A SURVEY ON DATA ENCRYPTION USING DNA TECHNIQUE

Prema T.Akkasaligar[1], Farhat Mulla[2]

[1]*Professor, Department of CSE, BLDEA's Dr.P.G.Halakatti College of Engineering and Technology, Bijapur, Karnataka, (India)*

[2]*M.Tech (CSE), Student of BLDEA's Dr.P.G.Halakatti College of Engineering and Technology, Bijapur, Karnataka, (India)*

ABSTRACT

The present paper focuses mainly on the review of literature of encryption techniques using DNA technology. The sensitive information such as financial transactions, medical records of patients and personal records are transmitted over the network is more vulnerable to attacks. The new concept is introduced called DNA computing, brings a new hope for unbreakable algorithms. The paper aims at extensive experimental study of implementation of various available DNA encryption techniques.

Keywords: Encryption, Decryption, DNA Computing, DNA Cryptography

I. INTRODUCTION

Now a day's, providing security is one of the great challenge because of the advancement in digital communication technology, growth of computer power and storage. The different encryption techniques such as symmetric encryption and asymmetric encryption enables security of sensitive information, but the code breakers have come up with various methods to crack these algorithms. A new concept of DNA computing is introduced. Deoxyribonucleic acid (DNA) represents the genetic blueprint of living creatures. DNA is unique for each individual. DNA contains instructions for assembling cells. Every cell in the human body has a complete set of DNA. DNA is a polymer made of monomers called deoxyribonucleotides. Each nucleotide consists of three basic items: deoxyribose sugar, a phosphate group and a nitrogenous base. There are two types of nitrogenous bases: purines (Adenine (A) and Guanine (G)) and pyrimidines (Cytosine(C) and Thymine (T)). Since nucleotide differs only in terms of their bases, we use the base abbreviations to identify them. Single-stranded DNA molecules are simply chains of nucleotides where two consecutive nucleotides are bounded together by a strong covalent bond along a sugar-phosphate "backbone". The most important feature of DNA is the Watson–Crick complementarity of bases. Bonding between single strands occurs by the pairwise attraction of bases; A bonds with T and G bonds with C. The pairs (A; T) and (G; C) are therefore known as complementary base pairs. The two pairs of bases form hydrogen bonds between each other, two bonds between A and T, and three bonds between G and C. In [1] authors have introduced the first trial of DNA based cryptography in which a substitution method using libraries of distinctly, one time pads, each of which defines a specific, randomly generated, pairwise mapping and an XOR scheme utilizing molecular computation and indexed, random key strings are used for encryption.

II. LITERATURE SURVEY

To study and analyze more about the encryption techniques, the following literature survey has been done.

In [2-3], authors have presented an encryption technique using DNA technology. Algorithm works on plaintext and provides more security using the technology of DNA synthesis, polymerase chain reaction (PCR) and DNA digital coding along with traditional cryptography. By applying the special function of primers to PCR amplification, the primers and coding mode are used as the key of the scheme. The traditional encryption method and DNA digital coding both together are used to pre-process the plaintext, which can effectively prevent attack from a possible word as PCR primers. The issues and difficulties of cryptography computing and biological difficulties, provide a double security safeguards for the scheme. The security analysis shows that the encryption scheme has high confidential strength [4]. Moreover, the cost of this encryption scheme will be cut greatly with the progress of biological technologies in the future.

In [5], authors have jointly developed a method called text encryption using DNA stenography where hiding of data is done by applying five different steps. The receiver applies the process of identifying and extracting the original message which is hidden in DNA reference sequence. The main goal is exploring characteristics of DNA molecules, searching for simple methods of realizing DNA cryptography, and laying the basis for future development.

At encryption end, shift each letter in the message to a new letter where the shifted value is k. The shifted message is converted into a binary string B using ASCII value conversion and then convert binary string B into segments, where each bit is of size k = 2, hence convert B into a fake DNA string. After converting into fake DNA, matrix A is constructed. Matrix can be constructed by converting each alphabet in the text into a fake DNA strand. Each DNA strand for the alphabets is taken as a column to construct a 4 x k matrix, where k is the length of the original message. Now obtain a new string by concatenation of the rows of A and send the new string obtained to the receiver. At the decryption end, obtain the matrix A using the new string which is obtained previously. Obtain the binary string B from the conversion table and get the shifted message from the ASCII table. Finally obtain original message using the shift key.

