A NEW ENCRYPTION TECHNIQUE FOR THE SECURED TRANSMISSION AND STORAGE OF TEXT INFORMATION WITH MEDICAL IMAGES

Prashant VAIDYANATHAN – Nitish MALHOTRA – Jagadish NAYAK

Abstract: *Modern day hospital management systems rely heavily on electronic data processing to maintain patient records. These electronic medical records (EMRs) must be maintained in an unaltered form by the creator. The need for a secure data handling method for the transmission and storage of text and digital media, comprising patient's diagnostic history, imaging, scans, etc., is indispensible. This paper presents a novel method of text encryption by means of symmetric key encryption technique, using variable length key derived from the encrypted text itself.*

Keywords: – encryption
– watermarking
– Fibonacci
– steganography
– security
– cipher

1. INTRODUCTION

The key to modernization of healthcare centers is to shift from clinical case record sheets handling to the use of digital media. Apart from providing the quality of healthcare, a good medical care relies mainly on a safe record keeping system. The information such as medical images (X-rays, MRI, and C.T Scans), personal information of patients and their medical history must be protected from malignant attacks and transformations. Any alteration to these records may lead to a misdiagnosis of condition, which may further have serious consequences, often proving fatal to the patient's life. An effective and efficient record handling system requires an enhanced level of security to keep these files safe [1].

Many symmetric key encryption techniques are available in literature. Some of them are encryption standards such as AES and DES [2, 3]. A symmetric key scheme utilizes a shared key for encryption and decryption. The message which is encrypted using this key can only be decrypted back using the same secret key. In this system, the key at both ends may be identical or easily derived from the original key by simple transformations of the encryption key. This key is shared between two or more parties to maintain a private info link. However, such systems are vulnerable to attacks and are easily compromised once the key has been discovered. One of the methods to keep the key a secret is to use a password based encryption algorithm in which the secret key is derived at both ends using a user supplied text string. Nevertheless, such a system also requires memorization or a safe handling of the password string. Another trivial method would be the safe transmission of the key exchange between the users by hand or by emails. The method of transmission or storage of patient text information with medical images has already been implemented by Rajendra et al [4]. Certainly, this paper does not account for the enhanced security of the patient text information, which needs to be protected from unauthorized access. This paper proposes a system where patient text information is encrypted and interleaved into a medical image.

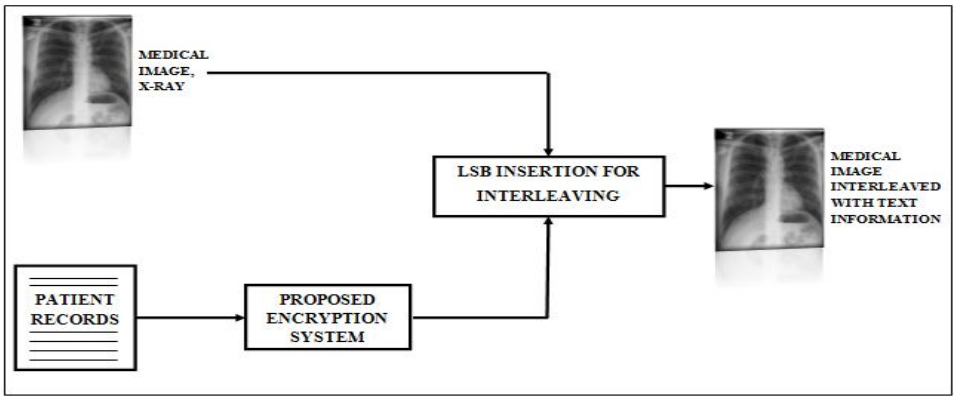

Figure 1. Block Diagram of Proposed System

2. PROPOSED SYSTEM

In this paper we propose an encryption technique, which is a blend of symmetric key encryption and steganography while eliminating the need for a separate key sharing method**.** The block diagram of proposed system is shown in Fig. 1. Based on a variant of the *Fibonacci Spiral* called 'The Whirling Squares'[5] algorithm, this encryption technique provides multilevel security and an easy and secure way of pass-key transmission by incorporating encryption with steganography. The encrypted text is subsequently interleaved into the medical images in the spatial domain using the bit insertion method.

