Obfuscrypt: A Novel Confidentiality Technique for Cloud Storage

L. Arockiam, Ph.D
Associate Professor,
St. Joseph's College,
Trichy, India

S. Monikandan
Research Scholar,
M S University,
Tirunelveli, India

P. D. Sheba K Malarchelvi, Ph.D
Professor & HOD of CSE,
JJ College of Eng & Technology,
Trichy, India

ABSTRACT
In today's IT industry, the more sophisticated data storage is cloud storage. Cloud storage mainly helps Small and Medium scale Enterprises (SMEs) to reduce their investments and maintenance of storage servers. Most of SMEs are outsourcing their data to cloud storage. Users' data that are sent to the cloud have to be stored in the public cloud environment. Data stored in the cloud storage might mingle with other users' data. This will lead the data protection issue in cloud storage. If the confidentiality of cloud data is broken, then it will cause loss of data to the industry. Security of cloud storage is ensured through confidentiality parameter. To ensure confidentiality, the most commonly used technique is encryption. But encryption alone doesn't give maximum protection to the data in the cloud storage. To have efficient cloud storage confidentiality, this paper uses encryption and obfuscation as two different techniques to protect the data in the cloud storage. Based on the type of data, encryption and obfuscation can be applied. Encryption can be applied to alphabets, alphanumeric and symbols. Obfuscation can be applied to numeric type of data. Applying encryption and obfuscation techniques on the cloud data will provide more protection against unauthorized usage. Confidentiality could be achieved with a combination of encryption and obfuscation.

Keywords
Cloud computing; Cloud Storage; Encryption; Obfuscation; Confidentiality;

1. INTRODUCTION
Cloud computing delivers massively scalable computing resources as a service with Internet based technologies. Resources are shared among a vast number of consumers allowing for a lower cost of IT ownership [1][2]. At present, cloud computing is widely discussed in academia and industry. With virtualization and distributed computing technology cloud computing integrates the computing, storage, networking and other computing resources, and then leases to users. Such mode could reduce the cost of enterprise information construction and accelerate the informatisation of enterprise. The Cloud storage is designed for virtualized computer environment. The cloud storage is implemented using cloud computing that means utilizing the software and hardware resources of the cloud computing service provider. Cloud computing is growing at a very high velocity in the IT industry around the world. While there are many advantages of cloud computing, the some enterprises are still waiting to use cloud computing. Hence the data security problem of cloud computing is not solved completely. Cloud Storage provides a virtual space to store bulk data. But the data owners have no control over their data. The cloud provider has full control on the user's data. This makes the user's mind to think about the data security in the cloud.

Data protection in the cloud storage is the core security problems. Data protection [3] is concerned with data confidentiality, integrity, authentication, availability and so on. Data confidentiality means that only authorized persons can use the data. Data integrity refers to information that has not been modified or remains untouched. Authentication refers to the process of verifying whether the incoming user is authorized or not. Data availability refers to the ability to guarantee to use data in time when needed and also refers to the availability of cloud service provider on-demand.

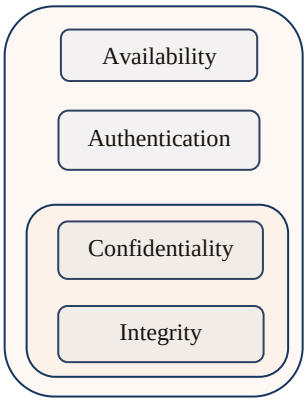

Fig 1: Layers of Data Security

Figure 1 represents the layers of data security in the cloud. First layer is availability, which ensures the availability of cloud computing resources or availability of the cloud providers needed when on-demand. Second layer is authentication, which helps to protect the unauthorized user's entry to the cloud. Third layer is confidentiality, which ensures that the data can be accessed by the privileged cloud users only. Last layer is integrity, which ensures that cloud data could not be modified by unauthorized cloud users. Authentication technique can be used to protect the data from outsiders attack. Confidentiality could be used to protect the data from outsiders as well as insiders attack. If the confidentiality of the data is ensured completely, then integrity will also be ensured. If the data in cloud storage can't be accessed by the intruders then it cannot be modified or altered by intruders. Even though the authentication mechanism is broken by the attackers, the data in the cloud is still be secured when an efficient confidentiality mechanism is used.

