

A Stego-Crypto-Based Technique for Multimedia Transmission

Dr. Emad S. Othman
High Institute for Computers
and Information Systems
Al- Shorouk Academy
Cairo – Egypt
e.mail: emad91@hotmail.com

Dr. Mohammed Sakre
High Institute for Computers
and Information Systems
Al- Shorouk Academy
Cairo – Egypt
e.mail: m_sakre2001@yahoo.com

Abstract

The main goal of this paper is to give new insights and directions on how to improve existing methods of hiding secret messages, possibly by combining steganography and cryptography. We start by a security background describing its history and a short comparison between cryptography and steganography. We then illustrate our approach that helps us achieve a higher level of secrecy and security. In such a way to make it harder for any steganalyst to retrieve the plaintext of a secret message from a stego-object if cryptanalysis were not used.

The MSPC has been exploited to bestow this proposed approach. MSPC relies for its security on the difficulty of factoring large prime numbers. The RSA generator is used as a Stego-key for astonishing randomization of embedding the ciphertext into an innocent image - which is the most popular cover objects used for steganography - to endow with another layer of protection. It is proved that extracting any information about the plaintext is hard for any eavesdropper with computational resources while maintaining the image quality. Cryptography and Steganography procedures are completely reversible at the receiver site.

Key Words: *Cryptography, Steganography, Public-Key, Probabilistic Encryption, Chinese Remainder Theorem, Quadratic Residues, Multimedia, RSA and the Extended Euclidean Algorithm.*

1. Introduction

Cryptography is the study of mathematical techniques related to aspects of information security which used by the Egyptians some 4000 years ago [1] while Steganography is the art and science of invisible communication [2]. Steganography differs from cryptography in the sense that where cryptography focuses on keeping the *contents of a message secret*, *steganography focuses on keeping the existence of a message secret*. Steganography and cryptography are both ways to protect information from unwanted parties but neither technology alone is perfect and can be compromised [3-7]. The strength of steganography can thus be amplified by combining it with cryptography as accomplished in the contribution of this paper.

In this paper a system is developed in which cryptography and steganography are used as integrated part along with newly developed enhanced security model. In cryptography the process of encryption is carried out using asymmetric cipher MSPC (Multimedia Staircase Probabilistic Cryptosystem) to encrypt a message and the obtained ciphertext is hidden in to the cover image which makes the system highly secured [8]. LSB technique is used for message hiding which replaces the least significant bits of pixel selected to the hide the information. The proposed work represents a heuristic approach to introduce the concept of Multi-Layer Data Security algorithm in the field of combined Cryptography and Steganography [9]. The algorithm, which is being proposed here, will secure the text message in multiple protection layers. Hence the concept of Cryptography and Steganography is as two Layers of Security and in between them an extra layer of security is also introduced. The distinction between cryptography and steganography is an important one, and is summarized in table 1.

Table 1: Cryptography VS. Steganography

Cryptography	Steganography
Known message passing	Unknown message passing
Encryption prevents an unauthorized party from discovering the contents of a communication	Steganography prevents discovery of the very existence of communication
Common technology	Little known technology
Most of algorithm known by all	Technology still being developed for certain formats
Strong current algorithms are currently resistant to attack, larger expensive computing power is required for cracking	Once detected message is known
Cryptography alter the structure of the secret message	Steganography does not alter the structure of the secret message

Sections 2, 3 and 4 contain description of the used Cryptography, Steganography algorithms and the proposed system respectively. A sample of the experiments and results is described in section 5 while section 6 stimulates the proposed system. Finally the paper is observed by a conclusion in section 7.

