ANALYSING THE EFFICACY OF GENETIC ALGORITHM TO MITIGATE THE DATA SECURITY VULNERABILITY ON CLOUD

VISHAL DUHAN
Department of Computer Science and Engineering,
Jaypee University of Information Technology, Waknaghat Solan(Himachal Pradesh)

ABSTRACT

In many organizations and institutions the use of cloud computing has rapidly increased. Cloud data storage is one of the main advantages of cloud computing, where the users cannot stored their data on own servers but the data is mainly stored on the CSP side. On this end, cloud computing security is one of the most analytical aspects due to the confidential information and the sensitive data is stored on the cloud by the users. This paper purposes a scheme to store the information or data securely on the cloud, by applying the Genetic algorithm on the data and split the data into various chunks and then stored on the CSP.

***Keywords**: CSP; DO; Genetic Algorithm; Data splitting; Cloud Storage; outsourced data*

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. Now a 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 issues, data protection is very essential security issue in cloud computing. The advantages of using cloud computing technology including easy scalability, cost saving, and high availability.

In cloud computing, there are three types of service models namely-

- **IaaS**- In IaaS (Infrastructure as a Service) users gets resources like CPU and power, network bandwidth, processing power and storage.

Once the customer gets the infrastructure he controls the OS, application, data host based security, services etc.

- **PaaS**- In Paas (Platform as a Service) users provides the hardware infrastructure, OS and network to make a hosting environment. From the hosting environment user can activate services and install his applications.
- **SaaS**- In SaaS (Software as a Service) users provides the access to an application. He has no restriction over the network, hardware, OS or security.

Cloud services can be typically deployed in various ways such as-

- **Public Cloud/External cloud -** It can be owned and operated by cloud provider and the services are offered over the internet.
- **Private Cloud/Internal cloud -** It is used when the cloud data centre is to be operated only for a specific business.
- **Hybrid Cloud/Mixed cloud-** It focus to made the cloud more secure and to provide the resource sharing and same services. It is the mixture of public and private cloud.
  Hybrid cloud = Private cloud + Public cloud
- **Community Cloud** – It is a joint effort in which the infrastructure is shared within several organizations with common concerns from a specific community.

The Cloud Computing (CC) has five main features they are- 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 profit by exploiting these cloud computing characteristics. This is the reason; industries are moving their business against cloud computing. In this way we can say that security of data is a major restriction in Cloud Computing. In present time the cloud computing security is one of the biggest critical issue in the environment of cloud computing due to the important data stored for users in the cloud .Cloud providers should consider security and privacy issues as high and essential priority.

The data of Data Owner (DO) is prepared and saved on outside servers. So, integrity, confidentiality as well as data access becomes more accessible. Since the outside servers are managed through monetary service providers, data owner can't expect about authority, as they can use the data of data owner being their profit as well as destroy the work of data owner. Even Data Owner (DO) cannot faith on clients or users at the time they can be spiteful. Confidentiality of data can breach over service providers and collusion attack of spiteful users. As the number of keys are more and number of keys handling mechanisms are also large in number , responsive rate of data is not achieved. In CP List (capability list) operations and recognized data for a client/user are described. It is superior to (ACL) access control list because it described users and their operation for each file and data. The approach has used MD5 to ensure data integrity.

RELATED WORK

There are two important security needs for outsourced the data that is access control and data confidentiality. Once, when we indicate further on security of data, we not be able to remember about systems performance (CSP, DO, Users). Remarkably, there is a strategy needed that not either provides data security although still maintains the work of the system. There are many strategy are recommended to appropriate these provision.

Scheme proposed in models and assumptions uses Third Party Auditor and Hash function and RSA. In this scheme, the third party auditor (TPA) is considered inactive and performs all the computations and verifications. It is known that we cannot fully trust on TPA's, that it can usage the data of Data Owners (DO) for own financial profit. Another improvement field in proposed scheme is, breaking the RSA much simple than factoring.

Scheme proposed using Shamir sharing algorithm with CRT (Chinese Remainder Theorem) which assign the key and shares the key among participant resource providers but proposed scheme in has no arrangement of managing data.

