ELLIPTIC CURVE CRYPTOGRAPHY TECHNIQUE ON DATA SEQUENCE

MAJOR PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF DEGREE OF

Master of Technology

In

Information Systems

Submitted By:

SAURABH KATIYAR

(2K13/ISY/22)

Under the Guidance

Of

Dr N.S. RAGHAVA

ASSOCIATE PROFESSOR

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(2013-2015)

CERTIFICATE

This is to certify that **Saurabh Katiyar (2k13/ISY/22)** has carried out the major project titled "**Elliptic Curve Cryptography Technique on Data Sequence**" in partial fulfilment of the requirements for the award of Master of Technology degree in Information Systems by **Delhi Technological University**.

The major project is bonafide piece of work carried out and completed under my supervision and guidance during the academic session 2013-2015. To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other University/Institute for the award of any degree or diploma.

Dr N.S. Raghava

Associate Professor

Department of Electronics and Communication

Delhi Technological University

Delhi-110042

ACKNOWLEDGEMENT

I take the opportunity to express my sincere gratitude to my project mentor Dr N.S. Raghava, Associate Professor, Department of Electronics and Communication, Delhi Technological University, Delhi, for providing valuable guidance and constant encouragement throughout the project. It is my pleasure to record my sincere thanks to him for his constructive criticism and insight without which the project would not have shaped as it has.

It humbly extends my words of gratitude to other faculty members of Computer Science and Engineering department for providing their valuable help and time whenever it was required.

Saurabh Katiyar

Roll No. 2k13/ISY/22

M.Tech (Information Systems)

E-mail: 100rabh.dtu@gmail.com

ABSTRACT

Communication is a meaningful exchange of information between two or more entities. Images and documents travel widely and rapidly, in multiple manifestations, through email and across the Internet. In this era of e-communication i.e. electronic transmission of information that has been encoded digitally (as for storage and processing by computers), the first concern is about the security of the content which is shared during communication. While the information is over net, it is next to impossible to keep a track of where the information or the copy of information is going through. Security is a continuous process via which data can be secured from several active and passive attacks. Several security techniques can be used to ensure the integrity, authentication and confidentiality of the information. Cryptography is one of the primitive way to secure the information from hackers or intruders. Encryption technique protects the confidentiality of a message or information which can be in the form of multimedia (text, image, and video).Since there are limited encryption algorithm but humongous key space, therefore, the secrecy of encryption depends on the secret key. In this work, a new symmetric image encryption algorithm is proposed based on ECC on image data. Intensity value of each pixel is mapped to a random DNA sequence which provided extra level of security in our scheme. In ECC, input data is mapped to points on elliptic curve. Different encoding techniques are used for this purpose. One of the best suited techniques is Koblitz's method. Our scheme has also used Koblitz's method for converting intensity of each pixel of the original image into points of elliptic curve. This is the pre-processing step of ECC. On the other hand, during decryption every point on elliptic curve is converted or deciphered into intensity value of original image. This is post processing step of ECC. Pre and post processing steps are performed using Koblitz's method.

Table of Contents

Title **Page no.**

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

To prevent the data from entering into the wrong hands are prevent its misuse, data security is crucial. The most commonly used techniques to ensure data security is cryptography. As will be shown later, elliptic curve cryptography aims to overcome the shortcomings of the currently used techniques. Cryptography can be used to store either stored data or data in transit over a network.

Security is a process in which basically we protect the data of our computer system from unauthorized access. An attacker is a person who gets the benefit of weakness in a computer system or network and attempt to achieve an unauthorized access of data or information. Both type of resources e.g. physical or non-physical need to protect from attackers. Non-physical resources comprise of data or information and physical resources contains peripherals or network[1].

1.1 CRYPTOGRAPHY

Cryptography is a technique in which any data is encrypted and decrypted by the help of some mathematical approach, which is transformation of readable form of some data into non-readable form and again conversation of non-readable form into readable form so that sender and receiver got the readable form of data and other than both will get non-readable form of data. Here the data which is readable is called plain text and the non-readable data is called cipher text. Transformation of readable data to non-readable data is called encryption process and reverse of this process is called decryption process. Apart from this , there is a requirement of key which known as secret key by the help of this key sender encrypt the message and further only that network user can decipher this message into plain text or readable form who have the secret key.

Cryptanalysis is a part of cryptography through which encrypted message sometimes can be broken, that also known as code breaking, although modern cryptography techniques are virtually unbreakable.

In cryptography suppose there are two users, A and B who wants to communicate to each other and another one is C who is adversary and have not permitted to access the confidential information. User A has some secret data and wants to share it with another user B but not want to disclose it any other user like C. Now the main issue is - how A can send the data so that it will not be disclosed to any other user? This can be possible in such a way that only user B can obtain information by De-scrambling it, which is scrambled by user A. Scrambling, is done at user A side who is the sender of information by the help of any encryption scheme. Encryption scheme requires a secret key that is used to both side at user A and user B. In case of user A secret key used to scramble the message or information and other side at user B to De-scramble the message or information to obtain original message.

In present era, the use of Internet and communication becomes extensive and due to this reason security is a prime concern over here. So cryptography is applied to secure the data , not only security feature it also provided the confidentiality of data that protected from alteration and stolen of data another feature that included by cryptography is authenticity. Much information as information of credit card or smart card, confidential information, electronic message, corporate data is protected by cryptography. Confidentiality feature is applied by encryption functions in cryptography that gives a secure communication environment, prevents discloser of store information and access of data by unauthorized users. There are various types of cryptographic techniques as authentication and digital signature can provide security against message forgeries and spoofing. Cryptography is an essential tool that required to make information secure and it's easily available on web to user. PGP (Pretty Good Privacy) is one of the cryptographic systems which are used on Internet because it is freely available and more effective.

1.2 VARIOUS ASPECTS OF CRYPTOGRAPHY:

- **Data Confidentiality:** In this aspect of cryptography, we simple scramble the data using encryption technique and De-scramble it by decryption technique so that only sender and receiver can read the data, no other one can't read it. Encryption technique is conversion of readable data into non-readable data which is cipher text and enables security and maintains the privacy while fetching the cipher text over any communication medium and by the help of decryption process, receiver can get the original message from the cipher text which is just the reversible process of encryption. Encryption and Decryption process requires secret key, in some cases secret keys are same at both side or may be different.

- **Authentication:** Authentication is a procedure which makes sure that the sender of the message is one who declared in the message. This can be made as this process. Suppose user A sends a message ans user B receive that message but user B does not know about the sender so user B requires a proof which proves the identity of sender that message is sent by the user A. For this it is necessary that user performs a action on the message which gives a proof of originator from where message was originated.
- **Integrity:** In communication medium there is one more problem that is the loss of integrity of message sent by sender. It's means that a message may be altered or modified by adversity over communication medium. This aspect ensures that no other user will alter the message so that the original message will reach at the end of receiver. Cryptographic hash is used to verify the integrity of message.

1.3 CRYPTOSYSTEM AND IT'S ASPECTS

A cryptosystem comprises an ordered list of all possible finite plaintexts, corresponding finite cipher texts, the set of all finite possible keys and the algorithms used in the encryption and decryption process for each key.

The aspects of a cryptosystem that characterize it and which are responsible for security are as follows.

1.3.1 Key Size

The key size plays an important role in determining the security of cryptography techniques. Small key size makes the cipher vulnerable to attacks such as the Brute Force attack. Large key sizes secure the system against Brute Force attack but slow down the encryption and decryption processes.

Improvements in technology have led to an increase in computing power which has increased the key sizes to provide acceptable security levels to extremely large values. The higher complexity of the Elliptic Curve Discrete Cryptography to provide comparable security levels, but with much smaller key sizes.

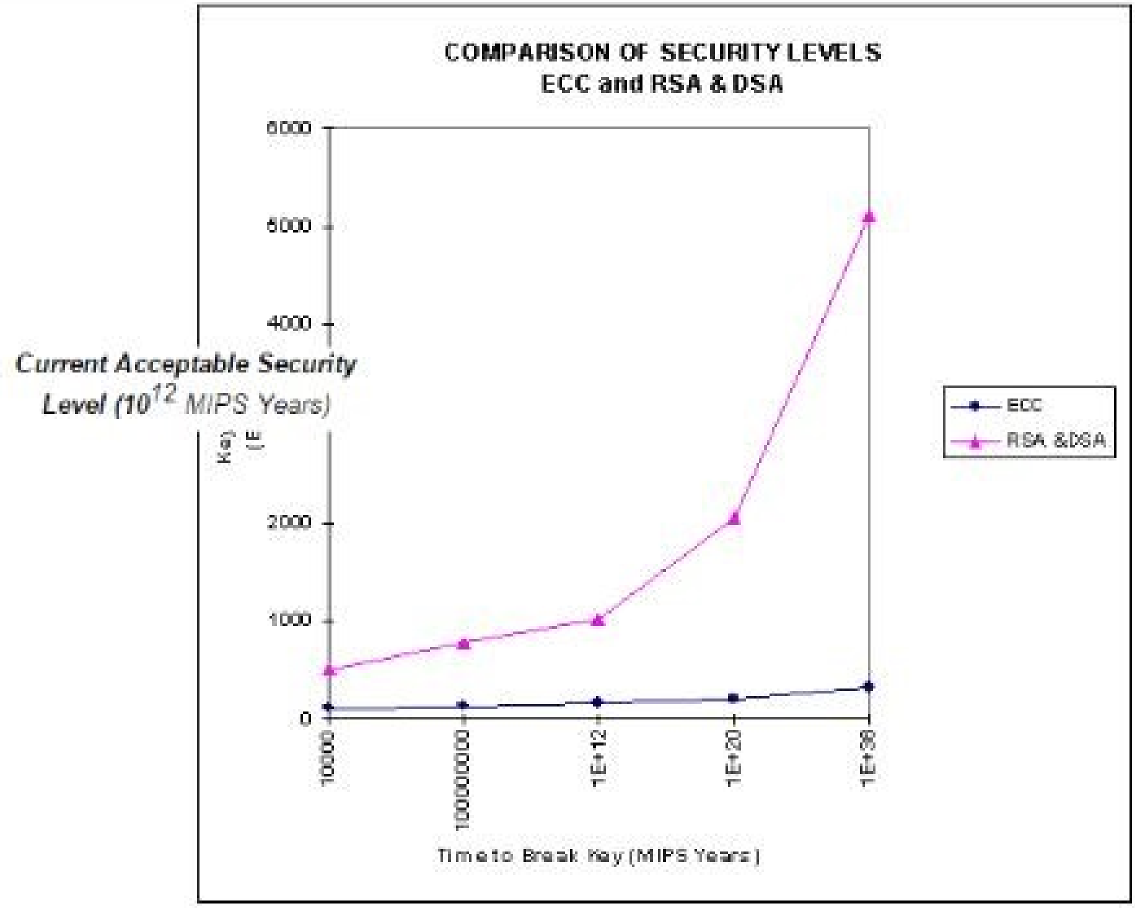


Figure 1.1 Key sizes for acceptable security levels (RSA vs ECC)

Figure shows a graph of the breaking time versus key sizes for currently used RSA and Elliptic Curve. As can be seen, acceptable key sizes in RSA IS 1024 while for ECC, it is 160 bit.

Also, as can be seen from Table, for comparable security levels, the key sizes for Symmetric Key Techniques, RSA/Diffie-Hellman and Elliptic Curve Cryptography such that they provide comparable levels of security are shown. While the key sizes for Symmetric Key techniques are the smallest, the problem of secret key transmission makes the method unviable.

Table 1.1 NSIT recommended Key sizes for comparable security levels

Symmetric Key Size (bits)	RSA and Diffie- Hellman Key (bits)	Elliptic Curve Key Size (bits)
80	1024	160
112	2048	224
128	3072	256
192	7680	384
256	15360	521

In case of asymmetric key techniques, the key sizes of ECC are less than one-tenth the key size required in the case of RSA. This is even more prominent as the key sizes increase in RSA. These long keys not only slow down the encryption and decryption process but are difficult to store. Thus, ECC is a much better alternative than RSA/Diffie-Hellman.

Thus, symmetric key technique are used in the case of encryption of long messages, while the secret key used in the encryption and decryption is encrypted using asymmetric key techniques before transferring from the sender to the recipient. This leads to faster encryption and decryption of the messages while simultaneously solving the key transmission problem and is the main idea behind hybrid cryptosystems.

1.3.2 Trapdoor Function

The main factor behind any mathematical construct being used in a cryptography lies in the existence of a trapdoor function. A trapdoor function is any function which can easily be solved in the forward direction, however reverse computation of the function is computationally infeasible or very expensive. The more expensive the reverse computation, the more secure cipher.

1.4 TYPES OF CRYPTOGRAPHY

Cryptographic techniques are classified on the basis of the types and number of keys involved in the encryption and decryption process.

1.4.1 Symmetric Key Cryptography

Encryption and decryption involve the use of the same key. This introduces the key transmission problem, which involves sending the shared key used by the sender for encryption, to the recipient. Interception of the message containing the key can provide a third party access to the shared key, who may now be able to decipher all the cipher texts in that particular exchange with ease.

Symmetric key ciphers mostly employ substitution and transposition techniques to encrypt and decrypt the data.

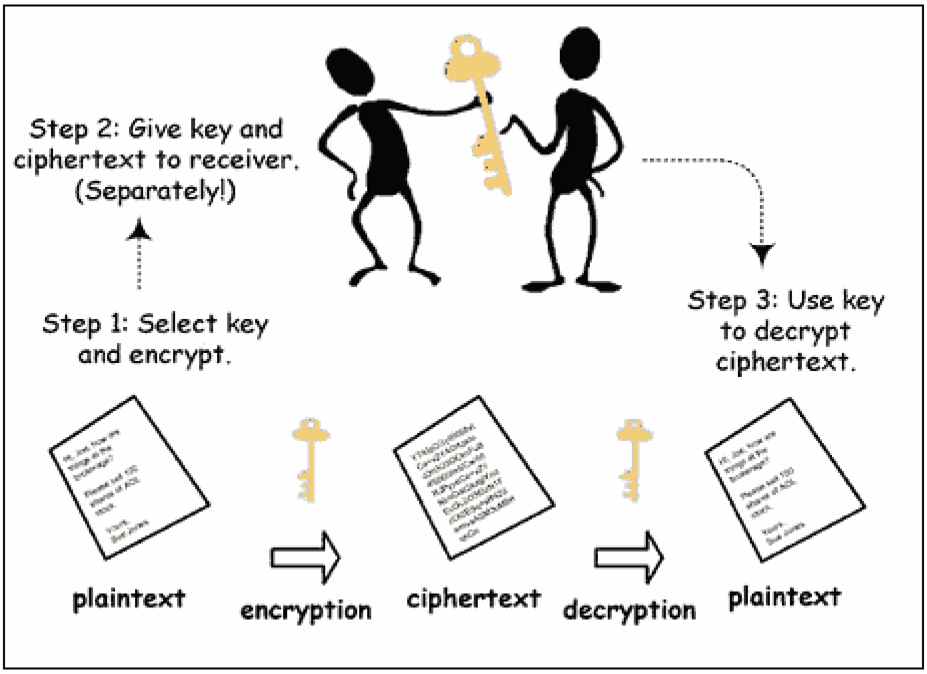

Step 2: Give key and ciphertext to receiver. (Separately!)
Step 1: Select key and encrypt.
Step 3: Use key to decrypt ciphertext.
plaintext
encryption
ciphertext
decryption
plaintext

Figure 1.2 Symmetric key cryptography

1.4.2 Public-Key (Asymmetric Key) Cryptography

In public-key cryptography, each entity has two keys: a public key, known publicly to all other entities and a private key, known only to the entity to which it belongs. The widely-known public key is used for encryption by the sender [19][20]. The receiver uses his private. The widely-known public key to decrypt the message, thus obviating the need for key transfer. A combination of the public and private keys is required to encrypt and decrypt the data, thus leading to an asymmetric cryptosystem. Since a private key can only decrypt messages encrypted using the corresponding public key and vice versa, the process is secure.