2. Cryptography Algorithm

Multimedia Staircase Probabilistic Cryptosystem (MSPC), presented a study about probabilistic encryption. Blum et al [10-13], presented an unpredictable number generator to produce a long sequence of bits in forward direction called the Blum-Blum-Shub pseudo-random generator (BBS). The implementation of the MSPC was accomplished by following some consequent steps described as follows,

- (1) Starting with the only piece of information available, the two broadcasting composites numbers "the Primary Public-key n_1 and the Secondary Public-key n_2 " whose each factorization is unknown and using the Euclidian's algorithm [14,15], we get the quadratic residues numbers s_j modulo the Public-keys for each (initial seeds s_0 's).
- (2) A multimedia file F , where each character is called x_i , is segmented into M segments. Each of the M seeds, obtained from the above step, initiates the BBS generator to produce unpredictable sequence of bits that are eXclusive-Ored with the T bits (where T is the length of the segment) of the plaintext segment to form the ciphertext block, where each character is called c_i .
- (3) Starting from the encrypted segment 2, each block is eXclusive-Ored with the original first segment, so one can't obtain any information about the plaintext unless reaches successively the top of the stairs.
- (4) For more complexity to any intruder, the last elements (L_1, L_2, \dots, L_M) of the encrypted segments are encrypted in a stair form such that L_1 is concatenated to the file extension x and they are together encrypted using the seed of segment 2 and the output is denoted y_1 . The second last number L_2 is concatenated to y_1 and they are together encrypted using the seed of segment 3 to form y_2 and so on. The encryption of L_M and y_{M-1} is done using a new seed established from the secondary public-key.

The above method makes an eavesdropper to be confused if trying to attack any part of the ciphertext. Quadratic residue numbers are selected using a pseudo random function to get completely random seeds. Hiding the multimedia file extension into the encrypted data was very essential, so one gets doubtful to know the file formats.

$3M$ encryption levels were applied, $2M$ levels were done on the plaintext and the third M levels had been obtained when the last elements were encrypted individually.

An extra encryption level was achieved when the file extension is encrypted. Since there are "many" possible encryptions of each plaintext, it is not feasible to test whether a given ciphertext is an encryption of a particular plaintext as shown in figure: 1.

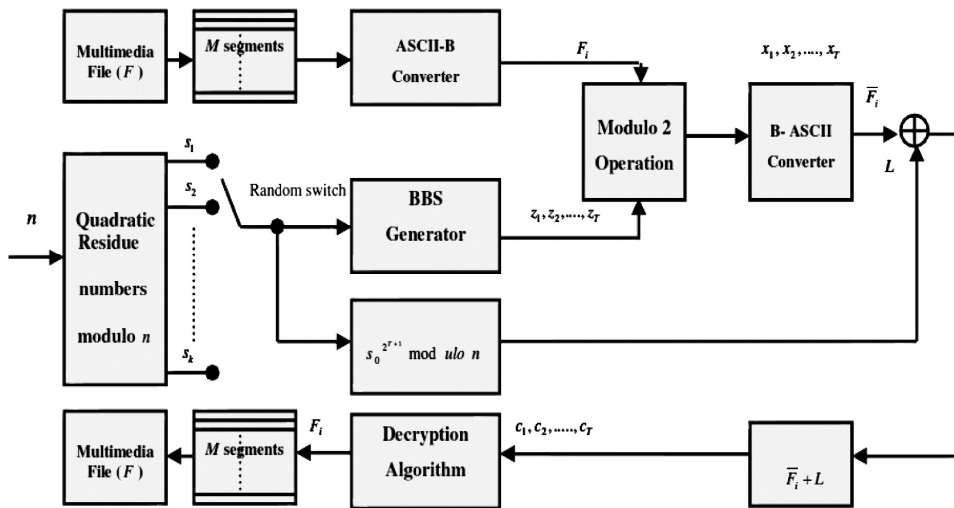


Figure 1: The MSPC Encryption Algorithm Block Diagram

The mathematical structure of MSPC can be implemented using the following formulas:

- (1) Compute the key stream z_1, z_2, \dots, z_T from initial seed s_0 using the BBS Generator.
- (2) Compute $s_{T+1} = s_0^{2^{T+1}} \bmod n$, where n could be n_1 or n_2 .
- (3) Compute $c_i = (x_i + z_i) \bmod 2$ for $1 \leq i \leq T$.
- (4) The cipher text can be defined as : $c = (c_1, c_2, \dots, c_T, s_{T+1})$.

