IJREAT International Journal of Research in Engineering& Advanced Technology, Volume 8, Issue 4, Aug - Sep, 2020
ISSN: 2320 – 8791 (Impact Factor: 2.317)
www.ijreat.org

VG2 CIPHER - ENCRYPTION SYSTEM FOR IMAGES

Akhil Kaushik[1], Dr. Vikas Thada[2]& Dr. Jaswinder Singh[3]

[1]PhD Scholar, CSE Deptt., Amity University, Gurugram, Haryana, India, akhilkaushik05@gmail.com
[2]Assistant Professor, CSE Deptt., Guru Jambhehwar University, Hisar, Haryana, India, vthada@ggn.amity.edu
[3]Associate Professor, CSE Deptt., Amity University, Gurugram, Haryana, India, jaswinder_singh_2k@rediffmail.com

Abstract— **Cryptographyis the elementary tool now-a-days where information and data security are the key to success either for the organization or its rivals. Cryptography has undergone myriad changes in the last few decades that include Asymmetric encryption, Elliptic Curve Cryptography and Quantum Cryptography to name a few. A contemporary crypto approach that arrived at horizon and has revolutionized the security sphere is DNA cryptography. DNA cryptography surfaced in mid 90s and still under a lot of research that is bound to give robust and secure ways for confidential data. This paper details one such fine work that defines how to safeguard a digital image using DNA encryption while sending it through insecure channel. This paper also discusses the various types of possible attacks and how the proposed cipher will counter these. The analysis of the proposed work demonstrates that it is not only stronger but also faster than the existing standards.**

Keywords— *DNA; DNA Cryptography; DNA Databases; One-Time Pad (OTP; Symmetric Encryption;.*

I. Introduction

The whole world is today under lockdown because of COVID-19, but the information has always been under lockdown since the medieval age. Anyone who has got the correct information and at the right time could do wonders whether in business or personal life. However, when corporate world is considered, the information acts as a lifeline to not only win against the competitors by gaining a momentous advantageover them, but also to gain the faith of customers. Nevertheless, the information age presently is not limited to the textual data and today loads of data is usually in the multimedia form. Consequently, the traditional ciphers have lost their meaning today and prove unworthy of dealing with the modern day eavesdropping and hacks. Hence, these conventional cryptographic algorithms must be amalgamated with some novel techniques especially DNA cipher.

DNA crypts arrived in mid 90s and has been in the limelight ever since. Leonard Adleman implemented this novel approach to solve the famous Hamiltonian Path Problem in 1994, followed by Lipton, Boneh, Dunworth, etc who solved another enigmatic mathematical problem using DNA computations[1]. Needless to say, the research area became popular among the fellow researchers and even today, loads of young aspirants are using this brilliant approach for seeking answers to NP-hard and NP-complete issues. Since last few decades, DNA cryptography and steganography have gained huge popularity due to several reasons especially while security of digital images. Some experts state that it is because DNA ciphers are based on biology rather than the modular arithmetic, hence the traditional cryptanalyst approaches are unable to break these perplex modern warfare tools[2]. However, the ciphers developed so far are still considered in the infancy stage and there is huge room for enhancement.

In this paper, section 2 highlights the biological background of DNA. Section 3 deliberates the mechanism of One-time pads. Section 4 deals with the basics of symmetric encryption. Section 5 confers the proposed message encryption algorithm. The proposed method with experimental results are listed in section 6. Subsequent section 7 discusses the security and time analysis of proposed cipher. Finally, the conclusions are drawn in section 8 and the future work is also discussed.

II. Biological Background

Primarily, the Deoxyribonucleic Acid (DNA) is defined as the genetic carrier that carries out the hereditary information among genes of all the living organisms from one generation to the subsequent ones. DNA is entailed of two stretched strands of nucleotides, each containing either of the four bases: Adenine, Guanine, Cytosine and Thymine. These stretched strands of nucleotides form a double-helix structure that was identified first by Watson and Crick in 1953[3].