Public-key cryptography is more widely used as it eliminates the need for a key to be transmitted. Disadvantages of public-key systems involve authentication of sender and messages, which involves the use of digital certificates or Trusted Third parties. Moreover, public-key cryptography techniques are more computationally expensive than symmetric key techniques, making them unsuitable for limited resource mobile devices, such as low battery life and power.

Unlike symmetric key techniques, asymmetric key techniques rely more on mathematical constructs and algorithms to encrypt and decrypt the data and do not use transposition and substitution.

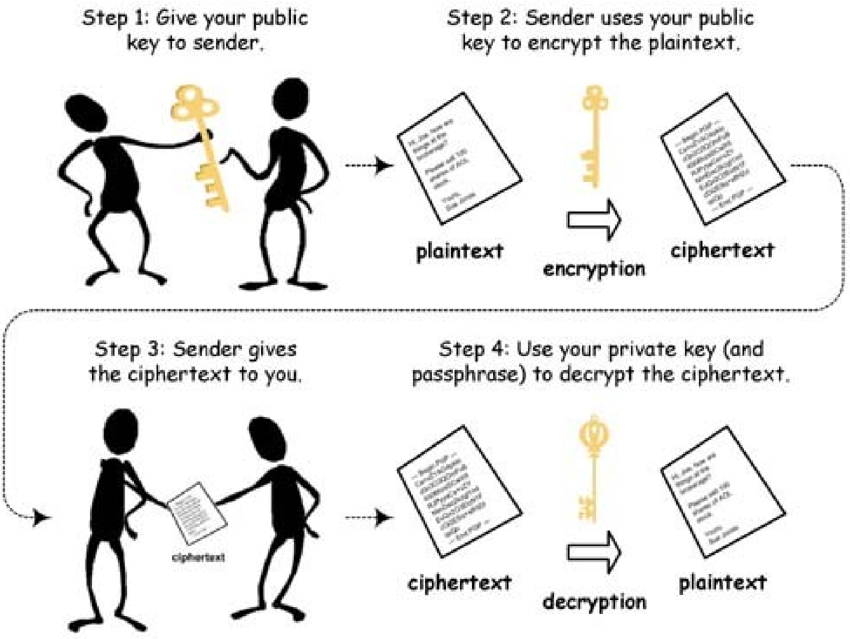

Step 1: Give your public key to sender.
Step 2: Sender uses your public key to encrypt the plaintext.
plaintext
encryption
ciphertext
Step 3: Sender gives the ciphertext to you.
ciphertext
Step 4: Use your private key (and passphrase) to decrypt the ciphertext.
ciphertext
decryption
plaintext

Figure 1.3 Public key cryptography

1.4.3 Identity-Based Cryptography

Identity-based cryptography is a type of public-key cryptography in which the keys for an entity are derived from some inherent characteristic of the entity. The public key is derived from some publicly known characteristic while the private key is derived from some information about the entity known only to that entity. The characteristic of the entity used to obtain the key is unique to that entity. Thus, no key authentication in the form of certificates is necessary as verification. The entity's attribute will serve as authentication.

Being a public-key system, no key transmission is necessary. The sender may simply compute the public key of the intended recipient from the publicly known unique attribute and use the computed key to encrypt. This ciphertext can only be decrypted using the private key, which is computed by the recipient based on some private attribute and is kept a secret, thus ensuring security of this mechanism. It also eliminates the need of a trusted third party to generate and transmit keys.

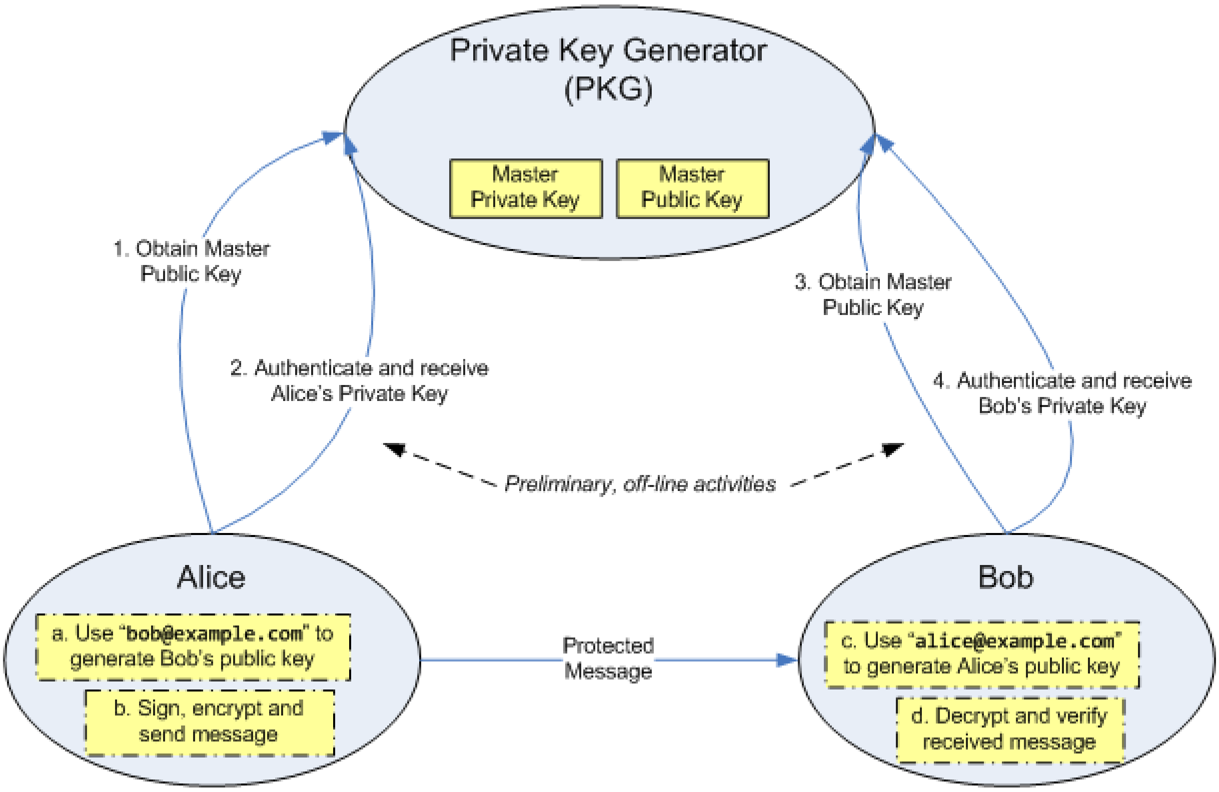

Private Key Generator (PKG)
Master Private Key
Master Public Key
1. Obtain Master Public Key
2. Authenticate and receive Alice's Private Key
3. Obtain Master Public Key
4. Authenticate and receive Bob's Private Key
Preliminary, off-line activities
Alice
a. Use "bob@example.com" to generate Bob's public key
b. Sign, encrypt and send message
Protected Message
Bob
c. Use "alice@example.com" to generate Alice's public key
d. Decrypt and verify received message

Figure 1.4 Identity based cryptography

1.4.4 Hybrid Mechanisms

A combination of any of the above 3 methods are used in encryption in order to leverage the advantages of the various techniques. For example, asymmetric key techniques are used to transmit the symmetric key used in encryption since symmetric key techniques are faster.

1.5 CRYPTANALYSIS AND CRYPTOGRAPHIC ATTACKS

Cryptanalysis or cryptographic attacks are aimed toward discovering and exploiting relationships between the plain text and the cipher text[18]. This could be used to guess the key or to detect patterns. Some of the most common attacks are as follows:

- **Brute Force Attack:**

The intruder tries out every possible key to determine the correct one. For a key size of n bits, there are 2^n possible keys. To the word such an attack key sizes must be of a very high order so that the number of possibilities is a high enough number to render the Brute Force attack infeasible.

- **Chosen-Plaintext Attack**

The attacker somehow obtains the cipher text for a plain text of his choice, encrypted using the secret key. Using this information, he tries to derive the secret key.

- **Known And Probable Plaintext Attack**

The attacker has access to cipher texts of one or more suspected plain texts, all encrypted using the same secret key. These plaintext-cipher text pairs are used to guess the secret key.

- **Eavesdropping**

The intruder watches the traffic of encrypted data, and uses this information for launching attacks at a later time. This is a form of passive attack and is most difficult to detect since the presence of the intruder is not known till a much later time when the attack is launched, by which time it might be too late.

- **Replay Attacks**

The intruder may intercept some messages in a conversation and replay these messages, thus tricking an entity into believing the message was sent by the other entity with whom it was exchanging messages, and not an intruder. Such attacks are harmful as the intruder may replay messages requesting sensitive information such as passwords. It is thus important to authenticate the recipient and the messages before sending responses.

- **Side Channel Attacks**

Rather than exploiting weaknesses in the algorithm, side channel attacks rely on the shortcomings of the Physical implementation of cryptosystem. Thus, knowledge of the algorithm is not a necessity for such attacks. Attacks launched without knowledge of the algorithm itself are termed as black box attacks.

Side channel attacks may be of several types as summarized below:

A. Timing Attack

The intruder uses such information as the timing between consecutive messages to draw conclusions about the encryption or computation time encryption time is proportional to the me e. Since in the normal case, the message length, this timing information can be used to guess the message length and in some cases, the message itself.

To thwart such an attack, padding techniques can be used to convert all the messages to equal length before encrypting or use pauses or time lags to confuse the attacker.

B. Power-Monitoring Attack

The power consumed by the hardware varies depending on the computation. Some information may be gained based on the power consumed while performing the computation, and this can be used to launch an attack.

C. Acoustic Analysis

This attack depends on the sounds released during computation and this information is exploited to launch an attack.

D. Differential Fault Analysis

By intentionally introducing faults in the computation, secrets are discovered, breaking the cryptosystem.

E. Data Remnance

This attack uses information that was not successfully deleted, and the user is thus unaware of the presence of the data. Such remnant data can give valuable information that could be used to break the entire cryptosystem.

F. Row Hammer

The storage mechanism can be exploited by using indexing to access adjacent memory locations that are actually off-limits to the intruder. Thus the intruder may gain access to data which he is not authenticated to access.

- **Pattern Analysis**

The cipher text may contain patterns. Observations of such patterns may give the intruder some idea of the plain text. For instance, if the plaintext is known to have a particular word occurring at two different positions and these are encrypted the same way both times it is encountered, then a pattern can be detected wherein the encrypted form of one word can be identified. Using this, the key may be guessed.

This could also be a major concern in images where patterns in the form of silhouettes or outlines of the components of the image may remain after encryption. To thwart such an attack, the encryption must be randomized to ensure patterns are never encrypted the same way.

- **Frequency Analysis**

An extension of pattern analysis is the frequency analysis. Statistics can be used to determine some patterns in a language. For instance, it is known that e is the most commonly occurring alphabet in the English language. If on encryption, e is always encrypted to the same value, a statistical analysis of the ciphertext will provide the most frequent character and this would be the encryption of e. This mapping can be used to identify the key.

The countermeasure is the same as in the case of pattern analysis that recurring alphabets are encrypted differently each time they are encountered.

- **Man In The Middle Attack**

The man in the middle attack involves the intruder stealing the identity of one of the entities in the conversation and pretending to be that entity, while the other entity continues to send messages without realizing it is in fact communicating with an imposter. This is a particularly serious attack in the case of El-Gamal encryption schemes since the public key of the imposter may be used instead of the public key of intended recipient and since the exchange of parameters takes place between the two entities, there is no way for one entity to know the true identity of the other entity.

This can be avoided through the use of certificates and Trusted Third Parties to authenticate both the entities before any data transfer occurs.

- **Active Attack**

The intruder deletes or modifies legitimate messages. These attacks are easy to detect since they lead to visible effect instantly. Deleted messages can be detected through the use of acknowledgements. A missing acknowledgement from intended recipient implies that the message was deleted. Modifications can be detected through the use of checksums appended to the message. Upon receiving a message, the recipient computes the checksum on the received message and compares it with the appended to the message. If they are equal, the message received it correct. If they are different, the message has been modified.

- **Birthday Attack**

The birthday attack is mainly used in the case of hash functions in which an intruder identifies two messages that hash to the same value and replaces one with the other. This attack can be avoided by the use of a strong hash function which results in unique mappings such that no two inputs map to the same value.

The main idea behind encryption is to convert the plaintext to unintelligible forms and make the process of mapping from plaintext to ciphertext as random as possible.

The process leverages some mathematical construct that is simple to compute in the forward direction, but is infeasible in the reverse direction, namely the trapdoor function.

Using multiple keys to encrypt the data might help improve security. However, it also slows down the encryption and decryption processes.

1.6 ASYMMETRIC (PUBLIC) KEY ALGORITHMS

Asymmetric key cryptography techniques depend on the infeasibility of some mathematical problem.Apart from their use in cryptography, asymmetric key cryptographic techniques can also be used for the authentication of entities. This works as follows:

The creator of a document signs the document using his private key. Any entity that wishes to check the authenticity of the document may use the public key of the expected creator to verify the signature. This will provide successful authentication only if the expected creator has signed the document since only the pair of corresponding public and private keys are complementary.

Thus, in the case of encryption, the entity uses the public key of the recipient to secure the data, while the recipient uses his own private key to decrypt the message.

However, in the case of authentication, the sender signs the document with his own private key, while the second entity verifies this signature using the public key of the expected creator.

There are three public key cryptosystems known today, namely RSA, El-Gamal Encryption and Elliptic Curve Cryptography.

1.6.1 RSA Algorithm

RSA, named after its three inventors Rivest, Shamir and Adleman, is the one of the most widely used cryptographic techniques today. Since Elliptic Curve Cryptography aims to overcome the shortcomings of RSA and become the standard of encryption, a basic understanding of RSA is crucial.

The trapdoor function used in the RSA Algorithm, known as the Integer Factorization Problem, is as follows:

While it is easy to compute the product of two extremely large primes, it is nearly impossible to compute the primes, given the product.

