International Journal of
Emerging Research in Management &Technology
ISSN: 2278-9359 (Volume-4, Issue-8)

Research Article | **August 2015**

Formation of Secure and Adaptable Cloud Data Encryption System in Cloud Computing

Anamika Sirohi, Vishal Shrivastava
Department of Computer Science and Engineering,
ARYA college of Engineering and IT, Jaipur, Rajasthan, India

Abstract:

Cloud computing is often considered the successor of grid computing. Cloud security is a multifaceted and highly complex issue. The data owners especially of large organizations fear possible data misuse by the cloud provider without their knowledge. This concern an issue related to data security in cloud environment including confidentiality, integrity, authentication and authorization our cloud security model plans to keep the most critical data security in cloud computing at different levels like user level, cloud service provider level, third party level and network intruder level. The Major issues are data security, data leakage, data privacy, data confidentiality and integrity. Due to which users are not able to fearlessly upload their data to cloud. To solve this problem we proposed a model which is highly secure and is based on data owner centric model i.e. data is under control of data owner. Encryption, Obfuscation, HMAC and Dual authentication and access management technique has been used which make the proposed model more reliable and effective to use it in real world.

Keywords: Cloud computing; Encryption; Hash MAC; Data Privacy Protection, Public and Private Cloud.

I. INTRODUCTION

Cloud computing has been coined as an umbrella term to describe a category of sophisticated on-demand computing services initially offered by commercial providers, such as Amazon, Google, and Microsoft. It denotes a model on which a computing infrastructure is viewed as a "cloud," from which businesses and individuals access applications from anywhere in the world on demand [2]. The main principle behind this model is offering computing, storage, and software "as a service.

Many practitioners in the commercial and academic spheres have attempted to define exactly what "cloud computing" is and what unique characteristics it presents. It have defined it as follows: "Cloud is a parallel and distributed computing system consisting of a collection of inter-connected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements (SLA) established through negotiation between the service provider and consumers." Vaquero et al. [3] have stated "clouds are a large pool of easily usable and accessible virtualized resources (such as hardware, development platforms and/or services). These resources can be dynamically reconfigured to adjust to a variable load (scale), allowing also for an optimum resource utilization. This pool of resources is typically exploited by a pay-per-use model in which guarantees are offered by the Infrastructure Provider by means of customized Service Level Agreements." A recent McKinsey and it [4] claims that "Clouds are hardware based services offering compute, network, and storage capacity where: Hardware management is highly abstracted from the buyer, buyers incur infrastructure costs as variable OPEX, and infrastructure capacity is highly elastic."

We can track the roots of clouds computing by observing the advancement of several technologies, especially in hardware (virtualization, multi-core chips), Internet technologies (Web services, service-oriented architectures, Web 2.0), distributed computing (clusters, grids), and systems management (autonomic computing, data center automation). Figure 1.1 shows the convergence of technology fields that significantly advanced and contributed to the advent of cloud computing. Some of these technologies have been tagged as hype in their early stages of development; however, they later received significant attention from academia and were sanctioned by major industry players. Consequently, a specification and standardization process followed, leading to maturity and wide adoption. The emergence of cloud computing itself is closely linked to the maturity of such technologies. We present a closer look at the technologies that form the base of cloud computing, with the aim of providing a clearer picture of the cloud ecosystem as a whole.

Three major milestones have led to cloud computing: mainframe computing, cluster computing and grid computing. Mainframes: These were the first examples of large computational facilities leveraging multiple processing units. Mainframes were powerful, highly reliable computers specialized for large data movement and massive input/output (I/O) operations. They were mostly used by large organizations for bulk data processing tasks such as online transactions, enterprise resource planning, and other operations involving the processing of significant amounts of data. Clusters: Cluster computing started as a low-cost alternative to the use of mainframes and supercomputers. The technology advancement that created faster and more powerful mainframes and supercomputers eventually generated an increased availability of cheap commodity machines as a side effect. Starting in the 1980s, clusters become the standard technology for parallel and high-performance computing.

