VOL. 9, NO. 11, NOVEMBER 2014 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences


www.arpnjournals.com

SECURE STORAGE AND TRANSMISSION OF IMAGES BASED ON A DUAL ENCRYPTION SCHEME

Grasha Jacob[1] and A. Murugan[2]
[1]Research and Development Centre, Bharathiar University, Coimbatore, India
[2]Department of Computer Science, Dr. Ambedkar Govt Arts College, Chennai, India
E-Mail: grasharanjit@gmail.com

ABSTRACT

Today there is an urgent need to provide and protect the confidentiality of images when stored in a cloud or transmitted over public insecure channel.. Various encryption schemes have been developed to make information intelligible only to the intended user. This paper proposes a dual encryption scheme which is a combination of Key dependent S-Box and DNA sequence based encryption imparting double fold security for the storage and transmission of images.

Keywords: images, confidentiality, S-Box, Key dependent S-Box, hamming weight, hamming distance, DNA sequence file.

INTRODUCTION

With the growth of computer networks and the latest advances in digital technologies, a huge amount of digital data is being exchanged over various types of networks. Telemedicine has become a common method for the transmission of images and patient data across long distances. The military relies heavily on secure image transmission to provide intelligence on enemy movements and for the safety of its soldiers on the ground. In business transactions, sensible data such as pin numbers are transmitted as images. All these information are either confidential or private and ensuring security has become increasingly challenging as many communication channels are intruded by attackers.

From time immemorial, the art of communicating secretly has been imperative. German's Enigma machine and American Indian Language Navajo used during World War II are concrete examples to explain everything about cryptographic techniques. A cryptosystem is a system that allows two parties to communicate secretly by transforming an intelligible message into unintelligible form and then retransforming that back to its original form. The main objective of cryptography is confidentiality. Even though cryptographic systems ensure security to sensitive information, attackers and intruders have come up with various methods to crack and crash the cryptographic systems that are developed. With the advancement of technology in both hardware and software, even the Data Encryption Standard published by NIST has been broken up by code breakers. Hence it is necessary to provide and protect the confidentiality of medical image data, sensible images such as pin numbers used in business transactions, military and navy images used in defense and highly confidential personal images when stored in databases or transmitted over networks of any kind.

NSA-approved Data Encryption Standard published in 1977 resulted in its quick international adoption and widespread academic scrutiny. However, DES was broken and its main disadvantage was its short key length [1]. Concerns about security and the relatively slow operation of DES in software motivated several researchers to propose a variety of alternative designs both mathematically and DNA sequence based. The Advanced Encryption Standard published by NIST in December 2001 then became the standard encryption technique. In AES, the static S-Box denotes SubByte transformation and provides non linearity and confusion, constructed by multiplicative inverse and affine transformation. DNA based Cryptography was first introduced by Gehani *et al*., in which a substitution method using libraries of distinct one time pads were used for encryption [2]. A pseudo DNA cryptography method to encrypt text files was introduced by Ning Kang [3]. A DNA based implementation of YAEA was proposed by Amin *et al*., [4]. A novel image encryption algorithm based on DNA subsequence operations was proposed by Qiang Zhang *et al*. that suffer from differential attack [5]. A hybrid encryption Scheme using DNA technology which has an overhead of having a key image with the same size as that of the original image and transmitting the key image through a secure channel was developed by Grasha and Murugan [6]. An Encryption Scheme with DNA Technology and JPEG Zigzag Coding for Secure Transmission of Images was also developed by Grasha and Murugan [7]. Combining traditional cryptography with biomolecular techniques will make it difficult for an attacker to decrypt the encrypted information. This paper proposes a dual encryption scheme which takes the advantage of both the Key dependent S-Box based on the static S-Box of AES and DNA sequence based encryption for the secure storage and transmission of images ensuring double-fold security.

DEFINITIONS

Definition 2.1 Crypto system: Let P be the plaintext space, C the ciphertext space, and K the key space. Let e_k be the encryption function and d_k be the decryption function. Then for each key $k \in K$, there is an encryption function and a corresponding decryption function such that $d_k(e_k(x)) = x$, for each element ($x \in P$).

