Research Article

An Information Security Using DNA Cryptography along with AES Algorithm

Varsha Kolate[1*], R.B. Joshi[2]

[1,2]Department of computer engineering, JSPM Rajarshi college of engineering, Pune, India
[1]varshakolate08@gmail.com
[2]ramjoshi.comp@gmail.com2

Article History: Received: 10 November 2020; Revised: 12 January 2021; Accepted: 27 January 2021; Published online: 05 April 2021

Abstract: Securing information is the most important need of not only the business world but also it's highly essential in all the other major sectors. The secured data storage capacity along with security during data transit is also an important factor. In this paper DNA based security technique is proposed as an information carrier, the new data securing method can be adopted by harnessing the advantages of DNA based AES. This technique will provide multilayer security. The proposed system aims to secure transactional data during communication as it is required when message or data transfer between sender and receiver should be confidential along with integrity and availability.AS the data hiding needs a carrier to hold the data, therefore in order to enhance data security and make the data more confidential effective encryption algorithm is proposed using DNA cryptography. DNA molecules, holds an ability to store, process and transfer data, stimulates the notion of DNA cryptography. This amalgamation of the chemical features of genetic DNA structures along with cryptography confirms the non-vulnerable communication. The current features with reference to DNA cryptography are reviewed and presented here.

Keywords: Information security, Time varying delay DNA cryptography, Data security, Encoding and decoding, AES.

1. Introduction

It's obvious that a new tactic to secure valuable information is required, if ecommerce and internet users would like to stay ahead of the invaders and more efficiently shield their intellectual property, files, client information and personnel then the employed strategies to secure the information must be virtuous and adequate to challenge the ever-changing data breaches. However this the scenario that demands secure ambiance for the information along with encryption of data which is static one or stored data and the data which is in transit over the network. As per the understanding from the various literatures, term cryptography stands for securing your information by writing it in some specific secret format to make it difficult to understand and retrieve the meaning just by simply reading it. The need of current generation to secure the huge amount of data produced and continuous transition over the network has raised the demand for securing this transit data from the hackers. Hence protecting the data which is static in the repository or data ware house and some data which is on the wire or transit needs to be protected is the biggest challenge for the corporate world and also for many organizations. Cryptography applies mathematical approach and techniques for securing this information as per the CIA triad of Confidentiality, Integrity and Availability. DNA cryptography is inspired from biological science. In biological science DNA is an information carrier from one generation to another. Security is concerned with the protection of information while transmitting over the network. In this paper DNA based AES algorithm is proposed to be used for the purpose of encrypting the information or data and provide protected secured original data to user. The security needs that include confidentiality, Integrity, Availability Non repudiation and Authentication are all together are implemented through this novel approach.

A: Benefits of DNA storage of data:

1)A gram of DNA contains 1021 DNA bases = 108 Terabytes of data.
2)Speed: Implement more complex crypto algorithms, it brings forward new hope to break unbreakable algorithms. This is because DNA computing offers more speed, minimal storage and power requirements.
3)Storage: DNA stores memory at a density of about 1 bit/nm3 where conventional storage media require 1012nm3/bit.
4)Power Requirements: No power required for DNA computing while the computation is taking place
5)Authenticity: Confirms that data is coming from right per- son.

1.1. Review of Literature

In [1] Author Raj, Bonny B; Sharmila and etc had suggested the use of DNA encoding methods. Use of

generic traditional approach to harness the power of cryptography along with DNA as an information carrier. In this approach the author highlighted the ability of De-oxyribo Nucleic Acid(DNA) for use as an upcoming technique. The use of DNA cryptography enhanced parallelism along with incomparable energy efficiency, storing and computing abilities.

