ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 2, March 2013

The Secure Data Storage in Cloud Computing Using Hadamard Matrix

Abhishek Mishra, Dileep Kumar Gupta, Dr. G. Sahoo
BIT Mesra Ranchi, Jharkhand, India

***Abstract:* Cloud computing is more popular because it can reduce the cost. It is not restricted to a particular location user can access the services from anywhere. But security becomes an important issue in the cloud computing. In this paper, we have focused on the storage as a service. Here, we have discussed the chaining technique that uses Hadamard transforms for encryption and decryption which has increased the level of security. This method is more effective when the key size will large and most of the elements in the key are distinct. But since the chaining of the Hadamard transforms can be as long as one pleases and since the choices as far as the formation of blocks that are converted into non-binary numbers is arbitrarily large, the effectiveness of this scheme is potentially very high. So the storage as a service becomes more effective when the level of security can be increased as per requirement.**

Index Terms— **Cloud computing, Hybrid cloud, Hadamard matrix, Hadamard transform., Private cloud, Public cloud.**

I. INTRODUCTION

Cloud computing is emerging as a prominent computing model. It provides a low cost, highly accessible alternative to other traditional high performance computing platforms. It has also many other benefits such as high availability, scalability, elasticity, and free of maintenance. Given these attractive features, it is very desirable if automated planning can exploit the large, affordable computational power of cloud computing. However, the latency in inter process communication in cloud computing makes most existing parallel planning algorithms unsuitable for cloud computing. Cloud computing is a type of computing, in which IT-related capabilities are provided “as a service” to end users. Users can access technology without knowledge of, experience of, or even control over the infrastructure that supports them. Basically, cloud computing offers three layers of services such as Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS). It is shown in Fig. 1[14].Organization of the paper is as follows: Types of cloud is discussed in section 2. Related work has been discussed in section 3. Storage as a service has been discussed in section 4. Types of information have been discussed in section 5. Classification criteria have been discussed in section 6. Cloud information security objectives have been discussed in section 7. Mathematical tools have been discussed in section 8. Algorithms have been discussed in section 9. Proposed model has been discussed in section 10. Conclusion and future work has been discussed in section 11.

A. SaaS (Software as a Service)

Software as a Service is a considerable change as we see software. There is no expenditure capital, only service cost. Software just like the processing power and storage is seen as a utility that clients can pay for only as needed. The goal is to centralize administrative tasks while improving scalability and workloads.

B. PaaS (Platform as a Service)

It offers a platform to clients for different purposes. For example, Windows Azure offers a platform to developers to build, test, and host applications that can be accessed by the end users. The end users may or may not know that the application is hosted on the cloud. The storage space for user data may be increased or decreased as per the requirement of the applications. As with the SaaS, users do not need to build the platform. Users just pay a nominal fee for using the service.

C. IaaS (Infrastructure as a Service)

It offers infrastructure on demand. The infrastructure can be anything from storage servers to applications to operating systems. Office 365 offers a combination of these infrastructure and falls under this category. With Office 365, user can get plenty of applications along with storage space. Buying infrastructure or renting it out in traditional models can be very expensive. When users opt for IaaS, they save a lot on expenses, space, and personnel required to set up and maintain the infrastructure. The cloud service provider takes care of setting up and maintaining the infrastructure. They just pay a fee to use it as per their requirements.

