A Review on DNA based Encryption and Steganography

Mumthas S[1], Lijiya A[2]

[1]M. Tech, Department of Computer Science and Engineering, NIT Calicut

[2]Assistant Professor, Department of Computer Science and Engineering, NIT Calicut

Abstract: ***Biotechnological methods can be used for cryptography to improve security of data. Steganography is the act of hiding messages inside an image. Combining these 2 methods is a topic of high relevance since secure communication is inevitable for mankind. This paper provides an overview of DNA based steganographic methods.***

Keywords: DNA, Steganography, Cryptography

1. Introduction

Security is the main concern of any type of communication. In secure communication aim is to improve the security of data being exchanged between 2 parties, say A and B. It can be accomplished using several methods. Cryptography or steganography can be used to improve the security. Steganography even hides the presence of a message. DNA based steganography is the act of using steganography along with DNA encryption. It has the advantage of increasing the randomness of message so that it cannot be extracted easily by a third party. General classification of steganographic methods is given in figure1.

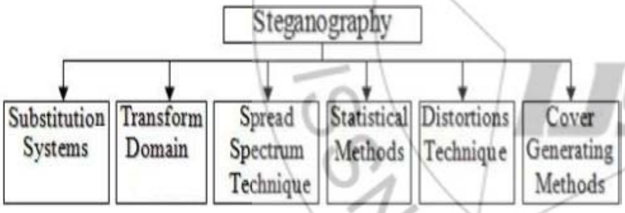


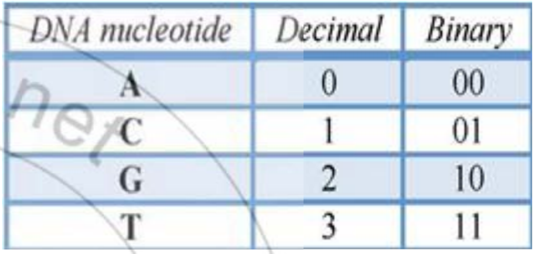
Figure 1: Classification of steganographic methods

2. DNA Computing

Till now we were using silicon based computers. Adleman in 1994, showed DNA(Dioxyribo Nucleic Acid) molecules can be used to create computer with great advantages like parallelism, less power and so on. In his pioneering work, he used chemical reactions with DNA and solved directed Hamiltonian path problem in an efficient way. It shows the power of DNA computing. Later several works were done in this field. It can be used in security domain also. Here we are planning to use the theoratical concepts of DNA computing for implementing steganography. Following are the basic biological terminologies used in this paper,

1) DNA - Its the carrier of genetic information. Its made up of four nucleoide bases, Adenine(A), Thymine(T), Cytocine(C) and Guanine(G). Generally A pair with T and C pair with G.
2) DNA encoding technique - Its a binary coding scheme for the ease of DNA computation. Binary to base mapping is given in table 1.

Table 1: DNA digital coding [11]



DNA nucleotide	Decimal	Binary
A	0	00
C	1	01
G	2	10
T	3	11

3) Codons - A codon triplet is made of three letters out of these four possible bases. So there are 64 total combinations.
4) Degenerative Codons - When two or more codons codes for the same amino acid these are called degenerative codons.

3. Current State of Literature

Today is the era of digital communication. Security and privacy is important for communication. Cryptography is the proess of hiding the mesage. Its known to exist from long time back.Steganography is derived from two words, stego means secret and graphy means writing. So its the act of secret writing. In this paper, we are discussing about only steganography using DNA.

Adleman is the father of DNA computation [1]. He done chemical reactions and shown how DNAs can be used for computations. We are going to discuss about the works done using theoretical DNA computing. Catherine Taylor [2], proposed an idea in which information is encoded into DNA strands, and then converted into microdots. A microdot is a highly reduced photograph of a typewritten page. Developed DNA based doubly steganographic method. First done DNA encryption and then reduced it to a microdot.Simple substitution cipher is used for encryption. Because of the huge possibilities of DNA nucleotides, it acts as a complex background for storing secret message. Random key is used for encryption. Disadvantage is its Expensive.

