Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 143 (2018) 765–775

Procedia Computer Science

www.elsevier.com/locate/procedia

8th International Conference on Advances in Computing and Communication (ICACC-2018)

A New Security Framework for Cloud Data

ShaluMall[a], Sushil Kumar Saroj[b]

[ab]*Department of Computer Science and Engineering, Madan Mohan Malaviya University of Technology Gorakhpur, UP, India*

Abstract

In many organizations and institutions, the use of cloud has increased rapidly. Cloud data storage is one of the main advantages of cloud computing, where the data owners do not store their data on own servers, but the data are mainly stored on the cloud by the data owners. On this end, cloud security is one of the most analytical aspects due to the confidential information and responsive data. This paper presents a new security framework which provides more data security and confidentiality. In the new security framework, a data is spilt in the blocks of bits. Genetic algorithm is applied on every two blocks of bits. Final output of every genetic algorithm operation is a ciphertext which is also two blocks of bits. Each ciphertext is stored on cloud at distinct location and location of the ciphertext is not fixed. So, it is difficult for an attacker to detect where ciphertext is. Also, genetic algorithm has no key concept due to which security of data increases. The new security framework applies genetic algorithm on smaller block size which increases the security. The framework also uses the capability list for secure and fine grain access of data.

*Keywords:*Genetic Algorithm; Crossover; Mutation; Data Splitting; Cloud Storage; Outsourced Data;

1. Introduction

Cloud computing is very popular and quick developing technology in institution as well as organizations because it gives computing services and storage of data at very attractive cost.

* Corresponding author. Tel.:8765143703
*E-mail address:*shalu.mall8@gmail.com

10.1016/j.procs.2018.10.397

The advantages of using cloud computing technology including easy scalability, cost saving, and high availability. Now days, the cloud is very essential and key aspect among every technology which includes the identity management virtualization security, application integrity, network security and data protections. In the above scenario, data protection is very essential in cloud
computing. In cloud computing, there are three types of service model namely-

- IaaS(Infrastructure as a Service)- In IaaS, users get resources like CPU time, network bandwidth, processing power and storage. Once the customer gets the infrastructure he may control the OS, application, data host-based security, services etc.
- PaaS(Platform as a Service)- In PaaS, users are provided the hardware infrastructure, OS and network to make a hosting environment. From the hosting environment, user can activate services and install his applications.
- SaaS (Software as a Service)- In SaaS, users are provided the access to an application. They have no restriction over the network, hardware, OS or security.

The Cloud computing has five major features: network access, on demand self-service, resource pooling, rapid elasticity, location independent. These all characteristics made the cloud significant. Institutions and industries are increasing their revenue and profited by exploiting these cloud computing characteristics [2]. This is the reason; industries are moving their business to cloud. But security of data is a major restriction in cloud computing. In present time, cloud security is one of the biggest critical issues in the environment of cloud computing due to the sensitive data of data owner (DO). So, cloud service provider (CSP) must consider security and privacy issues in highpriority.

The data of DO is prepared and saved on outside servers. So, integrity, confidentiality as well as data access control become more crucial and essential. Since, the outside servers are managed through monetary service providers, DO cannot trust on them as they can use its data for their profit and can spoil the business of DO. Even, DO cannot trust on clients or users associated to it, as they can be spiteful and malicious. Confidentiality of sensitive data can be breached over service providers. There are some frequent strategies provided to protect data although they are suffering from several issues. Here, we present a new security framework to secure the data of DO which is stored on cloud [2][8][9][10].

1.1. Why we should store data on cloud?

If, DO store own data at its side and not on the cloud then it may not be cost effective always. Because, DO are not individuals or smaller organizations always. It may be large organizations like universities, banks, hospitals, big companies and even government organizations. They have large amount of data. To store, maintain, distribute and secure the data by DO itself may not be cost effective. CSP are big organizations. They host bulk data. They are doing business to store, maintain and secure large amount of data every day. So, they provide storage facility at very low cost. If, DO store its data at its system then system can be theft, break, crash and destroy at any time due some reason (by attackers, flood, earthquake, fire, virus etc.) but CSP store data at multiple locations (data centres) by making replica of data as backup. Data can access from another data centres if any adverse situation arises. DO have to see the security as well as cost that is why DO store data at CSP not its side.

