International Journal of Engineering Research and General Science Volume 2, Issue 6, October-November, 2014
ISSN 2091-2730

Data Encryption Using DNA Sequences Based On Complementary Rules – A Review

Ms.Amruta D. Umalkar
Master of Engineering Department of Information Technology
. Sipna College of Engg. And Technology Amravati, India
amruta.umalkar2014@gmail.com

Prof. Pritish A. Tijare
Department of Information Technology
Sipna College of Engg. And Technology Amravati, India
pritishtijare@rediffmail.com

Abstract— With the quick development of net technology and data process technology, the knowledge is unremarkable transmitted via the net. The vital data in transmission is definitely intercepted by unknown person or hacker. So as to reinforce the knowledge security, encryption becomes a vital analysis is direction. A message cryptography formula supported deoxyribonucleic acid (Deoxyribo Nucleic Acid) sequence for presenting during this paper. The most purpose of this formula is to write the message with the premise of complementary rules deoxyribonucleic acid sequence.

Keywords— Data hiding; DNA Sequences; Complementary Rules, Secure Transmission and reception.

INTRODUCTION

The security of a system is essential nowadays. With the growth of the information technology power, and with the emergence of new technologies, the number of threats a user is supposed to deal with grew exponentially.

With the increasing growth of transmission applications, security has become a crucial issue on communication. DNA secret writing is rising as a brand new secret writing field wherever polymer is employed to hold the knowledge. The fascinating options concerning the structure of polymer square measure the complementary rule. These rules are used for proposing message encryption methods.

Message encryption is the process of transmitting the message stealthily. In the message encryption, the original message is transformed into an equivalent alternative by a definite encoding mechanism. This message is then send to the receiver. An encoding scheme by incorporating the important chemical characteristics of biological DNA (Deoxyribonucleic Acid) sequences or structure of purines and pyrimidines could serve as an effective stealth transmission of an message would be so secure that it could not be easily cracked. In the proposed algorithm, a DNA sequence or structure is initial randomly taken and complementary rules are framed so the secrete message to be sent is encoded at the sender's aspect. At the receiver's aspect, the decryption method is completed and therefore the original message is extracted out.

A DNA sequence is a sequence composed of four distinct letters, A, C, G and T. Each nucleotide contains a phosphate attached to a sugar molecule (deoxyribose) and one of four bases, adenine (A), cytosine (C), guanine (G), or thymine (T). It is the arrangement of the bases in a sequence, for instance like ATTGCCAT, that determines the encoded gene. The natural sequence pattern with complementary coding and chemical classification of the nucleotides can be used to shield the message.

