International Journal of Engineering, Applied and Management Sciences Paradigms
Volume 51, Issue 01, Quarter 04 (2017)
An Indexed and Referred Journal
ISSN (Online): 2320-6608
www.ijeam.com

Three Layer Architecture for DNA based Image Steganography

Garima Sharma[1] and Tarun Dalal[2]

[1,2]Department of Computer Science and Engineering
CBS Group of colleges, Jhajjar, Haryana (India)
[1]garima.dav.09@gmail.com

Publishing Date: October 16, 2017

Abstract

This paper discusses 3-layer architecture for the secure transmission of a message. The input message is compressed by using the dictionary based LZW compression. Then the compressed data is encrypted by light weight encryption technique i.e. one time pad. Then the encrypted message is coded as DNA codon by using DNA cryptography. The DNA codon of approximation part of the image is generated and the message DNA codon is embedded into image codon. This technique shows the high capacity and the high security analyzed over various images with different messages.

Keywords*: LZW, OTP, DNA, DWT.*

1. Introduction

The growing use of Internet among public masses and availability of public and private digital data and its sharing has driven industry professionals and researchers to pay a particular attention to information security. Internet users frequently need to store, send, or receive private information and this private information needs to be protected against unauthorized access and attacks. Presently, three main methods of information security being used: watermarking, cryptography and steganography.[1] In watermarking, data are hidden to convey some information about the cover medium such as ownership and copyright. Cryptography techniques are based on rendering the content of a message garbled to unauthorized people. Steganography techniques are based on hiding the existence of information by embedding the secret message in another cover medium. While all three are information security techniques cryptography and steganography are having wide application as watermarking is limited to having information particularly about the cover medium. DNA stands for Deoxyribo Nucleic Acid. DNA represents the genetic blueprint of living creatures. DNA contains "instructions" for assembling cells. Every cell in human body has a complete set of DNA. DNA is unique for each individual. DNA is a polymer made of monomers called deoxyribo nucleotides [1].

2. CODON

A codon is a triplet of three bases (T, A, C, G). With these four letters, 4^3= 64 combination are possible. With three exception, TAA, TAG, TGA, indicate codons STOP[1].

2.1 Operation for DNA

Addition and Subtraction operation can be performed on DNA sequences according to traditional addition and subtraction in the Z2 (mod 2). For example, 11+10 = 01, so of (C, A , T ,G)= (0 , 1 , 2 , 3), we have G+T = A. Table 1 & 2 shows addition and subtraction.

Table 1: Addition Operation for DNA Sequence

+	T	A	C	G
T	C	G	T	A
A	G	C	A	T
C	T	A	C	G
G	A	T	G	C

2.2 Degenerative CODON

When two or more codons (synonymous codons) code for the same amino acid they are calls degenerative codons. They typically differ in their 3rd base (Ex: CCT, CCA for cytosine). So by replacing third base, we can hide data without affecting the resulting amino acid. Figure shows the mapping of codon to amino acid with yellow ones being the degenerative codons.

Table 2: The Mapping of CODON to Amino Acid

	T		C		A		G		
T	TTT	Phe	TCT	Ser	TAT	Tyr	TGT	Cys	T
	TTC	Phe	TCC	Ser	TAC	Tyr	TGC	Cys	C
	TTA	Leu	TCA	Ser	TAA	Stop	TGA	Cys	A
	TTG	Leu	TCG	Ser	TAG	Stop	TGG	Trp	G
C	CTT	Leu	CCT	Pro	CAT	His	CGT	Arg	T
	CTC	Leu	CCC	Pro	CAT	His	CGC	Arg	C
	CTA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	A
	CTG	Leu	CCG	Pro	CAG	Gln	CGG	Arg	G
A	ATT	Ile	ACT	Thr	AAT	Asn	AGT	Ser	T
	ATC	Ile	ACC	Thr	AAC	Asn	AGC	Ser	C
	ATA	Ile	ACA	Thr	AAA	Lys	AGA	Arg	A
	ATG	Met	ACG	Thr	AAG	Lys	AGG	Arg	G
G	GTT	Val	GCT	Ala	GAT	Asp	GGT	Gly	T
	GTC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C
	GTA	Val	GCA	Ala	GAA	Glu	GGA	Gly	A
	GTG	Val	GCG	Ala	GAG	Glu	GGG	Gly	G

