International Journal of Enhanced Research in Management & Computer Applications
ISSN: 2319-7471, Vol. 4 Issue 11, November-2015

The Future of Data Security: DNA Cryptography

Anil Kumar

Programmer-Cum-Networking Engineer, Haryana Roadways, Transport Department, Govt of Haryana, India.

ABSTRACT

DNA cryptography is a new promising direction in cryptography research that emerged with the evolution in DNAcomputing field. DNA can be used not only to store and transmit the information, but also to perform computation. The extensiveparallelism and extraordinary information density inbuilt in this molecule are exploited for cryptographic purposes. The theoreticalanalysis shows this method to be efficient in computation, storage and transmission; and it is very powerful in certain attacks. DNA cryptography isproposed for a secure end to end communication due to the vast parallelismand extra ordinary information density that are inherent in any DNA molecule.In this paper various trends in DNA cryptographic approaches are surveyedand analyzed, highlighting the merits and demerits of each.

Keywords: Security, DNA, DNA computing, DNA cryptography.

1. INTRODUCTION

The security to a system is essential now a day! With the growth of the Information Technology and with theemergence of new techniques, the number of threats a user issupposed to deal with grew exponentially. It doesn't matter ifwe talk about bank accounts, social security numbers or asimple telephone call. Due to lot of information flow on the network. There are various adversaries who always try to break into the system in order to steal the crucial information or to destroy the integrity of data. So information security becomes necessity for modern computing systems. There are some sectors like government, banks, military who can't afford any leaks to their secret data.It is important that the information isknown only by the intended persons, usually the sender andthe receiver.

This is where the cryptography comes intopicture. Cryptography is the basis of security of all theinformation.Various cryptographic systems were developed in the pastyear. The recent development on this field is DNACryptography. This concept has emerged after the disclosureof computational ability of Deoxyribo Nucleic Acid (DNA).In this field of DNA Cryptography many research work isgoing on to make the computational process more complexto the unauthorized user. The main objective of this method is to encrypt the plaintext and hide it in the DNA digital form.

DNA cryptography enables the confidentiality of data more high then the modern methods with the use of one time pad (OTP) keys and its size. Also it is believed that in DNA cryptography the key can be generated for the huge length of data compared to the modern methods in which key are generated only for smaller length of the data. Also with the breaking up of modern cryptographic algorithm like DES and MD5 the new methods of information security are needed to protect our data. The concept of DNA computing plays an important role in the field of computer security which is assumed to be a more powerful and unbreakable cryptographic algorithm now a days.Well, presently it is in thedevelopment phase and requires a lot of work and researchto reach an established stage.

2. DNA

DNA stands for Deoxyribonucleic acid which store genetic information of the entire living organism ranging from human being to small viruses. It is also called as an information carrier and consists of long polymer of small units called nucleotides. Further nucleotides consist of three components: Nitrogenous base, five Carbon sugar and Phosphate group. Nitrogenous base consists of four bases: Adenine, Thymine, Cytosine and Guanine (A, T, C, G), all the complex information about organism are stored with the combination of these bases. Adenine and Guanine are called purines, whereas Thymine and Cytosine are called pyrimidines. DNA is a double helix structure as shown in the Figure below.

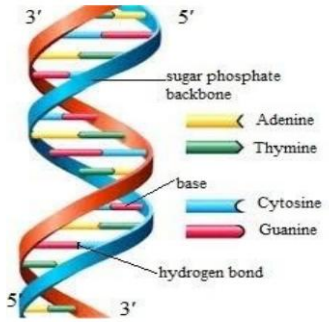

Figure 1: Double helical structure of DNA

DNA double helical structure was discovered by the two Nobel laureate Watson and Crick and therefore it is also called a Watson-Crick complementary structure, where A and T form hydrogen bond with each other, whereas C and G forms bond with one another. In this structure of DNA both the strands are antiparallel to each other, means if one strand starts from 3′ to 5′, then another strand is from 5′ to 3′

