An Efficient Secure Image Encryption Algorithm Using DNA Subsequence Operations

Naveen Kumar Kavartapu [#1], Dharmaiah Devarapalli [*2]

Department of Computer Science & Engineering,
Vignan's Institute of Information Technology,
Visakhapatnam, AP, India.

Abstract

This paper proposes a new Image Encryption using DNA Subsequence that uses the idea of chaotic systems to produce two Chaotic Maps, One is Logistic Map and the other is 2D logistic map, that produces secret key, which is used for encryption process. The input is an 8-bit grey image that is divided into Bit-planes using the pixel values. In this paper, we do not use biological operation to implement image encryption, but adopt the rule of DNA subsequence operation such as truncation operation, deletion operation, transformation operation and so forth, then combine DNA subsequence operation with chaos system to scramble the location and the value of pixel point from the image. A DNA subsequence operations, Complement Operations are Performed on the generated Sequence. Recombining the bit planes again we get the Encrypted image. The reverse of Encryption follows Decryption Process. By conducting several experiments on various inputs data we finally conclude that this DNA Subsequence operation is efficient in encrypting image securely.

Keywords

DNA Subsequence Operations, Chaotic and 2D Logistic map, Elongation Operation, Truncation Operation, Nucleic Acid bases.

1. Introduction

Subsequent to the discovery of DNA as the information-carrying blueprint for biopolymer assembly, the possibility has existed for its utilization to program molecular processes devised by man. DNA is an attractive material for several reasons. It provides very high information density: a micromolar solution of thousand-base DNA fragments can store 10^6 bits per femtoliter. The information is amplifiable, so that a single molecule can be copied to produce a measurable quantity of nucleic acid. A large collection of enzymatic tools (e.g., polymerases, helicases, recombinases, and restriction enzymes) and man-made tools (e.g., oligonucleotide synthesizers, thermal cyclers, and purification kits) exist to manipulate DNA. Several technologies take advantage of these facts. For example, patterned DNA fragments have been used to direct self-assembly of nucleic acid objects (Seeman 2003), to follow the fate of cells in complex populations (Shoemaker et al. 1996), to localize substrates and catalysts for "lab on a chip" experiments (Winssinger et al. 2002), and for DNA computing (Braich et al. 2002). More recently, the idea has been advanced that patterned DNAs could be used to direct small-molecule synthesis (Harbury and Halpin 2000; Gartner and Liu 2001), providing a genetic code for organic chemistry.

A fundamental difficulty in using DNA to program molecular events is transuding the information contained within a nucleic acid sequence into a corresponding physical outcome.

One general scheme to link DNA identity to a downstream process relies on sequence-specific partitioning. This self-separation is accomplished straightforwardly by hybridization of DNA molecules to immobilized oligonucleotide. Once spatially separated, the different pools of nucleic acid can be subjected to different processing steps. Thus, the sequence of a DNA fragment determines its fate. For multistep procedures, sequential hybridizations to multiple subsequences within a DNA molecule are required. Iterative partitioning of DNA molecules is equivalent to routing the molecules through a network, with each sequence taking a unique path.

The current system takes the initial Chaotic parameters and a Grey Colored image as input and encrypt that image based on DNA Subsequence operation as a Output [1]. Initially, the system takes six chaotic Parameters which is used as Secret Key and generates some chaotic sequences using 2D-logistic maps. Next, the image pixel values are converted into Binary form, and then subsequently into DNA Sequence. The Grey Colored image is encrypted by using DNA Subsequence Operations likes Elongation, Truncation and Deletion etc [2], [3]. The new implemented algorithm is Highly Secured algorithm due to large size of secret Key that restricts to exhaustive attacks.

2. Background Work

In this section we will discuss about the preliminaries that are used for implementing this new DNA sub Sequence method. Hence we will discuss the initial concepts that are required for implementing this proposed application in this section.