2.1. Encryption Method

The plain text such as patient information is given a primary layer of security by using a pattern based on the arrangement of characters within a matrix. A secondary layer of security is added by using symmetric key encryption to manipulate the rearranged data. The pass key is hidden in the text itself, based on a pattern known only to the users. This pattern, known as the Fibonacci spiral [or Golden Rectangle], widely present in nature, is based on the unusual properties of the Fibonacci series which was discovered by Leonardo Fibonacci [6]. This pattern based password storage eliminates the need to memorize a long string of bits or characters to use as the secret key. All the user is required to remember is the password hiding pattern and the algorithm to extract the key from the given cipher text. This also provides the system with a variable length key derived each time depending on the length and the characters of the plaintext. This key is finally used to encrypt the given randomized string using any of the proposed standards of Block Cipher Modes as set by National Institute of Standards and Technology (NIST) [7]. The proposed method used in this paper is inspired by Vigenere Ciphers. It X-ORs the key with the various blocks of data in the cipher text. The entire process comprises 3 steps Steganography & Encryption, Digital Watermarking and Extraction & Decryption.

2.2. Steganography & Encryption

The algorithm uses 3 levels of encryption to ensure that the data cannot be decrypted using basic brute force techniques. The following steps are followed to get the cipher text:

i. Use of 'Fibonacci Golden Rectangle' to store the data into a matrix.
ii. Use of a 'Chess Knight's Movement Pattern' algorithm to arrange the data in this rectangle.
iii. Use of the 'Golden Spiral Pattern or Whirling Squares Pattern' to store the pass-key.
iv. Select the nth Fibonacci number (F_x) such that [F_x . F_{x+1}]> [Size of text + 6] (the plus six ensures that the minimum number of rows is at least 3)
v. Generate the Golden Rectangle using "COLUMNS = F_{x+1} & ROWS = F_x".
vi. Using a row-wise allocation method, tag the last position to be occupied by the input string (assuming – sequential movement) and flag the consequent remaining positions. The flagged areas can be marked with any special ASCII character (character used here is, '#' – Fig. 2)

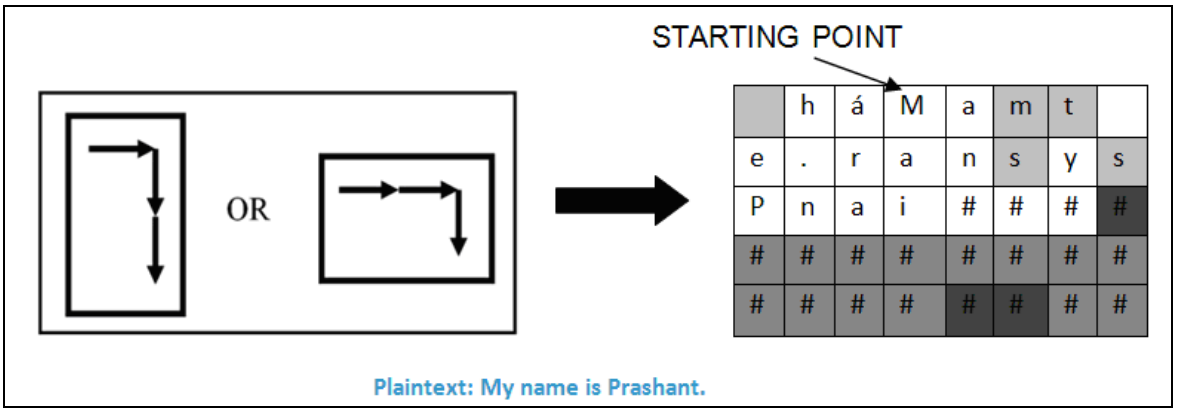

Figure 2. Knight Movement Pattern and its implementation in the golden rectangle

vii. Decide upon a certain value which is unique to every plaintext. One method would be to add the ACII value of every character of the plaintext and use a fixed fraction to obtain an index in the array and store the ASCII value of the character in that position in 'y'. There is no rigid or sophisticated algorithm required for choosing a value for 'y'. Any arbitrary formula can be used to find a unique value of 'y'. The method used in this paper is by finding out the value of $3/8^{th}$ of the total string length, and storing the ASCII value of the character in that index. For the purposes of decryption, the value of this ASCII character can be flagged by incrementing the value by 128.

