International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 09 | September - 2022 Impact Factor: 7.185 ISSN: 2582-3930

New technology to digitize genes and generate cipher text from image and using DNA code containing random numbers

Isha Yadav (Assistant Professor at NIMS University Rajasthan Jaipur)
Umashanker

Abstract

Visual cryptography is a cryptographic technique that allows visual information (pictures, text, etc.) to be encrypted in such a way that the decrypted information appears as a visual image.DNA encryption is a new field of cryptography; DNA is used as an information carrier and advanced biological tools. Technology adapts as a skill tool. DNA is fascinating Media for data storage may be due to very large amounts of data Stored in small amounts of DNA. DNA can be used for storage Transmits information in the field of cryptography. Nevertheless It is in a primitive stage and has shown its effectiveness in the field data transfer. A lot of research has been done for that The calculation process becomes more complicated for unauthorized users. This Using DNA as a computing device Molecular technology to distribute it. By using DNA makes our method inherently unbreakable due to its random nature of DNA. DNA has four types of nucleotide bases adenine (A), thymine (T), cytosine (C), guanine (G). Until then A computer model of DNA suggests that these bases are Binary A↔00, C↔01, G↔10, andT↔11. The parallelism in DNA is enormous.

Keywords: base64_encode() function, DNA, base sequence, code, PCR hardening, random number generator

1. Introduction

"DNA cryptography can be cryptography Each letter of the alphabet is generated a new unique binary combination of four ester bases, Forms star deoxyribonucleic acid (DNA). A bit of a deoxyribonucleic acid writing system from Composite message to be encrypted, thus hold on is pushed into the traditional human fragment Deoxyribonucleic acid of equivalent length. Tips The result is dried on paper and revolves around a small dot. as it is Approximately 30 single strands of deoxyribonucleic acid Messages can be included even if billions of messages are detected Encrypted messages are unlikely to exist. " The word cipher comes from ancient Greek. Encryption is a hodgepodge of her two words: (a) Cryptography (b) grafo means "to write"; that's why, The literal meaning of cipher is "hidden writing". that is due to the very old science of encrypting messages, Sender and receiver can understand it. one cipher The Science of Using Mathematics to Encrypt and Decrypt data. Encryption allows you to store sensitive data Send information over insecure networks (Internet), make it unreadable to anyone other than the intended recipient.

2. Biological background

DNA is found in the nucleus of all human cells Or Knowledge of DNA: (together with RNA) enters the cell. Creating new proteins that all confirm our

biology It is a personality that is inherited (copied) from one generation to the next follow. Therefore, carrying the style date between generations, Therefore, it describes genetic biological behavior. These DNA molecules contain all fabric styles or the factors an organism desires for growth, development and daily life. is the main source genetic information of organisms biosphere. DNA is made up of two long strands Nucleotides arranged in a kind of helix like bases Shown in the figure below.

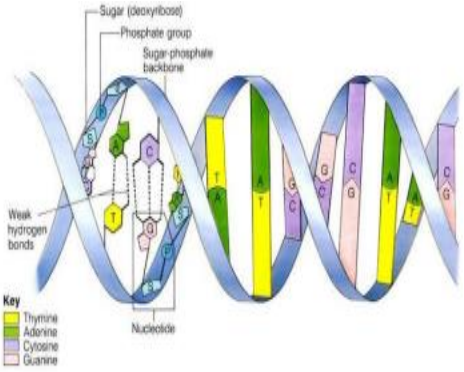

Figure 1 Basic structure of DNA.

2.1 Watson-Crick model

This is the basic and well-known DNA model. Francis HC and James D. Watson Crick derived the double helix In 1953 he elucidated the structure of DNA and was awarded the Nobel Prize in 1962. Adenine and Thymine are always paired together, Cytosine and guanine bind as a pair. Couple they are connected like the rungs of a ladder. Watson and Crick discovered that there are two sides to deoxyribonucleic acid or strands, these strands were twisted along the kind twisted ladder spiral Deoxyribonucleic acid is formed from chemical strands subunits called nucleotides that contain them all Gas based: A (A), T (T), C (C) or G (G) .

