Nat. Volatiles & Essent. Oils, 2021; 8(4): 5678-5687

Information Hiding through DNA Sequence Technology

Paspula Ravinder, M Geetha Yadav, Ms. K Rashmi, Mrs. Ch. Srividhya

Institute of Aeronautical Engineering

Abstract

Information flows through the intranet or internet of global scope. It is mandatory to secure that information to prevent unauthorized access by any node in the path. There are various large numbers of users and vast no of organizations who want to provide security for their crucial huge amount of data from unauthorized persons and hackers and also users need to provide privacy, integrity, and confidentiality of data flowing through the insecure communication channel. The proposed method consists of two rounds of process in the first-round intermediate cipher text will be generated with help of generated DNA encoding table, string matrix, and DNA digital encoding. In the second round of the process, the intermediate cipher will convert into a human mad DNA sequence and transfer to the receiver. This proposed method also provides the integrity with help of a message digest generated from MD12.

Key words: DNA Cryptography, human made DNA sequence, MD12, message digest, encryption, decryption.

1. Introduction

1.1 Introduction to Cryptography

Transmitting information through the insecure channel is in the form that is easily hacked by humans. That is the reason we need to protect the data. Using the Encryption process we convert the real text into an unknown form is known as cipher text. The encryption method is try to hide the sensitive information from anyone for whom it is not intended. The method of turning the cipher text into its actual text is termed as decryption. Cryptography is used to store sensitive information or use insecure networks, such as internet so that no one but the interested recipient can read it. Cryptanalysis is the sculpture of breaking cipher text and retrieving real text without knowing the proper key. Cryptography includes all aspects of providing privacy, authentication, digital signatures, electronic money, and other applications.

1.2 DNA Introduction

De-oxy-ribonucleic acid-DNA is located every cell of human body of all organisms in the worldand transfers genetic information. DNA consists of two parallel biopolymer strands in the form of a double helix. DNA is in the form of double-helical structure that includes four nucleus bases: Adenine, **G**uanine, **C**ytosine &**T**hymine [12]. The detailed information of all living things is stored in DNA bases [14].

1.3 DNA Cryptography

DNA base pairs will be used as the information carrier in DNA Cryptography. Compared with other methods used, the high computing power of DNA chips makes them a more advanced technology [18].

The need of cryptographic techniques has emerged as because of many conventional cryptographic methods such as DES, 3DES, and RSA, etc. are have already been and attacked broken

by many hackers. Much algorithms-based DNA computing has already been proposed for cryptography issues [15-17].Many cryptographic algorithms based on DNA cryptographyhave been developed which use private and public keys for protecting information [18]. Deoxyribonucleic acid-DNA is hereditary material for all living beings and carries genetic information; it consists of two ant parallel biopolymer strands overlapped around one another to generate a double helix form. DNA is a type of double-helical structure that consists of four nucleobases: Adenine, **G**uanine, **C**ytosine &**T**hymine [12]. The detail of any living thing is stored in DNA bases as shown in Figure -1[14]. The information of any living being is stored in DNA bases as shown in Figure -1.

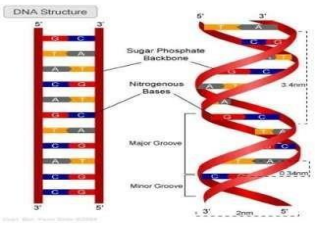


Fig – 1: DNA Strand

All these base pairs will be used as the data carrier in DNA Cryptography. It consists of huge processing power of DNA cells makes it a more advanced technique as compared to other techniques which are being used. As a result of DNA Technology in the future, which may enhance computer data processing. Many DNA computing algorithms are proposed for cryptography issues [18].

2. Literature Survey

P Ravinder, S.Laxman Kumar et al, proposed a method in which the DNA computing for converting plaintext to cipher-text with SSL protocol is used, which gives us three levels of security in WSN, as per this method the energy consumption problem for generating key pairs & generating certificates for sensor nodes are resolved to some extent by assigning key pairs & digital certificate before deploying sensor nodes in any environment and the public key & digital certificate sharing is done using the secure channel (SSL) thus the computation overhead for sensor nodes for generating the keys may be reduced which may in turn reduce the computation time leading to energy efficiency in sensor nodes[18].

