www.ijecs.in

International Journal Of Engineering And Computer Science ISSN:2319-7242

Volume 4 Issue 7 July 2015, Page No. 13017-13021

Secure Transmission of Data Using DNA Technology

Soumya V. Vastrad[1]

[1]Dept of CSE, B.L.D.E.A's CET, Vijayapur.
svvastrad5@gmail.com

Abstract*: The traditional encryption techniques use mathematical and theoretical concepts. With the advances in technology, the intruders are able to decrypt the private data and steal the secretive information easily. Thus cryptography using DNA technology can be viewed as a new hope for unbreakable algorithms. The text data is encrypted using DNA technology. A key is generated and encrypted. The encrypted text and key are then hidden in an image by applying Discrete Cosine Transform (DCT) or Discrete Wavelet Transform (DWT) to generate a DCT watermarked image or a DWT watermarked image. The watermarked image is transmitted to the receiver. Encryption followed by steganography further enhances the security of the data being transmitted. The proposed algorithm is evaluated using correlation coefficient, PSNR, MSE and embedding capacity.*

Keywords: DNA based cryptography, Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), steganography, watermarked image.

.

1. Introduction

Internet serves as an important means of communication. It is being increasingly used as an important tool for communication in the fields of medicine and e-commerce. Now a days many communication channels are being intruded by intruders and the attackers. Hence ensuring security has become very challenging. The cryptographic methods can be used in general to ensure security of the data to be transmitted. Steganography can also be used for the sake of security of the data to be transmitted. Steganography can be defined as the process of hiding information by embedding the information within the other seemingly harmless information. It works by replacing the bits of useless or unused data in regular computer files (such as images, graphics, sound, text, HTML) with bits of different, invisible information. This hidden information can be plain text, cipher text or even images. Steganography sometimes is used when the encryption is not permitted. The steganography is also used to supplement encryption process. Cryptography can be defined as a method of transmitting and storing data in a form that is only readable by the person for whom it is intended. The main objectives of cryptography are to maintain confidentiality, integrity and authentication of the data.

DNA stands for Deoxyribo Nucleic Acid. DNA is a polymer, which is composed of numerous monomers called deoxyribo nucleotides. Each nucleotide basically constitutes of 3 components: deoxyribose sugar, phosphate group and a nitrogenous base. There are 4 nitrogenous bases present. They are: Adenine (A), Guanine (G), Cytosine (C) and Thymine (T). The key characteristic of the structure of the DNA is its inherent complementary feature proposed by Watson and Crick which is well known as Watson - Crick Model. In this model, A binds with T and G binds with C. All DNA computing applications are based on Watson-Crick complementary model. DNA computing is concerned with the use of DNA molecules for the implementation of computational processes.

L.M. Adleman [1] proposed molecular computation of solutions to combinatorial problems. He is a pioneer for carrying computations using DNA molecules. Using the tools of molecular biology he came up with an idea to solve an instance of the directed Hamiltonian path problem. Gehani et al. [2] introduced the first ever trial of DNA based cryptography. The DNA based cryptographic approach used one-time-pads (OTPs) that are considered unbreakable. But this encryption technique offered a limited security. It could be broken with some reasonable assumptions on the entropy of the plain text messages. Martyn Amos et al. [3] proposed topics in the theory of DNA computing. This area is concerned with the theory, experiments, and applications of DNA computing. The theoretical developments are demonstrated by discussing various topics like operations on DNA which include synthesis, denaturing, annealing, ligation, gel electrophoresis etc. An introduction to the basic structure of DNA and the basic DNA processing tools has also been discussed. G.Z. Cui et al. [4] proposed an encryption scheme using DNA technology. The exceptional energy efficiency, vast parallelism and extraordinary information density which are inherent in DNA molecules are explored. These are used for data storage, computing and cryptography. An encryption scheme has been designed by using the technologies such as DNA synthesis, Polymerase Chain Reaction (PCR) amplification, DNA digital coding as well as the theory of traditional cryptography**.** The coding mode and the primers are used as the key of the encryption scheme. Monica Borda [5] proposed DNA secret writing techniques. The basic principles of bimolecular computation (BMC) and several algorithms for DNA cryptography as well as steganography are proposed. The concepts of OTP, DNA XOR OTP and DNA chromosomes indexing are used. The various works done in [6] - [11] point to the various opportunities in this field and the different methods of DNA computing in the field of cryptography.