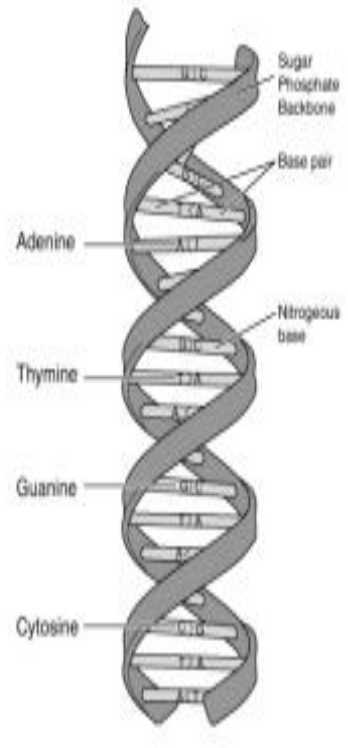

Fig 2 Watson & Crick Model of DNA

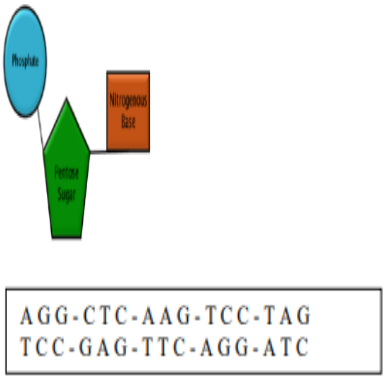

Fig.3Complimentry DNA strands

2.2 Biological manipulation

The biological model of the DNA code is the main activities distributed by the cells of our body, i.e. Integrated into the field of DNA computing. where we briefly describe the basic operations inherited

from area units of biochemical manipulation as an internal computational tool in the field of DNA encryption.

A. Synthesis

Synthesis is the design and manufacturing process. Reorganization of DNA sequence information shape. In DNA computing, design and Synthesize DNA sequence information forms are an important process. Beam design can lead to erroneous results.

B. Denaturation

A double-stranded DNA molecule is Since it is heated to 90 degrees Celsius, It is degraded into two single-stranded DNA.

C. Ligatures

DNA ligation is a process that combines the two Individual linear DNA fragments together. DNS Ligation involves creating a phosphodiester bond3'-hydroxyl groups of nucleotides and another 5'-phosphate.

D. **Hybridization**

What is hybridization Stranded DNA is formed by combining individual pieces of DNA .linear DNA sequence; joins nucleotides to complement them. In this process A always Pairs with T and G are always paired with C accordingly Add Watson Crick where appropriate condition.

E. Polymerase Chain Reaction (PCR)

PCR is a rapid amplification process amount of a particular DNA molecule in a particular set Solution using primer extension by polymerase.

F. Gel electrophoresis

Gel electrophoresis is a technique for sorting DNA. Strands based on length or weight by Agarose gel-like gels based on electric fields about the fact that DNA is negatively charged.

3 DNA code

Some of the key deoxyribonucleic acid technologies Analytics has evolved over the past few years and is gaining popularity. These techniques are polymerase chain reaction (PCR),Deoxyribonucleic acid synthesis and deoxyribonucleic acid digital encryption.

2.3 Random number generation

A random number is a "sequence or group of integers numbers that have nothing to do with each other Somewhere else in the sequence. all integers at any time they have the same probability of occurring and they occur once in a random manner.." A G G - C T C - A A G - T C C - T A GT C C - G A G - T T C - A G G - A T C

2.4 One Time Pad (OTP)

The intrinsic security of OTP (One Time Pad) is It's entirely because of the randomness of the keys. One-time pad is the only encryption method that appears this is said to be safe.

Table 1: Conversion of Binary Data to DNA Format and Vice Versa

DNA	Binary	ASCII- 7 bits Decimal	ASCII- 8bits Decimal
A	00	0	0+1=1
C	01	1	1+1=1
G	10	2	2+1=3
T	11	3	3+1=4

2.5 Diffi-Helman key exchange

Whitefield Diffie and Martien Hellman tackle Incredible reaction to a huge major agreement, or 1976 key exchange. This solution is Diffie-Hellman key exchange/negotiation algorithm.

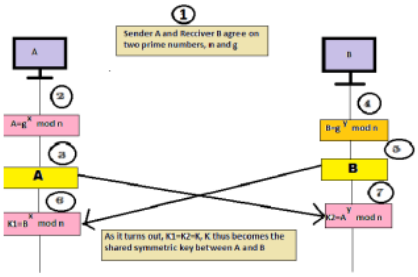

Fig 4 Diffie-Hellman key exchange illustration

4.System Design of Proposed Cryptography and decryption scheme

Suppose we have a sender A with an associated encoding. A key KA and a possible recipient B who owns the key KA Decryption key KB (KA = KB or KA ≠ KB). A uses KA to translate plaintext M into ciphertext C by translationE. B converts ciphertext C to plaintext using KB through translation M.

