An Authenticated Two-Tier Security on Transmission of Medical Image using Codebase Cryptosystem over Teeming Channel

Deo Brat Ojha
Deptt. Of Mathematics, R.K.G.I. T., Gzb., U.P. (India),

Ajay Sharma
Deptt. of I.T., R.K.G.I. T., Gzb., U.P.(India), (Research Scholar Singhania University, Raj. INDIA)

Abhishek Dwivedi
Deptt. of M.C.A., R.K.G.E.C., Gzb., U.P.(India), (Research Scholar Singhania University, Raj. INDIA)

Bhupendra Kumar
IIMT Meerut U.P., INDIA (Research Scholar Mewar University, Raj. INDIA)

Amit Kumar
Innovative College of Education & Technology, Greater Noida,U.P., INDIA (Research Scholar Mewar University, Raj. INDIA)

ABSTRACT

In this paper, we present an efficient and authentic transmission scheme without digital signature on an algebraic coding theory based public key cryptosystem of medical information like Image, Audio, Video etc. via the combination of stegnographic and cryptographic approach. Stegnographic approach facilitates data hiding, McEliece public-key cryptosystem to encrypts the data,.lossless compression provides us integrity and Fuzzy error correction provides error less message over noisy channel.

Keywords

McEliece cryptosystem, Cryptography, Stegnography, Compression, Image File, Error Correction Code, SEQUITUR .

1. INTRODUCTION

The necessity of fast and secure diagnosis is vital in the medical world to save the life of world creature. Nowadays, the transmission of images is a daily routine and it is necessary to find an efficient way to transmit them over the net [3, 4, 5]. For image transmission, two different approaches of technologies have been developed. The first approach is based on content protection through encryption [1], [2]. In this approach, proper decryption of data requires a key. The second approach bases the protection on digital watermarking or data hiding, aimed at secretly embedding a message into the data. In the current era, the transmission of Image over internet is so much challenging over the internet. In this manner, the better way to transmit the image over internet is encryption. Using the cryptography we secure the image as well as also better utilization of the communication channel with compression technique.

Cryptography is a tool of security that aims to provide security in the ciphers of any kind of messages. Cryptographic algorithms use encryption keys, which are the elements that turn a general encryption algorithm into a specific method of encryption. The data integrity aims to verify the validity of data contained in a given document. [2]

McEliece proposed the first public-key cryptosystem (the McEliece Scheme) based on algebraic coding theory in 1978[1]. The idea behind McEliece public-key cryptosystem is based on the fact that the decoding problem of an arbitrary linear code is an NP-hard problem [2].The McEliece scheme has the advantage of high speed encryption and decryption and this system employs probabilistic encryption [1,2], which is better than other type of deterministic encryption[9] in preventing the elimination of any information leaked through public-key cryptography.

It is point of remark [9] that the security comparison is made here for classical attackers. The picture changes drastically to the advantage of the McEliece system if we consider two systems.

It is point of remark [9] that the security comparison is made here for classical attackers. The picture changes drastically to the advantage of the McEliece system if we consider two systems. to offer the same level of security if breaking them requires quantum computers with the same number of qubits. This cryptosystem cannot be used for authentication because the encryption is not one to one and total algorithm is truly asymmetric.

In this current article, we describe SEQUITUR, an algorithm that infers a hierarchical structure from a sequence of discrete symbols. The ability to deal easily with long sequences has greatly extended the range of SEQUITUR's application. By introducing McEliece cryptosystem and SEQUITUR algorithm, an errorless image transmission with compression and encryption can be achieved. We estimate the errors which are generated during transmission and correct these errors through fuzzy error correcting code. This arrangement distributes in different phases and each phase plays an important role in manner.

2. PRELIMINARIES
2.1. Data Compression

A compression scheme can be employed what is known as lossless compression on secrete message to increase the amount of hiding secrete data, a scheme that allows the software to exactly reconstruct the original message [6].