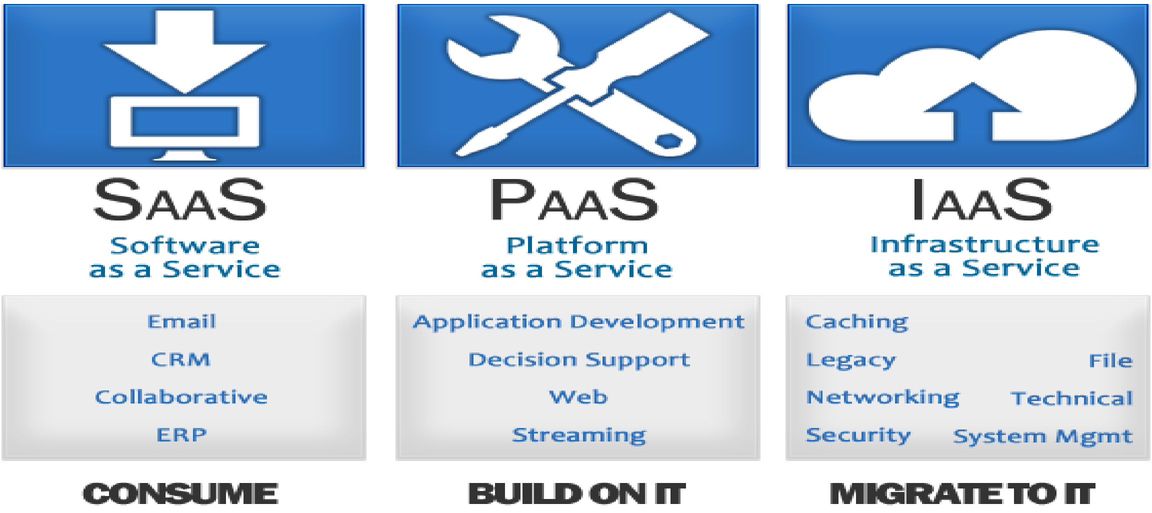

Fig.1: Types of cloud computing services [14]

II. TYPES OF CLOUDS

Following are the different types of clouds.

A. *Private cloud*

The cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party and may exist on premise or off premise.

B. *Community cloud*

The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g. mission, security requirements, policy, and compliance considerations). It may be managed by the organizations or a third party and may exist on premise or off premise.

C. *Public cloud*

The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.

D. *Hybrid cloud*

The cloud infrastructure is a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds).

III. RELATED WORK

Mukherjee and Sahoo [9] proposed the algorithm for encryption of sensitive data for C-governance along with their proper decryption to the authorized user by using of hadamard matrix. Yildiz et al. [10] have proposed a practical security model based on key security considerations by looking at a number of infrastructure aspects of Cloud Computing such as SaaS, Utility, Web, Platform and Managed Services, Service commerce platforms and Internet Integration. Itani et al. [11] have introduced a set of security protocols for ensuring the privacy and legal compliance of customer data in cloud computing architectures. Jaeger and Schiffman [12] have proposed in their paper to improve security of cloud architecture by building “verifiable base systems”. Zhang et al. [13] have presented information risk management framework for better understanding critical areas of focus in cloud computing.

IV. STORAGE AS A SERVICE (StaaS)

One of the primary uses of cloud computing is for data storage. With cloud storage, data is stored on multiple third-party servers, rather than on the dedicated servers used in traditional networked data storage. When storing data, the user sees a virtual server—that is, it appears as if the data is stored in a particular place with a specific name. But that place doesn’t exist in reality. It’s just a pseudonym used to reference virtual space carved out of the cloud. In reality, the user’s data could be stored on any one or more of the computers used to create the cloud. The actual storage location may even differ from day to day or even minute to minute, as the cloud dynamically manages available storage space. But even though the location is virtual, the user sees a static location for his data and it can actually manage his storage space as if it were connected to his own personal computer. Cloud storage has both financial and security advantages. Financially, virtual resources in the cloud are typically cheaper than dedicated physical resources connected to a personal computer or network. As for security, data stored in the cloud is secure from accidental erasure or hardware crashes, because it is duplicated across multiple physical machines; since multiple copies of the data are kept continually, the cloud continues to function as normal even if one or more machines go offline. If one machine crashes, the data is replicated on other machines in the cloud. In this paper, we have mainly focused on storage as a service in the cloud computing.

V. TYPES OF INFORMATION

For any organization the data is classified into the following categories.

A. Public

Information that is similar to unclassified information. All information of a company that does not fit into any of the next categories can be considered public. While its unauthorized disclosure may be against policy, it is not expected to impact seriously or adversely the organization, its employees and/or its customers.

B. Sensitive

This is the information that requires a higher level of security than public data. This information is protected from a loss of confidentiality as well as from a loss of integrity due to an unauthorized alteration. This classification applies to information that requires special precautions to ensure the integrity of the information by protecting it from unauthorized modification or deletion. It is information that requires a higher than normal assurance of accuracy and completeness.

C. Private

This classification applies to personal information that is intended for use within the organization. Its unauthorized disclosure could seriously and adversely impact the organization and its employees. For example, salary levels and medical information are considered private.

D. Confidential

This classification applies to the most sensitive business information that is intended strictly for use within the organization. Its unauthorized disclosure could seriously and adversely impact the organization, its stockholders, its business partners, and its customers.

This information is exempted from disclosure under the provisions of the Freedom of Information Act or other applicable federal laws or regulations. For example, information about new product development, trade secrets, and merger negotiations is considered confidential.

VI. CLASSIFICATION CRITERIA

There are several criteria’s that may be used to determine the classification of an information object.

A. Value

Value is the number one commonly used criteria for classifying data in the private sector. If the information is valuable to an organization or its competitors, it needs to be classified.

