International Research Journal of Computer Science (IRJCS)
Issue 5, Volume 2 (May 2015)
ISSN: 2393-9842
www.irjcs.com

An Efficient Hybrid Elliptic Curve Cryptography System with DNA Encoding

Prokash Barman*
Department of Computer Science & Engineering
Calcutta University

Banani Saha
Department of Computer Science & Engineering
Calcutta University

Abstract — Cryptography is the technique to encrypt and decrypt the data for secure communication. The cryptographic mechanism enables the entities of network to transmit secure data through insecure channel. So that, only the intended users can access the transmitted data. Two public key cryptographic techniques are mostly used for secure transmission of data, namely Elliptic Curve Cryptography & RSA. Among this two cryptographic system, ECC is more first due to its small key size. DNA-cryptography is a new technique, which also used for secure data transmission in recent years. The DNA based cryptography technique, derived from DNA computing. It uses DNA nucleotide sequence for cryptographic purpose. To develop more secure and stable cryptography technique, we propose a new hybrid DNA encoded Elliptic Curve Cryptography scheme in this paper. DNA encoded ECC cryptography uses smaller key size and less computation power with multilevel security. The main attraction of the proposed system is that it has two level of security. First is unknown DNA sequence based encoding and the second is Light weight ECC based encryption and decryption system.

Keywords—Elliptic Curve Cryptography (ECC), Rivest, Shamir and Adleman (RSA), Deoxyribonucleic Acid (DNA), Nucleotide, Kobliz's algorithm, Public Key Cryptography

I. INTRODUCTION

Elliptic Curve Cryptography is the most efficient public key cryptography technique independently proposed by Victor S. Millar and Neal Koblitz in 1986 and 1987 respectively [1]. Benefit of ECC is a smaller key size, which facilitates to reduce storage and transmission requirements; as a result an elliptic curve cryptography mechanism can provide same level of security which is offered by a RSA-based system with a large modulus and correspondingly larger key size. For example, a 256 bit ECC public key cryptography provides comparable security to a 3072 bit RSA public key. The main operation of ECC is point multiplication (scalar multiplication); which is done by elliptic curve arithmetic [2].

The concept of DNA cryptography was introduced after research in the field of DNA computing by Adleman (1994). DNA coding is a new area of cryptography which has appeared in recent years along with DNA computing research. DNA cryptography is built on DNA - which is an information carrier – and uses modern biotechnology for its tools. DNA cryptography achieves the encryption process by the use of the characteristics of DNA of massive parallelism and high storage density [3]. The reason of combining cryptography and molecular biology is the DNA encoded plaintext, which combines the mathematics and molecular biology techniques- to obtain the final cipher text [3]. In our proposed system DNA-computing is used to encode or decode plaintext and ECC is used for encrypt or decrypt. This makes our proposed hybrid crypto system of DNA-ECC. This paper is organised in the following sections. Section II describes Elliptic Curve Cryptography. DNA Cryptography and its Biological background have been discussed in Section III. In section IV, proposed DNA encoded ECC scheme is described. Advantage of the proposed system and conclusion are given in section V and VI respectively.

II.1. ELLIPTIC CURVE CRYPTOGRAPHY (ECC)

As the key length of RSA public key cryptography has increased recent years, it has put a heavier processing load on application. This problem has been reduced using smaller key length Elliptic Curve Cryptography (ECC). The ECC is based on Elliptic Curve. Elliptic Curves are not actually ellipses. The curve is so named because they are depicted by cubic equation, similar to those used for calculating the circumference of an ellipse. The locus of a point, whose coordinates conform to a particular cubic equation along with the point at infinity O (the point at which the locus in the projective plane intersects the line at infinity) is known as an elliptic curve [6]. An elliptic curve is defined in a two dimensional, standard, x, y Cartesian coordinate system; which is derived from the expressions (1) and (2) as given below

$y^2 = x^3 + ax + b$ ----------(1)

where $4a^3 + 27b^2 \neq 0$.

