IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-9, Sept 2019 E-ISSN: 2347-2693

Privacy Preserving Using AES-Mapping in Mix Column Encryption Algorithm: Cloud Approach

Anjali Kumari[1*], Varsha Namdeo[2]

[1,2]Dept. of Computer science and Engineering, RKDF Institute of Science and Technology (RKDFIST), Bhopal, India

**Corresponding Author: anjalikashyap533@gmail.com*

DOI: https://doi.org/10.26438/ijcse/v7i9.201206 | Available online at: www.ijcseonline.org

Accepted: 08/Sept/2019, Published: 30/Sept/2019

Abstract-In this paper we evaluate the efficient, scalable and practical method for privacy-preserving using k-nearest neighbors (KNN) classification method for EMR data. The approach enables performing the widely used k-NN classification method in complex scenarios where none of the parties reveal their information while they can still cooperatively find the nearest matches. To development AES- S.BOX mapping in mix column privacy preserving model used for preserving the privacy of the patients data in a cloud assisted system as the complex information is needed to be maintained confidential and should not be revealed to public users other than the physicians. As we know AES is based on several mathematical perform for security purpose substitutions, permutation and transformation.

Keywords: Cloud approach, privacy preserving, S.BOX mapping in mix column, KNN, MATLAB 2014a**.**

I.INTRODUCTION

The privacy preserving for the cloud assisted system is analyzed and the advantages of the protocol are determined. Privacy protection is an important aspect in the medical systems as there high risk of complex individual data being exposed to the public in an unauthorized way. The personal health information are collected from the patients with attributes such as heart beat rate, blood pressure, etc. during the medical treatment in terms of both text and images. Privacy preserving in cloud environments includes two aspects: data processing security and data storage security. Data processing security covers the issues of how to protect user privacy at runtime in a virtualized cloud platform. Data storage security covers the issues of guaranteeing user data privacy when the data is stored in data center.

The dynamic medical data mining and the image feature extraction are the only processes that require Privacy preserving data aggregation [3]. The privacy in data aggregation is achieved in this scheme by a tradeoff between the functionality and the optimized efficiency. Privacy concerns arise whenever complex data is outsourced to the cloud. By using encryption, the cloud server (i.e. its administrator) is prevented from learning content in the outsourced databases. But how can we also prevent a local administrator from learning the database content. And how can we avoid scenarios such as: employees using cloud applications may learn more than it is necessary to perform their respective duties? As an illustration, an organization may want to specify rules limiting request-per-day for call center employees to 100 client contacts. Such restrictions prevent download of the whole (customer) database. Our contribution in this paper is a system architecture that allows sufficient and flexible restriction writing. And in doing so, local administrators as well as cloud administrators are not able to change the access rules after an application is launched. The paradigm shift involves/results in the loss of control over data as well as new security and privacy issues [6]. For this reason caution is advised when deploying and using Cloud computing in enterprises. After all, the first big issue in data protection in Europe arose at the end of the 1960's, when a Swedish company decided to have its data processing done by a service bureau in Germany and the data protection legislations in both countries were not alike. With Cloud Computing quickly achievement approval, it is significant to focus the subsequent risks. As security and privacy issues are most important, they should be addressed before Cloud Computing establishes an important market share. In our work we proposed a new emerging concept namely KNN and AES- S.BOX mapping in mix column approach for data privacy preserving in cloud system [8, 9].

a. KNN in a privacy preserving

KNN-privacy preserving model for preserving the privacy of the patients in a cloud assisted system as the complex information is needed to be maintained confidential and should not be revealed to public users other than the physicians. Hence the privacy is modified by using k-nearest neighbor to develop KNN model in such a way that security of the medical data is improved. Instead of using a threshold value for the computed correlation function, the encrypted

© 2019, IJCSE All Rights Reserved

template (T) and the encrypted medical data (P) are processed to two non-colluding cloud service providers. The physician medical templates are encrypted and are outsourced to a cloud service provider while the secret key is stored in another cloud service provider.

