An Efficient Methodology for Storing Sensitive Data using Nested Cloud

Arjun Aggarwal
CSE Department
GCET, Greater Noida-201306

Abhijeet Mishra
CSE Department
GCET, Greater Noida-201306

Gaurav Singhal
CSE Department
GCET, Greater Noida-201306

Sushil Kr. Saroj
AP, CSE Department
GCET, Greater Noida-201306

ABSTRACT

In the period of revolutionary change in information technology, the two developments that are most far-reaching have been cloud computing and the mobile Internet. These two information technology revolutions diverge in mobile cloud computing. There are already immense risks with data hosted in-house, so it's no secret that data offsite sits at even higher risk. With Data on offsite, higher avenues for attack and the fact that it will be traveling more makes it easier to be hijacked. With the technology regularly improving, there are ways to make sure of greater security. However with technology regularly improving, there are always people out there lifting their hacking skills.

An effective methodology with prominent features of data integrity and confidentiality is proposed here to ensure the safety of offsite data or the data on cloud. In this paper we propose the mechanism which uses the concepts of ECC algorithm, Shamir's secret key sharing algorithm along with distinct cryptography tools such as threshold cryptography that enhances the security of data stored on clouds.

Keywords

Cloud, mobile cloud computing, offsite data security, threshold cryptography.

1. INTRODUCTION

We have been inspecting some significant technology trends for the last few years. **Mobile** and **Cloud Computing** are two such trends. Worldwide acceptances of these two trends are changing our lives, the way we do business and most of our day-to-day work. Cloud computing can be seen as the practice of using a network of remote servers hosted on the Internet to store, manage, manipulate and process data, rather than using a local servers or a personal computer. The data from Research and analysis shows how greatly these technologies have created a reverberation in the world of technology. Also with the explosion of mobile and handheld devices which also significantly contributes to world IP data traffic. To overcome such data demand, cloud computing looks to be the right choice because of its rapid scalability, robust disaster recovery, universal network access, on-demand self-service and other traits. [1].

Mobile cloud computing is one of the fastest growing stream. According to the latest study of Juniper Research, the number of mobile cloud computing subscribers is expected to grow very quickly in the next 5 years. Cloud-based mobile market has generated annual revenue of $400 million in 2009 to $9.50 billion in 2014, at an average annual increase of 88% [2].

The security of Cloud computing is one of the most demanding issue in cloud computing environment in present time due to the valuable information stored for clients in the cloud. Cloud providers should consider privacy and security issues as high and urgent priority. Many schemes are given to ensure these requirements but they are suffering from key handling mechanisms due to this drawback these schemes are incapable to secure very highly sensitive data such as data associated with armed forces, data of a billionaire organization, data related to any exam. As a result of the importance of data security and privacy in cloud computing, this paper target more on the issues related to the key handling aspect of cloud computing security. To address these issues the paper propose a scheme whose objective is to get a solution or come out with an integrated framework for achieving the data security on mobile cloud environment in various possible conditions, so that this technology may be used in applications of all round nature without any imperfection.

The remainder portion of this paper is organized as follows. In Section II, the paper reviews the related work. In section III, it outlines the model and assumption. In section IV, it presents the proposed scheme. In section V, it analyses the proposed approach. Finally, the conclusion of the paper is in section VI.

Table 1. List of Symbol used and their Description

Symbol	Description
DO	Data Owner
AES	Advanced Encryption Standards
DES	Data Encryption Standard
ECC	Elliptic Curve Cryptography
SP	Share Pool
RSA	Rivest Shamir Adleman

2. RELATED WORK

Data integrity and conservation of privacy are two of the major critical security issues related to user data. Many times, when a person accentuate more on performance of cloud that how quickly the cloud stores the confidential data, he forgets about the security aspects of it. For example, to secure data, a person sometime uses keys, he knows that keys itself are confidential, but to secure and maintain these keys he either uses the mobile device of user itself or he trusts on Third Party Auditor (TPA). But in both cases there are many loopholes to easily access the data in cloud using the keys stored in mobile device or TPA. So, there is a need of scheme that not only maintains the performance but also provides data security. Many schemes are suggested to meet these requirements.