The transmission of numerical images often needs an important number of bits. This number is again more consequent when it concerns medical images. If we want to transmit these images by network, reducing the image size is important. The goal of the compression is to decrease this initial weight. This reduction strongly depends of the used compression method, as well as of the intrinsic nature of the image. Therefore the problem is the following:

1. To compress without lossy, but with low factor compression. If you want to transmit only one image, it is satisfactory. But in the medical area these are often sequences that the doctor waits to emit a diagnostic.

2. To compress with losses with the risk to lose information. The question that puts then is what the relevant information is to preserve and those that can be neglected without altering the quality of the diagnosis or the analysis. The human visual system is one of the means of appreciation, although subjective and being able to vary from an individual to another. However, this system is still important to judge the possible causes of degradation and the quality of the compression [7].

2.1.1 The SEQUITUR Algorithm [8]

The SEQUITUR algorithm represents a finite sequence as a context free grammar whose language is the singleton set $\{\sigma\}$. It reads symbols one-by-one from the input sequence and restructures the rules of the grammar to maintain the following invariants:

(A) no pair of adjacent symbols appear more than once in the grammar, and

(B) every rule (except the rule defining the start symbol) is used more than once. To intuitively understand the algorithm, we briefly describe how it works on a sequence 123123. As usual, we use capital letters to denote non-terminal symbols. After reading the first four symbols of the sequence 123123, the grammar consists of the single production rule S ➔ 1, 2, 3, 1 where S is the start symbol. On reading the fifth symbol, it becomes S ➔ 1, 2, 3, 1, 2 Since the adjacent symbols 1, 2 appear twice in this rule (violating the first invariant), SEQUITUR introduces a non-terminal A to get

S ➔ A, 3,A   A ➔1, 2

Note that here the rule defining non-terminal A is used twice. Finally, on reading the last symbol of the sequence 123123 the above grammar becomes

S ➔ A, 3, A, 3   A ➔ 1, 2

This grammar needs to be restructured since the symbols A, 3 appear twice. SEQUITUR introduces another non-terminal to solve the problem. We get the rules

S ➔ B,B   B ➔ A 3   A➔ 1 2

However, now the rule defining non-terminal A is used only once. So, this rule is eliminated to produce the final result.

S ➔ B, B   B ➔ 1, 2, 3

Note that the above grammar accepts only the sequence 123123.

2.2 Error Correction Code:

A metric space is a set C with a distance function dist $: C \times C \to R^+ = [0, \infty)$, which obeys the usual properties(symmetric, triangle inequalities, zero distance between equal points)[12].

Definition : Let $C\{0,1\}^n$ be a code set which consists of a set of code words c_i of length n. The distance metric between any two code words c_i and c_j in C is defined by

$$dist(c_i, c_j) = \sum_{r=1}^{n} |c_{ir} - c_{jr}| \qquad c_i, c_j \in C$$

This is known as Hamming distance [12].

Definition : An error correction function f for a code C is defined as $f(c_i) = \{c_j / dist(c_i, c_j)$ is the minimum, over $C - \{c_i\}\}$. Here, $c_j = f\ c_i$ is called the nearest neighbor of c_i [12].

Definition :The measurement of nearness between two code words c and c' is defined by nearness $(c, c') = dist(c, c')/n$, it is obvious that $0 \le$ nearness $(c,c') \le 1$ [12].

Definition : The fuzzy membership function for a codeword c' to be equal to a given c is defined as[12]

$$FUZZ(c') = 0 \qquad \text{if nearness}(ç, c') = z \le z_0 < 1$$
$$= z \qquad \text{otherwise}$$

2.3. Cryptography

Cryptography is a branch of applied mathematics that aims to add security in the ciphers of any kind of messages. Cryptography algorithms use encryption keys, which are the elements that turn a general encryption algorithm into a specific method of encryption. The data integrity aims to verify the validity of data contained in a given document. [7]

2.3.1 McEliece Public-Key Cryptosystem[1,2]

It is assumed that McEliece public key(P_A) is duly certified and public. It can be described by its $k \times n$ generator matrix G. With the aid of a regular $k \times k$ matrix S and an $n \times n$ permutation

matrix P, a new generator matrix G' is constructed that hides the structure of G:

G' = S . G . P

The public key consists of G' and the matrices S and P together with g(x) are the private key(S_A).