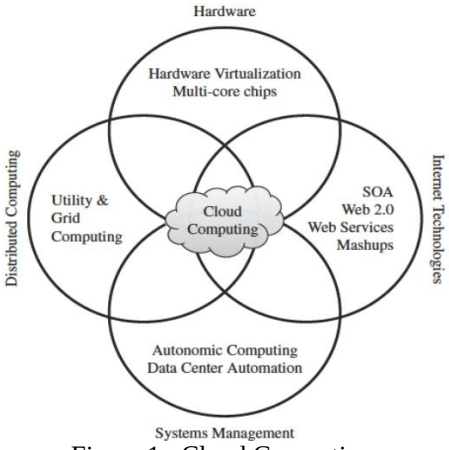

Figure 1: Cloud Computing

Cloud computing represents a distributing computing mechanism that by the utilize of the high speed network, data processing is moved from private PC or servers to the remote computer clusters (big data centers owned by the cloud service providers), any user has a potential super computer at hand and can access the data and get the computing capability at any time, from anywhere, you only need to pay for the resources which you have used, don't care about who provide the resources and in what way.

Actually, clouds [3] are Internet-based and it tries to disguise complexity for clients. Cloud computing refers to both the applications delivered as services over the Internet and the hardware and software in the datacenters that provide those services. Cloud providers use virtualization technologies combined with self-service abilities for computing resources via network infrastructure. In cloud environments, several kinds of virtual machines are hosted on the same physical server as infrastructure. In cloud, costumers must only pay for what they use and have not to pay for local resources which they need to such as storage or infrastructure.

HASH MESSAGE AUTHENTICATION CODE (HMAC)

Hash message authentication code (HMAC), a mechanism for message authentication using cryptographic hash functions. HMAC can be used with any iterative FIPS-approved cryptographic hash function, in combination with a shared secret key. The cryptographic strength of HMAC depends on the properties of the underlying hash function. (MSE).

An HMAC function is used by the message sender to produce a value (the MAC) that is formed by condensing the secret key and the message input. The MAC is typically sent to the message receiver along with the message. The receiver computes the MAC on the received message using the same key and HMAC function as was used by the sender, and compares the result computed with the received MAC. If the two values match, the message has been correctly received, and the receiver is assured that the sender is a member of the community of users that share the key.

H MAC uses the following parameters:

B-Block size (in bytes) of the input to the FIPS-approved hash function; e.g., for
SHA-1, B= 64.

H- FIPS-approved hash function, e.g., FIPS 180-1, Secure Hash Algorithm-1 (SHA-1).

Ipad- Inner pad; the byte x'36' repeated B times.

K- Secret key shared between the originator and the intended receiver(s).

K0-The key K with zeros appended to form a B byte key.

L- Block size (in bytes) of the output of the FIPS-approved hash function; for SHA-1,L= 20.

Opad- Outer pad; the byte x'5c' repeated B times.

T- The number of bytes of MAC.

Text- The data on which the HMAC is calculated; the length of the data is n bits, where the maximum value for n depends on the hash algorithm used.

X'N'-Hexadecimal notation, where each 'N' represents 4 binary bits.

||-Concatenation and

Exclusive-Or operation.

II. LITERATURE SURVEY

In this paper, I have made a review on my topic data security in cloud computing at different levels by reading different kinds of papers and analyzing different techniques which are being used in these papers published by authors which are discussed as follows:

Sood et.al [3] proposed approach to ensure data security in cloud computing. In this proposed approach key generation, encryption, indexing of data, user authentication and data integrity is performed by data owner itself. Unfortunately, there will be high overhead on data owner and hence time consuming too. Thilakanathan et.al [4] proposed scheme using proxy re-encryption for security of data. In this scheme data owner encrypt the data using his key piece then proxy encrypt the data using his key piece. Decryption is also carried in similar fashion. However, if proxy is fake then data