The scheme proposed in [3] used TPA with RSA & Hash function. In this scheme, the TPA is considered neutral and performs all the verifications and computations. It is known that TPA's cannot be fully trusted, as it can use data owner's data for its commercial benefits. Another area of improvement with proposed scheme [3] is the use of RSA cryptography algorithm, as breaking RSA is much easier than factoring [4].

The scheme proposed in [5] used Shamir sharing algorithm with Chinese remainder theorem which distribute the key shares among participant resource providers but scheme proposed in [5] has no provision of managing data.

The scheme proposed in [7] used AES algorithm with cryptographic data splitting mechanism which sliced the user encrypted file into two public clouds but there is no acknowledgement related to key.

Communication model of the proposed scheme somehow matches with [5], [7] but proposed scheme is more secure and take care both type of problems namely, data splitting and key shares in smart manner to provide better security. In this paper the proposed scheme keeps data and keys at different clouds and treat them differently.

3. MODEL AND ASSUMPTIONS

It is supposed that the model is composed of 3 entities: a DO, a Key Cloud and Data Cloud. Initially DO have the data which he wants to securely store on the cloud. So, Data is encrypted by the DO using ECC encryption algorithm. This generated cipher text and key, are sent to Data cloud and Key cloud respectively with the secured channel. Here it is assumed that all the network communications are secure as it is mainly concerned about the cloud security.

Different tasks are performed at Key cloud and the Data cloud. At Key cloud, the private key of cipher text is distributed to multiple parts at the time of storing the key and restoration at the time of retrieval of key using Shamir's Secret Key Sharing Algorithm.

At the Data cloud, cipher test is split into two or more sub-clouds which are changing their path dynamically and a directory is maintained in order to re-access these clouds when the user wants to retrieve the data from the cloud DO.

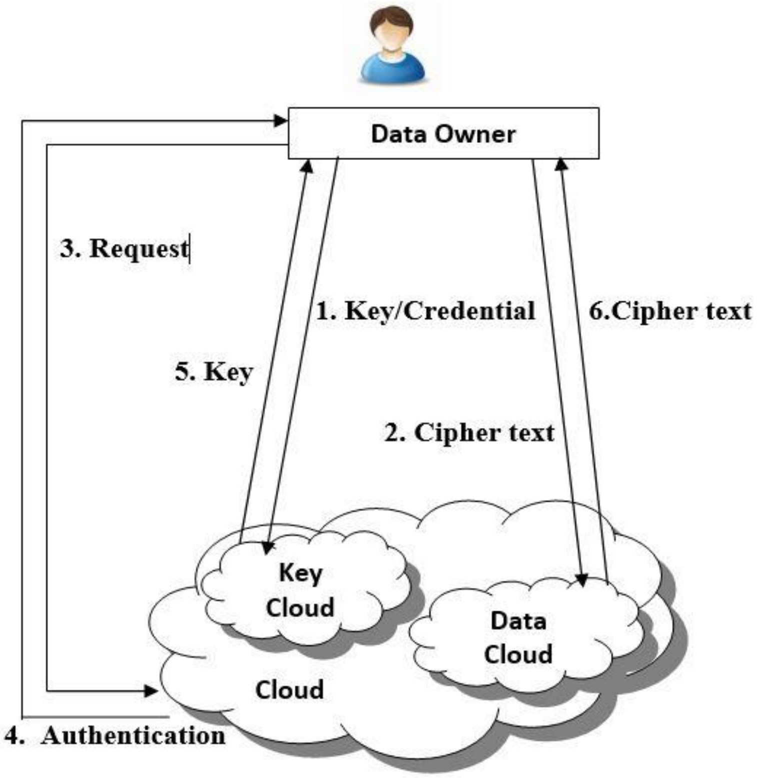

Fig 1: Communication Model in Proposed Scheme

4. PROPOSED SCHEME

In this section the paper presents a model for secure data storage and retrieval on clouds. It contains four algorithms used in the proposed scheme. Algorithm1 is Elliptic Curve Cryptography (ECC) which is a public key cryptography algorithm. It is known that a Public-key cryptography provides a mechanism for sharing keys between a numbers of participants. At first the data which is to be stored at the cloud is encrypted using ECC at the user's mobile machine after which the data is operated using algorithm 3 and the key is handled using algorithm 2 at the data server and key servers respectively. As on the user machine are not capable of performing higher computations so only ECC encryption and decryption is performed at this and then data and the keys are transferred to their respective servers for further processing. The algorithm used for encryption is described as algorithm 1 below.