$y^2 + xy = x^3 + ax^2 + b$ ----------(2)

The mathematical foundation of ECC is based on the above equations which may be depicted as follows

© 2014-15, IRJCS- All Rights Reserved

ISSN: 2393-9842
www.irjcs.com

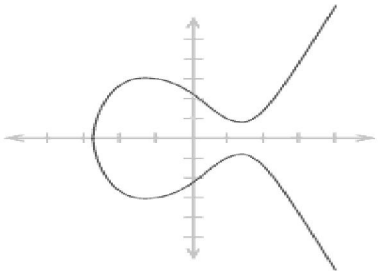


Figure 2: An Elliptic Curve

In ECC, the elliptic curve is used to define the members of the set over which the group is calculated and the operations between them which define how math works in the group. It is done by imagine a graph labelled along both axis with the numbers of a large prime field. The elliptic curve cryptosystem is one of the best public key cryptosystem among three cryptosystems currently in use i.e. (a) public key cryptography (PKC), (b) integer factorization system and (c) discrete logarithm system. The RSA cryptosystem is the best known example of the integer factorization problem while the Digital Signature Algorithm (DSA) cryptosystem is based on the discrete logarithm problem.

Public-key cryptography is based on the interoperability of certain mathematical problems. Before ECC, it assumed that Public-key systems are secure assuming that it is difficult to factor a large integer composed of two or more large prime factors. In case of elliptic curve based protocols, finding the discrete logarithm of a random elliptic curve element with respect to a publicly known base point is not feasible which is assumed. The size of the elliptic curve determines the difficulty of the problem. Benefit of ECC is a smaller key size, which facilitate to reduce storage and transmission requirements, as a result an elliptic curve group could provide the same level of security afforded by an RSA-based system with a large modulus and correspondingly larger key, i.e. a 256bit ECC public key should provide comparable security to a 3072bit RSA public key.

The mathematical operations of ECC is defined over the elliptic curve $y2 = x3 + ax + b$, where $4a3+27b2 \neq 0$. Each value of 'a' and 'b' gives a different elliptic curve. All points (x , y) which satisfies the above equation plus a point at infinity lies on the elliptic curve. A point in the curve is public key and the private key is a random number. By multiplying the private key with the generator point G in the elliptic curve, the public key is obtained. The domain parameter (Set of pre defined constant known by all the parties taking part in the communication) of ECC constitute by the generator point G, the curve parameters 'a' and 'b', along with few other constants.

II.II. ECC OPERATIONS

Various operations which are performed on ECC are given below in detail [6]:

1) Point Multiplication : The dominant operation in ECC cryptographic schemes is point multiplication method. Point multiplication is calculating the value of *nJ*, where *n* is an integer and *J* is a point on the elliptic curve defined in the prime field. In point multiplication, a point J on the elliptic curve is multiplied with a scalar n using elliptic curve equation to obtain another point K on the same elliptic curve, i.e. nJ=K.

The Point multiplication method is achieved by two basic elliptic curve operations.
• *Point addition*, adding two points J and K to obtain another point L i.e., L = J + K.
• *Point doubling*, adding a point J to itself to obtain another point L i.e. L = 2J.

a) **Point addition**: The Point addition method is the addition of two points J and K on an elliptic curve to obtain another point L on the same elliptic curve.

Consider two points J and K on an elliptic curve as shown in figure (a). If K ≠-J then a line drawn through the points J and K will intersect the elliptic curve at exactly one more point –L. The result of addition of points J and K gives the point -L. The reflection of the point –L with respect to x-axis gives the point L. That is on an elliptic curve L = J + K. If K = -J the line through this point intersect at a point at infinity O. Hence J + (-J)= O. This is shown in figure 2(b). Where, O is the additive identity of the elliptic curve group. The reflection of a point with respect to x-axis is the negative of that point.