viii. Prior to inserting the first character, iterate the input loop 'N + y' times (N – size of text) where, in the first 'y' runs of the loop, no text is filled in. At the $y+1^{th}$ iteration, the first character of the input text is filled in. This is done in order to protect the system against re-iterative attacks, where the encryption pattern is run N times to reproduce the original text.

ix. The text is filled into the Golden rectangular matrix using a Chess Knight's Movement pattern (demonstrated above – Fig. 2) while skipping the flagged positions.

x. To extract the password, the characters located at certain marked areas described by the "Golden Spiral or Whirling Squares Pattern" in the Golden Rectangle. (See Fig. 3) in the matrix are retrieved and stored into a text file.

xi. Using Cipher Block Chaining (CBC), encrypt the remaining text characters with the pass key string (keeping the Pass-key characters unaffected) [7]. In the example used in this paper we use a simple single X-OR operation. The ASCII value for '#' is taken as 7 for this paper.

xii. Replace all characters back to their original positions as they were (after step (x) with the pass-key) retained in their original form.

2.3. Digital Watermarking

We use the Least Significant Bit (LSB) insertion method to embed the given digital data into the LSBs of the cover, medical image in a sequential fashion [8, 9, 10].

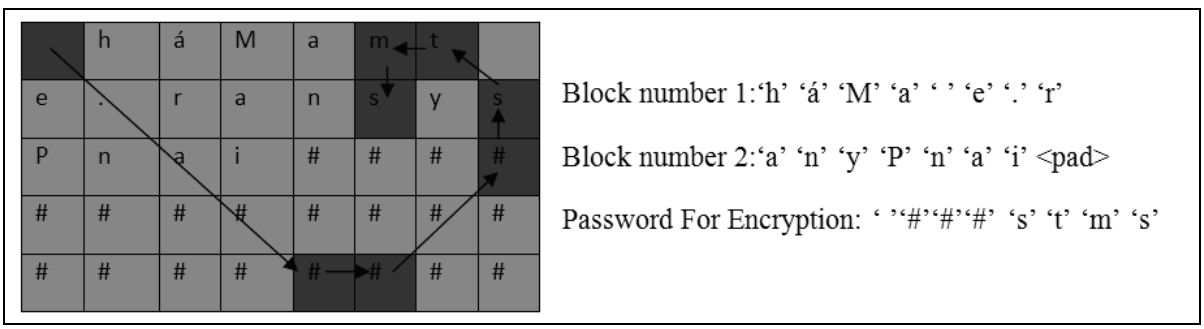

Figure 3. Fibonacci Spiral Pattern

2.4. Extraction and Decryption

i. The first password is found using an encoded character whose ASCII value is lesser than 0. This character is used to locate the starting point for the decryption process in the 'Golden Rectangle'. This password follows a set of rules as given: (N'+6) >= (F_x . F_{x+1}). Where – N' is the Size of the Encoded Text and Y=F_n (nth Fibonacci Number)

ii. This password helps form the matrix using the parameters found in the previous step, where F_x forms the number of Rows and F_{x+1} forms the columns.

iii. A linear search is performed by the software to find a character whose ASCII value is less than 0. The value of this ASCII is stored in an integer 'm' and to this value, 128 is added. This helps to find the position in the golden rectangle, where the software must start extracting the values while the iterations are running.

iv. The extracted text is sequentially filled into this rectangle and the extra spaces are tagged with a special character.