becomes insecure. Sharma et.al.[5] discussed different service model of cloud computing and highlights the key security issues, challenges and solution at different layers of cloud. Jingwei et.al.[6] discussed efficient model for secure data sharing in cloud. The proposed model consists of user, authority, hybrid cloud and owner. The data is stored at private cloud and data shared is encrypted Encryption technology used is keyword-based encryption. The keys are generated by authority and given to user group for encryption and decryption. The model has some issues like if authority is fake then data is insecure and also it is costly to use the model. Sood et.al [7] proposed the scheme to highly secure the data at cloud. They provided improved data security by using concept of hybrid cloud. In this scheme the sensitive data i.e. about 3%-5% is stored at private cloud and rest of the data at public cloud. This model is applicable to organisations whose sensitive data is about 3%-5%. If the sensitive data increases then this model will prove to be expensive. The white papers [8] of many organisations describes three types of data security models in cloud. First model
Consists of key generation and encryption on data is performed by data owner itself. However this model results in high overhead for data owner. Second model describes encryption performed by data owner and key generation by cloud service provider. Unfortunately, cloud service provider is fake then data is insecure hands. Third model encryption and key generation is control by cloud service provider. If cloud service provider is fake then data is endangered. Hwang et.al [9] proposed business model in which encryption/decryption service and storage as a service of user data were separated i.e. they were not provided by single operator. After encryption/decryption performed system should delete all the data. Varalakshmi et.al [10] proposed system consists of three entities cloud broker, client and cloud storage. Broker handles encryption, hash key, decryption and local database management. According to cloud space available the client files are partitioned into segment and hash values of segments has been generated. When the client needs its file it sends request to broker then broker download the file, partition the file into segments and then calculate the hash values. For checking the data integrity hash values before uploading to the after downloading are matched.

III. RESEARCH METHODOLOGY

Proposed model is based on main responsibility of data owner towards data protection at cloud. For data security, encryption and obfuscation technique is used which protect the data during transits as well as at rest. During transit of data, Data integrity plays vital role so for data integrity hash based message authentication. Code is calculated on encrypted data in order to have minimum overhead over data owner. So third party is involved.
The proposed model is divided into two categories i.e.
Phase-1(uploading)
Phase-2(downloading)
And involved four entities that is - data owner, CSP, third party and user.
1 Phase -1 (Uploading)

1.1 Key Generation and maintenance

In this model third party acts as key management infrastructure for key generation and storage. The third party generates the symmetric key and gives the key to data owner for further process. Data owner divide the key into key pieces. Owner keeps its one key piece for encryption and other key piece for corresponding user id for future use. Data owner encrypt these two keys by passcode and then sends to the third party. These key pieces are kept with third party and taken when needed.

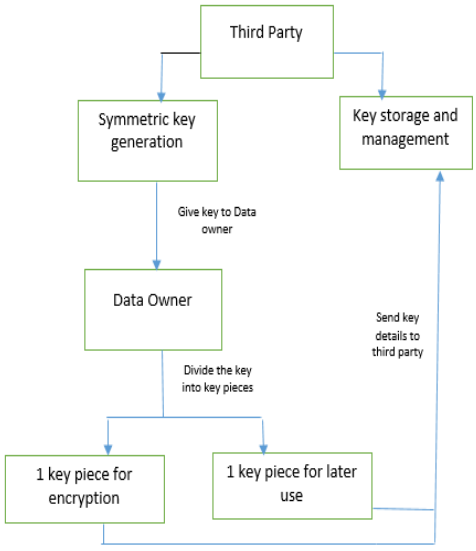

Figure 2. Key Generation and maintenance

1.2 Classification of data

The data can be of two types i.e. type 0 and type 1
Type 1:- when the data is of alpha numeric type
Type 0:- when the data is of numeric type

1.3 Encipher and Indexing

Based on classification of data, corresponding encipher technique is used. Data owner identifies the type of data. If data is of type 1 then encryption is used otherwise obfuscation is used. Before these techniques, indexing is performed. After indexing as well as data applied according to technique are uploaded to cloud.