© 2014-15, IRJCS- All Rights Reserved

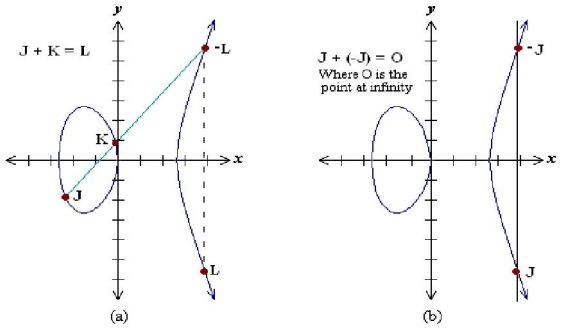


Figure 3: Addition of two points

b) ***Point doubling****:* Point doubling is the addition of a point J on the elliptic curve to itself to obtain another point L on the same elliptic curve. To find L = 2J which imply to double a point J to get the point L, consider a point J on an elliptic curve as shown in figure 3(a). If y axis coordinate of the point J is non zero then the tangent line at J will intersect the elliptic curve at exactly one more point –L. The reflection of the point –L with respect to x-axis gives the point L, which is the doubling result of the point J.

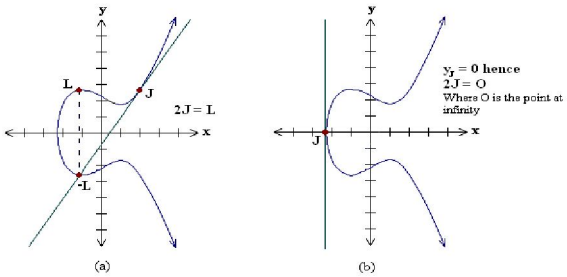


Figure 4: Doubling of Points

Thus L = 2J. If y coordinate of the point J is zero then the tangent at this point intersects at a point at infinity O. Hence 2J = O when yJ = 0. This is shown in figure.

II.III. SECURITY POLICY OF ECC (DISCRETE LOGARITHM PROBLEM)

The security of ECC [7] depends on Elliptic Curve Discrete Logarithm Problem. Let J and K are two points on an Elliptic Curve such that nJ=K, where n is a scalar. Given J and K, it is computationally infeasible to obtain n, if n is sufficiently large. n is the discrete logarithm of K base J. The main operation involved in ECC is point multiplication i.e. multiplication of scalar n with the point J on the curve to obtain point K on the curve. In Elliptic Curve Cryptography all Plain texts are converted to Plain text point. From each plain text point J, another point K on the elliptic curve is obtained with nJ=K. At the end, using ECC point addition or point doubling final encrypted cipher text obtained. Now, with given J and K, it is infeasible to obtain n. Without knowing the value of n, it will be hard to determine whether J and K are complementary points. So, determining n is the most important factor in ECC. Hence, n is called the discrete logarithm of K base J and determination of n is the main security of ECC.

III.I. DNA CRYPTOGRAPHY- BIOLOGICAL BACKGROUND

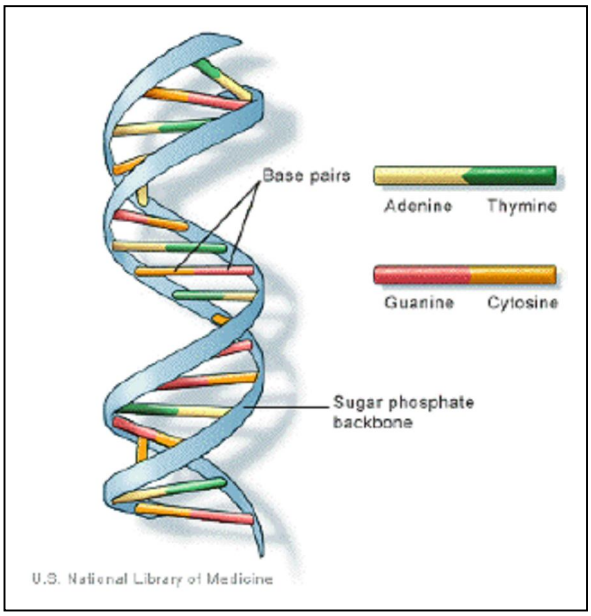
DNA is the genetic material of eukaryotes, with a double-helix molecular structure and two single-strands parallel to each other. DNA is something which is called a polymer, which composed of many small nucleotides. Each nucleotide consists of three parts:

1. The Nitrogenous bases;
2. Deoxyribose;
3. Phosphate.

DNA is the abbreviation of deoxyribonucleic acid which is the germ plasma of all life. DNA is a kind of biological macromolecule and is made up of nucleotide. Each nucleotide contains a single base and there are four kinds of bases, which are Adenine(A) and Thymine(T) or Cytosine(C) and Guanine(G) corresponding to four kinds of nucleotides.