2.1 About Bioinformatics

Bioinformatics is the application of computer technology to the management of biological information. Computers are used to gather, store, analyze and integrate the biological and genetic information which can then be applied to gene-based drug discovery and development. The need for Bioinformatics capabilities has been precipitated by the explosion of publicly available genomic information resulting from the human genome project.

The science of Bioinformatics which is the melding of molecular biology with computer science is essential to the use of genomic information in understanding human diseases and in the identification of new molecular targets for drug discovery. In recognition of this many Universities, Government institution and many Pharmaceutical firms have formed bioinformatics groups, consisting of computational biologists and computer scientists. Such groups will be key to unraveling the mass of information generated by large scale sequencing efforts underway in laboratories around the world. Computational biology and bioinformatics are multidisciplinary fields, involving researchers from different areas of specialty, including (but in no means limited to) statistics, computer science, physics, biochemistry, genetics, molecular biology and mathematics.

2.2 Encryption

The development of information technology and the rapid growth of computer networks allowed large files, such as digital images, to be easily transmitted in open networks such as the internet. Each type of data has its own aspects, and different techniques should be used to protect confidential image data from unauthorized access. Encryption is the process of transforming the information to ensure its security. With huge growth of computer networks and the latest advances in digital technologies, a huge amount of digital data is being exchanged over various types of networks. It's often true that a large part of this information is either confidential or private demanding different security technologies to be used to provide the required protection.

Although data encryption is widely used for ensure security, most of the available encryption algorithms are used for text data. Due to large data size and real time constrains, algorithms that are good for textual data may not be suitable for multimedia data [6], [7]. Even though Triple DES and IDEA can achieve high security, it may not be suitable for multimedia application and therefore encryption algorithms such as DES, AES, RSA and idea were built for textual data. However tests in many aspects and thus requiring different encryption algorithms.

3. Our Proposed DNA Sub-Sequence Approach

In this section we will be discussing the proposed DNA Sub sequence concepts that are required for encrypting images securely.

3.1. Chaotic sequence Generation

Pseudorandom sequences are produced by chaotic maps. A chaotic system improves the security of encryption systems [5]. We introduce the following two chaotic maps, one is logistic map, and the other is 2D logistic map. In the paper, to produce the eight parameters as the initial values and system parameters of four logistic maps, we use 2D logistic map [2][3]. Logistic map is an example for chaotic map, and it is described as follows:

$$x_{n+1} = \mu x_n(1 - x_n), \quad (1)$$

Where $\mu \in [0, 4]$, $xn \in (0, 1)$, and $n = 0, 1, 2, \ldots$

2D logistic map is described in (2) [9] as follows:

$$x_{i+1} = \mu_1 x_i(1 - x_i) + \gamma_1 y^2_i$$
$$y_{i+1} = \mu_2 y_i(1 - y_i) + \gamma_2(x^2_i + x_i y_i) \quad (2)$$

When $2.75 < \mu_1 \leq 3.4$, $2.75 < \mu_2 \leq 3.45$, $0.15 < \gamma_1 \leq 0.21$and, $0.13 < \gamma_2 \leq 0.15$, the system is in chaotic state and can generate two chaotic sequences in the region [0, 1].

3.2. DNA Sequence Encryption

We know that the binary number 0 and 1 are complements, so 00 and 11 are complements, and 01 and 10 are complements. Thus we can use these four bases: *A*, *T*, *G*, and *C* to encode 01, 10, 00, and 11, respectively [3][4]. Each pixel value of the 8 bit grey image can be expressed to 8 bits binary stream which can be encoded to a DNA sequence whose length is 4. Obtaining the DNA we follow five kinds of DNA subsequence operation as elongation operation, truncation operation, deletion operation, insertion operation, and transformation operation.