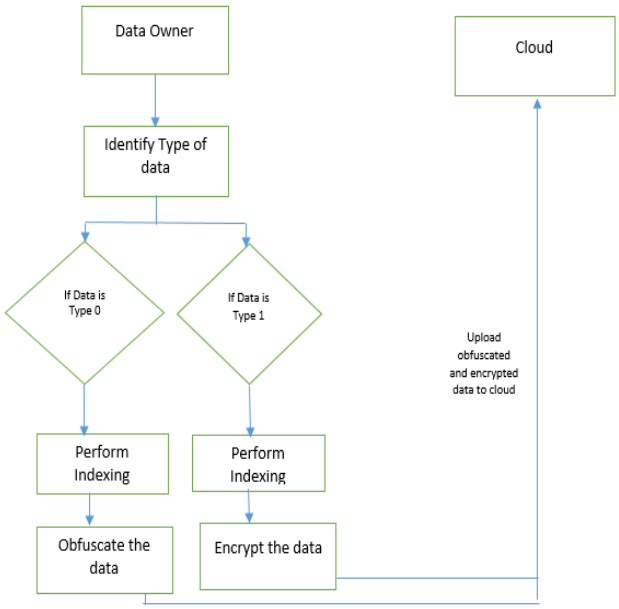

Figure 3. Encipher and Indexing

Message Authentication Code (MAC) Generation

Step1: Message Authentication Code (MAC) is generated on encrypted data using Message-Digest algorithm. Openssl dgst –md5 filename

```
[dataowner@server20 .ssh]$ openssl dgst -md5 file2
MD5(file2)= 0064c057046f128e32e0151e41c787a3
[dataowner@server20 .ssh]$ _
```

Figure4: Message Authentication Code (MAC) Generation

Step2: Encrypt the md5 output same as done before in data encryption

openssl rsautl –encrypt –pubin –inkey PUBLIC_KEY.pem –in md5 –out md

```
[dataowner@server20 .ssh]$ ls
file1  file2  md5  PUBLIC_KEY.pem
[dataowner@server20 .ssh]$ openssl rsautl -encrypt -pubin -inkey PUBLIC_KEY.pem
-in md5 -out md
[dataowner@server20 .ssh]$ ls
file1  file2  md  md5  PUBLIC_KEY.pem
[dataowner@server20 .ssh]$ cat md
x$P^a▪ct▪X9jT+▪▪_Ge▪
▪▪2C▪▪sh▪q▪▪▪▪at▪▪▪IUĎ▪▪ ▪S      ▪~▪▪▪0▪[▪=¡R▪1▪l▪}▪dF▪▪O@▪)B▪▪▪▪▪/▪u}▪▪>I▪▪{~▪▪F
▪▪▪/▪!▪(▪[dataowner@server20 .ssh]$
[dataowner@server20 .ssh]$ _
```

Figure5: Message Authentication Code (MAC) Encryption

Role-Based User Dual Authentication

Dual Authentication of user is carried by third party first and further by data owner.

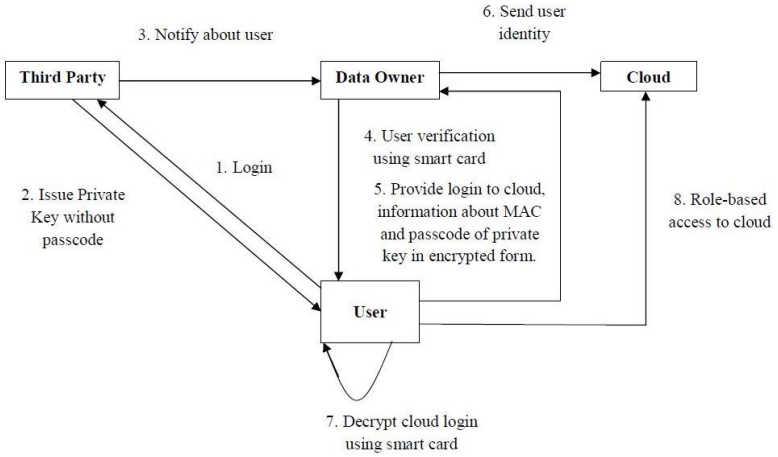

Figure 6: Role-Based Dual User Authentication

Message Authentication Code Verification

Now user has downloaded encrypted data and Message Authentication Code. Users decrypt Message Authentication Code first and calculate Message Authentication Code on encrypted data. If both the Message Authentication Code i.e. decrypted Message Authentication Code and calculated Message Authentication Code are same then decrypt the data and use it otherwise report to data owner about Message Authentication Code mismatch.