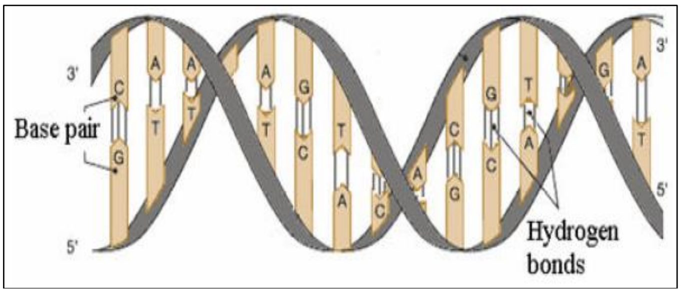


Fig. 1. Structure of DNA[4] {DNA secret writing techniques}

The nitrogenous bases are attached to each other via a hydrogen bond, where the complementary rule allows Adenine to bond with Thymine using a double bond(=), and Cytosine to combine with Guanine through a triple bond(≡). The main duty of a DNA molecule is to store the genetic information for long durations and carry it for forward to the next generation of species[5]. The DNA molecules also has instructions which are obligatory to construct other cell components like RNA molecules and proteins.

III. ONE-TIME PADS (OTP)

The One-time pad is a historical concept which introduced the timeless breed of ciphers which are provably unbreakable. It was first described by Frank Miller in 1882 but actuallypatented by Gilbert Vernam in 1919 and based on the idea of using a random key for encryption[6]. This key must possess four key properties: single usage only, length more than or equal to length of plaintext, truly random and kept completely undisclosed. Such a key is then paired (XORed) with the plaintext and the cipher text is obtained. As the key is used one time only, it is referred to as one-time pad. The word “pad” was added due to the paper pads, which were used for keys distribution. Once, the message is coded with the OTP, the OTP must be destroyed for safety reasons.

Fig. 2. A Russian One-Time Pad captured by MI5[8]

The OTP became popular in the two world wars where the Americans, British and Germans used them extensively for secure transmission of messages during wartime. Eventually, the final contributions were made by Information theorists Vladimir Kotelnikov and Claude Shannon. The former was a soviet citizen whose report remained secret (although submitted in 1941) and latter’s work was published in 1949 in Bell Labs Technical Journal[9]. The most striking feature of OTP is that it is used in pairs i.e. both communicating parties must share the OTP prior to communiqué either through a trusted channel or face-to-face. The additional copies of OTP also increase the likelihood of getting caught and revealing the secrecy of whole system to spies or eavesdroppers. Conversely, the key distribution problem can be handled easily by Quantum Cryptography, which ensures no one can eavesdrop and acquire knowledge about the keys[10]. There are other quandaries too that are associated with OTPs, like generating truly random numbers. Although variety of programming languages offer random number generation, but they are pseudo-random in particular. Hence, their usage is discouraged for cryptographic purposes, however the physical phenomenon like cosmic emissions are the only recognized ways to produce truly random numbers[11]. Though, the one-way hash functions like MD5 can offer much needed help in spawning randomness.

IV. SYMMETRIC ENCRYPTION

Symmetric or secret-key encryption is the conventional type of encoding which basically ruled the world of secrecy from 1940s to 1970s. It is also known as private-key or single-key cipher due to the fact that it uses a solitary key for encryption of plaintext on one side and decryption into plaintext on the other side. Another key aspect of symmetric enciphering is that the decryption process is reverse and identical to the encryption process. It is based on the postulation that even if the eavesdropper gains some plaintexts and their corresponding ciphertexts, he/ she is still unable to deduce the secret key or the algorithm behind it. Hence, a robust and strong enciphering algorithm is the foremost and primary prerequisite for symmetric encoding[12]. Another requirement is the secure transfer of key between the communicating parties, because if the eavesdropper detects the private key, the whole communication is compromised.