B. Age

The classification of information might be lowered if the information’s value decreases over time. In the Department of Defense, for example, some classified documents are automatically declassified after a predetermined time period is passed.

C. Useful life

If the information has been made obsolete due to new information, substantial changes in the company, or other reasons, the information can often be declassified.

D. Personal association

If information is personally associated with specific individuals or is addressed by a privacy law, it might need to be classified. For example, investigative information that reveals informant names might need to remain classified.

VII. CLOUD INFORMATION SECURITY OBJECTIVES

The Data and Analysis Center for Software (DACS) requires that service must exhibit the following three properties to be considered secure.

A. Dependability

This is the service that executes predictably and operates correctly under a variety of conditions including when under attack or running on a malicious host.

B. Trustworthiness

It is a service that contains minimum number of vulnerabilities or no vulnerability. It could sabotage the software's dependability. It must also be resistant to malicious logic.

C. Survivability (Resilience)

Survivability is a service that is resistant to or tolerant of attacks and has the ability to recover as quickly as possible with as little harm as possible. Beyond of this storage as a service is very useful for any organization. The organization spends lots of money to storing and maintaining their sensitive data. In the given paper, we give the model to provide secure storage as a service in cloud computing using hadamard matrix. But if we provide storage as service the security level may also varies according to the types of data and it must exhibit security objective.

VIII. MATHEMATICAL TOOLS

A. Hadamard matrix of order n

It is a N into N matrix having elements +1 and -1 such that any distinct row or column vectors are mutually orthogonal. A (normalized) Sylvester-Hadamard matrix [8] of size 2^{m}, m > 0, is a squared 2^{m} into 2^{m} matrix that is defined recursively by

$$H_{2^m} = \frac{1}{\sqrt{2}} \begin{pmatrix} H_{2^{m-1}} & H_{2^{m-1}} \\ H_{2^{m-1}} & -H_{2^{m-1}} \end{pmatrix}$$

Where recursion is initiated by $H_1 = (1)$

B. Hadamard transform

It may be defined either recursively, or by using binary representation. It is a generalize class of fourier transform[4] Its symmetric form lends itself to applications ranging across many technical fields such as data encryption [1], signal processing [2], data compression algorithms[3], randomness measures[6] and so on.

It is given by mapping T: R^{2m} to R^{2m} defined by T(x)=$H_2{}^m$ .x

C. Inverse Hadamard transform

Inverse matrix of Sylvester-Hadamard matrix is equal to its transpose So inverse Hadamard transform [8]is performed by applying $\mathbf{H^T{}_2{}^m}$

So x is given as

$$x= H^T{}_2{}^m .T(x) = H^T{}_2{}^m . H_2{}^m .x$$

IX. ALGORITHM

The encryption and decryption algorithms [7] used in the proposed model are given as follows.

A. Encryption Algorithm

1. In the encryption algorithm, first given binary input sequence and the key have considered. The given key can have n numbers such that each number say z in n is a prime then 2^z-1 is also a prime. Also, consider a two dimensional integer array. Here, number of rows is equal to number of elements in the array and number of columns is equal to number of input values at each row.

2. Next, first element in the key and group has considered the bits in the input sequence based on the number.

3. Now conversion of each group into corresponding decimal number has performed.

4. Divide the decimated sequence to equal length sub-sequences such that each sub-sequence length should be expressed as power of 2. This is depending on the length of input sequence
5. Append 0s at last to make the length equal to other sub sequences length. . If length of the sub-sequence cannot be expressed as power of 2
6. Corresponding index in the array is marked with 1 if a number in the sub-sequence is equal to 2^z-1, and remaining all indexes are marked with 0s.which is applied for every sub-sequence
7. Represent each sub-sequence as a column matrix. Multiply each sub-sequence matrix with the modified Hadamard matrix, such that the matrix of the form modulo 2^z-1 must be used to perform multiplication.
8. Perform the modulo 2^z-1operation on the resultant values obtained after multiplication.
9. Each decimal number in the sequence is converted into to corresponding binary values.
10. Now, use next element in the key and group the bits of the sequence obtained from the previous step based on the number.
11. Repeat 4 to 9. The sequence obtained after processing the last element in the key is the final encrypted message.

