International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

# IBADedup - Image Based Authentication and Deduplication Scheme in Cloud user Group

**T. A. Mohanaprakash, J.Andrews**

**Abstract:** ***In recent times the cloud storage services are widely used for storage and processing. This has made the increasing demand for the mechanism or methodology which will provide the solution for reducing the use of redundant data, to achieve better space and bandwidth requirements of a storage service. Cloud computing enables all the remote workstations are light weighted where the operating system, services and user data are centralized in cloud servers and are shared by other users of the cloud. This sort of data sharing causes collusion attacks on an unsecure environment. Thus a secure protection scheme is needed to encrypt the private information. This may prove to be a problem because the distribution and storage of a key is difficult in a cloud with dynamic users. The proposed system provides distribution of key without the need for a secure communication channel using a group manager. This system also performs data De-duplication and prevents unauthorized access to a file using graphical passwords. Here the user selects an image to authenticate his identity. This image is split into 16 parts by the administrator; these 16 parts are generated randomly using a pseudo random generator technique. Then, the administrator holds onto 4 parts of the entire image to authenticate the user at time of login using Attribute Based Encryption. This helps the administrator authenticate the user and also prevents a revoked user from gaining access to a file. The Group Manager ensures the privacy of the files shared within a group by updating the user information to the database and by using RC6block cipher technique to encrypt the documents at time of storage. Thus when the user needs to download a file from a group a request is sent to the group manager and the decrypted file is then downloaded.***

***Keywords: - Data deduplication, Image Based Authentication, Key distribution, Privacy-preserving.***

## I. INTRODUCTION

Cloud computing is considered to be a popular technology that comprises a higher level of computational power including a large storage capacity with the promise of lower expenditure on hardware and software. A cloud computing environment improves storage and retrieval of a file and other services over the internet by reducing the need for a well-equipped terminal at the user side. The key concept of cloud computing has helped us avail services and resources using a pay-as-you-go model [1]. However, this uniqueness has brought new challenges in security and the privacy of the files being accessed on a cloud.

The storage and distribution of keys are more problematic in the cloud because it requires a secure communication channel which is difficult for practice. The current cloud schemes use a Secret Sharing technique to store and distribute the key among a group of recipients [2]. The recipients with the key can regenerate the original secret sent by the sender. This technique suffers when revocation of user privileges are possible in the group. Thus the proposed system uses a Group manager and a Group admin to ensure key distribution and authentication in a cloud group [3]. This work is to develop a system that ensures protection against the existing issues of a cloud computing environment. These issues include security, privacy, file access, hacking and collusion attack. The scope of the paper is to analyze various anti collusion schemes used among the existing cloud model and to develop a system that provides image based authenticationand key distribution within a cloud group [4]. In the current times all the internet users are using huge amount of data, these users increasing not only web traffic but also increasing the demand for cloud data storage. All internet based companies, as well as social media users, are outsourcing their storage to cloud servers. This storage of records at one or more servers is leading to huge amount of data duplication and delaying of data storage and retrieve memory with identical copies of data. Cloud data storage lot of redundant data is untruthful occupying the most valuable storage space and creating a false unavailable storage space message to cloud users. This problem can be solved through erasing redundancy and cleaning the unwanted replicas of equal data, a highly advanced mechanism was introduced known as Deduplication(Fig-1). Data Deduplication technique is used to eliminate redundant data resulting in storage of unique data copies only [5]. It is increasing the better storage utilization and proves to be an efficient technique to handle related data. It is one of the best mechanisms which eliminates duplicate data, minimizes the unnecessary usage of bandwidth and reduce the storage cost.

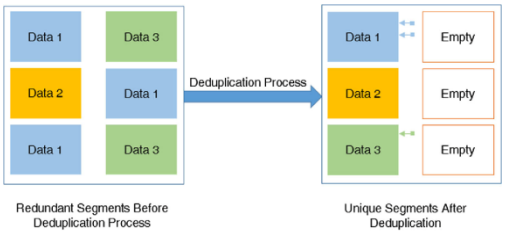



**Fig-1:DataDeduplication process**

Manuscript published on November 30, 2019.
* Correspondence Author
**T.A.Mohanaprakash***, Research Scholar, Department of CSE, Sathyabama Institute of Science and Technology, Chennai 600119, Tamilnadu, India.
**Dr.J.Andrews**, Associate Professor, Department of CSE, Presidency University, Rajanakunte, Yelahanka, Bengaluru-560064,Karnataka,India.

In this paper, various types of data deduplication techniques and anti-collusion mechanisms are discussed in Section II and comparative analysis has been organized in tabular form in section III. We have discussed proposed system model in section IV. Design and implementation explained in Section V. Experimental result shown in Section VI, We have analyzed system performance in Section VII, Conclusion and future scope has been entailed in Section VIII.

## II. BACKGROUND AND RELATED WORK

### A. Existing scheme

In existence the secure protection scheme used involves encryption using a secure key that is distributed by the means of a secure communication channel which is achieved using a secret sharing technique.This technique stores and distributes a distinct and randomly generated key among a group of recipients and this mechanism is feasible for N recipients in a cloud group. When a file is shared within the group the recipients with the key will be able to regenerate the secret information [6]. But, this privilege is reserved to only the members of the group M and all the other M-1 groups are excluded from this procedure. The existing system also performs encryption of the data before it is stored on the server. The chances of replication in the existing model are higher and thus the current model has a poor utilization of the cloud storage. The major issue in the existing system is,

1. The privileges provided to a revoked user.
2. Difficult to identify collusion attacks among cloud groups.
3. Deduplication is a challenging task.
4. Higher vulnerability.
5. Storage utilization is minimal.
6. Key distribution is challenging.

### B. Literature survey

Hiroaki Anada, Junpei Kawamoto, ChenyutaoKe, Kirill Morozov, Kouichi Sakurai, the author demonstrated across-group secret sharing technique for secure utilization of cloud data storage over different cloud service providers and section. This is simple secret sharing algorithm cross-group secret sharing (CGSS) is suitable for storing the data on cloud storage distributed over different groups. That is, different providers and regions. By combining an $\ell$ -out-of-m threshold secret sharing technique with a k-out-of-n edge secret sharing scheme using a symmetric-key encryption scheme, we construct the CGSS scheme that forces k shares to be collected from $\ell$ groups. “Cross-group secret sharing scheme for secure usage of cloud storage over different providers and regions” [3]. Hattim, May &Taha, Zahraa proposed a Secure And Hidden Text Using Aes Cryptography and Lsb Steganography. This system uses Advanced Encryption Standard (AES) algorithm tochange over content from its unique structure (plain text) to incoherent structure(figure content) then figure content is concealed in the picture by LeastSignificant Bit (LSB). The content is encoded with key 128 bit. Indeed, even withan alternate arrangement sorts and diverse sizes of the chosen pictures is utilized to cover up scrambled content. “Secure andhidden text using AEScryptography and lsbsteganography” [4].

