PROCEEDINGS

OF

INTERNATIONAL CONFERENCE ON CYBERSPACE (I2C)

CYBERNIGERIA

2020

Think, Imaginate, Innovate and Create

MAY 2021

Copyright Page

Table of Contents

A REVIEW OF DNA CRYPTOGRAHIC APPROACHES

Mohammed Awwal Iliyasu
Department of Computer Science,
Federal University of Technology,
Minna, Nigeria.
awwaliliyasu@gmail.com

Opeyemi Aderiike Abisoye
Department of Computer Science,
Federal University of Technology,
Minna, Nigeria.
o.abisoye@futminna.edu.ng

Sulaimon Adebayo Bashir
Department of Computer Science,
Federal University of Technology,
Minna, Nigeria.
bashirsulaimon@futminna.edu.ng

Joseph Adebayo Ojeniyi
Department of Cyber Security,
Federal University of Technology,
Minna, Nigeria.
ojeniyija@futminna.edu.ng

***Abstract*— Cryptography is described as the encryption analysis of data or secret data writing using logical and mathematical data protection principles. It is an information technology for banking, medical systems, transportation and other Internet of things applications. Cryptography has become more important, and it is subjected to growing security concerns. Each system is built with its own strength in cryptography; symmetric encryption provides an economical data protection solution without compromise but it is important to share or distribute the secret key during encryption and decryption process. In comparison, the asymmetric encryption addresses the issue of secret key distribution; however, the stand-alone technique is slow and needs more computing resources than the symmetric encryption approach. In this context, a study of papers relating to DNA cryptographic approaches are presented and the research was centered from 2015 to 2020. The primary sources of information are Science Direct and Research Gate publication platforms. The existing shortcomings of DNA cryptographic approaches were established and analysis was performed on the most frequently used encryption technique based on the literature. The significant findings of this research reviewed that DNA digital coding is the most adopted cryptographic technique used to improve information security and the most common limitations of the DNA cryptographic approaches are high time complexity and algorithm complexity which is possible to infer from the literature.**

Keywords—: Cryptography, DNA cryptography, One Time Pad, DNA Cryptographic approaches, DNA limitations.

I. CRYPTOGRAPHY

[1], [2] defined Cryptography as the study for encoding and decoding of data using logical and mathematical operations to secure information. This technique has evolved rapidly in providing security to computing technology applications such as medical, financial, transportation services among other Internet of Things (IoTs) applications. Cryptography encryption is mainly intended to protect sensitive data from unwanted changes. To guarantee a safe communication, it requires encrypting the data so that if an eavesdropper successfully intercepts an encrypted message, it will not be useful because an unauthorized person cannot possibly decrypt an encrypted message [3].

Auguste Kerckhoff formulated the first cryptographic engineering principle in 1883 [4]. He asserted that encryption technique maybe known publicly but decryption of encrypted information requires knowledge of the encryption key. This key is paramount at both encoding and decoding processes respectively, and without the key, encrypting or decrypting information is impossible even if the algorithm for the encryption is known. In recent years, the encryption framework is classified generally into symmetric and asymmetric algorithms based on the key roles for each algorithm. The Symmetric Encryption Algorithm (SEA), also known as Secret Key Encryption (SKE) require the sender and recipient of an information to be in possession of a unique private key for encrypting and decrypting of information. The Asymmetric Encryption Algorithm (AES) is also known as Public Key Encryption (PKE), it requires the sender and recipient of an information to be in custody of two keys (public and private key) before encryption and decryption operations can be performed successfully on the desired information. The two encryption techniques have ensured the securing of information against adversaries and vulnerabilities on insecure communication channels [4].

Cryptanalysis is the process of decoding information from encoded or concealed format to understandable form without any slight concept on how the information is transformed from plaintext to ciphertext. Encryption is the stages involved in encoding plaintext information into ciphertext while decryption is the reversed order of encryption process where encrypted information is transformed back to plaintext. Cryptography is mainly concerned with four objectives: confidentiality, integrity, non-repudiation and authentication of an information. The information preventive measures and rules that meet one or more of the objectives are known as Cryptosystems.

The traditional Cryptography security is centered on complex mathematical problems which uses mathematical theories. In advance cryptography, algorithms that are mainly accepted as substitutes of the security techniques are the vocal, elliptic, quantum and DNA encryption algorithms. Elliptic algorithms are techniques for portable devices that have limited processing ability, it uses simple algebra and relatively small ciphers. The quantum cryptography is a technique for creating and distributing of private keys. These techniques are yet vulnerable to the Man-in-the-Middle and DoS attack. The traditional cryptographic methods provide mathematical theory models that were exposed to cipher attacks, the embracing of DNA computing in cryptographic approaches has yielded the possibility of securing communication channels on modern technologies and raises new confidence for unbreakable algorithms [5].