Ravinder Paspula et al, proposed a method that consists of two rounds that work on the binary values of the message or plaintext the binary values or bits are read from the plaintext, the session key, a random number, and DNA sequence are shared through a secure channel between sender and receiver prior to communication establishment, the session key bears the information about the key that is used for encrypting the message[9].

Ravinder Paspula et al[10], For providing better security and reliable data transmission, a new method of the encryption process is proposed here. This proposed algorithm consists of two rounds that work on the binary values of the message or plaintext. In this algorithm, binary values or bits are read from the plaintext. A session key, a random number, and DNA sequence is shared through a secure channel between sender and receiver prior to communication establishment. The session key bears the information about the key that is used for encrypting the message. The round 1 key for encryption is computed based on the response of a random number generator and the

information about the key is send to the receiving side through a private channel. The sender will use a random number generator to generate a random number, and then this number along with the shared secret key will go through a function that will produce round 1 encryption key (KE).

Expected Output and Outcome of the Proposal

- Encryption time complexity, as well as decryption time complexity, is to be reduced. The time required for encryption and encryption process is trying to be reduced considerably.
- Security of data transmission is increased with better encryption and decryption processes applied for critical services like defense and external a_airs.3) The authorization of receiving end is to be improved extensively.

Sneha Javheri et al proposed an encryption scheme in which Level1 private key is computed by sharing the values between sender and receiver then Primary ciphertext (PCT) is obtained from plain text with help of level1 private key. Then after calculating the level2 private key, attach starting primers and ending primers to the primary ciphertext(PCT) and generate the final cipher text[11].

Ravinder P et al proposed an algorithm that contains two stages. Stage-1 uses encryption key -KE and Cipher Block Chaining mode CBC- is a mode of operation for a block cipher. The Encryption key is calculated based on a number generated randomly at the source side. After that convert plaintext into 8-bit binary blocks and then every block of plain-text and encryption key-KE will send it to CBC mode. The outcome of one block of CBC will be used as the input (key) for the next block. In stag-1 each and every block will generate an intermediate form of cipher-text also called level-1 cipher-text. Each block of intermediate form of cipher-text is converted into a DNA sequence by applying a DNA encoding scheme. And then apply the DNA complementary rule on generated DNA sequence and convert that DNA sequence into a binary format this is called level-1 cipher-text. In Stage-2, the sender has to choose a selectively reference DNA-Sequence randomly from publicly available DNA- sequences. The selected DNA sequence will act as one of the keys for the encryption stage-2. The receiver must have the information about the selected and used DNA sequence. Then this selected DNA sequence is converted into a binary string using a binary coding scheme. After that converts a binary string into K-bit blocks. Stage-1 and Stage-2 blocks are combined together in such a way that stage-1 blocks are to be appended in front of each k-bit block of stage-2. When the length of the k-bit block of stage-2 is less than the length of the stage-1 block, the stage-2 k-bit block will be repeated that is level-2 cipher-text. This level-2 cipher text converted into a faked-DNA- sequence using a binary coding scheme of DNA this faked sequence is referred to as human-made DNA-sequence [13].