Definition 2.2 Nonlinear function: A function with its corresponding vector is said to be highly nonlinear when the resulting vector y_i from a function f_i has a high Hamming distance with all the linear vectors in the set of B_n.

Definition 2.3 Hamming weight: The Hamming weight (H_w) of a binary vector v, is the number of 1's in v.

Definition 2.4 Hamming distance (H_d) between two binary vectors of equal length is the number of places for which the corresponding entries are different.

Definition 2.5 DNA Sequence Crypt function: A DNA Sequence Crypt function is a function that returns one of the many positions of the quadruple DNA sequence in the key DNA sequence file. A one to many DNA Sequence Crypt function is a one-to-many function d(x), which has the following three properties:

a) A pointer, d maps an input quadruple nucleotide sequence in x to one of the many positions obtained at random in the key DNA sequence file selected at run-time based on key.
b) Ease of computation: Given d and an input x, d(x) is easy to compute.
c) Resistance to guess: In order to meet the requirements of a cryptographic scheme, the property of resistance to guess is required of a crypt function with input x, x_1 and outputs y, y_1. As similar quadruple nucleotide sequence that occur in a plain text are mapped to different positions in the DNA nucleotide sequence file(one to many mapping), it is difficult for an attacker to guess the plain-text.

DUAL ENCRYPTION METHOD

In the proposed Dual Encryption method, a codeword is first generated based upon a 64 bit key. For simplicity, the key is denoted as Hex value in this paper. Once the codeword is generated, based upon the codeword, a Dynamic key-dependent S-Box is generated and a DNA sequence file is selected at run-time. Then the input image is transformed into a coded image based on the Dynamic key-dependent S-Box. The coded image is then encrypted using the DNA Sequence file and the encrypted image which is obtained by applying DNA Sequence Crypt function is sent to the receiver through any channel. The sender and the receiver should agree upon the 64 bit-key that is generated and the receiver can decrypt the encrypted image.

The Dual Encryption method consists of two phases. The decryption is the reverse of the encryption process. The steps involved in the two phases of the encryption method are summarized as follows:

Phase 1

CodeImage (inputimage)

Output: Coded image

a) Generate 64 bit key
b) Generate Codeword
c) Generate Key Dependent Dynamic S-Box
d) Convert the given image into a coded image based upon Key Dependent Dynamic S-Box

Phase 2

EncryptImage (codedimage)

Output: Encrypted image

a) Convert the coded image obtained in phase 1 into a DNA image
b) Select a DNA Sequence file at runtime based on the Codeword
c) For each quadruple DNA nucleotide sequence of the coded image, search for a match in the DNA Sequence file at a random position
d) If a match is found return the position of the match in the DNA Sequence file else repeat step 3 until a match is found.

Design criteria for a good S-box

(i) Bijection requires a one-to-one and onto mapping from input vectors to output vectors.
(ii) Strict avalanche criterion requires that if there are any slight changes in the input vector, there will be a significant change in the output vector.
(iii) Correlation-immunity means that output bits act independently from each other.
(iv) Nonlinearity requires that the S-box is not a linear mapping from input to output.
(v) Balance means that each Boolean vector responsible for the S-box has the same number of 0's and 1's.

Generation of codeword

The key used for encryption is considered to be 64 bits in length. From the key, a codeword of 8 bits ($C_8C_7C_6C_5C_4C_3C_2C_1$) is generated at run-time based upon the Hamming Distance and Hamming Weight. The key-dependent S-box generated at runtime based upon the codeword is non-linear in nature. The codeword is also used to randomly select a DNA sequence file. The codeword is generated as follows:

a) Each bit calculates the parity (Hamming weight) for some of the bits in the key.
b) Set a parity bit to 1 if the total number of ones in the positions it checks is odd or 0 otherwise.
c) C8: check all the bits (1-64)
d) C7: check 1 bit, skip 1 bit, check 1 bit, skip 1 bit, etc. (1, 3, 5, 7, 9, 11, 13, 15...)