II. DNA CRYPTOGRAPHY

DNA is an acronym for Deoxyribose Nucleic Acid (DNA) which is a hereditary property for the entire living organisms ranging from very tiny viruses to complex human beings. It is an information agent that transports information for all life forms. The DNA consists of double helical structure with 2 strands working in antiparallel form. DNA is a lengthy polymer of small units called nucleotides. The nucleotides consist of three components each; namely: nitrogenous base, five carbon sugar and a phosphate group. The nucleotides are four types and they depend on the nitrogenous type. The four different bases are A, C, T, G called Adenine, Cytosine, Thymine and Guanine respectively. The DNA saves very complex and large volume of information for an organism with the mixture of only these four letters A, C, T and G. The bases form the structures of DNA strands through formation of hydrogen bonds with each other to keep the two strands unbroken. A and T, C and G forms hydrogen bonds with one another [6].

[7] asserted that DNA cryptography is a new rapidly evolving approach within the cryptographic domain and it focused on the DNA sequences. The notion of DNA cryptography is inspired by the DNA molecule which has the ability to store, process and transmit information. It operates on the DNA computing principle which uses 4 bases to conduct computation [6], i.e. Adenine (A), Guanine (G), Cytosine (C), and Thymine (T).

III. DNA CRYPTOGRAPHIC TECHNIQUES

The essential techniques for DNA Encryption are DNA digital coding and Polymerase Chain Reaction (PCR) Amplification. These techniques have been leveraged on by many researchers while other encryption techniques are also used for encryption operations. Below is a brief description of the DNA encryption techniques:

A. DNA Digital Coding

DNA Digital Coding is a mapping technique which uses DNA bases (ACTG) on the notion of binary digital coding to encrypt and decrypt information. The technique plays a vital role in encrypting and decrypting information, it is used to encode information using the binary digits 0 and 1. DNA Digital Coding is basically deployed on four nucleotide bases A, C, T, G [9], [5]. This technology denotes the bases A, C, T, G as 00, 01, 10, 11 and also provide means for swapping the binary values with the bases. This coding procedure forms the basis of the encryption algorithms for DNA Digital coding approaches. The computation of DNA can be accomplished in two forms: through biological operations using human DNA and the other form involves simulation using Digital DNA and Pseudo DNA. DNA Cryptography is manufactured on the basis of DNA computations for encrypting and decrypting information which has been accepted and exploited in many research areas for both Symmetric and Asymmetric encryption technique. The technique has wider coverage due to the bio-computations and security nature of the algorithm [8].

B. Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) amplification is described by [5] as a molecular biological technique of DNA Amplification which is based on the concept of Watson Crick Complementary Model. Two different primers are used for the encoding of information in this technique. Primers are the tiny DNA fragments. The key which is used for PCR amplification is generated from two primer pairs. The plaintext which requires privacy is positioned between the primer pairs to obtain the new encoded sequence. The amplification of the encoded sequence is more once the PCR primer pairs are unknown. The accuracy of the primers of the sequence is needed in this technology to prevent generating different result due to difference in primers lengths as such the actual plaintext can't be ascertained. The biological PCR operation stages are:

First stage: PCR Amplification begins with Denaturation process, a double stranded molecule is divided into two single stranded DNA. For several minute a sample is heated for about 94 to 96 degree Celsius so as to denaturalize (separate) double strand into two single strands.

Second Stage: This stage is for processing of Primer Annealing, the temperature is cooled between 50 to 65 degree Celsius in minute(s) and the primers respective complementary sequences are attached. The essence of the primers is to amplify the DNA surroundings.

Third Stage: This is Primer Extension stage, in this stage temperature elevated to 72 degree Celsius for minute(s). In this phase, the polymerase enzyme augments the shorter strands with nucleotides base on the original DNA strand. The DNA strands between primers are amplified.