Yan, Zheng, Mingjun Wang, Yuxiang Li, and Athanasios V. Vasilakos and team introduced encrypted data management technique with data deduplication in Cloud Computing. This approach focused data deduplication and this is based on attribute-based encryption (ABE) to original copy of encrypted data stored in the cloud and it support secure data access control over communication network. Theanalyze and implementation shows that the scheme is secure, effective and efficient in the way to achieve better cloud management. “Encrypted Data Management with Deduplication in Cloud Computing” [5]. Tao Peng, Qin Liu, Guojun Wang, Central South University, China. Proposed a Multilevel Access Control Scheme in Transparent Computing (MACTC) to protect user data with various security stages, and provide more than one level access control and valid identity authentication. This work is effective in multilevel data security, flexible in authorized resource sharing, and secure against various malicious attacks. The experiment results verify the feasibility of the scheme. “A Multilevel Access Control Scheme for Data Security in Transparent Computing ” [6]. Tchernykh, Andrei &Babenkob, Mikhail &Chervyakovb, Nikolay& Miranda-Lópeza, Vanessa &Kuchukov, Viktor & Cortés-Mendoza, Jorge &Deryabin, Maxim &Kucherov, Nikolay&Radchenko, Gleb&Avetisyan, Arutyun, the authors showed a AC-RRNS technique based on modified threshold Asmuth–Bloom and Mignotte secret sharing schemes. This scheme achieves basic needs of computational security. This scheme ensures security under several types of attacks optimizes and data redundancy of encryption. “AC-RRNS: Anti-collusion secured data sharing scheme for cloud storage”. International Journal of Approximate Reasoning [8]. Wenhao Li, Yun Yang and Dong Yuan, Proposed a proactive replica checking method. PRCR guarantees reliable of the largecloud data with the minimum duplication.This method also provides low cost effective benchmark for replication based methodologies. The early used multi-replica approach in current cloudstorage occupies large storage space in cloud server and also increases storage price for data-intensive applications in the cloud computing. In order to reduce the cloud storage consumption this paper introduces cost-effective data reliability management mechanism named PRCR, this approach achieves the data reliability requirements and the system based on a comprehensive data reliability model. Author compared existing algorithm with PRCR, which can reduce from 33% to 66% of the cloud storage space consumption, hence significantly lowering the storage cost in a cloud. “Ensuring Cloud Data Reliability with Minimum Replication by Proactive Replica Checking” [14]. Li, J., Li, J., Xie, D., &Cai, Z. are demonstrated the Secure Auditing and deduplicating Data in Cloud mechanism. This concept developed using map reduces technology along with Hadoop. Here authors used two method SecCloud and SecCloud +. SecCloud, it checks the integrity auditing and secure deduplication of plain files,SecCloud+ performs integrity, auditing and deduplication on encrypted files .

This approach helps users to generate data tags before uploading file as well as data audit and data integrity information stored in cloud and also SecCloud allows secure data deduplication via presenting a PoW protocol and avoiding the leakage of side network information in data deduplication.

The author analysis SecCloud with earlier work, the execution by user in SecCloud is significantly reduced during the file uploading and auditing times. “Secure Auditing and DeduplicatingData in Cloud”, IEEE Transactions on Computers [36]. SayaleeShinde, Dr. S. S. Shaikh and team compared all techniques and methodologies which are already used to manage anti collusion in data sharing in cloud storage. Here this author analyzed from the Multi-Owner Data Sharing(MONA) to Cipher text-Policy Attribute-Based Encryption techniques. The papers is in the“A Survey on Secure Anti-Collusion Data Sharing Scheme for Dynamic Groups in the Cloud” [20]. Zhu, Zhongma& Jiang, Rui.Proposed a secure system for distributing a key in none secure messaging networks, and therefore the network users will confidently obtain their non-public keys from group manager. The author’s separated this process in to three major categories. First they focused algorithm to transmit the key without secures communication network. Secondly, author concentratedfull control access management which is performed where all the users within the group of members will use the key supply within the cloud and revoked users will not able to access the cloud data again after they're revoked. Third, the cloud storage is secure from any type of collusion attack, which suggests that revoked network users cannot get the early record though they combine with the un-trusted cloud. “A Secure Anti-Collusion Data Sharing Scheme for Dynamic Groups in the Cloud” [21]. G. Mercy Vimala, R. Vara Prasad, P. Rama Rao, G. Proposed a safe information sharing plan for element individuals in a cloud. Then, The authorized security information to the sharing data records will be given since they are subcontracted. The author mainly focused a safe route for key distribution with unsafe communication network channels. Then the clients can safely obtain their private keys from group administrator. The idea of the author is to accomplish powered access control and any client in the group can use the data resource in the cloud storage and rejected users cannot get access to the cloud again after they are forbidden. Finally this system guards trickery attack. “Secure Anti-Collusion Data Sharing Scheme for Groups in the Cloud ” [22]. NareshVurukonda, B. ThirumalaRao,are proposed the problems associated to the cloud data storage like data breaches, data theft, and inaccessibility of data stored in cloud. The cloud computing is innovative approach and it changing system to setup of hardware and software design and procurements. This approach provides integrated storage because of all the cloud users is storing important and valuabledata and application to cloud data centers. The Cloud service provider (CSP) in the position to give high leveldata availability, privacy, integrity and confidentiality.Sometime the Cloud service provider (CSP) is not in the position to providing reliable data
intensive services to the cloud users and vendors. Here the author study and analysis all problems and identifies the associated issues and trying to providing a possible solution to the respective issues in cloud data storage. “A Study on Data Storage Security Issues in Cloud Computing” [18]. AnanduJayan, Akash Nair Bhargavi R. Upadhyay, Supriya M.and team introduce the modified RC4 algorithm to overcome the drawbacks of existing simple RC4 algorithm. In the internet-based computing era, software, hardware and applications are involving a major role. With this role, the cryptography of this mechanism is also of a much more important. So this author’s concentrate on cryptographic algorithm. These techniques are complex and can consume lot of time and hence, it can benefit from parallelism. But parallelism is not that simple, as sometimes it is not possible to convert the whole algorithm into a parallel one. So author proposed a modified RC4 (MRC4) algorithm which can be made fully parallel. Usually the RC4 technique is separated into two category one is key scheduling algorithm (KSA) and pseudo random number generator algorithm (PRGA). This modified algorithm is compared with other cryptographic algorithms i.e., AES, DES and RSA for its speed and time complexity. This algorithm is also implemented in Cuda, MPI and OpenMP and the results are compared. “Performance Analysis of Modified Rc4 Cryptographic Algorithm Using Number of Cores in Parallel Execution” [24].Anurag Singh Tomar, Gaurav Kumar Tak, RuchiChaudhary, Proposed the system Image based authentication with secure key distribution approach in cloud computing. This system based on Image based authentication with secure key exchange between cloud user and Cloud service provider (CSP). User authenticated by CSP using Image based authentication after that key will be exchanged between cloud user and CSP and that key will change from session to session. “Image based authentication with secure key exchange mechanism in cloud”. [33] Waters, addressing the existing system draw backs in cipher text and attribute based encryption and proposed Ciphertext-Policy Attribute Encryption. This method allow all typeencryptor to require access control in terms of any access formulation over the attributes in the system. This approach having three section. First Parallel Bilinear Diffie-Hellman Exponent (PBDHE) technique used to view as a generalization of the BDHE assumption and other two decisional Bilinear-Diffie-Hellman Exponent and decisional Bilinear DiffieHellman assumptions used provide security to cloud data. “Ciphertext-policy attribute-based encryption: An expressive, efficient, and provably secure realization” in Proc. Int. Conf. Practice Theory Public Key Cryptography Conf. Public Key Cryptography” [34].