In[2] Authors Saijisha K S and etc had implemented the amalgamation of cryptography and steganography which delivers more security for the information thorough DNA encoding method and DNA based AES algorithm .This technique will enable to encrypt the data in a very complex. Here DNA is discovered as a new transporter for securing the information during transit since it accomplishes higher protection and prevailing security with high volume and low revision rate. A novel data security scheme can be established by capti- vating the benefits of DNA based AES (Advanced Encryption Standard) cryptography and DNA steganography. This method offers multilayer security to confidential information. In this approach initially text encoded to DNA bases then DNA based AES algorithm used over it. As a final point the encoded DNA will be masked in another DNA sequence. This hybrid technique provides triple layer security to the secret message. This hybrid technique provides triple layer security using DNA based algorithm as the secret message. They mention encryption algorithm proposed is based on the combination idea of DNA encoding and AES encryption.

In [3] Author SudiptaSingha Roy and etc had proposed and explained a novel encryption methods. It is proposed using delayed chaotic neural network with a posterior DNA cryptog- raphy. The binary sequence need to a perform XOR operation with message blocks to form a key by passing it through permutation function whose dependency is over the binary sequence made from chaotic neural network. The proposed method performs superior in field of security by including DNA cryptography and ensure secure between end to end users. The supplementary DNA cryptographic approach is castoff over the cipher text acquired from the first level encryption to strengthen the security of the proposed model.

In [4] authors K.KALAISELVI and etc has proposed methods to increase the performance of convential AES and using make the existing cryptosystem more complex and stronger against attacks. In traditional cryptosystems they uses block ciphers and also use Key-dependent ciphers for securing the data were found to be weaker in terms of efficiency as they rely on the security and the speed of the algorithm. In order to strength then encryption process by making them adaptive and dynamic so that they can tackle cryptanalytic attacks. Adding confusion and diffusion is one of the way to complicate the algorithm and avert the attacks. They enhanced AES cryptosystem by employing genetic algorithm because genetic operations are perform inconsequential and benefit of this algorithm were increase efficiency. The presented complication rises the execution time of the algorithm that tends to timing attacks. This paper proposed two improved AES cryptosystem by using Genetic algorithm (GA) in SPboxes and alteration of AES by employing nonlinear neural network (NN) in SP network to enhance security in contradiction of timing attack and reduce the computational time of the offered system. Both GA and NN are used in key expansion and key distribution of the AES algorithm.

In [5] Authors Panagiot is Papadimitratos suggested that wire- less secure data communication also protocols are widely applicable. They provide lightweight end to end security and features includes are collaborative support of basic networking function such as routing and data network functions also wire- less security protocol stop undesirable parties from connecting to your wireless network. They also addressed the problem of secure and fault-tolerant communication in the presence of adversaries across a multi-hop wireless network with frequently changing topology. In this approach to commendably handle with random nasty interruption of data transmissions, authors propose and assess the secure message transmission (SMT) protocol and its substitute, the secure single-path (SSP) protocol. Amongst the noticeable characteristics of SMT and SSP is their capability to function uniquely in an end-to-end method and without limiting rules on the network conviction and security associations.

In [6] authors Md. Rafiul Biswas and etc had proposed DNA cryptographic technique which is using dynamic DNA en- coding along with asymmetric cryptosystem for performance enhancement in terms of data security. By applying the math- ematical approach to divide the plaintext in the specific format of some fixed size length of text called chunk. Apply the algorithm on each of these chunks and merge the cipher text of each using dynamic DNA encoding. They applied the concept of converting the text into ASCII equivalent then separated it to a finite one. During encryption equivalent binary is considered for DNA bases. Finally to carry out the merging operation on each chunk, sufficient random strings are produced to diffuse and confuse. Fibonacci series is used for these random strings and the security levels are enhanced. An empirical analysis carried out by using RSA, ElGamal and Paillier cryptosystems.

1.2. Proposed Methodology

In the proposed system the DNA based key is used for encryption and decryption process. As the same key is used for both the operation so this concept works for symmetric key algorithms.

