Enhancing Security using Cryptography and Steganography and Providing Data Backup and Recovery in Cloud

K. Deerthana, B. Devi Saranya
Dept of Information Technology
Anna University Tiruchirapalli, India

R Jayamala, Asst Professor
Dept of Computer Science Anna University
Tiruchirapalli, India

***Abstract* –Cloud computing is a kind of Internet based computing, where shared resources, data and information are provided to computers and other devices on-demand [7]. Cloud computing allows us to Create, Configure and Customize the business applications in online.Security is headed against unauthorized access, to reduce risk of stealing. So there is a need to protect the data against unauthorized access. To provide data security, this paper introduces combination of Cryptography (Blowfish algorithm) and Steganography. To reduce time complexity, Blowfish algorithm is used. Blowfish has a better performance than other common encryption algorithms. Blowfish algorithm is to maintain data Integrity at the untrusted server. It is faster in encryption and decryption and the buffer space requirement is less as compared to other algorithms. Steganography is the art of hiding the file into other source like image, audio or video. Mostly digital images are popular because of their frequency on the internet. This paper also provides the compression technique to store the back-up of data.We have proposed a strong, formal model for data security on cloud and corruption detection using MD5 algorithm.**

Keywords: Blowfish, Steganography, JPEG Compression, MD5 algorithm.

I. INTRODUCTION

Cloud computing and storage solutions provide users and enterprises with various capabilities to store and process their data in third- party data centres [6]. Cloud Computing is not a technology, also called as distributed computing. Because, it can be able to run a same programs and applications on many connected computers at same time through the internet. Cloud computing include any payment based or pay-per-use service that, in real time over the Internet. Because of these benefits each and every organizations are moving their data to the cloud.So there is a need to protect the data against unauthorized users, modification or denial of services etc. Security goals of data include three points namely: Availability Confidentiality, and Integrity.

II. TYPES OF CLOUD

Private cloud and public cloud are the two types of cloud. The combination of these clouds is Hybrid cloud. Sensitive information's can be stores in Private cloud. All the information's can be stored in public cloud. Community cloud, this is a special type of cloud to share and reduce the cost of computing system. Data storage in cloud computing offers so many benefits to users [9][10].

III. CLOUD COMPUTING SERVICE MODELS

Cloud computing services can be divided into three services namely Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS).

A. Software as a Service:It provides different software applications over the Internet by Application Service Provider. Thus it eliminates tremendous load of software maintenance.

B. Platform as a Service:It provides an infrastructure with a high level of integration in order to implement and test cloud applications. The user does not manage the infrastructure.

C. Infrastructure as a Service:It refers to the sharing of hardware resources for executing services using Virtualization technology. Its main objective is to make resources such as servers, network and storage more readily accessible by applications and operating systems [8].

IV. SECURITY ISSUES IN CLOUD COMPUTING

1. Data /System Integrity
2. Authentication
3. Storage, Backup and Recovery of data
4. Data Confidentiality and privacy
5. Access control
6. Availability

Three aspects of security are confidentiality, integrity and availability. Confidentiality is hiding information and resources. Integrity refers to the trustworthiness of data or resources, which usually prevents any incorrect or unauthorized changes. Availability refers to the ability to use the information or resource. Thus security is a major challenge in cloud computing [10].

V. SERVICE PROVIDER

Usually service providers will give the cloud space to the data owners and users. But the service provider will give the cloud space for cost only. So that data owner and user should pay some money for this cloud space. Cost will depend on how much space they need. If suppose the cloud space is 1GB then it cost around RS 4500 per year, in our implementation use we are using 500MB. There are so many cloud service providers are available such as I cloud, goggle apps, open drive, drop box, Amazon.

VI. DATA OWNERS

Data owner can upload their data's in cloud. For that first the data owner got some space from the cloud provider. First the data owner should register their own details and buy some space from the service providers. Due to lot of security problems data owner uploading their data after encryption only by using some key to encrypt the plain text. Based upon the designation the encryption algorithm has been written.

VII. THIRD PARTY AUDITOR

Third party auditor (TPA) for given information to the users any type of modification can include our data. It provides different applications and services to use a shared pool of configurable computing resources. A Third Party auditor (TPA) is to check the integrity of another data add or not in user data. Third party auditor is to just a mediator for users, like an informer. The TPA is external third party, it is the only an individual and independent operator for the users. It is also can inform the any changes in the stored cloud data. The capabilities and access the serve to make auditing of users access getting data.