Table.1 DNA Encoding Table

	C	A	T	G
A	ACAT-a	AAAA-y	ATAA-W	AGAG-{
	ACTG-b	AATT-z	ATTT-X	AGTA-[
	ACCC-c	AACC-A	ATCG-Y	AGCG-}
	ACGA-d	AAGG-B	ATGC-Z	AGGG-]
T	TCAT-e	TAAT-C	TTAA-0	TGAA-|
	TCTG-f	TATG-D	TTTT-1	TGTT-\
	TCCG-g	TACC-E	TTCC-2	TGCG-+
	TCGT-h	TAGA-F	TTGG-3	TGGC-=
C	CCAG-i	CAAT-G	CTAT-4	CGAA-_
	CCTA-j	CATG-H	CTTG-5	CGTT--
	CCCG-k	CACG-I	CTCC-6	CGCC-)
	CCGG-1	CAGT-J	CTGA-7	CGGG-(
G	GCAA-m	GAAG-K	GTAT-8	GGAT-*
	GCTT-n	GATA-L	GTTG-9	GGTG-&
	GCCG-o	GACG-M	GTCG-<	GGCC-^
	GCGC-p	GAGG-N	GTGT->	GGGA-%
A	ACTC-q	AATA-O	ATTA-,	AGTT-$
	ACCG-r	AACG-P	ATCC-.	AGCC-#
T	TCTC-s	TATC-Q	TTTA-?	TGTA-@
	TCCC-t	TACG-R	TTCG-/	TGCC-!
C	CCTT-u	CATC-S	CTTC-:	CGTA-~
	CCCC-v	CACC-T	CTCG-;	CGCG-`
G	GCTA-w	GATT-U	GTTC-"	GGTC-€
	GCCC-x	GACC-V	GTCC-'	GGCG-£

Table.2 DNA Digital Coding

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

3. Proposed Method

3.1 Encryption Process

The encryption algorithm comprises of the subsequent steps for encrypting plaintext into Cipher Text: Before encryption starts, the encoding process is administered for plaintext to DNA sequence conversion

Step 1: DNA Encoding Table is generated for encoding of plain text into DNA Sequence using DNA Encoding algorithm.

Step 2: Divide the encoded plaintext into two equal halves. If the generated plaintext is not even, one randomly generated number is appended to both halves to make them both even. With the use of DNA Encoding Table-1, the first half of the plaintext is changed into a DNA sequence, and with the use of DNA, Encoding table-2 obtained from the receiver converts the rest of the half of the plaintext into DNA Sequence.

Forexample:Letusconsidertheplain-text“IARE”.

Dividethe plaintext into two halves equally as follows.

IA **RE**

CACGAACC **TACGTACC**

Step 3: This process includes two rounds applied on both left sideand right sideplain text.

Step 4: Round-1

- *Both left side part and right side parts are converted into string matrix using fillmat() functions[index]*
- *Convert both the string matrix into strings and then apply XOR operation on both parts with key-1*

Step 5: Round-2

- Interchange two parts and apply step4 again
- Convert the string matrices into intermediate cipher text

Step 6: Complimentary Rule

Apply the complimentary rule on intermediate cipher text and generate final cipher text this is called human made DNA Sequence, transfer it along with other DNA Sequences to receiver

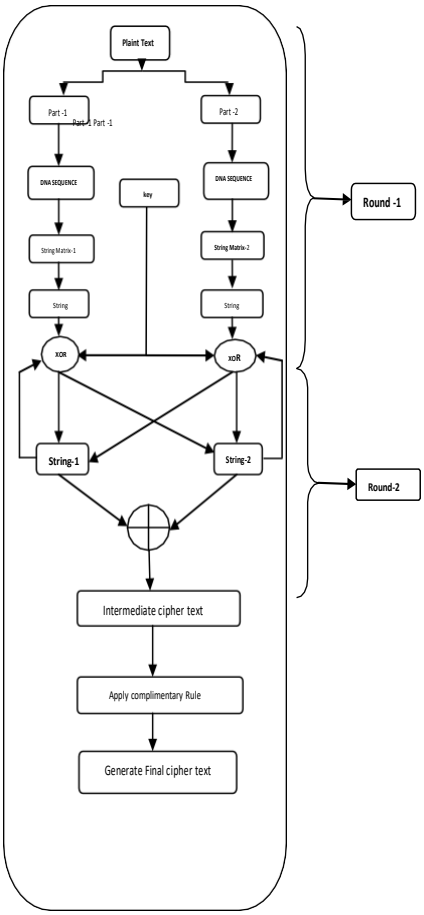


Fig-3 Shows the Encryption Process

3.1 Key Generation Phase

- In this phase that is in key selection phase a 128-bit key is chosen randomly to do the further encryption. This 128-bit key is then divided into two64-bit blocks which will be used as round-key1 and round-key2 at the time of encryption.
- Each block of sub-key is labeled with the DNA base namely A, T,C,G.Then randomly select any combination of these four bases out of the possible 24(4!) combinations without repetition, such as A,C,T,G; G,A,C,T etc. This key will be used in the round 1.