v. The text is still in the encrypted form, for which we extract the 2nd password by following the 'Golden Spiral/Whirling Squares' Pattern again to extract the password characters.

vi. This password string is then XOR-ed with the remaining characters using CBC as done in encryption to reproduce the original decrypted characters [7].

vii. The decrypted characters are now placed back into their prior locations with the password string occupying the original spiral position.

viii. The final step involves the usage of the 'Knight's Movement Pattern' to extract the text as it was done to hide it in the encryption process. The following steps are used for this extraction.

ix. Start from the position (0,0) and skip the first 'm' iterations (where m was found in step iii)

x. Start storing the values into a linear array with characters while skipping the flagged areas. Store the final array into a text file.

2.5. Example of Plaintext to Cipher Text

The following example will help understand the above mentioned encryption technique in a better manner.
Let the Plaintext be:

"My name is Prashant."

After the first layer of encryption, this plaintext yields the following matrix:

	h	á	M	a	m	t	
E	.	r	a	n	s	y	s
P	n	a	i	#	#	#	#
#	#	#	#	#	#	#	#
#	#	#	#	#	#	#	#

Hence, the first Cipher Text we get is:

" háMamt e.ransysPnai"

From the Golden Rectangle:
We obtain the key:

" ###stms"

And the following blocks:

1. "háMa e.r"
2. "anyPnai<PAD>"

Hence while X-ORing the blocks with the key:
(assume that the ASCII value of '#' is 7)

We get:

1. *"HºJfS(DCI)C(SOH)"*
2. *"Ai~W(GS)(NAK)(EOT)"*

(To understand how two characters are X-ORed, refer to Table 1.)

Thus, the final Cipher text is:

"HºJfmtS(DCI)C(SOH)Ais~sW(GS)(NAK)(EOT)"
where:
(EOT) – End of Transmission
(DC1) – Device Control 1
(SOH) – Start of heading
(GS) – Group Separator
(NAK) – Negative Acknowledgement

Table 1. CBC Encryption using X-OR

	1	2	3	4	5	6	7	8
		#	#	#	s	t	m	s
KEY	32	7	7	7	115	116	109	115
	00100000	00000111	00000111	00000111	01110011	01110100	01101101	01110011
	h	á	M	a		e	.	r
Block 1	104	160	77	97	32	101	46	114
	01101000	10100000	01001101	01100001	00100000	01100101	00101110	01110010
	01001000	10100111	01001010	01100110	01010011	00010001	01000011	00000001
X-ORed value	72	167	74	102	83	17	67	1
	H	°	J	f	S	(DCI)	C	(SOH)

3. CONCLUSION

A new algorithm has been developed to provide a secure way to transmit and store hospital data and medical records. The encryption method is made secure against re-iterative attacks by altering the starting point of the knight's pattern movement with every iteration and resulting encoded text. As Fig. 4 shows, the interleaved image looks exactly like the original image. After conducting a PSNR test on the two images, using the formula:

$$MSE = \frac{1}{MN} \sum_{y=1}^{M} \sum_{x=1}^{N} [I(x,y) - I'(x,y)]^2$$

$$PSNR = 20 * \log_{10} (255 / sqrt(MSE))$$

Where:
I: Original Image
I' : Interleaved Image
MSE: Mean Square Error
PSNR: Peak Signal To noise ratio
M: Number of pixels in each row of the image
N: Number of pixels in each column of the image
The PSNR obtained was 72.169 dB. When the MSE is lower, the PSNR value would be higher. In medical images, a PSNR ratio of at least 40-50dB is considered acceptable [11]. Hence 72.169 dB would imply that the image has very little noise.
Hence problems such as loss of precious data or addition of unwanted noise in medical images do not arise.

4. DISCUSSION

The textual data can be given a further level of security by improving on the LSB insertion method and by using truly random arrangement of data within the cover medical image. The robustness of the system can be improved by using ECC[12] prior to interleaving the data onto the image, making the data secure against channel noise and external interferences or single bit attacks.