The encryption process is as follows.
C = ECA(M)
The decryption process is as follows.
DKB(C) = DKB(EKA(M)) = M

4.1 Rough flow of the program

Here we describe the general process of encryption and the decoding scheme is: The scheme will first be completed in three phases. Generate, second to encrypt, third to decrypt step.

4.2 Key Generation

Then design a pair of PCR primers, exchanged over secure communication channels Diffe-Hellman, we can get the encryption key K_A. A pair of PCR primers and B's public key e, and Encryption key K_B that is a pair of PCR primer and B private key d.

so we have these keys

Encryption key **K_A= first pair of PCR primers**.

Decryption Key **K_B = second pair of PCR primers**.

e is B's public key

d is B's private key.

4.3 Encryption

Original text message edit or Data preprocessing. After information preprocessing get a completely different cipher text from one equivalent plaintext that can effectively thwart the attack of potential words as PCR primers.

4.4 Decryption

Recipient B then amplifies the secret message the polymer sequence M Sand from which he can get the plaintext From Sender A From reverse pre-processing (post-processing)An operation that sacrifices his private key d. This coding however, methods are more than just mathematical calculations further biological processes.

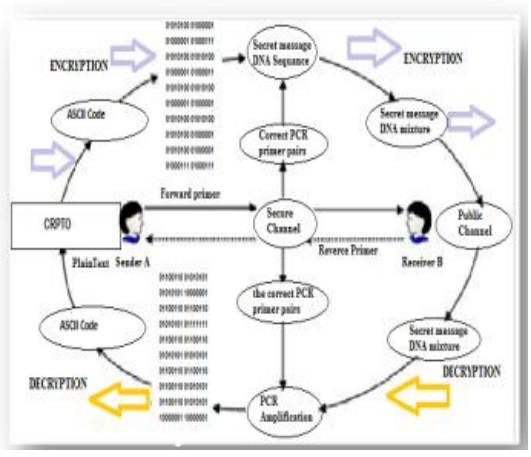

Fig 5 Flow chart of scheme

5.Proposed Algorithm

Algorithms for implementing data security in binary form the DNA sequence is Using random number generators and encryption Decryption Algorithm Based on Binary Method Rules for addition and binary subtraction.

5.1Convert An Image To base64 Encoding

The base64_encode() function is used to convert any data to base64 encoding. In order to Step1.Convert an image into base64 encoding firstly need to get the contents of the file. This can be done with the help of file_get contents () function.

Step2. Then pass this raw data to base64_encode() function to encode. Once the image is encoded with base64_encoding we use the encoded data and we use unicode2native to encrypt it.

Step3. Then a key generation process takes action.

Step4. Then Xor operation. The decryption of the Xor operation output takes place.

5.2 Encryption Algorithm

Step-1: Convert original message to ASCII code(converted to numeric format)

Step-2: Convert ASCII code to binary code (numerical value)conversion is applied)

Step-3: Convert binary code to DNA base Equivalents using Table 1 (DNA sequences)

Step-4: Generate a random number for each nucleotide of DNA sequence of the region of1-99.number if the random number is greater than 99 I need to subtract 99.

Step-5: Convert the DNA sequence from step 3 to a standard ASCII value. Then convert it to a binary value.

Step 6: Convert the random number from step 4 to binary Generate a number and generate a series of binary numbers.DNA sequence.

Step-7: Perform binary addition on the value of step 5 &step 6

Step-8: Repeat step 7 for the entire length.

Step-9: Output full binary encrypted message from original message.

Step-10: Convert the result of step 10 to ATCG format. using Table 2. This will print the encrypted message to the DNA sequence.

5.3 Decryption Algorithm

Step 1: Then we decode the Decrypted data with base64_decode() function which is used to decode data that is encoded in MIME base64.

Once these all steps are completed we are able to get the initially used image back again in the same format

Step 2: Read the sequence of the scrambled DNA mix A message containing random numbers. and isolate the random A number from the DNA sequence.

Step 3: Generate a sequence of decimal numbers for .Create a DNA sequence and generate the next sequence in parallel. Decimal number of DNA sequence using random number.

Step 4: Convert Both Decimal Numbers to Binary sequence.

Step 5: Perform Random Binary Subtraction a sequence from the binary sequence of the DNA code.

Step-6: Convert the result of Step 4 to decimal Convert to ATGC using Table 2.

Step 7: Convert the result of step 5 to a binary sequence Table 1 and .

Step-8: Perform byte 8-bit division on the result of Step 6 and convert to ASCII value and get the original text message.