- Then the key order of DNA bases is right shifted to 1 block that is round 1 key is right shifted to 64 bits and this process will be continued for 2 times and as a result 2keys are generated by shifting of 1 block (64 bit) in each round.

For example, if ACGT is the randomly chosen key combination where each base represents each block of 64 bit. So, within the second around the key is going to be TACG since a block is shifted right. Here a single was only chosen and the second key is induced from its previous one. In Figure 1 the key generation approach for two consecutive rounds used in the message encryption phase is depicted where the randomly chosen 128-bit key is split into two subparts each 64-bit and in each round, every subpart of 64-bit right-shifted to one block and as a result, in each round, a unique key is generated. So, in this way, 2 different keys are generated from a single randomly selected key. In this Figure 4, the dotted arrows signify each right shift operation for each block.

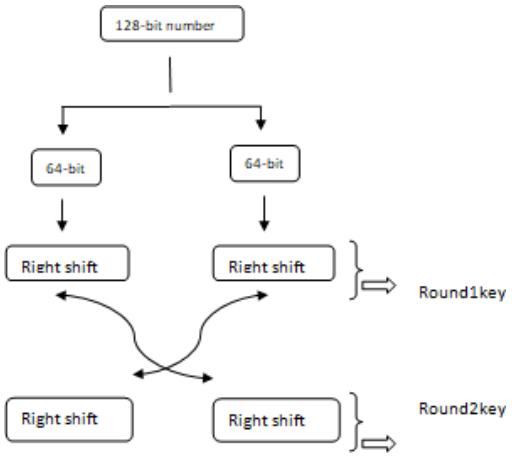


Fig 4 Key Generation

3.2 Ensuring Integrity

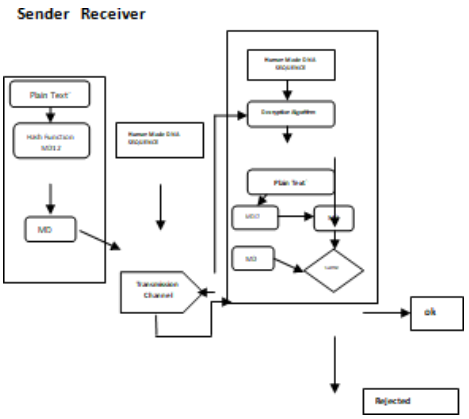


Fig-5Message Transmission Ensuring Integrity

Now to protect the data from modification, insertion, deletion, etc we have to ensure data integrity. For this, the Message digest is followed, which ensures the message's integrity. In this phase, the sender uses the MD12 as the hash function to create an MD send it along with the final ciphertext FCT to the receiver over an insecure channel. At the receiver side the received FCT is decrypted by following the reverse process of the encryption process used on the sender side and the plain text is retrieved. Then the receiver creates a new MD from the retrieved plain text using the same hash function and the hash key and compares the received MD and the new MD. If both of them are the same, then it ensures that the message has not been changed.

4. Experimental Analysis

Time Taken for DNA Encryption and Decryption Algorithm.

The Figure 3shows that time taken by encoding/decoding process and encryption/decryption algorithm.

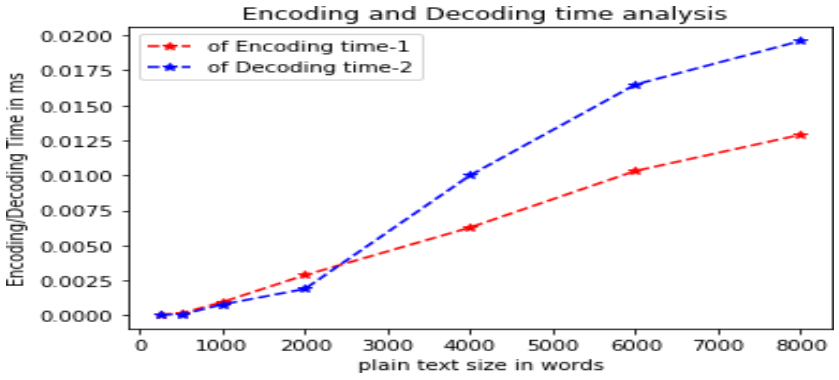
Tabel-2.Encoding and Decoding Time Analysis