A) Elongation Operation

Let P_1 be the original DNA sequence, and the subsequence P_2, of length l_1, is elongated to P_1. Then we get a new DNA sequence $P' = P_1P_2$. The expression is as follows:

$$P_1 + P_2 \longrightarrow P_1P_2. \quad (3)$$

B) Truncation Operation

Truncating the end of the subsequence P_2 from P_1P_2, we obtain a new DNA sequence $P' = P_1$. Hence, we have the sequence as

$$P_1P_2 - P_2 \longrightarrow P_1. \quad (4)$$

C) Deletion Operation

Let $P = P_3P_2P_1$. Deleting the subsequence P_2, we obtain a new DNA sequence $P' = P_1P_3$. The expression is as follows:

$$P_3P_2P_1 - P_2 \longrightarrow P_3P_1. \quad (5)$$

D) Insertion Operation

Let $P = P_3P_1$, inserting a subsequence P_2, whose length is l_2, into P. The expression is as follows:

$$P_3P_1 + P_2 \longrightarrow P_3P_2P_1. \quad (6)$$

E) Transformation operation

Let $P = P_5P_4P_3P_2P_1$. Transforming the locations of P_5 and P_3, we will get a new DNA sequence $P' = P_5P_2P_3P_4P_1$. For eg:

$$P_5P_4P_3P_2P_1 \longrightarrow P_3PP_5P_2P_1. \quad (7)$$

Since the inverse operation of elongation operation is truncation operation and the inverse operation of deletion operation is insertion operation, we have introduced five kinds of DNA subsequence operations. The insertion operation is used in the decryption process.

3.3 Proposed Methodology

The following indicates the methodology that is used for the proposed DNA subsequences in order to encrypt the images securely.

1. Generation of Chaotic Sequences

Initialize the Input state (x_0, μ_1, y_0, μ_2), by using 2D Logistic to produce eight parameters (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) by using following formulas to produce four groups of parameters:

$$x_1 = x_1,\ u_1 = 3.9 + 0.1 \times x_2,$$
$$y_1 = x_3,\ u_2 = 3.9 + 0.1 \times x_4,$$
$$z_1 = x_5,\ u_3 = 3.9 + 0.1 \times x_6,$$
$$q_1 = x_7,\ u_4 = 3.9 + 0.1 \times x_8. \quad (8)$$

Then, by using logistic chaotic map to generate four chaotic sequences under the condition that the four groups of initial values are ($x1$, $u1$), ($y1$, $u2$), ($z1$, $u3$), and ($q1$, $u4$), their length are $m \times n$, respectively.

2. Generation of DNA Subsequences

Step 1. Input A(m, n) , the original image of 8 bits , where m and n is rows and columns of the image.

Step 2. Convert image A into binary matrix A, whose size is (m, n × 8) and divide A' into eight bit-planes. Here, the first bitplanes and the eighth bitplanes, the second bitplanes and the seventh bitplanes, the third bitplanes and the sixth bitplanes, and the forth bit-planes and the fifth bit-planes are composed, respectively. Then we obtain four bit-planes.

Step 3. Carry out DNA encoding operation for the four bit planes, then we get four coding matrices P_1, P_2, P_3, P_4, with sizes (m, n).

Step 4. Convert P_1, P_2, P_3, P_4 into P_1', P_2', P_3', P_4' whose sizes are (1, (m × n)), which are divided into DNA subsequence; the average length of subsequences are $l_1 = 128$, $l_2 = 64$, $l_3 = 32$, and $l_4 = 8$, respectively.

The following are the conclusions:

$$P_1' = p_{11}\, p_{12} \cdot \cdot \cdot p_{1(mn/l_1)},$$
$$P'_2 = p_{21}\, p_{22} \cdot \cdot \cdot p_{2(mn/l_2)},$$
$$P'_3 = p_{31}\, p_{32} \cdot \cdot \cdot p_{3(mn/l_3)},$$
$$P'_4 = p_{41}\, p_{42} \cdot \cdot \cdot p_{4(mn/l_4)} \quad (9)$$

3. Deletion Operation

Step 1. Let there be a chaotic sequence X ={x_1, x_2 · · · x_{mn}/l_i }.

Step 2. If $x_i < 0.5$, delete the ith subsequence according, otherwise save the subsequence.