In this paper, the rest parts are systematized as follows. Here in section II, the related works are analysed. In section III, we give the communication model and presumptions. In section IV, we present the new security framework. In section V, the new security framework is analysed and simulated. Finally, in section VI we conclude the paper.

2. Related Work

Data access control and confidentiality are two necessary security measuresfor outsourced data. Once, when we

indicate further on security of data, we not be able to remember about systems performance (CSP, DO, Users). For example, sometimes we use more keys to secure the data. To store, maintain, secure and distribute the keys are total computational overhead. Generally, keys are either stored by DO or by Third-Party Auditor (TPA). But there are many alternatives to easily access data in cloud by using the keys stored in TPA. And we should also ensure the level of confidentiality and access control. Remarkably, there is a strategy needed that not only provides data security but also maintain the work of the system. There are many strategies are given below to secure the data.

Scheme proposed in [3] using Third Party Auditor, hash function and RSA. In this scheme, the third-party auditor (TPA) is considered in active and performs all the computations and verifications. It is known that we cannot fully trust on TPA's, that it can use the data of DO for own financial profit. Another improvement field in proposed scheme [3] is breaking the RSA much simple than factoring [4].

Scheme proposed in [5] is reliable to earn access control and confidentiality of data. In this strategy, the author encrypts data through secret- keys and these keys are only recognized to DO and corresponding data users. The encrypted files are stored at CSP. During the communication between CSP and the user, data are further encrypted through one-time private session key which is shared among CSP and the user through the modified Diffie-Hellman protocol. This scheme provides full data security but there is a key corresponding to every file and user but in some applications, files and users may be large in numbers. So, there may have a large number of keys.

Scheme proposed in [6] using Shamir sharing algorithm with CRT (Chinese Remainder Theorem) which assigns the key and shares the key among participant resource providers. Here, numbers of keys are reduced but the scheme has no arrangement of managing data.

Scheme proposed in [7] using the cryptographic data splitting mechanism with AES algorithm which divide the user's encrypted file into two parts and stored on public cloud but there is no acknowledgement related to key. The proposed scheme somehow matches with schemes [7][11][12] but is much safer.

3. Representations and Presumptions

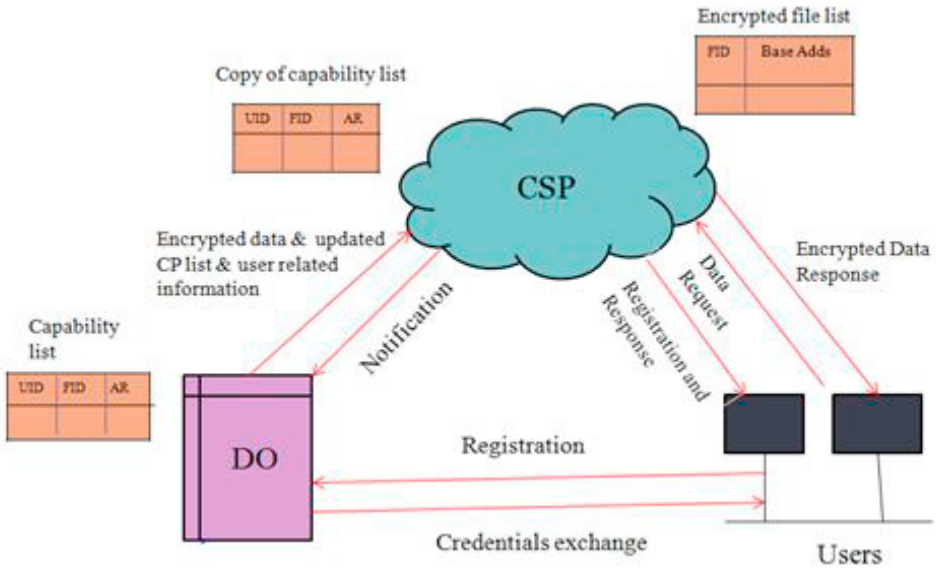


Fig.1. Representation of proposed communication method [1]

We assume that the proposed scheme consists of three entities DO, CSP and many users associated with DO. Firstly, every user is certified at DO. For the period of certification, users send their credentials/ required information to DO. We suppose that the user's credential is sent to DO safely during registration. And after successful registration, DO send required information (pseudorandom number and information related crossover and mutation operations) to the user safely in response.
We also assume that DO has some processing capability and space to store some data. DO stores its data to CSP. After successful authentication of the user to CSP, the user can retrieve data from CSP in some secret manner. We assume that CSP knows and has mechanisms to locate and store the data. CSP are big organizations they have such type of mechanisms. Fig.1 shows the communication among DO, CSP and Users.