Plain Text Size – in words	Time for conversion in MS	Time taken for Decoding in MS
250	0.000073	0.000055
500	0.000150	0.000097
1000	0.00958	0.000289
2000	0.002879	0.001297
4000	0.006278	0.007001
6000	0.010312	0.011473
8000	0.012894	0.014586

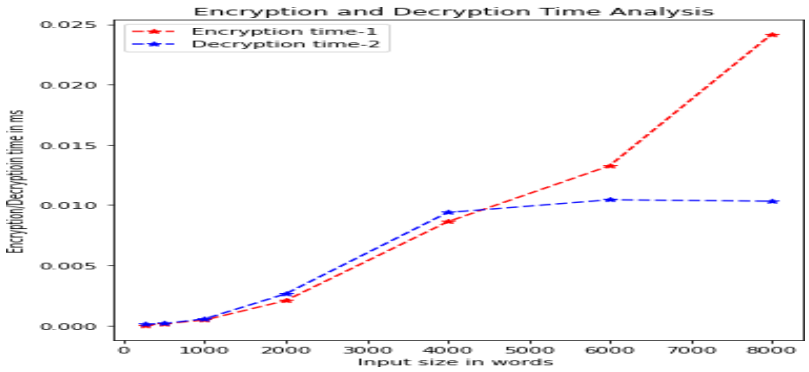
Table.-3 Encryption and Decryption Time Analysis

InputSize(word sincount)	Encryption Time (ms)	Decryption Time(ms)
250	0.000008	0.000103
500	0.000143	0.000161
1000	0.000453	0.000547
2000	0.002064	0.002624
4000	0.008603	0.009377
6000	0.013251	0.010432

8000	0.024160	0.010319



The Figure 6 shows that time taken by encoding/decoding process



The Figure 7 shows that time taken by Encryption/Decryption Algorithm.

4.1 The Frequency Analysis of Cipher Text

The Frequency analysis is that the study of the frequency of letters or groups of letters combination of cipher text. This strategy is used as an auxiliary tool for decoding numbers.This is usually because the correlation between cipher-text patterns is used as an effective tool for cryptanalysis, leading to cipher-text cracking. In order to show that the cipher-text generated using the proposed algorithm has the least correlation, thereby reducing the possibility of successful cryptanalysis, we used the two cipher-texts generated using the proposed algorithm in the following two scenarios.

- Cipher texts for 2 plain texts generated using the identical encoding tables
- Cipher texts for 2 plain texts generated using different encoding tables (because encoding tables are generated new for each interaction session between sender and receiver)

The two-cipher text relationship within the two scenarios is tested using Pearson's coefficient of correlation method. The results of the correlation analysis are that the 2 cipher texts

have weak relationships in both the above scenarios enabling reducing the chances of cryptanalysis and breaking the cipher. This can be depicted in figures 4&5.

5. Conclusion and Future Work

This article introduces a fast DNA encryption method for encryption and description. Safety and security play an important role in communication. For security reasons, this article proposes a method to quickly protect DNA using two layers. At the first level, the data is converted to the text again and other units are found. In the second level perform shifting operation and finally converted it into a human-made fake DNA sequence. At the first level, the data is converted to the text again and other units are found.

Here, we tested the encryption and decryption of different data sizes and obtained good performance.

In the future, one can transmit picture and video data into an insecure environment using the above-proposed method.

References

"A History of Cloud Computing". ComputerWeekly.

Daniela Hernandez (May 23, 2014). "Tech Time Warp of the Week". Wired.

Dr. A.Murugan "Cloud Storage Security Scheme using DNA Computing with Morse Code and Zigzag Pattern"IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI-2017)

W.Lizhe, Jie Tao, M.Kunze, A.C. Castellanos, D.Kramer, and W.Karl,"Scientific Cloud Computing: Early Definition and Experience,"

P.K.Paul and M.K.Ghose, "Cloud computing: possibilities, challenges and opportunities with special reference to its emerging need in the academic and working area of Information Science," In Procedia Engineering, vol. (23), pp.2222-2227, Jan 2012..