3. DNA COMPUTING (RELATED WORK)

In 1994, Adleman [1] laid the foundation of DNA computing by giving solutions to the combinatorial problems using molecular computation, one of which is "Hamiltonian path" problem. He solved the instance of graph containing seven vertices by encoding it into the molecular form by using an algorithm and then computational operations were performed with the help of some standard enzymes. This was solved by brute force method. In 1995, Lipton [8] extended the work of Adleman by solving another NP-complete problem called "satisfaction" by using DNA molecules in a test tube to encode the graph for 2 bit numbers. In 1996, Dan Boneh et al. [4] applied the approaches of DNA computing used by Adleman and Lipton, in order to break one of the symmetric key algorithm used for cryptographic purposes known as DES (Data Encryption Standard). They performed biological operations on the DNA strands in a test tube, such as extraction, polymerization via DNA polymerase, amplification via PCR and perform operations on the DNA strands which have the encoding of binary strings. Then DES attack is planned by generating the DES-1 solution, due to which key can be easily guessed from the cipher text and further evaluate the DES circuit, lookup table and XOR gates. By using their molecular approach they broke DES in merely 4 months. In 1997, Qi Ouyang et al. [12] applied the approaches of DNA molecular theory in order to generate the solution for maximal clique problem, which is another NP-complete problem. Thus shows the efficiency of DNA: to solve Hard-problems and vast parallelism inherent in it which makes the operations fast.

Advantages of DNA computing

- **Speed** – Conventional computers can perform approximately 100 MIPS (millions of instruction per second). Combining DNA strands as demonstrated by Adleman made computations equivalent to 10^9 or better, arguably over 100 times faster than the fastest computer.
- **Minimal Storage Requirements** – DNA stores memory at a density of about 1 bit per cubic nanometer, where conventional storage media requires 10^12 cubic nanometers to store 1 bit.
- **Minimal Power Requirements** – There is no power required for DNA computing while the computation is taking place. The chemical bonds that are the building blocks of DNA happen without any outside power source. There is no comparison to the power requirements of conventional computers.

4. DNA CRYPTOGRAPHY (RELATED WORK)

In 2003, Jie Chen [5] presented the DNA cryptographic approach based upon molecular theory, one-time pad and performed encryption/decryption of 2 dimensional image. In 2004, AshishGehani et al. [3] laid the foundation of DNA cryptography by using molecular approach and the concept of one-time pad which has perfect secrecy, according to Vernam's and Shannon: inventor of one-time pad. They have proposed a method of encryption and decryption which is based on DNA chip and one-time pad. So it is very hard for the adversary to guess the encrypted message. In 2005, Kazuo Tanaka et al. [9] proposed the DNA cryptographic approach based on Public Key (one way). In this approach they have clearly explained about the formation of public keys by using solid supports mixture for PKA and ODN mixture for PKB. After generating the keys, message is encoded in a DNA sequence with the help of one of the public key, which is further synthesized with the DNA synthesizer and then the encoded message sequence is ligated with the another public key. Now the outcome of the previous process is forwarded to the immobilization process and then for PCR amplification, where the amplification is done with the help of secret sequence, in order to decode the encoded DNA sequence.

In 2006, Sherif T. Amin et al. [14] proposed the DNA cryptographic approach based on symmetric key, where key sequences are obtained from the genetic database and remain same at both ends (sender and receiver). In 2008, Anil Verma et al. [2] proposed a novel paradigm for secure routing in Mobile Adhoc Networks (MANETs) that uses Pseudo DNA cryptography approach in order to secure the Adhoc networks. Adhoc network is a wireless network which has no fixed infrastructure and where each node act as a host and router and there is no centralized authority which makes them vulnerable to the security attacks present in the networks. Pseudo DNA cryptography approach they have used is based on the central dogma of molecular biology. Concept of how messages are stored in DNA and then transfer to the mRNA (transcription), and then to the proteins (translation) which is our ciphertext. Ciphertext is send through the secure channel to the intended receiver and symmetric key with one-time pad is used at both the ends (encryption and decryption). In 2008, Guangzhao Cui et al. [6] proposed the public key encryption technique that uses DNA synthesis, DNA digital coding and PCR amplification to provide the security safeguard during the communication. This encryption scheme has high confidential strength.