In [6], authors have worked on RGB image encryption algorithm based on DNA encoding combined with chaotic map. RGB image has a high pixel correlation in spatial domain, but traditional encryption algorithm is mostly used to process it on the R, G and B layers, respectively. It is difficult to eliminate the pixel correlation in spatial domain. Aiming to the characteristics of RGB image, authors have used binary DNA encoding to make the mathematical problems into biological problems and introduce the biological knowledge of DNA computing into the proposed algorithm. The algorithm security is decided by chaotic system and DNA operation, to obtain dual security. The algorithm firstly carries out DNA encoding for R, G and B components of RGB image. The addition of R, G, and B are realized by DNA addition. After that complementation operation is carried out using the DNA sequence matrix controlled by logistic. Three gray images are obtained after decoding, finally the encrypted RGB image is reconstructed which uses image pixels disturbed by logistic chaotic sequence.

This encryption algorithm is effective, simple to implement and has a large secret key space, strong secret key sensitivity. Meanwhile, it can resist exhaustive attack, statistical attack, and thus it is suitable for RGB image encryption. In addition, the algorithm also has certain reference value for encryption of video, audio and other multimedia data. The speed performance of the proposed algorithm is not ideal, but authors have used

mathematical model to simulate the proposed algorithm in electronic computer and with the development of the DNA chip technology, it is not difficult to use ultra-large-scale parallel computing, power of DNA computing to implement the algorithm.

In [7], authors have worked on DNA based cryptographic Techniques. DNA encryption comes from DNA computing, initiated with the idea of "computing using DNA not on DNA". A lot of work has been done in the area and many researchers have done encryption based on different techniques like DNA digital encoding, PCR amplification, DNA synthesis, electrophoresis etc. Here an attempt is made for the message encryption along with the idea of adding authentication and message integrity. Encryption can be applied before or after authentication to maintain data confidentiality and data integrity so that only intended receiver can read or modify the data.

In [8], author has presented the method for cloud security. The cloud can be provided more security using DNA cryptography. The major concerns in cloud are the lack of confidentiality, integrity and authentication among the cloud users and service providers. The author in this paper proposed a new techniques for security schemes, to ensure data confidentiality, integrity, authentication and also DNA cryptographic algorithms are adopted for the optimization of data security in cloud computing. Although in its primitive stage, DNA cryptography is shown to be very effective. Theoretical analysis should be performed before its real applications, because it requires a high tech lab and computational limitations, as well as the labor-intensive extrapolation means so far. This makes the efficient use of DNA cryptography, difficult in the security world. The concept of DNA cryptography is used for very powerful and unbreakable encryption technology.

In [9], authors have worked on the technique called DNA based cryptography using random key generation scheme. They have presented a new DNA encryption technique based on mathematical matrix manipulation where they have used a secure generation algorithm for encryption process. The benefit of key generation scheme is, always get a new cipher text for same plaintext and same key. So it provides a good security layer which does not give any hint about plaintext. DNA binary strands support feasibility and applicability of DNA based cryptography. The security and the performance of the DNA based cryptographic algorithms are satisfactory for multilevel security applications of today's network. Certain DNA algorithms can resist exhaustive attack, statistical attack and differential attack. DNA computing is viable and DNA authentication methods have shown great promise in the marketplace of today and it is hoped that its applications will continue to expand. DNA cipher is the beneficial supplement to the existing mathematical cipher. If the molecular word is controlled then it may be possible to achieve vastly better performance for information storage and security .

In [10], authors have presented a paper, DNA based cryptography using permutation and random key generation. Initially plaintext is converted into ASCII code, ASCII code is again converted into binary form to get the data in 0's and 1's. These binary values are encoded in DNA sequences to nucleotide conversion where each of the four bases is represented by combinations of 0's and 1's. A DNA sequence is selected as a key and grouped into the blocks in which each block is of 4 characters. Then a table is created based on the positions of each character in the key sequence. Based on table and the randomly selected DNA sequence, text gets converted into encrypted form. Finally the encrypted sequence with the key is sent to the receiver. The DNA sequence in decryption process gets decoded into binary then that binary is converted into ASCII and finally ASCII to the plaintext. The method explains how traditional cryptography differs from the emerging DNA cryptography.