B. Decryption Algorithm

1. In this algorithm, first, consider the ciphered message, sequence of key which is reversed and the array used in encryption.
2. Then, consider the first element in the key and group the bits in the encrypted sequence based on the number.
3. Now convert each group to corresponding decimal number
4. Depending on the length of input sequence, divide the decimated sequence to equal length sub-sequences such that each sub-sequence length should be expressed as power of 2.
5. Represent each sub-sequence as a column matrix. Now, multiply each sub-sequence matrix with the modified Hadamard matrix, such that the matrix of the form modulo 2^z-1 must be used to perform multiplication.
6. Now, multiply two modulo 2^z-1 matrices and find the divisor such that the resultant matrix obtained is thus represented as an Identity matrix.
7. Calculate modulo multiplicative inverse for the divisor that is, a*y mod 2^z-1 =1[5] where y is the divisor and a is modulo multiplicative inverse.
8. Multiply the resultant matrix obtained in 7 with a.
9. Apply modulo 2^z-1 on the resultant values obtained after multiplication.
10. For every 0 in the decimal sequence check for the corresponding array index, if the index has an element 1 then replace 0 with 2^z-1.
11. Convert each decimal number in the sequence to corresponding binary values.
12. Eliminate all successive 0s at end of the sequence in such a manner; the resultant sequence has a length equal to power of 2.
13. Now, use next element of the key and group the bits of the sequence obtained from the previous step based on the number.
14. Repeat 4 to 12. For the sequence obtained after processing the last element in the key and obtain original message.

C. *Illustration*

1. Encryption

Let the input sequence (original message) is 110010001110111110 and Key is {3}.
To encrypt this input sequence, find the resultant matrix P =H_8 mod 7, where H_8 is 8 X 8 Hadamard matrix.
At this level every 3 bits in the input sequence are grouped and corresponding decimal sequence is given as follows {3,6,2,5,7,3,0,6}.
The maximum value corresponding to the input sequence is stored in the array
A[0][0]={0,0,0,0,1,0,0,0}
Multiply the resultant matrix P and decimal sequence and perform modulo 7 operation which gives the sequence as {4,6,6,3,0,3,5,4}.
After converting the above decimal sequence into binary, we get the encrypted sequence as

{100110110011000011101100}.

2. Decryption

The encrypted message is given as {100110110011000011101100}.
Now the decimal sequence is given as {4,6,6,3,0,3,5,4}.
Multiplying with the resultant matrix P and the decimal sequence, we find the following sequence {31,111,121,131,91,101,91,111}
Then, we find a multiplicative inverse using the formula a*y=1mod7, here a=1.
After multiplying in the above sequence by ‘a’ and taking modulo7 . We get the sequence {3,6,2,5,0,3,0,6}.
After comparing with stored array 5th element replaces with 7. We get the sequence {3,6,2,5,7,3,0,6}.
Now, after converting it into binary, we get {1100100011101111110} which is the original message.

X. PROPOSED MODEL

The proposed model is shown in the Fig. 2. [15]

1. In the given model, using above encryption algorithm in private cloud to encrypt the data.
2. The data is stored in the public cloud.
3. Only authorize user can access the original data after decryption of the stored data from the private cloud.

So in such a way storage as a service is provide to the user and data is stored into the public cloud in encrypted form so unathorize person cannot access the data.

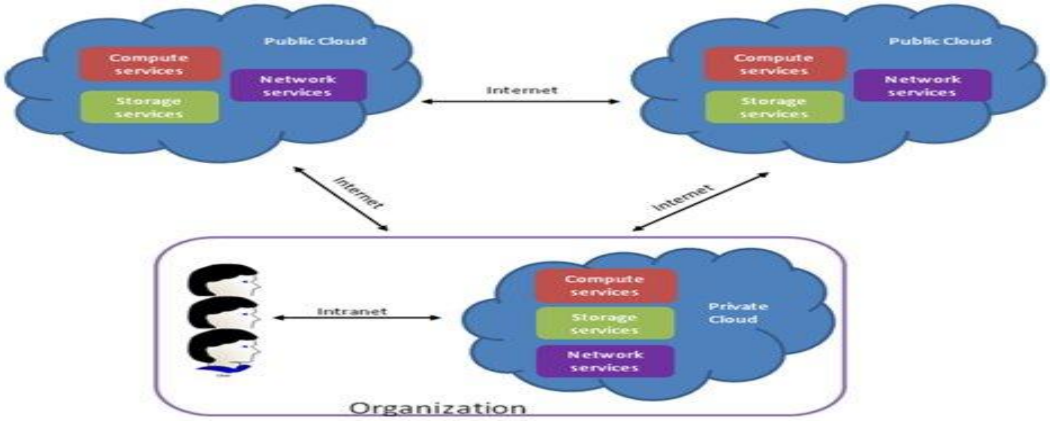

Fig.2: Proposed model [15]

XI. CONCLUSION AND FUTURE WORK

This paper presents an approach to encryption of data using a chain of Hadamard transforms. We have presented implementation of this system for binary and non-binary message sequence. An algorithm is given for both encryption and decryption. Clearly, this method is more effective when the key size is large and most of the elements in the key are distinct. But since the chaining of the Hadamard transforms can be as long as one pleases and since the choices as far as the formation of blocks that are converted into non-binary numbers is arbitrarily large, the effectiveness of this scheme is potentially very high. The proposed algorithm is very useful to increase the level of security as the types of data and we can also achieve cloud security objective. So it is very useful to provide storage as a service. But we feel that job scheduling, availability etc are serious challenges in cloud computing which we intend to do in our forthcoming endeavor.