2. Methodology

In the proposed work the encrypted text and the key are embedded in the watermarked image. In the proposed work both the methods i.e. cryptography and steganogrphy are combined and the watermarked image is finally generated to protect the data from illegal manipulations. E.coli genome sequence is used as the reference sequence shared between the sender and receiver securely. The methodology of the proposed work consists of the security operations performed at the sender and receiver end. The Fig.1 shows the steps performed in the proposed method at the sender end.

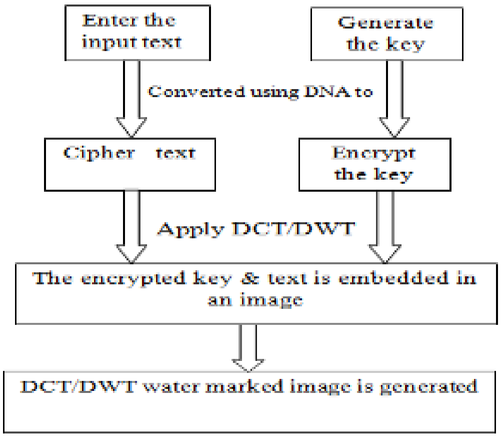


Fig.1 Operations performed at the sender end.

Algorithm 1: Security operations at the sender end.

Input: Cover image, key and text.

Output: Watermarked image.

Steps

Step 1: Key values of range 16 are generated by the random process with the values from 0 to 9.

Step 2: The key values are scrambled.

Step 3: Bit wise XOR is performed of the key generated and the scrambled key to generate the encryption key.

Step 4: The user enters the input in the form of text.

Step 5: Text is converted to its corresponding ASCII values.

Step 6: ASCII values are converted into 8 bit binary values.

Step 7: The 8 bit binary sequence is converted to 4 nucleotide sequence depending on the DNA conventions (A=00, T=01, G=10, C=11).

Step 8: The position and the frequency of the occurrence of four nucleotide sequences in the reference DNA sequence is located and it is stored. This is referred to as cipher text.

Step 9: The cipher text and the key generated is hidden in the image by applying DCT or DWT to generate DCT/DWT watermarked image.

Step 10: The watermarked image is transmitted to the receiver by the sender.

At the receiver end the inverse of DCT/DWT is applied on the watermarked image. The encrypted key and the cipher text are decrypted. The Fig.2 shows the steps performed in the proposed method at the receiver end.

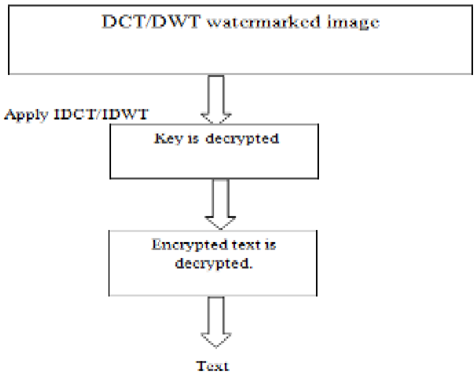


Fig.2 Operations performed at the receiver end.

Algorithm 2: Operations at the receivers end.

Input: watermarked image.

Output: Plain text, key.

Step 1: Inverse of DCT i.e. IDCT or DWT i.e. IDWT is applied on the watermarked image.

Step 2: The key is decrypted.

Step 3: The DNA sequences are retrieved from their positions.

Step 4: Four nucleotides are taken together and are converted to 8 bit binary.

Step 5: Binary numbers are converted to ASCII values.

Step 6: ASCII values are converted to input text.

3. Discrete Cosine Transform

Discrete Cosine Transform is used to hide the data securely in a cover image. DCT based watermarking scheme offers higher resistance to image processing attacks such as JPEG compression, noise, rotation, translation etc. DCT is a general orthogonal transform for digital image processing and signal processing with advantages such as high compression ratio, small bit error rate, good information integration ability and good synthetic effect of calculation complexity. DCT allows an image to be broken up into three frequency bands namely the high, middle and low frequency bands thus making it easier to choose the band in which the secret data is to be embedded. We select certain bytes into which to embed the message using a random number generator. Then resample the bytes to pixel mapping to preserve the color scheme. In the case of an image, information is hidden in the coefficients of the discrete cosine wavelet transform of an image. 2-D DCT is used in the proposed work. DCT is performed on JPEG images to transform successive 8 x 8 pixel blocks of the image, into 64 DCT coefficients each. Then quantized DCT coefficient can be used to hide information. The DCT has a strong energy compaction property.