In 2010, Lai Xuejia et al. [10] proposed a DNA public key cryptosystem which is based on DNA microarray/chip technology in which DNA chip is fabricated with probes. One set of probes are used for encryption process and another set for decryption process. In 2011, Deepak Kumar and Shailendra Singh [7] proposed a new secret data writing techniques based on DNA sequences. They have explained this algorithm by using a simple example of "HELLO" as a plaintext and generate assDNA one-time pad key of 350 bits which is 70 times longer than the plaintext and perform encryption and decryption on the plaintext using symmetric key cryptography. So to find the exact key, adversary has to search among 4310 different ssDNA string which is impossible for the adversary. In 2012, SabariPramanik and Sanjit Kumar Setua [13] proposed a new parallel DNA cryptography technique using DNA molecular structure and hybridization technique which certainly minimize the time requirement. They have explained how message is exchanging safely between sender and receiver with an example.

In 2012, Yunpeng Zhang et al. [15] proposed a DNA cryptography based on DNA fragment assembly. In their algorithm they have clearly mentioned how sender converts the plaintext into binary sequence and then into long chain of DNA, which is further fragmented into small DNA chains. Key of short chain implantation takes place in the fragments and forward to the receiver as a ciphertext and then receiver deciphers it and starts fragment reassembly to obtain the plaintext. In 2013, Olga Tornea and Monica E. Borda [11] proposed a DNA based cipher which is based on DNA indexing. They take the random DNA sequence from the genetic database and use as a one-time pad key, which is send to the receiver by a secure communication channel. The encryption mechanisms takes place by converting the plaintext into its ASCII code and then convert it into the binary format which is converted into the DNA sequence (A, C, G, and T). Now DNA sequence formed is search in the key sequence and writes down the index numbers. The array of integer numbers obtained are our ciphertext which is decrypted by the receiver only using the key and index pointer.

Table 1. DNA Cryptography (Related Work in Chronicle Order)

Year	Algorithm Title	Cryptographic Approach	Technology Used
2003	A DNA-based, Bio molecular Cryptography Design [5].	Symmetric key	Molecular, one-time pad.
2004	DNA-Based Cryptography [3].	Symmetric key	Molecular, DNA chip, onetime pad.
2005	Public-key system using DNA [9].	Asymmetric key	Molecular, DNA synthesis, PCR amplification.
2006	YAEA DNA Encryption [14].	Symmetric key	Substitution, one-time pad.
2008	DNA Cryptography: secure routing in MANETs [2].	Symmetric key	Central dogma of molecular biology, one-time pad.
2008	Encryption Scheme Using DNA [6].	Asymmetric key	DNA synthesis, DNA digital coding, PCR amplification
2010	Asymmetric Encryption and Signature with DNA [10].	Asymmetric key	DNA chip technology, Hybridization.
2011	Secret Data Writing Using DNA Sequences [7].	Symmetric key	one-time pad.
2012	DNA Cryptography [13].	Symmetric key	Hybridization, one-time pad
2012	DNA Cryptography Based on Fragment Assembly [15].	Symmetric key	DNA Fragment assembly.
2013	Security and Complexity of DNA Based Cipher [11].	Symmetric key	DNA Indexing, one-time pad,