6. Algorithmic Analysis

Despite some differences between Deoxyribonucleic acid and ancient codes, all of

them It has a function equivalent to encryption Or Security requirements should be based solely on confidentiality decryption key.

Three keys are used in this scheme. These are K_A and K_B and a pair of PCR primer pairs. This system offers two functions. security level. The first level is **biological difficulty** Used, the second level is the **mathematical difficulty** used.

• First stage - difficult biological collateral.

•Second stage - difficult mathematical safeguards.

6.1 First level securities

First, the other party must have sufficient knowledge biology and chemistry, and cryptography calculation. Because of this scheme, we take advantage of the complexity of difficult biological problems and manipulations. Very Difficult to amplify encrypted messages without knowledge Two correct primer pairs.

6.2 Second Level Securities

The second level of securities is mathematically difficult Problems used in traditional cryptography encryption. Our Algorithms Offer Strong Mathematical Security Comparable to similar previously available algorithms, This is because the algorithm is based on Random-One-Time Pad. This means that the private or random key is Used only once in a transaction. after transaction this key will be destroyed.

6.3 Analysis of Key Strengths

Our algorithm uses random numbers Nucleotide as key. Key strength can be calculated as follows: substitution formula. Use 4 random numbers for each nucleotide Range from 1 to 99.

The permutation formula is

$$ {}^{n}P_{r} = \frac{n!}{r!\,(n-r)!} $$

Fig 6 Permutation Formula

7. Conclusion

In this research paper, we proposed a new Algorithm scheme called " **New technology to digitize genes and generate cipher text from image and using DNA code containing random numbers**". This scheme is based on a random one-time pad key cipher.

One of the main issues when using networks is data safety. This paper focuses on data security Questions about providing secure and effective encryption and Decryption method using a random number key generation. We look at the properties of our DNA,A random key for realizing new ideas in data security. This Schema uses DNA digital coding technology DNA. Synthesis and PCR amplification, random numbers Generative and Arithmetic Operations, and Conventional Operations encryption.

Planned Algorithm Program Scheme is still valid It's far from an ideal scheme, but it has some peculiarities It has its advantages and satisfies the cryptography principle. but i hope this scheme is useful in some or specific cases Technique. A "new cipher text generation" is expected. Technology to digitize the genetic DNA code "random number" cryptographic algorithm the paper makes a positive contribution data security, communications, data protection, and Data storage and transfer efficiency.

References

1. Adleman, L., Molecular computation of solutions to combinatorial problems. Sci. 266:1021–1024, 1994.

2. Stallings, W., Network security essentials, Prentice Hall, Fourth edition, 2011

3. Delman, B.,Genetic Algorithms in Cryptography, MS Thesis 2004.

4. Mislovaty, R., Klein, E., Kanter, I. and Kinzel, W. Security of neural cryptography, Proceedings of the 2004 11th IEEE International Conference on Electronics, Circuits and Systems, 2004. ICECS 2004, 2004, pp. 219–221.

5. Anurag Roy and Asoke Nath, "DNA Encryption Algorithms: Scope and Challenges in Symmetric Key Cryptography", IJIRAE 2016

6. John H Reif, Michael Hauser, Michael Pirrung and Thomas LaBean, "Application of Biomolecular Computing to Medical Science: A Biomolecular Database System forStorage, Processing & Retrieval of Genetic Information & Material", Duke University, 2006

7. Junling Sun, "Sequence Splicing Techniques and Their Applications For Information Encryption", International Conference on Advanced Mechatronic Systems, Tokyo, Japan, September I S-21, 2012

8. V. M. M. Shyam, N. Kiran, "A novel encryption scheme based on DNA computing," In 14th IEEE International Conference, Tia, India, Dec. 2007

9. Yunpeng Zhang and Liu He Bochen Fu, "Research on DNA Cryptography", College of Software and Microelectronics, Northwestern Polytechnical University, Xi'an,China

10. Risca, V.I., DNA-based steganography. Cryptologia, Tylor and Francis. 25(1):37–49, 2001.

11. Kaur H, Ahmed J, Scaria V, Computational analysis and In-silico predictive modeling for inhibitors of PhoP regulon in S. typhi on high-throughput screening bioassay dataset., Interdisciplinary Sciences: Computational Life Sciences (a Springer SCI Journal),2016.

12. Kaur, H., Chauhan, R., Wasan, S. K. A Bayesian network model for probability estimation, Encyclopaedia of Information Science and Technology, IGI Global, Third Edition, 1551–1558, 2015.