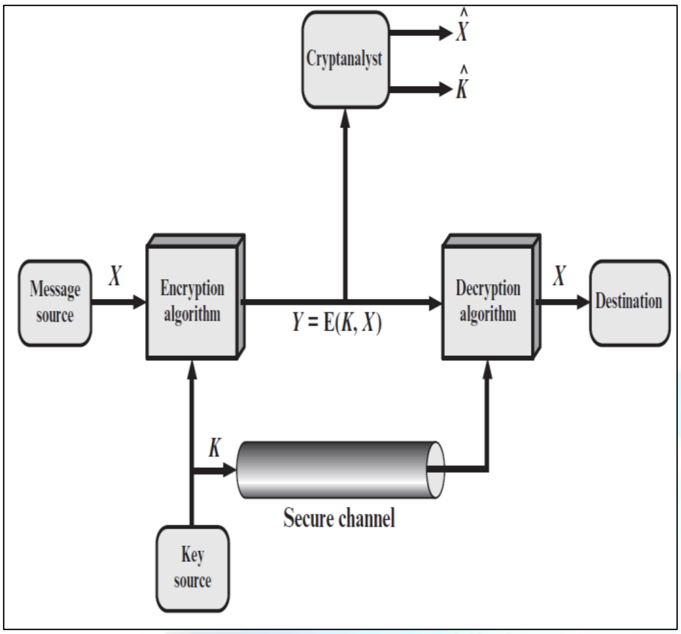


Fig. 3. Symmetric Encryption Model[13]

As shown in the image above, the ciphertext can be send over the insecure channel, but the secret key must be send secretly through secure channel. This eventually means that only the encryption key must be kept secret, but the secrecy of encryption algorithm is not mandatory. It is a big reason why numerous symmetric ciphers were implemented in hardware and their low-cost chips were manufactured. This also highlights one major feature of symmetric cipher i.e. key distribution problem, which must be handled carefully for robust message passing between users.

Another quandary in private key cryptography is the key management. As a single key is used to converse between a pair of users, the upsurge in communicating parties for a user means exponential increase in number of keys and that has to be maintained carefully[14]. For example: if a user interconnects with 10 users, he need to maintain 10 keys i.e. one for each recipient and if all 10 users interact with each other, they need (10*9)/2 or 45 keys. Now, imagine the situation of interaction among 100,000 users. It is obligatory as the chances of using the wrong key for interaction are manifold.

V. Proposed DNA Encryption Algorithm

The proposed DNA enciphering algorithm here is based on One-time pad encoding primarily and is implied for digital images here. The encoding depends on the two keys: primary and secondary key, each of different length. The primary key is applied on the pixel level and its length depends upon the bit representation of a pixel i.e. 24 bits (8 bits each of Red, Green and Blue color). This key is deduced from the random DNA sequences selected from a publicly accessible DNA databases like Genbank, FASTA, ASN.1, etc[15]. Usually codons (triplet of 3 nucleotide bases) form the primary key to match the binary representation of a pixel.

However, the secondary key is applied beforehand the primary key on a block level. The secondary key composes of 24-bits and is generated as a combination of characters (A-Z, a-z), digits (0-9) and special characters like ~, !, #, $, %, ^, &, @, etc. As the whole image is alienated into 9 blocks, there is requirement of 9 randomly produced secondary keys. As the secondary keys are randomly formed, hence they can also be called as Secondary Session Keys (as shown in Figure 8).

A. Message Encryption

Following are the steps of the encryption procedure of the proposed algorithm:

i. Generation of OTP as the primary key from publicly available DNA sequences. It will act as the session primary key.
ii. Generation of the secondary session key from the secondary key as shown in figure 4.
iii. Select the digital image to be enciphered.
iv. Divide the image into 9 blocks of equal size.
v. Shuffle the blocks randomly, so that the image is distorted.
vi. Now, for the first 3 blocks, select red component of 1st block, green color of 2nd block, blue hue of the 3rd block and transform them into binary.
vii. Apply binary operations (AND, XNOR and OR) on the output of step vi using the secondary session key.
viii. Repeat the steps vi and vii for blocks in 2nd and 3rd row also, but with different order of these binary operations.
ix. After the block level encoding, RGB contents of a pixel is converted into binary and shift-left operation is applied 10 times.
x. Then, the primary key (OTP) produced from the DNA sequences are applied on each pixel for further renovation.
xi. Finally, the output of step x can be either converted back into image.

