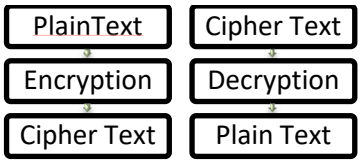
Substitution Based Encryption Model for Cloud Data Storage

Akhilesh Deep Arya[1], Gaurav Kumar Ameta[2]
[1, 2]*Department of Computer Science and Engineering, Pacific Institute of Technology, Udaipur, Rajasthan, India*

***Abstract*: - Cryptography is the method of converting plain readable text into non-readable, and thus achieves security by it. Technique of messages conversion from a non-readable format back to readable format without knowing how they were converted from readable format to non-readable format is known as cryptanalysis.**

In technical terms, the transformation of plain text into cipher text message is called as encryption, while the reverse process is known as decryption. Encryption and decryption process have mainly two elements: the key used for encryption and decryption and algorithm.



According to the key used, cryptography is divided into two parts:

1. **Symmetric key cryptography: Symmetric key technique uses similar key for both encryption and decryption. The key is secretly exchanged between the sender and the receiver and message transformation took place. Only the authenticated receiver is allowed to decrypt the cipher text with the help of the secret key. Since only one key is used for both encryption and decryption, this technique is also known as private key cryptography.**
2. **Asymmetric key cryptography: Asymmetric key cryptography involves the usage of one key for encryption and another, different key for decryption. This is also known as Public key Cryptography, as two different keys (which form a key pair) are used. One key is used for encryption and other for decryption. No other key can decrypt the message- not even the original key which is used for encryption. Public key cryptography can be used for digital signing as it supports authentication of users.**

In this paper an attempt has been made to generate an algorithm which provides security to data transmitted over internet. The algorithm considers a random matrix key which on execution by a series of steps generates a sequence. This sequence is used as a sub key to build encryption model.

***Keywords:* Cryptography, Encryption Model, Substitution based model.**

I. LITERATURE REVIEW

Ayesh et al. [7] developed a framework using multi-agent systems for Internet security. The system architecture of this approach is composed of three different agent types classified on their functionalities. The first type is responsible for intrusion detection; the second type is responsible for encryption and decryption of messages, while the third type can act as the combination of the previous two types. Although this approach has provided useful security system, it does not address some other important issues such as authentication, authorization, digital signature, and verification security services.

Balakrishnan et al. [7] worked on the problem of ensuring the integrity and security of data storage in Cloud Computing. Security in cloud is achieved by signing the data block before sending to the cloud. They used BLS algorithm for signing which is more secure compared to other algorithms. To ensure the correctness of data, they introduced the effective third party auditor (TPA), on behalf of the cloud client, to verify the integrity of the data stored in the cloud as well as to audit user's outsource data when needed [9]. They utilized public key based homomorphic authenticator with random masking to achieve privacy preserving public auditing protocol [10]. They used the technique of bilinear aggregate signature to achieve batch auditing, i.e. multiple auditing tasks simultaneously [8]. Batch auditing reduces the computation overhead. As the data in the cloud is used by many industries, modification of data cannot be avoided. The new scheme further supports secure and efficient dynamic operations on data blocks, including: data update, delete and append. They explored the efficient technique for error correction called Reed Solomon technique which ensures the correctness of data.

Sherif et al. [12] developed an evaluation for selected eight modern encryption techniques namely RC4, RC6, MARS, AES, DES, 3DES, Two-fish, and Blow-Fish. This evaluation was performed for those encryption algorithms according to randomness testing and using the NIST statistical testing in both Cloud Computing and traditional desktop environments. [13] The performance of evaluation was tested by measure encryption speed for those encryption algorithms in both Cloud Computing and traditional desktop environments. The selected eight modern encryption techniques use a random number generator to get some critical data similar to keys and initial vectors [14] [15]. They focused on evaluation of eight modern encryption algorithms namely. The evaluation was

implemented as Pseudo Random Number Generator (PRNG). The evaluation was used to determine the most suitable technique. In addition the evaluation analysis the performance selected modern encryption techniques. Cryptography algorithms were implemented using Java Cryptography Extensions (JCE). Simulation results were shown to demonstrate the effectiveness of each algorithm.