Algorithm 1.1: Key Generation

One of the necessary procedures is the generation of key in which it generates both private and public key

Step-1: Encryption will be performed using receiver's public key at the sender side and decryption at receiver's side by his private key.

Step-2: Now, a number 'd' is selected within the provided range of 'n'.

Step-3: Public key (Q) is generated using the given equation

$\mathbf{Q = d * P}$

Where d is a random number which is selected within the range of (1 to n-1). P is the point on the curve.

'Q' is the public key and 'd' is the private key.

Algorithm 1.2: Encryption

Step-1: If the user wants to send the message 'm' then he has to represent this message on the curve. These have in-depth implementation details.

Step-2: Consider *'m'* has the point *'M'* on the curve *'E'*.
Step-3: 'k' is selected from [1 – (n-1)].
Let C1 and C2 are the two cipher texts that will be generated.

$$\mathbf{C1 = k*P}$$
$$\mathbf{C2 = M + k*Q}$$

C1 and C2 will be sent.

Algorithm 1.3: Decryption

To get the original message 'm' the user has to perform decryption using cipher texts C1 and C2 with equation

$$\mathbf{M = C2 – d * C1}$$

M is the original message that the user has send.

The second algorithm that is used is Shamir's secret key sharing algorithmin which the private key is distributed amongst a group of participants as a share of the secret. These shares are further split into shares as second level of sharing along with a threshold value that is the minimum number of shares required for reconstructing the key when required. At the time of retrieval of the key the user has to have at least the threshold number of key shares to reconstruct the key.

The algorithm replicates each share of secret into k (let) numbers so that if one resource provider goes offline or compromised then that share can be accessed from other resource provider. To determine the number of shares m each resource provider pair from the share pool SP which is created beforehand. It is intended to create multiple share pools and place one or two of them in each cluster of the nested cloud. The CRT solution generates a number m which decides the number of shares to split and reconstruct the key in the second level. Shamir's threshold key sharing is applied on the private key obtained by encrypting the data at the user mobile machine, and at the time of retrieval this key is downloaded at the user machine for decryption of the data.

Algorithm 2.1: Share Pool Generator

Input: S - Secret key, n - number of resource providers in cloud, t - threshold value (where $t \leq n$). From this share pool each resource provider generates a number out of a selected pair.

```
for i = 1 at least n do
    Generation of a random series of pair wise relatively prime positive integers, Pi = pi1, pi2,.,
    Generation of a random series of m arbitrary integers Ni = $n_{i1}$, $n_{i2}$,.,$n_{im}$.
    Place these two series Pi along with Ni, represented as (Pi, Ni) in SP.
end for
```

Algorithm 2.2: Split Algorithm

```
User performs encryption on the file f with secret key S.
Split the secret key S into n number of shares, $S_1$, $S_2$, $S_3$, $S_4$, ..,$S_n$.
To reconstruct S at least t number of shares are required where t is threshold value
for each share i of S, where i∈(1, n) do
    Duplicate the share $S_i$ into k number of replicas. Where k>= 1
    Distribute these replicas among n different resource providers.
end for
for each $R_i$ Resource provider, where i∈(1, n) do
    Choose a pair ($P_i$, $N_i$) from SP.
    User saves (i, $P_i$) and $R_i$ saves $N_i$.
    Get a unique solution m = $x_i$ from ($P_i$, $N_i$) using Chinese Remainder Theorem (CRT)[5].
    Split the share of secret Si into m number of shares, $S_{i1}$, $S_{i2}$, $S_{i3}$, $S_{i4}$, ..,$S_{im}$ .
    Generate j number of dummy shares SD ij , where $j \geq m$.
end for
```