This paper proposes an efficient cloud storage confidentiality technique by using encryption and obfuscation technique. Encryption is the process of converting the readable text into unreadable form using an algorithm and a key. Obfuscation is same like encryption. Obfuscation is a process which disguises illegal users by implementing a particular mathematical function or using programming techniques.

Normally, confidentiality is ensured by encryption technique, but for the cloud environment encryption alone is not enough for data protection. Encryption is integrated with obfuscation technique. Obfuscation technique alone is also not enough to adopt for complete confidentiality of data in cloud storage because the user can find values through reverse engineering or by using brute force technique, which may compromise cloud data security. This paper uses encryption and obfuscation techniques in an integrated manner to protect the data from the attackers (insiders and outsiders). In the proposed technique, users should encrypt and obfuscate the data whatever they want to send to the cloud storage. Encryption and obfuscation could be done from user's side.

2. DATABASE MANAGEMENT IN THE CLOUD

Outsourcing of database management [4] is a necessary component of cloud computing. Due to the rapid advancements in network technology, the cost of transmitting a terabyte of data over long distances has decreased significantly in the past decade. In addition, the total cost of data management is five to ten times higher than the initial acquisition cost. This will lead the enterprises to outsource their data with cloud storage provider with minimum rate [5]. Database outsourcing model will enable the cloud users for better utilization of their data.

Despite the advantages offered by cloud-based DBMS, many people still have apprehensions about them. This is most likely due to the various security issues that have yet to be dealt with. Security becomes a serious problem with cloud DBMS when there are many Virtual Machines (which might be accessing databases via any number of applications) that might be able to access the database without being noticed or setting off any alerts. In this type of situation, a malicious person could potentially obtain pertinent data or cause serious harm to the integral structure of the database, putting the entire system in risk. This is the main problem with the cloud data storage.

Thus, an efficient security model needs to address this issue in the cloud data storage. According to this paper, even through the data is accessed by malicious persons it cannot be read or modified by them.

3. RELATED WORK

Ensuring confidentiality of user's data in cloud storage is the main research problem around the cloud computing. Cloud storage providers store users critical data; it needs to be secured. Cloud computing has a recent success in information technology and will dominate the IT industries in the coming years. Cloud computing also faces the overwhelming challenges. To ensure the proper physical, logical and personnel security controls, especially in cloud data storage are more significant. Moreover, while moving such large volumes of data, the management of the data may not be fully trustworthy. This section describes the research works which are related to ensure the confidentiality of data in cloud storage.

Database as a service (DaaS) is highly appreciated in business community because it saves hardware cost, cost of the technical people required to manage the database and it also saves the license cost of the database. Moreover, it offers reliable services and people can access their data 24 x 7 from anywhere provided the internet connection is available. Despite of all these advantages enterprises are reluctant to adopt DaaS, because of two types of threats that are associated with it. Firstly, it can be attacked by the hacker and secondly the privacy of data can be compromised by the administrators, managing the cloud database environment. To address this issue, Atiq ur Rehman [6] proposed a model that encrypt and obfuscate data on client side before sending to the cloud database. In addition paper offered a mechanism to query over encrypted and obfuscated data on server side. Once the required data is filtered on server side, it is transferred on client side where the de-obfuscation and decryption is performed. Experiment results are also highlighted showing the enhancement in performance due to obfuscation factor. This paper considered the SaaS application. But this proposal is adapt only to PaaS model.

In [7] Dr. Nashaat el-Khameesy and Hossam Abdel Rahman proposed a security policy and procedures explicit to enhance the Data storage security in the cloud. They had a Control Access Data Storage (CADS) that included the necessary policies, processes and control activities for the delivery of each of the Data service offerings. The collective control Data Storage encompasses the users, processes, and technology needed to maintain an environment that supports the effectiveness of specific controls and the control frameworks. The security, correctness of data made in cloud storage, and availability of the data files being stored on the distributed cloud servers. It must be guaranteed by providing security policy and procedure for Data Storage, Defense in Depth for Data Storage in the cloud, Correctness Verification and Error Localization computing. All these recommendations are only theoretically proposed.

R. Anitha et al. proposed a method for providing protection to the data stored at the data server through metadata in [8]. This process provides protection using a cipher key which is created from the features of metadata. In this model, the time required for generating the cipher key is proportional to the number of attributes in the metadata as well the algorithms used for cipher key generation. Their plan enforced safety by providing two novel features; 1. Security is provided by their proposed design, where the encryption and decryption keys cannot be compromised without the involvement of data owner and the Metadata Data Server (MDS). 2. The cipher key generated using the modified feistel network holds good for the avalanche effect as each round of the feistel function depends on the previous round value. This approach is time consuming for generation of cipher key.