II. RELATED WORK

Privacy-preserving cloud computing solutions have been developed from theoretical recommendations to concrete cryptographic proposals. There are many works which deal with general security issues in cloud computing but only few works deal also with user privacy. The authors [11] explore the cost of common cryptographic primitives (AES, MD5, SHA-1, RSA, DSA, and ECDSA) and their viability for cloud security purposes. The authors deal with the encryption of cloud storage but do not mention privacy preserving access to a cloud storage The work [12] establishes requirements for a secure and anonymous communication system that uses a cloud architecture (Tor and Free net). Nevertheless, the author does not outline any cryptographic solution. Only validate user can write on the cloud and invalid user doesn't get access to the cloud this work is presented in paper [13]. Another cryptographic solution ensuring user privacy in cloud scenarios is presented in [14]. The authors propose an algorithm which reduces the risk of the leakage of user private information. The authors of paper [15] presented a onetime password to authenticate the users. This paper presents a secure structure to the cloud. The issue of data security is one of the most important problems to be solved in paper [16].

III. SYSTEM MODEL

The *k*-nearest neighbor (*k*-NN) technique, due to its interpretable nature, is a simple and very intuitively appealing method to address classification problems. However, choosing an appropriate distance function for *k*-NN can be challenging and an inferior choice can make the classifier highly vulnerable to noise in the data. The great optimal of *k* depends upon the data; usually, greater values of *k* decrease the consequence of noise on the sorting, but create boundaries between classes less distinct. A good *k* can be selected by various heuristic techniques. In binary classification difficulties, it is cooperative to select k to be an odd no. as this escapes tied votes. The K-Nearest Neighbor algorithm is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its *k* nearest neighbors (*k* is a positive integer, typically small) [1]. Typically Euclidean distance is recycled as the distance metric; though this is only appropriate to constant variables. In cases such as text classification, another metric such as the overlap metric or Hamming distance, for example, can be used.

KNN is a modest procedure that supplies all obtainable cases and categorizes new cases based on a correspondence measure (e.g., distance functions). KNN has been used in statistical estimation and pattern recognition already in the beginning of 1970's as a non-parametric technique.

K nearest neighbor algorithm is very simple. It mechanism based on smallest distance from the query example to the training examples to control the K-nearest neighbors. The data for KNN algorithm consist of several attribute names that will be used to classify. The data of KNN can be any measurement scale from nominal, to quantitative scale [1, 2].

The KNN algorithm is shown in the following form:

Input: D, the set of k training objects, and test object z= (x', y').

Process: Compute d(x', x), the distance between z and every object, (x, y) ∈ D. Select $D_z \subseteq D$, the set of k closet training objects to z.

Output: $y' = \text{argmax}_v \sum_{(xi, yi) \subseteq Dz} I(v = y_i)$

- v is a class label
- y_i is the class label for the i^{th} nearest neighbors
- I (.) is an indicator function that returns the value 1 if its argument is true and 0 otherwise.

In this system, KNN algorithm is used the suitable result by mixing the Euclidean distance among the various kinds of distance metric. The Euclidean distance is as shown in below:

$$d_{ij} = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \cdots + (x_{ip} - x_{jp})^2}$$

Where

d_{ij} = the distance between the training objects and test object

x_i = input data for test object

$x_{j=}$ data for training objects stored in the database

IV. PROPOSED METHOD

AES-128

We confine to depiction of a characteristic round of advance encryption algorithm. Every round include of four sub-processes [7]. The 1st round procedure is represented below:-

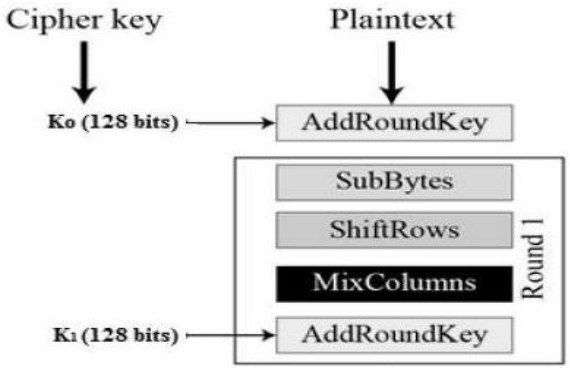

Fig. 1: Operation process of AES-128

© 2019, IJCSE All Rights Reserved

Byte Substitution (Sub Bytes)

The 16 input bytes are replaced in observing up a secure table (S-box) assumed in strategy. The consequence is in a matrix of 4 rows and 4 columns.