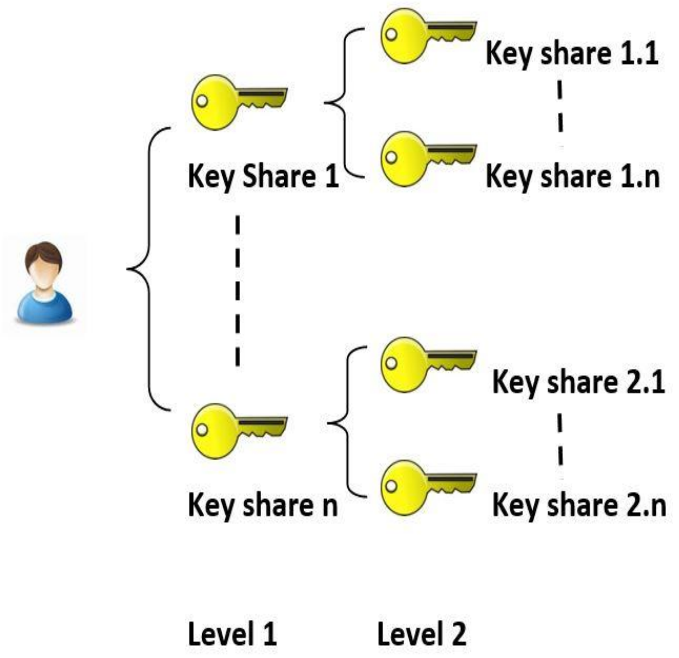

Fig 2: Distribution of private key using Threshold cryptography

Algorithm 2.3: Reconstruction Algorithm

```
for each $R_i$ Resource provider, where $i \in (1, n)$ do
    $R_i$ has $N_i$ and asks for $(i, P_i)$ pair from user.
    Get a unique solution $m = x_i$ from $(P_i, N_i)$ using Chinese Remainder Theorem (CRT)[5].
    Reconstruct the share of secret $S_i$ from m number of shares, $S_{i1}, S_{i2}, S_{i3}, S_{i4}, .., S_{im}$ .
end for
for each $R_i$ Resource provider, where $i \in (1, n)$
    Collect the share $S_i$ from each resource provider.
end for
Reconstruct the secret S from Si where i = 1, ..., t.
```

The third algorithm used is for splitting of encrypted data and private key before storing them on cloud servers. And the algorithm includes the steps for cryptographic data splitting and changing the logical path of both file slices with respect to time. In the algorithm firstly the data is sliced into two parts and stored on different clouds. Now in next phase the logical paths to these slices are changed with respect to time. These access paths are changed internally to the cloud not in between the clouds. While downloading the data required the user have to authenticate him along with a request after which servers retrieves the data stored in respective clouds. These retrieved slices of data then merged and decrypted to obtain original data and then stored on server cache. This algorithm is implemented at the data server by the CSP after receiving it from user's machine. At the time of retrieval the dynamic url is used for downloading of the data at the user's machine.

Algorithm 3: Data Splitting

```
Begin
    DivideE($U_{fi}$) into two randomized parts $U_{fpb}U_{fpZ}$;
    Store these parts in different memory locations
    While
        Generate two random numbers $R_{b,}R_z$;
        Store E ($R_b$), E ($R_z$) in private cloud;
        If ($R_b$ not exists) then
            Make directory $R_b$ in memory;
        Else if ($R_z$ not exists) then
            Make directory $R_z$ in memory;
        End if
        Move $U_{fbb}$ into $R_b$ and $U_{fpz}$ into $R_z$
        Delay loop with time T.
    End while
End
```

The proposed scheme works as using these algorithms collectively to provide security to the data that is stored on clouds. Initially the data set that is required to be stored is encrypted using ECC algorithm and as a result we get a private key and the encrypted set of data. Now they are handled one by one, at first on the keys Shamir's secret key sharing algorithm is applied to split the keys into several partitions or shares so and set up a threshold, so that it can be reconstructed when required by the user.

On the other hand the encrypted data is split into two parts and stored at different directories, which are dynamically changed at regular time intervals. A record is kept of those access paths of the directories used and stored in cloud after encryption and on request from the user after authentication the server retrieves file slices from the clouds. The retrieved data are merged and decrypted and stored in the server as cache. This data along with the regenerated keys is decrypted and then viewed as the whole document.