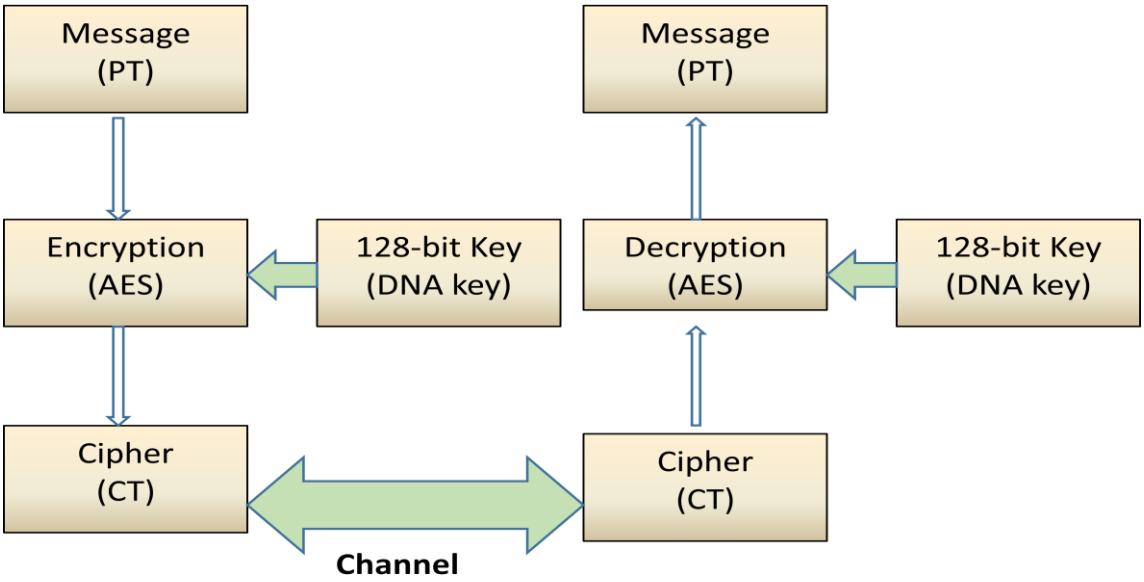


Figure 1. DNA Cryptography block diagram

DNA sequences can be represented using the DNA binary strands. Here text message of varying length begin and end with domains termed as s and e respectively which are nothing but predetermined terminators. The coded binary strands are in the form of s{0 | 1}e. Two different types of partially double stranded DNA oligo-nucleotides with sticky ends are used for representation of 0-DNA bit and 1-DNA bit. Terminator domains also have sticky ends. DNA binary strands are formed by repeated concatenation of the oligonucleotides encoding bits through the complementary sticky ends. Fig. 2 shows the DNA binary strands which are the representative of the corresponding digital binary strings.

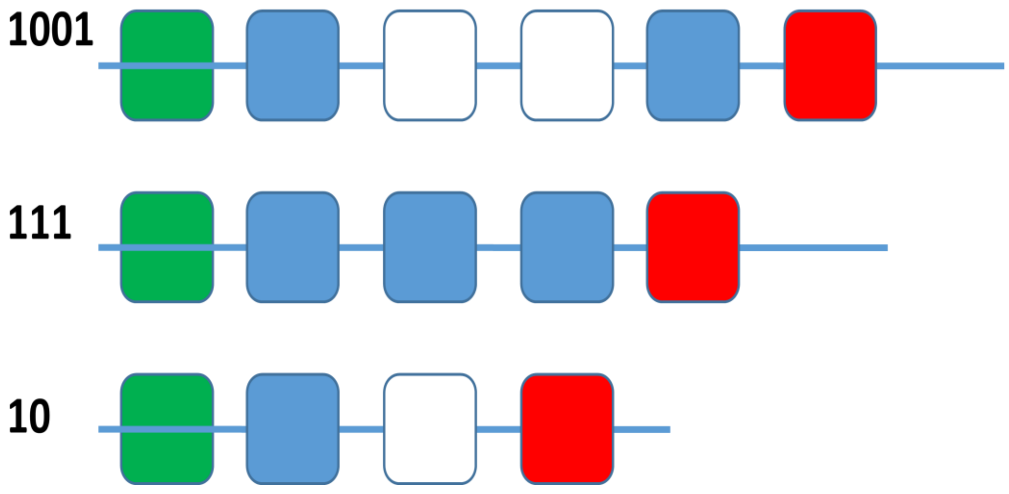


Figure 2. Encoding DNA binary strand

The cryptosystem based on DNA is explained through the following steps.