The above is true only for extremely large prime numbers. The algorithm can be summarized as follows:

STEP 1: Select two extremely large, random primes p and q.

STEP 2: Compute $n = p*q$.

n is the modulus for both the public and private keys.

STEP 3: Compute Euler's totient function $Ø(n)=(p-1)(q-1)$.

STEP 4: Select an integer e $(0 < e <Ø(n))$ which is co-prime to $Ø(n)$, that is $gcd(e, Ø(n)= 1)$.

e is now public key exponent.

STEP 5: Compute d such that e*d= 1 (mod Ø(n)), d is the private key component. Thus, the public key is (n,e) and the private key is (n,d).

Encryption

cipher text c = m^e (mod n). (1)

Decryption

Plaintext m=c^d (mod n). (2)

Example:

Let p = 7 and q = 13

n = 7 * 13 = 91

Ø(n) = 6 * 12 = 72

Let us choose e = 5 since gcd(5,72) = 1

Calculating d such that ed = 1(modØ))

That is, ed + Øk = 1

for some integer k

Thus, d = 29

Thus, for a message M,

Encryption function: E(M) = M^5 mod 91

Decryption function: D(M) = M^{29} mod 91

Consider a message M = 10

E(M) = E(10) = 10^5 mod 91 = 82

D(E(M)) = D(82) = 82^{29} mod 91 = 10

Shortcomings of RSA

The RSA algorithm has been in use for many years. However, with improvement in technology, machines are now able to factorize much larger products into their corresponding primes. Thus, larger keys are needed to ensure security. Larger keys increase the computation time for the encryption and decryption process and hence, alternative methods are under scrutiny. As per current acceptable standards for security, RSA requires keys of 1024 bits to be secure.

1.6.2 El-Gamal Encryption

Since the proposed algorithm is an extension of the El-Gamal encryption scheme but incorporating Elliptic Curves, the following section provides an overview of the El-Gamal algorithm.

El-Gamal Encryption algorithm is a method based on the Diffie-Hellman key exchange protocol, explained in section 2.3.2.1. It is a method of public-key cryptography that uses asymmetric-key cryptographic techniques.

The El-Gamal encryption consists of three phases: Key Generation, Encryption and Decryption.

Key Generation:

1. Alice computes a finite cyclic group G of order q using a generator g.

2. Alice chooses a random x from (1,2... q-1). This is her private key.

3. Alice computes $h = g^x$.

4. Alice publishes h, G, g, q as the public key.

Encryption:

Let the plain text message be m, which needs to be sent from Bob to Alice.

1. Bob selects a random value y from the group G described by Alice in her public key.

2. Bob then computes the shared secret as shown by the Diffie-Hellman key exchange protocol as:

$s = h^y$ where h is Alice's public key.

3. Bob converts his plain text m to a member of the group G, namely m'.

4. Bob computes the cipher text c_2 = m's.

5. Bob sends a combination as the ciphertext c = $(c_1,c_2)= (g^y, m' h^y) = (g^y,m' (g^x)^y)$

y is called an ephemeral key and is changed for every message. This is because once m' is known, h^y by can be computed from the ciphertext c_2.

Decryption:

Alice receives the ciphertext c.

1. Alice computes the shared secret s = $c_l^{x.}$

2. She then computes m' = c_2 s^{-1}, which is then converted to rn.

This works because:

$$c_2s^{-1} = m' (g^x)^y (g^{yx})^{-1} = m' g^{xy} g^{-xy} = m'$$

El-Gamal encryption is used in hybrid cryptosystems. That is, the cryptosystems in which a symmetric key is used for encryption and decryption and the shared key is transmitted from the sender to the receiver using El-Gamal encryption techniques. This is mainly because of the processing time of El-Gamal, which is on the higher side for entire messages.

The security of the El-Gamal technique depends on that of the underlying group G.

One of the most prominent attacks on El-Gamal encryption is the man-in-the-middle attack. This makes entity authentication by certificates or Trusted Third Parties an important aspect of El-Gamal scheme implementations.

Diffiee-Hellman Key Exchange Protocol

The Diffie-Hellman key exchange protocol is one used for the exchange of cryptographic keys between two entities. This method involves some information kept a secret and some publicly known information. The security behind this method comes from the fact that the two entities never exchange the private information as it is. A certain combination of the public and secret information, which is computationally infeasible to break even by the modern day supercomputers in a reasonable amount of time is transferred back and forth between the two entities. Some further arithmetic on these values by the entities using their private information generates a common key between the two entities that will be known only to them.

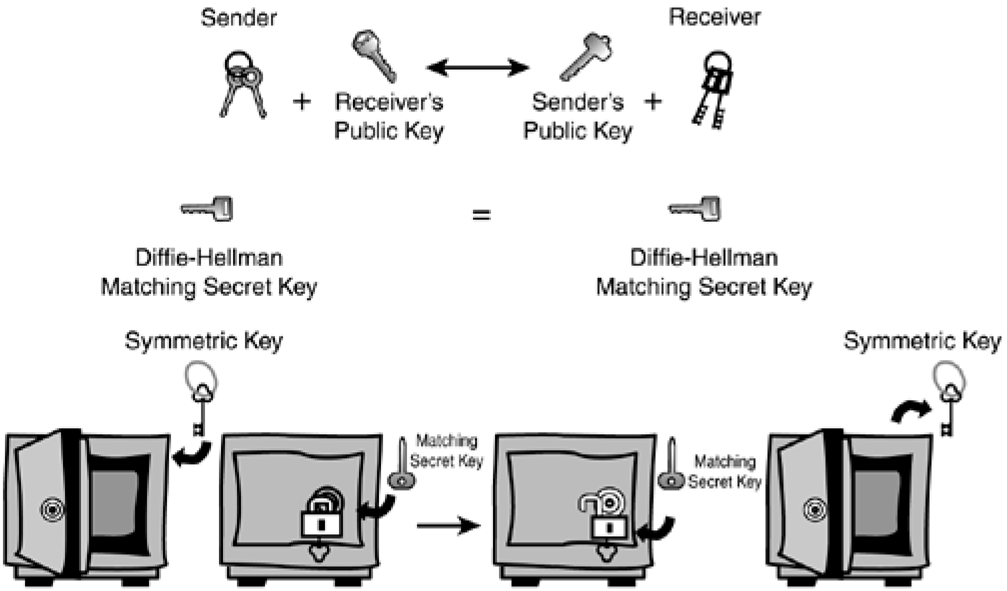


Figure 1.5 Diffie-Hellman Process

The algorithm for Diffie-Hellman Key Exchange is as follows:

1. A publicly known value (base) g is chosen.

2. Another public value p is chosen as the modulus for modular arithmetic.

3. Alice and Bob decide on their personal keys, a and b respectively.

4. Alice sends to Bob : $A = g^a \bmod p$

5. Bob sends to Alice: $B = g^b \bmod p$

6. Alice computes

$$s = B^a \bmod p = g^{ab} \bmod p.$$

7. Bob computes

$$s = A^{b} \bmod p = g^{ab} \bmod p.$$

8. This value s is now a shared secret between Alice and Bob. Also, since a and b are private and not known to anyone else, s cannot be computed by any party other than Alice and Bob. Also, knowing s and g, a third party could identify ab. However, using the same trapdoor function as that of RSA, it is practically infeasible to obtain either a or b from ab, thus ensuring that the above technique is secure.

Example:

Given that Alice and Bob share the following parameters: p = 23 and g = 5.

1. Alice selects the value a = 6 as her private key. She computes A =g^{a} mod p = 5^{6} mod 23 = 8 and sends it to Bob.

2. Bob selects a value b = 15 as his secret key. He computes B = g^{b} mod p = 5^{15} mod 23 = 19 and sends it to Alice.

3. Alice on receiving B computes s = B^{a} mod p = 19^{6} mod 23 = 2.

4. Bob on receiving A computes s = A^{b} mod p = 8^{15} mod 23 = 2.

5 The value 2 is now the shared secret between Alice and Bob.

1.6.3 Elliptic Curve Cryptography

Elliptic Curve Cryptography refers to the techniques of encryption and decryption developed using Elliptic Curves as the base of the algorithm [2][3]. Elliptic Curve Techniques can be employed for cryptography as well as for digital signatures used for entity authentication. Elliptic Curve Cryptography is considered in detail in the rest of this work.

Chapter 2
DNA SEQUENCE

2.1 INTRODUCTION

The science of DNA sequencing deals with identifying the order of nucleotides, which remains same as per DNA molecule. It compromises of several methods or technology that aims to determine the order of the four bases— cytosine, adenine, guanine, and thymine—in a strand of DNA. In the early 1970s, academic researchers obtained the first DNA sequence based on two-D chromatography. The sudden evolution of DNA sequencing and Sequence Alignments methods has greatly induced the field of biological research. There is many application area of DNA sequencing such as forensic biology, virology, biotechnology and biological system. The ability of DNA to carry massive parallel information has simulated the use of DNA to hide the information and utilizing then for cryptographic purpose. Thus, giving rise to new domain 'DNA cryptography' and which is continuously growing with the exploring of DNA computing.

2.2 DNA CRYPTOGRAPHY

There is lot of research going on since mid 90's, in the field of DNA Computing. Today, the science of DNA computing posses a high level computational ability and have the potentiality to solve huge and complex mathematical problems. The massive parallel nature of DNA and ability to bear the extraordinary information density is a key feature which has been efficiently utilized by the researchers for data hiding purposes and all sort of cryptographic purposes, giving rise to new field known as DNA Cryptography. There is very fast evolution in DNA cryptography and it continuously growing with the advent of fields such as DNA computing.

2.2.1 Biological Background of DNA

DNA is acronym for deoxyribonucleic acid which is germ plasma of all lifestyle. It is a biological macromolecule and is made up of nucleotide. There are Nucleic acids, which consists of a chemical string of interlinked attribute called nucleotides. Each nucleotide comprises of three things: a sugar (ribose in RNA, deoxyribose in DNA) a phosphate group and which act as integral unit of nucleic acid strand, and attached to the sugar is one of a set of nucleobases. These nucleobases are accountable for double helical structure of DNA which participates in base pairing of DNA strands to form higher-level structure as secondary and tertiary [5][6].

The four nucleotide bases of DNA strand is shown in the table 1 representing the four nucleotide bases of a DNA strand, which is covalently linked to a phosphodiester backbone.

Table 2.1 DNA Nucleotide Base

.

A	Adenine
T	Thymine
G	Guanine
C	Cytosine

From figure 2.1(as per their chemical structure), the four nucleotides A, C, G, T can be divided into two classes:

1. Purine R= {A, G} and Pyrimidine Y= {C, T}

2. Amino Group N = {A, C} and Keto Group K= {G, T}.

Apart from these divisions, further bifurcation can be made on the basis of the hydrogen bonds, i.e. how strong the bond is? Strong H-bonds S= {G,C} and Weak H-bonds W={A,T}.[8]

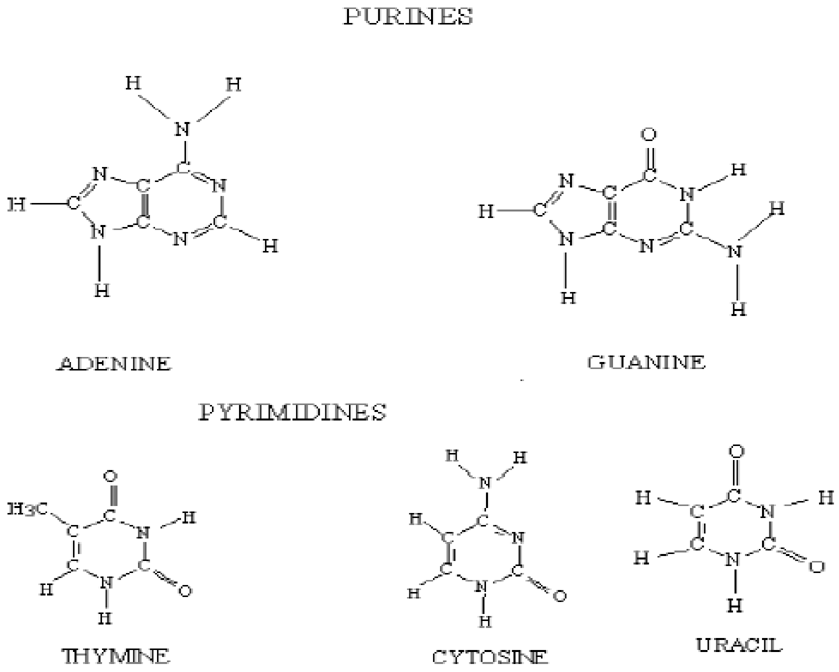


Figure 2.1 Structure Of Purines and Pyrimidnes

A gene is DNA sequence which carries the genetic information. Within a gene, a messenger The RNA sequence can be defined on the basis arrangement of bases in a DNA strand. The translation and transcription technique are collectively known as genetic coding. [9][10] It defines the relationship between the amino-acid sequences of proteins and the nucleotide sequences of genes. The word "Genetic Code" comprises of 'words' of three-letter called codons composed from a sequence of four nucleotides bases (e.g. ACT, CAG, TTT). In transcription, a DNA segment is transformed into messenger RNA (mRNA) by RNA polymerase, which exits the nucleus and enters into the body of a cell. In translation, the encoded information in mRNA is decoded by ribosome and assembles amino acid into protein chains.

2.2.2 DNA Sequencing

DNA sequencing is the process of determining the precise order of nucleotides within a DNA molecule. Any method or technology that defines the ordering of four nucleotide bases i.e. adenine, guanine, cytosine, and thymine in a DNA strand. Sequencing methods of DNA has greatly fastens the medical and biological research. The field of natural sequence pattern along with chemical classification and complementary genetic coding is used to protect or hide the message.

Two sequences can be said as complementary sequence, if their base position is complementary to each other and also when the order is reverse. The arrangement of series of codon in a mRNA molecule is shown in Fig 2.2 describing the complementary properties of nucleotide base. For example, the complementary sequence of ATGC is TACG. Thus, if one is a sense strand then the other is antisense strand and shows complementary behavior to the other strands.

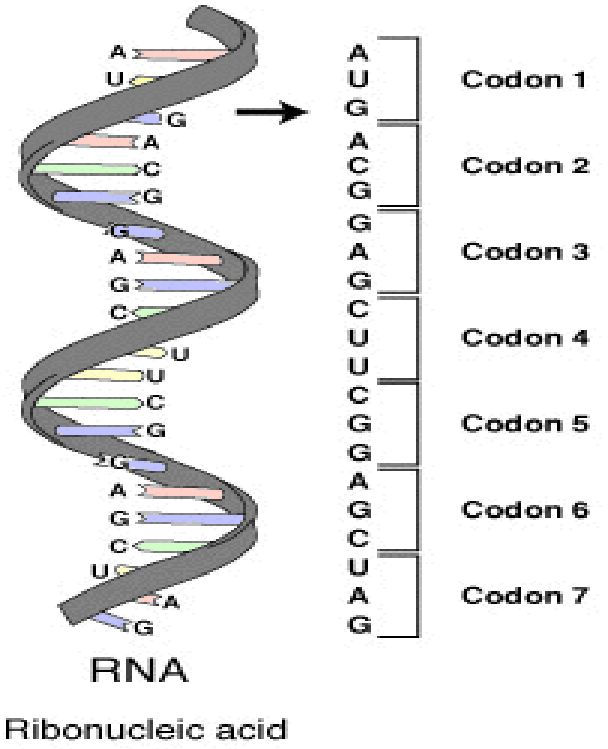


Figure 2.2 Series of codon in mRNA molecule.