A single-stranded DNA is constructed with orientation: one end is called 5', and the other end is called 3'. Usually DNA exists as double-stranded molecules in nature. The two complementary DNA strands are held together to form a double helix structure by hydrogen bonds between the complementary bases of A and T (or C and G). The double-helix structure was discovered by Watson and Crick; thus the complementary structure is called Watson Crick complementary structure [4].

© 2014-15, IRJCS- All Rights Reserved



DNA Structure

III.II. DNA CRYPTOGRAPHY TECHNIQUE

DNA cryptography is a study of DNA computing about how to use DNA as an information carrier and to use mathematical operation to transfer plaintext into cipher text. In this section we will describe some DNA Cryptography techniques used in [8]. In DNA cryptography, the sequence of DNA base is used for encryption or decryption operation. A DNA base sequence consists of four alphabets: A, C, G and T. Each alphabet is related to a nucleotide. The DNA sequence is usually quite long. For our example purpose we shall use part of the long DNA sequence. The following are two DNA sequences are taken for our example.

The first one is a segment of DNA sequence of Litmus, its real length is with 2856 nucleotides long:
ATCGAATTCGCGCTGAGTCACAATTCGCGCTGAGTCACAATTCGCGCTGAGTCACAATTGTGACTCAGCCG
CGAATTCCTGCAGCCCCGAATTCCGCATTGCAGAGATAATTGTATTTAAGTGCCTAGCTCGATACAATAAA
CGCCATTTGACCATTCACCACATTGGTGTGCACCTCCAAGCTCGCGCACCGTACCGTCTCGAGGAATTCCTG
CAGGATATCTGGATCCACG AAGCTTCCCATGGTGACGTCACC [10].

The second one is a segment of DNA sequence of Balsaminaceae, its real length is with 2283 nucleotides long:
TTTTTATTATTTTTTTTCATTTTTTTCTCAGTTTTTAGCACATATCATTACATTTTATTTTTTCATTACTTCTAT
CATTCTATCTATAAAATCGATTATTTTTATCACTTATTTTTCTAATTTCCAATATTTCATCTAATGATTATATT
ACATTAAAGAAATCGGTTAAAAGCGACTAAAAATCAATCTGGAACAAGGCTTAGTTTATTTAATATATTAT
TTTATGTAATTTCTATTGAAAAAT TAGTTAAAAGGCAAGTATTTGAGAT [10].

There are a large number of DNA sequences publicly available in various web-sites, such as [10]. Publicly available DNA sequences are to be around 55 million [10]. Three DNA based encryption methods are designed in [8]. These methods secretly select a reference sequence S from publicly available DNA sequences. Only the sender and the receiver are aware of this reference sequence. The sender transform this selected DNA sequence S into a new sequence S' by incorporating the DNA sequence S with the secret message M. This transformed sequence S' is sent by a sender to the receiver together with many other DNA sequences. The receiver examine all of the received sequences, identify S' and recover the secret message M.[8].

III.III. DNA CRYPTOGRAPHY OPERATION

Three DNA cryptography methods as in [8] are generally used. The methods are as follows

a. *Insertion method.*
b. *Substitution Method.*
c. *Complementary Pair approach.*

In all of above the approaches, a common method of encoding and decoding is used. The plaintext is converted to binary numbers. The converted the binary numbers are converted to equivalent DNA nucleotides sequence. Then one of the DNA-cryptographic methods of [8] is used for encryption or decryption.