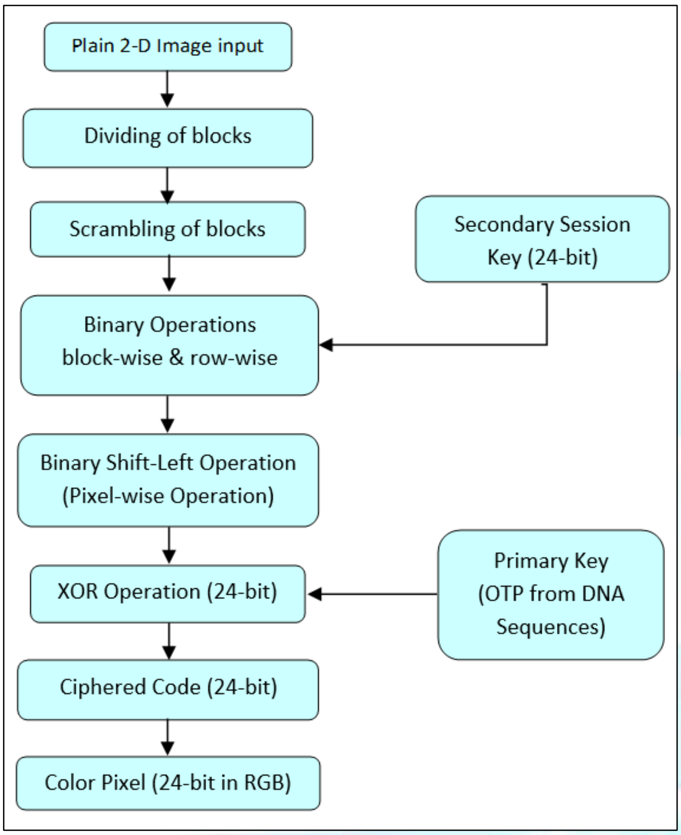

Plain 2-D Image input
Dividing of blocks
Scrambling of blocks
Secondary Session Key (24-bit)
Binary Operations block-wise & row-wise
Binary Shift-Left Operation (Pixel-wise Operation)
XOR Operation (24-bit)
Primary Key (OTP from DNA Sequences)
Ciphered Code (24-bit)
Color Pixel (24-bit in RGB)

Fig. 4. Encryption Procedure of VG2 Cipher

B. Message Recovery

As the VG2 Cipher is based on symmetric encryption, the process of decryption is exactly reverse of the encryption procedure. Before starting the deciphering over the image, the primary and secondary key must be received using secure transmission. The DNA sequence used to generate the primary key could be conveyed over the secure telephone or in physical contact. Another alternative can be the usage of secure key transfer mechanism like the Diffie-Hellman key exchange protocols.

The steps of decryption are explained as follow:

i. The ciphered image is received over the insecure channel.
ii. The primary key is identified through the same DNA sequences available publicly.
iii. The session key is also received through secure transmission and the secondary session key is calculated.
iv. The primary key is applied on the pixel level at the ciphered image.
v. Afterwards, binary shift-right operation is applied 10 times on pixel level binary code.
vi. The secondary session key is applied to do the reverse binary operations on blocks row-wise.
vii. The blocks are descrambled and unified into a single image
viii. The plaintext in the image form is obtained.