The steps involved in DCT technique are

1. Preprocessing: To compress a colour image in the JPEG format, the first step is to convert the red, green, blue colour channels to YCbCr space. After that, everything we do for a gray scale image is done to the Y, Cb, and Cr channels. We next partition the image into blocks of size 8 x 8 pixels.

2. Transformation : The DCT tends to push most of the high intensity information (larger values) in the 8 x 8 block to the upper left-hand of the image with the remaining values in the image taking on relatively small values. The DCT is applied to each 8 x 8 block.

3. Quantization: The next step in the JPEG process is the quantization step. Here we will make decisions about values in the transformed image - elements near zero will converted to zero and other elements will be shrunk so that their values are closer to zero. All quantized values will then be rounded to integers.

4. Encoding: The last step in the JPEG process is to encode the transformed and quantized image

4. Discrete Wavelet Transform

The dwt2 command performs single-level two dimensional wavelet decomposition with respect to a particular wavelet filter.

[cA,cH,cV,cD] = dwt2(X,′wname′) computes the approximation coefficients matrix cA and detailed coefficients matrices cH, cV, and cD (horizontal, vertical, and diagonal, respectively), obtained by wavelet decomposition of the input matrix X. The ′wname′ string contains the filter name. Fig. 3 shows the sub-bands in the frequency domain of two dimensional DWT applied on an image.

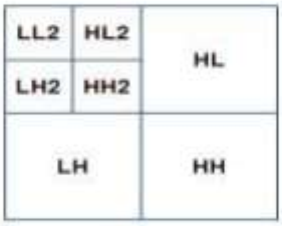


Fig. 3 The sub-band regions in the frequency domain of 2-D DWT applied on an image

5. Performance parameters

The performance parameters determine whether the work carried out is as per the standards or meet the standards in all aspects. The performance parameters to analyze the performance of our work are discussed below.

5.1 Correlation coefficient

It is used to measure the similarity between the original image and the watermarked image. If the value is 1, it means both the images are absolutely similar. The two-dimensional correlation coefficient (r) between two the matrices A, B is calculated as

r = corr2 (A, B)

5.2 Mean Square Error (MSE)

It is an error metric, used to compare image compression quality. It represents the average squared error between the watermarked image and the original image. The lower the value of MSE, the lower is the error occurred. When the two images are identical the MSE is zero.

$$\mathrm{MSE} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(X[i,j] - Y[\,i,j])^{\,2}}{MN}$$

Where X (i, j) and Y (i, j) represent pixel value of original image and watermarked image respectively. M and N represent size of the image.

5.3 Peak Signal to Noise Ratio (PSNR)

The PSNR is an error metric which computes the peak signal-to-noise ratio in decibels, between two images. This ratio is often used as a quality measurement between the original and the watermarked image. The higher the PSNR value, the better is the quality of the compressed or reconstructed image. It represents a measure of the peak error.

PSNR=$10\log_{10}$ peakval/MSE)

where peakval is either specified by the user or taken from the range of the image datatype.

5.4 Embedding Capacity

It determines the amount of secret data that can be safely and securely hidden in a cover image without being noticed or detected. Higher the embedding capacity, higher is the amount of data that can be embedded.

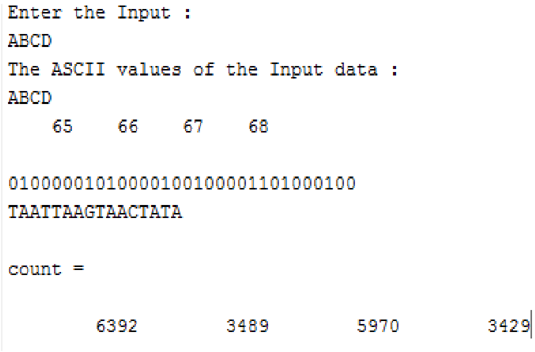
6. Experimental results

The experiment is carried out selecting a cover image in which cipher text and encrypted key is to be embedded by the application of DCT/DWT techniques**.** Matlab 2010 version is used to carry out the experiment.

Colour image of lena is selected and is converted to gray scale.