Nomenclature

GA	Genetic Algorithm	FID	File Identity
DO	Data Owner	AR	Access Right
CSP	Cloud Service Provider	CP List	Capability List
AES	Advanced Encryption Standards	TPA	Third Party Auditor
RSA	Rivest Shamir Adleman	PRN	Pseudo Random Number
E	Encryption	PU_{CSP}	Public Key CSP
D	Decryption	PR_{CSP}	Private Key of CSP
UID	User Identity	PR_{DO}	Private key of DO

4. Proposed Method

4.1. The new security framework

In the proposed method, we provide the new security framework for data that is stored on the cloud. In this security framework, a data is converted into ASCII values. Then these ASCII values are converted into binary bits. Binary bits then divided into blocks of bits of some size. Block size may be 8 bits, 4 bits, 2 bits etc. Here, genetic algorithm (GA) operations (crossover and mutation) are used for encryption and decryption process. GA operations (crossover and mutation) are applied on each pair of blocks. The final output of each GA operations is a ciphertext which also pair of blocks of bits. Each ciphertext is stored on cloud at distinct location and, location of ciphertext is not fixed (for example, output1 stored in the cloud has location l1 at time t1, may have location l2 at time t2) by some mechanisms.Since, encrypted data parts are stored on cloud, CSP can't see the data. Before sending to CSP, these encrypted data parts are further encrypted with private key of DO for DO authentication and, then encrypt the encrypted data with public key of CSP again so that attacker can't see the data and CPList. The whole procedure of the new scheme is shown in Fig.2.

Here, there are some terms or functions which are used in encryption or decryption process of data. And, also describe here how GA works.

4.2. Genetic Algorithm

Genetic Algorithm (GA) [11] is a collection of three processes i.e. replacement, selection and genetic operation (crossover, mutation). In this paper, only genetic operations (crossover, mutation) and pseudorandom number are used in encryption process of data which are described as follows:

- Pseudorandom Number-It is a random number which is used in decision to which crossover function should select for crossover operation.

- Pseudorandom Number Generator-There are a variety of techniques by which a random number is generated, however the generally used technique is multiplicative congruential generator. Following is the function which is used to generate pseudo random number-

$$x_{i+1} = x_i.c \text{ (modulo m)}$$

where x_{i+1} is the subsequent Pseudo Random Number(PRN) of x_i, c and m are +ve integer number, c is frequently multiply by x_i and the outcome is $x_i.a$ and it is divided by m. As far as the remainder comes less than m.x_0 is the first number by which we start calculating the PRN. In pseudo random number the new number is generated from previous one.Output of modulo operation on generated pseudo number decides which crossover operation should apply on two selected chromosomes or blocks of data [11].

4.2.1. Crossover

It is the process in which two blocks or chromosomes are taken to generate a new offsprings or children. There are mainly three crossover operations which are used on binary coded GA.

- One-point crossover-In one-point crossover two blocks (P1 and P2) are given, randomly two chromosomes are chosen and broken the blocks into half then tails of two chromosomes is exchanged to obtain new off springs (OS1 and OS2).

- Two-point or Multi-point crossover-It is related to one-point crossover excepting two cut points are created instead of one, then one part of every block or chromosome is exchanged to form new block or chromosome.

- Uniform crossover-In uniformcrossover, the bits are copied randomly from the 1st and 2nd point. In this operator we do not divide the blocks into pieces and the random mask is generated, and the mask determine which bit is copied from 1st point and which from 2nd point.

Random mask generated randomly i.e. 11010110

4.2.2. Mutation

Basically, the mutation is based on random changes; it changes 0 to 1 and vice versa. It is performed on two selected chromosomes or blocks.