Step 1: The encryption key i.e. the unique identification sequence is shared between the sender and receiver of the secret message through secure communication channel. The key sequence can be terminator domain of the

binary strand.

Step 2: The secret message i.e. the digital binary string is encrypted in the form of DNA sequence. The key sequence is legated to the encrypted strand.

Step 3: A certain number of dummy DNA strands are generated which has similar binary format as the encrypted strand. This is because of the fact that the encrypted strand follows a particular linguistic structure (for example, English); but if the dummy strands are generated randomly, then, the adversary may take benefit of this particular dissimilarity between the encrypted strand and the dummy strands.

Step 4: The dummy strands and the encrypted strands are mixed in equimolar amounts.

Step 5: The resultant solution is sent to the intended receiver through open communication channel.

Step 6: The encrypted message can only be decoded by the receiver who knows the encryption key. Using the key sequence as one of the primers and the corresponding 0-DNA bit or 1-DNA bit as another primer PCR is performed.

Step 7: Gel electrophoresis is performed using the amplified sequences. The encrypted strand is extracted from gel and decrypted.

Though it has been assumed that the adversary has the same technical potentials as the sender and the receiver, but, the possibility to differentiate between the encrypted strand and the dummy strands by the adversary is very low. The only line of attack is to guess the key sequence or extract the encrypted strand coincidentally, which is also very rare. If the security of the proposed cryptosystem is σ, then probability of randomly selection the encrypted strand is (1-σ), which is very low.

1.3. Structure of DNA

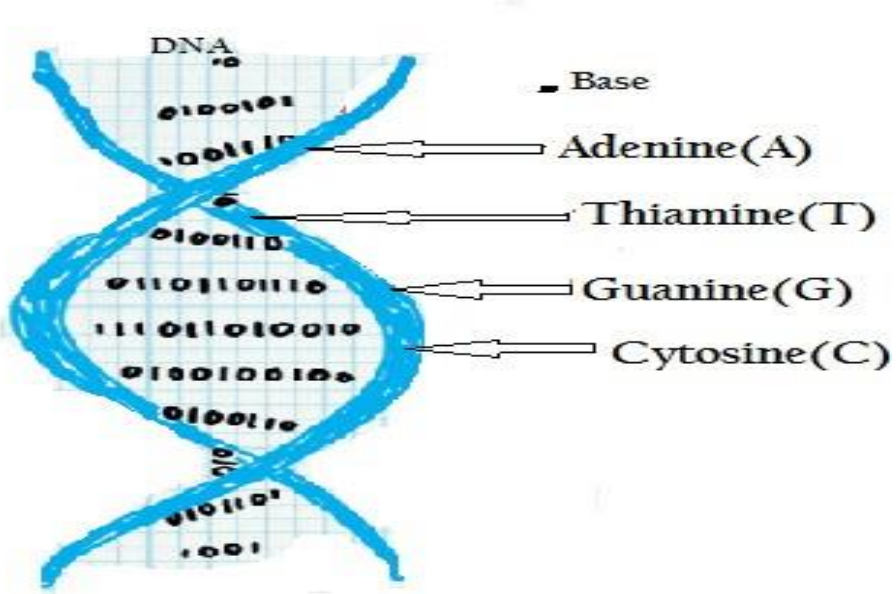


Figure 3. Structure of DNA

DNA stands for De-oxyribo nucleic acid is a thread like chain of molecules known as nucleic acid. They are used for transmitting genetic instructions which in turn is used in growth, development, functioning and reproduction of all living organism [1]. One of the major advantage of DNA molecule is that, it is a combination of four bases:

A. Table DNA digital encoding

These four bases combines in different order to form: Purines (Guanine and Adenine) and Py-rimidines (Thymine and Cytosine). These bi- strands of DNA molecules are anti parallel and they can moves in the reverse directions also DNA molecule are converted into two bit binary value[1].