DNA cryptography, a new branch of cryptography utilizes DNA as an informational and computational carrier with the aid of molecular techniques. It is relatively a new field which emerged after the disclosure of computational ability of DNA [5]. DNA cryptography gains attention due to the vast storage capacity of DNA, which is the basic computational tool of this field. One gram of DNA is known to store about 108 tera-bytes. This surpasses the storage capacity of any electrical, optical or magnetic storage medium [4], [5]. Traditional cryptographic systems have long legacy and are built on a strong mathematical and theoretical basis. Traditional security systems like RSA, DES or NTRU are also found in real time operations. So, an important perception needs to be developed that the DNA cryptography is not to negate the tradition, but to create a bridge between existing and new technology. The power of DNA computing will strengthen the existing security system by opening up a new possibility of a hybrid cryptographic system.

3. LZW Compression

LZW compression is named after its developers, *A. Lempel and J. Ziv*, with later modifications by *Terry A. Welch*. It is the [2] foremost technique for general purpose data compression due to its simplicity and versatility. Typically, it can be expected to compress text, executable code, and similar data files to about one-half their original size. LZW also performs well when presented with extremely redundant data files, such as tabulated numbers, computer source code, and acquired signals. Compression ratios of 5:1 are common for these cases. LZW is the basis of several personal computer utilities that claim to "double the capacity of our hard drive."

When the LZW program starts to encode a file, the code table contains only the first 256 entries, with the remainder of the table being blank. This means that the first codes going into the compressed file are simply the single bytes from the input file being converted to 12 bits. As the encoding continues, the LZW algorithm identifies repeated sequences in the data, and adds them to the code table. Compression starts the second time a sequence is encountered. The key point is that a sequence from the input file is not added to the code table until it has already been placed in the compressed file as individual characters (codes 0 to 255). This is important because it allows the decompression program to reconstruct the code table directly from the compressed data, without having to transmit the code table separately.

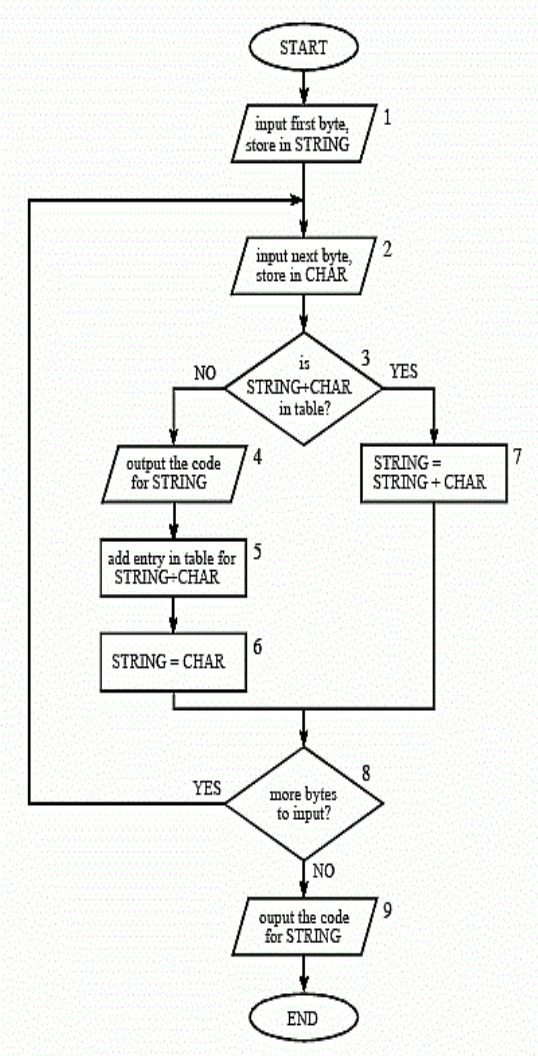

Figure 1: Flow chart for LZW Compression algorithm

4. One Time Pad

One Time Pad encryption is a very simple, yet completely unbreakable cipher method. The One Time Pad encryption method is a binary additive stream cipher, where a stream of truly random keys is generated and then combined with the plain text for encryption or with the ciphertext for decryption by an 'exclusive OR' (XOR) addition. It is possible to prove that a stream cipher encryption scheme is unbreakable if the following preconditions are met:

- The key must be as long as the plain text.
- The key must be truly random.
- The key must only be used once.