Andre Leier et.al. [3] proposed cryptography using DNA binary strands. They proposed two different DNA based cryptographic techniques. In method 1, initially mix the binary encoded plaintext with dummy strands in equimolar amounts. Here decryption is done using Polymerase Chain Reation(PCR). In method 2, encryption same as before. Here they used gel image of dummy pool as the key. Decryption is done by graphical method. Method 2 has the advantage of easy encryption, but resolution of gel is a problem.

Jie chen [4], used the random nature of DNA for making the cryptographic system unbreakable. Here they used carbon nanotubes as a medium for message transmission. Plaintext messages are converted to cipher text by adding message with one time pads. Here DNA sequences act as one time pads. But this method is expensive.

Pak chung wong et.al [5], proposed an idea of DNA memory prototype. Today, we use magnetic media and silicon chips to store our data. All these storage media can easily destroyed by people or natural disasters. So they proposed an alternate storage mechanism. Here initially Encode meaningful information as artificial DNA sequences. Then transform the sequences to living organisms. Allow the organism to grow and multiply. Extract the information back from organisms. Success of this method depends on finding good storage medium to ensure adequate protection for the encoded DNA strands. Host with embedded information must be able to grow and multiply. Advantage is that it has enormous potential capacity. Disadvantage is that mutation of organism may affect the integrity of embedded messages.

Monica Borda [8], published a paper on DNA secret writing. Steganography using DNA hybridization has five steps, plaintext message given in ASCII is converted to binary. Evaluate required length for ssDNA OTP. If each bit is encoded with 10 nucleotide, OTP of length > 10*n. The encrypted message is placed between two primers and hidden in a microdot. Perform decryption using PCR.

Qiang Zhang et.al. [9], published a paper on Image encryption using DNA addition combining with chaotic maps. Here initially encode the original image to obtain DNA sequence matrix. Divide this matrix to equal blocks and then carry out DNA sequence addition operation. Find the DNA sequence complement using 2D logistic maps. Decrytion done as reverse of above.

Deepak Kumar [10], proposed the idea of secret data writing using DNA sequences. Here DNA OTP method is used for defining the new security algorithm. DNA coding is necessary because we cannot process the DNA molecules as in form of alphabets, so change alphabets to ASCII. Almost same as Monica Borda's algorithm.

Amal Khalifa [11], proposed a steganography algorithm to exchange data secretly. Its implemented in mainly 2 levels. In first level, encryption is done using DNA based play fair cipher. In second level, encrypted message is hidden to some reference DNA using substitution. The performance of presented algorithm is also analysed with respect to robustness against attaks as well as hiding capacity. Construct the play fair matrix using the secret key and then apply traditional play fair encryption process to get the chains of aminoacids that corresponds to the ciphertext. Map the aminoacids back to DNA codons and append all ambiguity numbers in the form of nucleotide. Apply the substitution method to hide encrypted sequence to some reference DNA. Advantage is that it provides high level of security. Disadvantage is that imple substitution method is used.

Algorithm 1 : DNA encryption[11]
1) Convert plaintext to binary using 8 bits encoding.
2) Code binary data into a DNA sequence using some binary coding coding rule.
3) Map the codons of the DNA sequence into aminoacids and save a 2 bit ambiguity number.

Table 2 shows mapping of aminoacids to 26 character alphabets.

Shyamasree C M, Sheena Anees [12], proposed highly secure DNA based audio steganography. Here a highly secure method to hide the messages is proposed. It works in three levels. Single level of encryption and 2 levels of steganography is used. Following are the 3 levels. In level 1, DNA based play fair cipher encryption is used.In level 2, DNA encrypted

message is hidden in randomly generated DNA sequence. In level 3, audio steganography is used. Initial steps are same as proposed by Amal Khalifa. Do audio steganography as a last step. Advantage is that three level of security.