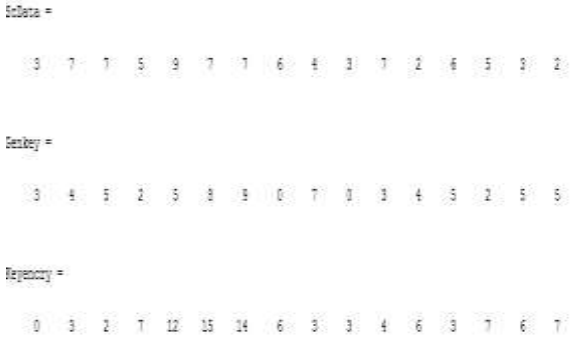
Fig.4 Gray scale image

Text input is given and it is converted to the corresponding ASCII values. The ASCII values are converted to 8 bit binary values. Binary values are converted to DNA sequence (‘A’=00, ‘T’= 01, ‘G’=10, ‘C’=11). The frequency of occurrence of four nucleotide sequence which represents a single character in the reference E.coli genome sequence is displayed as “count” and its position is stored. This is referred to as cipher text.



```
Enter the Input :
ABCD
The ASCII values of the Input data :
ABCD
    65     66     67     68

01000001010000100100001101000100
TAATTAAGTAACTATA

count =

       6392         3489         5970         3429
```

Random key is generated, scrambled and then encrypted.



```
ScData =

  3   7   7   5   9   7   7   6   4   3   7   2   6   5   3   2


Senkey =

  3   4   5   2   3   8   9   0   7   0   3   4   5   2   5   5


Keyencry =

  0   3   2   7  12  15  14   6   3   3   4   6   3   7   6   7
```

The encrypted text and key are embedded in the gray scale image to generate DCT watermarked image.

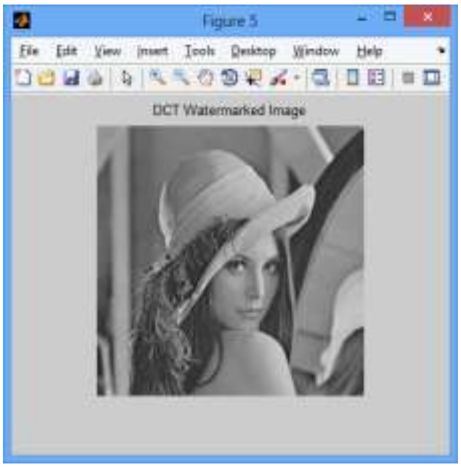
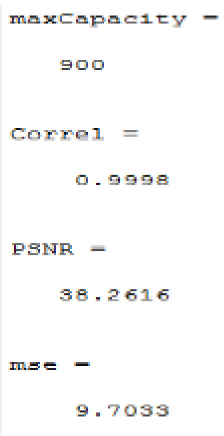


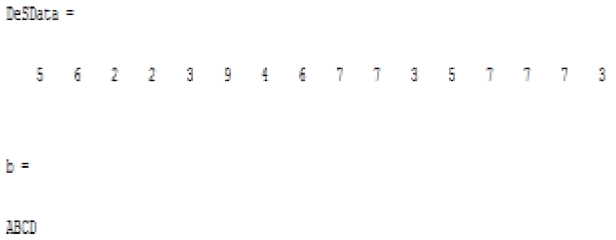
Fig.5 Text & key are embedded in the gray scale image to generate the DCT watermarked image

The performance parameters: maximum embedding capacity of the watermarked image, correlation of the gray-scale and watermarked image, PSNR and MSE values are calculated and displayed. The correlation of gray-scale image and DCT watermarked image is 0.9998≈1. Hence DCT has a good embedding capacity and also ensures security to the data embedded.



```
maxCapacity =

   900


Correl =

    0.9998


PSNR =

   38.2616


mse =

    9.7033
```

Finally the IDCT is applied on the watermarked image. Initially random key is decrypted. Then the input text is decrypted.

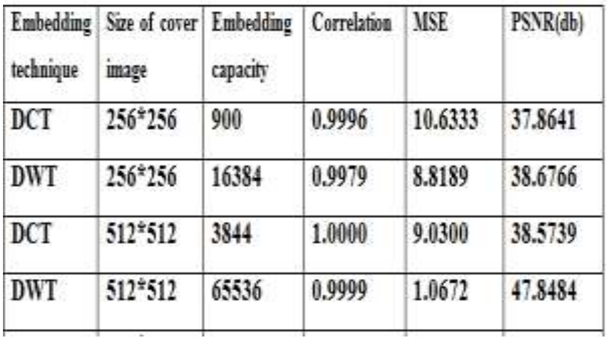


```
DeSData =

   5   6   2   2   3   9   4   6   7   7   3   5   7   7   7   3


b =

ABCD
```

Similar procedure can be followed by applying DWT technique for embedding the text and the key.