The steps followed in One Time Pad encryption are as follows:

- To encrypt plain text data, the sender uses a key string equal in length to the plain text. The key is used by mixing (XOR-ing) bit by bit, always a bit of the key with a bit of the plain text to create a bit of cipher text.
- This cipher text is then sent to the recipient.
- At the recipient's end, the encoded message is mixed (XOR-ed) with the duplicate copy of the One Time Key and the plain text is restored.
- Both sender's and recipient's keys are automatically destroyed after use, to ensure re-application of the same key is not possible.

M: $\{0, 1\}^l$, where ℓ is the message length.
K : $\{0, 1\}^l$.

$c = Enc(k,m) = m \oplus k$, for $m \in M$, $k \in K$, where "$\oplus$" stands for a bit-wise xor

$m = Dec(k, c) = c \oplus k$, for $m \in M$, $k \in K$.

Decryption works because $(c \oplus k) = ((m \oplus k) \oplus k) = m \oplus (k \oplus k) = m$

5. Existing Work

The system performs the encryption by using the images as primary (main) cover media for transferring of message from one to another end. The DNA strand is extracted from the image that acts as secondary cover media. The data is hided by selecting the image pixels. The 2 bit of data is hided into one single pixel of image. The process can be explained by using the following algorithm:

Existing Algorithm

1. Input image & cover image

2. Check capacity with a set of 2d map parameter & same cover. Generate ssDNA
3. Convert encrypted message(m) byte to 8 bit binary
4. Select next 2 bit & encode to nucleotide
5. Choose next degenerative codon from ssDNA and replace 3rd base with nucleotide
6. Perform dna addition with primer.
7. Resulted m is then embed behind the cover image using LSB
8. Stego image with same quality is obtained.

The data can be easily embedded to the image by using the above the algorithm. Then this cover media can be easily transmitted. At the receiver end the message can be extracted by using the reverse procedure only if users have the parameters and key.

6. Proposed System

The existing algorithms performs the encryption by using the images as primary(main) cover media for transferring of message from one to another end. The DNA strand is extracted from the image that acts as secondary cover media. The data is hided by selecting the image pixels. The 2 bit of data is hided into one single pixel of image. The present work enhances the message capacity of the existing technique. In this technique the input data is firstly compressed by using the LZW compression techniques i.e. a dictionary based compression of data. It reduces the size of data depending upon the repetition in the data. The compressed data is encrypted by using a light weight encryption method. Then the image pixel is selected by using the DNA strand to hide the data into mage by using DNA. The whole process can be easily understood by using the following algorithm:

Proposed Algorithm

1. Input image & cover image
2. Take DWT of image to decompose the image into LL, LH , HL , HH
3. Apply LZW compression technique to compress the input message(m).
4. Apply modified OTP encryption technique to encrypt the compress m.
5. Check capacity with a set of 2d map parameter & same cover. Generate ssDNA
6. Covert encrypted m byte to 8 bit binary
7. Select next 2 bit & encode to nucleotide
8. Choose next degenerative codon from ssDNA and replace 3rd base with nucleotide
9. Perform dna addition with primer.
10. Resulted msg is then embed behind the cover image
11. Stego image with better quality is obtained.
12. Generate Setgo image.

The above algorithm cab be used to embed the message. The extraction is just reverse process used at receiver end. The receiver can decrypt the message only if corresponding parameters and key is available. The above algorithm cab be implemented by using the MATLAB discussed in next section.