1)Encryption: The plaintext is sent to encryption process and number of steps to produce DNA encrypted form.

2)Decryption: The encrypted ambiguity sequence is first encrypted using AES to require key sequence .After using this key the amino sequence is decrypted to sequence. This is converted to binary, then corresponding ASCII values.

2. Related Work

The proposed algorithm were developed by researchers not only to ensure data security but also to enhance the performance. The researchers suggested that using DNA based encryption algorithm it's possible to accomplish the goal. When DNA Based encoded data received then apply PCR amplification (polymer chain reaction). Which is often used to examine extremely small amount of sample and test the results.[1] Due to an added security features Advanced Encryption Standard (AES), usage became widespread in the field of commercial transactions, e-business also it support and provide security for wireless communication and encrypted data storage etc. AES is more safe and quicker as compare to triple DES both in hardware and software. The flexibility provided in terms of key size and number of rounds makes it more viable solution as compared to other symmetric key ciphers. There are 10 rounds for 128-bit key, 12 rounds for 192-bit key and 14 rounds for 256-bit key. Different round keys, obtained from AES key are utilized round wise. AES algorithm considers bytes for the block of data so in case of 128 bits of plain text is considered as 16 bytes.

The authors suggested that DNA based AES algorithm provide triple layer security. They also discuss about methods and steps involved in the proposed DNA encryption and decryption[2].Authors explain a cryptographic model, which is proposed for text messages by using chaotic neural network along with transmogrify delay for encryption to first step DNA cryptography[3].

A. Algorithm Explanation

DNA based AES algorithm DNA base algorithm(AES) The solution will be a series of four bases. The first priorty is an AES algorithm takes data in sections of 64 bases. The key of 128 bit [64 DNA]is used for encoding. [2]

2.1. Working of AES Algorithm-

AES algorithm support the message length of 128 bits as an input sequence and referred as a block whereas the Cipher Key for the AES algorithm is a sequence of 128, 192 or 256 bits. All input sequences are chopped into bytes. All byte values in the AES algorithm will be presented as the concatenation of its individual bit values (0 or 1) between braces in the order {b7, b6, b5, b4, b3, b2, b1, b0}. The polynomial representation for example, {01110011} identifies the specific finite field element $x^6 + x^5 + x^4 + x + 1$.

It is also convenient to denote byte values using hexadecimal notation with each of two groups of four bits being denoted by a single character as

0000->0, 0001->1……1111->f;

Hence the element {01110011} can be represented as {73}, where the character denoting the four-bit group containing the higher numbered bits is again to the left.

Bytes representations: an = {input8n, input8n+1, …, input8n+7}.

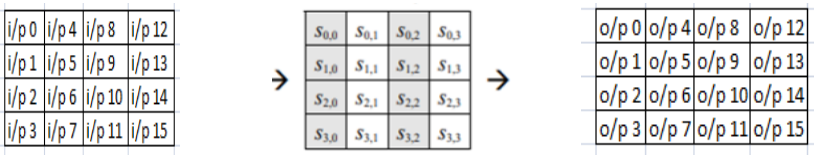
A two-dimensional array of bytes called the State is used for AES operations. The State consists of four rows of bytes, each containing Nb bytes, where Nb is the block length divided by 32. In the State array denoted by the symbol s, each individual byte has two indices, with its row number r in the range 0 £ r < 4 and its column number c in the range 0 £ c < Nb. This allows an individual byte of the State to be referred to as either sr,c or s[r,c].