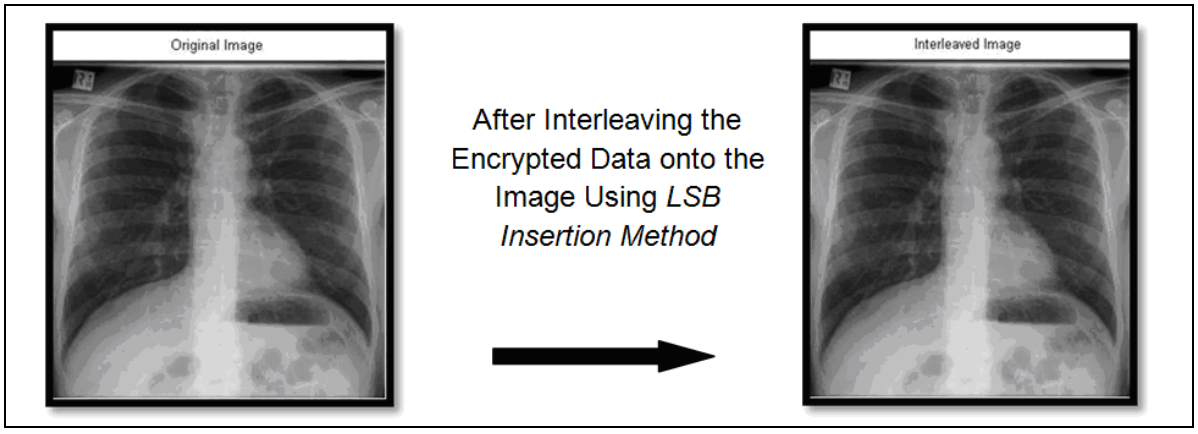

Figure 4. Original Image and the interleaved Image

The CBC encryption provides better security [7] when methods like checking the frequency of a particular character in an encoded text is used, it will be difficult to break the code without knowing the encryption password. It can be observed in the example in the Section 2.5 that although the 2nd block had repeated characters, the final cipher text block does not have any repeated characters. Moreover, other methods such as trial and error as well as brute force techniques in which characters are substituted or transposed do not yield fruitful results of decrypting the cipher text as any slight change in the input text changes the pattern in which data is arranged significantly.

5. ANALYSIS

To gauge the strength of the first layer of security where the cipher text is simply made up of data scrambled using the Knight's pattern in the Golden Rectangle (without using CBC Encryption), a survey was conducted by sending plaintext and the corresponding cipher text as well as the program code to a 5 programmers, fellow students and professors who were talented in cryptography.
The survey was conducted in 2 stages.

5.1. Analyzing the Cipher

Six students talented in cryptography and two professors were sent a file containing 10 plaintext files and its corresponding cipher text.
Using these samples, the cryptographers were asked to recognize any patterns in the encryption if they did see any. They were also given 5 cipher texts and were asked to find the original plain text of each cipher text.

5.2. Analyzing the Algorithm

An executable file was made in such a way that the user could put the input in a plaintext and the program would give the cipher text as the output. This executable file was sent to 5 programmers. They were asked to use this executable file to figure out the algorithm of the program. The programmers were allowed to team up with the cryptographers to figure out the algorithm. The executable file was treated as a black box; the source code of the executable was not viewable.

5.3. Results from the survey

Almost every single student analyzing the crypt said that the encryption changed with the length of the plaintext. However it was noted that all brute force techniques that were tried did not yield any conclusive pattern. Since the encryption pattern changed with the change of any character, or with the change of length,, it was very difficult to find the manner in which the code was being arranged in the cipher box. It was also noted that it was increasingly difficult to break the code when the cipher text had multiline inputs.