7. Implementation

The proposed work is implemented using the MATLAB. The work is done by using the script file; no toolbox is used for the work. The implementation occurs as follow:

1. Input image

Figure 2: Input Image

2. Take DWT of image to decompose the Image

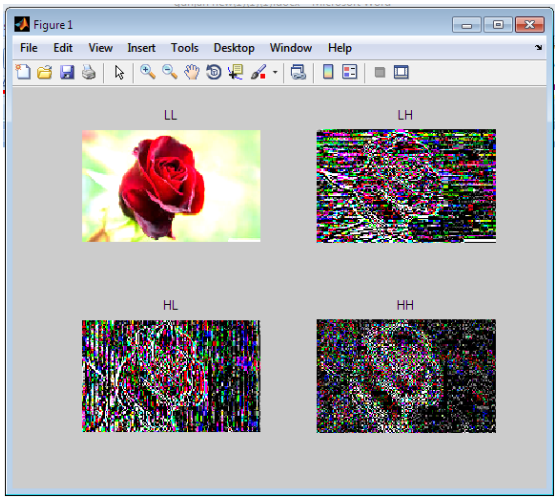

Figure 3: DWT of Input Image

3. Input Text
Msg=' hellohellobye'

4. The message in integer form:

msg =
Columns 1 through 11
104 101 108 108 111 104 101 108 108 111 98
Columns 12 through 13
121 101

5. The message after compression

msg =

102 109 109 112 257 259 112 99 122 102

6. The message after The encryption is

msg= 154 150 157 158 160 2 4 160 147 171 150

7. Generated Codon

GCTGAAAAAAAGCCCTGACCGTAACAT GCACTCCATGCAACG
ACAAC

8. Image After Hiding the data

Figure 4: Steago Image

The results of the above implementation are discussed in next section.

8. Results and Discussion

This section describes the results and comparison of the results with other existing system by using various parameters. The work encrypt the compressed data by using the DNA and the hides this encrypted data in to the image. The parameters evaluated are PSNR, MSE and the total percentage of bit change.

PSNR and MSE are defined as follows :

$$\text{PSNR} = 10\log_{10}\frac{255^2}{\text{MSE}} \quad \text{and} \quad \text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(I_m(i) - I_s(i)\right)$$

Where I_m and I_s are the original and stego image, respectively, n is the number of pixels. Higher the PSNR means better image quality.

The results obtained compared graphically on different images and different messages. The graphical comparison is show in the figure 5, 6, 7, 8:

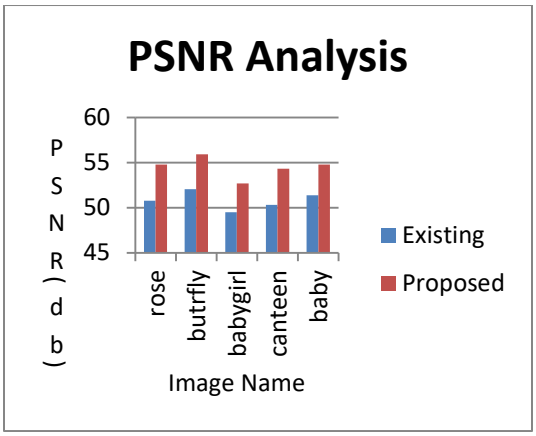

Figure 5: PSNR Analysis on Different Images

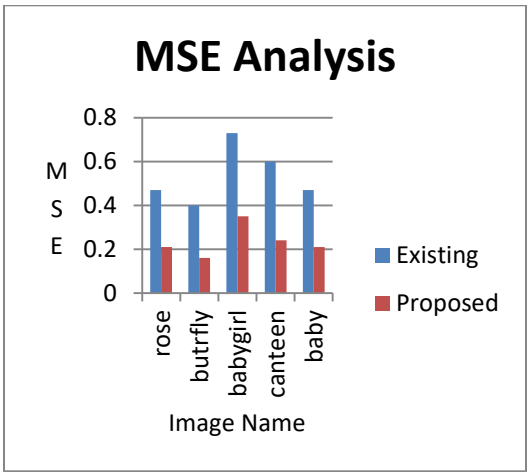

Figure 6: MSE Analysis on Different Images

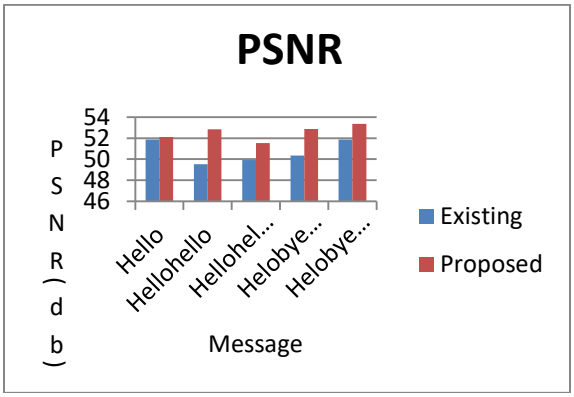

Figure 7: PSNR Analysis on Different Message

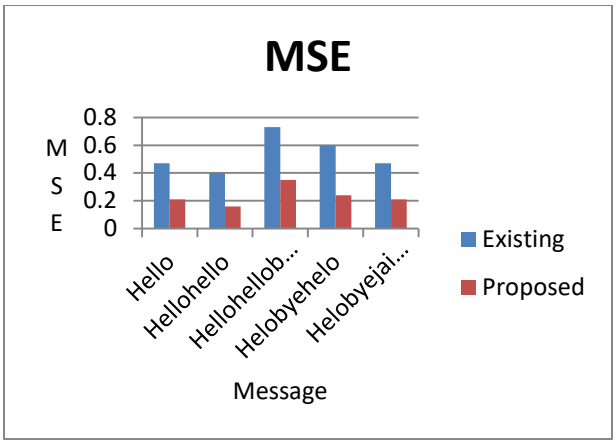

Figure 8: PSNR Analysis on Different Message

The result comparison clearly shows that the amount of bit change reduced drastically. This signifies the importance of the proposed technique. The higher values of the PSNR also show the effectiveness of the technique.

9. Conclusion and Future Scope

DNA cryptography is relatively a new field which emerged after the disclosure of computational ability of DNA. DNA cryptography gains attention due to the vast storage capacity of DNA, which is the basic computational tool of this field. The existing algorithms performs the encryption by using the images as primary(main) cover media for transferring of message from one to another end. The DNA strand is extracted from the image that acts as secondary cover media. The data is hided by selecting the image pixels. The 2 bit of data is hided into one single pixel of image. The present work enhances the message capacity of the existing technique. In this technique the input data is firstly compressed by using the LZW compression techniques i.e. a dictionary based compression of data. It reduces the size of data depending upon the repetition in the data. The compressed data is encrypted by using a light weight encryption method. Then the image pixel is selected by using the DNA strand to hide the data into mage by using DNA. The result comparison clearly shows that the significance of the proposed technique. The higher values of the PSNR also show the effectiveness of the technique. In future following

work can be done: The work can be extended to use other medium to hide the DNA.

References

[1] Das, P., Deb, S., Kar, N., & Bhattacharya, B. (2015). An Improved DNA Based Dual Cover Steganography. *Procedia Computer Science*, *46*, 604-611.

[2] Prof. Bandyopadhyay S K , Banik B G, "LSB Modification and Phase Encoding Technique of Audio Steganography Revisited". International Journal of Advanced Research in Computer and Communication Engineering Vol. 1, Issue 4, 2012

[3] Kajal, I., Rohil, H., & Kajal, A. LZW based Image Steganography using Kekre's Algorithm. (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 2643-2648

[4] Rajyaguru M H, "CRYSTOGRAPHY-Combination of Cryptography and Steganography with Rapidly Changing Keys". International Journal of Emerging Technology and Advanced ISSN 2250-2459, Volume 2, Issue 10, 2012.

[5] Divya S S and Reddy M R M, "Hiding Text in Audio Using Multiple Lsb Steganography and Provide Security Using Cryptography". International Journal of Scientific & Technology Research Volume 1, Issue 6, 2012.