There is one special property for DNA sequences i.e. the original DNA sequence and the faked DNA sequence will almost look like the same. And, there are also a large number of DNA databases which are publicly available. By using these facts, in this paper a new methodology is formed for encrypting messages using DNA sequences.

DNA sequences offer a unique method of encrypting messages or information. The main advantage of DNA sequences is they are composed of letters which are meaningless for most people. The DNA sequence is a combination of A, C, G and T base pairs. [26]

Using these properties of DNA sequences, three complementary rules can be formed and subsequently be used to generate fake DNA sequences. DNA sequence serves as an ultra-compact information storage medium which stores a large amount of data in compressed form. A single gram of DNA contains 1021 DNA bases = 108 tetra bytes.

Thus, these characteristic of DNA:

- Massive parallel computing
- Large data storage
- Information carrier (mRNA)
- Genetic coding
- Generation of faked and random DNA sequences
- Searching complexity of a particular DNA sequence in large database

Has increased the possibility of using DNA and gives a prominent direction in cryptographic research.

There are several DNA-based algorithms that have been practically applied for cryptography purposes. Kang Ning proposes a method in which sender uses the original DNA sequence to encode its secret message and performs transcription and translation obtaining a protein which act as a public key for the receiver. Ning also proves that this cryptography method is secure against many intruder attacks like replay attack, brute force attack though he acknowledges that the encryption complexity increase with key size. [10]

Debnath Bhattacharyya proposed a new data hiding methods based upon DNA complementary rules and message indexing. In this indexing of a random DNA sequence is done which is used as a reference for encoding the message in DNA sequence and during decryption process one requires the DNA string and Index mapping to obtain the original message. [11]

In 2010, H.J. Shiu, K.L. N proposed a more robust method for hiding data.[27] He introduced the insertion method, substitution method and complementary pair method for hiding data in DNA sequence. In the insertion method both the reference DNA sequence and the secret message are assembled from the sequence and the secret message after decomposing. In the Complementary Pair Method the complementary rules are used to encode the secret message. In the substitution method, another letter is substituted for an existing letter decided by the algorithm substitution rule.

The DNA-crypt algorithm has also been used for image cryptography. Qiang Zhang, Ling Guo proposed a new scheme of encrypting image using DNA sequences.[37] The proposed approach defines two new mathematical operation on DNA sequences i.e. addition and subtraction operation which helps in combining the encoded matrix block of DNA sequences and then using complementary rules for the output of added matrix block by using chaotic dynamical system(Logistic Map).

Jin-Shiuh Taur et al. [38] proposed a method for improving the effectiveness of the substitution method, known as Table Lookup Substitution Method (TLSM), this methods enhance the message hiding capacity twice. In TLSM, the extended the complementary rule definition and introduce a 2-bit rule table instead of 1-bit rule table. Thus, while encoding allowing two bits of secret message to be encoded.

2.3 SEQUENCE ALIGNMENT

In bioinformatics, sequence alignment is one of the integral techniques. It implies the arrangement of the DNA, RNA, or protein in order to judge the similarity among the sequences based upon functional, structural or evolutionary relationship. The evolutionary analogy of a nucleotide or amino acids sequence is given by the degree of sequence conservation, whereas the degree of variation shows the deviation that have occurred in the form of insertions, deletion and substitution during evolution. We need a minimum of two sequences to be aligned known as pairwise alignment and maximum varies to thousand. The Alignment of sequences of nucleotide and amino acid are illustrated in row major matrix form.

Sequence homology is an important phenomenon in sequence alignment analysis. Two sequences are said to be homogonously related when they are descended from a same evolutionary origin. A related but different term is sequence similarity, which is the percentage of aligned residues of a nucleotide or amino acid that are alike in physiochemical properties such as charge, size and hydrophobicity. Sequence homology is an illation obtained from sequence similarity comparative analysis about common ancestral relationships. Whereas, similarity is a direct result of experimental observation from the sequence alignment. Sequence similarity is quantitative measure and homology is qualitative inference. For example, if two nucleotide sequences share 45% similarity, then we cannot say they share 45% homology. They are either homologous or nonhomologous. Inferring homologous relationships from a particular similarity level depends upon type of sequence and sequence length. Length of nucleotide or amino acid sequence is also an important factor. Longer sequences needs lesser cut-offs for inferring homologous than shorter sequences.

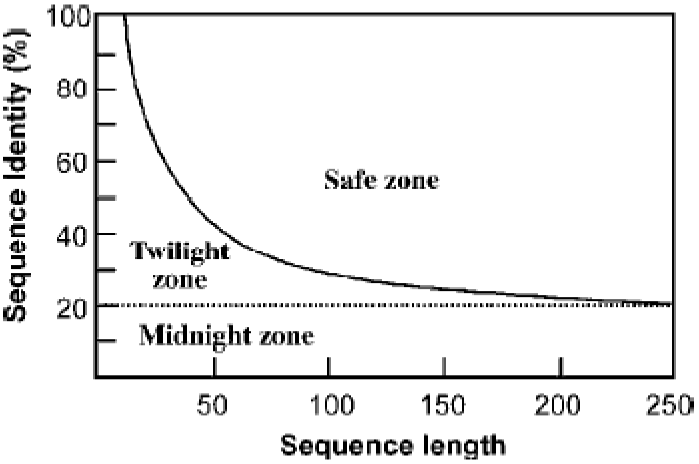


Figure 2.3 Three zones of nucleotide sequence alignments.

The zones of nucleotide sequence alignment is shown in the Fig 2.3, if the percentage sequence identity falls in safe zone, then the nucleotide sequences are homologous. If the percentage sequence identity falls in twilight zone, then the homologous relationships are less certain and in the midnight zone, it cannot be determined reliably.

Chapter 3

ELLIPTIC CURVE CRYPTOGRAPHY

Elliptic curve cryptography is a type of asymmetric key cryptosystem which uses the algebraic phenomena of elliptic curve having finite field to provide the security of data. Elliptic curve point arithmetic has been used to convert a message or an input into a pair of coordinates that lies on Elliptic curve [16][17].

3.1 ELLIPTIC CURVES

Mathematically Elliptic curves are represented as,

$$Y^2=X^3+aX+b \qquad (3)$$

here different value of 'a' and 'b' gives the different elliptic curve[4]

Equation having the discriminant

$$\Delta = 4a^3+27b^2 \mathrel{!}= 0 \qquad (4)$$

The graph of an elliptic curve is as shown in figure

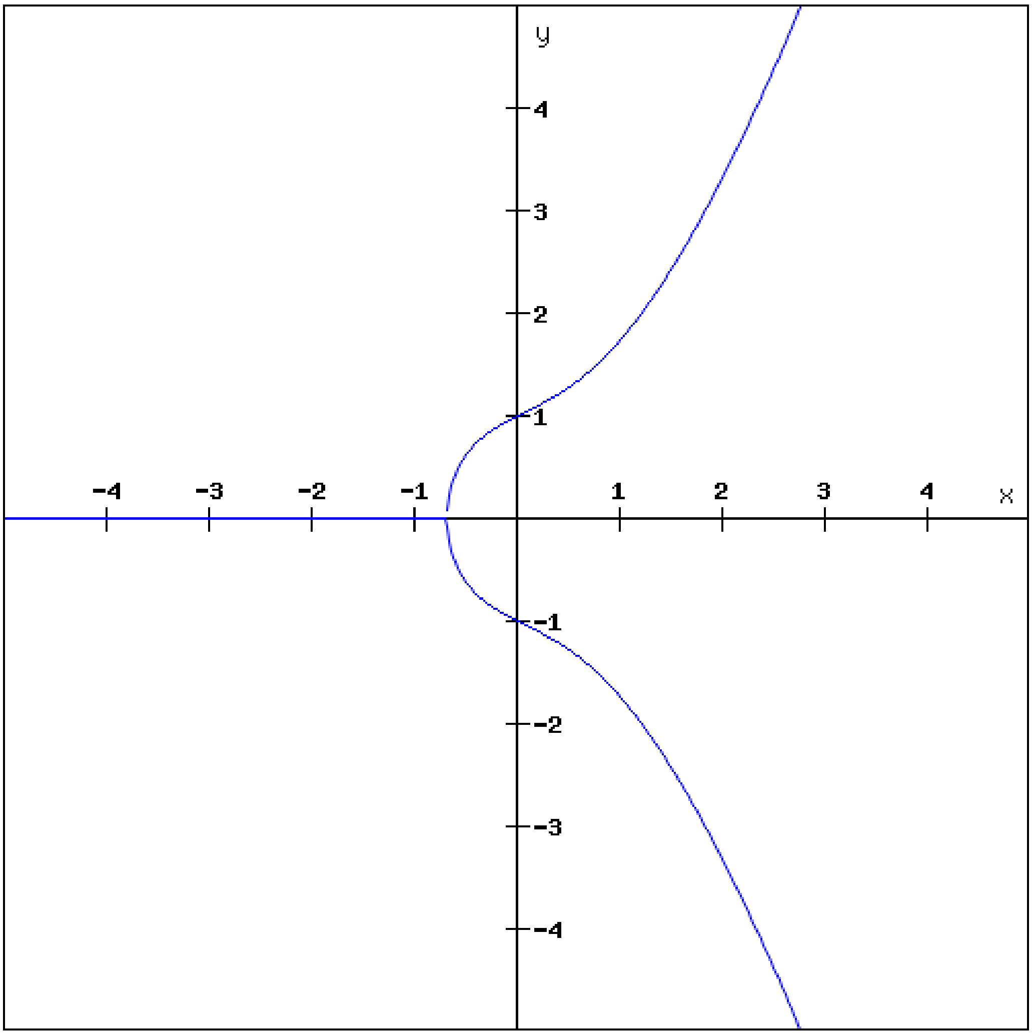


Figure 3.1 Graph of Elliptic curve $Y^2=X^3+X+1$

3.1.1 Characteristics Of Elliptic Curves

Elliptic curve has the following properties that make it suitable to use in cryptography[21]

I. If a line intersects at two points of curve, it surely intersects at third point.

II. If a line is tangent to the curve, it will intersect another point.

III. Those lines whose slope is infinity will intersect the point at infinity.

IV. If we increases the slope of curve and becomes infinity, it also intersects at infinity.

3.1.2 Protocols That Involve In Elliptic Curve

I. Elliptic Curve Digital Signature Algorithm (ECDSA)

II. Elliptic Curve Integrated Encryption Scheme (ECEIS)Elliptic Curve Diffie Hellman Key Agreement Scheme

III. ECMQV key agreement scheme

IV. ECQV implicit certificate scheme

3.1.3 Applications Of Elliptic Curves

I. Digital Signatures

II. Pseudo-random number generators

III. Integer factorization algorithm used in cryptography, eg. Lenstra Elliptic Curve Factorization.

IV. Encryption.

3.2 TRAPDOOR FUNCTION IN ELLIPTIC CURVE CRYPTOGRAPHY

Elliptic Curve Discrete Logarithm Problem is the trapdoor function which is used in Elliptic Curve Cryptography.

In Elliptic Curve Discrete Logarithm Problem there is base point which is publicly known and arithmetically it is not feasible to compute discrete logarithm of a arbitrary elliptic curve element regarding that base point

Or in different manner, if we have a point on an Elliptic Curve (i.e. original point) and multiplication operation is performed with a fixed prime number, which is referred to as product point. It is hard to compute the prime number by which the point has been operated. This is the trapdoor function because it is easy to perform multiplication of a prime number with a point but computationally hard to perform reverse of such operations. This ensures secrecy of the algorithm.

Suppose if there is a private key p and a point is I, we can compute the operation I'=p*I easily but with the help of I and I' it is not easy to compute the value of p.

3.3 ADVANTAGES AND DISADVANTAGES OF ECC

- **Advantages**

Today commonly used technique for encrypting the data is RSA but in this technique there are some drawbacks. To provide high level of security RSA uses key of large sizes. It is highly recommended to use keys of larger size because of the availability of systems having high computational power which can easily break integer factorization problem for smaller numbers.

Elliptic Curve Cryptography can overcome this shortcoming of RSA by using keys of smaller size which makes Elliptic Curve Cryptography better than RSA.

The benefit of shorter key size

I. Encryption and Decryption takes less time.
II. Small key size saves bandwidth.
III. Now this technique uses small certificates.
IV. Consumption of power is low.
V. Storage Mechanisms is more effective.

Now if we concern these advantages of ECC , it provides a feasible environment for mobile devices. The reason behind this, Mobile devices have some limitation of power and resources which are not able to support resource and power through encryption and decryption process.

Other scenario Elliptic curve discrete logarithm problem having more secure than Integer factorization problem which is used in other encryption technique such as RSA over ECC more secure and a more operable substitute.

- **Disadvantages**

The major drawback of ECC is that the size of encrypted message is notably longer than the other encryption techniques.Further ECC algorithm have more complication than other algorithm, which may compromise in terms of security.

3.4 APPLICATIONS OF ELLIPTIC CURVE CRYPTOGRAPHY

- **Bitcoin**

Bitcoin is a type of cryptocurreny which is used in digital currency. By the help of bit coin we makes on line payment between two parties without the help of any financial institution or payment gateway. Transactions of bit coin are stored in a form which is known as Bitcoin block chain. Elliptic Curve Digital Signature Algorithm is used to recognize the account of bit coin user.

Suppose we transfer a bit coin from user A to user B it takes basically two parameters to perform this transaction, namely, Digital Signature using A's private key of hash value of the previous transaction and another ,information about B's public key further signature will be verified using public key of A from the last transaction.

- **Transport Layer Security (TLS)**

When we implement transport layer security protocol ,it uses elliptic curve at many locations. Elliptic curve diffie-hellman key exchange policy is also used in the cipher suits. There are two types of ECDH keys : 1) Ephemeral Keys ,in this always a new ECDH keys is used to generate for every key exchange. 2) Long-term keys, in these different key exchanges reuse the same ECDH key.

- **Secure Shell (SSH)**

For user authentication, Elliptic curve cryptography may be used in secure shell in one of three ways:

A. Session Key Negotiation:

With the help of ephemeral Elliptic Curve Diffie-Hellman key exchange method used in SSH,SSH-2 performs negotiation of session keys.

B. Server Authentication

In server authentication process server authenticates the clients by the help of host keys of client, it may be ECDSA public keys. In the process of key exchange ,the host key of server is transmitted towards client from the end of server further client matches the server's host key to the key which is stored at client's end to complete the verification process. After this process, a transcript of the key exchange is signed by server that helps to authentication.