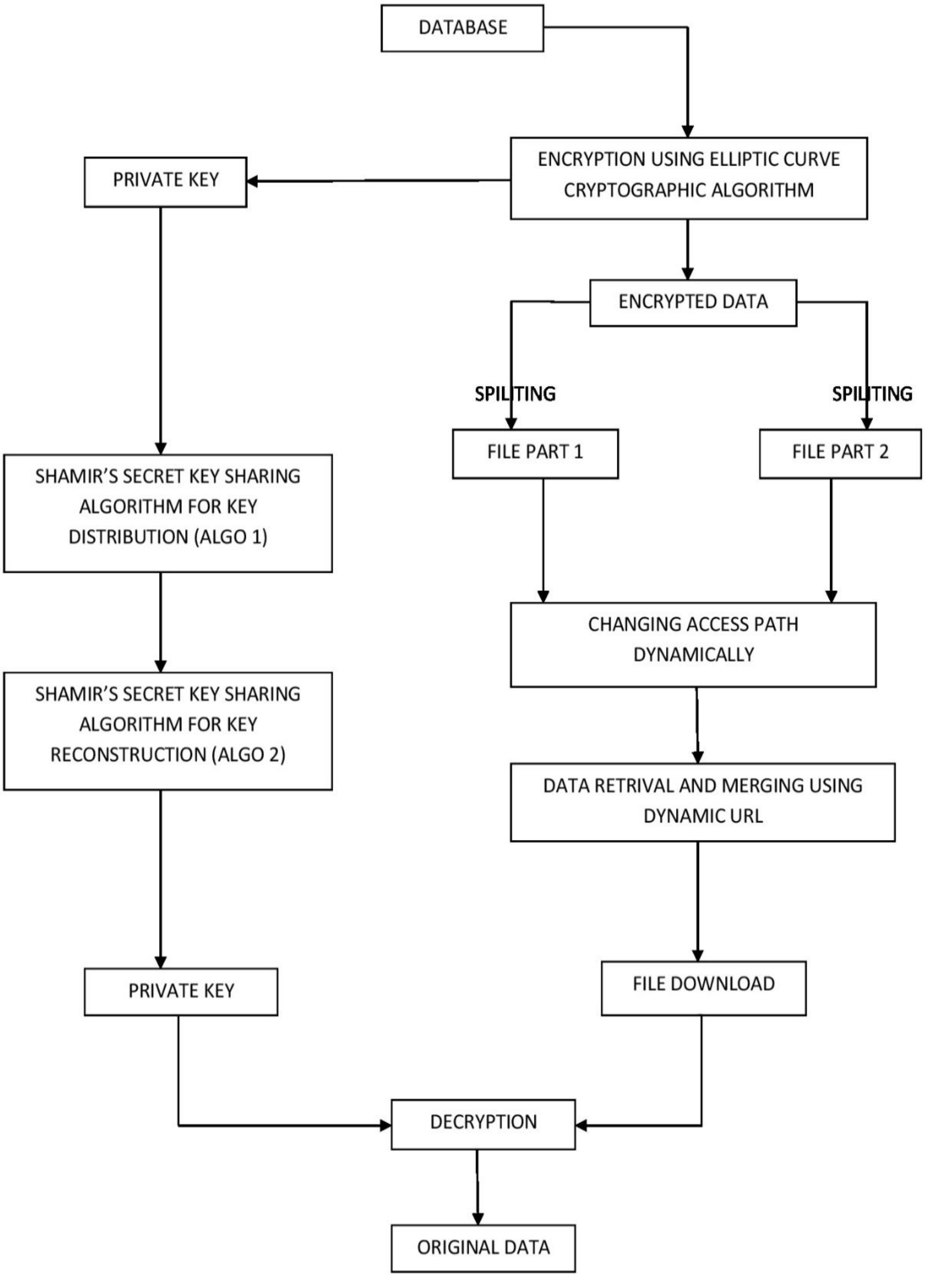
DATABASE
ENCRYPTION USING ELLIPTIC CURVE CRYPTOGRAPHIC ALGORITHM
PRIVATE KEY
ENCRYPTED DATA
SPILTING
SPILTING
FILE PART 1
FILE PART 2
SHAMIR'S SECRET KEY SHARING ALGORITHM FOR KEY DISTRIBUTION (ALGO 1)
CHANGING ACCESS PATH DYNAMICALLY
SHAMIR'S SECRET KEY SHARING ALGORITHM FOR KEY RECONSTRUCTION (ALGO 2)
DATA RETRIVAL AND MERGING USING DYNAMIC URL
PRIVATE KEY
FILE DOWNLOAD
DECRYPTION
ORIGINAL DATA