Shift rows

All of the 4 rows of the matrix are removed to the left-hand. Some accesses that 'fall off' are re-inserted on the correct crosswise of row. Shift is approved out as surveys −

- 1st row is not removed.
- 2nd row is removed 1 (byte) location to the left.
- 3rd row is removed 2 (byte) locations to the left.
- 4th row is removed 3 (byte) positions to the left.
- The consequence is a novel matrix containing of the similar 16 bytes but removed w.r.t each other.

Mix Columns

Every column of 4 bytes is currently altered using a singular exact purpose. This purpose taking as input the four bytes of one column and productions four entirely new bytes, which change the unique column. The consequence is alternative novel matrix containing of 16 novel bytes. It must be well-known that this stage is not achieved in the previous round [8].

Add round key

The 16 bytes of the matrix are presently dignified as 128 bits and are XORed to the 128 bits of the round key [9]. In case this is the most recent round formerly the productivity is the cipher text. Then, the subsequent 128 bits are construed as 16 bytes and we instigate additional comparable round.

AES with S.BOX MODIFICATION MAPPING IN MIX COLUMN

AES key that will be used is 128-bit with 16 length of character. Character of key and plaintext will be transformed into hexadecimal that shown on table 1. Key characters will call state key when this key will be proceeding on first round and data characters would called state data.

Table 1: AES Key and Sample Data

Key	b	i	n	a	n	u	s	a	n	t	a	r	a	1	2	3
Hex	62	69	6E	61	6E	75	73	61	6E	74	61	72	61	31	32	33
Plain Text	T	w	o		O	n	e		N	i	n	e		T	w	o
Hex	54	77	6F	20	4F	6E	65	20	4E	69	6E	65	20	54	77	6F

The hexadecimal public key will be transformed into 4 x 4 matrix dimension in Equation 1 as follows:

$$\begin{bmatrix} K_{0,0} & K_{0,1} & K_{0,2} & K_{0,3} \\ K_{1,0} & K_{1,1} & K_{1,2} & K_{1,3} \\ K_{2,0} & K_{2,1} & K_{2,2} & K_{2,3} \\ K_{3,0} & K_{3,1} & K_{3,2} & K_{3,3} \end{bmatrix} = \begin{bmatrix} 62 & 6E & 6E & 61 \\ 69 & 75 & 74 & 31 \\ 6E & 73 & 61 & 32 \\ 61 & 61 & 72 & 33 \end{bmatrix}$$

Key will be processed with four steps transformation stage12 that mention above:

1. ***Sub Bytes*** -Transformation process for a non-linear byte substitution using S.box lookup table

2. *Shift Rows* - Cyclical shifting process for key matrix in each row

3. ***Mix Column*** - Dot matrix operation combine with XOR using matrix finite field $GF(2^8)$ and Galois Field.

4. ***Add Round Key*** - XOR addition operation for round key with state data

Sub Bytes are steps of byte to byte substitution using Rijndael's S.Box lookup table13. The table has 16x16 dimension where its contain hexadecimal characters. The hexadecimal is a replacement byte for a given input key state byte. Sub-Bytes process can be seen on Fig. 1. For example, matrix element $K_{0,1}$ has 6*E* value. The 6E value used as a row coordinate and the *E* value used as column coordinate on S.Box lookup table. The lookup result for the value is 9*F*. It is important that key and plaintext characters have to transform to hexadecimal for fit AES encryption and decryption process. A Shift Row is arranging elements of state key matrix which performs a circular shift each row. The circular shift length is different every each row. The first row is never moved over. Second row move one first element to the right at last element. Third row move two first elements to the right at last element and the last row move three first elements. Shift row process can be seen on Fig 2. For example, element *K*1,0 to be switched to the last element and three others element will come forward to the left, so the composition in the new state of matrix at second row will be {*K*1,1, *K*1,2, *K*1,3, *K*1,0}.