## III. COMPARATIVE ANALYSIS OF VARIOUS ANTI COLLUSION SCHEMES AND DEDUPLICATION SCHEMES

| TITLE | AUTHORS | TECHNIQUE USED | METHODOLOGY |
|---|---|---|---|
| Encrypted Data Management with Deduplication in Cloud Computing [5] | Yan, Zheng, Mingjun Wang, Yuxiang Li, and Athanasios V. Vasilakos | Attribute Based Encryption (ABE) algorithm | Author used CP-ABE method with the pairing based cryptography (PBC). This system uses three separated section CSP,data owner and data holder. This system input is collected fron data holder or holder PDA (such as a smart mobile device) and preprocess data collected by IoT devices. |
| Boafft : Distributed deduplication for Big Data Storage in the cloud [32] | G. Zhang, C. Wu, S. Luo | An algorithm of Data routing on the basis of similar index | Boafft - Data routing which is based on resemblance of information. By identifying the locations and this process reduce network overhead. In this an in-memory similarity index is maintained in every data server which reduces number of random disk read/write operations this process speed up the local data deduplication. |
| DedupDUM: Secure and Scalable Data Deduplication With Dynamic User Management [26] | Yuan, Haoran& Chen, Xiaofeng& Jiang, Tao & Zhang, Xiaoyu& Yan, Zheng& Xiang, Yang | DedupDUM-which supports cloud user revocation and new cloud user joining by exploiting the re-encryption techniques | DedupDUM contains two entities: the cloud user and the cloud server. DedupDUM=( Setup, Encrypt, Group-Encrypt, Group-Decrypt, Decrypt ) In the cloud environment, the cloud server updates the ownership list of the cloud users frequently. The unauthorized cloud users should be prevented from accessing the data stored in the cloud server before uploading their own data or after deleting or updating the data. |
| A similarity-aware encrypted deduplication scheme with flexible access control in the cloud [27] | Zhou, Yukun&Feng, Dan &Hua, Yu & Xia, Wen & Fu, Min & Huang, Fangting& Zhang, Yucheng, | To protect data security,users encrypt data with message-locked encryption (MLE) to enable deduplication over ciphertexts.. | EDedup groups files into segments and performs server-aided MLE at segment-level, which exploits similarity via a representative hash to reduce computation consumption. |
| ClouDedup: Secure deduplication with Encrypted Data for Cloud Storage [10] | Puzio, P., M olva, M., &Loureiro, S | Block-level deduplication | This system architecture uses a low complicated design in parallel IO system and distributed file system. Storing huge size data files of approximately terabytes of cloud data. |

| | | | |
|---|---|---|---|
| A new content-defined chunking algorithm for data deduplication in cloud storage. [13] | Widodo, Ryan & Lim, Hyotaek&Atiquzzaman, Mohammed. | Rapid Asymmetric Maximum (RAM) | Rapid Asymmetric Maximum uses bytes value method to declare the cut points. This mechanism uses a fixed-sized window and a variable-sized window method to find a maximum-valued byte. It is calculated as cut point. The light-valued byte is included in the chunk and located at the edge of the chunk. |
| Public Integrity Auditing for Shared Dynamic Cloud Data with Group User Revocation [6] | Tao Peng, Qin Liu, Guojun Wang, Central South University, China | Diffie-Hellman assumption ‘ | Dynamic public integrity auditing scheme with group user revocation |
| Efficient Cross-User Chunk-Level Client-Side Data Deduplication with Symmetrically Encrypted Two-Party Interaction[34] | Chia-Mu Yu | XDedup, based on Merkle puzzle | XDedup is the first brute-force strong encrypted data deduplicationwith only symmetrically cryptographic twousers interaction. |
| A Secure Anti-Collusion Data Sharing Scheme for Dynamic Groups in the Cloud [21] | Zhongma Zhu and Rui Jiang | Delov-Yao model | Scheme including method called key distribution, data confidentiality, access control and efficiency |
| Enhancing Security of Cloud Computing by using RC6 Encryption Algorithm [28] | Salim Ali Abbas , Malik Qasim Mohammed | Rivest cipher 6 – RC6 | This approach uses secure key encryption method with key distribution algorithm to produce the client provided key. Additionally this author uses extended array S to fill up the variety of t random binary words. |
| An efficient AES and RC6 based cloud-user data security with attack detection mechanism [29] | ShristiBhute, Siddhartha Kumar Arjaria | AES and RC6 | In this approach the data can be uploaded on the any one of the available server from the total three servers. One server is used for backup and data pre-processing by the cloud server. So there are total of four servers. Three servers for data uploading and one server are dedicated for the cloud server |
| A Lightweight Virtual MachineImage Deduplication Backup Approach in Cloud Environment[37] | Xu, J., Zhang, W., Ye, S., Wei, J., & Huang, T | Finger print clustering and sampling method in virtual machine deduplication | Native deduplication method for speeding up the process and progress of VM. Numerical method is used for computing size of sample space. |

**Table- i: Comparative analysis of various anti collusion and data deduplication schemes**

## IV. SYSTEM MODEL

The proposed system performs authentication of the user based on the password and the security image provided by the user at a time of registration. The user selects four parts of the randomly split image. This image is verified every time the user requests to login. Once the user gets into the account, the user may requests the Group Manager to gain access to a group. When a user wants to access a group, the client sends a request to the group manager along with a private key generated by the system. Once the key is verified by the group manager, then the user’s key can be activated to use within the group.

This system uses a fine-grained access control to prevent revoked users from gaining access to confidential files. The group manager executes the followingoperation when a fresh user newly joins the group or when a user has out of the particular group,

1. Update the all the user name-list.
2. Generate a secure key and encode the key without activation and send to the updated user list.
3. Update the user rights on the cloud server.

Advantages of our works are the implementation of de-duplication concepts and the reduction of the storage expenses in a cloud for the data owners can be done by integrating algorithms/techniques.

Secure protocol for anti-collusion attack reduces the processing time of load balancer and effective and efficient usage of cloud storage space[22].This system performs authentication on two levels with both password and a security image. The image is used as a key using Attribute Based Encryption (ABE) [5] with the image as the attribute (Fig-2). This prevents hackers and other malicious programs from accessing the user's account. Once the user has access to the user login, he/she may request access to a group within the cloud to upload/download a file of choice. After verifying the request, the group manager responds with a group key that has to be submitted by the user to upload/download a file within the group such that the file is encrypted/decrypted using RC6 algorithm. The user needs to use the key within the group to decrypt the file before downloading it.

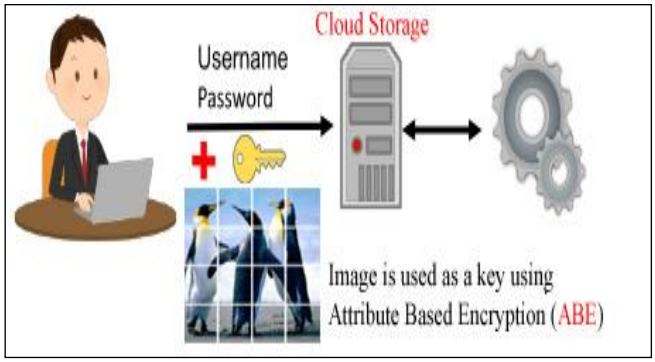



**Fig 2: Two levels of authentication with ABE.**

To perform de-duplication, the system checks for 30 percent uniqueness in among the files uploaded in a group and informs the user when the file is redundant. The below figure illustrates the architecture of the system.