After the encrypted file is being sent to the owner of the public-keys, she/he is the only person who is capable to decrypt this file and her/his ultimate goal is to obtain which initial seeds had been selected during the encryption procedure as shown in figure 2.

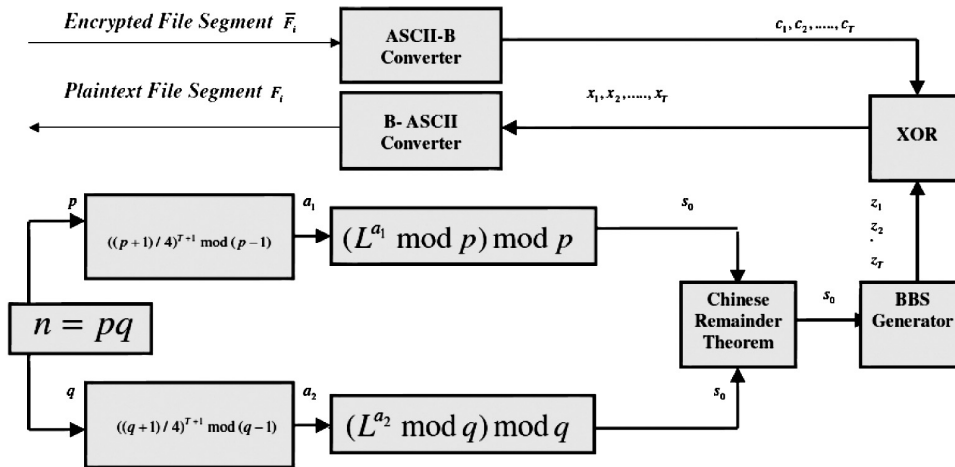


Figure 2: The MSPC Decryption Algorithm Block Diagram

To decrypt, one must perform the following sequence of steps backwards correctly to reconstruct the original plaintext:

- (1) Compute $a_1 = ((p+1)/4)^{T+1} \bmod (p-1)$,
 $a_2 = ((q+1)/4)^{T+1} \bmod (q-1)$, and $n = pq$, where p and q are the largest prime odd integer numbers.
- (2) Compute $b_1 = s_{T+1}^{a_1} \bmod p$, and $b_2 = s_{T+1}^{a_2} \bmod q$.

- (3) Using the Chinese remainder theorem to solve this system of congruence and discover the elected initial seed s_0 :

$$\{s_0 = b_1 \bmod p \text{ and } s_0 = b_2 \bmod q\}.$$
- (4) Using the obtained initial seed s_0 to compute the key stream z_1, z_2, \dots, z_T (BBS Generator).
- (5) To get plaintext $x = (x_1, x_2, \dots, x_T)$, compute $x_i = (c_i + z_i) \bmod 2$ for $1 \leq i \leq T$.

3. Steganography Algorithm and the Collision Problem

Usually the secret information is concealed by the use of an innocuous cover as to not arouse suspicion if hostile agents discover the cover. As an example, the cover text:

I'm feeling really stuffy. Emad's medicine wasn't
strong enough without another febrifuge.

hides the sentence ***“Meet me at nine”*** if the reader retains the second letter of each word in sequence. Embedding data, which is to be hidden, into an image requires two files. The first is the innocent-looking image that will hold the hidden information, called the cover image [16]. The second file is the ciphertext as a bit stream. When combined, the cover image and the embedded encrypted message make a stego image as shown in figure3.