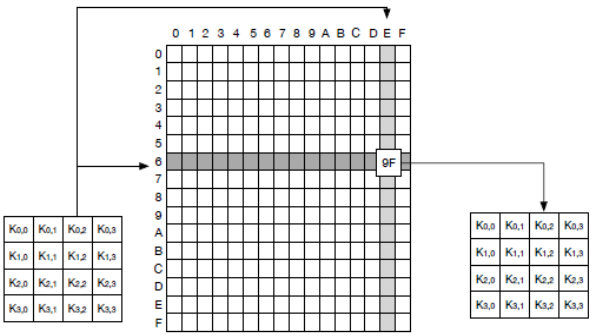

Fig.2: Sub Byte Process

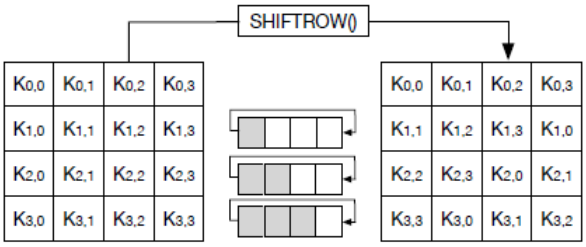

Fig.3: Shift Row Process

© 2019, IJCSE All Rights Reserved

MixColumnisalineartransformationprocess.Eachelementofstatecharactersmultipliedagainstelementsofmultiplicationmatrixthatcamefromtwofour-termpolynomialwhichhasthecoefficientelementof$GF(2^8)$.FirstpartofMixColumnismatrixMultiplicationbetweenfinitefield matrixandkeystatematrix.For
Example, the Equation 2.as follows:

$$\begin{bmatrix} 02\ 03\ 01\ 01 \\ 01\ 02\ 03\ 01 \\ 01\ 01\ 02\ 03 \\ 03\ 01\ 01\ 02 \end{bmatrix} x \begin{bmatrix} K_{0,0}\ K_{0,1}\ K_{0,2}\ K_{0,3} \\ K_{1,1}\ K_{1,2}\ K_{1,3}\ K_{1,0} \\ K_{2,2}\ K_{2,3}\ K_{2,0}\ K_{2,1} \\ K_{3,3}\ K_{3,0}\ K_{3,1}\ K_{3,2} \end{bmatrix} = \begin{bmatrix} 2*K_{0,0}+3*K_{0,1}+1*K_{0,2}+1*K_{0,3} \cdots \cdots \cdots \\ 1*K_{0,0}+2*K_{0,1}+3*K_{0,2}+1*K_{0,3} \cdots \cdots \cdots \\ 1*K_{0,0}+1*K_{0,1}+2*K_{0,2}+3*K_{0,3} \cdots \cdots \cdots \\ 3*K_{0,0}+1*K_{0,1}+1*K_{0,2}+2*K_{0,3} \cdots \cdots \cdots \end{bmatrix} \quad (2)$$

The strong AES algorithm lies within in the last transformation, Add Round Key. The key must be expanded every round. Every time Adds Round Key process is triggered, every key element *Kr*,*c* XOR against the data.

Forexample,$K\_3,2=K3,2 \oplus D3,2$.Table.2isaaddroundkeyprocessforeachelements.

Table2: Add Round Key Transformation

$K_{0,0}$	$K_{1,0}$	$K_{2,0}$	$K_{3,0}$	$K_{0,1}$	$K_{1,1}$	$K_{2,1}$	$K_{3,1}$	$K_{0,2}$	$K_{1,2}$	$K_{2,2}$	$K_{3,2}$	$K_{0,3}$	$K_{1,3}$	$K_{2,3}$	$K_{3,3}$
$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$	$\oplus$
$D_{0,0}$	$D_{1,0}$	$D_{2,0}$	$D_{3,0}$	$D_{0,1}$	$D_{1,1}$	$D_{2,1}$	$D_{3,1}$	$D_{0,2}$	$D_{1,2}$	$D_{2,2}$	$D_{3,2}$	$D_{0,3}$	$D_{1,3}$	$D_{2,3}$	$D_{3,3}$

Table: 1 explain that every time the data bytes increase by 1024 bytes, then time computation increase by three milliseconds. Our experiment shows that AES algorithm will be slower as bytes of data growth. Therefore, several improvements can be applied to increase performance algorithm. Our methods improve shift row transformation by using array shift mapping regardless move and rotate each element. Our improvement also in Mix Column transformation process by modifies S.Box and using it in its transformation and also make Sub Byte process unnecessary.

Table 3: Performance of AES Algorithm Implementation

Bytes of Data	Average Times (ms)
1024	3.045
2048	6.570
3072	9.806
4096	13.851
5120	17.284

V. RESULT DISCUSSION

This work presents the hybrid cryptography of the S.BOX mapping in mix column Encryption and AES-128 Set of rules. In this investigation we reviewed the best collective approaches in the cryptography of a slab cipher system.
The resultant of Public-Key Processes is symmetric, that is to approximately use to encode the text or given text by user is different from the key used to decrypt the message. The encryption key, identified as the Public key which used to encode a communication, but the message can only be deciphered through the information that has the decryption key, recognized as the private key.