Table 1. Bytes Representation

I/pBitSeq.	Byte No	Bit No.In Byte
0	0	7

1		6
2		5
3		4
4		3
5		2
6		1
7		0
==========		
8	1	7
9		6
10		5
11		4
12		3
13		2
14		1
15		0
….		….
120	15	7
121		6
122		5
123		4
124		3
125		2
126		1
127		0

At the start of the Cipher and Inverse Cipher The i/p array copied into the State array and its given as i/p to the algorithm to obtain the cipher. The Cipher or Inverse Cipher operations are then conducted on this State array, after which its final value is copied to the output – the array of bytes o/p0,o/p1, … o/p15 as shown below:-



i/p bytes **State Array** **o/p bytes**

Figure 5. State array representation

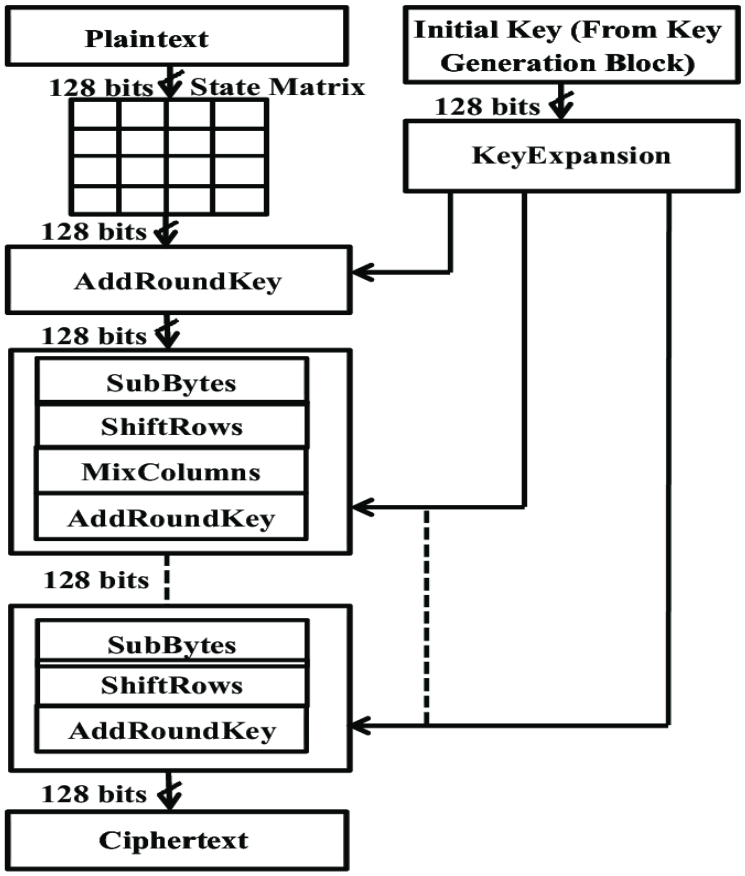


Figure 6. AES operations (Working)

In the above diagram the main building blocks are as follows:

Key Generator block: The original key of size 128 bit will be passed through this block and 11 keys are generated which will be used for the 10 rounds.

There are mainly four sub-processes used in encryption as listed in each round:

Byte Substitution (Sub Bytes): - State formation using substitution s-box.

Shift rows:- Rows shifted to Left However first one is not shifted whereas second row goes to position one to the left ,third to second and fourth to third. A new matrix of 16 bytes but shifted with respect to each other.

Mix Columns:- a special mathematical function is applied here which takes as input the four bytes of one column and outputs four completely new bytes, which replace the original column.

Add round key:- Xor operation between the state and 128 bits of the round key. During decryption process the reverse order is followed in each round –

1)Add round key
2)Mix columns
3)Shift rows Byte substitution

Since sub-processes in each round are in reverse manner

A. DNA Encryptions steps:

1)The message is first convert to digital from its ASCII numbers.

2)The ASCII numbers are grouped into blocks. Apply 4- [binary digit] processing principle to transform the data in a digital sequence of DNA.

3)The encrypted sequences of binary numbers spilt into pair like 00,01,10,11.

4)Perform exchange in four base form of the data and next transform it into DNA. Order to apply DNA processed AES encryption [2], the processed DNA form requires to be divided into sections or states (every state contains 64 DNA bases).

5)Applying DNA processed AES algorithm.

6)Transform the DNA coded text to binary digital form[2].

B. DNA decryption steps:

1)The binary form converts to coded text to DNA Also acquire the DNA digital sequence of key extract. Then perform key process.

2)The DNA coded information decrypt using with DNA bayes algorithm [AES].