In [11], authors have presented a paper on enhanced information security using DNA cryptographic. The DNA cryptography is a new and promising area to achieve higher information security, using the characteristics of human DNA. Lots of DNA based encryption methods are proposed by several researchers. In [12], authors have given an idea using some special properties of DNA sequences to encrypt data. The method secretly selects a reference DNA sequence for encryption. In [13], authors have presented a method of data hiding where the data is encrypted using amino acid; DNA based playfair cipher and also use complementary rules to hide the resultant cipher text in a DNA sequence. In [14], authors have used a sort of indexing method over the complementary DNA sequence.

The present algorithm is more secure and uses a couple of 28 bit DNA sequence to generate the secret encryption key after a number of computations. Moreover a better level of message encryption technique is proposed where two rounds of encryption has been carried out among the plain text to generate two secret keys and produce a cipher DNA sequence with appending some extra information within it. As approximately 55 million publicly available DNA sequences are available, it is almost impossible for an intruder to predict the sequence. The message encryption approach is also better than the available cryptographic algorithms based on the DNA due to using some special operations performed on the data. Thus it is very much difficult for the intruders to apply different cryptanalysis on the cipher text.

In [15], authors have presented a paper on secure transmission of plaintext using DNA based message encoding. In the recent year few works on qualitative and quantitative analysis on DNA based cryptography as well as many new cryptographic techniques are proposed by the researchers. Bibhash Roy et al. [16-18] have proposed a DNA sequence based encryption and decryption process. The authors have proposed a unique cipher text generation procedure as well as a new key generation procedure. But the experimental result shows that the encryption process requires high time complexity. In [19], authors have designed a DNA encryption technique based on 4*4 matrix manipulations and using a key generation scheme which makes data much secure. In [20], authors have presented a theoretical and empirical based analysis on application of DNA cryptography.

In [21], author has designed a new method by integrating DNA computing in IDEA. Such conceptual works can be useful in the development of this new born technology of cryptography to fulfil the future security requirements. In this paper; a proposal is given where the concept of DNA is being used in encryption and decryption process. The theoretical analysis shows this method to be efficient in computation, storage and transmission; and it is very powerful in certain attacks. This paper also presents a secured symmetric key generation scheme which generates primary cipher and this primary cipher is then converted into final cipher using DNA sequences, so as to make it again more complicated in reading. Finally, the implementation methodology and experimental results are presented.

In actual scenario, DNA cryptography is far away from realization because in current time it can be performed only in labs using chemical operations. In order to provide better security and reliable data transmission an effective method of DNA based cryptography is presented. In this method the mixture of mathematical and biological concepts are used to get the encrypted data in the form of DNA sequences. The benefit of the scheme is that it makes difficult to read and guess about plaintext. The proposed algorithm has two phases in consequence: these are primary cipher text generation using substitution method followed by final cipher text generation using DNA digital coding.

In [22], authors have presented a symmetric key cryptosystem based on the DNA symmetric cryptosystem using index. As a result of applying the block cipher, the cryptosystem can be standardized and synchronized. Besides, by applying the method that ciphers plaintext strings have a proper computing with the result come from pseudo random number generator to create cipher text, it presents a proper random key sequence to improve security.

The DNA based encryption provides security but some additional level of security is added using index. The algorithm encodes each character into ASCII codes. According to the nucleotide sequence, the author should convert it to the DNA coding. Besides, the author selected the special DNA sequence as the encryption index, and likewise, the pre-treated plaintext will be divided into different groups. Next, the key created by the chaos key generator based on the logistic mapping and initialized by the number x_0 and μ will take XOR operation with the block-plaintext. The algorithm stores the position as the cipher text. The validity of the algorithm can be proved through simulation and the theoretical analysis, including bio-security and math security. The algorithm has a huge key space, high sensitivity to plaintext, and an extremely great effect on encryption. The algorithm provides an excellent performance of its encryption and an anti-attack ability.