DNA-based cryptography is a research method that employs biological structure to encode data. Scientists are rapidly doing research in this area which is based on using DNA computing to depict binary information in various forms. DNA encryption method involves the use of DNA sequence to convert plaintext into ciphertext. [7] has acknowledged only four methods of DNA encryption and [5] emphasized that despite PCR and DNA digital coding techniques are the most vital DNA Cryptographic techniques, there are other four DNA Encryption Techniques, which are:

C. DNA Random One Time Pad (OTP) Based

This technique operates using a randomly, ordered and unique sequence to implement a one-time pad (OTP), and once the OTP is used to generate a ciphertext, it can't be use again. This method increases information security. In this notion, the plaintext size is equivalent to OTP size [5]. DNA OTP scheme is used often to transform short plaintext or part of a plaintext message to ciphertext. The plaintext is substituted using a random and special codebook. Despite the hardware limitations on modern computers, this approach is applicable to short messages. For large size of messages, it uses DNA mapping for complexity and high execution time [10].

D. DNA Chip Based

Microarray is referred to a DNA chip, this DNA chip is designed with nucleic acid and electronic circuit and it's made of semiconductor. This technology depicts outstanding evolvement in DNA cryptography. A DNA-chip is used to store, process and uphold a high volume of genome and other biological information [6], [5]. Information is encoded using biochemical processes, the main setback of this technique is that any environmental change can result to physical factor shift which often yield negative outcome [10].

E. DNA Steganography

The word steganography originates from a Greek word “stegos” implies “cover” and “grafia” signifies “writing”, defining it as “covered writing” [11]. DNA Steganography is a method of encoding messages inside digital media such as a photograph, audio, or a video, to preserve huge volumes of information, although data can get lost as a result of sudden changes in an environment [4], [10]. The aim of steganography is to conceal secret information inside a digital media so that only the intended recipient can access the information. This technique doesn’t change the information structure, but it will conceal the information inside a digital media so than only the intended recipient can access the information [4].

F. DNA Fragmentation

DNA fragmentation is an approach that uses DNA sequence to build libraries. This technique segments the DNA sequence into bits. Often encryption algorithms uses DNA fragmentation as a second security layer and it is also applied in key encryption [10].

IV. DNA CRYPTOGRAPHIC OPERATIONS

Most biological operations can be operated on DNA molecules to help in solving mathematical and complex computational hitches. The most frequently used arithmetic and logical operations implemented on DNA are as follows.

A. Arithmetic Operations

The most basic arithmetic operations that can be imposed on DNA nucleotides as stated by [9] are:

1) Addition Operation: This operation performed addition on DNA nucleotides based on the traditional binary addition rules. For instance, addition of two binary numbers 10 and 11 will result to 11 binary digits. Furthermore, the four DNA nucleotide bases A, T, C and G are depicted as 00, 01, 10 and 11 respectively and this can be deduced that addition of C and G, will result to T.

2) Subtraction Operation: This operation involved subtraction on DNA nucleotides based on the traditional binary subtraction rules. For example, if 01 is subtracted from 11 the result is 10. Therefore, this can be deduced that subtraction of T from G, will result to C.

B. Logical Operations

The various logical operations that can be applied on DNA sequences are:

1) NOT Operation: This operation is used for inverting DNA sequences. It’s referred to as an inverter or negation operation and it’s among the simplest DNA-based logic operation. This operation required the supply of a single input while the output is the corresponding complimentary of the sequence. The output will result to true if the input supplied is false and the base combination received or supplied are the representative of true sequence. DNAs are provided to abolish any single stranded sequence. Once a double stranded sequence is detected from the input unit then the result is true, else the result is false.

2) OR Operation: In OR operation, the result is true if at least one of the input provided is true. DNA is used to terminate any single stranded sequence but once a double stranded sequence is detected then the result will be false.

3) AND Operation: The AND operation will result to true if both inputs are true. DNA will put an end to any single stranded sequence in the combination, also if double stranded sequence is noticed, it will lead to true and else it’s going to be false.

4) XOR Operation: The XOR operation provides true if and only if one of the input of the sequence is true. In binary, XOR is described as true if the input values are opposite. DNA based XOR logical operation is the simplest method because no base sequence is required or provided to the XOR operation. Opposite input sequences are termed “complementary” and will blend together to form a double stranded sequence. Once the inputs are not opposite, it will lead the sequences not to bind with each other and DNA will terminate the two input sequences.

5) XNOR Operation: XNOR operation evaluates true if the two inputs are the same, it is produced through application of NOT operation on the inputs result of the XOR operation. This operation is like the earlier logical operations, the result is true in the presence of a double stranded sequence while false in the absence of a double stranded sequence.

6) NAND Operation: The NAND operation is used to produce a true result if the inputs (i.e. two or more input) are not true. This operation is similar to the OR operation as depicted above, but the base sequence comprises of the sequence representing the true value somewhat than false. One of the inputs have to be false before it can form a double stranded sequence. The DNA will destroy any single stranded sequence in the combination if a double stranded sequence is observed, the result will be true otherwise the result will be false.