C. Client Authentication

In client authentication process, same algorithm is applied at the end of client by the help of their ECDSA public keys.

- **AUSTRIAN E-ID**

In this era the use of physical smart card is increasing and the main issue is to authenticate this card in a very short of time at any service point.

Basically cards store the private keys for encryption process and signatures. Some hardware modules are used to perform cryptographic computation. With the help of some decreased computational complexity, ECC may be a better substitute of RSA or large prime order groups for these physical smart card.

The national e-id cards in Austria contain the RSA or ECDSA public key which enables the provision of legally binding digital signature.

3.5 KOBLITZ'S METHOD

If we compare RSA and ECC on the basis of security, ECC is more secure for keys of smaller size.[14] In ECC, data is mapped to points on elliptic curve.[12] Different encoding techniques are used for this purpose. One of the best suited techniques is Koblitz's method. Our scheme has also used Koblitz's method for converting intensity of each pixel of the original image into points of elliptic curve. This is the pre-processing step of ECC. On the other hand, during decryption every point on elliptic curve is converted or deciphered into intensity value of original image.[13] This is post processing step of ECC. Pre and post processing steps are performed using Koblitz's method.

Koblitz's method is briefly explained below:

1. Choose an elliptic curve which is represented using following equation.

$$y^2 \bmod p = x^3 + a * x + b \bmod p$$

Where p is any prime number, x and y are the coordinates of an elliptic curve. a and b are constants that must satisfy following equation

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

2. Choose a parameter k such that if $x = m * k + i$ where i vary from 1 to k-1, then for at least 1 value there must exist a value for y which satisfies the equation of elliptic curve. Here m is input value which is converted into a point on curve (x,y).
3. For converting point (x,y) back to the input data following operation is performed and floor value of m is taken.

$$m = (x-1)/k$$

Whole Algorithm is explained with the help of following example:

Consider k=30 and prime number selected be 223 (p=223) and the input that is converted into curve points is 5 (m=5). Elliptic curve equation used in the encoding process be

$$y^2 \bmod 223 = x^3 + x + 1 \bmod 223$$

We get a point(178,22) which satisfies the above equation at i=28. For decoding

$$m = (178-1)/30$$

$$m = floor(5.9)$$

$$m = 5$$

Thus we obtained original input back. Method is illustrated in the following Figure 3.2

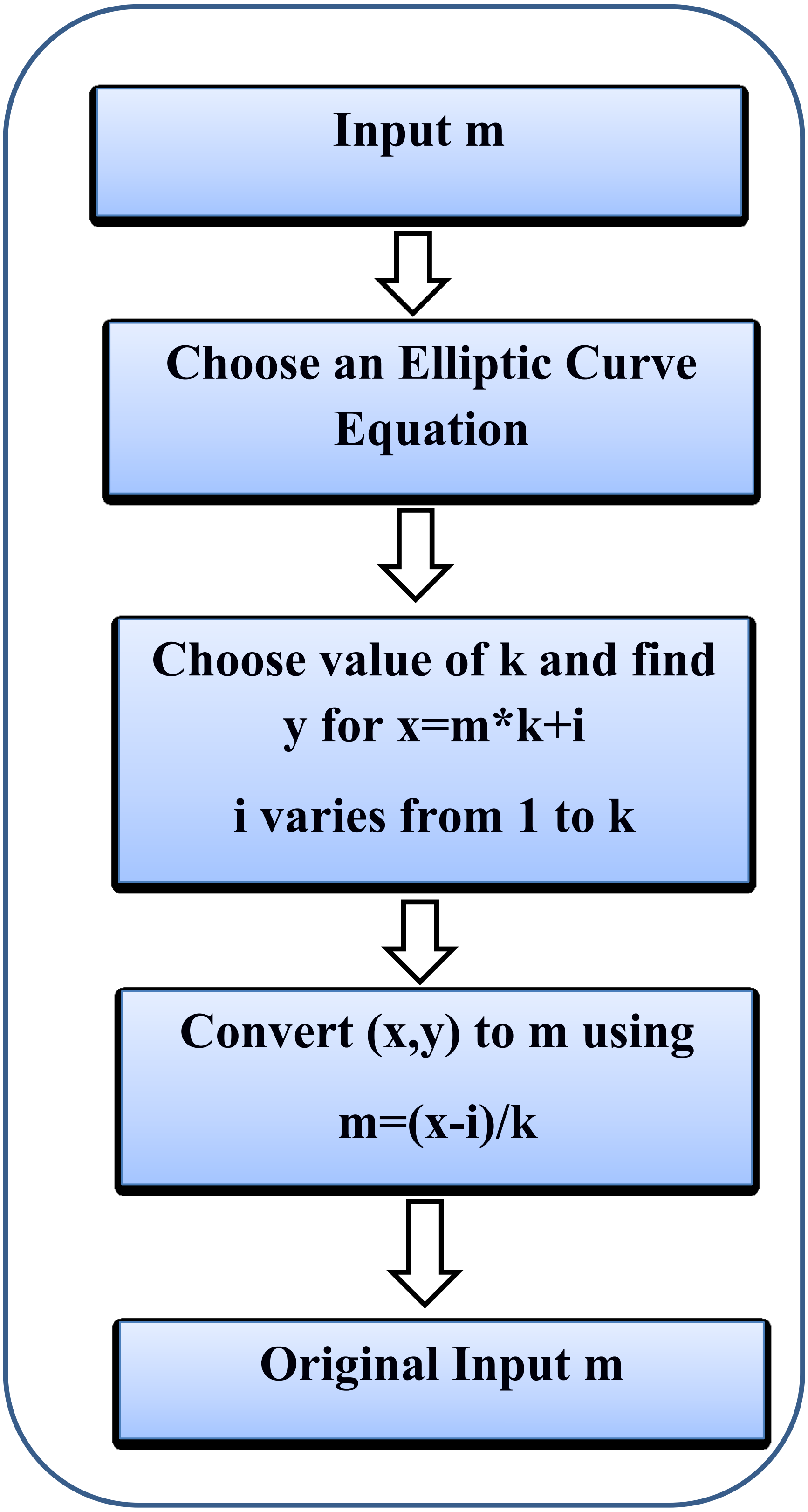


Figure 3.2 Diagrammatic representation of Koblitz's method

Chapter 4

MATHEMATICAL REVIEW

4.1 MODULAR ARITHMATIC

Modular Arithmetic is used to perform arithmetic operation within a finite field. If the limiting number is prime number, the finite field is known as a prime field F_p . Prime field consists a range of number from 0 to p-1 ,where p is a prime number and the result of arithmetic operation lies between the same range which is [0,p-1].

This technique of affine the numbers in a specific range by the help of arithmetic modular is known as wrapping around of value. This is able to be used for arithmetic operation in case of prime field.

Mathematically it represented as 'a mod b' which signifies that always the remainder of the operation a/b can never exceed the value of b. It means the value of a mod b will be mapped in the range of [0,b-1].

Examples:

Suppose p=29 ,a=20 and b=15

- **Addition:**

(a+b) mod p= 35 mod 29 =6

- **Subtraction:**

(a-b) mod p=5 mod 29 = 5

- **Multiplication:**

(a*b) mod p=300 mod 29 = 10

- **Multiplicative Inverse:**

Multiplicative inverse of a number b with repeat to mod p is defined as

$b^* b^{-1}$ mod p=1

By the help of Extended Euclidean Algorithm,we can compute the multiplicative inverse.

- **Division:**

a/b mod p= a* b^{-1} mod p

where b^{-1} is the multiplicative inverse of b with respect to p.

thus it is necessary to compute inverse of b before to complete the division operation.

- **MODULAR SQUARE ROOT:**

In a prime field F_p , Modular square root of a number n is defined as b such that

b*b mod p=n

Tonelli Shanks Algorithm is used to compute efficient modular square root.

4.2 POLYNOMIAL ARITHMETIC

In Elliptic curve arithmetic, it involves integers of length m bits over a prime field F2m because binary representation constitutes a binary field.

Suppose if there is m-bit of integer which is in binary form then it can be represented in the form of binary polynomial of degree (m-1).

The binary string ($a_{m-1....}a_1a_0$) can be expressed as a polynomial as this way:

$$a_{m-1}\, x^{m-1}+a_{1}x+a_0$$

where a_i is 0 or 1. Let's take a example, a 5 bit binary number 11001 can be expressed in polynomial form as x^4+x^3+1

Here the value of coefficients of the binary polynomial is either 0 or 1, suppose at any instance if the value becomes more then 1, at that condition we use modulo 2 operation to convert it into either 0 or 1.

Same way in modular arithmetic, the maximum degree of a polynomial will be m and if at any instance after applying any operation the value of degree will be more than m ,it will be reduce in range of [0, m-1] using the irreducible or reduction polynomial.

We can perform all arithmetic operation on binary integers by converting binary integers into polynomial equations and after performing the arithmetic operation on these polynomial equations, we can again convert it into corresponding binary integers.

Irreducible Polynomial:

An irreducible polynomial is a type of polynomial in which an arithmetic operation of two polynomial results, in a polynomial of degree more than the maximum degree. Here the degree of polynomial can be never greater than maximum, the main use of irreducible polynomial is to reduce the degree of polynomial in the range which is [0, max-1].

Examples

Let A= $1101_2 = X^3+X^2+1$

Let B= $0110_2 = X^2+X$

- **Addition**

$A+B=X^3+2X^2+X+1$

Using modulo 2 arithmetic,we can convert the coefficients as:

$A+B=X^3+X+1=1011_2$ =A XOR B.

- **Subtraction**

$A-B=X^3-X+1$

Taking mod 2 over coefficients

$A-B=X^3+X+1=1011_2$=A XOR B.

- **Multiplication**

$A*B=X^5+X^3+X^2+X$

Taking mod 2 over coefficients

$A*B=X^5+X^3+X^2+X$

since m=4 the results to be reduced to a degree less than 4 by irreducible polynomial X^4+X+1

such as $X^5+X^3+X^2+X$ (mod f(x))

$=(X^4+X+1)X+X^5+X^3+X^2+X$

$=2X^5+X^3+2X^2+2X$

$=X^3$

Converting the coefficient into the range [0,1] using modulo-2

$A*B=X^3=1000$

- **Division**

A/B (mod f(x))= $A*B^{-1}$(mod f(x))

where B^{-1} is the multiplicative inverse of B over f(x).

- **Multiplicative Inverse**

Multiplicative inverse of a polynomial equation is computed by the $B*B^{-1}$ (mod f(x))=1 with the help of any irreducible polynomial. We can use extended Euclidean algorithm to compute it and this computation is too expensive.

4.3 ELLIPTIC CURVE ARITHMETIC

4.3.1 Graphical Approach:

4.3.1.1 Point Addition:

The steps which are required to add two points are as follows:

STEP 1: A line is drawn joining the two points that need to be added.

STEP 2: Since a line intersecting an elliptic curve will also intersect it at a third point, we find the third point of intersection between the line and the elliptic curve.

STEP 3: The reflection of this point about the X-axis gives us the point that represents the sum of the required points.

An example of point addition is shown graphically in figure 2.

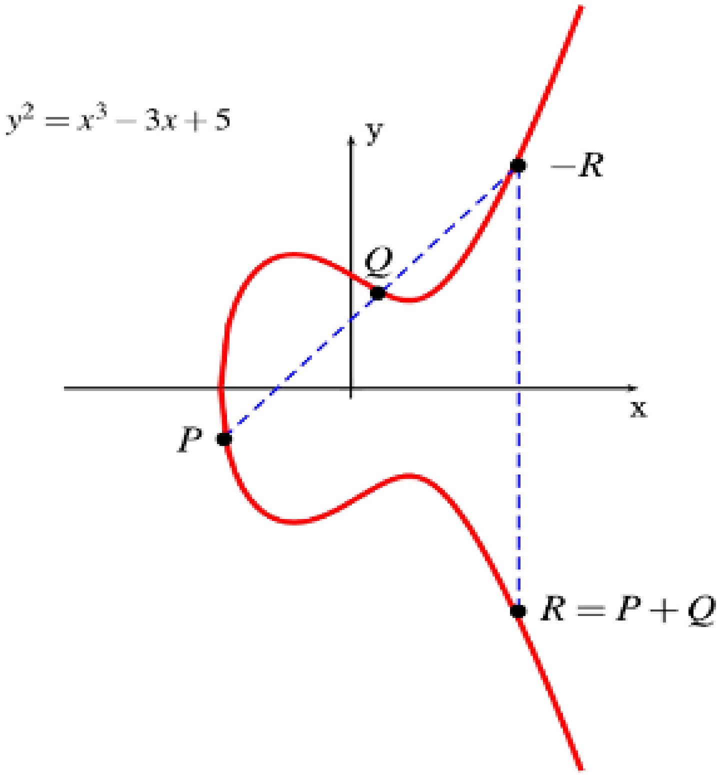


Figure 4.1 Graphical representation of point addition in elliptic curves

4.3.1.2 Point Doubling

The steps involved in point doubling are as follows: [15]

STEP 1: The tangent to the elliptic curve is drawn at the point that needs to be doubled.

STEP 2: Since we know that if a line is a tangent to the curve, it also intersects the curve at another point, we find this point of intersection.

STEP 3: The reflection of this point of intersection about the X-axis gives the double of the required point.