II. PRESENT WORK

An algorithm is developed. The algorithm is based on substitution method which uses a ternary vector and a random matrix as key, by multiplying both and applying sign function on to it we get a sequence. This sequence will be used to generate a model of substitution technique. This algorithm uses a single key that is to be shared with both sender and receiver in some secure fashion, algorithm is considered to be substitution algorithm. The new encryption algorithm is based on the concept of Polyalphabetic cipher [4] which is an improvement over the mono alphabet [5].

III. DESIGN OF ALGORITHM

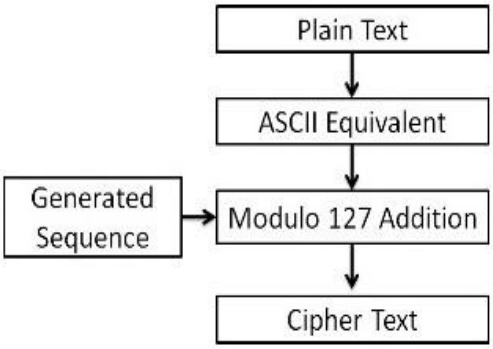
3.1 Adoptability of some mathematical functions in Cryptography

Sign Function [1]: This function is when applied on values of matrix, converts all the positive values to 1, negative values to -1 and zero with 0. The advantage of using this function in cryptography is it cannot be a reversible process, i.e., it is impossible to get back the original matrix by applying the process in reverse order.

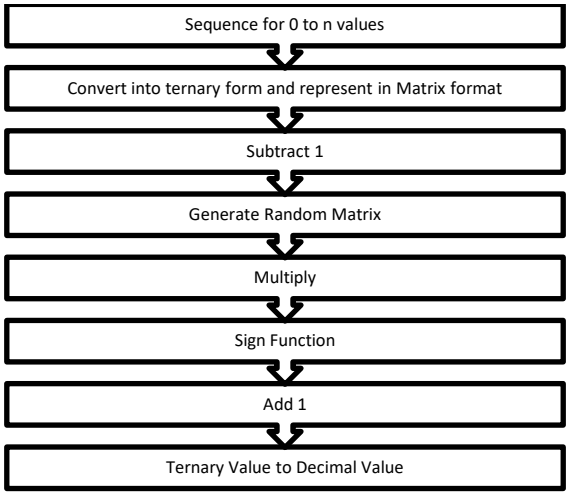
Modular Arithmetic [2]: One more function that is widely used in cryptography is modular arithmetic of a number with a base value. It will generate the remainder of a number with respect to the base value. This function is mostly used in public key cryptography.

3.2 Design Process

1. Encryption



Algorithm for generating the sequence



Example:

Step1: Consider the sequence for n= 0 to 26 values.

Step2: Convert the sequence to ternary form of a 3 digit number.

i.e.

0 ------ 000
1 ------ 001
2 ------ 002
. .
. .
. .
26 ------ 222

Step3: Represent above ternary form in 27x3matrix.

R=

0	0	0
0	0	1
0	0	2
0	1	0
0	1	1
0	1	2
0	2	0
0	2	1
0	2	2
1	0	0
1	0	1
1	0	2
1	1	0
1	1	1
1	1	2
1	2	0
1	2	1
1	2	2
2	0	0
2	0	1
2	0	2
2	1	0
2	1	1
2	1	2
2	2	0
2	2	1
2	2	2

Step 4: Subtract 1 from each element of the above matrix and the resultant matrix R is

R=

-1	-1	-1
-1	-1	0
-1	-1	1
-1	0	-1
-1	0	0
-1	0	1
-1	1	-1
-1	1	0
-1	1	1
0	-1	-1
0	-1	0
0	-1	1
0	0	-1
0	0	0
0	0	1
0	1	-1
0	1	0
0	1	1
1	-1	-1
1	-1	0
1	-1	1
1	0	-1
1	0	0
1	0	1
1	1	-1
1	1	0
1	1	1