7) NOR Operation: Finally, the NOR operation produces true result when both inputs are false. This operation is implemented through application of NOT operation to the output of the OR logical operation.

V. METHODOLOGY

This study aim to provide a review of DNA cryptographic approaches proposed by various researchers, and the research was centered on the recent literatures. Two research questions were formulated in this study, these are: What are some of the major and most frequent limitation in DNA cryptographic approaches? 2. Which of the DNA encryption technique is proposed most frequent in recent researches? Three research objectives were formulated based on this research questions. The first objective is to review the most recent DNA cryptographic approaches. The second objective is to identify some common limitations of DNA cryptographic approaches. The third objective is to discover the most commonly used encryption method in DNA cryptography research domain. To achieve the first and second objectives, 18 publications on DNA cryptographic approaches were reviewed from 2015 to 2020. The Keyword “DNA cryptographic approaches” and “DNA cryptographic techniques” were searched in Science Direct and Research Gate publication sources (platforms). 36 papers related to DNA cryptographic approaches where found from both sources (platforms) within 2015 to 2020 and eighteen papers were selected randomly. In this research, the most recently proposed DNA cryptographic approaches are reviewed and their limitations were identified. The third objective was accomplished through analyzing the encryption techniques of the selected papers which are derived in tabular form. After tabulation, the most common techniques used are identified using a chart based on the selected papers reviewed and the result is displayed using a chart in section VI.

A. DNA Cryptographic Approaches

A new DNA-based encryption technique was proposed by [12]. The system combined traditional cryptography and modern methods to improve data security. The plaintext is initially converted into ASCII value, followed by binary strings. The binary strings are also translated into hexadecimal values and a 128-bit key is generated simultaneously with MD5 algorithm. The key is translated into a 32-character hexa-decimal string which is mapped to 16 dynamic values. With support of a mapping table, the binary values are encoded. After encoding, certain mathematical and logical operations occur. For data transmission, an unreliable transmission channel is used for the experiment. It is a very fast and effective technique, the algorithm is implemented on Java. This algorithm does not provide support for multilevel applications.

A new technology for safe transmission of data was developed [13] using XOR operation, One Time Pad (OTP) and DNA cryptography. Here, the OTP technique is used with an appropriate method that has certain specification. Between the OTP and the binary form, XOR operation will be applied. By denoting 00, 01, 10 and 11 for A, T, C and G respectively, then binary numbers are converted into a DNA sequence. After complementing DNA bases, the DNA sequence is reversed from right to left. Then the result of the encrypted data is send to the receiver. This algorithm offers 3 protection levels, i.e. Exclusive OR, OTP and a DNA complementary rule. Further-more, the method is very straightforward and highly safe because it is very difficult for an attacker to guess the randomly generated OTP. Since there are other prerequisite-sites, the system is not so easy to use, as users must take care of the prerequisites before choosing the OTP.

Cloud computing is becoming common due to its features, these includes economic accessibility, sharing and ease of use. However, one of the key issues is the protection offered by cloud data. [14] proposed a new DNA encryption method for enhancing cloud data security. A Symmetric Key is used for the encryption algorithm. The plaintext is initially encrypted and then translated into binary text using a key. DNA sequences are selected and converted to the appropriate cipher for the DNA base pair. Although it looks simple, the algorithm is secure but it may be vulnerable to brute force attack.

[15] presented a method which provides a safe data transmission medium using symmetric algorithm of DNA cryptography. Initially, the input data is converted to ASCII value and then it is again converted into its binary equivalent bits. This now transforms the binary value into DNA code. The resultant DNA code is assigned randomly to the ASCII code based on a private key. In conclusion, the information is encoded with the DNA code and a private key is used to conduct clinical permutation. The implementation of the system is on Java programming and is known as a modern technique for symmetric encryption. DNA chromosomes are to be used with a deep algorithm analysis during data transmission. Conclusively, this method em-ploys an encoding system more efficiently than traditional cryptographic techniques. This method can be adopted to enhance the wireless network security process. The use of DNA chromosome raises the total expense of the algorithm's implementation.

A DNA cryptographic algorithm was proposed based on one-way public key technique. The keys are obtained using ODN mixture and solid mixture for PkB and PkA respectively, denoted as public keys A and B. The plaintext is stored using one public key in a DNA sequence. In addition, both the DNA synthesizer and the remaining public key are synthesized and linked. PCR amplification is done using a coded sequence to decipher the DNA content. It is a highly safety asymmetric system, but it is very costly to deploy [16].