An example of point doubling is presented in this figure 4.2

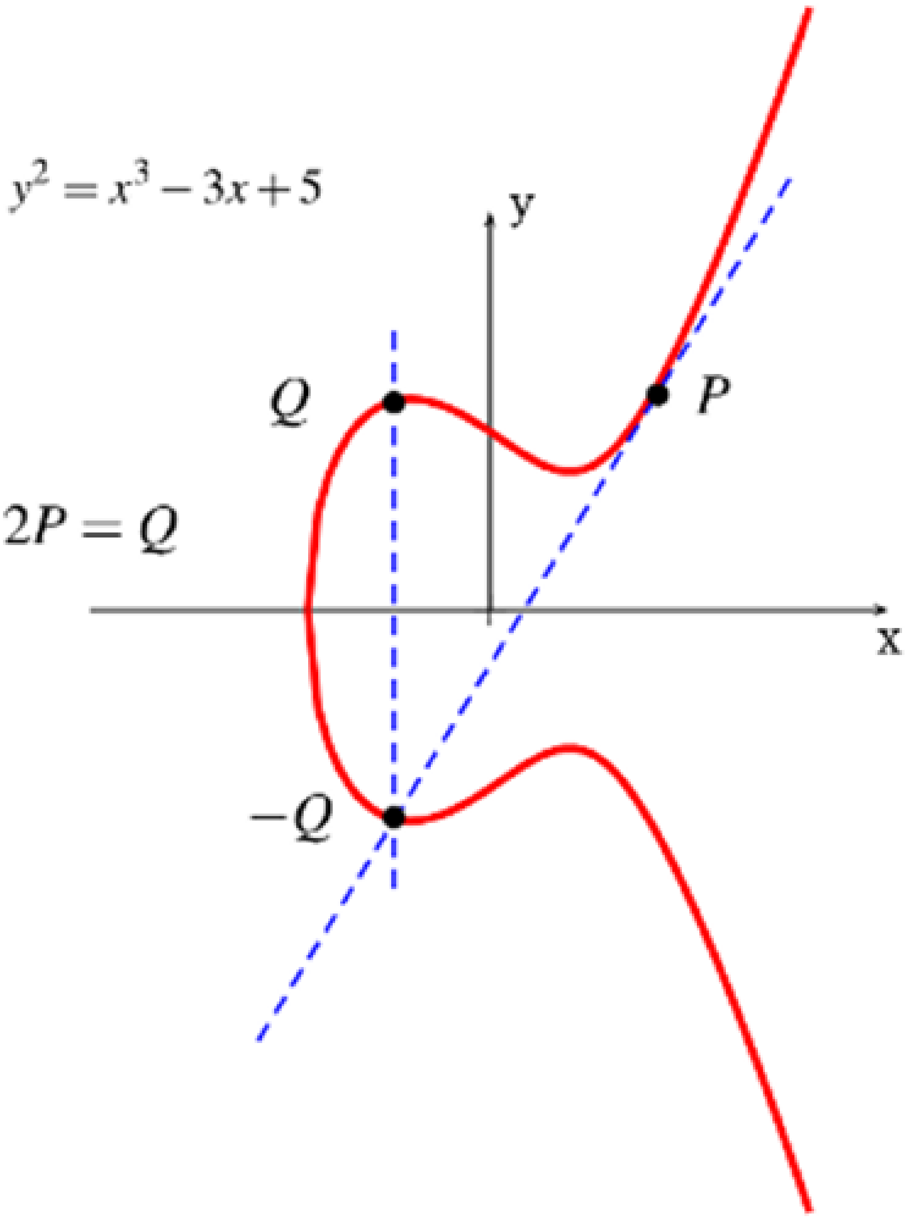


Figure 4.2 Graphical representation of point doubling in an elliptic curve

4.3.1.3 Point Multiplication

Multiplication of a point is mostly only done with a scalar constant. The process of multiplication can be done in two ways:

Assuming that a point is to be multiplied by a constant k:

Method 1: Repeated addition.

Efficiency:

To perform multiplication using this method, we use a loop that runs k times. Thus,the time complexity of this method is O(k).

Method 2: Double-and-add method.

For example, consider multiplying a point P by a value k=13.This multiplication can be broken down a series of addition and doubling as shown below:

Table 4.1 Double-and-add method of point multiplication

P	Operation
2P	Doubling
3P	Addition
6P	Doubling
12P	Doubling
13P	Addition

Efficiency:

On splitting a number k into addition and doubling, we get an average of log(k) addition and doubling.

To perform multiplication using this method, we use a loop. However the loop only runs log(k) times, making the time complexity of this method O(log(k)).

Thus, the second method is more efficient and used in the cryptography process suggested later.

4.3.2 Mathematical Approach

4.3.2.1 Point Addition

Let the two points to be added be represented as $P_1=(X_1,Y_1)$ And $P_2=(X_2,Y_2)$.

Let the sum of the two points be represented as $P_3=P_1+P_2=(X_3,Y_3)$.

Let the slope of the line joining P_1 and P_2 be represented by m.

Then,

$$x_3= m^2 - x_1 - x_2 \qquad (5)$$

$$\text{and } y_3=m*(x_1 - x_3) - y_1$$

where $m = (y_2 - y_1) / (x_2 - x_1)$

Proof:

we know that the slope of the line joining the points P_1 and P_2 is given by:

$$m = (y_2 - y_1) / (x_2 - x_1)$$

using slope intercept form:

$y - y_1 = m * (x - x_1)$

$y = mx - m*x_1 + y_1$

let $p = y_1 - m*x_1$

$y = mx + p$

Squaring to both side

$y^2 = (mx + p)^2$

we know that equation of the elliptic curve is given curve is given by $y^2 = x^3 + ax + b$

$(mx + p)^2 = x^3 + ax + b$

$0 = x^3 - m^2x^2 - 2mxp - p^2 + ax + b$

The above equation is a cubic in x with roots x_1,x_2 and x_3.

Thus , sum of roots = –(coefficient of x^2)

$x_1 + x_2 + x_3 = m^2$

Thus, $x_3 = m^2 - x_1 - x_2$.

And $y_3 = m * (x_1 - x_3) - y_1$.

4.3.2.1 Point Doubling

Point Doubling is a special case of point addition in which the two poitns to be added are the same . Thus,

$$x_3 = m^2 - 2x_1. \qquad (6)$$

and $y_3 = m * (x_1 - x_3) - y_1$

where $m = (3x^2 + a) / 2y$

Proof:

Since we consider the tangent to the elliptic curve at tha point that needs to be doubled, the slope is given by the differential of the curve at that point.

$d(y^2) / dx = d(x^3 + ax + b) / dx$

$2y\ dy / dx = 3x^2 + a$

$m = dy / dx = (3x^2 + a) / 2y$

Proceeding in the same way as for addition but applying $x_1 = x_2$ and $y_1 = y_2$,

Obtained these values :

$x_3 = m^2 – 2x_1$

and $y_3 = m * (x_1 – x_3) – y_1$

where $m = (3x^2 + a) / 2y$.

Special Cases:

Point At Infinity:

The point at infinity O is a point assumed to lie on every vertical line.

In case of point addition:

while computing the sum of two vertically aligned points, the line joining the two points does not intersect the curve at a third point. We assume the third point to be O.

The point at infinity is the additive identity of elliptic curves.

Given a point A and point at infinity O,

A + O = A

And A + (-A) = O

In case of point multiplication:

Let the multiplication of a point P, say yP = O. The tangent to that point is a vertical line which does not intersect the curve at a second point. We thus assume the point of infinity to be second point of intersection on the curve.

Then (y+1)P = yP + P = O + P =P

(y+1)P = O , (y+3)P = P, (y+4)P = O and so on.

Chapter 5

PROPOSED METHODOLOGY

The proposal of this methodology contains three approach DNA sequencing, koblitz's method and elliptic curve cryptography and applies on the image file. This methodology basically overcomes the shortcomings of existing methodology which is DNA computing with RSA and the proposed methodology is containing DNA computing with elliptic curve cryptography along with koblitz's method which is used here for data encoding. One level of security achieves by mapping of DNA nucleotide with the intensity value of image file. DNA mapping can store large amount of data within few nucleotide. Thus combination of elliptic curve cryptography and DNA sequence provides a higher level of security having less computation and complexity. The other advantage of this methodology is key size of ECC. ECC key size is smaller than RSA key size e.g. ECC-224 has same security level as RSA-2048. One more thing which makes it more efficient that is koblitz's method it converts mapped DNA intensity into elliptic curve points whereas existing scheme uses a simple database of intensity vs elliptic curve points.

5.1 METHODOLOGY PARAMETERS

5.1.1 MAPPING TABLE

In this mapping table we mapped each intensity value (0,1,2….,255) of a gray image to a DNA sequence which contains value of DNA nucleotide A,G,C,T.

This DNA nucleotide converts to a decimal number by another mapping as A=1,C=2,G=3,T=4.

Let

I= intensity value

D=DNA nucleotide

And N=Decimal number

Table 5.1 Mapping table of DNA nucleotide and Decimal Number with intensity value of image.

I	D	N	I	D	N	I	D	N	I	D	N
0	AAAA	1111	64	CAAA	2111	128	GAAA	3111	192	TAAA	4111
1	AAAC	1112	65	CAAC	2112	129	GAAC	3112	193	TAAC	4112
2	AAAG	1113	66	CAAG	2113	130	GAAG	3113	194	TAAG	4113
3	AAAT	1114	67	CAAT	2114	131	GAAT	3114	195	TAAT	4114
4	AACA	1121	68	CACA	2121	132	GACA	3121	196	TACA	4121
5	AACC	1122	69	CACC	2122	133	GACC	3122	197	TACC	4122
6	AACG	1123	70	CACG	2123	134	GACG	3123	198	TACG	4123
7	AACT	1124	71	CACT	2124	135	GACT	3124	199	TACT	4124
8	AAGA	1131	72	CAGA	2131	136	GAGA	3121	200	TAGA	413
9	AAGC	1132	73	CAGC	2132	137	GAGC	3122	201	TAGC	4131
10	AAGG	1133	74	CAGG	2133	138	GAGG	3123	202	TAGG	4132

11	AAGT	1134	75	CAGT	2134	139	GAGT	3124	203	TAGT	4133
12	AATA	1141	76	CATA	2141	140	GATA	3131	204	TATA	4144
13	AATC	1142	77	CATC	2142	141	GATC	3132	205	TATC	4141
14	AATG	1143	78	CATG	2143	142	GATG	3133	206	TATG	4142
15	AATT	1144	79	CATT	2144	143	GATT	3134	207	TATT	4143
16	ACAA	1211	80	CCAA	2211	144	GCAA	3211	208	TCAA	421
17	ACAC	1212	81	CCAC	2212	145	GCAC	3212	209	TCAC	4214
18	ACAG	1213	82	CCAG	2213	146	GCAG	3213	210	TCAG	4211
19	ACAT	1214	83	CCAT	2214	147	GCAT	3214	211	TCAT	4212
20	ACCA	1221	84	CCCA	2221	148	GCCA	3221	212	TCCA	4223
21	ACCC	1222	85	CCCC	2222	149	GCCC	3222	213	TCCC	4224
22	ACCG	1223	86	CCCG	2223	150	GCCG	3223	214	TCCG	422
23	ACCT	1224	87	CCCT	2224	151	GCCT	3224	215	TCCT	4221
24	ACGA	1231	88	CCGA	2231	152	GCGA	3231	216	TCGA	4232
25	ACGC	1232	89	CCGC	2232	153	GCGC	3232	217	TCGC	4233
26	ACGG	1233	90	CCGG	2233	154	GCGG	3233	218	TCGG	4234
27	ACGT	1234	91	CCGT	2234	155	GCGT	3234	219	TCGT	4231
28	ACTA	1241	92	CCTA	2241	156	GCTA	3241	220	TCTA	4242
29	ACTC	1242	93	CCTC	2242	157	GCTC	3242	221	TCTC	4243

30	ACTG	1243	94	CCTG	2243	158	GCTG	3243	222	TCTG	4244
31	ACTT	1244	95	CCTT	2244	159	GCTT	3244	223	TCTT	4241
32	AGAA	1311	96	CGAA	2311	160	GGAA	3311	224	TGAA	4312
33	AGAC	1312	97	CGAC	2312	161	GGAC	3312	225	TGAC	4313
34	AGAG	1313	98	CGAG	2313	162	GGAG	3313	226	TGAG	4314
35	AGAT	1314	99	CGAT	2314	163	GGAT	3314	227	TGAT	4311
36	AGCA	1321	100	CGCA	2321	164	GGCA	3321	228	TGCA	4322
37	AGCC	1322	101	CGCC	2322	165	GGCC	3322	229	TGCC	4323
38	AGCG	1323	102	CGCG	2323	166	GGCG	3323	230	TGCG	4324
39	AGCT	1324	103	CGCT	2324	167	GGCT	3324	231	TGCT	4321
40	AGGA	1331	104	CGGA	2331	168	GGGA	3331	232	TGGA	4332
41	AGGC	1332	105	CGGC	2332	169	GGGC	3332	233	TGGC	4333
42	AGGG	1333	106	CGGG	2333	170	GGGG	3333	234	TGGG	433
43	AGGT	1334	107	CGGT	2334	171	GGGT	3334	235	TGGT	4334
44	AGTA	1341	108	CGTA	2341	172	GGTA	3341	236	TGTA	4341
45	AGTC	1342	109	CGTC	2322	173	GGTC	3342	237	TGTC	4342
46	AGTG	1343	110	CGTG	2333	174	GGTG	3343	238	TGTG	4343
47	AGTT	1344	111	CGTT	2344	175	GGTT	3344	239	TGTT	4344
48	ATAA	1411	112	CTAA	2411	176	GTAA	3411	240	TTAA	4411

49	ATAC	1412	113	CTAC	2422	177	GTAC	3412	241	TTAC	4412
50	ATAG	1413	114	CTAG	2433	178	GTAG	3413	242	TTAG	4413
51	ATAT	1414	115	CTAT	2444	179	GTAT	3414	243	TTAT	4414
52	ATCA	1421	116	CTCA	2421	180	GTCA	3421	244	TTCA	4421
53	ATCC	1422	117	CTCC	2422	181	GTCC	3422	245	TTCC	4422
54	ATCG	1423	118	CTCG	2423	182	GTCG	3423	246	TTCG	4423
55	ATCT	1424	119	CTCT	2424	183	GTCT	3424	247	TTCT	4424
56	ATGA	1431	120	CTGA	2431	184	GTGA	3431	248	TTGA	4431
57	ATGC	1432	121	CTGC	2432	185	GTGC	3432	249	TTGC	4432
58	ATGG	1433	122	CTGG	2433	186	GTGG	3433	250	TTGG	4433
59	ATGT	1434	123	CTGT	2434	187	GTGT	3434	251	TTGT	4434
60	ATTA	1441	124	CTTA	2441	188	GTTA	3441	252	TTTA	4441
61	ATTC	1442	125	CTTC	2442	189	GTTC	3442	253	TTTC	4442
62	ATTG	1443	126	CTTG	2443	190	GTTG	3443	254	TTTG	4443
63	ATTT	1444	127	CTTT	2444	191	GTTT	3444	255	TTTT	4444

5.1.2 Base Point G

Usually the point with the smallest coordinates, but can be any randomly generated point. It is used to compute the value of kG at the sender, which is sent to the receiver to help compute the shared key. It is also required by the receiver to compute his public key from his private key. Thus, both communicating entities (sender and receiver) need to know the value of G.

5.1.3 Modulus Prime P

Randomly generated prime number used to define the finite prime field. The number generated should be large enough so that there is sufficient number of affine points, thus ensuring security. Since the prime field used by both the participating entities must be the same, this value should be known to both the sender and the receiver.

5.1.4 Private Key Of Sender K

It is a 128-bit prime number generated by the sender, used by sender for encryption. It is kept private by the sender.

5.1.5 Private Key Of Receiver Nb

It is a 128-bit prime number, generated by the receiver, used by receiver to compute the public key and for decryption. It is kept private by the receiver.

5.1.6.1 Public Key Of Receiver Pb

It is computed using nB as PB = nB * G. It is transmitted to the sender, who uses it for encryption.

The security of the proposed algorithm is further enhanced by obviating the need to transmit any private keys. However, we need certificates or trusted third parties for the sender to authenticate the public key of the recipient.

5.2 ALGORITHM

5.2.1 Setup

I. Take a Gray image and take its intensity value in a matrix form.

II. Map this intensity value to corresponding DNA nucleotide and then into a decimal number with the help of given Table

III. Consider an elliptic curve.

IV. Generate a prime p which is used to obtain the set of affine points and to calculate the modulus.

V. Compute the set of affine points for the elliptic curve.

VI. Assume the point in the set of affine points with the smallest coordinates to be the base point G.

5.2.2 Key Generation:

I. Randomly generate a prime number n_B at the receiver's end. This value is now the private key of the receiver and must be kept a secret.