The encoding and decoding operation based on the following facts:

© 2014-15, IRJCS- All Rights Reserved

For DNA there are four basic units which are encoded into binary in the following manner:

DNA Nucleotide Base	Binary equivalent
Adenine (A)	00
Thymine(T)	01
Guanine(G)	10
Cytosine(C)	11

Table 1, DNA Nucleotide to Binary conversion table

In the encoding phase of insertion method of DNA cryptography, plain texts are converted into binary. A DNA sequence is taken from publicly available sequence. Convert the DNA sequence into binary as per Table -1. Divide the binary DNA sequence into segments, where each segment contains a randomly selected number of bits. Randomly selected number of bits will be greater than 2. Each bit of binary plain text is then inserted at the beginning of segmented binary DNA sequence. The inserted sequences are then concatenated to obtain encoded binary sequence. A new fake (not found really, only obtained from operation) binary sequence is obtained by converting the encoded binary sequence into nucleotide using table -1. The derived

IV. PROPOSED DNA ENCODED ELLIPTIC CURVE CRYPTOGRAPHY SCHEME

IV.I. ECC-DNA CRYPTOSYSTEM

The elliptic curve cryptography and DNA cryptography are the most recent cryptographic mechanism. For better security, characteristics of both the systems may be combined. A highly secured cryptographic system may be developed using this two recent techniques. Our proposed cryptographic scheme has advantages of both ECC and DNA based computing. For encoding of plaintext we propose to use the insertion method of DNA cryptography as used in [8]. The plaintext is converted to its equivalent ASCII value. The ASCII values are converted into binary. A known DNA nucleotide sequence is taken from publicly available sequences. Both the sender and the receiver should know the chosen DNA sequence. The DNA nucleotide sequences are converted into binary using table – 1. In this stage we will get several pairs of binary numbers. All the pairs are concatenated and a long binary number will be generated. The binary number is then segmented into several parts. Here an arbitrary number of bits, greater than 2 (it was already pairs of bits in early stage) will be taken in each segment. Now each bit of binary plaintext is inserted into the beginning of the binary segments of nucleotide sequence. The segments are concatenated again and converted to Nucleotide letter.(A,T,G &C). The new sequences are converted into decimal as per table – 2. These are the steps for encoding of our proposed system.
For encryption the decimal numbers are converted into elliptic curve point using Koblitz method [8]. This point is called plain test point. The points are encrypted into another elliptic curve point using ECC encryption expression (1). ECC encryption is done with the help of its generated keys. The encrypted points from the elliptic curve are in the form of cipher text points. This point is send to the receiver. In the reverse process the receiver will derive the plaintext message.

$\{kG, Pm + k\ PB\}$ (Expression-1)

Where,

G - Generated Points
P_m - Plaintext points
k - Random number chosen by user
P_B - Public key of another user

$P_m + kP_B - n_B(kG) = P_m + k(n_B)G - n_B(kG) = P_m$ (Expression-2)

The cipher text points are deciphered using ECC decryption expression (2). Deciphered points are converted into numbers using Koblitz's method. These numbers are decoded to an unknown DNA nucleotide sequence. From the unknown DNA sequence, known DNA nucleotide sequence (S) is decoded to obtain plaintext.

A – 10	T – 20	G – 30	C – 40

Table – 2. Nucleotide to number conversion table

IV.II. PROPOSED HYBRID DNA-ECC CRYPTOGRAPHIC ALGORITHM

Input: Plain text (P), number of bits of DNA sequence segment (k), known DNA sequence(D)
Output: Cipher text points(C)

© 2014-15, IRJCS- All Rights Reserved

Procedure
Begin
Input P
Convert P into Binary P'
Convert D into Binary D'
Segment D' with k bit in a segment
Insert each bits of P' into the beginning of each segment of D'
Concatenate segments of D'
Convert D' into DNA nucleotide DN (where A=00, T=01, G=10, C=11)
Convert DN into decimal N (as in table-1)
Call koblitz to convert N into ECC point and cipher text C
End
End Procedure

Example for converting plain text
Plaintext message (P): "t"
ASCII message: 116
Binary message(P'): 1110100

DNA sequence (D):
ATCGAATTCGCGCTGAGTCACAATTCGCGCTGAGTCACAATTCGCGCTGAGTCACAATTGTGACTCAGCCG
CGAATTCCTGCAGCCCCGAATTCCGCATTGCAGAGATAATTGTATTTAAGTGCCTAGCTCGATACAATAAA
CGCCATTTGACCATTCACCACATTGGTGTGCACCTCCAAGCTCGCGCACCGTACCGTCTCGAGGAATTCCTG
CAGGATATCTGGATCCACG AAGCTTCCCATGGTGACGTCACC