A modified Shamir Secret Algorithm, DNA based encryption and decryption technique was proposed by the researcher [17] which involved a group rather than a single user in the recipient end. The algorithm includes some added stability. In this technique, all clients have to be involved in the decryption process before the secrete message can be decoded. To translate the message into ASCII values, mathematical computations are performed. The ASCII values are then modified to form DNA bases. The message is sent to the entire clients involved via a group platform, then the message is decrypted with DNA encoding to improve message transmission security for multicast applications. Python and Java can be used to implement the proposed protocols and the method may be used in future in trust-based image encryption. The method is only suitable and also appropriate for one party. Furthermore the message cannot be decrypted if any client is absent from the group.

[18] proposed a technique to improve data hiding security with double sequences of DNA. The main concept behind the built framework is that secret message should be encrypted to ensure protection and robustness. The encryption takes place in two steps, the DNA reference sequence covers the encrypted message. Generally, a new data encryption algorithm which focused on DNA sequences was recommended. Hiding of data with repeated characters reduces the alteration rate of the encryption algorithm. After studying the security measures of this algorithm, the attacker can find it hard to identify the private message but yet if the intruder somehow succeeded in accessing the transmitted secret message then the message can be broken.

A new DNA algorithm technique has been proposed for encryption [19]. This encryption technique combined the XOR operations with the symmetric key exchange. The algorithm is very simple and efficient, where plaintext messages are encrypted to DNA cipher. The message will be reviewed at the end of the recipient to increase the security. Sender uses the symmetric key method to encrypt plaintext into a DNA sequence. The message is then sent via various insecure channels such as the Internet to the recipient. At the recipient unit, cipher text is obtained by decrypting the message through the means of the DNA. DNA hybridization principles and matrix computation are performed to reduce the complexity of the running time. DNA sequences have the ability to store large volume of messages in solid form, which is one of these algorithms' key advantages. Application of this approach is extremely uneconomical.

DNA computer-based cryptography algorithm is proposed for encryption and decryption of plaintext message [20]. The algorithm consists of two stages, namely: encryption and decryption. In the first stage, data is converted and then sends to the receiver in an appropriate ciphertext form. The code is deciphered to the original data at the receiver end. The plaintext is provided through PS 2 keyboard to the Field Programmable Gate Array (FPGA). The message is interpreted with FPGA as codes values in ASCII form. Then a table of codons given by the researcher is used to convert the ASCII values. For encryption of codon, Vigenere cipher processing method is used. The algorithm contributes a notion of a symmetrical key for double layer security. The main distribution is not mentioned here, which means that the algorithm might have a hectic problem.

[21] proposed a unique procedure for the production of ciphertext and a new key generation method. There are two rounds in the key generation process. An intermediate ciphertext is produce using traditional cryptographic technique and the intermediate cipher is converted into the DNA cipher text finally. The approach misled an attack with an invented false DNA sequence. The algorithm increases time and space complexities unreasonably because the application requires a single security layer.

In [22], a proposed biotic pseudo DNA encryption system, for splicing, a device is used in the technique to improve encryption algorithm security. A random technique is used to generate the key of the algorithm, which increases its degree of uncertainty, making it difficult to decipher the resulting ciphertext. The method is robust and analysis demonstrates that the method is more secured from common ciphertext attacks. High-tech bio-computing labs is required to implement this algorithm.

[23] have proposed a new approach for developing a hybrid DNA encryption technique. The technique consists of traditional encryption of DNA and Elliptic Curve Cryptography (ECC). The plaintext is first translated to ASCII and then to binary. A DNA nucleotide is derived from publicly available sender-and-receiver sequences. The DNA encoding scheme transforms these bases into binary form. Many pairs of binary numbers are generated, all of which are combined to produce a long binary number. Encoding is achieved through several tables provided by the researcher. The Koblitz method is used as an elliptical curve points for converting the decimal numbers. With the help of the ECC encryption these points are again encrypted at another elliptical curve point. The encrypted points are ciphertext points sent to the receiver. This hybrid approach of DNA ECC is safer than the existing techniques of DNA encryption. This has a limited key size and two layers of protection simultaneously. On FPGA-based embedding framework, the method proposed can be achieved. Due to small key size, the algorithm may not be secured much in terms of brute force attack.

[24] proposed a modern and secure encryption algorithm based on DNA which uses big data. An unauthorized person can access the message(s) ciphertext in this system without it being necessary that no one can read or understand this cipher. Using big data, this algorithm is used to encrypt a great deal of data. The encryption process employs DNA encoding table and PHP language in this system. This algorithm solved important data problems and the study of big data.