VII. LITERATURE SURVEY

1) Shirole and Sanjay [4], "Data Confidentiality in Cloud Computing with Blowfish Algorithm", propose a system that uses encryption technique to provide reliable and easy way to secure data for resolving security challenges. Scheduler performs encryption .On plain data into cipher data followed by uploading of ciphered data on the cloud. When the data is to be retrieved from the cloud, it is obtained in plain data format and is stored on the system.

2) Garima and Naveen [5], "Triple Security of Data in Cloud Computing", this paper proposes two cryptographic algorithms viz. DSA (Digital Signature Algorithm) and AES (Advanced Encryption Standard) and Steganography. DSA is used for authentication purpose, AES is used for encrypting the data and Steganography is used.

3) Parsi and Sudha [6], "Data Security in Cloud Computing using RSA Algorithm", to provide data security in cloud environment, RSA algorithm has been implemented to provide the same. RSA stands for Ron Rivest, Adi Shamir and Len Adleman. RSA is public key cryptography. In the proposed system, RSA is used for encryption as well as decryption of data. The process involves that the data is encrypted and then uploaded onto the cloud. For decryption of data, data required is downloaded from the cloud, cloud provider authenticates the user and then the data is decrypted. RSA is used to provide authenticated access to intended user only and hence makes the system secure.

VIII. EXISTING SYSTEM

In the existing system, they provide security to data by implementing algorithms like AES, Steganography together to cloud network. The above mentioned scheme revolves around the problem of data security and with the help of data safety but also simple in its implementation and hence usage. Disadvantage of this paper is "Time Complexity" because of using AES algorithm and the image produced by Steganography requires more storage space [1] [4].

IX. PROPOSED WORK

The proposed work is based on blowfish encryption algorithm with Steganography in order to make a secure system. The encryption/decryption process of this algorithm is complex and cannot be broken by any intruder.

Then apply Steganography that hiding the data in the image and that image is stored in the server in the cloud. Image compression helps to improve storage and data recovery process is also done by remote server in case of any data corruption. This is accomplished by MD5 algorithm by generating hash value.

A) Blowfish Encryption:

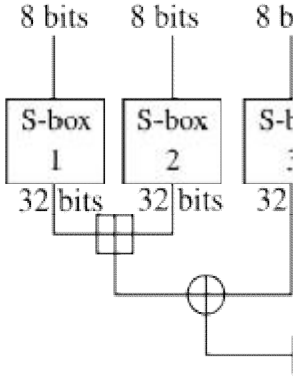
In 1993, Bruce Schneider [1993] published the Blowfish block cipher. At this time, the current Data Encryption Standard (DES) was known to be vulnerable to crypto analysis and brute-force attacks. Other cryptographic algorithms were available to replace DES, but many of the cryptographic algorithms were either protected by patents or considered proprietary. Schneier developed Blowfish to be a publicly available cryptographic algorithm with the potential to replace DES.It is a Feistel Network, iterating a simple encryption function 16 times. The block size is 64 bits, and the key can be any length up to 448 bits. Although there is a complex initialization phase required before any encryption can take place, the actual encryption of data is very efficient on large microprocessors.

Blowfish consists of two parts: key-expansion and data encryption. During the key expansionstage, the inputted key is converted into several sub key arrays total 4168 bytes. There is the Parray,which is eighteen 32-bit boxes, and the S-boxes, which are four 32-bit arrays with 256entries each. All of these boxes are initialized with a fixed string, the hexadecimal digits of pi.After the string initialization, the first 32 bits of the key are XORed with P1 (the first 32-bit box in the P-array). The second 32 bits of the key are XORed with P2, and so on, until all 448, or fewer, key bits have been XORed. Cycle through the key bits by returning to the beginning of the key, until the entire P-array has been XORed with the key. Advanced versions of blowfish are two fish and three fish. They are symmetric key block cipher with a block size of 128 bits and key sizes up to 256 bits. It was one of the five finalists of the Advanced Encryption Standard contest, but it was not selected for standardization. Two and three fish is related to the earlier block cipher Blowfish [11][12].

B) Blowfish Algorithm:

a) Divide x into two 32-bit halves: xL, xR
b) For i = 1 to 16:
c) xL = xL XOR Pi
d) xR = F(xL) XOR xR
e) Swap xL and xR
f) Next i
g) Swap xL and xR (Undo the last swap.)
h) xR = xR XOR P17
i) xL = xL XOR P18
j) Recombine xL and Xr

S-Box Genereation



C) Steganography:

The main goal of steganography is to hide the information using some covered media[13]. In case of cryptography the user can able to see the contents of message but can't comprehend the information. On the other hand, in steganography the existence of information will not be noticed by viewer because it is embedded inside some medium. This medium is also called as carrier or cover object. It may be an image, video, texts, sound or any music file.