B. Raja Sekhar et al. [9] introduced the Ciphertext policy attribute-based encryption (CP-ABE) which is a promising cryptographic solution to ensure the data security and integrity in cloud storage. It allows data owners to define their own access policies over user characteristics and enforce the policies on the data to be distributed. It provides a way of defining access policies based on various characteristics of the requester, background, or the data object. Especially, ciphertext-policy attribute based encryption (CP-ABE) enables an encryptor to define the attribute set over a universe of attributes that a decryptor needs to possess in order to decrypt the ciphertext, and enforce it on the contents. Thus, each user with a different set of attributes is allowed to decrypt several pieces of data as per the security policy.

Siani Pearson et al. described a privacy manager [10] for cloud computing, which reduces the risk to the cloud computing user of their private data being stolen or misused. As a first line of defense, the privacy manager uses a feature called obfuscation. The idea is that instead of being present unencrypted in the cloud, the user's private data is sent to the cloud in an obfuscated form. The obfuscation method uses a

key which is chosen by the user and known by the privacy manager, but which is not communicated to the service provider. Thus the service provider is not able to de-obfuscate the user's data, and this data is not present on the service provider's machines, reducing the risks of theft of this data from the cloud and unauthorized uses of this data. Moreover, the obfuscated data is not personally identifiable information, and so the service provider is not subject to the legal restrictions that apply to the processing of the unobfuscated data. However, it is not practical for all cloud applications to work with obfuscated data. For applications for which users have to upload some private data to the cloud, the privacy manager contains two additional features, called *preferences* and *personae*, which help the users to communicate to service providers their wishes for the use of this personal data, and thus assist the service providers to respect privacy laws requiring users' consent.

4. PROPOSED CONFIDENTIALLY TECHNIQUE

Cloud computing provides an efficient storage setting to store and retrieve the cloud users critical data. Ensuring data security is a vital role to cloud users as well as cloud providers. This paper uses the confidentiality parameter to address the data security problems. Figure 2 represents the cloud storage confidentiality protection system using encryption and obfuscation technique. All the data must be encrypted or obfuscated before it is sent to the cloud database. Based on the type of the data, encryption or obfuscation can be used. Once the data is applied by proposed confidentiality technique, then the data is submitted to the cloud storage. Encryption and obfuscation of cloud data is done in the user side. The key used for encryption algorithm is generated in the user environment.

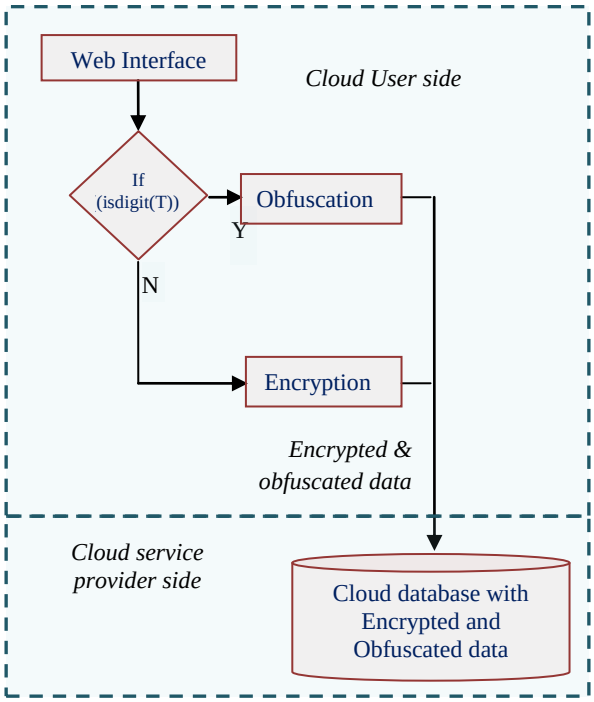

Fig 2: Proposed Technique for cloud storage confidentiality using Encryption and obfuscation

Generally, Confidentiality is ensured by encryption algorithm. For cloud data storage, Symmetric encryption is best choice, because symmetric encryption has the speed and computational efficiency to handle encryption of large volumes of data [11]. Along with the encryption algorithm [12], this paper also uses the obfuscation technique to improve the data confidentiality in cloud storage.