This type of encryption has a quantity of advantages over usual symmetric Ciphers.
It means that the recipient can create their public key approximately available- someone deficient to send them a communication usages the procedure and the receiver's public key to do so. A viewer may have both the procedure and the public key, but will still not be capable to decode the text. Individual the receiver, with the private key can decrypt the message.

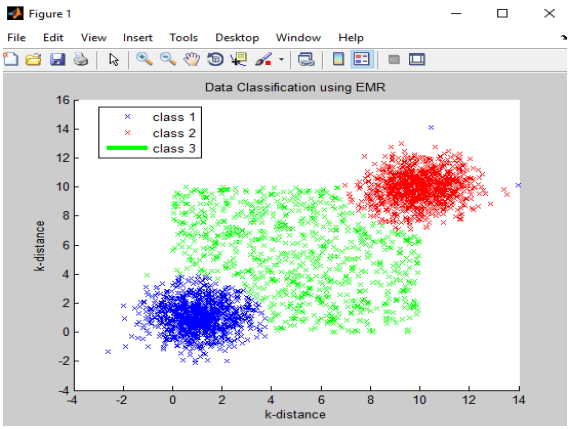

Figure 4: Data classification using KNN for EMR data

Figure 4 shows the classification of EMR data in three part class 1, class 2 and class 3. It also measured the k-distance between one clusters to other cluster.

A circumstance is confidential through a majority vote of its neighbors, with the case being allocated to the class utmost mutual between its K nearest neighbors measured by a distance function. If K = 1, then the case is simply assigned to the class of its nearest neighbor. Distance calculated by below function.

Euclidean $\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$

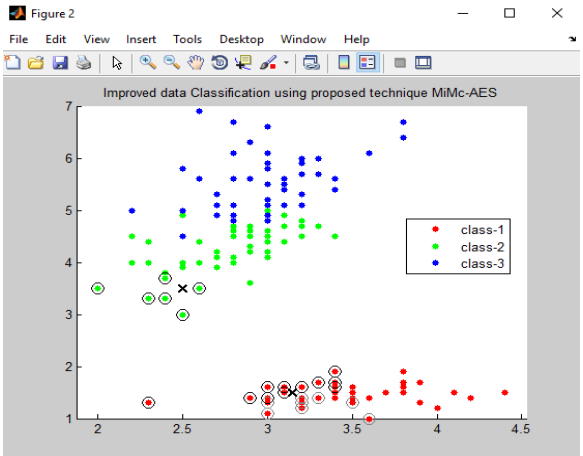

Figure 5: Data classification using KNN for EMR data

© 2019, IJCSE All Rights Reserved

As above figure we can see that data classification is improved as compare to figure 5.

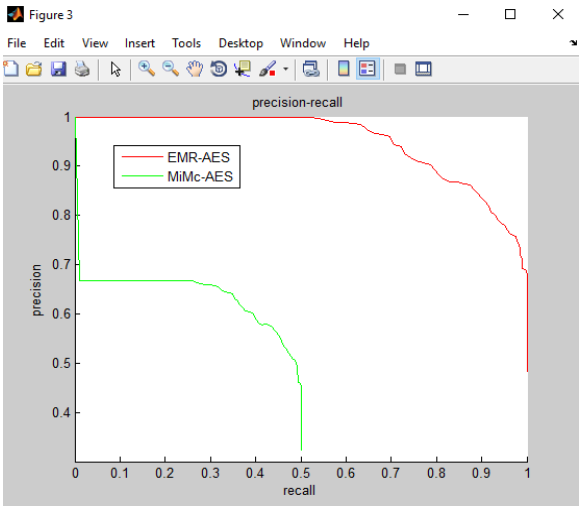

Figure 6: precision and recall using EMR-AES and Proposed algorithm

Precision and recall are the basic measures used in evaluating search strategies.

RECALL approach is the fraction of the no. of applicable records recovered to the total no. of appropriate records in the database. It is typically communicated as a percentage.