13. Kaur, H., Chauhan, R., and Ahmed, Z., Role of data mining in establishing strategic policies for the efficient management of healthcare system–a case study from Washington DC area using retrospective discharge data. BMC Health Services Research. 12(S1):P12, 2012.

14. Chauhan, R., Kaur, R. Predictive Analytics and Data Mining: A Framework for Optimizing Decisions with R Tool, Advances in Secure Computing, Internet Services, and Applications, Springer, 73–88, 2014. 15. Kaur, H., Chauhan, R., and Alam, M.A., Spatial Clustering Algorithm using R-tree. Journal of Computing. 3(2):85–90, 2011.

16. Hermans, M. and Schrauwen, B. Training and analysing deep recurrent neural networks. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, Advances in Neural Information Processing Systems 26, pages 190–198. Curran Associates, Inc., 2013.

17. Chen, C., Xiang, H., Qiu, T., Wang, C., Zhou, Yang., Chang, V. A rear-end collision prediction scheme based on deep learning in the Internet of Vehicles, Journal of Parallel and Distributed Computing, 2017. 18. Liao, X., Yin, J., Guo, S., Li, X., & Sangaiah, A. K. (2017). Medical JPEG image steganography based on preserving interblock dependencies. Computers & Electrical Engineering.

19. Zheng, H. T., Wang, Z., Ma, N., Chen, J., Xiao, X., & Sangaiah, A. K. (2017). Weaklysupervised image captioning based on rich contextual information. Multimedia Tools and Applications, 1–17.

20. Zhang, R., Shen, J., Wei, F., Li, X., & Sangaiah, A. K. (2017). Medical image classification based on multi-scale non-negative sparse coding. Artificial Intelligence in Medicine.

21. Diffie, W., and Hellman, M., New directions in cryptography. IEEE Transaction on Information Theory. 22(6):644–654, 1976.

22. EI Gamal T., A public-key cryptosystem and a signature scheme based on discrete logarithms. IEEE Transactions on Information Theory. 31(4):469–472, 1985.

23. Borda M. & Tornea O. DNA secret writing techniques [C]. In COMM(2010), Chengdu: IEEE, June 10-12, 2010: 451–456 Table 3 Time Taken for DNA Encryption and Decryption process No. of Characters Encryption (ms) Decryption (ms) 500 0.00016401500 0.00026401450 1000 0.00035178900 0.00048178658 1500 0.00067315500 0.00077815800 J Med Syst (2018) 42:17 Page 11 of 12 17

24. Hongjun Liu, Xingyuan Wang and Abdurahman Kadir, “Image encryption using DNA complementary rule and chaotic maps”, ScienceDirect, 2012

25. Martin JAVUREK and Marcel HARAKAL’, “Cryptography And Genetic Algorithms”, Science & Military, 2016

26. Tornea, O., and Borda, M.E., DNA Cryptographic Algorithms, MEDITECH 2009. IFMBE Proceedings. 26:223–226, 2009.

27. U.Noorul Hussain, T. Chithralekha and A.Naveen Raj, G.Sathish, A.Dharani, “A Hybrid DNA Algorithm for DES using Central Dogma of Molecular Biology (CDMB)”, International Journal of Computer Applications, 2012

28. K. Li, S. Zou, and J. Xv, Fast parallel molecular algorithms for DNAbased computation: Solving the elliptic curve discrete logarithm problem over gf(2 n), Journal of Biomedicine and Biotechnology, Hindawi., vol. 2008, pp. 1–10, Apr. 2008

29. Fastest DNA Computer. Science, 2005, 308: 195

30. Roweis, S., Winfreel, E., Burgoyne, R., et al., A sticker based model for DNA computation. Journal of Computational Biology. 5(4): 615–629, 1998.

31. Tornea, O., and Borda, M.E., DNA Cryptographic Algorithms. IFMBE Proceedings. 26:223–226, 2009. 32. Goyat, S.: Cryptography Using Genetic Algorithms (GAs). In: IOSR Journal of Computer Engineering (IOSRJCE), 1(5), pp. 06- 08 Identification of Common Molecular Subsequences. J. Mol. Biol. 147, 195-197. 2012.

33. Mishra, S., and Bali, S., Public key cryptography using genetic algorithm. Int. J. Recent Technology and Engineering. 2(2):150– 154, 2013. 34. A. J. Bagnall, "The Applications of Genetic Algorithms in Cryptanalysis", School of Information Systems, University Of East Anglia, 1996.

35. Kaur, H., & Tao, X. (Eds.). ICTs and the millennium development goals: A United Nations perspective. New York, Springer, US, 2014.