2.4 Stegnography

Steganography is a technique used to embed secret information into non-secret information, preventing the message from being detected by non-authorized people.[14]

The purpose of steganography is to hide the very presence of communication by embedding messages into innocuous-looking cover objects, such as digital images. To accommodate a secret message, the original cover image is slightly modified by the embedding algorithm to obtain the stego image. The embedding process usually incorporates a secret stego-key that governs the embedding process and it is also needed for the extraction of the hidden message [15].

There are three basic views behind hiding information. The first is capacity, which is the amount of information that can be embedded within the cover file. An information-hiding algorithm has to be able to compactly store a message within a file. Next is security, which refers to how a third-party can detect hidden information within a file. Intuitively, if a message is to be hidden, an ideal algorithm would store information in a way that was very hard to notice. High security layers have been proposed through three layers to make it difficult to break through the encryption of the input data and confuse steganalysis too. Various encryption techniques like cryptography, digital watermarking, steganography etc have already been introduced in attempt to address these growing concerns [16].

Steganography have four application areas:

- Copyright Protection. It has security, invisibility and robustness requirements. Watermark techniques fit in this area.
- Authentication. It has security and invisibility requirements. Digital signature fits in this area.
- Secret and Invisible Communication. It has requirements for security, invisibility and insertion of high volumes of secret data. [16]

3. PROPOSED SCHEME

In our proposed scheme, we use sequitur as a compression technique and McEliece as an encryption technique, this cryptosystem cannot be used for authentication because the encryption is not one to one and total algorithm is truly asymmetric. Sequitur is a single-pass hierarchical algorithm that builds a context-free grammar for a string. The resulting grammar compactly represents the original structure and has the interesting property that the compressed format itself contains useful information about the string. The McEliece scheme has the advantage of high speed encryption and decryption and this system employs probabilistic encryption. Here we concatenate with give us authenticity of sender without digital signature.

3.1 Algorithm for encrypting the confidential message

This cryptosystem cannot be used for authentication because the encryption is not one to one and total algorithm is truly asymmetric so we process as shown below to add authenticity without digital signature.

1. Message m in the form of bitstring to which she wishes to send information.

2. Alice generates a secret pseudo q-bit random vector r.

3. Alice has a identifier Id_A of p-bit random vector.

4. Alice concatenate her identifier Id_A with secret pseudo q-bit random vector r which give us a vector R= $Id_A \square r$.

Here $h(m) = mP_A$ where $h(m) \subseteq GF(2^n)$,

Encryption: $C = mP_A \oplus e$, where $e = g(R)$, here g is an invertible function which maps R in to an n-bit error vector of weight α .

Then after acceptance ,Bob decrypt the massage as first m can be recovered by using the decryption algorithm in the original scheme. In the meantime, the value $g(R)$ can also be obtained. Then the receiver computes $R = g^{-1}(g(R))$, where g^{-1} is the inverse of g .Ffinally Bob calculates $f(c')(SGP)^{-1}$ and get the message. Here Bob get the Id_A from the R to know the authenticity of the sender.

3.2 Algorithm for compress the confidential message

Step:

Perform the lossless compression technique (sequitur) on cipher text to increase the amount of hiding secrete data.

3.3 Process for convert cover image file

Step 1: Generating $n \times n$ blocks

In RGB space the image is split up into red, blue and green images. The image is then divided into 8×8 blocks of pixels and accordingly the image of $w \times h$ pixels will contain $W \times H$ blocks. Where, $W = w/8$, $H = h/8$.