[25] introduced a new cryptographic technique system that focus on encrypting client-side data until they are processed on the cloud. This is a symmetric key cryptography scheme that uses cryptography based on DNA. In addition to presenting the detailed nature of this approach and contrasting it with the existing symmetric-key algorithms (DNA, AES, DES, and Blowfish), the experimental results show that this method leaves behind the conventional algorithms based on ciphertext size, encryption time, and transmission. This new method is therefore much more effective, and performs better.

[26] proposed cryptography based on DNA, which uses hamming code and a block cipher to protect a key. Its critical symmetric cryptography, used to refine a technique based on DNA. The maximum-length matching technique was also developed in this technique to protect against various attacks.

[27] proposed a scheme where ECC was given the DNA mapping technique. In this system the DNA code is random, and alphabets are distributed with non-repetitive subsections. These alphabets are then used at the two ends for encoding and decoding. This system was successfully used in the internet of things apps and used in real-time but not reliable due to algorithm complexity.

[28] suggested a DNAA mapping technique using Elliptic Curve Cryptography. This technique adopted random and non - repetitive allotted of subsections to alpha-bets. At both encryption an decryption end, the alphabets are used for encryption and decryption of information. This method has been deployed and used effectively in real - time internet of things devices.

[29] proposed a form of encryption that has two rounds. This scheme is the same as the latest technique called the algorithm Data Encryption Standard (DES). In this step, the plaintext is encoded using two keys. These two keys consist of the Elliptic Curve Cryptography (ECC), and the Gaussian Kernel Function (GKF) and another key is generated on the second characters replicated in the first key based on random injective mapping. Finally, in the second DNA sequence, the encryption message arbitrarily hides, based on GKF numbers.

VI. RESULT AND DISCUSSION

After achieving the first objective in the previous section, a summary of the reviewed DNA cryptographic approaches are presented in this section. Table 1 present a tabular form of the summary. The detailed summaries consist of research year, algorithm used for the approach, cryptographic method, DNA encryption technique and the limitation of the research.

Table 1: Summary of the DNA Cryptographic Approaches

S/N	Publication year	Algorithm used	Cryptographic method	Technique used	Limitations
1	2016	DNA based, MD5 algorithm [12]	Symmetric cryptography	DNA Digital coding	Lack support for multi-level application
2	2016	XOR, OTP, DNA complimentary rule [13]	Symmetric cryptography	OTP based	Not flexible to implement due to algorithm complexity

3	2016	DNA encryption [14]	Symmetric cryptography	DNA Digital coding	Vulnerable to brute force attack
4	2016	DNA encryption [15]	Symmetric cryptography	DNA Digital coding	High implementation cost
5	2015	DNA encryption [16]	Symmetric cryptography	PCR Amplication	High implementation cost
6	2016	DNA encryption [17]	Symmetric cryptography	DNA Digital coding	Message can't be decrypted
7	2015	DNA encryption [18]	Symmetric cryptography	DNA Steganography	Not secured once a third party is aware of the message medium
8	2015	XOR, DNA Hybridization, Matrix computation [19]	Symmetric cryptography	DNA Digital coding	High execution time, algorithm complexity
9	2016	DNA based, Vigenere [20]	Symmetric cryptography	DNA Digital coding	Inappropriate documentation for implementation, algorithm complexity
10	2017	Traditional algorithm, DNA based [21]	Symmetric cryptography	DNA Digital coding	High time and space complexity
11	2016	OTP. DNA based [22]	Symmetric cryptography	Pseudo DNA	Required high tech bio-computing for implementation
12	2015	Traditional encryption, DNA ECC [23]	Asymmetric cryptography	DNA Digital coding	Vulnerable to brute force attack
13	2015	DNA based [24]	Symmetric key cryptography	DNA digital coding	High time complexity
14	2018	DNA based [25]	symmetric-key cryptography	Binary DNA	High time complexity
15	2017	DNA based Cryptography [26]	Symmetric key cryptography	Hamming code and a block cipher mechanism	High time complexity
16	2017	DNA based cryptograph	Symmetric key cryptograpy	DNA Digital coding	Algorithm complexity
17	2018	DNA - Based ECC for IoT Devices [28]	Asymmetric key cryptography	DNA Elliptic curve cryptography (ECC)	High time complexity
18	2019	Artificial DNA sequences based on Gaussian kernel function(GKF) [29]	Asymmetric key cryptography	ECC, and Gaussian kernel function (GKF) cryptography	High time complexity

Figure 1 present the DNA cryptographic approaches limitations and their frequency.