Algorithm #1 is used to find out the type of data (T) which is ready to be stored in the cloud storage. Based on the type of data, encryption or obfuscation is applied on the data before forwarding to the cloud. If the data (T) are digits, then obfuscation technique is applied, if the data are alphabets or alphanumeric or special symbols then encryption is applied on the data. This algorithm will call the corresponding algorithm based on the data type of (T).

Algorithm #1

1. *start*
2. *T=plaintext*
3. *if (isdigits(T)) then*
   obfuscation_digits(T)
4. *else*
   encryption_text(T)
5. *end if*
6. *end*

Algorithm #2 is used for obfuscation. This algorithm is used for numeric data type. Obfuscation is a technique by applying through specific mathematical functions or by using programming techniques. Obfuscation doesn't use any key to mask the user's data. But deobfuscation is only happened with a key which is generated during the process of obfuscation. Several obfuscation techniques are available in the literature. This paper uses two mathematical function namely pow() and round() functions. The pow() function is used to convert T into its square value (D). The round() function is used to rounded up the value of D into its nearest whole value (RD). From the rounded value (RD), find the printable ASCII value by subtracting 256 from RD value until the RD value reaches the range between printable ASCII values (32-127). Finally the resultant value is converted into ASCII character value. The example result of obfuscation algorithm is shown in table 3.

Algorithm #2

1. *obfuscation_digits(T)*
2. *start*
3. *for i=1 to size of(T)*
4. *D(i)=pow(T(i),2)*
   //find square value of plain text value
5. *RD(i)=round(D(i))*
   // Round off the square value
6. *loop*
7. *for RD(i)<32||RD(i)>126*
   //Find the printable ASCII value for RD by subtract
   //256 until its value comes in between 32 to 126
8. *RD(i)=RD(i)-256*
9. *count=count+1*
10. *loop*
    //count the no. of subtraction happened; this count
    //value will be the key for deobfuscation
11. *Check the resultant value in the range of printable ASCII value*
12. *Convert the value into ASCII code (C)*
    // C-Cipher Text
13. *end*

Algorithm #3 is used for encryption. This algorithm is used for alphabets or alphanumeric or special symbols. This is a symmetric encryption algorithm. The algorithm uses substitution and transposition technique to convert the plain text into cipher text. ASCII codes of the plain text are used throughout the algorithm. It uses four keys for encryption, and same keys are used for decryption also. The given plain text characters are converted into ASCII values. A square matrix is formed based on the number of character in the plaintext. The square matrix is divided into three matrices called upper (*UMAT*), diagonal (*DMAT*) and lower (*UMAT*) matrix. Apply the encryption to the matrices *UMAT*, *DMAT* and *UMAT* individually by using the key K_1, K_2, K_3 respectively. Form a square matrix with encrypted value. Now read the text by column by column, order of reading the column is based on the Key K_4. Finally the ASCII code values are converted into character value, this value is cipher text.

Algorithm #3

1. *encryption_text(T)*
2. *start*
3. *Convert (T) into ASCII code*
4. *N= count (T)*
   // N-no.of character in T
5. *Based on the value of N, form a square matrix MAT [MXM] > N,M=M*
   // M-order of matrix
6. *Apply T into the matrix from left to right*
7. *Divide the Matrix MAT into three matrices called UMAT,DMAT,LMAT*
   //UMAT-Upper Matrix
   //DMAT-Diagonal Matrix
   //LMAT-Lower Matrix
8. *Read the Text T by UMAT(U), DMAT(D) and LMAT(L)matrix*
   //U, D, L- text of upper, diagonal and lower matrix respectively
   // generate the random number for keys
9. *for i=1 to 3*
10. *Ki=random_num_gen()*
11. *loop*
12. *Get three random integer number as KEYS K_1, K_2, K_3 for each matrix.*
13. *Apply the key K_1, K_2, K_3 for U, D, L*
    // [U-K_1, D-K_2, L-K_3]
14. *Apply the resultant text (from step 10) into another matrixMAT_1 [MXM]*
15. *K_4 =random_str_gen()*
    //Generate random string value as a key called K_4
16. *Order the matrix based on the key K_4*
17. *Read the matrix by column using the order of key K_4*
18. *Resultant text from step 14 is converted from the ASCII code into character(C)*
    //C-Cipher Text
19. *end*

Algorithm #4 is used for generation of random integer number. This algorithm generates a random value each time it invoked. Algorithm#3 invokes algorithm#4 for three times to get three random numbers as keys K_1, K_2 and K_3.