PRECISION is the ratio of the number of relevant records retrieved to the total number of irrelevant and relevant records retrieved. It is usually expressed as a percentage.

In the diagram beyond, the two outlines might signify the performance of dissimilar search schemes. While the exact slope of the curve may vary between systems, the general inverse relationship between recall and precision remains.

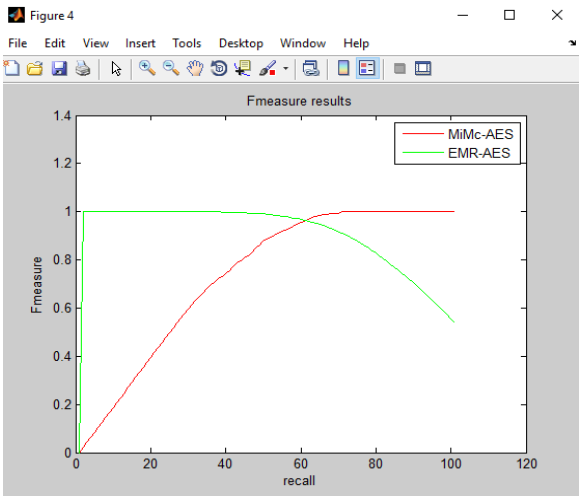

Figure 7: F-measure using EMR-AES and MiMc-AES

The F (Frequency)-measure can be viewed as a compromise between recall and precision. It is high only when both recall and precision are high. It is equivalent to recall when $\alpha = 0$ and precision when $\alpha = 1$. The F-measure assumes values in the interval [0, 1]. It is 0 when no relevant data have been retrieved, and it is 1 if all retrieved data are relevant and all relevant data have been retrieved.

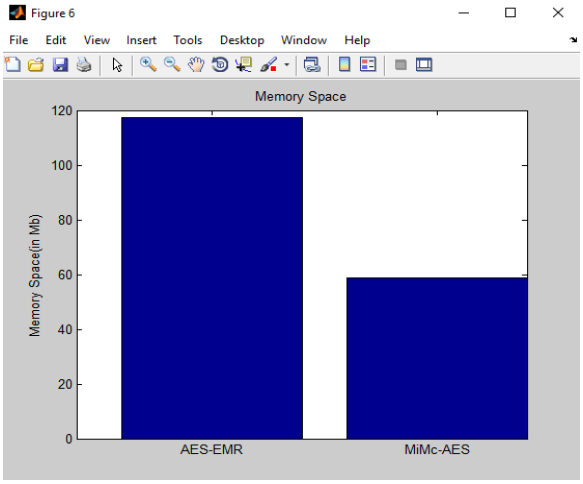

Figure 8: Memory space for AES-EMR and AES-S.BOX mapping in mix column

Above figure shows the AES-EMR has taken more memory space as compare to proposed method i.e. ***AES-S.BOX mapping in mix column***. In x=1, y=117 for EMR and x=2, y=58.8 for our proposed algorithm as we proposed by our experimental work with the help of our proposed method.

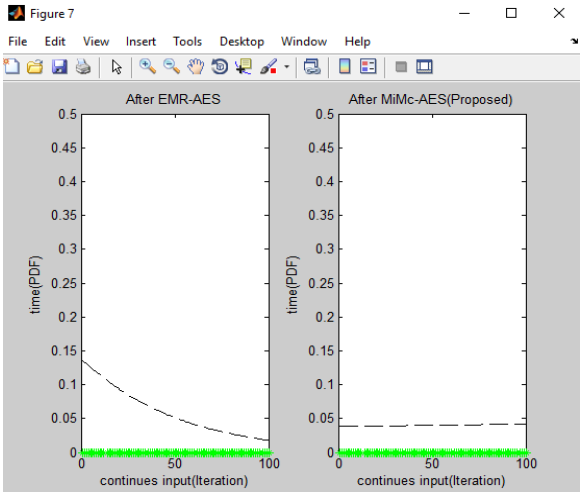

Figure 9: Comparison between AES-EMR and *AES-S.BOX mapping in mix column*

The PDF is used to specify the probability of the random variable falling *within a particular range of values*, as opposed to taking on any one value. This probability is given by the integral of this variable's PDF over that range—that is, it is given by the area under the density function but

above the horizontal axis and between the lowest and greatest values of the range. The probability density function is nonnegative everywhere, and its integral over the entire space is equal to one.