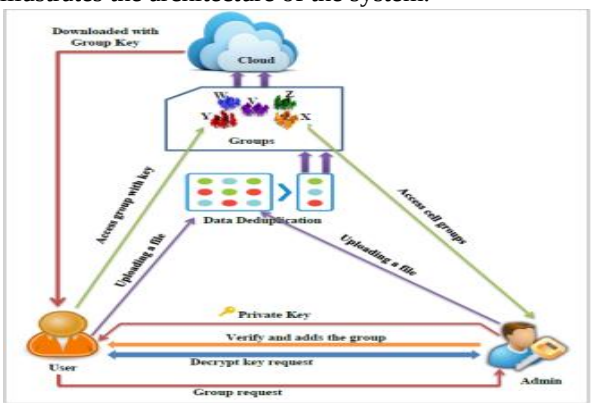



**Fig 3: System Architecture**

This system architecture (Fig–3) demonstrates a direct interaction between the users and the group manager,

## V. DESIGN AND IMPLEMENTATION

The Implementation of the proposed system is separated into modules to achieve an efficiency that satisfies the objective of the product development. The system modules are as follows

A. Image Based Authentication
B. Privacy-preserving
C. Key Distribution & Access Control
D. De-Duplication

### A. Image Based Authentication

Image based authentication[33] follows Attribute Based Encryption(Fig–4) a new user assigns an image for authentication at time of registration. The assigned image is split into 16 parts using a Pseudo Random Generator Technique and each part is given a numeric value to reduce complexity in authentication using image. The user then selects four parts of the image, which is used by the administrator to authenticate a valid user every time a file is uploaded or downloaded from the cloud.

**Algorithm: Image Based Authentication with ABE**

**Key in Values :*Idvi, pubk, vi***
**Output :** Decrypted message ( ***ENCpk (KEY, vj)*** )

1. Input ***Idvi, pubk, vi***
2. Initialize *Idvi, pubk, vi* send to group manager
3. **do**
4. Group manager generates a random number ***r of size n=8.***
5. Encryption done using ***ENCpk (KEY, vj)***and stored in disk
6. **if** *ENCpk (KEY, vj)==ENCsk (Idi, vi, ac)*
7. Decrypt the message
8. **else**
9. Don't decrypt the message
10. **break**
11. **endif**
12. **while** input is not empty
13. Outputdecrypt the message

**Fig 4: Image Based Authentication with ABE**

- In ABE, the user initially sends ***Idvi, pubk, vi*** as a request to the group manager, where ***Idvi*** is the identity value of the respective user, ***pubk***is the generated public key.
- The group manager generates a random number ***r of size n=8.***
- The encryption is then done using ***ENCpk (KEY, vj)*** and is stored in the local storage space.
- The received **ReIdi** message is compared with the identity **Idi** by decrypting ***ENCsk (Idi, vi, ac).***
- At last, the user gets to decrypt the message ***ENCpk (KEY, vj)*** by using his private key.

### B. Privacy-preserving

The privacy-preserving module establishes a reasonably secure protocol to provide the following requirements. This module maintains user privacy and data confidentiality by making sure that any M+1 user in a group of M participants is denied the access to a file within the group (Fig-5).

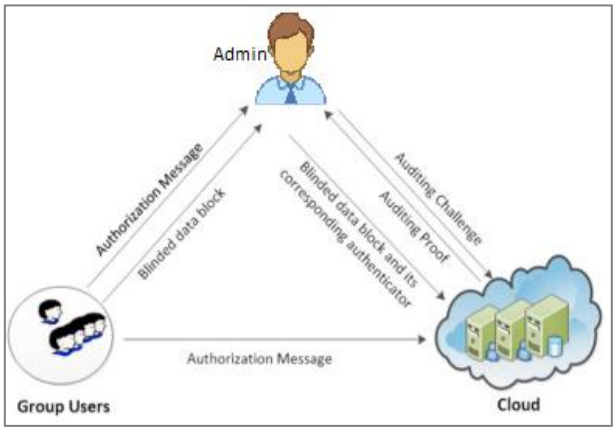



**Fig 5: User privacy and data confidentiality**

1. **Authentication**- A legal user can access their own data field, only the authorized partial or entire data fields can be identified by the legal user, and any forged or tampered data fields cannot deceive the legal user.
2. **Data Privacy**- Any unrelated user will not be able to identify the exchanged information and message state even it intercepts the exchanged messages via an open channel.
3. **User Privacy**- Unauthorized users cannot know or guess a user's id or password, which represents a user's interest in another user's authorized data fields. If and only if the both users have mutual interests in each other's authorized data fields, the cloud data server will inform the two users to understand the access permission sharing.
4. **Forward Security**- Any challenger cannot correlate two communication sessions to derive the prior cross-examinations according to the presently captured messages.

### C. Access Control&Key Distribution

Group manager having control for generate system parameters, user registration, and user revocation. In this application, the group manager typically is the controller of the group entire group (Fig-6). Hence, the group manager is reliable by the other parties. The group Manager is responsible for generating a random key to perform encryption and decryption of user data using RC6 algorithm.

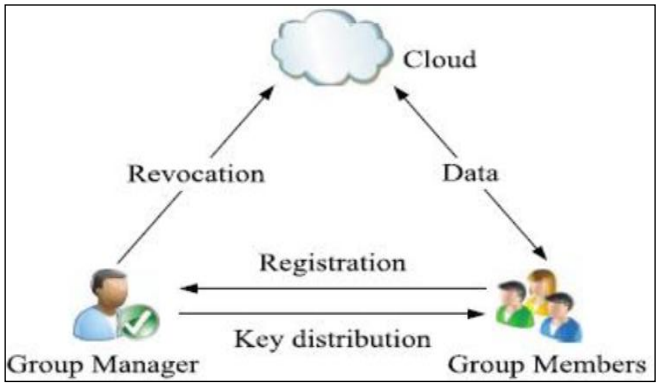



**Fig 6: Key distribution and Access control**

The group manager also checks the group for any revoked user and updates a new key within the group and activates the authorized users. RC6 is a symmetric key block cipher that follows data-dependent rotations, modular addition, and XOR operations. Our system uses two parallel encryption processes and although RC6 [28,29] use an extra multiplication operation which is not present in RC5 in order to produce the rotation dependent on every bit in a word, and not just the (LSB)least significant bit to achieve encryption and decryption (Fig-7). Normally RC6 is uses key size 256 bits and block size 128 bits with the time complexity of O(n). After the user is revoked, the group administrator creates the new encryption key for the particular group and transfers it to the members of the group. Secondly, the group manger updates the entire data list in the cloud server database. Thirdly, the group administrator updates the user list and activates the key for the user access.

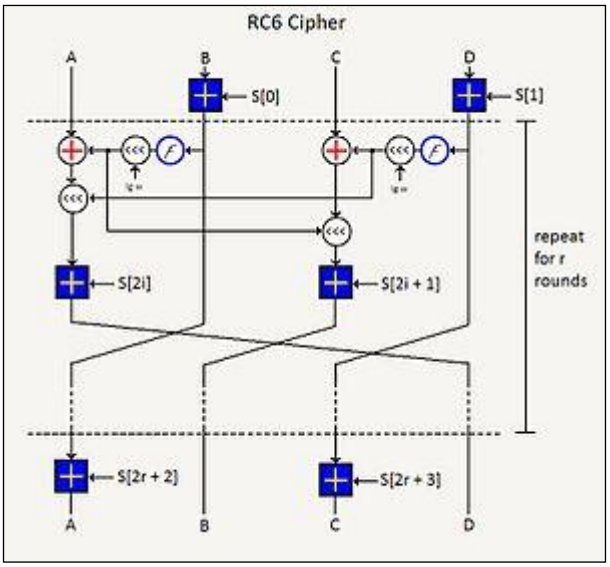



**Fig 7:General Diagram of RC6 Algorithm**

This proposed algorithm is divided into three phases, which are:

**A.** The key expansion algorithm
**B.** Encryption process
**C.** Decryption process

### A. The key expansion algorithm

Thus the key expansion algorithmis generates the user provided key to fill an array S.Here the notation ***S*** like a multiplicity of ***T***random binary words, more important user mustto give a key of ***B*** bytes, where $0 \le B \le 255$, and from which words($2r$+4) are indirect and put in a S(round key array), Zero bytes are attached to provide the key length equivalent to a "Non -zero integral number".

| ***RC6-Key expansion algorithm*** |
|---|
| **INPUT:**<br>*User-supplied b byte key preloaded into the c-word array L[0, ..., c - 1]*<br>**Number r of rounds**<br>*Pw = Odd((e – 2)2w)*<br>*Qw = Odd((ø – 1)2w)* |
| **OUTPUT:**<br>w-bit round keys S[0,…, 2r + 3]<br>**Procedure:**<br>S[0] = Pw<br>for i = 1 to (2r + 3) do<br>S[i] = S[i _ 1] + Qw<br>A = B = i = j = 0<br>v = 3 x max{c, 2r + 4}<br>for s = 1 to v do<br>{<br>A = S[i] = (S[i] + A + B) <<< 3<br>B = L[j] = (L[j] + A + B) <<< (A + B)<br>i = (i + 1) mod (2r + 4)<br>j = (j + 1) mod c<br>} |

**Fig 8:Key expansion algorithm**

The key bytes are then arranged into a cluster ***L*** of size ***c***:*when b= 0, c=1 and L[0]=0, e ="2.718281828459" and ø = "1.618033988749", $P_w$ and $Q_w$* are "Magic Coefficients" and Odd(x) is the least odd integer >= x , The (2r+4) determined words are put in array ***S*** for later encryption and decryption , Below figure illustrates the algorithm of the key expansion utilized in RC6 (Fig-8).

**B. Encryption process**

Encryption process starts after completion of key expansion process, when the users demand to encrypt their data our system will applied encryption algorithm and stores the encrypted data of the users in database, the algorithms of this phase are shown in details below (Fig-9).

| *RC6-Encryption algorithm* |
|---|
| **INPUT:**<br>Plaintext stored in four w-bit input registers A,B,C,D<br>Number r of rounds<br>w-bit round keys S[0,…,2r + 3] |
| **OUTPUT:**<br>Cipher text stored in A,B,C,D<br>**Procedure:**<br>B = B + S[0]<br>D = D + S[1]<br>for i = 1 to r do<br>{<br>t = (B x (2B + 1)) <<< log2 w<br>u = (D x (2D + 1)) <<< log2 w<br>A = ((A t) <<< u) + S[2i]<br>C = ((C u) <<< t) + S[2i+ 1]<br>(A,B,C,D) = (B,C,D,A)<br>}<br>A = A + S[2r + 2]<br>C = C + S[2r + 3] |

**Fig 9: Encryption algorithm of RC6**

**C. Decryption process**

The system will use decryption algorithm when user demand to decrypt the data to retrieve the plain text (Information) from the encrypted data that stored in database. The algorithms of this phase are shown in details below (Fig-10).

| *RC6- Decryption algorithm* |
|---|
| **INPUT:**<br>Cipher text stored in four w-bit input registers A,B,C,D<br>Number r of rounds<br>w-bit round keys S[0,…,2r + 3] |
| **OUTPUT:**<br>Plaintext stored in A,B,C,D<br>**Procedure:**<br>C = C - S[2r + 3]<br>A = A - S[2r + 2]<br>for i = r down to 1 do<br>{<br>(A,B,C,D) = (D,A,B,C)<br>u = (D x (2D + 1)) <<< log2 w<br>t = (B x (2B + 1)) <<< log2 w<br>C = ((C - S[2i + 1]) >>> t) u<br>A = ((A - S[2i]) >>> u) t<br>}<br>D = D - S[1]<br>B = B - S[0] |

**Fig 10: Decryption algorithm of RC6**

**D. Data De-Duplication**

Our system done data deduplication to removing duplicate copies of redundant data in cloud storage. The new mechanism is offers higher storage consumption and it reduce network load in computer network. Instead of storing more than one data copies with the duplicate content, data deduplication eliminates redundant data by keeping only one physical copy of data and mentioning others are duplicated data to that copy. Deduplication done in two levels, one is file level and another one is block level. First the file level data deduplication removes duplicate copies of the same file. Secondly block level data deduplication removes duplicate blocks of data that occur in congruous files which is stored in cloud server. The block data in the block-level deduplication can be measured as file data in the file-level deduplication and the idea of file-level deduplication can simply be used to strategy then block-level scheme. File level data deduplication is easy to index and very less CPU usage. Thus, we only consider the file-level deduplication in this work (Fig-11).

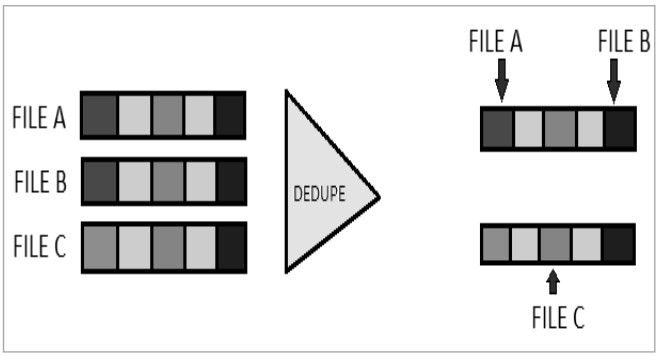



**Fig 11: File level Deduplication**

**Deduplication** = (Setup, Encrypt, Group-Encrypt, Group-Decrypt, Decrypt)consists of the following algorithms:

**Setup:** The setup algorithm chooses a cyclic multiplicative group G with prime order p . Let g ∈ Z p be a generator of G .

Hence ***h: {0,1} *→ {0,1} n*** be a cryptographic hash function. Let the public parameters be PP = (p, g, G ,h) .

**Encrypt (M) :** The encryption algorithm takes the plaintext M as input, then the cloud user runs the key generation algorithm to get the key K ←RCE.KeyGen ( M ), runs the opening algorithm to compute C = C 1 || C 2 || T ← RCE.Encrypt ( K, M ) and uses the ciphertext C to generate tag T ←RCE.TagGen ( C ). In addition, the cloud user randomly chooses an element x ∈ Z p as his secret key and computes X = g x mod pas his public key. Finally, the encrypt algorithm outputs ciphertextC , tag T and public key X .

**Group-encrypt (C) :** For the ciphertext C = C 1 || C 2 || T , the cloud server firstly generates the group key G . Then the cloud server re-encrypts ciphertext C 1 by C1_ = E H(G ) (C1) and constructs C _ = C_ ||C2||T .

**Group-decrypt (U, C_):** The cloud user firstly uses auxiliary information U to recover group key G = U × X mod p. Then the cloud user uses group key G to decrypt C1_ by C1 = D h(G ) (C 1_ ) . Now the cloud user gets the cipher text C = C 1 || C 2 || T.

***Decrypt* (K, C*) :*** The cloud user runs the opening algorithm to generate the tag T ← RCE.TagGen ( C ) and the

original data M ←RCE.Decrypt ( K, C ). If the decryption algorithm returns M , the cloud user accepts M ; if the algorithm returns ⊥ , the cloud user drops the message.