Algorithm 2 : Audio steganography[12]
1: Read the embedded DNA sequence and audio file in binary format.
2: A password is provided in addition to embedded DNA sequence to provide additional layer of security.
3: Read the password.
4: Encrypt the password and the embedded DNA sequence.
5: Sample the audio file.
6: Encode the length of the cipher in lower half of first 32 audio samples.
7: Encode the cipher in lower half of remaining audio samples.

Prasenjit Das [14], proposed DNA based image steganography. Proposed algorithm uses images as primary cover media for message transfer between two interested parties. In the embedding algorithm following are the main steps. In step 1, the secret message bits are hidden inside the ssDNA, which in turn is hidden inside the cover image. In step 2 , check the validity of both message and cover. In step 3, check capacity with a set of 2D map parameters and same cover. In step 4, generate ssDNA . In step 5, encrypt secret data. In step 6, embed encrypted secret data into ssDNA. In step 7, generate and embed header . In step 8, generate stego image with modified pixel values. Advantage is that DNA is attributed by pixel propertes of the image. Hence more secure. Figure 3 shows the procedure.

Fasila K.A. et al [15], proposed the idea of multi phase crypto system. Here a hybrid cryptography based on RGB colours is proposed. Convert the plaintext to matrix form, pass it through a number of manipulation steps. Security is further enhanced by using a strong key which is encapsulated using DNA steganography method. As a next layer of security encryption is done using DNA bases and aminoacids. Encryption algorithm has following steps. In step 1, matrix generation and manipulation. In step 2, mini cipher generation .In step 3, key encapsulation . In step 4, cipher text generation . In step 5, conversion to colurs.

Sreeja C.S et al [16], proposed a DNA symmetric algorithm based on the pseudo DNA cryptography and central dogma of molecular biology. The suggested algorithm uses splicing and padding techniques along with complementary rules which make the algorithm more secure as its an additional layer of security than conventional cryptography techniques. Shweta et al [17], proposed paper on cascaded DNA cryptography and steganography. Initially it performs DNA cryptography and then its hidden in a random frame of video.

Prasenjit Das et.al. [18], proposed an algorithm which improves the existing dual cover steganography by reducing the noise bits. P.Vijayakumar et.al. [19], proposed a DNA steganographic algorithm using hyper elliptic cryptography. It ensure high level of security to image file and also assure digital media security.Proposed algorithm which provides high level of security to image file. Main findings are embedding capacity is increased, mean square error is reduced. PSNR value is increased.

```
Algorithm 3: Steganography using HECC[19]
1: Get an image file of known size.
2: Convert the pixel value into DNA nucleotide.
3: Convert DNA nucleotide to binary.
4: Binary digits of stego and cover image subject to XOR
   operation.
5: Convert XOR values into decimal value.
6: Convert the decimal numbers into HEC points using
   Koblitz method.
7: These points are encrypted using HECC encryption
   algorithm and obtain the ciphertext points.
```

4. Comparison of Existing Methods

Table 2 below shows comparison of methods existing in literature so far.

Table 2: Comparison of existing methods

Year	Author	Key used	Level of security	Cover medium	Method	Paper
1999	Catherine	Random	2	DNA	Substitution	[2]
2000	Andre Leier	Random	2	DNA	Substitution	[3]
2003	Jie Chan	Random	2	DNA	Substitution	[4]
2003	Pak Chung	Random	2	Organism	Substitution	[5]
2010	Monica	Random	2	DNA	Substitution	[8]
2011	Deepak	Random	2	ssDNA	Substitution	[10]
2011	Amal	Random	2	DNA	Substitution	[11]
2013	Shyamasree	Random	3	Audio	Substitution	[12]
2014	Prasanjit Das	Random	2	Image	Substitution	[14]
2014	Fasila KA	Securely generated	1	DNA	RGB colors	[15]
2014	Sreeja CS	Random	1	DNA	Splicing and padding	[16]
2015	Shweta	Random	2	A frame of video	DNA with RGB colors	[17]
2016	VijayaKumar	Public key	2	Image	HECC	[19]

5. Conclusion

Steganography protects secret information. We have seen several steganographic methods using DNA. This method has various advantages like speed, minimal storage requirements and minimal power requirements.