This paper studied various symmetric key encryption algorithms which were proposed earlier and identified the best method for secure storing of files. AES 256 bit encryption is found to be best for fast and secure transfer and storage of files. AES 256 bit takes less time as compared to its variations. The test is performed on various format of different file sizes and found that AES 256 bit encryption and decryption is fast and secure. On the basis of this study a model for secure transfer, storing and sharing of files in the cloud is implemented which results in secure transfer of files

VI.CONCLUSION

Our proposed approach defined an emerging scheme in which two methods, AES (advanced encryption encryption) and block based tiny which offers a robust support for its safety and the method to protected data or message with verification and signature verification in our hybrid method which goes to modify the innovation of the records files into encrypted form using Tiny-AES-128 encryption procedure that variations it into an illegible cipher text and plaintext is cryptography using the processes from mixed (orthogonal) arithmetical collections and an enormous amount of circles to attain security with easiness. At two, sixty-four (64) Feistel rounds, an entire number of rounds are used in the AES-128 and S.BOX mapping in mix column encryption process with smallest time next encoded, the encoded files is embedding in a random text by using the idea of cryptography and formerly this text file directed via user and Processing time for block cipher, response time for senders, minimizing space consumption of s-box simulate in MATLAB.

REFERANCES

[1] A. Andoni and P. Indyk. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. In FOCS, 2006.
[2] M. Bellare, V. Tung Hoang, Sriram K., and P. Rogaway. Efficient garbling from a fixed-key block cipher. In S&P. IEEE, 2013.
[3] J. Boyar and R. Peralta "Concrete multiplicative complexity of symmetric functions in MFCS" Springer, 2006
[4] Brenner, perl, and Smith. hcrypt Secure Function Evaluation (SFE) project. https://hcrypt.com/sfe/.
[5] H. Carter, C. Lever, and P. Traynor. Whitewash: Outsourcing garbled circuit generation for mobile devices. In ACSAC. ACM, 2014.
[6] H. Carter, B. Mood, P. Traynor, and K. Butler. Secure outsourced garbled circuit evaluation for mobile phones. In USENIX Security. USENIX, 2013.
[7] D. Demmler, T. Schneider, and M. Zohner. Ad-hoc secure two-party computation on mobile devices using hardware tokens. In USENIX Security. USENIX, 2014.
[8] C. Gentry. "Fully homomorphic encryption using ideal lattices". In STOC, 2009.
[9] A. Bessani, M. Correia, B. Quaresma, F. Andr´e, and P. Sousa, "Depsky: Dependable and secure storage in a cloud-of-clouds," in *Proceedings of the Sixth Conference on Computer Systems*, 2011, pp. 31–46.
[10] E. Stefanov, M. van Dijk, A. Juels, and A. Oprea, "Iris: A scalable cloud file system with efficient integrity checks," in Proceedings of the 28th Annual Computer Security Applications Conference, 2012, pp. 229–238.
[11] Y. Chen and R. Sion, "On securing untrusted clouds with cryptography," in Proceedings of the 9th annual ACM workshop on Privacy in the electronic society. ACM, 2010, pp. 109– 114.
[12] R. Laurikainen, "Secure and anonymous communication in the cloud," Aalto University School of Science and Technology, Department of Computer Science and Engineering, Tech. Rep. TKK-CSE-B10, 2010.
[13] Sushmita Ruj, Milos Stojmenovic and Amiya Nayak, "Privacy Preserving Access Control with Authentication for Securing Data in Clouds"in proceedings of the 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing
[14] Debajyoti Mukhopadhyay, Gitesh Sonawane, Parth Sarthi Gupta, Sagar Bhavsar, Vibha Mittal." Enhanced Security for Cloud Storage using File Encryption" Department of Information Technology Maharashtra Institute of Technology
[15] Kawser Wazed Nafi1,2, Tonny Shekha Kar2, Sayed Anisul Hoque3, Dr. M. M. A Hashem4, " A Newer User Authentication, File encryption and Distributed Server Based Cloud Computing security architecture"in proceeding of the (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 3, No. 10, 2011
[16] Du meng." Data security in cloud computing" in yhe proceeding of the The 8th International Conference on Computer Science & Education (ICCSE 2013) April 26-28, 2013. Colombo, Sri Lanka

© 2019, IJCSE All Rights Reserved