D) Information Hiding Using Steganography:

- The secret message (M), which is going to be hidden, it may be a plain text or cipher text or any type of data.
- The Cover media (C) , which hold the message
- The Stego-Function (Fe) and its inverse (Fe-1)
- The Stego-Key (K) or Password use to hide or unhide the data.

E) Compression Techniques:

The process of reducing the size of data is known as "data compression". In digital signal processing, data compression, source encoding and low bit rate reduction involves encoding information using fewer bits than the original representation. Compression is useful because it helps reduce resource usage, such as data storage space or transmission capacity.

F*) JPEG Compression:*

The acronym JPEG stands for the Joint Photographic Experts Group, a standards committee that had its origins within the International Standard Organization

(ISO)[14].JPEG, is a lossy compression algorithm for images.The algorithm behind JPEG is relatively straightforward and can be explained through the following steps:
1. Take an image and divide it up into 8-pixel by 8-pixel blocks. If the image cannot be divided into 8-by-8
Blocks, then you can add in empty pixels around the edges, essentially zero-padding the image.
2. For each 8-by-8 block, get image data such that you have values to represent the color at each pixel.
3. Take the Discrete Cosine Transform (DCT) of each 8-by-8 block.
4. After taking the DCT of a block, matrix multiply the block by a mask that will zero out certain values fromthe DCT matrix.
5. Finally, to get the data for the compressed image, take the inverse DCT of each block. All these blocks are combined back into an image of the same size as the original. The data is compressed and store at remote server which reduces the storage at remoteserver to large extent. Remote server provides the backup of users' data in case of cloud crash.

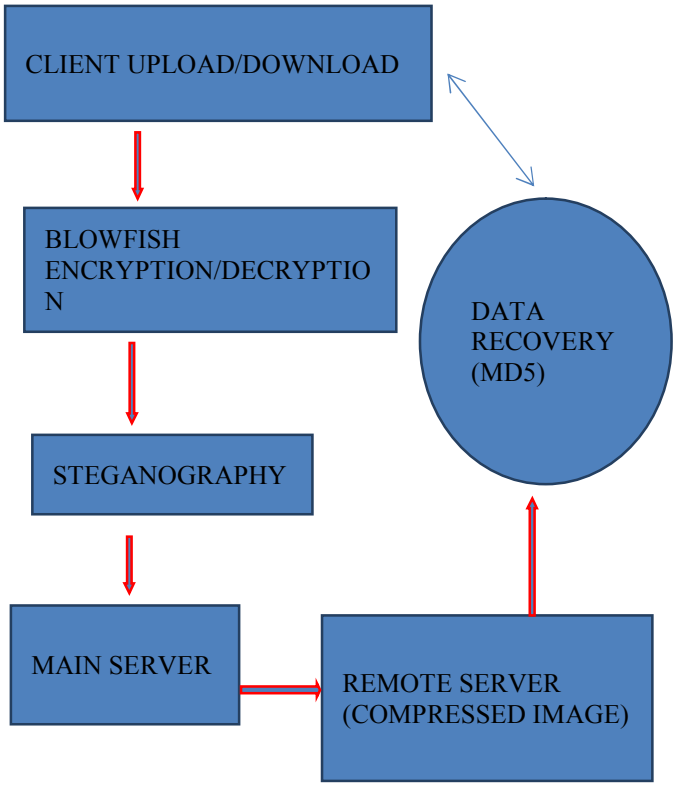
G) File Corruption Detection and Recovery:

To recover infected or corrupted resources /files.We have implemented MD5 algorithm to check all files are stored as it is or any files are modified due to impact of any mishap. In case of data corruption or missing file part, we can recover original data from cloud server by using MD5 hash value.By using MD5, we have calculated hash value and stored on database.When user wants to download his/her file then again hash value of current file is calculated and verified with old hash value. If both hash values are matched then user gets his original file. If hash values are not same then we can say that file parts are corrupted or infected. When we know that file is corrupted then we can recover this file[15].