References

[1] Leonard M Adleman. Computing with DNA. Scientific American, pages 34–41, August 1998.

[2] Catherine Taylor Clelland. Hiding Messages in DNA Microdots. Nature, 399:533–534, June 1999.

[3] Andre Leier. Cryptography with DNA Binary Strands. BioSystems, 57:13–22, April 2000.

[4] Jie Chen. A dna-based, biomolecular cryptography design. In Circuits and Systems, 2003. ISCAS ’03. Proceedings of the 2003 International Symposium on, volume 3, pages III–822–III–825 vol.3, May 2003.

[5] Pak Chung Wong. Organic Data Memory using DNA Approach. In Communications of the ACM, volume 46, pages 95–98, January 2000.

[6] VenkatramanS, Ajith Abraham, and M. Paprzycki. Significance of steganography on data security. In Information Technology: Coding and Computing, pages 347–351, April 2004.

[7] X. Wang and Q. Zhang. Dna computing-based cryptography. In Fourth International Conference on Bio-Inspired Computing, pages 1–3, Oct 2009.

[8] M. Borda and O. Tornea. Dna secret writing techniques. In Communications (COMM), 2010 8th International Conference on,, pages 451–456, June 2010.

[9] Qiang Zhang. Image encryption using dna addition combining with chaotic maps. Elsevier, Mathematical and Computer Modelling, 52(1112):2028 – 2035, 2010. The BIC-TA 2009 Special IssueInternational Conference on Bio-Inspired Computing: Theory and Applications.

[10] D. Kumar and S. Singh. Secret data writing using dna sequences. In Emerging Trends in Networks and Computer Communications (ETNCC), 2011 International Conference on,, pages 402–405, April 2011.

[11] Khalifa and A. Atito. High-capacity dna-based steganography. In Informatics and Systems (INFOS), 2012 8th International Conference on,, pages BIO–76–BIO– 80, May 2012.

[12] M. Shyamasree and S. Anees. Highly secure dna-based audio steganography. In Recent Trends in Information Technology (ICRTIT), 2013 International Conference on,, pages 519–524, July 2013.

[13] P.VijayaKumar and V.Vijayalakshmi. Enhanced level of security using DNA computing technique with hyperelliptic curve cryptography. Network Security, 4, 2013.

[14] P. Das and N. Kar. A DNA based image steganography using 2d chaotic map. In Electronics and Communication Systems (ICECS), 2014 International Conference on,, pages 1–5, Feb 2014.

[15] F. K. A. and D. Antony. A multiphase cryptosystem with secure key encapsulation scheme based on principles of dna computing. In Advances in Computing

and Communications (ICACC), 2014 Fourth International Conference on,, pages 1–4, Aug 2014.

[16] S. C. S, M. Misbahuddin, and M. Hashim N. P. Dna for information security: A survey on dna computing and a pseudo dna method based on central dogma of molecular biology. In Computer and Communications Technologies (ICCCT), 2014 International Conference on,, pages 1–6, Dec 2014.

[17] Shweta and S. Indora. Cascaded dna cryptography and steganography. In Green Computing and Internet of Things (ICGCIoT), 2015 International Conference on,, pages 104–107, Oct 2015.

[18] Prasenjit Das, Subhrajyoti Deb, Nirmalya Kar, and Baby Bhattacharya. An improved dna based dual cover steganography. Elsevier, Procedia Computer Science, 46:604 – 611, 2015.

[19] P. Vijayakumar, V. Vijayalakshmi, and G. Zayaraz. An improved level of security for dna steganography using hyperelliptic curve cryptography. Springer, Wireless Personal Communications, 89(4):1221–1242, 2016.

[20] Seyyed Mohammad Reza Farschi and H. Farschi. A novel chaotic approach for information hiding in image. Springer, Nonlinear Dynamics, 69(4):1525–1539, 2012.