Step 5: Consider a random matrix

$$A=\begin{bmatrix}-2 & 3 & 1\\ 3 & -3 & 3\\ 4 & -2 & -3\end{bmatrix}$$

Step 6: R = R x A

R=

-5	-10	-32
-1	0	-1
3	10	30
-2	-4	-11
2	6	20
6	16	51
1	2	10
5	12	41
9	22	72
-7	-16	52
-3	-6	-21
1	4	10
-4	-10	-31
0	0	0
4	10	31
-1	-4	-10
3	6	21
7	16	52
-9	-22	-72
-5	-12	-41
-1	-2	-10
-6	-16	-51
-2	-6	-20
2	4	11
-3	-10	-30
1	0	1
5	10	32

Step 7: Convert all positive values to 1, negative values to -1 and zero to 0 of the resulting matrix in step 6.

-1	-1	-1
-1	0	-1
1	1	1
-1	-1	-1
1	1	1
1	1	1
1	1	1
1	1	1
1	1	1
-1	-1	1
-1	-1	-1
1	1	1
-1	-1	-1
0	0	0
1	1	1
-1	-1	-1
1	1	1
1	1	1
-1	-1	-1
-1	-1	-1
-1	-1	-1
-1	-1	-1
-1	-1	-1
1	1	1
-1	-1	-1
1	0	1
1	1	1

Step 8: Add 1 to each element of matrix R.

R=

0	0	0
0	1	0
2	2	2
0	0	0
2	2	2
2	2	2
2	2	2
2	2	2
2	2	2
0	0	2
0	0	0
2	2	2
0	0	0
1	1	1
2	2	2
0	0	0
2	2	2
2	2	2
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
2	2	2
0	0	0
2	1	2
2	2	2

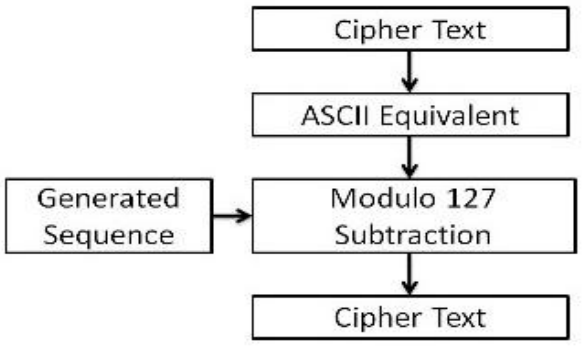
Step 9: Convert each row of the matrix R to decimal form to generate sequence, i.e., 0 0 2 will form
$0 * 3^2 + 0 * 3^1 + 2 * 3^0 = 2$

The sequence formed is =

0	3	26	0	26	26	26	26	26	0
0	26	0	13	26	0	26	26	0	0
0	0	0	26	0	23	26			

Plain Text	2	4	3	1	6	4	6
ASCII equivalent	50	52	51	49	54	52	54
Key	0	3	26	0	26	26	26
Add	50	55	77	49	80	78	80
Mod 127	50	55	77	49	80	78	80
Cipher text	2	7	M	1	P	N	P

2. Decryption



Cipher Text	2	7	M	1	P	N	P
ASCII equivalent	50	55	77	49	80	78	80
Key	0	3	26	0	26	26	26
Subtract	50	52	51	49	54	52	54
Add 127 if negative	50	52	51	49	54	52	54
Plain text	2	4	3	1	6	4	6

IV. ANALYSIS OF THE PROPOSED MODELS

Avalanche effect [11], is defined as the process through which slight change in the input plain text or in the key, make effective changes in the cipher text. Cipher is something which changes the actual data into some other form which is understood only by the sender and receiver. Higher the Avalanche effect, better is the algorithm. Both the models are compared for their performance in terms of avalanche effect along with some classical and modern encryption techniques.

Avalanche Effect can be calculated using the formula

$$\text{Avalanche Effect} = \frac{\text{No. of flipped bits in the ciphered text}}{\text{No. of bits in the ciphered text}} \text{X } 100$$

We use plain text as “ENCYCLOPEDIA” . The key used is “DISASTER”. Changing one bit from input, we have “DISCSTER” (changing A (01000001) to C (01000011)).