II. Compute the public key of the receiver as $P_B = n_B*G$ and publish it.

III. Randomly generate another prime number k at the sender's end. This value is the private key of the sender and is used by the sender in the encryption process.

5.2.3 Mapping Table Construction:

I. The first column contains the values 1-256.

II. Fill in the affine points into the table column-wise. The mapping table simplifies the conversion of data to a DNA nucleotide and a decimal number.

5.2.4 Encryption

I. Take the intensity value matrix of gray image

II. Convert intensity value into a DNA nucleotide and corresponding number.

III. Convert this decimal number into a point of elliptic curve by the help of koblitz's method.

IV. Compute kG. This value is sent to the receiver and is used in the decryption process. Since the value is common for the entire data, it can be sent only once.

V. Compute kP_B.

VI. Take a random Elliptic Curve point R_p and add it to that point which is generated by koblitz's method. Result of this addition is known by P_m

VII. Compute $P_{m1} = P_m + kP_B$

VIII. Perform an mapping of P_{m1} in a graph. This is done as follows:

A. Take the coordinate of P_{m1} and store into 2 matrix as matrix X and matrix Y which is our encrypted matrices.

B. Graph it into a curve using matrix X and matrix Y

C. Steps 5.2.4.7 and 5.2.4.8 are repeated for the entire data file.

2 encrypted matrix are obtained which are sent to the receiver.

ENCRYPTION:

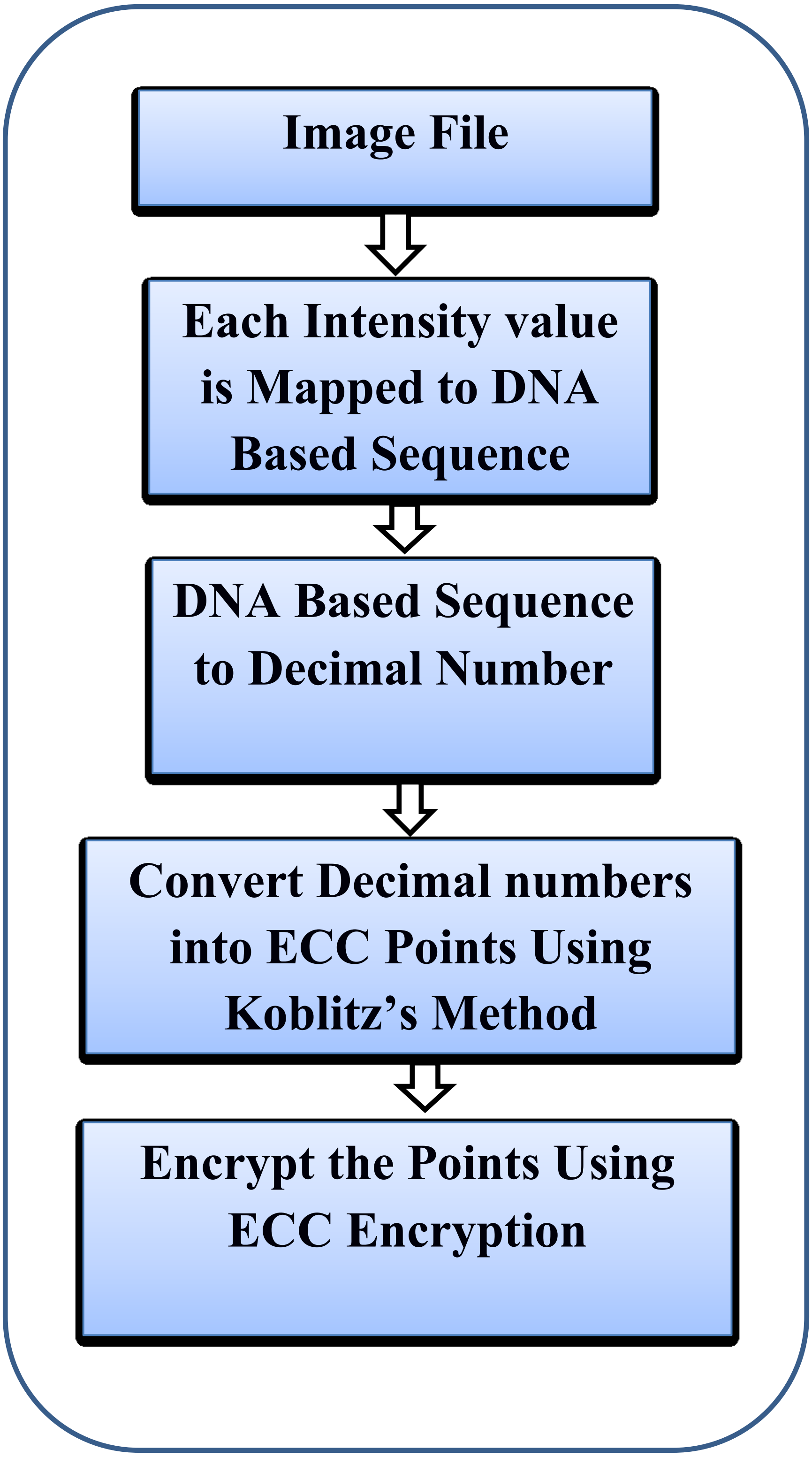


Image File
Each Intensity value is Mapped to DNA Based Sequence
DNA Based Sequence to Decimal Number
Convert Decimal numbers into ECC Points Using Koblitz's Method
Encrypt the Points Using ECC Encryption

Figure 5.1 Encryption Flow

5.2.5 Decryption:

Upon receiving kG and the 2 encrypted matrix. The following steps are for decrypying image

I. Multiply kG with private key of the receiver n_B to obtain kP_B. $kG*n_B = k*n_BG = kP_B$

II. Take the data from the 2 matrices into a point on the curve which is P_{ml}.

III. Compute $P_m = P_{ml} – kP_B$.

IV. Again compute original elliptic curve point by perform a operation P_m -R_p

V. Repeat step 5.2.5.2 to 5.2.5.4 for entire data.

VI. Take the point which is generated at step 5.2.5.4 and now convert it into a decimal number by the help of reverse koblitz's method

VII. Now mapped it into intensity of image and store it into a matrix for all points

This is our original image.

DECRYPTION:

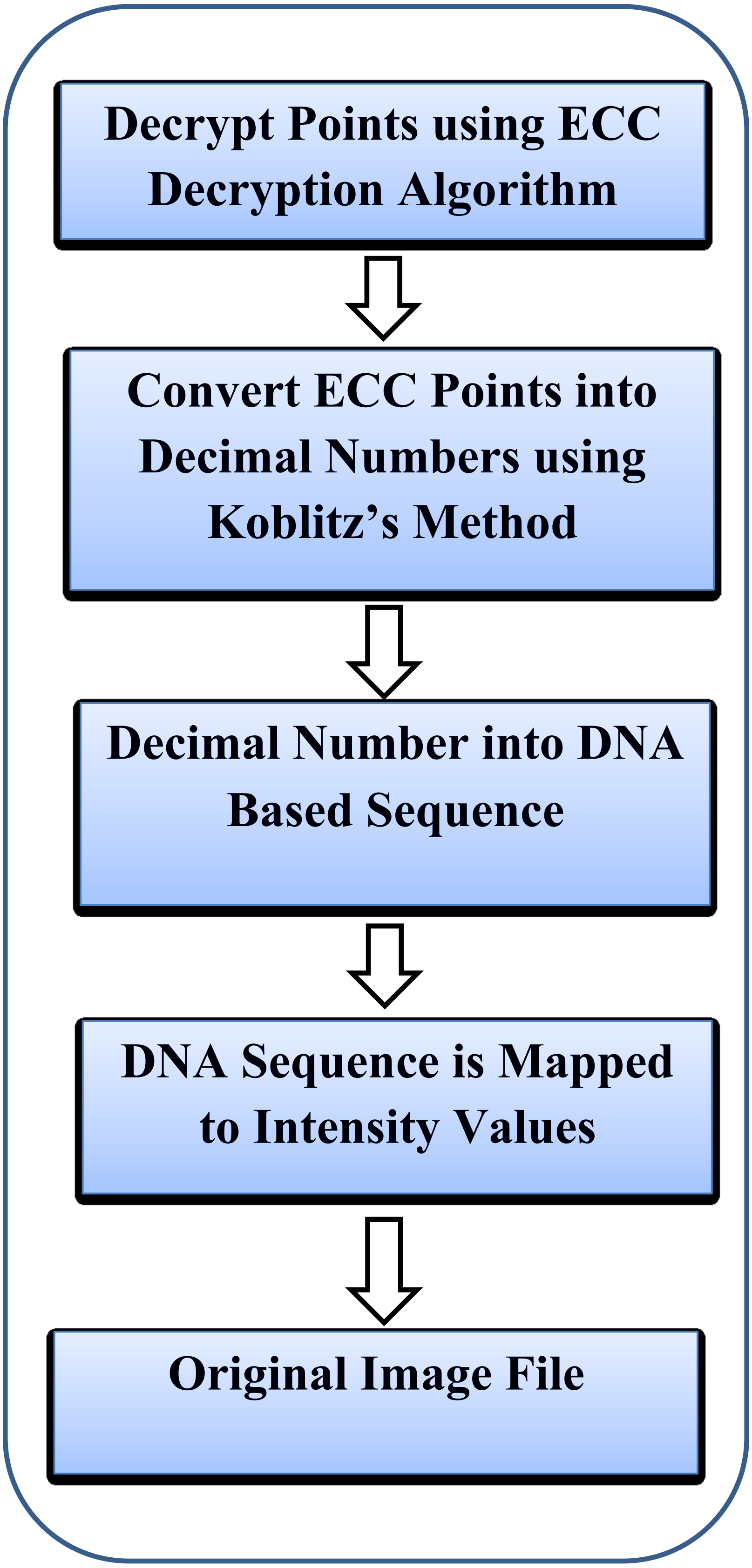

Decrypt Points using ECC Decryption Algorithm
Convert ECC Points into Decimal Numbers using Koblitz's Method
Decimal Number into DNA Based Sequence
DNA Sequence is Mapped to Intensity Values
Original Image File

Figure 5.2 Decryption Flow

5.3 TOY EXAMPLE FOR ALGORITHM

Lets take an elliptic curve equation

$$y^2 \bmod p = x^3+ax+b \bmod p$$

where assume a=1 and b =1 such that $4a^3+27b^2 \mathrel{!}= 0$

Encryption

assume $p = 15554017$, $N_B=2$, $k= 5$, $G= (0,1)$ and intensity value= 255

where p=modulus prime

N_B= private key of receiver

k = private key of sender

G = base point

Take the nucleotide decimal number from mapping table of intensity which is equal to 4444. Say m to this number and also take a random k. Let k is equal to 3000.

Now convert it into elliptic curve points by the help of koblitz's method that is (13333819,1520) which is our data point $D_{p.}$

Calculate public key of receiver

$$P_B= N_BG$$

$$= 2(0,1)=(11665513,1944251)$$

this is just doubling of point G.

The doubling formula is

$$x_3= m^2 - 2x_{1.}$$

and $y_3 = m * (x_1 - x_3) - y_1$

where $m = (3x^2 + a) / 2y$

Take a random point R_p= (72,611)

Calculate $P_m= R_p+ D_{p.}$

$P_m= (72,611) + (13333819,1520)$

$P_m= (6997299,4930323)$

The addition formula is

$x_3= m^2 – x_1 – x_2$

and $y_3=m*(x_1 – x_3) – y_1$

where $m = (y_2 – y_1) / (x_2 – x_1)$

Calculate $k\ P_B= 5(11665513,1944251)$

$k\ P_B= (4606449,2971607)$

This can be obatined by two times doubling and one time addition to self value means

$2(2P_B)+ P_B$

Now find $P_{m1}= P_m + k\ P_B$

$P_{m1}= (6997299,4930323) + (4606449,2971607)$

$P_{m1}= (3386389,10408953)$

Calculate kG= 5(0,1)

kG= (2983568,4190878)

Store P_{m1} into matrixes X and Y .

Decryption

Take kG point and multiply it by N_B which just equal to k G N_B and it is $kP_{B.}$

So $kGN_B =$ (4606449,2971607)

Subtract it by P_{m1} and find the value of P_m means

$P_m= P_{m1} - kGN_B$

$P_m= (3386389,10408953) - (4606449,2971607)$

P_m= (6997299,4930323)

Now to find the D_p ,we subtract R_p from P_m

D_p= P_m - R_p

D_p= (6997299,4930323) - (72,611)

D_p= (13333819,1520)

Now apply reverse koblitz's method to find the mapped point which was represented by m

so m= greatest integer of ((x-1)/k)

m= greatest integer of ((13333819-1)/3000)

m= greatest integer of (13333818/3000)

m= greatest integer of (4444.606)

m= 4444

now we can just mapped m to mapping table and then find the intensity value is 255.

Store it for all values into a matrix and this is our original image.

5.4 ALGORITHM DETAILS FOR IMAGE DATA

The algorithm supports the encryption and decryption of image data. The specific information relating to file formats and encryption and decryption methodology is summarized in this section.

Images are considered to be arrays of pixels. In the implementation, we use the RGB model for representing images.

Each intensity component is mapped to a decimal number using a DNA nucleotide table, using the decimal value and further convert this number into some points. The generated points help ensure that silhouettes of objects in the image are not visible after encryption and there are no patterns in the encrypted image either. This is particularly important in the case of images that the intensity value lies between a range of 0 to 255 but matrices of points have so higher value of this range so this value can not be mapped into a image. It will only possible and performing some operation.

Encryption of a particular intensity component of a pixel produces the encrypted value of the corresponding intensity component in the corresponding pixel of the 2 encrypted matrices.

This process is repeated for the three intensity components and for each pixel in the image leading to two encrypted matrices and a curve.

The input files can be either PNG or JPEG files.

The encrypted images are stored in PNG (Portable Network Graphics format). This is because a compression of the encrypted image results in loss of information, which results in distortion of the decrypted image.

The decrypted images are stored in the JPEG format to result in compressed images having smaller file size. However, this can be made user specified by small modifications in the program.

An overview of the PNG and JPEG file formats is provided in next section

RGB COLOR MODEL

According to the RGB model, each pixel is represented as a triplet of intensity values; the first value represents the red color component, the second represents the green component and the third value represents the blue component.

Each intensity value lies in the range 0 to 255 and is represented by 8 bits. They may also be represented in terms of their hex values, which range between 00 and FF.

Thus, each pixel is represented by 24 bits or as a series of three hex values put together.

For example, black is represented as (0,0,0) or 000000.

White is represented as (255,255,255) or FFFFFF.

All other colors have their components varying between the limits.

Red can be expressed as (255, 0, 0) or FF0000 since the blue and green components do not exist for red. Similar color codes may be obtained for all the other colors.