Step 3. Those deleted subsequences are moved to the end of the saved subsequences.

Step 4. Perform Transformation Operation

Step 5. Perform Elongation and Truncation Operation followed by complement operation.

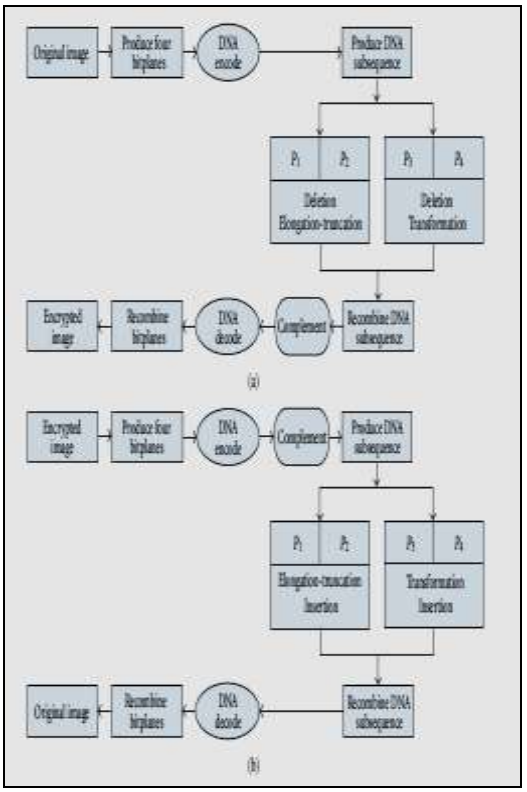


Fig. 1 Represents the Architecture of DNA Sub Sequence

4. Proposed Algorithm

In this section the proposed algorithm is explained in detail. This section mainly describes the Image encryption and Image Decryption algorithms in a secure or secret manner.

1. Encryption Algorithm

The Encryption Algorithm works as follows. It is explained in detail according to the expected output and actual input. The detailed explanation of this algorithm is

Input: An chaotic parameters (x_0, μ_1, γ_1, y_0, μ_2, γ_2) and 8-bit image A,

Output: The encrypted image B,

(1) $[S_1, S_2, S_3, S_4]$:= four DNA sequences obtained from image A;

(2) $[P_1, P_2, P_3, P_4]$:= four groups of DNA subsequences obtained from image $[S_1, S_2, S_3, S_4]$;

(3) $[x_1, u_1, y_1, u_2, z_1, u_3, q_1, u_4]$:= eight chaotic parameters obtained by 2D Logistic map under chaotic initial parameters (x_0, μ_1, γ_1, y_0, , μ_2, γ_2);

(4) $[X,Y, Z,Q]$:= four chaotic sequences obtained by Logistic map under the chaotic parameters (x_1, u_1, y_1, u_2, z_1, u_3, q_1, u_4);

(5) $A_1 = Deletion(P_1,X)$;

(6) $A_2 = Deletion(P_2,Y)$;

(7) $[E_1, E_2] = Elongation - truncation(A_1,A_2)$;

(8) $A_3 = Deletion(P_3, Z)$;

(9) $A_3' = Transformation\ (A_3, Z)$;

(10) $A_4 = Deletion(P_4,Q)$;

(11) $A_4' = Transformation\ (A_4,Q)$;

(12) $[B_1, B_2, B_3, B_4] = Recombine - subsequence(E_1, E_2,A_3',A_4')$;

(13) $[B_1', B_2', B_3', B_4'] = Complement(B_1, B_2, B_3, B_4)$;

(14) B := carry out DNA decoding and recombining binary bitplanes for B_1', B_2', B_3', B_4';

2. Decryption Algorithm

The Decryption Algorithm works as follows. It is explained in detail according to the expected output and actual input. The detailed explanation of this algorithm is

Input: The decrypted image B and chaotic parameters (x_0, μ_1, γ_1, y_0, μ_2, γ_2)

Output: The encrypted image A

(1) $[B_1, B_2, B_3, B_4]$:= four DNA sequences obtained from image B;