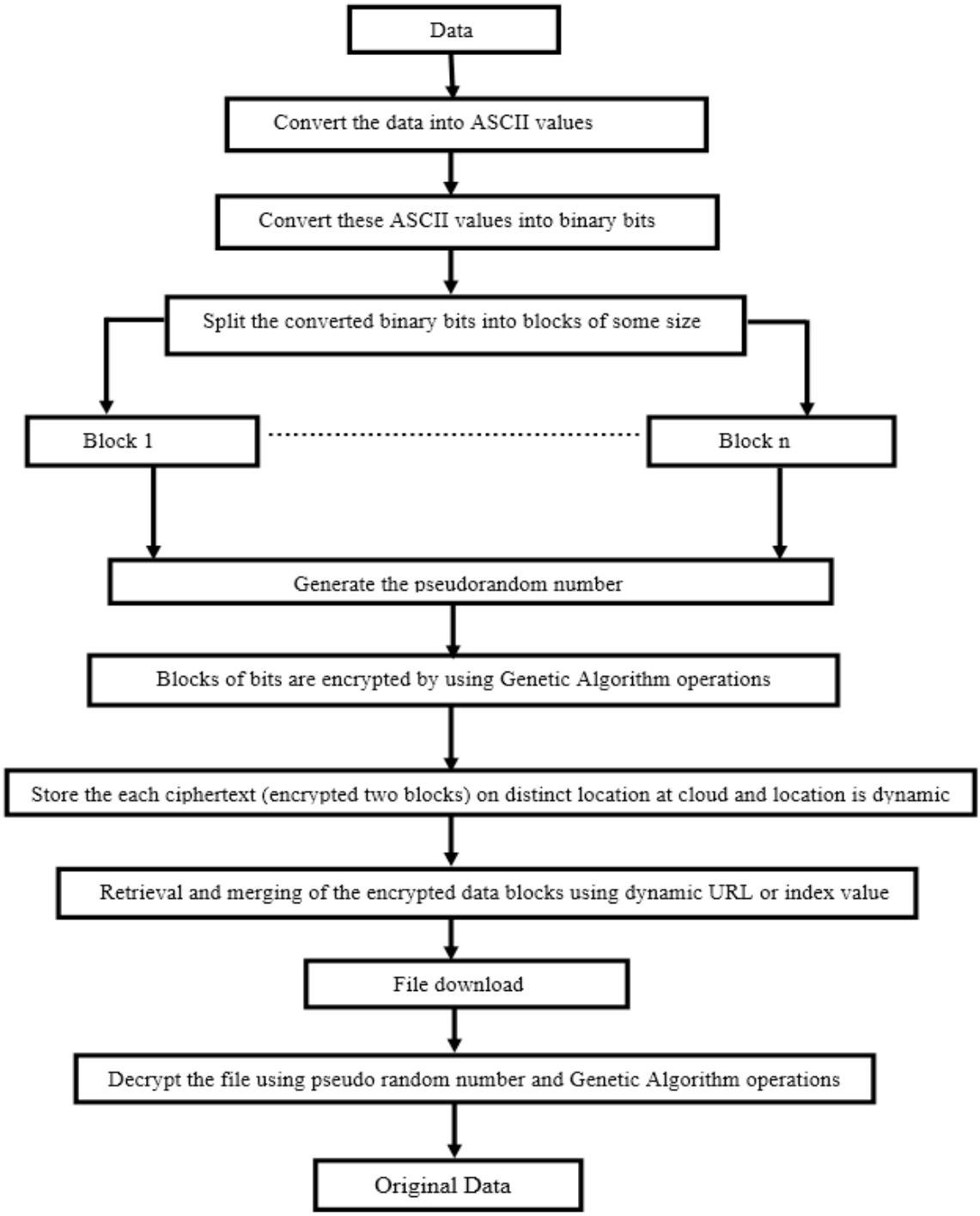


Fig.2. Security framework of proposed scheme

Algorithm 1.1: User Registration Process

Step-1: The User sends registration request to DO with his details

Send (User details)

Step-2: After successful registration, DO sends required information (random number, UID, FID, AR and few information about GA) to the User

Step-3: DO updates the capability list at its end

CPList = Add (CPList, (UID, FID, AR))

Step-4: DO encrypts the data and CP List and send it to the CSP

Send ($E_{PU_{CSP}}(E_{PR_{DO}}(E_{GA}(Data) \| CP\ List))$)

Step-5: CSP decrypts the received information from DO using his private key

$E_{GA}(Data) \| CP\ List = D_{PR_{CSP}}(D_{PU_{DO}}(E_{GA}(Data) \| CP\ List))$

Step-6: CSP updates the capability list at its end with entries sent by the DO

CPList = Add (CPList, (UID, FID, AR))

and store the encrypted data parts (each ciphertext) at distinct location on the cloud and the location where each ciphertext is stored, is dynamic

Step-7: Now, the User can precisely connect to CSP to fetch his data.