Binary DNA sequence (D'):00 01 11 10 00 00 01 01 11 10 11 10 11 01
Segmented Binary DNA sequence (where k=3): 000 111 100 000 010 111 101 110 110 1

Insert each bits of P' into beginning of each segments of D'.
1-000 / 1-111 / 1-100 /0-000 / 1-010 / 0-111/ 0-101 / -110 /-1

Concatenate the segments of D'.
1000111111000000101001110101110 1

Convert D' to nucleotide DN
10 00 11 11 11 00 00 00 10 10 01 11 01 01 11 01
G A C C C A A A G G T C T T C T

Convert DN to ASCII

G	A	C	C	C	A	A	A	G	G	T	C	T	T	C	T
30	10	40	40	40	10	10	10	30	30	20	40	20	20	40	20

Covert ASCII(DN) to Elliptic Curve point which is encrypted cipher text point

V. ADVANTAGES OF PROPOSED SYSTEM

The proposed DNA-ECC hybrid crypto system is more efficient than the existing DNA cryptography and Elliptic Curve Cryptography in terms of security. Because it uses DNA cryptographic mechanism in its encoding phase and for encryption it uses Elliptic Curve Cryptography mechanism. So, the proposed system has two level of security first one in the encoding phase and the other is in the encryption phase. It uses small key size for encryption (as ECC is used for encryption, so the key size is the ECCs key size) comparative to other cryptographic system in cyber space. It would be hardly breakable by eavesdropper as it contains two level of security.

VI. CONCLUSION

This paper describes a novel cryptographic scheme by combining DNA computing theory with ECC algorithm. Elliptic curve cryptography with DNA Computing offers major advantages over traditional systems. In the form of increased speed, less memory and smaller key size. In addition, the proposed system has multi level security one in DNA encoding and other in ECC encryption steps. A DNA-ECC embedded system based on the principles described in this paper may be developed using FPGA based embedded system to verify the practical implementation of the proposed cryptographic scheme.

© 2014-15, IRJCS- All Rights Reserved

REFERENCES

[1] N. Koblitz, "Elliptic Curve Cryptosystem, Mathematics of Computation", Vol A8, 1987, PP 203-209
[2] Prokash Barman & Banani Saha, " E-Governance Security using Public Key Cryptography With special focus on ECC", International Journal of Engineering Science Invention, Vol-2, Issue 8, August 2013, PP 10-16.
[3] Yunpeng Zhang and Liu He Bochen Fu (2012). "Research on DNA Cryptography, Applied Cryptography and Network Security", Dr. Jaydip Sen (Ed.), ISBN: 978-953-51-0218-2, InTech, Available from: http://www.intechopen.com/books/applied-cryptography-and-network-security/research-on-dna-cryptography
[4] Er.Ranu Soni, Er.Vishakha Soni and Er.Sandeep Kumar Mathariya," Innovative field of cryptography: DNA Cryptography", Computer Science & Information Technology (CS & IT), PP 162-179.
[5] P.Vijayakumar, V.Vijayalakshmi and G.Zayaraz, "DNA Computing based Elliptic Curve Cryptography", International Journal of Computer Applications, December'2011, Vol-36, No-4, PP 18-21.
[6] Arun Kumar, Dr. S.S. Tyagi, Manisha Rana, Neha Aggarwal, Pawan Bhadana - "A Comparative Study of Public Key Cryptosystem based on ECC and RSA" - International Journal on Computer Science and Engineering (IJCSE), Vol 3, No. 5, May-2011, pp. 1904-1905.
[7] William Stallings, Cryptography and Network Security, Fifth Edition, Pearson. pp. 344.
[8] H. Z. Hsu and R. C. T. Lee, "DNA Based Encryption Methods", The 23rd workshop on Combinatorial Mathematics and Computation Theory, pp. 145-150.
[9] P. Vijayakumar, V. Vijayalakshmi & G. Zayaraz, "DNA Computing based Elliptic Curve Cryptography", International Journal of Computer Applications (0975 – 8887), Volume 36– No.4, December 2011, pp 18-21.
[10] European Bioinformatics Institute, URL: http://www.ebi.ac.uk/.

© 2014-15, IRJCS- All Rights Reserved