Encryption / decryption time taken by RSA algorithm is less as compared to the time taken by the AES algorithm. By analyzing the Figure3, is shows the scalability power of the two algorithms (i.e. AES and RSA) and it is noticed when executed with the different block sizes from 512kb to 5120kb. The RSA algorithm does not perform well when increasing the block size from 2048kb onward because it takes more time to complete the encryption/decryption process, but the AES algorithm delivers better results when increasing the block size from 2048kb onwards. It takes less time when compared to the time taken by the RSA algorithm. The different mechanisms are used to authenticate the authorized users in the cloud and protect the audio files and video files from the cloud storage in future work.

6. REFERENCES

[1] A. Anusha Priya , Dr. S. Saravanan,Protecting Healthcare Database by Access Control Method On Cloud Computing Technique - A Survey, international journal of scientific research.2015.

[2] B. Dhivya, L.M.Nithya, Privacy Preserved Secure and Dependable Cloud Data Storage, in International Journal of Computer Science and Management Research NCNICS 2013 Issue.

[3] Cong Wang, Qian Wang, Kui Ren, and Wenjing Lou,‖, Privacy- Preserving Public Auditing for Data Storage Security in Cloud Computing‖ in IEEE INFOCOM 2010, San Diego, CA, March 2010.

[4] Deepika Verma1, Er. Karan Mahajan, To Enhance Data Security in Cloud Computing using Combination of Encryption Algorithms, International Journal of Advances in Science and Technology (IJAST) Vol 2, Issue 4, 2014.

[5] Dr. Prerna Mahajan & Abhishek Sachdeva, A Study of Encryption Algorithms AES, DES and RSA for Security, Global Journal of Computer Science and Technology Network, Web & Security Volume 13 Issue 15, 2013.

[6] F. M. Shelke, and P. D. Soni,, An enhanced authentication strategy for multiservice authorization over mobile cloud, International Journal on Recent and Innovation Trends in Computing and Communication, vol. 3, issue 3, pp. 1669-1672.

[7] Hafiz Zahid Ullah Khan, Comparative Study of Authentication Techniques, International Journal of Video& Image Processing and Network Security IJVIPNS-IJENS Vol 10 No.04, pp: 9-13.

[8] N. Saravanan, A. Mahendiran, N. Venkata Subramanian and N. Sairam, An Implementation of RSA Algorithm in Google Cloud using Cloud SQL,Research Journal of Applied Sciences, Engineering and Technology 4(19): 3574-3579, 2012.

[9] Nasrin Khanezaei, Zurina Mohd Hanapi, A Framework Based on RSA and AES Encryption Algorithms for Cloud Computing Services, Proceedings of 2014 IEEE Conference on Systems, Process and Control (ICSPC 2014), 12 - 14 December 2014, Kuala Lumpur, Malaysia, pp 58-62.

[10] P. Kalyani Karule , Neha V. Nagrale Comparative Analysis of Encryption Algorithms for Various Types of Data Files for Data Security, International Journal of Scientific Engineering and Applied Science (IJSEAS) – Vol-2, Issue-2, February 2016

[11] Parsi Kalpana and Sudha Singaraju, Data Security in Cloud Computing using RSA Algorithm, International Journal of Research in Computer and Communication technology, IJRCCT, ISSN 2278-5841, Vol 1, Issue 4, September 2012, pp 143-146.

[12] Voorsluys, William; Broberg, James; Buyya, Rajkumar (2011) "Introduction to Cloud Computing". Cloud Computing: Principles and Paradigms. New York, USA: Wiley Press.pp.1– 44.ISBN978-0-470-88799-8.