In [23], authors have presented an asymmetric DNA mechanism, a more reliable and more powerful encryption than the OTP DNA symmetric algorithm. The purpose of this paper is to compare the time required to complete the encryption/decryption in the case of the DNA cipher with the time required by other classical encryption algorithms. The DNA cipher requires a longer time for encryption and decryption, comparatively to the other ciphers. Authors would expect these results because of the platform used for developing this algorithm. Java cryptography architecture contains the classes of the security package Java 2 SDK, including engine classes. The methods in the classes that implement cryptographic services are divided into two groups. The first group is represented by the APIs (Application Programming Interface) and the second group is represented by the SPIs (Service Provider Interface). Each SPI class is abstract. In order to implement a specific service, for a specific algorithm, a provider must inherit the corresponding SPI class and implement all the abstract methods. All these methods process array of bytes while the DNA cipher is about strings. The additional conversions from string to array of bytes and back, makes this cipher to require more time for encryption and decryption then other classic algorithms. To emphasize the difference between DNA and classical algorithms a dedicated application (smart cipher) is developed.

The dedicated application shows the encryption and decryption time. Based on this criterion and the strength of the cipher, the user can estimate the efficiency of the used algorithm. The authors have compared the execution time of the DNA symmetric cipher with the time required by other classical encryption algorithms. The algorithm is tested on a random text of 360 characters, which is in string format. To be able to compute the time required for encryption and decryption, authors have used the public static nanoTime() method from the system class which gives the current time in nanoseconds. It is important to understand that the execution time varies depending on the OS used, the memory load and on the execution thread management. Authors have therefore measured the execution time on 3 different machines one machine is Intel Core 2 Duo 2140, 1.6 GHz, 1 GB RAM, Vista OS and second one is Intel Core 2 Duo T6500, and finally third machine is 2.1 GHz, 4 GB RAM, Windows 7 OS. The first machine and second machine (with Windows OS) have larger time variations for the encryption and decryption processes. The third machine, based on the linux platform, offers a better stability, since the variation of the execution time is smaller.

In [24], authors have presented a paper on enhanced level of security using DNA computing technique with hyper elliptic curve cryptography. DNA based elliptic curve cryptographic technique require larger key size to encrypt and decrypt the message resulting in increased processing time, more computational and memory overhead. To overcome the above limitations, DNA strands are used to encode the data to provide first level of security and HECC encryption algorithm is used for providing second level of security. HECC is better than the existing public key cryptography technique such as RSA, DSA, AES and ECC in terms of smaller key size. DNA cryptography is a next generation security mechanism, storing almost a million gigabytes of data inside DNA strands. Hence this proposed integration of DNA computing based HECC provides higher level of security with less key size of HECC-80 bits than ECC-160 bits and with less computational and memory overhead.

In this present method first level of security is provided by converting original text message into DNA nucleotide which can able to store millions of data in a single DNA strands. Further encoded nucleotide is converted into numbers. Second level of security is provided by converting numbers into points using Koblitz method. These points act as plaintext for encryption using hyperelliptic curve cryptography. MATLAB simulation tool is used to simulate the proposed cryptographic scheme for different key size and processing time. Recent research shows that HECC are well suited for various software and hardware platforms and their performance is compatible to that of ECC.

In [25], authors have presented a DNA based implementation of YAEA encryption algorithm, used to enhance the security of cryptography. The investigation conducted in this paper is based on a conventional symmetric encryption algorithm called “Yet another Encryption Algorithm” (YAEA) developed by Saeb and Baith. It was Adleman, with his pioneering work [Adleman, 1994], who set the foundation for the new field of bio-computing research. His main notion is to use actual chemistry to solve problems that are either unsolvable by conventional computers, or require a massive amount of computation.

The present method is effortlessly scalable for large digital information products. The algorithm is effective at encrypting and decrypting digital information from biological DNA strand. The algorithm utilizes a sequential search algorithm in order to locate and randomly return one of the many positions of quadruple DNA nucleotides sequence representing the binary octets of plain-text characters. The decryption process is achieved by using the pointer file and the same random binary file that is available to both sender and receiver in advance. The algorithm is a symmetric cipher consisting of recording pointers to the randomly selected locations of the file in the searchable DNA strand for each plaintext character. Authors have conducted a test by recording the time needed to encrypt the “Uncle Tom’s Cabin” novel using six different DNA strands of different lengths. Utilizing a dedicated 3.2 GHz CPU employing 3G RAM.