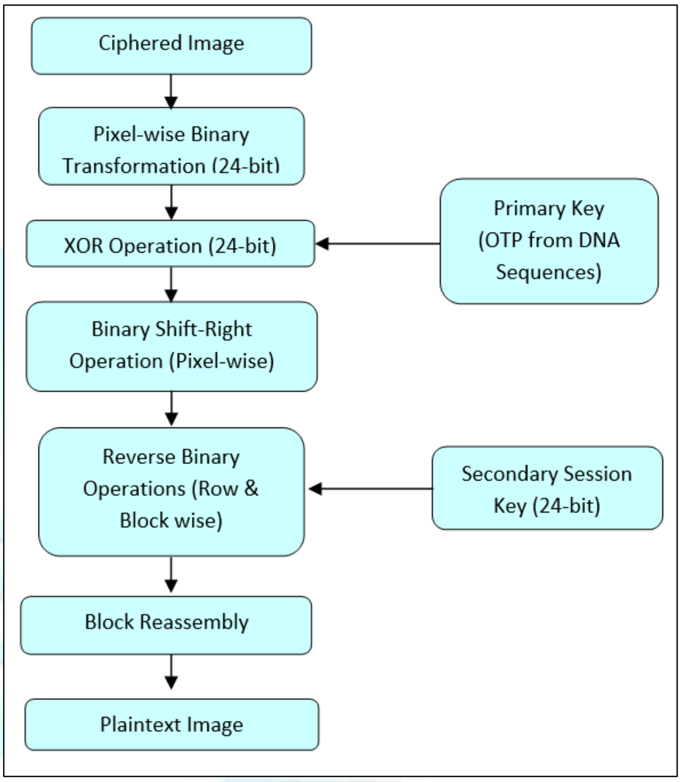

Ciphered Image
Pixel-wise Binary Transformation (24-bit)
XOR Operation (24-bit)
Primary Key (OTP from DNA Sequences)
Binary Shift-Right Operation (Pixel-wise)
Reverse Binary Operations (Row & Block wise)
Secondary Session Key (24-bit)
Block Reassembly
Plaintext Image

Fig. 5. Decryption Process of VG2 Cipher

VI. EXPERIMENTAL RESULTS

The OTP-based encoding algorithm – VG2 cipher is implemented in the Python language using Jupyter Notebook on a machine with Intel Core i5 – 10th generation processor with 8GB RAM Hewlett Packard built. Here, is an example of VG2 cipher implementation that shows the screenshots of plain-image and the produced ciphered-image.

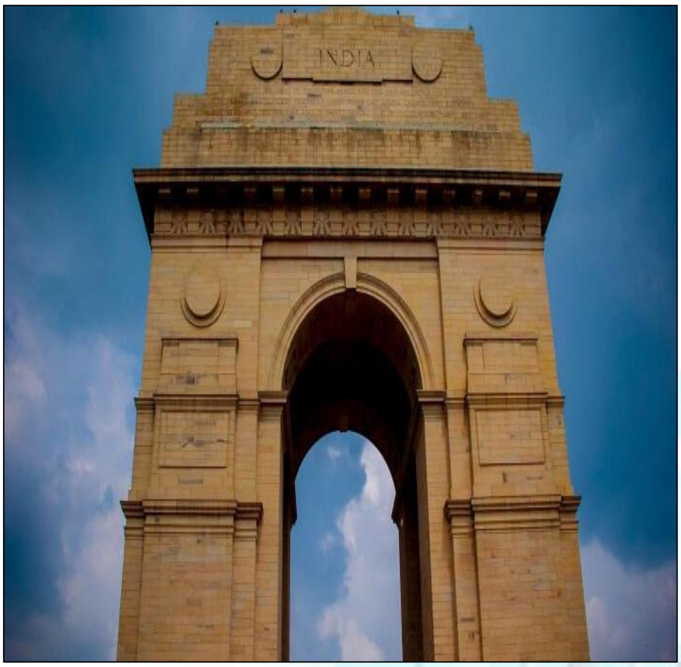


Fig.6. The Input Image for VG2 Cipher[16]

The Image is divided into 9 blocks and then scrambled to add distortion. The sample image produced after this step is:

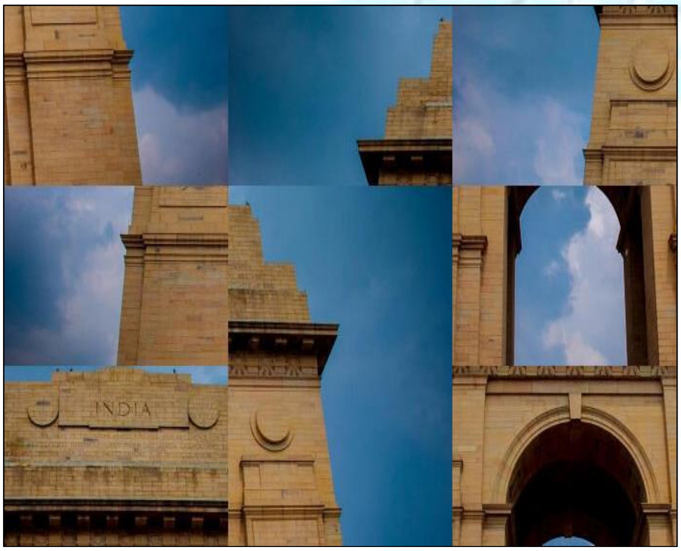
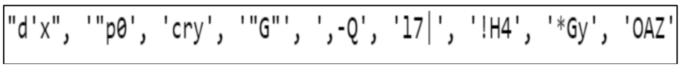