(2) $[B_1', B_2', B_3', B_4'] = Complement(B_1, B_2, B_3, B_4)$;

(3) $[P_1, P_2, P_3, P_4]$:= four groups of DNA subsequences obtained from image $[B_1', B_2', B_3', B_4']$;

(4) $[x_1, u_1, y_1, u_2, z_1, u_3, q_1, u_4]$:= eight chaotic parameters obtained by 2D Logistic map under chaotic initial parameters (x_0, μ_1, γ_1, y_0, μ_2, γ_2);

(5) $[X,Y, Z,Q]$:= four chaotic sequences obtained by Logistic map under the chaotic parameters(x_1, u_1, y_1, u_2, z_1, u_3, q_1, u_4);

(6) $[E_1, E_2] = Elongation - truncation(P_1, P_2)$;

(7) $M_1 = Insertion(E_1,X)$;

(8) $M_2 = Insertion(E_2,Y)$;

(9) $E_3 = Transformation\ (P_3, Z)$;

(10) $M_3 = Insertion(E_3, Z)$;

(11) E_4 = *Transformation* (E_4,Q);

(12) M_4 = *insertion*(E_4,Q);

(13) [A_1,A_2,A_3,A_4]=*Recombine subsequence*(M_1,M_2,M_3,M_4);

(14) A := carry out DNA decoding and recombining binary bit-planes for A_1,A_2,A_3,A_4;

The Proposed System will uses Simple Biological Operations and has wide key Space range is large enough to resist the exhaustive attacks. The algorithm uses six secret keys (x_0, μ_1, γ_1, y_0, μ_2, γ_2) of 256-bit are used for Encryption Process.

The scope of the system is to accept any Grey colored image as input and key, the sender will encrypt the image with the key, and the receiver will decrypt with the same Key value.

- Select/Browse the Grey Colored image.
- Take initial Chaotic Parameter Values.
- Generate Chaotic Sequences Using 2D-logistic Map.
- Encrypt the input image based on DNA Subsequence Operation.
- Use initial Chaotic Parameter Values as Secret Key.
- Authenticate the Sender and Decrypt the Encrypted image.

The proposed encryption algorithm includes three steps: first, produce four groups of DNA sequences P_1, P_2, P_3, and P_4, where P_i (i = 1, 2, 3, 4) is made up of many DNA subsequences. Then, generate chaotic sequences and DNA subsequence operations (such as elongation operation, truncation operation, deletion operation, transformation, etc.), to disturb the position and the value of pixel points from image by combining the logistic map [6]. At last, by DNA decoding and recombining bit-planes the encrypted image is obtained. The block diagram of the proposed algorithm is shown in Figure 2.We can see that the procedure of image decryption is inverse procedure of image encryption from Figure 2.