5.4. Examples

(The Encrypted text in all the examples shows the cipher text after the first level of encryption, but do not show the CBC encryption)

1)Text:
"Hello World. My name is Prashant"

Encrypted Text:
"de Mlr saonrti.lPyoanWhmaelHsá"

2)Text: (changing just 1 word)
"Hello World. My name is Ameya"

Encrypted Text:
"l lyn momyos.HAMliäaaWeere" t

3)Text (Multiline input):
"Name: Robin
Age:44
Sex: Male
Date: 20/12/2011
Condition: Liver Cirrhosis"

Encrypted Text:
"Ni
::
r24h
CMatDR A /
oSsoamiaoL2gC2i0einleotbi0e/:r1xsdå:neivne14r1:"

5.5. Crypt Analysis of Final Cipher Text (including CBC block encryption)

When cryptographers were given the final cipher text which was to be digitally watermarked in the image, they found it impossible to use techniques such as frequency analysis to crack the original

code. Since the size of the blocks, and the key used for CBC encryption was within the code, using brute force to try every single permutation and combination proved to be too much time consuming and inconclusive.

REFERENCES

[1] Ragib Hasan, Marianne Winslett and Radu Sion, "Requirements of Secure Storage Systems for Healthcare Records", Secure Data Management Lecture Notes in Computer Science, Springet, Verlag, 2007, Volume 4721/2007, 174-180

[2] J.Daemen and V.Rijmen, AES Proposal: Rijndael, AES Algorithm Submission, September 3, 1999, available at AES page which can be accessed at the following website http://www.nist.gov/CryptoToolkit.

[3] Westlund, Harold B. (2002). "NIST reports measurable success of Advanced Encryption Standard". Journal of Research of the National Institute of Standards and Technology.

[4] U. Rajnendra Acharya, D. Anand, P. Subbanna Bhat, U.C.Niranjan, Compact storage of medical images with patien information, IEEE Trans. Information Technol. Biomed. 5 (4)(2001) 320—323.

[5] Ball, Keith M. (2003). "Chapter 8: Fibonacci's Rabbits Revisited". Strange Curves, Counting Rabbits, and Other Mathematical Explorations. Princeton University Press. ISBN 0691113211.

[6] H.E.Huntley, " The Divine Proportion: A Study of Mathematical Beauty" page 157, Dover Publications; First edition. edition (June 1, 1970)

[7] M. Dworkin. Recommendation for block cipher modes of operation. Special Publication 800-38A, NIST, 2001.

[8] Johnson, N. F. and Jajodia, S, "Exploring steganography: Seeing the unseen", IEEE Computer Magazine, pp. 26-34, February 1998.

[9] S.K.Bandyopadhyay, Debnath Bhattacharyya, Poulumi Das, S. Mukherjee, D. Ganguly, "A Tutorial Review on Steganography", IC3 Noida, pp. 106-114, August 2008.

[10] Jagadish Nayak, P Subbanna Bhat, Rajendra Acharya U, M. Sathish Kumar, "Reliable Storage And Transmission Of Retinal Fundus Images with Patient Information Using Reversible Watermarking Technique And Error Control Codes", Journal of Medical Systems , Springer , Netherlands,Volume-33, Page: 163-171, April-2009.

[11] K.Chen and T.V. Ranmabadran, "Near-Lossless Compression of Medical Images Through Entropy-Coded DPCM" IEEE Trans. Medical Imaging, Vol. 3, No. 3, pp 538-548, 1994

[12] Lin, Shu; Costello, Daniel J. Jr. (1983). Error Control Coding: Fundamentals and Applications. Englewood Cliffs NJ: Prentice–Hall. ISBN 0-13-283796-X.

Received: 11. 11. 2011.

Accepted: 31. 01. 2012.

Authors' address
Prashant Vaidyanathan
BITS, Pilani – Dubai Campus
Phone: +971554036510, +919821692218
Nitish Malhotra
University of Pennsylvania
Phone: +1(215)429-0771
Dr Jagadish Nayak ,
ECE Department , BITS PILANI Dubai Campus
Dubai , UAE
Phone: 00971 55 4907979, Fax: (009714) 4200555.
vprashant1@gmail.com
nitish.malhotra@gmail.com
jagadishnayak@bits-dubai.ac.ae