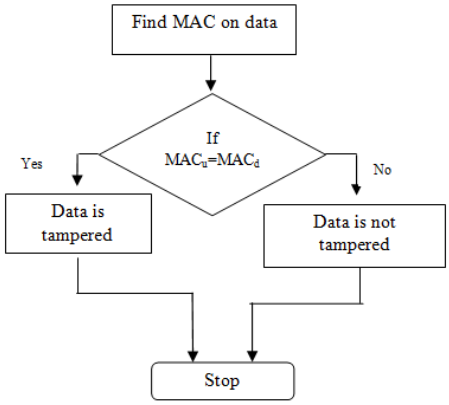

Figure 7: Data Integrity

Role Based Access to Cloud

Step1: Data-owner have full control over cloud i.e. data owner act as administrator. Data-owner can add new user, delete user and revoke user anytime.

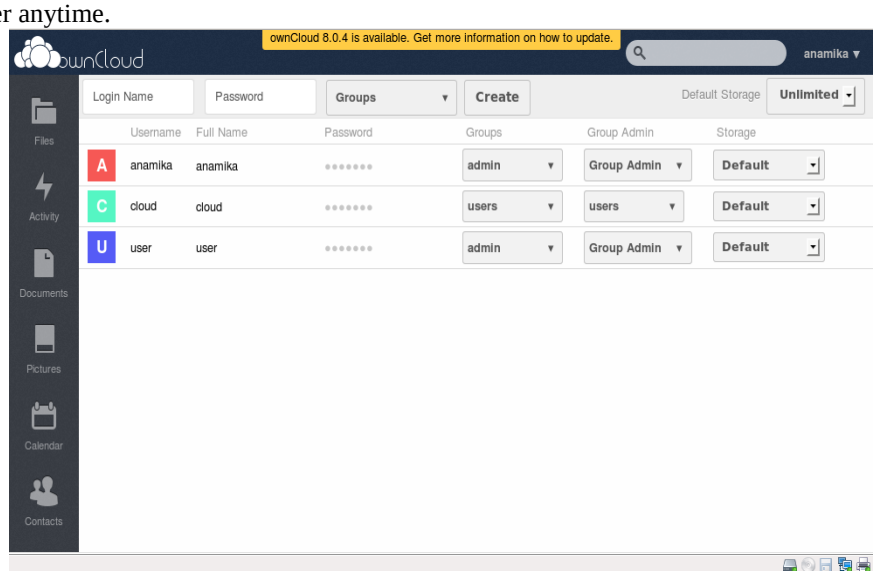

Figure8: User Creation and Deletion by data owner on cloud

Step2: Data-owner can add no of files according its wish and can upload to maximum file size 2GB.

Step3: As it is role based access data owner can share its file with restriction of modification. File2 and md are shared with user.