III. PERFORMANCE PARAMETERS

In [26], authors have presented a paper on hybrid encryption using DNA technology. They defined a set of parameters, based on which the performance can be evaluated and compared with the existing encryption technique using DNA technology such parameters are as follows.

3.1 Statistical Analysis

The encrypted image should not have any statistical similarity with the original image to prevent the leakage of information

3.2 Histogram Analysis

To get the good performance, histogram of original image and encrypted image should not be similar.

3.3 Correlation Coefficient Analysis

In most of the plaintext-images, there exists a high correlation among adjacent pixels, while there is a little correlation between neighbouring pixels in the encrypted image. It is the main task of an efficient image encryption algorithm to eliminate the correlation of pixels. Two highly uncorrelated sequences have approximately zero correlation coefficient [27].

3. 4 Differential Attacks

Attackers often make a slight change for the original image, use the proposed algorithm to encrypt the original image before and after changing, and compare two encrypted images to find out the relationship between the original image and encrypted image.

3.5 Known-Plaintext and Chosen Plaintext Attacks

For encryption with a higher level of security, the security against both known-plaintext and chosen-plaintext attacks are necessary. Chosen/Known-plain text attacks are such attacks in which one can access/choose a set of plain texts and observe the corresponding encrypted texts.

3.6 Brute Force Attack

Brute force attack or exhaustive key search is a strategy that can be used against any encrypted data by an attacker who is unable to take advantage of any weakness in an encryption system that would otherwise make his task easier. It involves systematically checking all possible keys until the correct key is found [26].

IV. CONCLUSION

Now a days, the security for the data has become highly important since sensitive information such as financial transactions, medical and personal records are transmitted through public communication facilities and also transmission of digital products over the open network occur very frequently. In this paper, the survey is done on existing works on the encryption techniques using DNA. The different DNA encryption techniques are studied and analyzed well to promote the performance of the encryption methods also to ensure the security proceedings. To sum up, all the techniques are useful for real-time encryption. Each technique is unique in its own way, which might be suitable for different applications. Everyday new DNA encryption technique is evolving hence fast and secure conventional DNA encryption techniques will always work out with high rate of security.

REFERENCES

[1] Gehani, Ashish La Bean, Thomas H. Reif, H.John, “DNA based cryptography”, Dimacs Series in Discrete Mathematics and Theoretical ComputerScience, 2000, pp.162-167.

[2] J. Kawai and Y. Hayashizaki, DNA book. Genome Res, Vol. 13, 2003, pp.1488–1495.

[3] T. Kamei, “DNA-containing inks and personal identification systemusing them without forgery”, Jpn. Kokai Tokkyo Koho, 2002, p.8.

[4] Guangzhao Cui , Limin Qin , Yanfeng Wang , Xuncai Zhang, "An encryption scheme using DNA technology", Computer Engineering and Applications,2008, pp.37-42.

[5] M. Yamuna, Nikhil Bagmar, Vishal, "Text encryption using DNA stenography", International Journal of Emerging Trends & Technology in Computer Science, 2013, Volume 2, Issue 2, ISSN 2278-6856,pp.231-233.

[6] Lili Liu, Qiang Zhang, Xiaopeng Wei, "A RGB image encryption algorithm based on DNA encoding and chaos map", Computers and Electrical Engineering, 2012, www.elsevier.com/locate/compeleceng, pp.1-9.

[7] Kritika Gupta, Shailendra Singh, "DNA based cryptographic techniques: a review", international Journal of Advanced Research in Computer Science and Software Engineering, 2013, Volume 3, Issue 3, ISSN: 2277 128X , pp.607-610.

[8] Anup R. Nimje, "Cryptography in cloud-security using DNA (Genetic) techniques", 2012, Vol. 2, Issue5, ISSN: 2248-9622, pp.1358-1359.

[9] P.Surendra Varma, K.Govinda Raju, "Cryptography based on DNA using random key generation scheme", International Journal of Science Engineering and Advance Technology, IJSEAT ,2014, Vol 2, Issue 7, ISSN 2321-6905, pp.168-175.