Fig 3: Data and Key Handling Mechanism

5. RESULT AND ANALYSIS

In this section the security provided by algorithms of the proposed work is analyzed. One of the major threats of using single cloud storage is that if the security is compromised for that cloud then whole of the data of that cloud will be accessed. So to overcome this difficulty from [3][5], the paper uses multiple clouds storage concept in the proposed scheme that makes it difficult to access all the stored data because a person has to have access of all the clouds for getting whole of the data of a user.

In the proposed work, the dynamic access path changing mechanism is used on split data files within the clouds, which provides additional security because attackers don't know where encrypted files are actually kept.

The proposed work [7] encrypts the data and further processes the encrypted data but it is not described how keys are handled. In this scheme it ensures the additional processing of keys. The scheme uses separate key cloud server and employed the Shamir's algorithms to manage the keys.

Using Shamir's threshold secret sharing, the scheme assures the distribution of the shares of secret keys among the shareholders to be free from collusion attack. Because these individual shares have no individual value but a threshold numbers of shares collectively are capable of regenerating the keys whenever it is required.

In the paper [3], data are transmitted between user and the TPA after encryption with RSA. RSA algorithm is considered secure by assuming that factoring a number which is multiplication of two large prime numbers is difficult. In ECC algorithm it is assumed that finding discrete logarithm of an element on elliptic curve with respect to given element is infeasible. Computing discrete logarithm is much more difficult to computing factoring. Also, ECC algorithm provides better security than RSA or any other algorithms of similar class having same key size. That's why, the ECC algorithm is employed instead of RSA.

6. CONCLUSION

The proposed scheme in this paper enhances the security of the sensitive data stored on clouds as compared to the existing methodologies. It provides better handling of key and data by distributing key using threshold cryptography and splitting data dynamically in nested cloud. This approach may not be economically efficient but at the point of handling sensitive data the economic aspects may be ignored as the security of highly sensitive data cannot be compromised at the cost-effective level.

Yet there is always a scope for improvement in a certain approach. This approach may be enhanced at some points as at the level of data splitting and access paths maintenance.

7. REFERENCES

[1] Shamim Hossain, "*What is mobile cloud computing*?". [Online].Available:http://wedowebsphere.de/blogpost/what-mobile-cloud-computing.

[2] M. Padma, and M. Lakshmi Neelima, "*Mobile Cloud Computing: Issues from a Security Perspective*," International Journal of Computer Science and Mobile Computing (IJCSMC), v.3, i.5, p.972-977, May. 2014.

[3] Preeti Garg, and Dr. Vineet Sharma, "*An Efficient and Secure Data Storage in Mobile Cloud Computing through RSA and Hash functions*," International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014 IEEE, vol., no., pp.334-339, 7-8 February 2014.

[4] Don Boneh, and RamarathnamVenkatesan, "*Breaking RSA May Be Easier Than Factoring,*"Lecture Notes in Computer Science, 1998EurocryptSpringer-Verlag,vol. 1233, no., pp.59-71, March 1998.

[5] Doyel Pal, Praveen kumar Khethavath, Johnson P. Thomas, and Tingting Chen, "*Multilevel Threshold Secret Sharing in Distributed Cloud*,"Third International Symposium on Security in Computing and Communications (SSCC), 2015Springer,vol. 536, no., pp.13-23, 10-13August 2015.

[6] Bithin, "*Simple explanation for Elliptic Curve Cryptographic algorithm (ECC)*". [Online].Available: https://bithin.wordpress.com/2012/02/22/simpleexplanation-for-elliptic-curve-cryptography-ecc

[7] Balasaraswathi V.R., and Manikandan.S, "*Enhanced Security for Multi-Cloud Storage using Cryptographic Data Splitting with Dynamic Approach,*" International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), 2014 IEEE, vol., no., pp.1190-1194, 8-10May 2014.