KEY 1: DISASTER

01000100 01001001 01010011 01000001 01010011

01010100 01000101 01010010

KEY 2: DISCSTER

01000100 01001001 01010011 01000011 01010011

01010100 01000101 00110010

PLAIN TEXT: ENCYCLOPEDIA

01000101 01001110 01000011 01011001 01000011

01001100 01001111 01010000 01000101 01000100

01001001 01000001

CIPHER TEXT 1: YbCpWLidY^IA

01011001 01100010 01000011 01110000 01010111

01001100 01101001 01100100 01011001 01011110

01001001 01000001

CIPHER TEXT 2: MUI`IOUVWLQT

01001101 01010101 01001001 01100000 01001001

01001111 01010101 01010110 01010111 01001100

01010001 01010100

Difference is **33** bits that is Avalanche Effect is **34.38%**

V. CONCLUSION

There is always a need of security when your data is stored on cloud or transferring from cables from one place to another. Researchers always work on new algorithms to come up with a solution to this.

This algorithm is one of such effort to implement security to the data stored in cloud. In this work a ternary system with a 3 digit number is used. So the sub key generated is a 33, i.e., a 27 digit number. By considering a ternary vector with a four digit number or five digit number, the length of the sub key can be increased by 34, 35. Similarly by considering n –ary vector the length of the subkey generated can still be

increased. Thus by increasing the length of subkey, security of cipher system can be increased still further.

REFERENCES

[1]. PanditS.N.N, "A New Matrix Calculus", Journal of Society for Industrial and Applied Mathematics, Volume 9, Issue 4, pp. 632-639, Dec., 1961.

[2]. Suter, B.W.; Honeywell, Inc. "The Modular Arithmetic of Arbitrarily Long Sequences of Digits", IEEE Transactions on Computer, Volume: C-23, Issue: 12, pp.1301-1303, Dec. 1974.

[3]. Akanksha Shukla: "Algorithm for generating sub-keys/basins from a New Substitution Block Cipher Algorithm" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 03 Issue: 11 | Nov -2016

[4]. Omran, S.S.; Al-Khalid, A.S.; Al-Saady, D.M. "A cryptanalytic attack on Vigenere cipher using genetic algorithm", IEEE Conference on Open Systems (ICOS), pp. 59-64, Sep. 2011

[5]. Omran, S.S.; Al-Khalid, A.S.; Al-Saady, D.M. "Using Genetic Algorithm to Break a Mono-Alphabetic Substitution Cipher", IEEE Conference on Open Systems (ICOS), pp. 63-67, Dec. 2010

[6]. Ayesh A., Bechkoum K., "Framework of Multi-agents internet security system". Appl Inform, Nov. 1999

[7]. Balakrishnan.S, Saranya.G, Shobana.S&Karthikeyan.S, "Introducing Effective Third Party Auditing (TPA) for Data Storage Security in Cloud", International Journal of Computer Science and Technology IJCSTVol.2, Issue 2, 2011.

[8]. Craig Gentry, Dan Boneh, "Aggregate and verifiably encrypted signatures form bilinear maps", 2004.

[9]. Q. Wang, C. Wang, J. Li, K. Ren, W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing", in Proc. of ESORICS'09, Saint Malo, 2009.

[10]. M. A. Shah, R. Swaminathan, M. Baker, "Privacy preserving audit and extraction of digital contents", Cryptology ePrint Archive, 2008.

[11]. SriramRamanujam, MarimuthuKaruppiah, "Designing an Algorithm with high Avalanche Effect", IJCSNS International Journal of Computer Science and Network Security, Volume: 11 No.1, pp. 106-111, January 2011

[12]. Sherif el-etriby, EmanM.Mohamed and Hatem S. Abdelkader published "Modern Encryption Techniques for Cloud Computing Randomness and Performance Testing" Third International Conference on Communications and Information Technology ICCIT 2012

[13]. Andrew Rukhin, Juan Soto, James Nechvatal, Miles Smid, "A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications", 2010.

[14]. Carolynn Burwick C, Don Coppersmith, "The MARS Encryption Algorithm"

[15]. Daemen, J., and Rijmen, V. "Rijndael:The Advanced Encryption Standard", Dr. Dobb's Journal, Vol.26, No.3, pp.137-139, March 2001.