[10] Bonny BRaj, Panchami, "DNA based cryptography using permutationand random key generation method, International Conference On Innovations & Advances In Science, Engineering And Technology",2014, Volume 3, Special Issue 5, ISSN (Online) : 2319 – 8753, ISSN (Print) : 2347 – 6710, pp.263-267.

[11] AbhishekMajumdar, Meenakshi Sharma, "enhanced information security using DNA cryptographic Approach", International Journal of Innovative Technology and Exploring Engineering, 2014, Volume-4 Issue-2, ISSN: 2278-3075, pp.72-76.

[12] H.Z. Hsu and R.C.T.Lee, "DNA based encryption methods", The 23rd Workshop on Combinatorial Mathematics and Computation Theory, National Chi Nan University Puli, NantouHsies, Taiwan 545, April 2006, pp.145-150.

[13] AmalKhalifa and Ahmed Atito. "High-capacity DNA-based stegano graphy", The 8th International Conference and informatics and Systems (INFOS2012), IEEE, May.2012, pp.76-80.

[14] Mohammad Reza Abbasy, PouryaNikfard, Ali Ordi, Mohammad RezaNajaf Torkaman, "DNA base data hiding algorithm", in international Journal on New Computer Architectures and TheirApplications.2012, ISSN: 2220-9085, pp.183-192.

[15] SnehalJavheri, Rahul Kulkarni, "Secure data communication and cryptography based on DNA based message encoding", International Journal of Computer Applications, 2014,Volume 98– No.16, pp.35-40.

[16] Bibhash Roy, GautamRakshit, PratimSingha, Atanu Majumder, Debabrata Datta, "An improved symmetric key cryptography with DNA based strong cipher", ICDeCom, 2011, Feb' 24-25'2011, pp.1-5.

[17] Bibhash Roy et al, "A DNA based symmetric key cryptography", ICSSA, 2011, 24-25 Jan'11.

[18] Bibhash Roy, GautamRakshit, PratimSingha, Atanu Majumder, Debabrata Datta, "An enhanced key generation scheme based cryptography with DNA logic", IJICT, 2010-11, Volume 1 No. 8, Dec' 2011.

[19] Miki Hirabayashi, Akio Nishikawa, "Analysis on secure and effective applications of a DNA based cryptosystem", IEEE computer Society, 978-0-7695-4514-1/11, 2011, pp.205-210.

[20] Nucleotide base pairing of strands, http://dedunn.edblogs.org, 2012.

[21] TusharMandge, Vijay Choudhary, "A DNA encryption technique based on matrix manipulation and secure key generation scheme", ICICES Journal, 2013,Print ISBN:978-1-4673-5786-9, pp.47-52.

[22] Zhang Yunpeng, Zhu Yu, Wang Zhong, Richard O.Sinnott, "Index based symmetric DNA encryption algorithm", Proceedings of the 4th International Congress on Image and Signal Processing, 2011, http://hdl.handle.net/11343/32713 , pp.2290–2294.

[23] Er.RanuSoni, Er.VishakhaSoniand Er.Sandeep Kumar Mathariya, "Innovative field of cryptography: DNA cryptography", CS & IT-CSCP, 2012, pp. 161–179.

[24] P.Vijayakumar, V.Vijayalakshmi, G.Zayaraz, "Enhanced level of security using DNA computing technique with hyperelliptic curve cryptography", ACEEE Int. J. on Network Security, 2013, Vol. 4, No. 1,pp.1-5.

[25] T. Amin, MagdySaeb, Salah El-Gindi, "A DNA-based implementation of YAEA encryption algorithm", IASTED International Conference on Computational, 2006.

[26] Grasha Jacob1, A. Murugan, "A hybrid encryption scheme using DNAtechnology", The International Journal of Computer Science and Communications Security, 2013, Volume 3, pp.61-65.

[27] Hiremath P.S., Prema T.Akkasaligar and Sharan Badiger, "Speckle noise reduction in medical ultra sound image, advancement and breakthrough in ultra sound images", in Tech Publisher, Crortia, 5th june 2013, pp.201-241, (DOI:10.5772156519).