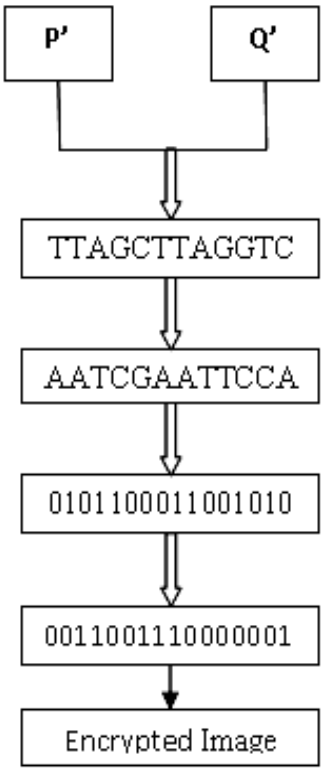


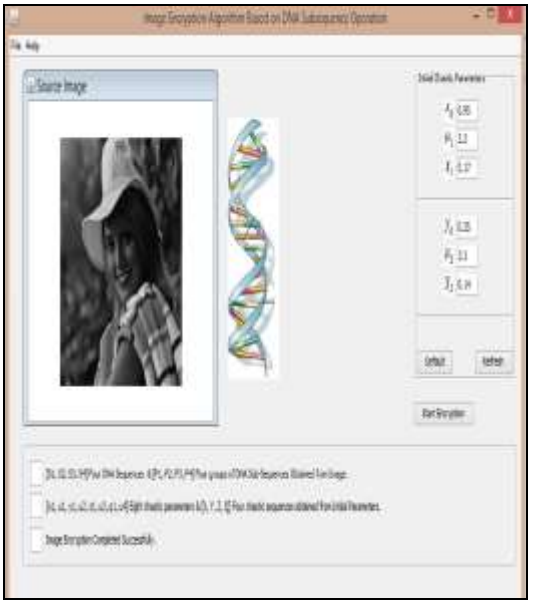
Fig. 2 Represents the Procedure of Image Encryption and Decryption

5. Experimental Results

We have conducted simulation on this DNA sub sequence approach in order to encrypt and decrypt valuable images. This simulation is done by using Java Technology as programming language in which we have used Java Swings as **Front End** User Interface. And for storing the back end images we use **My-Sql** as Back End database for storing and retrieving the image which is stored to the data base. As we are using Java as Programming language for implementing this proposed application it can run in any type of operating system despite of environment dependency. As we are using mysql as back end data base so we have been used Heidi Sql or WAMP as front end GUI tool for accessing the data base tables. This GUI tool helps the user in reducing time wastage while accessing the records.

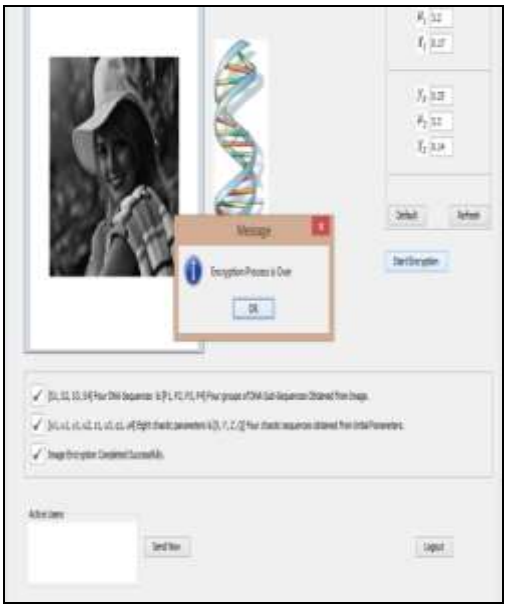
5.1 Main Window

This is the main window that represents the image is browsed in order for encryption.



5.2 Encryption Window

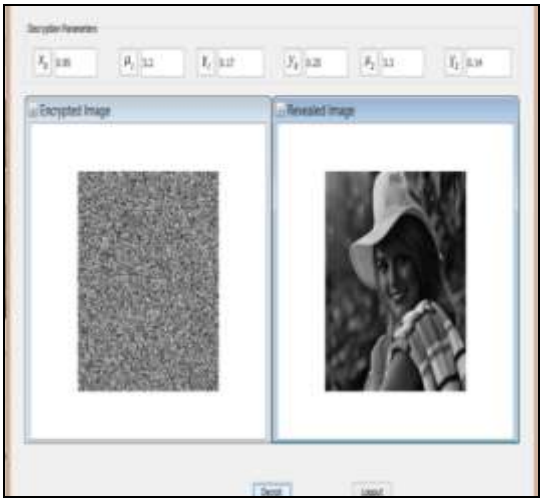
The below window clearly represents the encryption window.



In this paper, for standard 256 × 256 gray image Lena, we use Java to simulate experiment. In our experiment, we set x_0 = 0.95, μ_1 = 3.2, γ_1 = 0.17, y_0 = 0.25, μ_1 = 3.3, γ_2 = 0.14. The original image is shown

5.3 Output Revealed Image

The below window clearly represents the output revealed from the encrypted image. This clearly shows that our proposed application is very accurate in hiding valuable images through DNA Subsequence operation.