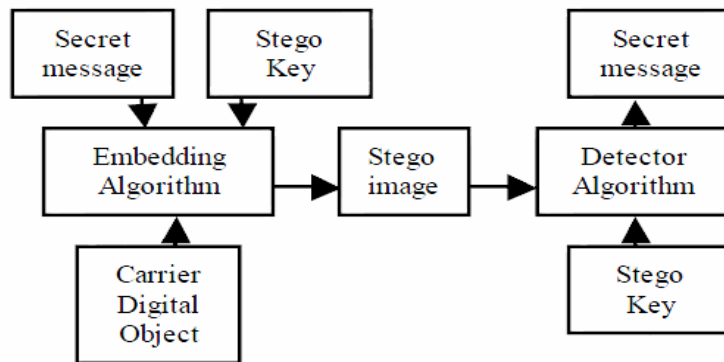


Figure 3: The Model of Steganography

The RSA Generator, which is based on the fundamental problems of factorization (as it relates to the RSA public-key cryptosystem) and the Discrete Logarithm problem, is used to get pseudo-random selection pixel to embed the ciphertext bits into a pseudo randomly selected subset of the cover image. The message bits are then embedded into the elements using the Least significant bit (LSB) technique. Pseudo-random selection rules typically offer better security than sequential rules [17].

RSA Generator is based on the presumed difficulty of factoring large integers, the factoring two large prime numbers problem. RSA stands for Ron Rivest, Adi Shamir and Leonard Adelman [18], who first publicly described it in 1978. **The mathematical structure of RSA Generator can be implemented using the following theory:**

Let p, q to be two $(k/2)$ bit primes, and define $n = pq$. Let b be chosen such that $\text{GCD}(b, \phi(n)) = 1$. As in RSA Cryptosystem, n and b are public while p and q are secret.

A seed s_0 is any element of \mathbb{Z}_n^*

So s_0 has k bits. For $i \geq 1$. Define $s_{i+1} = s_i^b \text{ mod } n$,

And then define $f(s_0) = (z_1, z_2, \dots, z_l)$

Where $z_i = s_i \text{ mod } 2, 1 \leq i \leq l$.

Then f is a (k, l) - RSA Generator.

Many steganography experts recommend using images featuring 256 shades of gray. Gray-scale images are preferred because the shades change very gradually from byte to byte, and the less the value changes between palette entries, the better they can hide information. Least significant bit (LSB) insertion is a common, simple approach to embedding information in a cover image [19]. The least significant bit (in other words, the 8th bit) of some or all of the bytes inside an image is changed to a bit of the secret message. One can store up to 128 KB message in 1024×1024 Gray-Scale Picture.

On average, only half of the bits in an image will need to be modified to hide a secret message using the maximum cover size [20]. Since there are 256 possible intensities of each primary color, changing the LSB of a pixel results in

small changes in the intensity of the colors. These changes cannot be perceived by the human eye - thus the message is successfully hidden. With a well-chosen image, one can even hide the message in the least as well as second to least significant bit and still not see the difference.

The RSA generator seeded with a value known to both sender and receiver is used to randomly select the pixels which the data will be embedded in where q is the grayscale value of the pixel as follows:

Let M denote a binary encrypted secret message sequence, $M = \{m_i | m_i \in \{0, 1\}, i = 0, 1, \dots, t-1\}$, where t is the message length.

Let $f_I(x, y)$ denote the grayscale value at (x, y) in cover-image I , and

Let $LSB_I(x, y)$ denote the LSB of the grayscale value at (x, y) .

The embedding algorithm is as follows:

Input: cover-image I , binary encrypted message sequence M .

Output: stego-image I' .

Step 1:

Set $I' = I$.

Step 2:

Use RSA- PRNG to randomly select t pixels from I' .

Let (x_i, y_i) denote the coordinate of the selected pixel. $i = 0, 1, \dots, t-1$.

Step 3:

Let $q_i = f_I(x_i, y_i)$ denote the grayscale value of pixel (x_i, y_i) .

Let m_i denote the message-bit to be embedded in pixel (x_i, y_i) .