In the proposed scheme, we used four algorithms. Algorithm 1.1 describes the registration process [1]. Algorithm 1.2 describes the splitting process of data [12]. Algorithm 1.3 describes the encryption process of data using GA and Algorithm 1.4 describes the decryption process of data using GA [11].

Algorithm 1.2: Split Algorithm

Step-1: Data is converted into ASCII values first.

Step-2: ASCII values of data then converted into binary bits

Step-3: DO splits the binary bits into n blocks of some size

Step-4: Make the new folder and save these blocks of bits

Step-5: Generate pseudo random number

Step-6: Select two stored blocks of bits for encryption

Step-7: Apply GA on selected two blocks of bits for encryption

Step-8: Then store the encrypted data parts(ciphertext) on distinct location of cloud

Algorithm 1.3: Encryption Process

Step 1: Choose two blocks of bits from storage

i.e. s_1=10001110 s_2=10001011

Step 2: Generate pseudorandom number using the pseudorandom function for selected two blocks of bits. Apply modulo operation (mod of 3) on the generated pseudorandom number

Step 3: Crossover function is chosen according to the output of the mod function

If output = 0 then Single point crossover is applied

If output = 1 then two-point crossover is applied and,

If output = 2 then uniform crossovers is applied

Example: 23%3=2 (Uniform crossover), 63%3=0 (Single point crossover)

Step 6: Apply the chosen crossover function on selected two blocks of bits

Step 7: Apply mutation function on the output of crossover function i.e. also two blocks of bits

Step 8: Output of mutation function is ciphertext (two blocks of bits) which is stored on cloud at distinct location and the location where ciphertext is stored, is not fixed

4.3 .GA operations with smaller block size

Here, we apply encryption and decryption on blocks where each block size is 8 bits. We can also apply encryption and decryption process on smaller block size (4 bits or 2 bits etc.). If, we work with smaller block sizethen number of blocks corresponding to a data increase. Number of GA operations increase as number of blocks increase. So, to encrypt a data, more number of GA operations will require. So, ciphertexts corresponding to a data have more number of random bits. Hence, confidentiality of data increases as randomness increases.

4.4. Data Retrieval from cloud

When a user wants to access a data, he should first send data request to CSP. The user sends data request with information (UID, FID, AR) to CSP. CSP matches sent (UID, FID, AR) with stored (UID, FID, AR) for a data. If matches, authentication is successful. CSP then retrieves the data parts from location 1, location 2……. location n using dynamic URL. Then whole parts of a data are merged and stored in the user's server as cache, from where data is downloaded and decrypted.

4.5. Capability list

The new scheme uses the capability list (CP List) for secure and fine grain access of cloud data. Basically, CP List has UID, FID and AR entries corresponding to each data. It is basically row-based decomposition of access matrix. In CP List, operations and recognized data for a user are described.DO have a right to execute the activity and CSP read this activity for the objective of securely access the data.

Algorithm 1.4: Decryption Process

Step 1: One ciphertext is selected from the storage of ciphertexts of a data which are sent by CSP to the user after authentication

Step 2: Apply mutation function on the selected ciphertext in reverse order

Step 3: Read the random number that is sent to the user corresponding to that ciphertext and perform modulo function (mod 3) on the random number

Example: 63%3=0 (Single point crossover), 23%3=2 (Uniform crossover)

Step 4: According to the output of mod function, perform the crossover function on the output of mutation function in reverse order

Step 5: Output of crossover function is a plaintext (two blocks of bits)

Step 6: The process from Step 1 to Step 4 are repeated until all the ciphertexts corresponding to a data are not converted into plaintexts

Step 7: Convert the binary bits of data into ASCII values

Step 8: Convert the ASCII values into texts that is the original data

5. Performance and security Analysis

5.1. Analysis of Security

Here, we discuss about the strength of the new scheme and, security of outsourced data [8][9][10].

- **Data Confidentiality-**In the new scheme, DO encrypts its data itself with GA. Since, GA information are known only to DO and corresponding data user, only corresponding data user can see the data. CSP and attackers can't see the data. DO again encrypt that encrypted data with public key of CSP using public key cryptography so that attacker can't see the encrypted data as well as CPList. Here, CSP can see only the CPList. Here, data confidentiality is increased due to double encryption. In related works, there are so many schemes suggested to secure the data, but all have used the cryptographic schemes which have keys for all types encryptions. In this scheme, data is encrypted with GA which has no concept of key which increases data confidentiality more because key is as important as data. If, key is compromised data may be decrypted.