Phase 2: DCT

All values are level shifted by subtracting 128 from each value. The Forward Discrete Cosine Transform of the block is then computed. The mathematical formula for calculating the DCT is:

$$T(u,v) = \sum_{x=0}^{n} \sum_{y=0}^{n} f(x,y) . g(x,y,u,v)$$

Where, $g(x, y, u, v) = \frac{1}{4}\alpha(u)\alpha(v)\cos\left[\frac{(2x+1)u\pi}{2n}\right]\cos\left[\frac{(2y+1)v\pi}{2n}\right]$

Where $\alpha(u) = \begin{cases} \frac{1}{\sqrt{2}}\dots\dots\dots\dots\dots\dots\dots\dots\dots.for.u = 0 \\ 1\dots\dots\dots\dots\dots\dots.for.u = 1,2,\dots.N-1 \end{cases}$

Phase 3: Quantization

Quantization is the step where the most of the compression takes place. DCT really does not compress the image, as it is almost lossless. Quantization makes use of the fact that, the high frequency components are less important than the low frequency components. The Quantization output is

$Q_{DCT} = round(\frac{T(u,v)}{Z(u,v)})$

The $Z(u, v)$ matrix could be anything, but the JPEG committee suggests some matrices which work well with image compression.

Phase 4: Compression using SEQUITUR

After quantization, the scheme uses a filter to pass only the string of non-zero coefficients. By the end of this process we will have a list of non-zero tokens for each block preceded by their count.
DCT based image compression using blocks of size 8x8 is considered. After this, the quantization of DCT coefficients of image blocks is carried out. The SEQUITER compression is then applied to the quantized DCT coefficients.

3.4 Algorithm to embed confidential message into cover image file.

Algorithm to embed confidential message into cover image file named inFile generate new file with embedded message file named outFile.

Encoded-Message (msg,inFile on input-mode, outFile on output-mode)

Step 1:

Read offset bytes from input inFile and writes to output File outFile

Step 2:

Calculate message length and write it into output file by embedding using XOR function it in last two bits for every byte. Suppose, Message length being 16 bits, will be stored in 8 pairs of 2 bits.

Step 3:

Embed each byte of message in 4 pairs of 2 bits each is embedded in 4 byte of input file and written into output file named outFile.

Step 4:

Write the remaining bytes of the input file into output file.

3.5 Algorithm for generate of message from Image

The picture is received at receive side. This function decode message from a file named outFile open on output mode. Decode Message (outFile on Input-mode)

Step 1:

Read offset bytes from the input file and apply again XOR function, Generate message bit.

Step 2:

Read last 2 bits of consecutive 8 bytes and concatenate them to get the message length.

Step 3:

Read last 2 bits from input file in pairs of 4 and concatenate them to get message of 1 byte.

Step 4:

Repeat step 3 until the message is extracted of calculated length.

Step 5:

Decompress & Decrypt the message.

3.6 Procedure for detecting and correcting error

If any error occurred during the transmission of message, we can detect and correct using fuzzy error correcting code.

Receiver check that $dist(t(c)c') > 0$, he will realize that there is
an error occur during the transmission. Receiver apply the error correction function f to $c' : f(c)$.
Then receiver will compute nearness $(t(c), f(c')) = dist(t(c)f(c'))/n$

$$FUZZ(c') = 0 \quad \text{if } nearness(c, c') = z \le z_0 < 1$$
$$= z \quad \text{otherwise}$$

4. CONCLUSION

In the new scenario the Health Insurance Portability and Accountability Act (HIPAA) [11] requires that medical providers and insurance companies implement procedures and policies to protect patient's medical information.

In this paper attempt has been made to explain an authentic transmission scheme without digital signature on an algebraic coding theory based public key cryptosystem which relay on the difficulty of decoding and proposed by McEliece in 1978. The main feature of this approach is randomness of the error vector

concatenate with identifier, here identifier provide the authenticity with randomness.