Algorithm #4

1. *int random_num_gen()*
2. *start*
   // generation of random using random() class
3. *ran=new random().nextInt(100)*
4. *return ran*
5. *end*

Algorithm #5 is used for generation of random string value. This algorithm generates a string and returns it to encryption algorithm #4. This string value is used as key K_4.

Algorithm #5

1. *String random_str_gen()*
2. *Start*
3. *for i=1 to 3*
4. *ran=new random().nextInt(126)+32*
   // generate the random number
5. *ran_str_buff=(char)ran*
   //convert the number into equivalent char
6. *loop*
   // return the string buffer value by converting it into string
7. *ran_str=ran_str_buff.toString*
8. *return ran_str*
9. *end*

Proposed cryptography technique uses encryption and obfuscation for the different type of data. Integration of obfuscation technique with encryption technique has given more confidentiality than when they are used separately. Confidentiality of cloud data is ensured by using this technique. This can protect the data in cloud storage from insiders as well as outsiders attack.

For simple understanding of the proposed cryptographic technique, consider a sample transactional table as shown below Table 1, to be stored in the cloud storage. This Table 1 values are encrypted and obfuscated by the proposed cryptographic technique.

Table 1. Transactional table with plain text

Trans_Id	Cust_Id	Item_Name	Quantity	Total_P
TId_1003	A230kum	Lux	20	200
TId_923	B301sus	Himalaya	17	1500
TId_2304	C100mon	Bovonto	3	145
TId_9087	B002lav	Laxmi	9	100
TId_0012	G123aro	Medicine	12	60
TId_9999	X987ren	Chocolate	50	15

Proposed cryptography technique is implemented in JAVA running in windows 7 operating system. Table 1 represents the plain text values of sample transactional table of a shopping mall. Table 2 represents the encrypted values of the table 1 using algorithm #3 only. Table 3 represents the encrypted and obfuscated values of the table 1 using algorithm #2 and #3.

Table 2. Transactional table with plain text using encryption

]zUdbju|p	L!pvL|xh"	RymhQnwhy	Zz!dlj x}&	]m\wb}ru{
]k<g3mL:?	J<y6x<5t>	U,{{k"x|	;9>.3.	;;<.3.
]k?g5mLBU	K= 3x96|?	Qm&pdvlum	:@=.:.	:9<.8.
]k<g6mL;@	L<z3r94v=	K{{yw rw{	<..	::A.7.
]kDg3mLBC	K>#3d93u<	Uy%{drdry	B..	::<.3.
]k=g3mL9>	P?{5u;4j=	Vuzglmhlq	:;=.5.	?9B.3.
]kEg<mLBE	aCz;hA<{E	Lomronkx!	>9A.3.	:>=.8.

Table 2 shows the encrypted values of all columns in Table 1. In this table encryption algorithm alone is applied for any

types of value. Encrypted values in Table 2 shows similar values for same plain text occurred in different places of Table 1.

Based on the proposed technique, encryption and obfuscation can be applied on the Table 1. Alphabets and alphanumeric types of data are encrypted, and numeric types of data are obfuscated. The Table 3 shows the cipher text value of Table 1 using encryption and obfuscation.

User's data like the transactional Table 1 is submitted to the cloud storage in the form of encrypted and obfuscated shown in Table 3. This will increase the data security in the cloud storage.

Table 3. Transactional table with cipher text using encryption and obfuscation

]zUdbju|p	L!pvL|xh"	RymhQnwhy	Zz!dljx}&	]m\wb }ru{
]k<g3mL:?	J<y6x<5t>	U,{{k"x|	2	@
]k?g5mLBU	K= 3x96|?	Qm&pdvlum	!	j
]k<g6mL;@	L<z3r94v=	K{{yw rw{	)	!
]kDg3mLBC	K>#3d93u<	Uy%{drdry	Q	0
]k=g3mL9>	P?{5u;4j=	Vuzglmhlq	2	0
]kEg<mLBE	aCz;hA<{E	Lomronkx!	F	c

By the observation of the above three tables, it is evident that the integration of encryption and obfuscation will give more security to the data stored in the cloud. Table 2 shows only encrypted value, if the hackers find logic, they will apply it to the entire field in table to get the whole record. But in case of proposed technique same logic is not applicable for all the fields in the table. So hackers will not be able to get the actual values from table. This definitely increases the confidentially of the data stored in the cloud.