Also, the new scheme does not encrypt whole data or file at once. Here, data is first divided into number of blocks of bits. Two blocks of bits are selected for encryption process at a time. Each pair of blocks of bits is encrypted with GA. And, each pair of blocks may have different GA operations. So, to decrypt a whole data, there may need to perform many and different GA operations. So, it is difficult for an attacker to decrypt a whole file or data. The output of each GA is ciphertext (two blocks of bits). Each ciphertext is stored on the cloud at distinct

location, which increases confidentiality of data. Also, the location of ciphertext is dynamic, so attacker can't guess where ciphertext is stored. Hence, security of the data increases tremendously. If a data is divided into smaller blocks, then number of blocks increases corresponding to a data. Hence, more number of distinct GA operations are required to encrypt a whole data. So, obtained ciphertexts of a whole data has more random bits. Hence, confidentiality of a data increases as randomness increases.

Also, most of the existing schemes have the single location for storage ona cloud data center. The major disadvantage of using single location for storage on a cloud data center is that, if attacker attacks on cloud the whole datawill be accessed easily. So, to mitigate this security issue, the new scheme stores the encrypted data parts on different locations on CSP.Since, data are encrypted by use of pseudorandom number which is only known to DO and respective user, DO and respective user can decrypt the data.

- **Entity Authentication-** In this scheme, user is authenticated at DO when he forwards his own details to DO during the registration. DO and CSP have authenticated each other at CSP when DO sends encrypted data and CP List to the CSP because DO encrypt the data using his private key. The user is authenticated at CSP when he requests for data by sending his UID, FID, AR and CSP compare it with related stored UID, FID, AR of a data.

- **Data Access Control-** The new scheme uses the CP List for secure and fine grain access of cloud data. In CP List, operations and recognized data for a user are described. Some schemes have used the Access Control List (ACL). But CP List is superior to ACL because ACL describes users and their authorised operations for each data and it is practically inefficient that two users require same data and have same operations on it. ACL lacks the scalability and fine grain access of data [5].

5.2. Analysis of Performance

In the new scheme, DO has moved its maximum computation and load to CSP and only did few important things by itself. The new scheme has reduced the additional computation time by using GA because there is no key concept used in GA. As, we know that to store, maintain, secure and distribute the keys securely is difficult, challenging and total computational overhead.

6. Simulations and Results

Here, there is simulation and results. We implement the scheme in C# language using visual studio tool. Following screenshot shows the user registration, login, encryption, decryption and downloading of file pages.

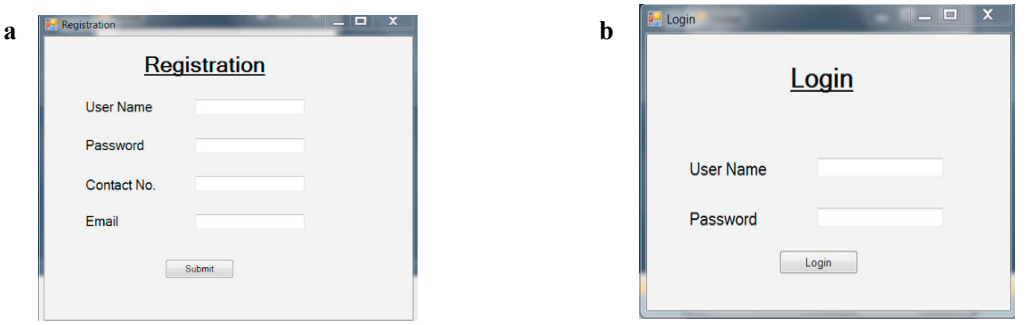
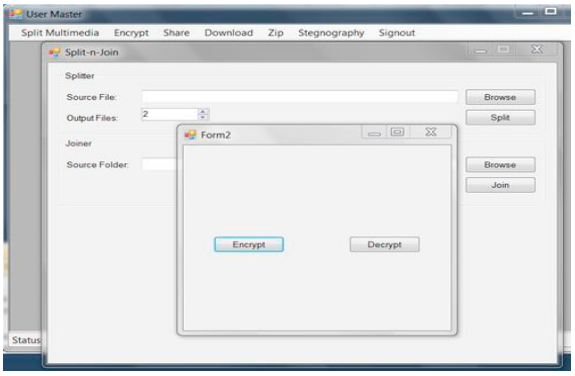