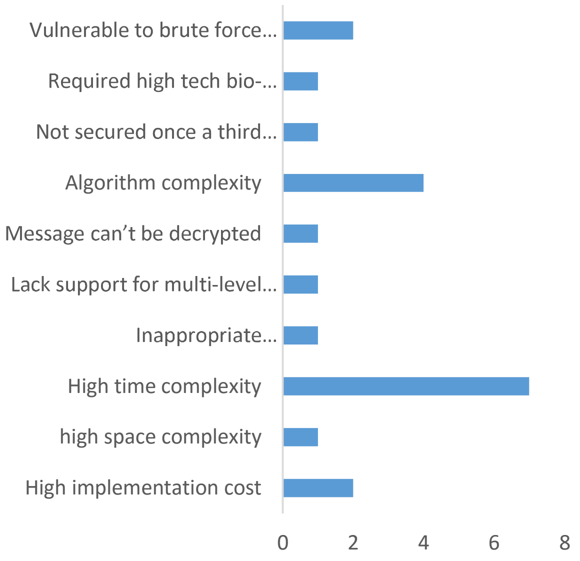

Figure 1: Summary of reviewed DNA cryptographic approaches limitations and their frequency

As depicted in figure 1, this study has investigated and identified high time complexity and algorithm complexity as the most common limitations of DNA cryptographic approaches. Other limitations are high implementation cost, unreasonable expansion of encrypted messages, Lack of support for multi-level application, algorithm vulnerable to brute force attack, and so on.

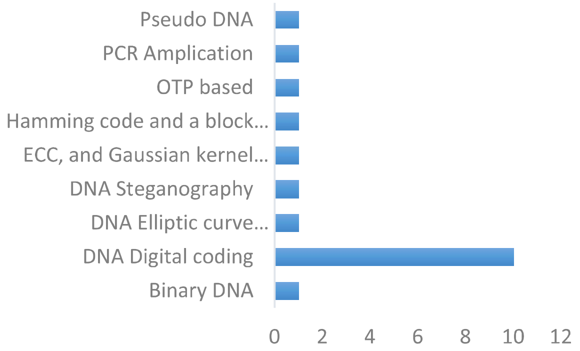

Figure 2: Summary of reviewed DNA Cryptographic techniques and their frequency

Based on Figure 1, it can be concluded that the DNA cryptographic approaches are mostly implemented in DNA Digital coding encryption technique and the technique is implemented often with OTP encryption technique and/or XOR operation.

VII. CONCLUSION

Most articles on DNA OTP based encryption technique did not present security analysis or mathematical proven methods for the encryption approaches. The researchers mostly demonstrated the proposed system model and developed web page for testing the approaches. Notwithstanding, our literature reveals that high time complexity and algorithm complexity are the major limitations of DNA encryption approaches, as presented in section VI. This study will serve as a reference for further research in future. Therefore, researchers can exploit this paper to improve more on the current DNA cryptographic limitations.

REFERENCES

[1] M. A. Saleh and H. Hashim, "HYBRID CRYPTOGRAPHIC APPROACH FOR INTERNET OF," vol. 3, no. 3, pp. 279–319, 2020.

[2] A. Sharma, "Security and Information Hiding based on DNA Steganography," vol. 5, no. 3, pp. 827–832, 2016.

[3] M. Rathi, S. Bhaskare, T. Kale, N. Shah, and N. Vaswani, "Data Security Using DNA Cryptography," vol. 5, no. 10, pp. 123–129, 2016.

[4] M. S. Taha, M. Shafry, and M. Rahim, "Combination of Steganography and Cryptography : A short Survey Combination of Steganography and Cryptography : A short Survey," 2019, doi: 10.1088/1757-899X/518/5/052003.

[5] M. A. Athitha, M. A. Akshatha, and B. Vandana, "A Review on DNA Based Cryptographic Techniques," vol. 3, no. 11, pp. 2819–2824, 2014.

[6] B. Anam, W. Yorkshire, W. Yorkshire, W. Yorkshire, and W. Yorkshire, "Review on the Advancements of DNA Cryptography."

[7] Y. Niu, K. Zhao, X. Zhang, and G. Cui, "Review on DNA Cryptography," in Communications in Computer and Information Science, 2020, vol. 1160 CCIS, pp. 134–148, doi: 10.1007/978-981-15-3415-7_11.

[8] S. C. Sukumaran and M. Mohammed, "DNA Cryptography for Secure Data Storage in Cloud," vol. 20, no. 3, pp. 447–454, 2018, doi: 10.6633/IJNS.201805.20(3).06.