5. CONCLUSION

Cloud computing is profitable computing services to an individual and enterprise customers. But due to some of security problem in it, people might be reluctant to use it. Once the issues are resolved, cloud computing will be the trillion dollars business in the computing world. The Data storage on un-trusted cloud makes data security as a challenging problem. Data security in the cloud is ensured by the confidentiality of sensitive data. This paper proposed a new cryptographic technique which is applied to address this problem. Encrypted data are stored on storage servers while secret key(s) are retained by data owner; access to the user is granted by issuing the corresponding data decryption keys. Along with encryption, obfuscation technique is used to increase the confidentiality of data. Algorithms are proposed for encryption and obfuscation technique. The user data are encrypted or obfuscated before they are forwarded to the cloud storage. Encryption only or obfuscation only is not sufficient for cloud data storage. Hence in this paper, a new confidentiality technique namely Obfuscrypt techniques, has been proposed and implemented. From the results obtained, it is observed that this technique could offer better security to data stored on a cloud than the existing techniques that are based on encryption, obfuscation alone.

6. REFERENCE

[1] Uegebe Ikechukwu Valentine and Omenka Ugochukwu Enyinna, " Building Trust and Confidentiality in Cloud computing Distributed Data Storage", West African Journal of Industrial & Academic Research, Volume 6 Issue 1 March 2013, pp78-83.

[2] Dr. L. Arockiam, S. Monikandan, G.Parthasarathy, Cloud Computing: A Survey, International Journal of Internet Computing, ISSN No: 2231 – 6965, Volume 1, Issue 2, pp. 26-33, 2011.

[3] Xiaojun Yu, Qiaoyan Wen, "A View about Cloud Data Security from Data Life Cycle", International Conference on Computational Intelligence and Software Engineering (CiSE), IEEE, Dec 2010, pp 1-4.

[4] Masayuki Okuhara et al, "Security Architecture for Cloud Computing", FUJITSU Sci. Tech. J., Volume 46, Issue 4, October 2010, pp. 397-402.

[5] Atiq ur Rehman, M.Hussain, "Efficient Cloud Data Confidentiality for DaaS", International Journal of Advanced Science and Technology Volume. 35, October 2011, pp 1-10.

[6] Yvette E. Gelogo and Sunguk Lee, " Database Management System as a Cloud Service, International Journal of Future Generation Communication and Networking Volume 5, Issue 2, June 2012, pp 71-76.

[7] Nashaat el-Khameesy, Hossam Abdel Rahman, " A Proposed Model for Enhancing Data Storage Security in Cloud Computing Systems", Journal of Emerging Trends in Computing and Information Sciences, Volume 3, Issue 6, June 2012, pp 970-974.

[8] R. Anitha, P. Pradeepan, P. Yogesh, and Saswati Mukherjee, "Data Storage Security in Cloud using Metadata", 2nd International Conference on Machine Learning and Computer Science(IMLCS'2013), Kuala Lumpur (Malaysia), August 2013, pp 26-30.

[9] B. Raja Sekhar, B. Sunil Kumar, L. Swathi Reddy, and V. Poorna Chandar "CP-ABE Based Encryption for Secured Cloud Storage Access", International Journal of Scientific & Engineering Research, Volume 3, Issue 9, September 2012, pp 1-5.

[10] Siani Pearson, Yun Shen and Miranda Mowbray, "A Privacy Manager for Cloud Computing", CloudCom '09 Proceedings of the 1st International Conference on Cloud Computing, Springer-Verlag Berlin, Heidelberg, 2009, pp 90 – 106.

[11] Tim Mather, Subra Kumaraswamy, and Shahed Latif, "Cloud Security and Privacy", O'Reilly Media, Inc., chapter 4, September 2009, pp 61-71.

[12] Dr. L. Arockiam, S. Monikandan, "Data Security and Privacy in Cloud Storage using Hybrid Symmetric Encryption Algorithm", International Journal of Advanced Research in Computer and Communication Engineering, Volume 2, Issue 8, August 2013, pp 3064-3070.

[13] S Kamara, K Lauter, "Cryptographic Cloud Storage", IFCA/ LNCS 6054, Springer-verlag, Berlin Heidelberg, 2010, pp 136-149.