Fig.7. The Intermediate Image after Block Shuffling

The Image is then subjected to the randomly generated secondary key and consequently to primary key (OTP) created from the DNA databases. One such secondary key produced for each block is demonstrated below:



```
"d'x", '"p0', 'cry', '"G"', ',-Q', 'l7|', '!H4', '*Gy', 'OAZ'
```

Fig.8. Sample of Secondary Key Produced Randomly

The final enciphered image produced at the end of encryption process is shown below:

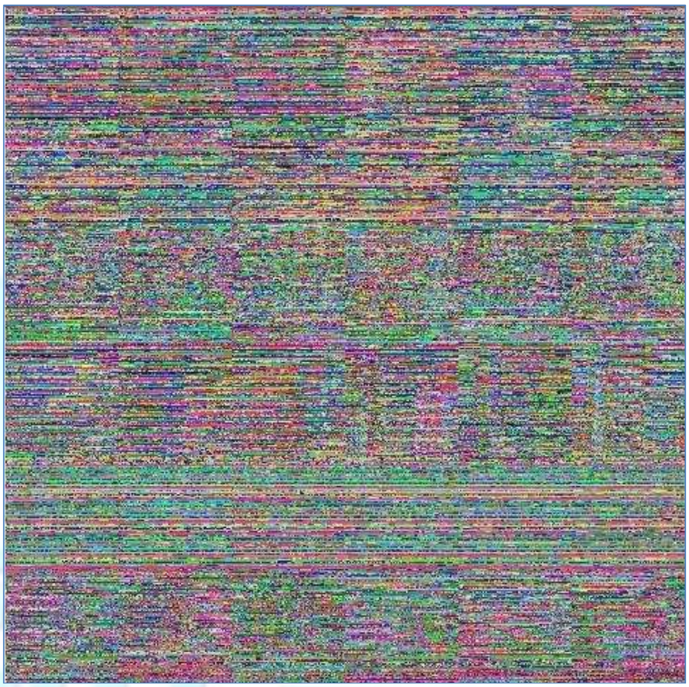


Fig 9. Final Output Image of VG2 Cipher

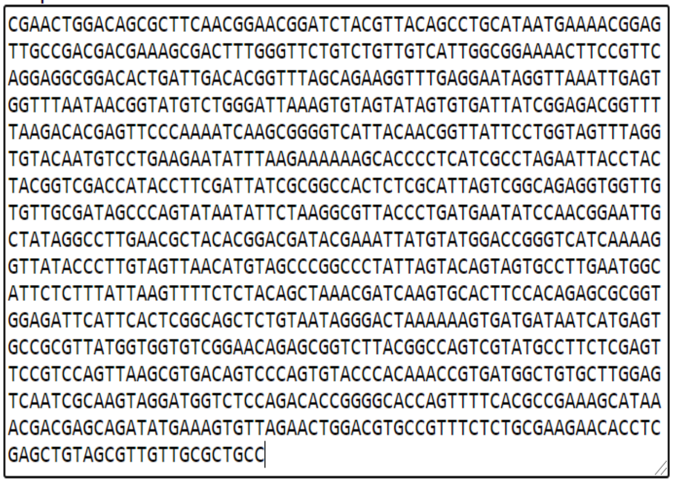
VII.PERFORMANCEEVALUATION

Cryptography is the binding technique that combines privacy and confidentiality to ensure message security between trusted users but with untrusted communication channel. For any cryptographic algorithm, the topmost striking characteristics are security and speed[17]. The equilibrium between both these features must be maintained, so that its performance supersedes the counterparts. The following section discusses both of the parameters one by one:

A. *Security Analysis*

The VG2 cipher proposed in the paper has laid the foundation of robustness both in the algorithm and the security keys. The encryption algorithm has perplex steps and encodes both at grid level and pixel level of the image, which adds another feather of mystique.