[9] A. Hazra, S. Ghosh, and S. Jash, "A Review on DNA Based Cryptographic Techniques," vol. 20, no. 6, pp. 1093–1104, 2018, doi: 10.6633/IJNS.201811.

[10] P. Dixit, M. C. Trivedi, A. K. Gupta, and V. K. Yadav, "Video Steganography using Concept of DNA Sequence and Index Compression Technique," no. 5, pp. 408–417, 2019.

[11] G. Hamed, M. Marey, S. A. El-sayed, and M. F. Tolba, "Comparative Study for Various DNA Based Steganography Techniques with the Essential Conclusions about the Future Research," no. December, 2016, doi: 10.1109/ICCES.2016.7822003.

[12] L. Gehlot, R. Shinde, "A survey on DNA-based cryptography," International Journal of Advanced Research in Computer Engineering and Technology (IJARCET'16), vol. 5, no. 1, pp. 107-110, 2016.

[13] N. Gulati, S. Kalyani, "Pseudo DNA cryptography technique using OTP key for secure data transfer," International Journal of Engineering Science and Computing, vol. 6, no. 5, pp. 5657-5663, 2016.

[14] A. Kumar, V. K. Pant, "DNA cryptography a new approach to secure cloud data," International Jour- nal ofScientific and Engineering Research, vol. 7, no. 6, pp. 890-895, 2016.

[15] T. Mahalaxmi, B. B. Raj, J. F. Vijay, "Secure data transfer through DNA cryptography using symmet- ric algorithm," International Journal of Computer Applications, vol. 133, no. 2, pp. 19-23, 2016.

[16] A. Okamoto, I. Saito, K. Tanaka, "Public key system using DNA as a one way function for distribution," Biosystem, vol. 81, no. 1, pp. 25-29, 2015.

[17] T. Purusothaman, K. Saravanan, "DNA-based secret sharing algorithm for multicast group," Asian Jour- nal of Information Technology, vol. 15, no. 15, pp. 2699-2701, 2016.

[18] H. M. Abdelkader, F. E. Ibrahim, M. I. Moussa, "Enhancing the security of data hid- ing using double DNA sequences," in Industry Academia Collaboration Conference (IAC'15), 2015. (https://www.researchgate.net/publication/ 278028006_Enhancing_the_Security_of_Data_ Hiding_Using_Double_DNA_Sequences).

[19] T. Anwar, A. Kumar, S. Paul, "DNA cryptography based on symmetric key exchange," International Journal of Engineering and Technology (IJET'15), vol. 7, no. 3, pp. 938-950, 2015.

[20] M. Bhavithara, A. P. Bhrintha, A. Kamaraj, "DNA- based encryption and decryption using FPGA," In- ternational Journal of Current Research and Modern Education (IJCRME'16), pp. 89-94, 2016.

[21] K. Chiranjeevi, S. L. Kumar, R. Paspula, "Hidden data transmission with variable DNA technology," International Journal ofElectronics and Information Engineering, vol. 7, pp. 41-52, 2017.

[22] E. S. Babu, M. H. M. K. Prasad, C. N. Raju, "In- spired pseudo biotic DNA based cryptographic mech- anism against adaptive cryptographic attacks," In- ternational Journal of Network Security, vol. 18, no. 2, pp. 291-303, 2016.

[23] P. Barman, B. Saha, "An efficient hybrid elliptic curve cryptology system with DNA encoding," In- ternational Research Journal of Computer Science, vol. 2, no. 2, pp. 33-39, 2015.

[24] S. S. Basha, I. A. Emersonand, R. Kannadasan. Survey on molecular cryptographic network DNA (MCND) using big data. In Procedia Computer Science of 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15), vol. 50, pp. 3-9, 2015

[25] Sohal and Sharma. BDNA-A DNA inspired symmetric key cryptographic technique to secure cloud computing. Journal of King Saud University in Computer and Information Sciences, 2018. Available online 29 September

[26] Z. Yunpeng, L. Xin, M. Yongqiang, C. C. Liang. An Optimized DNA Based Encryption Scheme with Enforced Secure Key Distribution. Springer Science+Business Media, LLC, 2017.

[27] H. D. Tiwari, Jae Hyung Kim "Novel Method for DNA ‑ Based Elliptic Curve Cryptography for IoT Devices.ETRI Journal, Volume 40, Number 3, June 2018.

[28] E. I. Abd El-Latif & M. I. Moussa "Information hiding using artificial DNA sequences based on Gaussian kernel function" Journal of Information and Optimization Sciences ISSN: 0252-2667, 2019.