3)Transformed the output acquired from the earlier step to based on DNA sequences. Previous the translation verifies whether the DNA processed form acquired can be split in DNA sequence code.

4)Execution overturns exchange in DNA form thus acquired.

5)Transform code bases of DNA with the help out ofdigital bits, then modify to binary.

6)Acquire the real data or information from ASCII number after that ASCII alteration of binary. [2]

Both encryption and decryption steps help to avoid problems

Table 2. DNA encoding

Characters	**DNA Triple**
A	CGA
B	CCA
C	GTT
D	TTG
E	GGT
…	…
…	…
V	CCT
W	CCG
X	CTA
Y	AAA

3. Implementation

The plaintext message is encrypted with the AES algorithm. The security of this algorithm is given by the computational difficulty of factoring large numbers. To be secure, very large numbers must be used as primes, 100 decimal digits at the very least. A product of such large prime numbers is an easy mathematical operation, but reverse process is a very hard task. It is extremely difficult, nearly impossible, to determine the original values of the product, at least it will take a lot of time. The encryption process uses a set of specially derived keys called round keys. These are applied along with other operations, on an array of data to be encrypted. These are the following steps of encryption for a 128-bit block.

1.Derive the set of round keys from the cipher key.

2.Initialize the state array with the block data (plaintext)

3 Add the initial rounds key to the starting state array.

4 Perform nine rounds of state manipulation.

5 Perform the tenth and final round of state.

6. Copy the final state array out as the encrypted data(ciphertext). The encrypted message with AES is a set of numerical values. These numbers will be converted using substitution in artificial DNA strand. All resulted peaces of DNA strands are bound together using a special ligase protein and the complementary strand as a template. The encrypted message can be transmitted in a compact form on DNA chip.

Algorithm steps:

Step 1: Binary data, text or image, are visualized like ASCII cod or brightness levels.

For example original message: “my secret !” in ASCII will be: 109 121 32 115 101 99 114 101 116 33.

Step 2: This numeric values are arranged in a string and taken by several digits at once,the number of digits rise together with the public keys length. In this example, we’ll take seven digits at once and obtain: 1091213 2115101 9911410 111633.

Step 3: These numbers, seven digits long will be encrypted with the public key and the result is another set of numbers:

417310496328959, 129126952185213
373906236380070, 367568882589235.
Step 4: Encrypted sequence is transformed into binary form :
417310496328959Æ 0101111011100010101010101111100101000011001111111
Step 5: Binary sequence using substitution is transformed in DNA sequence:
A – 00
C – 01
G – 10
T – 11
0101111011100010101010101111100101000011001111 1111
⇨ CCTGTGAGGGGGTTGCCAATATTTT
Step 6: All sequences are bind together in a single strand, the cipher text:
CCTGTGAGGGGGTTGCCAATATTTTCTCCC- TAAG TCGACTCGGTCCTTCCTCCCTAAGTCGACTCG-GTCC TTCCCCCAACAATCCACGCGACATGGCGCCC- CAAC
AATCCACGCGACATGGCGCCATGCATCCATAG-GTC CCTGATAT.

Decryption is a reverse process: the DNA strand is cleaved in original peaces using restriction enzymes and transformed in numerical values using the same substitution as for encryption. The last step of decryption is done using the private AES key.

4. Result

The encryption and decryption process is tested on various file sizes and it has been observed that the time required using DNA based key with AES algorithm is almost similar as the file size increase the difference in time will tends to null.