H) MD5 (Message-Digest 5 algorithm):

It is a widely used cryptographic function with a 128- bit hash value. MD5 has been employed in a wide variety of security applications, and is also commonly used to check the integrity of files. An MD5 hash is typically expressed as a 32-digit hexadecimal number MD5 processes a variable-length message into a fixed-length output of 128 bits.

SYSTEM ARCHITECTURE



X. WORKING OF PROPOSED WORK

A) Data uploading:

The data to be secure is uploaded to the cloud main server by user. The third party auditor (TPA) will receive this data before stored in cloud to provide security. He/she will do encryption by blowfish algorithm and create stegano image using Steganography. Then the secured image is stored in cloud main server by TPA. It is also stored in remote server as a compressed image by JPEG compression technique to reduce the size of the image and also size of remote server. Now the data is secured and also the back up of data is to be stored. In case of main server is crashed, this data will be used.

B) Data downloading:

When the user wants to download their data, the TPA will decrypt the file from stegano image in the remote server and blowfish decryption will be done to get an original data. To check data integrity the user will produce hash value by MD5 algorithm. Already the original data before uploading has some hash value while comparing both the hash values the data corruption is detected. When it is not same it is concluded that data is

corrupted/changed. So by using this algorithm data recovery is flexible and data integrity is achieved.

XI. CONCLUSION AND FUTURE SCOPE

From this paper the Blowfish algorithm is fast, so it reduces time complexity and it provides strong encryption and decryption. Steganography gives more security to the file stored in the cloud. Here we are using Remote server to back up the compressed data in case of cloud crash. In this paper MD5 algorithm is used to identify the data corruption. So the integrity of data is checked using MD5. In future we can try this security mechanism in multi cloud and also in advanced blowfish algorithms like two fish and three fish.

REFERENCES

[1] Jawahar takur, Nagesh Kumar, "DES, AES, and Blowfish: Symmetric key algorithms simulation based Performance analysis", International Journal of Emerging Technology and Advanced Engineering, Volume1, Pg no.6-12, 2011.
[2] Prachi Jain, Prof.Shubhangi Kharche, "Effectuation of Blowfish Algorithm using Java Cryptography" , international journal of scientific & engineering research, volume 4, pgno.13911400, 2013.
[3] Poonam M. Pardeshi, Prof. Bharat Tidke, "Improving Data Integrity for Data Storage Security in Cloud Computing", International Journal of Computer Science and Information Technologies, Vol. 5 (5), Pg No.6680- 6685, 2014.
[4] Govinda K, Mythili D, Geetha Priya S, "Data Security in Cloud using Blowfish Algorithm", International Journal for Scientific Research & Development, Vol. 2, Pg no.523-52, 2014.
[5] Rajesh M. Devakate, Amol B Rajmane, "Dynamic Resource Allocation and Data Security for Cloud", International Journal of Innovative Research in Advanced Engineering, Volume 1, Pg.no.272-276,2014.
[6] Hassan, Qusay (2011). "Demystifying Cloud Computing" (PDF). The Journal of Defense Software Engineering (Crosstalk) 2011 (Jan/Feb): 16–21.
[7] From this link, www.smartlogic.com/glossary/cloud-computing
[8] Kurdi, Heba; Li, Maozhen; Al-Raweshidy, H. S. (2010). "Taxonomy of Grid Systems". In Antonopoulos, Nick. Handbook of Research on P2P and Grid Systems for Service-Oriented Computing: Models, Methodologies and Applications. IGI Global research collection. IGI Global. p. 34. ISBN 978-1-61520-687-2. Retrieved 2015-07-29.
[9] "Self-Run Private Cloud Computing Solution — GovConnection". govconnection.com. 2014. Retrieved April 15, 2014.
[10] Foley, John. "Private Clouds Take Shape". Information Week. Retrieved 2010-08-22.
[11] Z. Xiao and Y. Xiao. 2013. Security and Privacy in Cloud Computing, Communications Surveys and Tutorials, IEEE.15 (2): 843-859. Doi: 10.1109/SURV.2012.060912.00182
[12] From this link https://en.wikipedia.org/wiki/Blowfish_%28cipher%29
[13] Bruce Schneier (1993). "Description of a New Variable-Length Key, 64-Bit Block Cipher (Blowfish)". Fast Software Encryption, Cambridge Security Workshop Proceedings (Springer-Verlag): 191–204.
[14] Fridrich, Jessica; M. Goljan; D. Soukal (2004). "Searching for the Stego Key" (PDF). Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VI 5306: 70–82. Retrieved 23 January 2014.