The comparison of DCT and DWT techniques for different parameters is given in table 3.1 below. The values obtained for various parameters are tabulated.

Table 3.1 Compares DCT and DWT techniques used for embedding the data in a cover image.



Embedding technique	Size of cover image	Embedding capacity	Correlation	MSE	PSNR(db)
DCT	256*256	900	0.9996	10.6333	37.8641
DWT	256*256	16384	0.9979	8.8189	38.6766
DCT	512*512	3844	1.0000	9.0300	38.5739
DWT	512*512	65536	0.9999	1.0672	47.8484

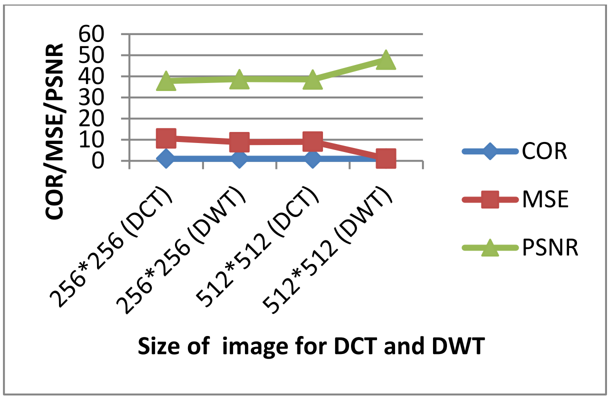


Fig.6 The above graph shows the comparison of the DCT and DWT techniques for the performance parameters correlation, MSE and PSNR applied on the cover images of sizes 256*256 and 512*512.

From the above Fig.6, it is clear that DWT is better than DCT for embedding data in the image.

Conclusion

In today's technological era, data travel widely and very rapidly, also in multiple manifestations, through email and across the Internet. Now-a-days, telemedicine has become a common field for the transmission of images. Internet has a vital role to play in dealing with business transactions. In business transactions, sensible data such as pin numbers are encrypted before transmission. Corporations have a very little visibility into exactly where their documents are being accessed or by whom. Thus protecting sensitive information is an ethical and legal requirement. DNA based encryption helps in the secure transmission of confidential data. The proposed encryption scheme with DNA technology is easy to implement. It can resist brute-force, statistical and differential attack. It is therefore suitable for multilevel security applications of today’s network. Here DNA based cryptography is employed to encrypt the sensitive data. Then a key is generated and encrypted. The data and key are then embedded in a cover image securely by using DCT or DWT techniques to generate DCT/DWT watermarked image. This watermarked image is securely transmitted to the receiver. A comparative study of DCT and DWT techniques is also performed.

References

[1] L. M. Adleman, “Molecular Computation of Solutions to Combinatorial Problems”, Science, 266, November 1994, pp. 1021-1024.

[2] Gehani, Ashish La Bean, Thomas H. Reif, H.John, “DNA Based Cryptography”, Dimacs Series In Discrete Mathematics and Theoretical Computer Science 2000, pp. 54:233-249.

[3] M. Amos, G. Paun, and G. Rozenberg, “Topics in the theory of DNA computing”, Theoretical Computer Science 2002, 287, pp. 3-38.

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

[5] Monica BORDA, “DNA secret writing Techniques”, IEEE conferences 2010, pp.451-456.

[6] J. Chen, “A DNA-based, bimolecular cryptography design”, in IEEE International Symposium on Circuits and Systems (IS-CAS), 2003, pp.822–825

[7] Souhila Sadeg “An Encryption algorithm inspired from DNA”, IEEE, November 2010, pp.344 –349.

[8] G.Xiao, M. Lu, L. Qin, X. Lai, "New field of cryptography: DNA cryptography," Chinese Science Bulletin, vol. 51, Jun. 2006, pp. 1139-1144.

[9] Ning Kang, “A pseudo DNA cryptography Method”, http:// arxiv.org/abs/ 0903.2693, 2009.

[10] G. Z. Cui, L. M. Qin, Y. F Wang and X. C. Zhang, “Information Security Technology Based on DNA Computing,” 2007 IEEE International Workshop on Anti-counterfeiting Security, Identification. 2007, pp. 288–291.

[11] Grasha Jacob, Murugan A, “A Hybrid Encryption Scheme using DNA Technology”, IJCSCS Vol 3, Feb 2013.