CONCLUSION

DNA cryptography is in its infancy. Only in the last few years has work in DNA computing seen real progress. DNA cryptography is even less well studied, but ramped up work in cryptography over the past several years has laid good groundwork for applying DNA methodologies to cryptography and steganography.DNA cryptography based on DNA hybridization, DNA synthesis, DNA microarray/chip technology, Central dogma, PCR amplification and one-time pad makes it unique from the traditional cryptographic techniques which are unbreakable by the adversaries, due to molecular computation inherent in it. Further with the addition of one-time pad in DNA symmetric key cryptography makes it more strong and secure and protect from brute force attacks. It also offers high confidential strength and large storage density inherent in it, as compare to the traditional storage devices. With the invention of energy efficient DNA computer chip by IBM, opens up the way for the bright inventions by the researchers now days in the field of computing, information security. Though DNA has many positive aspects in different fields which fascinate the researchers, some of the aspects like requirement of bio molecular labs, environment impact, and quantum attacks still an issue for its smooth progress.Researches and studies are being carried out to identify a better and unbreakable cryptographic standard. A number of schemes have been proposed that offer some level of DNA cryptography, and are being explored. At present, work in DNA cryptography is centered on using DNA sequences to encode binary data in some form or another. Though the field is extremely complex and current work is still in the developmental stages, there is a lot of hope that DNA computing will act as a good technique for Information Security.

REFERENCES

[1]. Adleman. M. L (1994), Molecular Computation of Solutions to Combinatorial Problems, Science, vol. 266, pp. 1021- 1024.
[2]. A.K. Verma, Mayank Dave, C. Joshi (2008), DNA cryptography: a novel paradigm for secure routing in MANETs, Journal of Discrete Mathematical Sciences and Cryptography, Vol. 11, No. 4, pp. 393-404.
[3]. AshishGehani, LaBean Thomas and John Reif (2004), DNA-based cryptography, In Aspects of Molecular Computing, Springer Berlin Heidelberg, pp. 167-188.
[4]. Boneh. D (1996), Breaking DES using Molecular computer, American Mathematical Society, pp 37-65.
[5]. Chen Jie (2003), A DNA-based bio molecular cryptography design, Proceedings of IEEE International Symposium, Vol. 3, pp. III-822.
[6]. Cui Guangzhao, Limin Qin, Yanfeng Wang, and Xuncai Zhang (2008), An encryption scheme using DNA technology, In Bio-Inspired Computing: Theories and Applications, BICTA, 3rd IEEE International Conference on, pp. 37-42.
[7]. Deepak Kumar, and Shailendra Singh (2011), Secret data writing using DNA sequences, In Emerging Trends in Networks and Computer Communications (ETNCC), IEEE International Conference on, pp. 402-405.
[8]. J. Lipton. R (1995), Using DNA to solve NP Complete problems, Science, Vol. 268, pp. 542-545.
[9]. Kazuo Tanaka, Akimitsu Okamoto, and Isao Saito (2005), Public-key system using DNA as a one-way function for key distribution, Biosystems 81, 1, pp. 25-29.
[10]. Lai XueJia, MingXin Lu, Lei Qin, Han JunSong, and Fang XiWen (2010), Asymmetric encryption and signature method with DNA technology, Science China Information Sciences 53, no. 3, pp. 506-514.
[11]. Olga Tornea, and Borda E. Monica (2013), Security and complexity of a DNA-based cipher, In Roedunet International Conference (RoEduNet), 11th IEEE International Conference, pp. 1-5.
[12]. Ouyang Qi, D. Peter Kaplan, Liu Shumao and Albert Libchaber (1997), DNA solution of the maximal clique problem, Science 278, 5337, 446-449.
[13]. PramanikSabari and Kumar SanjitSetua (2012), DNA cryptography, In Electrical & Computer Engineering (ICECE), 7th IEEE International Conference, pp. 551-554.
[14]. Sherif T. Amin, MagdySaeb and El-Gindi Salah (2006), A DNA-based implementation of YAEA encryption algorithm, In Computational Intelligence, pp. 120-125.
[15]. Yunpeng Zhang, Bochen Fu, and Xianwei Zhang (2012), DNA cryptography based on DNA Fragment assembly, Information Science and Digital Content Technology (ICIDT), 8th IEEE International Conference, Vol. 1, pp. 179-182.