e) C6: check 2 bits, skip 2 bits, check 2 bits etc. (2, 3, 6, 7, 10, 11, 14, 15,...)
f) C5: check 4 bits, skip 4 bits, check 4 bits etc. (4, 5, 6, 7, 12, 13, 14, 15,...)
g) C4: check 8 bits, skip 8 bits, check 8 bits etc. (8-15, 24-31, 40-47,...)
h) C3: check 16 bits, skip 16 bits, check 16 bits, skip 16 bits, etc. (16-31, 48-56)
i) C2: check 32 bits, skip 32 bits, check 32 bits etc. (32-63)
j) C1: check 1 and 64

The codeword for the key A451 B672 90F7 DE38 and A451 B672 90F7 DE39 are 10111001 and 00111000, respectively.

Key dependent dynamic S-Box function

With the key dependent Dynamic S-Box function, it is a single simple function applied over and over again to each byte of the input image, returning a byte. Each of the 256 possible byte values is transformed to another byte value with the key dependent SubBytes (S-Box) transformation, which is a full permutation, meaning that every element gets changed, and all 256 possible elements are represented as the result of a change, so that no two different bytes are changed to the same byte. Figure-1 represents the key dependent Dynamic SBox generated at runtime, when the key is A451B67290F7DE38. The Dynamic S-Box thus generated satisfies all the criteria for a good S-Box.

36	c7	77	b7	2f	b6	f6	5c	03	10	76	b2	ef	7d	ba	67
ac	28	9c	d7	af	95	74	0f	da	4d	2a	fa	c9	4a	27	0c
7b	df	39	62	63	f3	7f	cc	43	5a	5e	1f	17	8d	13	51
40	7c	32	3c	81	69	50	a9	70	21	08	2e	be	72	2b	57
90	38	c2	a1	b1	e6	a5	0a	25	b3	6d	3b	92	3e	f2	48
35	1d	00	de	02	cf	1b	b5	a6	bc	eb	93	a4	c4	85	fc
0d	fe	aa	bf	34	d4	33	58	54	9f	20	f7	05	c3	f9	8a
15	3a	04	f8	29	d9	83	5f	cb	6b	ad	12	01	ff	3f	2d
dc	c0	31	ce	f5	79	44	71	4c	7a	e7	d3	46	d5	91	37
06	18	f4	cd	22	a2	09	88	64	ee	8b	41	ed	e5	b0	bd
0e	23	a3	a0	94	60	42	c5	2c	3d	ca	26	19	59	4e	97
7e	8c	73	d6	d8	5d	e4	9a	c6	65	4f	ae	56	a7	ea	80
ab	87	52	e2	c1	6a	4b	6c	8e	dd	47	f1	b4	db	b8	a8
07	e3	5b	66	84	30	6f	e0	16	53	75	9b	68	1c	d1	e9
1e	8f	89	11	96	9d	e8	49	b9	e1	78	9e	ec	55	82	fd
c8	1a	98	d0	fb	6e	24	86	14	99	d2	f0	0b	45	bb	61

Figure-1. Dynamic SBox when the key is A451B67290F7DE38.

RESULTS AND SECURITY ANALYSIS

To prove the validity of the proposed algorithm, experiments were performed using two hundred different images of different sizes. The experimental results and security analysis reveal that the proposed algorithm is easy to be implemented, can get good encryption effect, has strong sensitivity to the key and DNA sequence file used, and has the abilities of resisting exhaustive, statistical and differential attacks (strong avalanche effect).

Figure-2 depicts an example of an original image, its corresponding coded, encrypted and decrypted image.

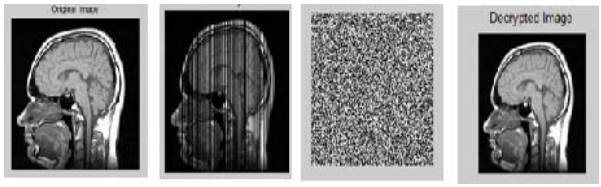

Figure-2. a) Original Image b) Coded image c) Encrypted Image d) Decrypted image.