The message transmission is usually done through cables, radio channels, fiber-optics, etc. and if the encryption keys are also sent through these channels, there are copious chances of keys falling in the hands of attackers. However, the use of DNA cryptography brings in the layer of biological refuge. This extra biological layer is enormous in size thus providing colossal advantage over traditional techniques. Moreover, the primary key used in VG2 cipher is a One-Time Pad (OTP) which is shaped from the random DNA sequences available from open DNA databases like NCBI as shown in Figure 10.



```
CGAACTGGACAGCGCTTCAACGGAACGGATCTACGTTACAGCCTGCATAATGAAAACGGAG
TTGCCGACGACGAAAGCGACTTTGGGTTCTGTCTGTTGTCATTGGCGGAAAACTTCCGTTC
AGGAGGCGGACACTGATTGACACGGTTTAGCAGAAGGTTTGAGGAATAGGTTAAATTGAGT
GGTTTAATAACGGTATGTCTGGGATTAAAGTGTAGTATAGTGTGATTATCGGAGACGGTTT
TAAGACACGAGTTCCCAAAATCAAGCGGGGTCATTACAACGGTTATTCCTGGTAGTTTAGG
TGTACAATGTCCTGAAGAATATTTAAGAAAAAAGCACCCCTCATCGCCTAGAATTACCTAC
TACGGTCGACCATACCTTCGATTATCGCGGCCACTCTCGCATTAGTCGGCAGAGGTGGTTG
TGTTGCGATAGCCCAGTATAATATTCTAAGGCGTTACCCTGATGAATATCCAACGGAATTG
CTATAGGCCTTGAACGCTACACGGACGATACGAAATTATGTATGGACCGGGTCATCAAAAG
GTTATACCCTTGTAGTTAACATGTAGCCCGGCCCTATTAGTACAGTAGTGCCTTGAATGGC
ATTCTCTTTATTAAGTTTTCTCTACAGCTAAACGATCAAGTGCACTTCCACAGAGCGCGGT
GGAGATTCATTCACTCGGCAGCTCTGTAATAGGGACTAAAAAAGTGATGATAATCATGAGT
GCCGCGTTATGGTGGTGTCGGAACAGAGCGGTCTTACGGCCAGTCGTATGCCTTCTCGAGT
TCCGTCCAGTTAAGCGTGACAGTCCCAGTGTACCCACAAACCGTGATGGCTGTGCTTGGAG
TCAATCGCAAGTAGGATGGTCTCCAGACACCGGGGCACCAGTTTTCACGCCGAAAGCATAA
ACGACGAGCAGATATGAAAGTGTTAGAACTGGACGTGCCGTTTCTCTGCGAAGAACACCTC
GAGCTGTAGCGTTGTTGCGCTGCC
```

Fig. 10. Random DNA Sequences Generated as OTP Key

As the primary key is acting as OTP key (random and used one time only), it makes the attacker's hit and trial outbreak nearly impossible. Moreover, the ciphered image can be converted into DNA sequences also. The biological quandaries of DNA generated OTP and ciphertext in form of DNA sequences will create enormous overhead for the cryptanalyst, where he needs to crack both algorithm and keys to deduce the plaintext. Another point to consider here is that the secondary key is also fashioned randomly thus producing a different secondary key for 9 blocks each and they also change with every session. The security is also enhanced due to the fact that the proposed cipher is not fully dependent on either of the keys – primary or secondary. Thus the practice of using multiple keys and that too random generated with every session increases the robustness and safekeeping of the proposed cipher.

B. Timing Analysis

The proposed cipher is designed to produce faster encoding and decoding, so that the performance of VG2 cipher is not compromised. The timing for enciphering and deciphering of the recommended encoding system is given in the following table:

Table 1: Timing Analysis of VG2 Cipher

Input Size (in Pixels)	Encryption Time (in Seconds)	Decryption Time (in seconds)	Total Time for Execution (in seconds)
480 * 480	2.72	2.6	5.32
720 * 720	5.15	4.95	10.1
960 * 960	8.6	8.5	17.1