Table I. DNA Based Coding

DNA Base	code
A	00
C	01
G	10
T	11

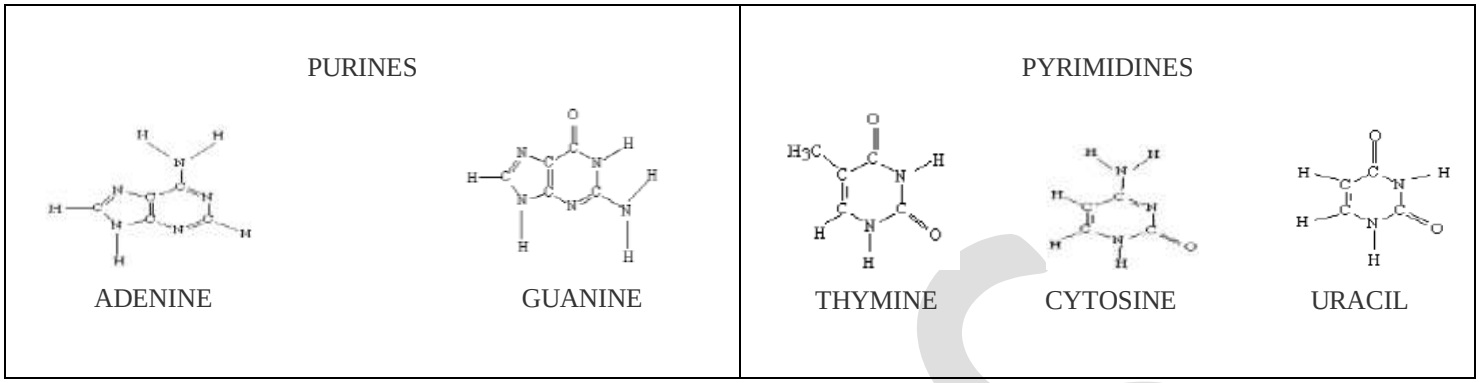

Figure 1. Structure of Purines and Pyrimidines

LITERATURE SURVEY

Message encryption using DNA sequence is a very new technique still evolving and tried out for secure transmission and reception of hided messages. The method is deemed to be so secure that it would be very difficult for any intruder to break the encrypted message and retrieve the actual message. Only the intended receiver can decrypt and receive the original message.

The following is some of the prospective DNA based messages encryption and data hiding schemes reported recently.

K. Menaka [1], proposed a data hiding method where the algorithm first randomly selects a DNA sequence. The message to be encoded is then taken and each letter in the faked DNA sequence. Each letter in the message is converted into its ASCII equivalent and they are then converted into equivalent binary form. Each two digits in the converted binary sequence are converted as per Table 1. Then, the message index position (first position of each letter) in the faked DNA sequence is applied to each letter of the converted sequence. Each digit in the resultant sequence is replaced with its equivalent three digit binary value and the equivalent alphabet value is replaced for the binary value. For example, if the obtained binary value is 010 011 101 …, then it will be replaced as C D F… where A has the value 000, B has 001 and so on. The resultant sequence of alphabets is transmitted over to the receiver. In the receiver side, the reverse process is done in which the original receiver knows the complementary rules and the randomly selected DNA sequence. The message to be sent is then encoded with the fake DNA sequence.

Debnath Bhattacharyya [2] developed an algorithm for data encryption using DNA sequencing. In their algorithm, they have used the concept of indexing the DNA Sequencing and transmitting the message to the receiver. They have not used any complementary rules.

Jin-Shiuh Taur et al. [3] proposed a way referred to as Table Lookup Substitution methodology (TLSM) that might double the capability of message activity. In TSLM, they need replaced the complementary rule with a rule table. The key plan of the TLSM is to increase the 1-bit complementary rule into a 2-bit rule table so every conversion of letters will represent 2 bits of the secret message.

In the method by Cheng Guo, Shiu, [4] the hiding procedure substitutes another letter for an existing letter on a special location set by the algorithm. The embedding algorithm encompasses a conversion operates that converts a given letter with a selected letter outlined by the complementary rule. For example, if a complementary rule is outline as (AC)(CG)(GT)(TA), then the result of θ(G) are going to be T, and therefore the result of θ(T) are going to be A. To boot, the substitution methodology can convert the letter s into s (unchanged), θ(s) and θ (θ (s)) once the secrete message is 0, 1 and no data, respectively.

Mohammad Reza Abbasy, et al. proposed [5] an information hiding methodology wherever data was efficiently encoded and decoded following the properties of DNA sequence. Complementary combine rules of DNA were employed in their methodology.

Kritika Gupta* Shailendra Singh[6] has been projected a DNA Based Cryptological Techniques for an encryption algorithm based on OTP (one-time-pad) that involve data encryption using traditional mathematical operations and/or data manipulating DNA techniques. However once an encryption algorithm has been applied and therefore the data is transmitted on the transmission media: there's a clear stage that the data, although within the cipher type gets manipulated by any interceptor.

Snehal Javheri, Rahul Kulkarni[7] proposed an algorithmic program has two phases in consequence: these are Primary Cipher text generation using exploitation substitution methodology followed by Final Cipher text generation exploitation DNA digital secret writing.

In the Primary Cipher text generation phase, the coding algorithmic program uses OTP (one-time-pad) key generation theme, since nearly one key for one piece of data is sufficient to supply voluminous strength in coding technique. The projected methodology uses indiscriminately generated symmetrical key of 8 bits size by the supposed receiver and provided to the sender. Therefore the sender can have partial information of the personal key solely and so it generates the remainder part of the keys to cipher the data.