Security analysis

Security analysis is defined as the technique of finding the weakness of a cryptographic scheme and retrieving whole or a part of the encrypted image without knowing the decryption key or the algorithm. For an encryption scheme to be good, it should be robust against statistical and brute-force attacks, and therefore the proposed method was examined for these attacks.

Statistical analysis

The encrypted image should not have any statistical similarity with the original image to prevent the leakage of information. The stability of the proposed method is examined via statistical attacks - the histogram and correlation between adjacent pixels.

Histogram analysis

The histograms present the statistical characteristics of an image. If the histograms of the original image and encrypted image are different, then the encryption algorithm has good performance. An attacker will find it difficult to extract the pixels' statistical nature of the original image from the encrypted image and the algorithm can resist a chosen plain image or known plain image attack. Figure-3 reveals that the histograms of the encrypted images are fairly uniform and

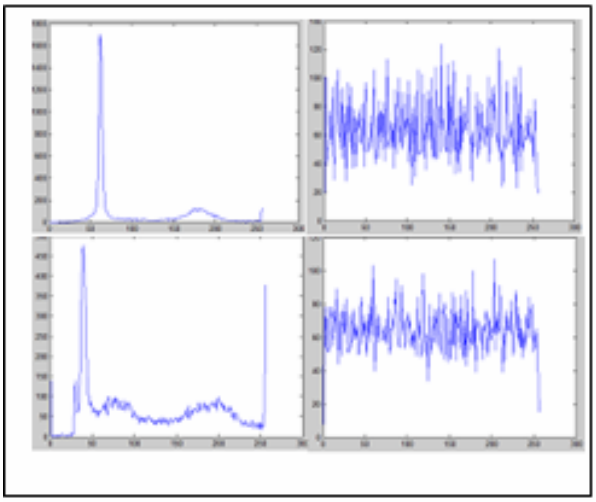

Figure-3 (a) and (b). Histogram of two images-original and encrypted.

significantly different from the original image. As the encrypted image does not provide any information regarding the distribution of gray values to the attacker, the proposed algorithm can resist any type of histogram based attacks and strengthens the security of the encrypted images significantly.

Correlation immunity

Bit independence criterion or correlation-immunity requires that there should not be any statistical dependencies between output bits from the input vectors. Therefore, two highly uncorrelated images have approximately zero correlation coefficient.

The Pearson's Correlation Coefficient determined using the formula:

(1)

where x and y are the gray-scale values of two adjacent pixels in the image and N is the total number of pixels selected from the image to analyze the correlation is used to find the correlation immunity.

Figure-4 (a) to (f) represents the correlation between the adjacent pixels of the original and encrypted images column-wise, row-wise and diagonal-wise. The correlation of the adjacent pixels are highly correlated and almost uniformly distributed in the case of original and encrypted images respectively. The uniformly distributed correlation of encrypted image gives no information to the attacker regarding the nature of the original image that is being transmitted.

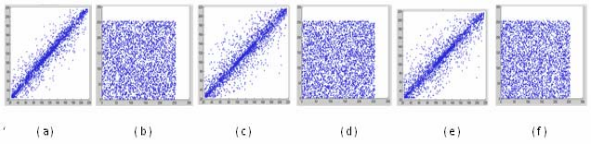

Figure-4. a) Correlation coeff columnwise of original image. b) Correlation coeff column-wise of Encrypted Image c) Correlation coeff rowwise of original image d) Correlation coeff rowwise of Encrypted Image e) Correlation coeff diagonalwise of original image f) Correlation coeff diagonalwise of Encrypted Image.

Avalanche effect

Avalanche effect analysis is to determine the sensitivity of encryption algorithm to slight changes. If an attack is made to create a bit change in the key to observe the results, this manipulation causes a significant change in the codeword and hence the key dependent S-box and the DNA sequence file and the opponent will not be able to find a meaningful relationship between the original and encrypted image with respect to diffusion and confusion. The codeword for the keys A451 B672 90F7 DE38 and A451 B672 90F7 DE39 are 10111001 and 00111000 respectively. A bit change made to the key will result in a completely different S-box coding and a different sequence file and hence a completely different encrypted image proving that the algorithm is highly sensitive to slight changes.