Table1. List of Symbols and their Description used in this paper

Symbols	Description
GA	Genetic Algorithm
DO	Data Owner
CSP	Cloud Service Provider
AES	Advanced Encryption Standards
RSA	Rivest Shamir Adleman
Ek	Encryption
Dk	Decryption
UID	User Identity
FID	File Identity
AR	Access Right
CP List	Capability List

Scheme proposed used cryptographic data polarizing mechanism with AES algorithm which bifurcate the user encrypted file of public cloud but there is no corresponding confirmation related to key.

Scheme proposed reliable to earn access control and confidentiality of data. In this strategy, files or data are encrypted through symmetric keys and these keys are only known to corresponding data users and data owner. The encrypted files are stored at CSP. During the communication among CSP and user to secure the data from outsiders data are further encrypted through one time secrete session key which is shared among CSP and user through the modified Diffie Hellman protocol. Here the keys are increases as there is a key corresponding to every user and users are large in number..

In proposed scheme communication model is matches with previous researches, but the proposed scheme is much more secure.

MODELS AND ASSUMPTIONS

We assume that our model is composed of three entities a DO, a CSP and users linked to that DO. Initially, certification of every user is done at DO, during certification users send their credentials to DO. We suppose DO received the credential of users securely.

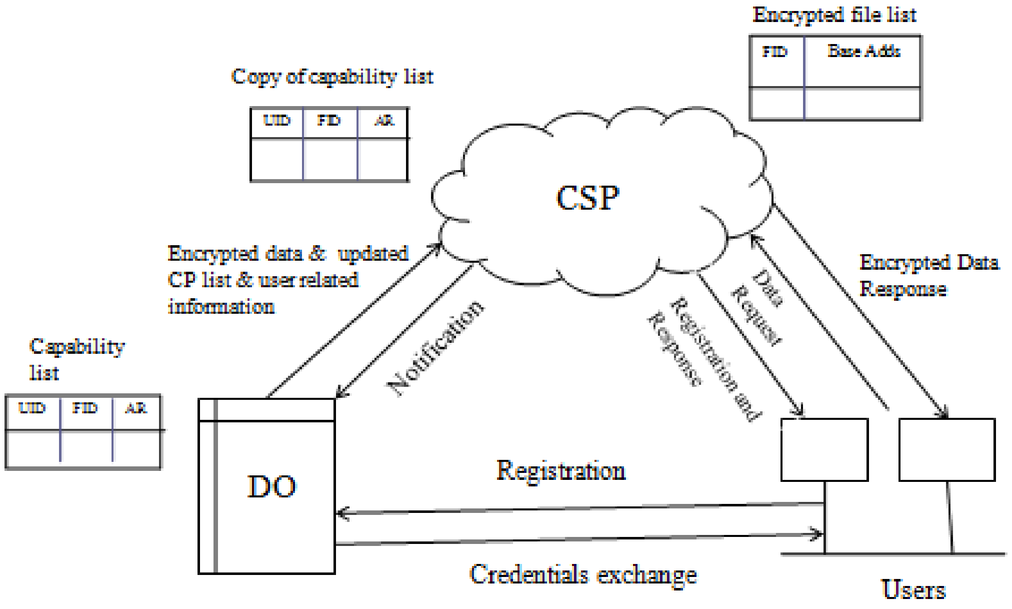

Fig.1 Communication model for proposed scheme

To segregate the user apply MD5 algorithm for securely communicate to group of users for response of registration. After successful authentication of CSP a user can retrieve data from CSP in secret manner. DO have space to store some data or files. Genetc algorithm stores the user's data on cloud by encrypting the data using genetic algorithm and then splitting the data in n parts.

PROPOSED SCHEME

In this part, we present a model for securely communicate with different entities and securely access the data. In our proposed model we used three algorithms. Algorithm 1 describes new user registration. Algorithm 2 describes encryption process in which data is securely communicates between DO and CSP. Algorithm 3 describes decryption process with genetic algorithm. This algorithm ensures authentication and data confidentiality of CSP and User.