The Byte values are extracted from the input data or message. The additional secret writing method works on unsigned byte values of the input data or text referred to as plain text. These byte values are replaced by combination of alphabets and special symbols exploitation substitution methodology. And so this substitution worths are regenerate into its binary value. So as to embed lots of security additional bits are padded at each ends of the first cipher text. These additional bits are nothing however the file size information that is provided to the receiver through key. So the secret key, the data of primer pairs are shared between sender and receiver through the secret key channel.

In the DNA digital secret writing section, the Ultimate Cipher text is generated from Primary Cipher text exploitation DNA digital encryption technique. From a process purpose of read, cannot process the DNA molecules as in sort of alphabets, therefore the DNA sequence encryption is employed during this methodology through that the binary knowledge is regenerate into DNA format and it's vice versa.

Guangzhao Cui #1, Limin Qin #2, Yanfeng Wang #3, Xuncai Zhang *4 [8]proposed a secrete writing theme by exploitation the technologies of DNA synthesis, PCR amplification and DNA digital secrete writing additionally because the theory of ancient cryptography. The supposed PCR two primer pairs was used because the key of this theme that not severally designed by sender or receiver, however severally designed by the entire cooperation of sender and receiver. This operation might increase the safety of this secrete writing theme. The standard secretes writing methodology and DNA digital cryptography is wont to preprocess to the plaintext. Through this preprocess operation will get fully different ciphertext from the identical plaintext, which might effectively stop attack from a potential word as PCR primers. The quality of biological troublesome issues and cryptography computing difficulties give a double security safeguards for the theme. And therefore the security analysis the secrete writing theme has high confidential strength.

Ritu Gupta, Anchal Jain [9] symmetric-key encoding algorithmic rule supported the DNA approach is projected. The initial key sequence is enlarged to desire length victimization projected key growth technique guided by the pseudo random sequence. The advantage is that there's no need to send an extended key over the channel. The variable key growth in encoding method combined with DNA addition and complement makes the technique sufficiently secure. A DNA sequence consists of four nucleic acid bases A (adenine), C (cytosine), G (guanine), T (thymine), wherever A and T are complementary, and G and C are complementary. Also use C, T, A and G to denote 00, 01, 10, 11 (the corresponding decimal digits are "0123"). By victimization this encoding technique every 8-bit component worth of the gray scale image is pictured as a nucleotide string of length four. Reciprocally to decrypt the nucleotide string will get a binary sequence simply. In total 4! = 24 forms of writing, there are only 8 of them will meet complementary rule , for instance, the decimal digits "0123" (the corresponding binary range is "00011011") will be encoded in to one of them, like "CTAG", "CATG", "GATC", "GTAC", "TCGA", "TGCA", "ACGT" or "AGCT". There are total six legal complementary rules [3] that are as follows:
(AT)(TC)(CG)(GA), (AT)(TG)(GC)(GA), (AC)(CT)(TG)(GA), (AC)(CG)(GT)(TA),(AG)(GT)(TC)(CA), (AG)(GC)(CT)(TA).
Any one of them for instance, (AG) (GC) (CT) (TA) is applied to projected methodology.

FOLLOWING TABLE SHOWS THE WORK DONE BY RELATED AUTHORS ALONG WITH RESPECTIVE YEARS:

Author	Proposed Work	Year
Guangzhao Cui, Limin Qin, Yanfeng Wang, Xuncai Zhang	Proposed a secrete writing theme by exploitation the technologies of DNA synthesis, PCR amplification and DNA digital secrete writing	2008
Jin-Shiuh Taur	Proposed a way referred to as Table Lookup Substitution methodology (TLSM) that might double the capability of message activity. The key plan of the TLSM is to increase the 1-bit complementary rule into a 2-bit rule table so every conversion of letters will represent 2 bits of the secret message.	2010
Mohammad Reza Abbasy	Proposed an information hiding methodology wherever data was efficiently encoded and decoded following the properties of DNA sequence.	2011
Cheng Guo, Shiu,	Proposed the hiding procedure substitutes another letter for an existing letter on a special location set by the algorithm. The embedding algorithm encompasses a conversion operates that converts a given letter with a selected letter outlined by the complementary rule.	2012
Debnath Bhattacharyya	Developed an algorithm for data encryption using DNA sequencing. They have used the concept of indexing not used any complementary rules.	2013
Kritika Gupta* Shailendra Singh	Projected a DNA Based Cryptological Techniques for an encryption algorithm based on OTP (one-time-pad) that involve data encryption using traditional mathematical operations and/or data manipulating DNA techniques.	2013
K. Menaka	Proposed a data hiding method where DNA based complementary rules are used for hiding the data.	2014
Snehal Javheri, Rahul Kulkarni	Proposed an algorithmic program has two phases in consequence: these are Primary Cipher text generation using exploitation substitution methodology followed by Final Cipher text generation exploitation DNA digital secret writing.	2014
Ritu Gupta, Anchal Jain	Projected a symmetric-key encoding algorithmic rule supported the DNA approach.	2014

ANALYSIS OF PROBLEM

The message encryption algorithm has many steps to break and to get the original message. The random selection of DNA sequence can be increased to many numbers. The complementary rules which are formed based on the properties of DNA could also be increased since DNA sequence has many biological properties and using those properties also some more complementary rules can be formed.

The complementary rules which are formed based on the properties of DNA could also be increased since DNA sequence has many biological properties and using those properties also some more complementary rules can be formed for message encryption.

CONCLUSION

The entire proposed algorithm has many steps to break and to get the original message. So, any intruder who receives the intermediate message will never be able to retrieve the original message as intended by the sender. The random selection of DNA sequence can be increased to many numbers. The complementary rules which are formed based on the properties of DNA could also be increased since DNA sequence has many biological properties and using those properties also some more complementary rules can be formed.

Message encryption using DNA sequence is a very new technique still evolving and tried out for secure transmission and reception of hided messages. The method is deemed to be so secure that it would be very difficult for any intrude to break the encrypted message and retrieve the actual message. Only the intended receiver can decrypt and receive the original message.

REFERENCES

[1] K.Menaka"Message Encryption Using DNA Seqence" 978-1-4799-2977-4. 201 I EEE.

[2] Debnath Bhattacharyya, Samir Kumar Bandyopadhyay, "Hiding Secret Data in DNA Sequence", International Journal of Scientific &Engineering Research Volume 4, Issue 2, February-2013 ISSN 2229-5518.

[3] H. J. Shiu, K. L. Ng, J. F. Fang, R. C. T. Lee and C. H. Huang, "Data hiding methods based upon DNA sequences",Information of Science, vol.180, no.11, pp.2196-2208,2010.

[4] Cheng Guo, Chin-Chen Chang and Zhi-Hui Wang "A New Data Hiding Scheme Based On DNA Sequence" International Journal of Innovative Computing, Information and Control ICIC International Volume 8, Number 1(A), January 2012.

[5]Mohammad Reza Abbasy, Azizah Abdul Manaf, and M.A.Shahidan, "Data Hiding Method Based on DNA Basic Characteristics", International Conference on Digital Enterprise and Information Systems, July 20-22, (2011), London, UK,pp. 53–62.

[6] Kritika Gupta* Shailendra Singh **"**DNA Based Cryptographic Techniques**"** International Journal of Advanced Research in Computer Science and Software Engineering Volume 3, Issue 3, March 2013 ISSN: 2277 128X

[7] Snehal Javheri, Rahul Kulkarni "Secure Data communication and Cryptography based on DNA based Message Encoding" International Journal of Computer Applications (0975 – 8887) Volume 98– No.16, July 2014

[8] Guangzhao Cui #1, Limin Qin #2, Yanfeng Wang #3, Xuncai Zhang *4 "An Encryption Scheme Using DNA Technology" 978-1-4244-2724-6/08/2008 IEEE.

[9] Ritu Gupta, Anchal Jain "A New Image Encryption Algorithm based on DNA Approach" International Journal of Computer Applications (0975 – 8887) Volume 85 – No 18, January 2014