Brute force attack

A Brute Force Attack is a strategy that involves systematically checking all possible keys until the correct key is found. In the proposed method, the encrypted image depends on the key and hence the codeword, the key dependent S-Box generated at run time and the DNA sequence file. A combination of $2^8!$ x $2^8!$ ways will be required to check all the possible keys as there are 2^8 different codewords and hence dynamic S-B0xes and DNA Sequences. Moreover, the encrypted image is actually randomly generated pointers to the DNA sequence file and rarely there is less probability for more than one quadruple nucleotide sequence pointing to the same position in the DNA sequence file. Moreover, the aspect of bio-molecular environment is more difficult to access as it is extremely difficult to recover the DNA digital code without knowing the correct coding technology used. An incorrect coding will cause biological pollution, which would lead to a corrupted image. Since there are many web-sites and roughly 55 million publicly available DNA sequences, it is virtually impossible to guess the key sequence.

CONCLUSIONS

As the criterion for a good S-Box is satisfied by the proposed Key Dependent S-Box, it enhances the security. .Moreover, the complexity and randomness of DNA based encryption provides a great uncertainty which makes encoding of data in DNA format better than other

mechanisms of cryptography. Integrating key dependent S-Box coding and DNA sequence based encryption helps in double fold secure storage and transmission of confidential images. The proposed Dual Encryption Scheme is easy to implement and can resist brute-force, statistical, differential attack and is highly suitable for the secure transmission and storage of images.

REFERENCES

[1] E. Biham and A. Shamir. 1991. Differential Cryptanalysis of DES like Cryptosystems. Journal of Cryptology. 4(1): 3-72.

[2] A. Gehani, T. H. Reif and H. John. 2000. DNA - Based Cryptography. Dimacs Series in Discrete Mathematics and Theoretical Computer Science. 54: 233-249.

[3] K. Ning. 2009. A pseudo DNA cryptography Method. http://arxiv.org/abs/0903.2693.

[4] S. T. Amin, S. Magdy and S. El-Gindi. 2006. A DNA based Implementation of YAEA Encryption Algorithm. IASTED International Conference on Computational Intelligence.

[5] Q. Zhang, X. Xue and X. Wei. 2012. A Novel Image Encryption Algorithm based on DNA Subsequence Operation. The Scientific World Journal. Article Id: 286741.

[6] J Grasha and Murugan A. 2013. A Hybrid Encryption Scheme using DNA Technology. IJCSCS. Vol. 3.

[7] Grasha A. Murugan. 2013. An Encryption Scheme with DNA Technology and JPEG Zigzag Coding for Secure Transmission of Images. http:// arxiv.org/abs/1305.1270.

[8] W. Diffie and E.M. Hellman. 1977. Exhaustive Cryptanalysis of the NBS Data Encryption Standard. Computer. 10(6): 74-84.

[9] B. H. Westlund. 2002. NIST reports measurable success of Advanced Encryption Standard. Journal of Research of the National Institute of Standards and Technology.

[10] B. Monica. 2010. DNA secret writing Techniques. IEEE conferences.

[11] G.Z. Cui, L.M. Qin and Y.F. Wang. 2008. An Encryption Scheme using DNA Technology. Computer Engineering and Applications. pp. 37-42.

[12] S. Sadeg. 2010. An Encryption algorithm inspired from DNA. IEEE. pp. 344-349.

[13] G. Xiao, M. Lu, L. Qin and X. Lai. 2006. New field of cryptography: DNA cryptography. Chinese Science Bulletin. 51: 1139-1144.

[14] G. Cui and L. Qin. 2007. Information Security Technology Based on DNA Computing. IEEE International.

[15] G.N. Krishnamurthy and V. Ramaswamy. 2008. Making AES Stronger: AES with Key Dependent S-Box. IJCSNS International Journal of Computer Science and Network Security. 8: 388-398.