Table 3. AES and DNA based AES Encryption

I/p Msg. SIZE in (KB)	AES (ms)	AES WITH DNA	Time Diff in ms
128	**682**	**722**	**40**
256	**710**	**742**	**32**
512	**732**	**758**	**26**
960	**762**	**786**	**24**
1024	**766**	**790**	**22**

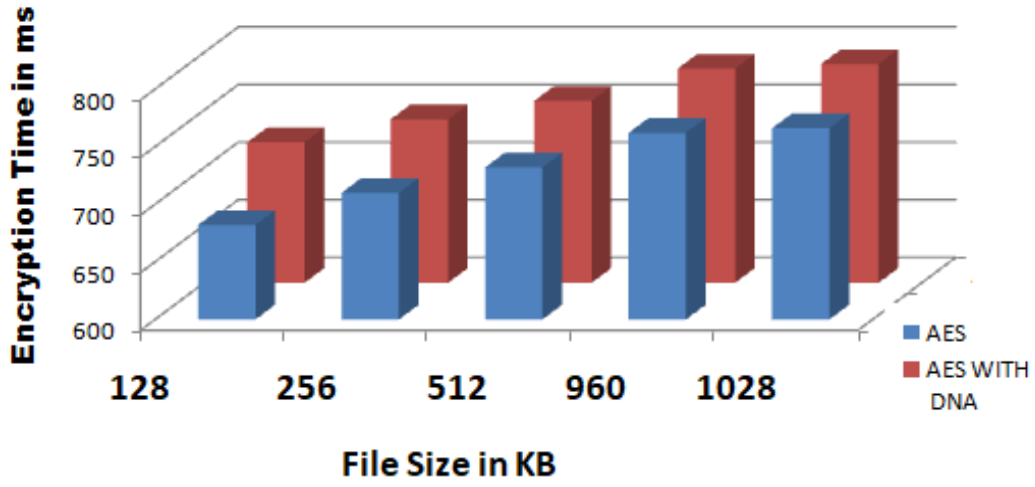


Figure 7. Comparing Encryption using Time difference Graph

Observation in the experimental setup are mainly on the file size as a input, throughput of the system, encryption and decryption time. However, System has taken many more attribute for the input purpose but here mainly the accuracy, timeliness, storage and energy cost these parameters obtained in the system. Based on attributes comparing following analytical result for our system with respect to existing system.

5. Conclusion

DNA cryptography is a favorable and fast developing arena in data security. The uses of four bases A,T,G and C for encoding the messages helps in to improve the performance in terms of parallelism and also huge capacity to store the data. .A secured DNA based cryptographic algorithms provides multi-levels of security along with DNA based AES encryption. Compression techniques can also be applied with DNA cryptography using AES. It can be used to protect sensitive data like military purposes. Main purpose of DNA cryptography is to securely share and receive business information.

6. Future Work

The big tech giants, may take an initiative to commercialize DNA computers in near future. Hopefully, in years the virtually un-hackable DNA cryptography techniques will be an effective alternative to classical cryptosystem. The security of real time information flow among the distributed network system will be area of research.

References

1. Bonny B.Raj, V. Ceronmani sharmimila," An Survey on DNA Based Cryptography" IEEE 2018 International Conference on Emerging Trends and Innovations In Engineering And Technological Research (ICETIETR) - Ernakulam (2018.7.11-2018.7.13)] 2018.
2. Saijisha K S,S.Mathew," An encryption based on DNA cryptography and steganography"IEEE 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)COIMBATORE, India (2017.4.20-2017.4.22)] 2017 .
3. S.Roy, Sudipta Singha, Shahriyar, Shaikh Akib, Asaf-Uddowla, Md, Alam, Kazi Md. Rokibul; Morimoto, Yasuhiko"A novel encryption model for text messages using delayed chaotic neural network and DNA cryptography[IEEE 2017 20th International Conference of Computer and Information Technology (ICCIT) - Dhaka, Bangladesh(2017.12.22-2017.12.24)].
4. K.KALAISELVI "Enhanced AES Cryptosystem by using Genetic Algorithm and Neural Network in S-box 978-1-5090-1936-6/16/$31.00 ©2016 IEEE.
5. Panagiotis Papadimitratos". Secure Data Communication in Mobile Ad Hoc Networks", IEEE journal on selected areas in communications 0733-8716.
6. Md .Rafiul Biswas,Kazi Md.Rokibul Alam, Ali Akber,Ya Suhiko Mori-moto"A DNA cryptographic technique based on dynamic DNA encoding and asymmetric cryptosystem" Published in 2017 4th International Conference on networking system and security(NSysS)IEEE.