For all randomly selected pixels (x_i, y_i) ,

if $LSB_{I'}(x_i, y_i) = m_i$,

Do nothing;

if $LSB_{I'}(x_i, y_i) \neq m_i$,

Force change with the new value;

Step 4:

Output I'

So, in short, RSA pseudo randomly generator is used to choose randomly selected pixels of the cover image to embed the encrypted message bits inside it as shown. Then the receiver can reconstruct the message by repeating the RSA pseudo randomly selected pixels and extracting the LSBs of pixel values directly.

This steganography embedding scheme does not consider collision of the selected and used pixels since the operated RSA generator which used for selecting pixels is deterministic random bit generator (DRBG), and will not repeat any of the selected pixels once again. Therefore collision problem of reusing the same pixel in the cover image more than once is avoided and sidestepped.

RSA generator produces random numbers within (truncated to) the dimensions of the cover image to be used for the pixels injection, so these selected pixels are saved (kept) in a temporary indexed-table. A quick comparison will be accomplished between the new generated number and the old contents of that indexed-table and no collision occurred. Some experiments were performed to enforce embedding an encrypted message that can occupy the whole pixels of the cover image and collision didn't happen. But just in a bad case if the new generated number is repeated, the RSA generator will reproduce a new number and use the comparison mentioned above more and more to ensure that there is no same selected pixels collision.

Cryptographic applications require the output to be unpredictable, and more elaborate designs, which do not inherit the linearity of simpler solutions, are needed. More recent instances of PRNGs with strong randomness guarantees are based on computational hardness assumptions, and include the Blum Blum Shub, Fortuna, and Mersenne Twister algorithms [21].

4. Proposed System

In this proposed system we have the software for data encryption and then embed the outputted ciphertext in an image with help of stego key (RSA Generator). This algorithm combines the effect of these two methods to enhance the security of the data. The proposed algorithm encrypts the data with the MSPC crypto algorithm and then embeds the encrypted text in an image file.

This algorithm improves the security of the data by embedding the encrypted text and not the plain text in an image. This combined chemistry will satisfy the requirements such as capacity, security and robustness for secure data transmission over an insecure communication channel [22-24].

A pictorial representation of the combined concept of cryptography and steganography is depicted in the figure 4.

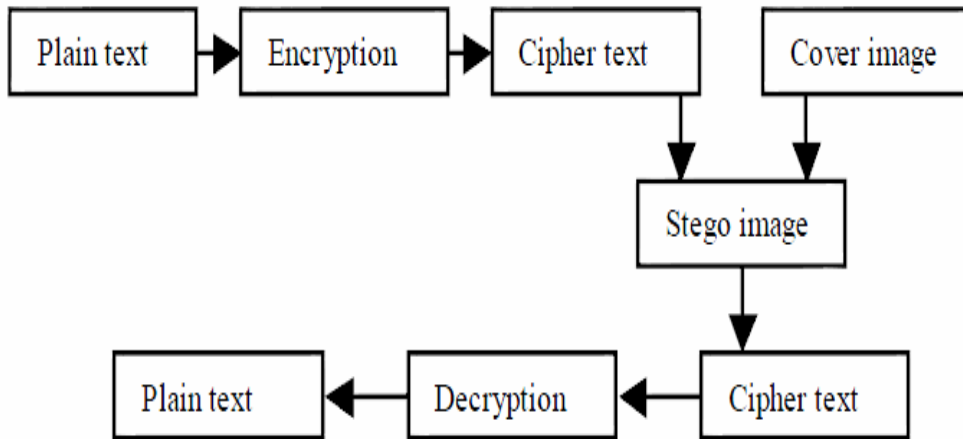


Figure 4: Block-Diagram of the Proposed Technique

If an attacker were to defeat the steganographic technique to detect the message from the stego-object, he would still require the cryptographic decoding key to decipher the encrypted message.

5. Experiments and Results

In this section, some experiments are carried out to prove the efficiency of the proposed scheme. The proposed method has been simulated using the MATLAB 7 program on Windows 7 platform. The proposed high secured system using cryptography is tested by taking the encrypted message and hiding it into the **Toronto city** image. The result thus obtained from the experiment is recorded and is being summarized in the following figures. A sample of plaintext and its corresponding ciphertext after applying the MSPC cryptosystem is shown in Table 2 while its histogram is shown in Figure 5.