## VI. EXPERIMENTAL RESULT & DISCUSSION

This section, the experimental result shows the evaluation of our proposed technique. In the registration phase the user i choose his identity ($ID_i$), password ($PW_i$) and other detail with security image for registration (Fig-12).

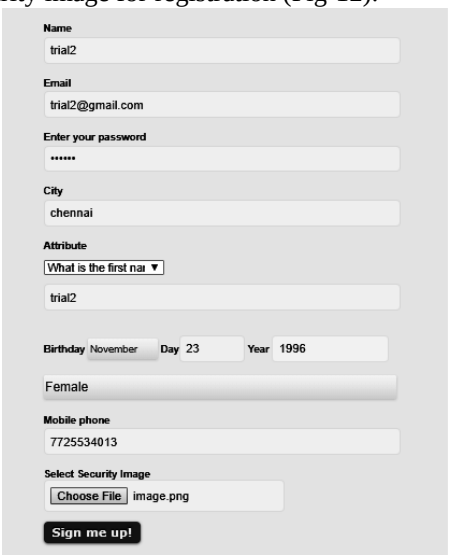



Fig 12: Registration phase

The Fig-13 illustrates image split using the image attributes such as the image width and the image height. The user selects 4 parts as a key for authentication using ABE.

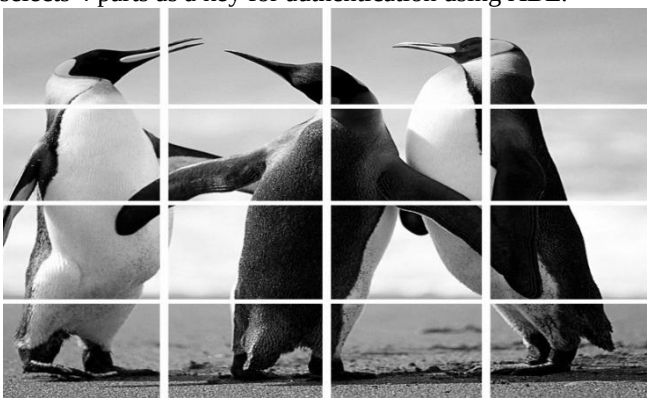


Fig 13: Image Split

After logged into group the user have to request group login request to group manager.


**Fig 14: Group Login Request**

The Fig-14 illustrates the request to access a group by the user. The group manager verifies the user and activates a group access key for the user. The user can use this activated key to gain access to a group using a group login.

The Fig-15illustrates the group login page where the activated group key is the password. The group key is of length n=8 generated at random to ensure security. When the fields of the group login match the entries of the database the user logs in successfully.

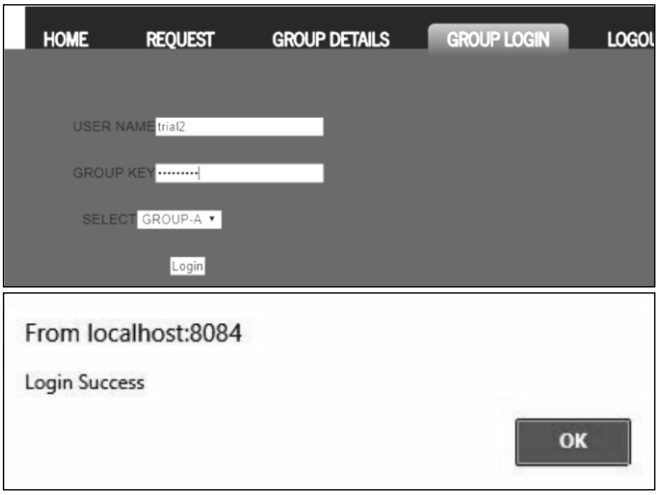



**Fig 15: Group Login**

The file upload and deduplication shown in Fig 16, it could be understood that only unique files get uploaded in a group. The file uploaded is encrypted using RC6. When the user tries to upload the same file within a group the following warning is seen.

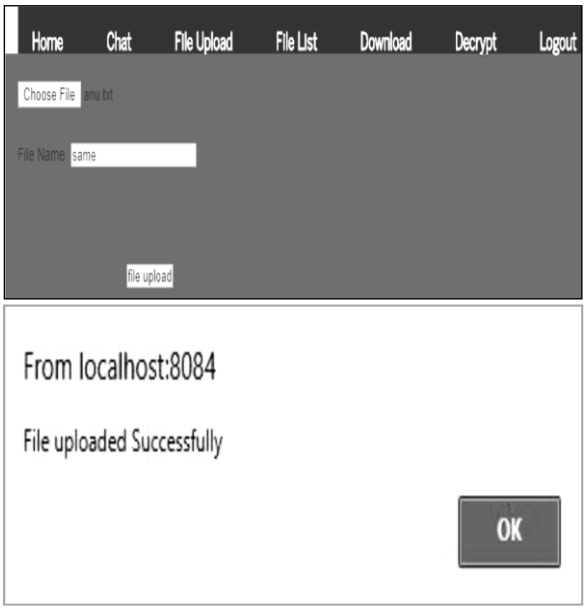
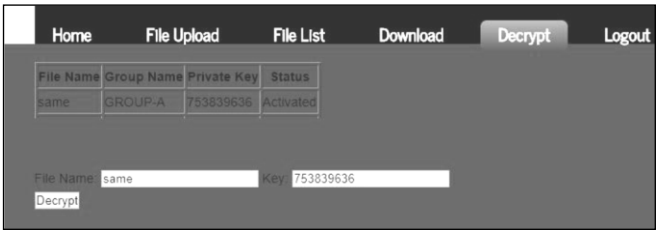



**Fig 16: File Upload is encrypted using RC6**

The De-duplication algorithm checks for 30 percent uniqueness in the documents being uploaded in the group.



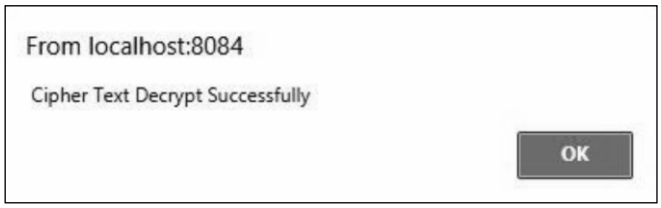



**Fig 15: File Decryption is encrypted using RC6**


**Fig 16: File Download**

## VII. PERFORMANCE EVALUATION

In this section, we compare the evaluation performance of the proposed Image based authentication with data deduplication scheme with previous works. Each cryptographic process is implemented by using the Openssl library ver. 1.1.0 and Crypto++ library ver. 8.2.

### A. Security Performance Comparison

Security Performance Comparison of various schemes like Mona, RBAC, ODBE, Secure Anti Collusion Data Sharing, DedupDUM with our scheme (Table-ii), the data like secure data confidentiality, data deduplication, access control key distribution, secure user revocation and anti-collusion attack,. The ✓ in the blank means the approach can achieve the corresponding goal. Our scheme Image based authentication with deduplication schemeachieves the entire corresponding goal after testing of our work.