Algorithm 1.1: New user registration

Step-1: User send registration request to DO.

Send (user details)

Step-2: DO updates table or CP List its end

Update (UID, FID, AR)

Step-3: DO encrypts the data and send the encrypted data to CSP

Send Ek (data) CSP

Step-4: CSP updates the copy the CP List

Updates (UID, FID, AR)

Step-5: Now, the user can precisely connect to CSP for fetch his details.

Algorithm 1.2: Encryption Process

Step-1: File should be in ASCII form which ultimately need to be in binary form.

Step-2: binary chunk is divided into block of 8 bits and the value of 8 is also stored as a secret key.

B is equal to number of block which is equal to length of binary string divided by 8.

Blocks of 8 bits that are generated are considered as B1,B2,B3….Bn stc.

Step-3: for every tow selected blocks generate a random number which in turn acts as a key. Key is then mod by 3 and hence the value is then used to choose crossover.

for single point crossover value of key is 0

for two-point crossover value of key is 1

for multipoint crossover value of key is 2

Step-4: Apply this method on blocks and blocks is recognised as N1,N2,N3…..Nn etc.

Step-5: Mutation is applied on selected block , blocks are recognised as M1, M2, M3…….Mn etc.

Step-6: Repeat these steps till the end of chunk.

Step-7: Now convert this form into ASCII then ultimately into corresponding file type.

Step-8: Store the result in a file..

Algorithm 1.3: Decryption Process

Step-1: File should be in ASCII form which ultimately need to be in binary form.

Step-2: Binary chunk is divided into block of 8 bits.

Step-3: Mutation process is applied on blocks M1,M2,M3….Mn etc.

Step-4: Crossover method is applied on blocks N1,N2...Nn etc. and key of the selected block is mod by 3.

Step-5: Repeat these steps till the end of chunk..

File Part 1 File Part n

PERFORMANCE AND SECURITY ANALYSIS

I. *Analysis of Security*

Here, we discuss about the strength, scalability as well as security.

a) Data Confidentiality- In this scheme, Data Owner (DO) stores the own data in encrypted form on the cloud service provider (CSP). As, data are encrypted by Genetic Algorithm and then split into n parts, then stored on CSP. Here the confidentiality of outsourced data or file is increased.

b) Entity Authentication- In this scheme, user is validated at DO, when user sends his personal detail to DO in encrypted form during registration. DO is validated at Cloud Service Provider (CSP) when it send data and capability list as well as encrypted Message Digest (MD) to the Cloud Service Provider (CSP).When password and users ID both are match with user ID and password which is stored on the CSP database, then user is validated at Cloud Service Provider.

c) Data Integrity- In proposed work, MD5 is use for the calculation of Message Digest (MD) of the data. With the help of genetic algorithm data owner encrypts the data and also calculate the message digest of the data with which is only familiar to respected group of users. User brings Message Digest (MD) and encrypted data through CSP, at that condition firstly user decrypts the data and also calculates the message digest of gathered data. When calculated data and gathered data both is matched, then we can say that integrity of data is insured.

d) Data Access Control- In proposed work, for ensuring the access control of data we can use Capability List (CP List). Basically capability list contains AR, FID and UID. DO have a right to execute the activity and CSP read this activity for the objective of securely access the data.

Analysis of performance

We analyse that DO removed its maximum computation and load to CSP and only did important things by itself. Data confidentiality and security is increased in the proposed scheme by using the genetic algorithm and also reduced the additional computation time.

CONCLUSION

Our proposed model gives the security of data which is stored on Cloud Service Provider (CSP).Many methods are presented for security of data but these are suffering from collusion attack. By employing the Genetic Algorithm we secure the information or data from collusion attack. As, DO stored its data at CSP in the form of encryption, here the data confidentiality is ensured. The strategy has used capability list to ensure the access control of data and MD5 ensure the data integrity and the entity authentication respectively. Genetic Algorithm ensures an effective encryption and decryption of data or files and hence security is maintained.