6. Conclusion

This Algorithm improves security because of using large size of key. The experimental result shows that the encryption algorithm is effective, easy to be realized, has larger key space, and is sensitive to the secret key. Algorithms for image Encryption due to different storage formats of image like (Ex: jpeg, gif, bmp… etc.) and has less Security while transfer through Network. Algorithms like DES, AES, and IDEA up to 128 – bit key only. Different types of Algorithms proposed by "Clelland" and "Gehani", but uses complex biological operations. Our algorithm can also resist statistical analysis and exhaustive attacks. Furthermore, it avoids complex biological experiment in traditional DNA cryptography. This can be applied to color pixels as wall which will be our future work.

7. References

[1] Ren H, Shang Z, Wang Y, Zhang J. A chaotic algorithm of image encryption based on dispersion sampling. Proceedings of the 8th International Conference on Electronic Measurement and Instruments (ICEMI '07); August 2007; pp. 836–839.

[2] Fu C, Zhu Z.A chaotic image encryption scheme based on circular bit shift method. Proceedings of the 9th International Conference for Young Computer Scientists (ICYCS '08); November 2008; pp. 3057–3061.

[3] Ren HE, Zhang J, Wang XJ, Shang ZW. Block sampling algorithm of image encryption based on chaotic scrambling. Proceedings of the International Conference on Computational Intelligence and Security Workshops (CIS '07); December 2007; pp. 773–776.

[4] Zhang YH, Kang BS, Zhang XF. Image encryption algorithm based on chaotic sequence. Proceedings of the 16th International Conference on Artificial Reality and Tel existence—Workshops (ICAT '06); December 2006; pp. 221–223.

[5] Lian S. Efficient image or video encryption based on spatiotemporal chaos system. Chaos, Solitons and Fractals. 2009; 40(5):2509–2519. 6. Lian S, Sun J, Wang Z.

[6] Shanshan LI, Bayi QU. A novel image encryption scheme based on chaotic sequence index. Journal of computational information systems 9: 13 (2013) 5237-5244.

[7] Musheer Ahmad, M.Shamsher Alam. A new algorithm of encryption and decryption of image using chaotic mapping. International journal on computer science and engineering, Vol.2 (1), 2009, 46-50.

[8] Shoaib Ansari, Prof.Neelesh Gupta, Prof.Sudhir agarwal. A review on chaotic map based cryptography. International journal of science engineering and technology (IJSET), Volume No.1, Issue No.4, pp:24-27.

[9] J. D. Watson and F. H. C. Crick, "Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid," *Nature*,vol. 171, no. 4356, pp. 737–738, 1953.

8. About the Authors

Mr. Kavartapu Naveen Kumar is currently pursuing his 2 Years of M.Tech (CSE) in Department of Computer Science and Engineering at Vignan's Institute of Information Technology, Visakhapatnam. He is Received B.Tech (CSE) From Laki Reddy Bali Reddy College of Engineering, Mylavaram, Krishna District; A.P, India. Current Research will be on DNA Subsequence Image encryption Algorithm (jpeg, jpg, gif and etc). His area of interests includes Bioinformatics, Data Mining and Information Security, Cloud Computing and Image processing.

Dr.Dharmaiah Devarapalli is currently working as an Associate Professor in Computer Science and Engineering department, Vignan's Institute of Information Technology, Visakhapatnam and the teaching and research experience of about 10 years. He also guided various dissertation works for both UG and PG students of VIGNAN'S IIT and other Colleges. He has taught various subjects of Computer Science and Applications for both of UG & PG students such as C programming, Data Structures, Java, Operating Systems, Compiler Design, Linux, UNIX and Bioinformatics. He has a very good reputation among the students and faculty community for his proficiency in subjects. He is a life member of CSI. He has published many papers in National, International conferences and leading Journals. His area of interest is Bioinformatics, Neural Networks, Data mining and computer networks.