### B. Encryption and decryption time Comparison

We test the time cost of data encryption and decryption algorithm with RC6 where the key is 256-bit. The data size ranges from 10MB to 100MB. We set CK = 256 bits, CT = 256 bits, prime order p = 2048 bits. The testing environment is Intel(R) Core(TM) i5 CPU 3.40 GHz 16.0GB RAM, Ubuntu v12.04.2 16.0GB RAM, Octa-core processor and 20.0GB Hard disk .

| Scheme | Secure key distribution | Secure user revocation | Anti-collusion attack | Data confidentiality | Data Deduplication | Access control |
|---|---|---|---|---|---|---|
| Mona | | | | | ✓ ✓ | |
| RBAC scheme | ✓ | ✓ | | | ✓ ✓ | |
| ODBE ✓ | ✓ | ✓ | ✓ | | | |
| Secure Anti Collusion Data Sharing | | | | | | |
| ✓ DedupDUM ✓ | ✓ ✓ | ✓ | ✓ | ✓ ✓ ✓ | | |
| Our Scheme | | | | | | |

**Table – ii: Security Performance comparison with our scheme.**

Our scheme Image based authentication with deduplication is based on the RC6encryption and decryption. In the encryption stage, the key generation time in our scheme is longer than the Secure Anti Collusion Data Sharing scheme. However, the RBAC scheme and other schemes are not considering the ownership changes in the ownership list. Compared with previous works, the Secure Anti Collusion Data Sharing scheme and our scheme resolve the dynamic ownership management problem. By using the rekeying mechanism, the backward secrecy and forward secrecy are protected. The detailed data encryption and decryption time for different data sizes (ranging from 10MB to 100MB) are shown in Fig-17.

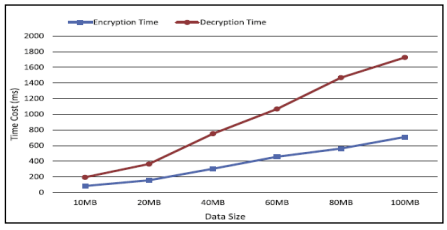



**Fig 17: Encryption and decryption time**

### C.Computation time for upload and download

The file upload and download processes with the size of 10 and 100 Mbytes computation time of the cloud is illustrated in Fig-18 and Fig-19. The computation time of the cloud is mainly determined by the key verification operation. Then, the cost increases with the number of revoked users. This scheme verifies the identities like image, user name and password.

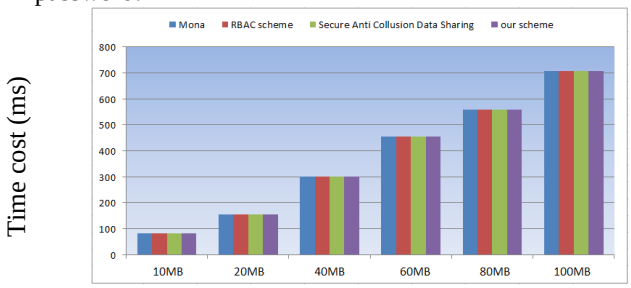



**Fig 18: Computation time for upload**

Then, the calculation cost of the cloud for file upload and download is unrelated to the number of the revoked cloud users. The reason for the large execution time of the cloud storage in RBAC scheme performs some operations to help the user to decrypt data files. So this RBAC algorithm takes additional computation time compare to all other scheme. In addition, the computation time is independent with the size of the file, since both the key in Mona and the encrypted information or data in our algorithm are not matched to the size of the requested file and the operations of cloud for decryption. In RBAC scheme is also unrelated to the size of the encrypted data files. The Secure anti collusion data sharing scheme and our scheme file uploading and downloading nearby same but Image based authentication and deduplication techniques additionally added with this process.

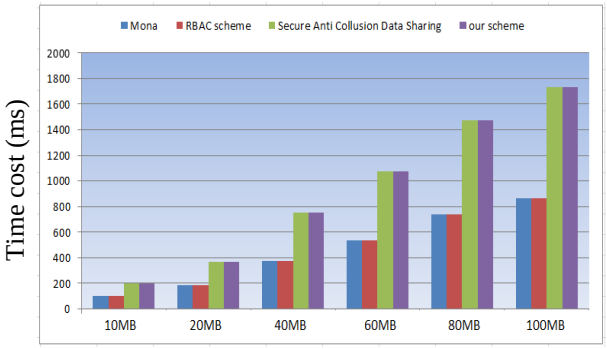



**Fig 19: Computation time for download**

## VIII. CONCLUSION

In this paper, we design an Image based authentication with deduplication scheme for cloud user group, our scheme provides security against various threats in a cloud. It accomplishes user authentication using image, de-duplication of the files being uploaded & prevents collusion by updating the data base whenever a user is revoked thus generating a new key for encryption and decryption. De-Duplication in cloud groups enables various industries to remove replication of data and thus improve storage capacity. Providing the user with an image based authentication adds a security feature to user authentication making it difficult for unauthorized users to gain access to a file. Removing the need for a secure channel to distribute the encryption key makes it less complicated and provides confidentiality.

## REFERENCES

1. Subramanian Nalini and Andrews Jeyaraj. Recent security challenges in cloud computing. Computers & Electrical Engineering 71 (2018): 28-42.
2. Attasena, Varunya&Darmont, Jérôme&Harbi, Nouria, Secret Sharing for Cloud Data Security : A survey, The VLDB Journal. 10.1007/s00778-017-0470-9, June 2017.
3. Hiroaki Anada , Junpei Kawamoto , ChenyutaoKe , Kirill Morozov , Kouichi Sakurai, Cross-group secret sharing scheme for secure usage of cloud storage over different providers and regions, The Journal of Supercomputing, v.73 n.10, p.4275-4301, October 2017
4. Hattim, May &Taha, Zahraa. (2019). Secure And Hidden Text Using Aes Cryptography And Lsb Steganography. Vol. 14, No. 3 (2019) 1434 - 1450 © School of Engineering, Taylor's University.
5. Yan, Zheng, Mingjun Wang, Yuxiang Li, and Athanasios V. Vasilakos "Encrypted Data Management with Deduplication in Cloud Computing" IEEE Cloud Computing 3.2, pp no: 138-150. Web, 2016.
6. Tao Jiang, Jianfeng Ma, &Xiaofeng Chen, "Public Integrity Auditing for Shared Dynamic Cloud Data with Group User Revocation", IEEE Transactions on Computers, Volume: 65, Issue: 8, 2015.
7. Tao Peng, Qin Liu &Guojun Wang, A Multilevel Access Control Scheme for Data Security in Transparent Computing, IEEE Trans. Computing in Science & Engineering, Vol. 19, Issue 1, 2017.
8. Tchernykh, Andrei &Babenkob, Mikhail &Chervyakovb, Nikolay& Miranda-Lópeza, Vanessa &Kuchukov, Viktor & Cortés-Mendoza, Jorge &Deryabin, Maxim &Kucherov, Nikolay&Radchenko, Gleb&Avetisyan, Arutyun. (2018). AC-RRNS: Anti-collusion secured data sharing scheme for cloud storage. International Journal of Approximate Reasoning. 102. 10.1016/j.ijar.2018.07.010.
9. Tomar, Anurag& Shankar, Shashi Kant & Sharma, Manmohan&Bakshi, Aditya. (2016). Enhanced Image Based Authentication with Secure Key Exchange Mechanism Using ECC in Cloud. 625. 63-73. 10.1007/978-981-10-2738-3_6.
10. Puzio, P., M olva, M., &Loureiro, S., "ClouDedup: Secure deduplication with Encrypted Data for Cloud Storage", IEEE 5th International Conference on Cloud Computing Technology and Science, PP No:363-370. doi:10.1109/cloudcom.2013.54, 2013.
11. Lorena González-ManzanoAgustinOrfila "An efficient confidentiality-preserving Proof of Ownership for deduplication" Journal ofNetworkandComputerApplications 50 (2015) 49–59
12. V. Miranda-López, J. M. Cortés-Mendoza, A. Tchernykh, M. Babenko, G. Radchenko, S. Nesmachnow, Z. Du, Experimental Analysis of Secure and Reliable Schemes for Cloud Storage based on RNS, in: High Performance Computing. CARLA 2017 (CCIS), vol. 796, 2018, pp. 370-383.
13. Widodo, Ryan & Lim, Hyotaek&Atiquzzaman, Mohammed. (2017). A new content-defined chunking algorithm for data deduplication in cloud storage. Future Generation Computer Systems. 71. 10.1016/j.future.2017.02.013.
14. Wenhao Li, Yun Yang, and Dong Yuan, "Ensuring Cloud Data Reliability with Minimum Replication by Proactive Replica Checking", IEEE Transactions on Computers, Volume: 65, Issue: 5, May 2016.
15. Mazhar Ali, Eraj Khan, and RevathiDhamotharan, "SeDaSC: Secure Data Sharing in Clouds", IEEE Systems Journal, Volume: PP, Issue: 99, 2015.
16. Rahul S. Nandanwar, Vijendrasinh P. Thakur, "Review On Secure Anti-Collusion Data Sharing For Dynamic Group In The Cloud", IJECS, Volume 5 Issues 6, Page No. 16886-16888, June 2016.
17. Cheng, Minquan& Miao, Ying. (2011). On Anti-Collusion Codes and Detection Algorithms for Multimedia Fingerprinting. Information Theory, IEEE Transactions on. 57. 4843 - 4851. 10.1109/TIT.2011.2146130.