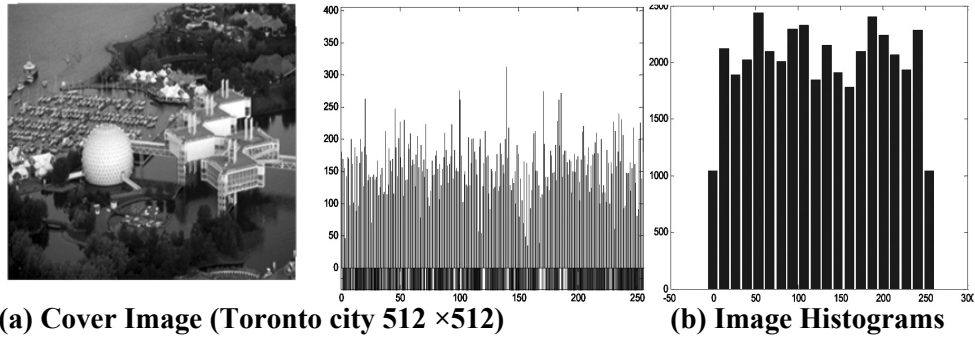


Figure 6 : Cover Image and its histograms

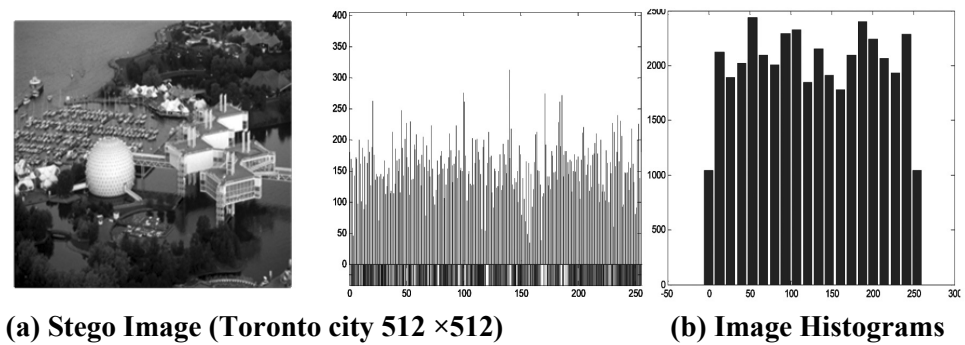


Figure 7: Stego Image and its Histograms

Modulating the least-significant bit does not result in a human-perceptible difference because the amplitude of the change is small. An advantage of the proposed scheme is that the extracting algorithm is simple and easy to implement. When receiving a stego-image, the receiver uses the same stego-keys to generate the same stego-tables and table indices (using the RSA generator) as those used in the embedding process to extract the originally encrypted data. Then MSPC decryption procedure is applied to obtain the plaintext. There are infinite number of steganography applications for digital image including copyright protection, feature tagging, and secret communication. This paper explores a tiny fraction of the art of steganography. It goes well beyond simply embedding text into an image.

6. Generalized Chi-Square Attack

Andreas Pfitzmann and Andreas Westfeld [25] introduced a method based on statistical analysis of Pair of Values (PoVs) that are exchanged during sequential embedding. This attack works on any sequential embedding type of

stego-system such as EzStego and Jsteg. The chi-square attack can be applied to any steganographic technique in which a fixed set of Pairs of Values (PoVs), or other fixed groups of values, are flipped into each other to embed message bits. For example, the PoVs can be formed by palette indices that differ in their LSBs. Before embedding, in the cover image the two values from each pair are distributed unevenly. After message embedding, the occurrences of the values in each pair will have a tendency to equalize (this depends on the message length).