Kandukuri, Balachandra Reddy, and Atanu Rakshit, "Cloud security issues," In Services Computing, 2009. SCC'09. IEEE International Conference on IEEE, pp. 517-520, Sep 2009.

L. Adleman, "Molecular computation of solutions to combinationalproblems". American Association for the Advancement of Science,pp.1021-1024, 1994.

P Ravinder. "Secure communication using DNA cryptography with secure socket layer (SSL) protocol in wireless sensor networks." IOSR Journal of Computer Engineering (IOSR-JCE) , vol. 19, no. 5, 2017, pp. 20–23.

Ravinder Paspula," A Symmetric Key Encryption Method with a Reference DNA Sequence Technology", International Journal of Engineering Trends and Technology (IJETT) – Volume-41 Number-5 - November 2016*e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 19, Issue 5, Ver. IV (Sep.- Oct. 2017), PP 20-23*

Ravinder Paspula," Hidden Data Transmission with VariableDNA Technology I.J. of Electronics and Information Engineering, Vol.7, No.2, PP.96-106, Dec. 2017 (DOI: 10.6636/IJEIE.201712.7(2).06)

Snehal Javheri et al, "Secure Data communication and Cryptography based on DNA based Message Encoding", International Journal of Computer Applications (0975-8887) Volume 98- No.16, July 2014.

Sadeg S, Gougache M, Mansouri N, Drias H. An encryption algorithm inspired from DNA. Machine and Web Intelligence (ICMWI), 2010 International Conference on 2010. p. 344 - 349

Ravinder Paspula,"Transmission of Data in secure manner with DNA Sequence",APTIKOM Journal on Computer Science and Information Technologies", Vol. 5, No. 2, 2020, pp. 214~220 ISSN: 2528-2417 DOI: 10.34306/APTIKOM.J.CSIT.39

Guozhen X, Mingxin L, Lei Q, Xuejia L. New field of cryptography: DNA cryptography. Chin. Sci. Bull. 2006; 51 (12): 1413- 1420.

Yick J, Mukherjee B, Ghosal D. Wireless sensor network survey. Computer Networks, 2008; 52 (12): 2292-2330.

Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E. A survey on sensor networks. Communications Magazine, IEEE 2002; 40: 102– 114.

Chen X, Makki K, Yen K, Pissinou N. Sensor Network Security: A Survey. IEEE Communications Surveys & Tutorials 2009; 11 (2): 52-73.

Patel MM, Aggarwal A. Security attacks in wireless sensor networks: A survey. Intelligent Systems and Signal Processing (ISSP), 2013 International Conference on 2013. p. 329–33

Monikaa*, Shuchita UpadhyayaaSecure at all, "Secure communication using DNA cryptography with secure socket layer (SSL) protocol in wireless sensor networks",4thInternational Conference on Eco-friendly Computing and Communication Systems, p-808 – 813

MAJUMDAR, ABHISHEK, and MEENAKSHI SHARMA. "A NEW APPROACH TOWARDS INFORMATION SECURITY BASED ON DNA CRYPTOGRAPHY." *International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR)* 4.4: 59-68.

Wadhwani, Priyanka, Akanksha Gaur, and Vipin Jain. "Cryptanalytic JH and Blake Hash Function for Authentication and Proposed Work Over Blake-512 on C Language." *International Journal of Computer Science Engineering and Information Technology Research* 4.3 (2014): 187-198.

Aggarwal, Anupriya, and Praveen Kanth. "DNA encryption." *International Journal of Computer Science and Engineering (IJCSE)* 3.3 (2014): 51-66.

Nafea, Sally Safaa, and Mohmood Khalel Ibrahem. "Cryptographic Algorithm based on DNA and RNA Properties." *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* 7.11 (2018).

Bhar, Sreya. "Encryption Key Generation by Using Modified Hand-Geometry Based Cryptosystem to Secure SMS in Android." *International Journal of Computer Science and Engineering (IJCSE)* 4.5 (2015): 17-26.