REFERENCES

[1] L. J. Yan and J. S. Pan, Generalized discrete fractional Hadamard transformation and its application on the image encryption, International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IEEE Computer Press, pp. 457-460, 2007.

[2] C. C. Gumas, “A century old fast Hadamard transform proves useful in digital communications”, Chip Center quest link, 2006 .

[3] W. Ouyang, W.K. Cham, “Fast algorithm for Walsh-Hadamard transform on sliding windows”, IEEE Trans. on Pattern Analysis and Machine Intelligence 32, 165-171, 2010.

[4] K. J, Horadam, “A generalized Hadamard transform”, in A. Grant (ed.) Proceedings of the 2005 IEEE International Symposium on Information Theory, Adelaide, Australia, 4-9 September 2005, pp. 1006-1008.

[5] T. Koshy, “Elementary Number Theory with Applications”, 2nd Ed, Elsevier Inc., Academic Press Publications, Burlington, MA, 2007, pp. 346.

[6] S. Kak, “Classification of random binary sequences using Walsh-Fourier analysis”. IEEE Trans. On Electromagnetic Compatibility EMC-13, pp. 74-77, 1971.

[7] Rohith Singi Reddy “Encryption of Binary and Non-Binary Data Using Chained Hadamard Transforms”, OK 74078 .

[8] Katerina Tepla, “Hadamard transform” Charles University in Prague, 24th March 2012.

[9] Mukherjee & Sahoo, “Security Mechanism for C-Governance using Hadamard Matrices”, Computer and Communication Technology (ICCCT), 2011 2nd International Conference on 15-17 Sept. 2011.

[10] Mehmet Yildiz, Jemal Abawajy, Tuncay Ercan, Andrew Bernoth,"A Layered Security Approach for Cloud Computing Infrastructure," Parallel Architectures Algorithms, and Networks, International Symposium on, pp. 763-767, 2009 10th International Symposium on Pervasive Systems, Algorithms, and Networks, 2009.

[11] Fr Wassim Itani, Ayman Kayssi, Ali Chehab, "Privacy as a Service: Privacy-Aware Data Storage and Processing in Cloud Computing Architectures," Dependable, Autonomic and Secure Computing, IEEE International Symposium on, pp. 711-716, 2009.

[12] Trent Jaeger, Joshua Schiffman, "Outlook: Cloudy with a Chance of Security Challenges and Improvements," IEEE Security and Privacy, pp. 77-80, January/February, 2010.

[13] Xuan Zhang, Nattapong Wuwong, Hao Li, Xuejie Zhang, "Information Security Risk Management Framework for the Cloud Computing Environments”, Computer and Information Technology, International Conference on, pp. 1328-1334, 2010 10th IEEE International Conference on Computer and Information Technology, 2011.

[14] http://www.azureadvantage.co.uk/aboutazure/cloudcomputing/Pages/default.aspx

[15] http://stack.nil.si/ipcorner/CoreCloud/#chapter2.

AUTHOR BIOGRAPHY

Abhishek Mishra received the B-Tech (2010) degree in Electronics and Communication Engineering, Uttar Pradesh Technical University, Lucknow, India. He is currently a M-Tech student in the Birla Institute of Technology, Mesra, Ranchi, India. His research interests include cloud computing, information security.

Dileep Kumar Gupta received the B-Tech (2009) degree in Information Technology, Uttar Pradesh Technical University, Lucknow, India. He is currently a M-Tech student in the Birla Institute of Technology, Mesra, Ranchi, India. His research interests include cloud computing, software testing, information security.

Dr. G. Sahoo received the M.Sc degree in Mathematics, and the PhD degree in Mathematics from The Indian Institute of Technology, Kharagpur, India. He is a professor and head in the Birla Institute of Technology, Mesra, Ranchi, India. His current research interests include theoretical computer science, sequential & parallel computing, soft & evolutionary computing, pattern recognition & image processing, distributed & grid computing, cloud computing, Cryptography & Data Security.