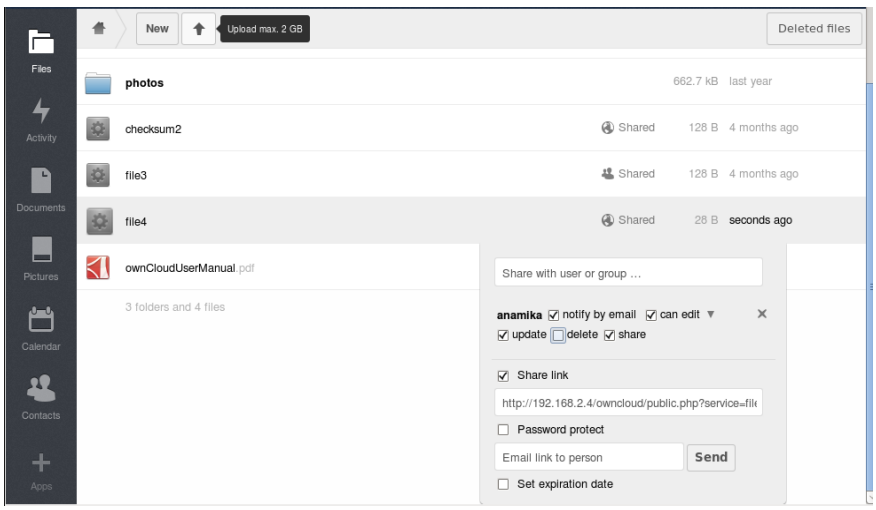

Figure9: Sharing of files with Users by data owner

Step4: User decrypts the md file for data integrity and generates message authentication code on encrypted file2. Matches both MAC, if same then decrypt the file2 and use the data otherwise report to data owner.
Step5: User login to cloud with login id and password as provided by data owner.

IV. RESULTS ANALYSIS

Proposed model has been organized so that it give throughout data security in cloud computing at different levels. The different levels are: user level, cloud service provider level, third party level and network intruder level. Data is remains private against all level. This model is data owner centric with least overhead and highly secure to adopt in real life while storing and retrieving data from cloud. This model is designed in way to protect data from every aspect. Security analysis of model has been performed at different levels. The proposed model has been evaluated with implementation. This model has been verified using OpenSSL tool [17] in red hat Linux and own Cloud[18]. Figure 5 shows that after implementation various security parameters i.e. encipher and indexing, classification of data, HMAC and dual user authentication. HMAC provides less data security than classification of data and this classification of data provide less security than encipher and indexing technique. Basically if we combine all security parameters i.e. HMAC, classification of data, encipher and indexing and dual user authentication, it results in highly secured data owner centric approach for uploading data to cloud. It results in highly secured proposed model used in various cloud environment which is denoted as peak value as shown in figure.

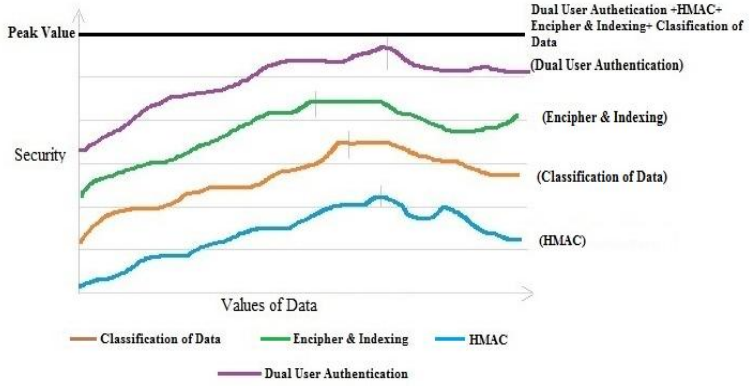

Figure 10. Security Evaluation

V. CONCLUSION & FUTURE SCOPE

The proposed model provides a way to protect the data, check the integrity and authentication by following the best possible industry mechanisms. In these model data security is checked where ever there is possibility of data threat. Recent trend in cloud computing shows that data security is major issue which has been an obstacle for adopting the cloud. To overcome this issue, model has been proposed. The model is highly secure and protects the data during transit as well as data at rest. It also secures the data against all threats i.e. insight as well as outsight. It helps the user to fearlessly upload the data at cloud without any hesitation of data being lost or steeled. Although the model is secured but in future we will try to add on more security parameter so as to make model more secure and efficient.