Fig.3.(a) User Registration; (b) User Login.

c



d

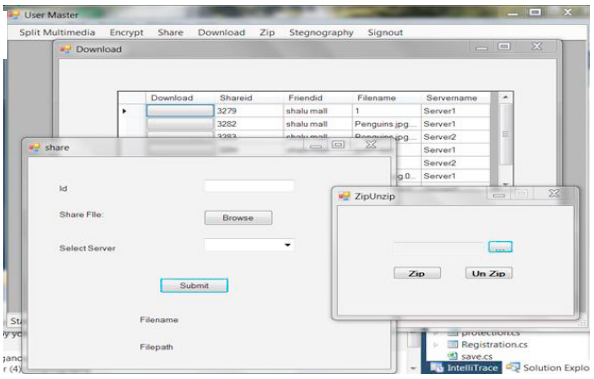


Fig.3. (c) Process of File Splitting, Encryption and Decryption; (d) Process of Downloading, Sharing and Compressing the File

7. Conclusion

The new scheme ensures the security of data which are stored on CSP. Many schemes are presented for security of data, but they have some issues such as vulnerability of attack, lack of fine grain control of access and system performance. Due to large number of keys, confidentiality of data and performance of system decreases. But, the new scheme uses GA operations (crossover and mutation) which has no key concept. In the scheme, GA is applied in unique way and data are stored at distinct locations on the cloud in secure manner. GA ensures the data confidentiality. The scheme has used the capability list to ensure the fine grain control access.

References

[1] Sushil Kumar Saroj, Sanjeev Kumar Chauhan, Aravendra Kumar Sharma and Sundaram Vats. (2015) "Threshold Cryptography Based Data Security in Cloud Computing." *IEEE International Conference on Computational Intelligence & Communication Technology*: 202-207.

[2] Jeong-Min Do, You-Jin Song and Namje Park. (2011) "Attribute Based Proxy Re-encryption for Data Confidentiality in Cloud Computing Environments."*First ACIS/JNU International Conference on Computers, Networks, Systems and Industrial Engineering,*:248-251 .

[3] Preeti Garg and Vineet Sharma. (2014) "An efficient and secure data storage in Mobile Cloud Computing through RSA and Hash function." *IEEE International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*:334-339.

[4] Don Boneh,and Ramarathnam Venkatesan. (1998) "Breaking RSA May Be Easier Than Factoring."*Lecture Notes in Computer Science*: 59-71.

[5] Sunil Sanka, Chittaranjan Hota and Muttukrishnan Rajarajan. (2010)"Secure data access in cloud computing." *IEEE 4th International Conference on Internet Multimedia Services Architecture and Application*:1-6.

[6] Doyel Pal, PraveenkumarKhethavath, Johnson P.Thomas, and Tingting Chen. (2015) "Multilevel Threshold Secret Sharing in Distributed Cloud."*Springer International Symposium on Security in Computing and Communication Cham*:13-23.

[7] Balasaraswathi V. R. and Manikandan S. (2014) "Enhanced security for multi-cloud storage using cryptographic data splitting with dynamic approach."*IEEE International Conference on Advanced Communications, Control and Computing Technologies*:1190-1194.

[8] W.Stallings. "Cryptography and network security."*LPE Fourth Edition*.

[9] Mrinal Kanti Sarkar and Sanjay Kumar. (2016) "A framework to ensure data storage security in cloud computing."*IEEE 7th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*:1-4.

[10] Ahmed Albugmi, Madini O. Alassafi, Robert Walters and Gary Wills. (2016) "Data security in cloud computing."*Fifth International Conference on Future Generation Communication Technologies (FGCT)*: 55-59.

[11] P Srikanth, Abhinav Mehta, Neha Yadav, Sahil Singh and Shubham Singhal. (2017) "Encryption and Decryption Using Genetic Algorithm Operations and Pseudorandom Number." *IJCSN - International Journal of Computer Science and Network* **6(3)**:455-459.

[12] Arjun Aggarwal, Abhijeet Mishra, Gaurav Singhal and Sushil Kr Saroj. (2016) "An Efficient Methodology for Storing Sensitive Datausing Nested Cloud."*International Journal of Computer Applications* **142 (10)**: 0975 – 8887.