The efficiency and security of McEliece cryptosystem comparatively better than the RSA cryptosystem also[9,10]. Here we use sequitur as a compression technique and Sequitur has the ability to read a stream in reverse and Stegnographic approach facilitates data hiding which provide user two-tier security. So our approach is more appropriate, secure and futuristic than previous literature of medical data transmission over un-secure and noisy channel. .

5. REFERENCES

[1] .R.J.McEliece, " A public-key cryptosystem based on algebraic coding theory," DSN Progress Report, 42-44, 1978, pp.114-116.

[2]. E.R.Berlekemp, R.J.McEliece, and H.C.A. vanTilborg, " On the inherent intractability of certain coding problems," IEEE Transactions on Information Theory, vol.24, No.5, 1978, 384-386.

[3] G. Lo-varco,W. Puech, and M. Dumas. "Dct-based watermarking method using error correction codes", In ICAPR'03, International Conference on Advances inPattern Recognition, Calcutta, India, pages 347–350, 2003.

[4]. R. Norcen, M. Podesser, A. Pommer, H.P. Schmidt, and A. Uhl. "Confidential storage and transmission of medical image data", Computers in Biology and Medicine, 33:277–292, 2003.

[5] Diego F. de Carvalho, Rafael Chies, Andre P. Freire, Luciana A. F. Martimiano and Rudinei Goularte, "Video Steganography for Confidential Documents: Integrity, Privacy and Version Control" , University of Sao Paulo – ICMC, Sao Carlos, SP, Brazil, State University of Maringa, Computing Department, Maringa, PR, Brazil.

[6].Nameer N. EL-Emam, "Hiding a Large Amount of Data with High Security Using Steganography Algorithm" Applied Computer Science Department, Faculty of Information Technology, Philadelphia University, Jordan

[7]. Borie J., Puech W., and Dumas M., "Crypto-Compression System for Secure Transfer of Medical Images", 2nd International Conference on Advances in Medical Signal and Information Processing (MEDSIP 2004), September 2004.

[8]. N.Walkinshaw, S.Afshan, P.McMinn "Using Compression Algorithms to Support the Comprehension of Program Traces" Proceedings of the International Workshop on Dynamic Analysis (WODA 2010) Trento, Italy, July 2010.

[9] Johannes Buchmann, Carlos Coronado, Martin D¨oring, Daniela Engelbert,Christoph Ludwig, Raphael Overbeck, Arthur Schmidt ,Ulrich Vollmer, Ralf-Philipp Weinmann , "Post- Quantum Signatures",eprint.iacr.org/2004/297.

[10].Canteaut and N. Sendrier, "Cryptanalysis of the original McEliece Cryptosystem, Advances in Cryptology" - ASIACRYPT '98 Proceedings, Springer-Verlag ,1998, pp.187–199.

[11] "Health Insurance Portability and Accountability Act (HIPAA) and Its Impact on IT Security," Regulatory Compliance Series 3 of 6, Apani Networks White Paper Compliance Series. May 12, 2005. http://www.apani.com.

[12] V.Pless, "Introduction to theory of Error Correcting Codes", Wiley, New York 1982.

[13] Ajay Sharma, Deo Brat Ojha , "Application of coding theory in Fuzzy Commitment Scheme" in Middle-East Journal of Scientific Research 5 (6) 445-448, 2010

[14]Nameer N. EL-Emam, Hiding a Large Amount of Data with High Security Using Steganography Algorithm Applied Computer Science Department, Faculty of Information Technology, Philadelphia University, Jordan

[15]Alain C. Brainos, A Study of Steganography and the Art of Hiding Information, East Carolina University, http://www.infosecwriters.com/text_resources/pdf/steganographyDTEC6823.pdf

[16]Niels Provos, Peter Honeyman, Hide and Seek: Introduction to Steganography, IEEE Security and Privacy, Volume 1, Issue 3 (May 2003), Pages: 32 – 44.