18. NareshVurukonda& B. ThirumalaRao, "A Study on Data Storage Security Issues in Cloud Computing", ICCC, page no-128-135, 2016.
19. XuAn Wang, ZhihengZheng&FatosXhafa, "Identity Based Proxy Re-Encryption Scheme (IBPRE+) for Secure Cloud Data Sharing", International Conference on Intelligent Networking and Collaborative Systems (INCoS), 2016.
20. SayaleeShinde& Dr. S. S. Shaikh, "A Survey on Secure Anti-Collusion Data Sharing Scheme for Dynamic Groups in the Cloud", IJIRCCE, Vol. 4, Issue 12, December 2016.
21. Zhu Zhongma& Jiang Rui. (2015). A Secure Anti-Collusion Data Sharing Scheme for Dynamic Groups in the Cloud. IEEE Transactions on Parallel and Distributed Systems. 27. 1-1. 10.1109 / TPDS.2015.2388446, 2016..
22. G.MercyVimala, R.Vara Prasad, P.RamaRao, "A Secure Anti-Collusion Data Sharing Scheme for Dynamic Groups in the Cloud", International Journal of Computer Engineering In Research Trends, Volume 2, Issue 12, December-2015, pp.1113-1118.
23. T.D.B Weerasinghe, Software Engineer, IFS R&D International, "Analysis of a Modified RC4 Algorithm", International Journal of Computer Applications (0975 – 8887) Volume 51– No.22, August 2012.
24. Jayan, A. & Nair, A. &Upadhyay, Bhargavi&Muthuraman, Supriya. (2016). Performance analysis of modified RC4 cryptographic algorithm using number of cores in parallel execution. International Journal of Control Theory and Applications. 9. 225-231.
25. PushpendraVerma, Dr. JayantShekhar, Preety, AmitAsthana, "A Survey for Performance Analysis Various Cryptography Techniques Digital Contents", International Journal of Computer Science and Mobile Computing, Vol.4 Issue.1, January- 2015, pg. 522-531.
26. Yuan, Haoran& Chen, Xiaofeng& Jiang, Tao & Zhang, Xiaoyu& Yan, Zheng& Xiang, Yang. (2018). DedupDUM: Secure and Scalable Data Deduplication With Dynamic User Management. Information Sciences. 456. 10.1016/j.ins.2018.05.024.
27. Zhou, Yukun&Feng, Dan &Hua, Yu & Xia, Wen & Fu, Min & Huang, Fangting& Zhang, Yucheng, A similarity-aware encrypted deduplication scheme with flexible access control in the cloud. Future Generation Computer Systems. 84. 10.1016/j.future.2017.10.014.
28. Salim Ali Abbas , Malik Qasim Mohammed , Enhancing Security of Cloud Computing by using RC6 Encryption Algorithm, International Journal of Applied Information Systems (IJAIS) , Volume 12 – No. 8, November- 2017.
29. ShristiBhute, Siddhartha Kumar Arjaria , An efficient AES and RC6 based cloud-user data security with attack detection mechanism, International Journal of Advanced Technology and Engineering Exploration, Vol 3(21), 2016.
30. Prof. Dr. Salim Ali Abbas, Malik Qasim Mohammed, Improving Data Storage Security in Cloud Computing Using RC6 Algorithm, IOSR Journal of Computer Engineering (IOSR-JCE), Volume 19, Issue 5, (2017), PP 51-56.
31. Aditya Poduval1, AbhijeetDoke, Hitesh Nemade, RohanNikam, Secure File Storage on Cloud using Hybrid Cryptography, International Journal of Computer Sciences and Engineering , Vol.7(1), Jan 2019.
32. Luo, Shengmei, Guangyan Zhang, Samee Khan, Chengwen Wu, and Keqin Li. "Boafft: Distributed Deduplication for Big Data Storage in the Cloud." IEEE Transactions on Cloud Computing (2015) ,pp no: 1- 13. Web.
33. Tomar, Anurag&Tak, Gaurav&Chaudhary, Ruchi. (2015). Image based authentication with secure key exchange mechanism in cloud. 428-431. 10.1109/MedCom.2014.7006046.
34. Waters, Brent. (2008). Ciphertext-Policy Attribute-Based Encryption: An Expressive, Efficient, and Provably Secure Realization. IACR Cryptology ePrint Archive. 2008. 290. 10.1007/978-3-642-19379-8_4.
35. C.-M. Yu, Efficient cross-user chunk-level client-side data deduplication with symmetrically encrypted two-party interactions, in: Poster of Proceedings of ACM CCS, Vienna, Austria, 2016, pp. 1763–1765
36. Li, J., Li, J., Xie, D., &Cai, Z., " Secure Auditing and Deduplicating Data in Cloud", IEEE Transactions on Computers, 65(8), 2386-2396. doi:10.1109/ tc.2015. 2389960, 2016.
37. Xu, J., Zhang, S., Wei, J., W., Ye, & Huang, T.," A Lightweight Virtual Machine Image Deduplication Backup Approach in Cloud Environment" IEEE 38th Annual Computer Software and Applications Conference, 503-508. doi:10.1109 / compsac. 2014.73,2014.


## AUTHORS PROFILE

Mr.T.A.Mohanaprakash is a Research Scholar in Sathyabama Institute of Science and Technology. He received her M.Tech degree in Information Technology from Sathyabama University, Chennai, India in 2009. He has 14 years of teaching experience. He is currently working in Panimalar institute of Technology, India. His area of research interest includes cloud computing, network security and Web Technology.

Dr.AndrewsJeyaraj received Ph.D degree in 2014 from Sathyabama University in the area of code optimization. He has published more than 42 research papers in referred international and national journals. His research interest includes machine learning, compiler design, operating system and Deep learning networks. He works currently as a Associate Professor in the Department of CSE, Presidency University, Yelahanka, Bengaluru, Karnataka, India and has more than 16 years of teaching experience.