Since swapping one value into another does not change the sum of occurrences of both indices in the image, we can test for the statistical significance of the fact that the occurrences of both values in each pair are the same. If, in addition to that, the stego-technique embeds message bits sequentially into subsequent pixels/indices/coefficients starting, for example, in the upper left corner, we will observe an abrupt change in the statistical evidence as we encounter the end of the message.

For a sequentially embedded message, this test enables us not only to determine with a very high probability that a message has been embedded, but also calculate its length. If the message-carrying pixels in the image are selected randomly rather than sequentially, this test becomes less effective unless majority of pixels (i.e., more than 97%) have been used for embedding. This attack though, does not work for pseudo-random type of embedding as in the presented model and the model is invulnerable.

The larger Peak Signal to Noise Ratio (PSNR) indicates the higher the image quality i.e. there is only little difference between the cover-image and the stego-image. On the other hand, a smaller PSNR means there is huge distortion between the cover-image and the stego image, where the measurement of the quality between the cover image f and stego-image g of sizes $N \times N$ is defined as:

$$PSNR = 10 \text{ Log} \left(\frac{255^2}{MSE} \right)$$

$$\text{Where } MSE = \frac{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} (f(x,y) - g(x,y))^2}{N \times N}$$

Where $f(x,y)$ and $g(x,y)$ means the pixel value at the position (x, y) in the cover-image and the corresponding stego-image respectively. The PSNR is expressed in dB's.

7. How Secure is the Proposed System?

The proposed system is approvingly secure since:

- a- It's a combination of two highly secured techniques
 - i. MSPC for cryptography and
 - ii. LSB manipulation for Steganography.
- b- Number of Keys: This system contains total of 6 keys
 - i. Two large public keys for MSPC Encryption algorithm.
 - ii. One seed used by the BBS Generator.
 - iii. Another one seed used by the RSA Generator.
 - iv. Two extra private keys for MSPC Encryption algorithm and retrieving the original message.

These two extra private keys make the system highly secured.

As a Stego-Analysis the Chi-Square Attack which based on statistical analysis of Pair of Values (PoVs) does not work for pseudo-random type of embedding as in the presented model so the model is invulnerable against the Chi-Square Attack.

8. Conclusions and Future Works

Steganography, especially combined with cryptography, is a powerful tool which enables people to communicate without possible eavesdroppers even knowing there is a form of communication in the first place. The proposed method provides acceptable image quality without any distortion in the image. If an intruder detect the partial part of the hidden message from the stego image it will be totally meaningless for him and moreover until the complete set of keys are available getting the original message is impossible.

Encrypting the payload is not always done solely to make recovery of the payload more difficult. The proposed cipher has the desirable property of making the payload appear indistinguishable from uniformly-distributed noise, which can make detection efforts more difficult, and save the stenographic encoding technique the trouble of having to distribute the signal energy evenly.

Using the RSA generator for selecting pixels is deterministic random bit generator (DRBG) so the collision problem of reusing the same pixel in the cover image more than once is avoided and sidestepped.

By this, we are going to achieve two important goals:

- a) Make the detection of message harder to gain to stricter security.

- b) Create a stego-key for embedding encrypted message.

Development in covert communications and steganography will continue, as will research in building more robust algorithms. The more information that is made available on the Internet, the more owners of such information need to protect themselves from theft and false representation. Steganography, like cryptography, will play an increasing role in the future of secure communication in the “digital world.”