REFERENCES

[1] Mrinal RajkumarBuyya, Christian Vecchiola and S. ThamaraiSelvi, Mastering Cloud Computing Foundations and Applications Programming.Morgan Kaufmann, USA.
[2] Jing Huang Jing, LI Renfa, and TangZhuo, "*The Research of the Data Security for Cloud Disk Based on the Hadoop Framework*", IEEE,2013.
[3] Sandeep K. Sood, "*A Combined Approach to Ensure Data Security in Cloud Computing*", Submitted to Journal of Network and Computer Applications, Elsevier Ltd, 2012.
[4] Danan Thilakanatha, Shiping Chen,Surya Nepal, Rafael A. Calvo and Leila Alem, "*A platform for secure monitoring and sharing of generic health data in the Cloud"*, Elsevier Ltd, 2013.
[5] Pardeep Sharma, Sandeep K. Sood, Sumeet Kaur, "*Cloud Implementation Issues and What to Compute on Cloud*", International Journal of Advances in Computer Networks and its Security, vol.1, no. 1, pp. 130-135, 2011.
[6] Jingwei Li, Jin Li, Zheli Liu and Chunfu Jia "*Enabling efficient and secure data sharing in cloud computing*" Concurrency Computat.:Pract Exper.,John Wiley & Sons, Ltd.,2013.
[7] Sandeep K. Sood, "*A Highly Secure Hybrid Security model for Data Security at Cloud* ", Submitted to Security and Communication Networks, John Wiley and Sons (Interscience), Special Issue on Trust and Security in Cloud Computing, 2012.
[8] Amazon Web Services.: "*Encrypting Data at Rest in AWS*", https://aws.amazon.com/whitepapers.
[9] Jing-Jang Hwang, Hung-Kai Chuang, Yi-Chang Hsu and Chien-Hsing Wu, "*A Business Model for Cloud Computing Based on a Separate Encryption and Decryption Service"*, National Science Council of Taiwan Government.

[10] P.Varalakshmi and Hamsavardhini Deventhiran, "*Integrity Checking for Cloud Environment Using Encryption Algorithm"*, IEEE, 2012.
[11] Eman M.Mohamed and Sherif EI-Etriby, "*Randomness Testing of Modem Encryption Techniques in Cloud Environment*", 8th International Conference on Informatics and Systems, 2012.
[12] Zhiqian Xu and Keith M. Martin, "*Dynamic User Revocation and Key Refreshing for Attribute-Based Encryption in Cloud Storage*", International Conference on Trust, Security and Privacy in Computing and Communications, IEEE, 2012.
[13] Kuan-Ying Huang, Guo-Heng Luo and Shyan-Ming Yuan, "*SSTreasury+: A Secure and* E*lastic Cloud Data Encryption System*", International Conference on Genetic and Evolutionary Computing, IEEE(2012).
[14] Chul Sur, Youngho Park, Sang Uk Shin, Changho Seo and Kyung Hyune Rhee, "Certificate-Based Proxy Re-Encryption for Public Cloud Storage", International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, IEEE, 2013.
[15] NarendraChandel, Sanjay Mishra, Neetesh Gupta and AmitSinhal, "Creation of Secure Cloud Environment using RC6", IEEE,2013.
[16] Miranda Mowbray and Siani Pearson, "Protecting Personal Information in Cloud Computing", Springer Verlag, 2012
[17] Chun-I Fan and Shi-Yuan Huang, "Controllable Privacy Preserving Search Based on Symmetric Predicate Encryption in Cloud Storage", International Conference on Cyber Enabled Distributed Computing and Knowledge Discovery, IEEE, 2011.
[18] Keiko Hashizume, David G Rosado2, Eduardo Fernández-Medina2 and Eduardo B Fernandez, "An analysis of security issues for cloud computing", Springer,2013.
[19] Swetha Reddy Lenkala, KaiqiXiong and Sachin Shetty, "Security Risk Assessment of Cloud Carrier", IEEE,2013.
[20] Marten van Dijk and Ari Juels, "On the Impossibility of Cryptography Alone for Privacy-Preserving Cloud Computing", ACM,2010.
[21] Balachandra Reddy Kandukuri, Ramakrishna Paturi V and AtanuRakshit, "Cloud Security Issues", IEEE 2009