VIII. Conclusion

This paper talks a novel encrypting algorithm – VG2 cipher which is useful for image cryptography. The proposed algorithm is based on symmetric encryption, DNA cryptography and One-time pad key concept. The novel algorithm takes advantage of biological methods and their randomness is ensured due to its wide obtainability from DNA databases. The primary key is extracted from these DNA sequences and an additional secondary key is also used to add superfluous security layer that is also produced randomly. Furthermore, the encryption is done at two levels – grid level and pixel level, thus making it more multifarious. Also, the same size of plain-image and ciphered-image proves least distortion in the images. The experiments reveal that the encoding and decoding timings of VG2 cipher are also quite low, thus enhancing the performance. Security and performance analysis of the proposed cipher demonstrates ultimate offence against the attackers and still at an excellent pace. The future work includes hardware realization of encoding algorithm and extending it to audio and video files encryption.

References

[1] Kaundal, A.K. and Verma, A.K, Extending Feistel structure to DNA Cryptography. Journal of Discrete Mathematical Sciences and Cryptography, Vol. 18. No. 4, pp.349-362, 2015.

[2] Anwar, T., Paul, S. and Singh, S.K., Message transmission based on DNA cryptography. International Journal of Bio-Science and Bio-Technology, Vol. 6, No. 5, pp.215-222, 2014.

[3] Ghosh, A. and Bansal, M., A glossary of DNA structures from A to Z. Acta Crystallographica Section D: Biological Crystallography, Vol. 59, No. 4, pp.620-626, 2003.

[4] Borda, M. and Tornea, O., 2010, June. DNA secret writing techniques, in Proc. 8th International Conference on Communications, pp. 451-456, IEEE 2010.

[5] Grosse, I., Herzel, H., Buldyrev, S.V. and Stanley, H.E., Species independence of mutual information in coding and noncoding DNA. Physical Review E, Vol. 61, No. 5, pp.5624, 2000.

[6] Sherwood, A. (2013), A Brief History of One Time Pads [Online], Available: https://learn.adafruit.com/raspberry-pi-thermal-printer-one-time-pads/a-crash-course-in-one-time-pads.

[7] Zheng, G., Fang, G., Shankaran, R. and Orgun, M.A., Encryption for implantable medical devices using modified one-time pads. IEEE Access, Vol. *3*, pp.825-836, 2015.

[8] Ranum, M.J., (1995), One-Time-Pad (Vernam's Cipher) Frequently Asked Questions [Online], Available http://www.ranum.com/security/computer_security/papers/otp-faq/.

[9] Balanici, D., Tomsa, V., Borda, M. and Malutan, R., Full Duplex OTP Cryptosystem Based on DNA Key for Text Transmissions, in Proc. International Conference for Information Technology and Communications, pp. 39-48, Springer, Cham, June 2015.

[10] Upadhyay, G. and Nene, M.J., One time pad generation using quantum superposition states,in Proc. IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), pp. 1882-1886, May 2016.

[11] Atsumi, O., Hayashida, S. and Maruta, R., Unified OTP Cryptosystem with Authentication and Secrecy, Frontiers in artificial intelligence and applications, *147*, pp.287, 2006.

[12] Singh, G., A study of encryption algorithms (RSA, DES, 3DES and AES) for information security. International Journal of Computer Applications, Vol. 67, No. 19, 2013.

[13] Stallings, W., Cryptography and Network Security: Principles and Practice, 7th ed., Pearson Education India, February 2016.

[14] Ghani, A., Mansoor, K., Mehmood, S., Chaudhry, S.A., Rahman, A.U. and Najmus Saqib, M., Security and key management in IoT-based wireless sensor networks: An authentication protocol using symmetric key. International Journal of Communication Systems, Vol. 32, No. 16, p.e4139. 2019.

[15] Coordinators, N.R., Database resources of the national center for biotechnology information. nucleic acids research, 46(Database issue), D8, 2018.

[16] Methar, A. [India Gate Image]. Retrieved May 30, 2020 fromhttps://www.pexels.com/photo/brown-concrete-india-gate-789750/

[17] Wu, Y., Wu, M. and Shu, H., November. Research on Optimal Combination of Secondary Hybrid Encryption Algorithm Based on K-Means Clustering Algorithm,in Proc. International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 134-141, Springer, Cham, 2018.