References

- [1] Alfred J. Menezes, Paul C. van Orschot and Scott A. Vanstone, *Hand book of Applied Cryptography*, Page # 1, CRC Press, 2009.
- [2] William Stallings; *Cryptography and Network Security: Principals and Practice*, Prentice Hall international, Inc.; 2010.
- [3] Douglas R. Stinson, *Cryptography Theory and Practice*, CRC Press, 1995.
- [4] Shafi Goldwasser and Silvio Micali, *Probabilistic Encryption*, Journal of Computer and System Sciences 28, PP. 270-299, 2008.
- [5] Emad Osman , *Digital Image Steganography Based on EHMC "Data Sneaking Between Pixels"* Proceedings of the 1st International Conference on Computer Science from Algorithms to Applications (CSAA), JW Marriot Mirage City, Cairo, EGYPT, December 8-10, 2009.
- [6] Jessica Fridrich et al, *Digital Watermarking and Steganography*, Morgan Kaufmann publications, 2008.
- [7] Dorothy Elizabeth Robling Denning, *Cryptography and Data Security*, Addison-Wesley Publishing Company, 2010.
- [8] Osman et al, “*Multimedia Staircase Probabilistic Cryptosystem (MSPC)*”, ICAIA’ 99, proceedings of the 7th International Conference on Artificial intelligence & its Applications, pp. 294 – 298, Cairo, February 1999.
- [9] Lorrie Cranor, Simson Garfinkel, *Security and Usability*, O'Reilly Publisher, 2005.
- [10] Fred Piper and Sean Murphy, *Cryptography: A Very Short Introduction*, Oxford University Press, 2002.
- [11] Neha Sharma, J.S. Bhatia and Dr. Neena Gupta, “ *An Encrypto-Stego Technique Based secure data Transmission System*”, PEC, Chandigarh, 2010.
- [12] I. Venkata Sai Manoj, “*Cryptography and Steganography*”, International Journal of Computer Applications (0975 – 8887), Volume 1 – No.12, 2011.

- [13] Alan Siper, Roger Farley and Craig Lombardo, “*The Rise of Steganography*”, Proceedings of Student/Faculty Research Day, CSIS, Pace University, May 6th, 2005.
- [14] B. B Zaidan, A.A Zaidan, A.K. Al-Frajat and H.A. Jalab, “*On the Differences between Hiding Information and Cryptography Techniques: An Overview*”, Journal of Applied Sciences 10(15): 1650-1655, 2010
- [15] Domenico Bloisi and Luca Iocchi, “*Image Based Steganography and Cryptography*”, Sapienza University of Rome, Italy, 2010.
- [16] Kallam Ravindra Babu, Dr. S.Udaya Kumar, Dr. A.Vinaya Babu, “*A Survey on Cryptography and Steganography Methods for Information Security*”, International Journal of Computer Applications(0975-8887), Volume 12 – No. 2, November 2010.
- [17] Dipti Kapoor Sarmah, Neha bajpai, “*Proposed System for Data Hiding Using Cryptography and Steganography*”, International Journal of Computer Applications (0975 – 8887), Volume 8 – No. 9, October 2010.
- [18] Eiji Kawaguchi and Richard O. Eason, “*Principle and applications of BPCS-Steganography*”, Kyushu Institute of Technology, Kitakyushu, Japan, University of Maine, Orono, Maine 04469-5708.
- [19] Ross J. Anderson, Fabien A.P. Petitcolas, “*On The Limits of Steganography*”, IEEE Journal of Selected Areas in Communications, 16(4): 474-481, May 1998. Special Issue on Copyright & Privacy Protection. ISSN 0733-8716.
- [20] Sashikala Channalli and Ajay Jadhav, “*Steganography An Art of Hiding Data*”, International Journal on Computer Science and Engineering Vol.1(3), 2009, 137.
- [21] M. Fadhil, “*A Novel Steganography-Cryptography System*”, Proceedings of the World Congress on Engineering and Computer Science 2010 Vol I WCECS 2010, October 20-22, 2010, San Francisco, USA
- [22] <http://www.xdp.it>
- [23] http://en.wikipedia.org/wiki/Pseudorandom_number_generator
- [24] <http://www.codeproject.com/KB/library/ArisFFTDFTLibrary.aspx>
- [25] Westfeld, A. and Pfitzmann, A. “*Attacks on Steganographic Systems*”, 3rd International Workshop. Lecture Notes in Computer Science, Vol.1768. Springer-Verlag, Berlin Heidelberg New York, 2000.