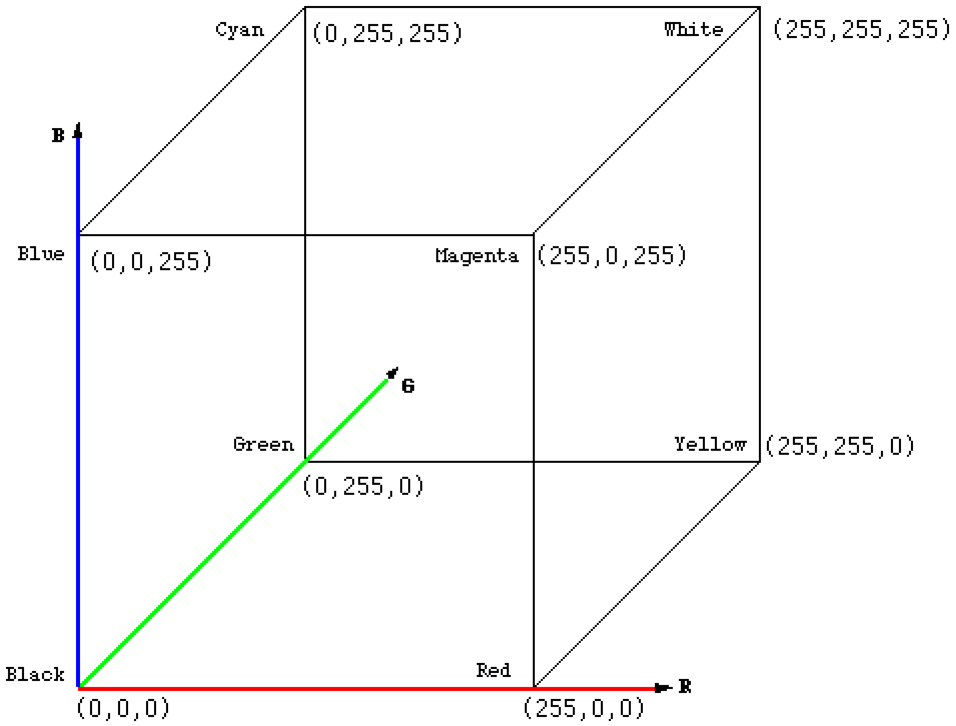


Figure 5.3 RGB Color Model

PNG (Portable Network Graphics)

PNG was developed as a replacement for GIF (Graphics Interchange Format). PNG is a raster graphics file format in which the image is stored as a dot matrix data structure. The matrix represents a rectangular grid representing pixels. The grid structure stores the color of each pixel. A raster graphics file is characterized by the number of pixels along the width and the height of the image. A bitmap, single bit raster, is a bit-for-bit representation of the image stored in the format in which storage device stores data or a device independent representation. PNG performs two-step lossless data compression — First step involves a pre-compression process called filtering or prediction followed by a compression algorithm called DEFLATES.

PNG file format contains the following components:

Header

The header of a PNG file contains the following information:

I. High bit is set to detect systems that do not support 8 bit data transmissions and to prevent the misinterpretation of the PNG file as a text file.

II. The letter PNG in ASCII (hex) to state the file format.

III. End of line characters. This is crucial since in the storage, the pixels are stored sequentially, while they represent a 2D array. Thus the End of line character is required to distinguish the rows.

IV. End of file character to detect the end of file.

Chunks

The chunks contain control information and are of two main types, namely critical and ancillary chunks. The critical chunks must be understood by all systems reading the file while the ancillary chunk information may be ignored if the system cannot understand it.

The chunk consists of four parts:

I. Chunk length

II. Chunk name/type

III. Chunk data

IV. Cyclic Redundancy Check (CRC) or checksum, computed over the chunk type and data.

The critical chunks are as follows:

IHDR

Contains the height, width, bit depth and color type of the image.

PLTE

In the case when the image data is stored in a palette and the numbers in channels represent the index into the palette, the palette is stored in the PLTE chunk.

IDAT

This chunk contains the actual image data which is the output stream of the compression algorithm. It may be split into several IDAT chunks. While the splitting increases file size, it enables generation of the PNG in a streaming manner.

TEND

End of image file.

Ancillary chunks may store such information as the histogram of the image, time the image was last changed, chromaticity, etc.

PNG file format consists of a header, followed by a series of chunks that contain control information, followed by the pixel data.

Each pixel is divided into channels. The number of channels depends on whether the image is a gray scale image, and whether it has the alpha channel that represents opacity.

Each channel refers to a number. All these numbers are encoded in the same format. Depending on the representation, the number could represent either an in an index into a palette, stored in the PLTE chunk or the data itself, encoded as between one to four numbers.

The bit depth represents the number of bits in the encoded unsigned integral value, while the color depth refers to the number of bits in each pixel.

JPEG (Joint Photographic Experts Group)

JEPG is a technique of lossy compression used for digital images. It allows for a tradeoff between quality and storage size by allowing adjustments to the degree of compression. A JPEG file consists of segments. Each segment begins with a marker signifying what the segment represents.

Chapter 6
RESULT AND ANALYSIS

6.1 RESULT

Elliptic curve cryptography is performed by converting the intensity value into a DNA sequence. A general Elliptic Curve is taken that is represented by the following equation:

Y^2 mod p=(X^3+aX+b) mod p

Where X,Y are elements of GF(p) and a,b are integers modulo p , satisfying :

$4a^3+27b^2$!= 0

Generate elliptic curve values for which a=1,b=1 and p=15554017

Then generate private and public key pair for user A and user B

Let private key for user A is k=5 , and

Private key of user B is N_b=2, and public keys of user A and B are a multiplication of private keys and base point of curve.

Base point G =(0,1)

And take another point that is affine point R_p =(72,611)

Now D_p is the point which is generated by koblit'z method after converting a intensity value into respective DNA sequence.

Now cipher text mean unreadable form of data C_f=(P_m + k P_B)

Where P_m =R_p+D_p and k P_B=kN_BG.

Encryption

assume p = 15554017,intensity value= 255

where p=modulus prime

Take the nucleotide decimal number from mapping table of intensity which is equal to 4444. Say m to this number and also take a random k. Let k is equal to 3000.

Now convert it into elliptic curve points by the help of koblitz's method that is (13333819,1520) which is our data point $D_{p.}$

Calculate public key of receiver

P_B= N_BG

= 2(0,1)=(11665513,1944251)

Take a random point R_p= (72,611)

Calculate P_m= R_p+ $D_{p.}$

P_m= (6997299,4930323)

Calculate k P_B= 5(11665513,1944251)

k P_B= (4606449,2971607)

Now find P_{m1}= P_m + k P_B

P_{m1}= (3386389,10408953)

Calculate kG= 5(0,1)

kG= (2983568,4190878)

Store P_{m1} into matrices X and Y .

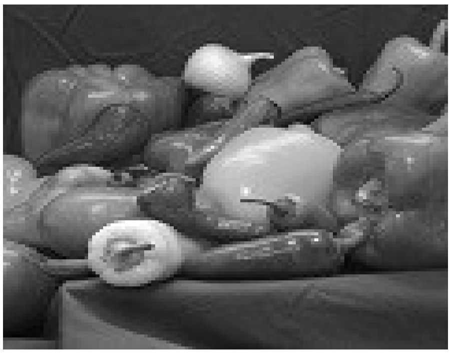
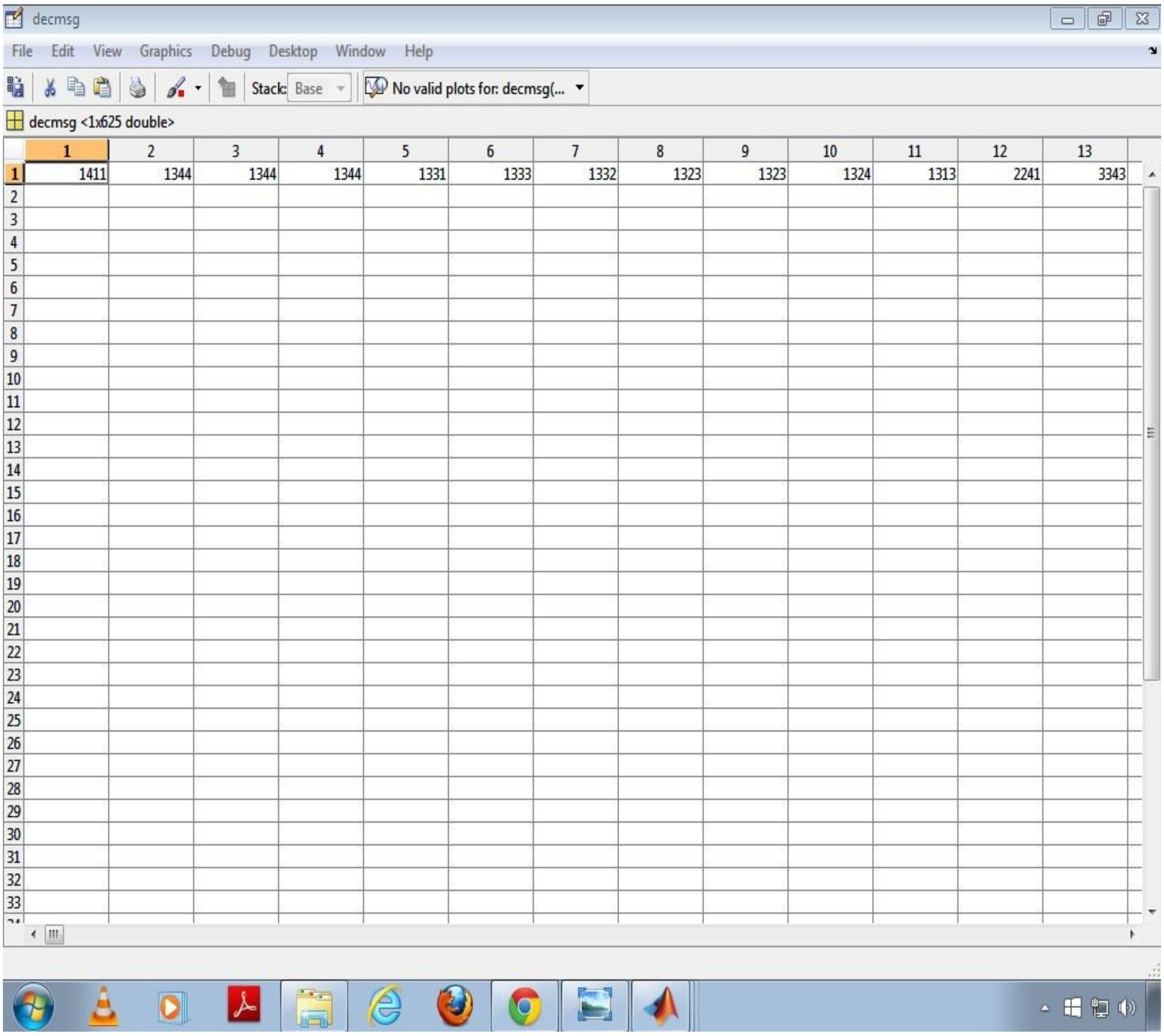


Figure 6.1 Original Image

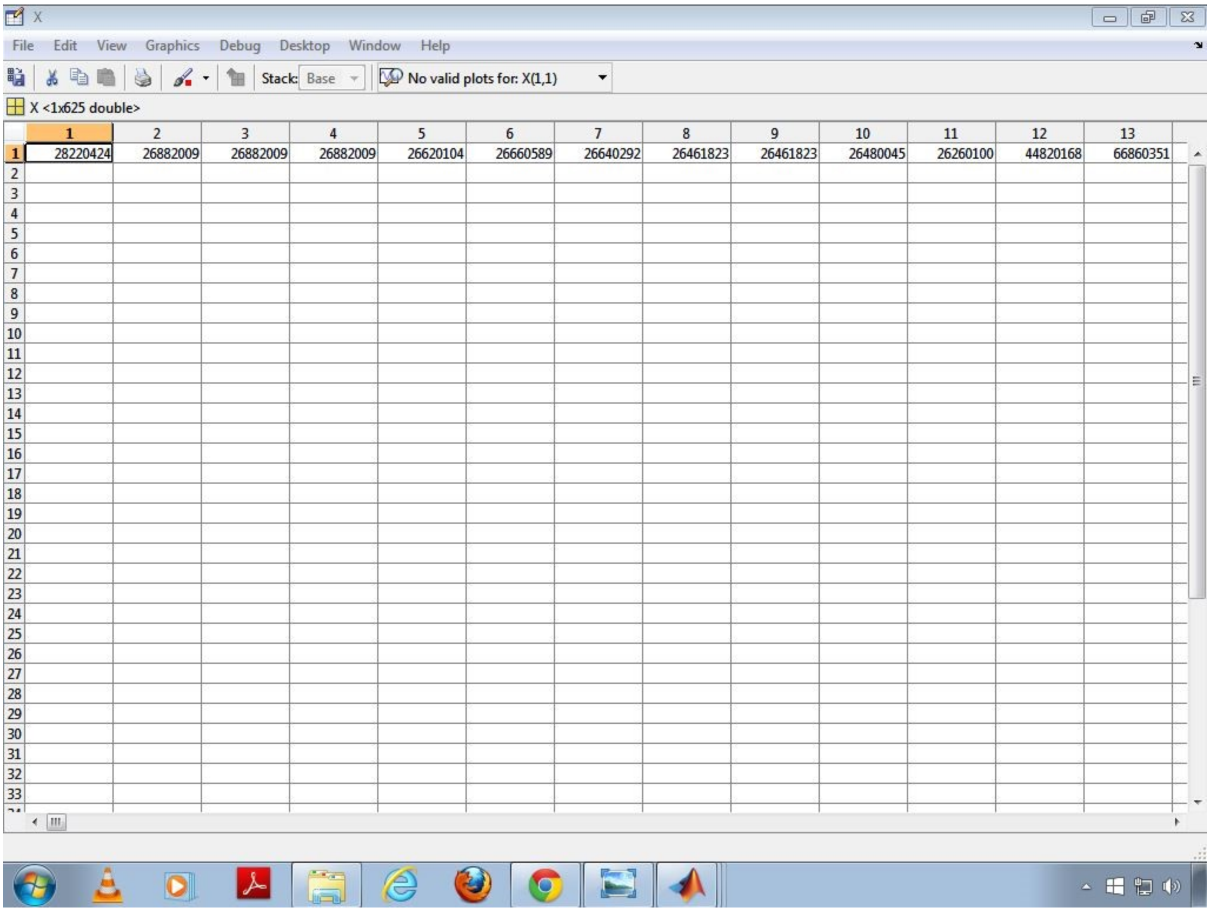

decmsg
File Edit View Graphics Debug Desktop Window Help
Stack: Base
No valid plots for: decmsg(...

decmsg <1x625 double>

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1411	1344	1344	1344	1331	1333	1332	1323	1323	1324	1313	2241	3343
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													
32													
33													

Figure 6.2 Mapping of intensity value of image to DNA sequence value

In this Figure 6.2, intensity value of image is mapped into DNA sequence value which is already define in Table 5.1.



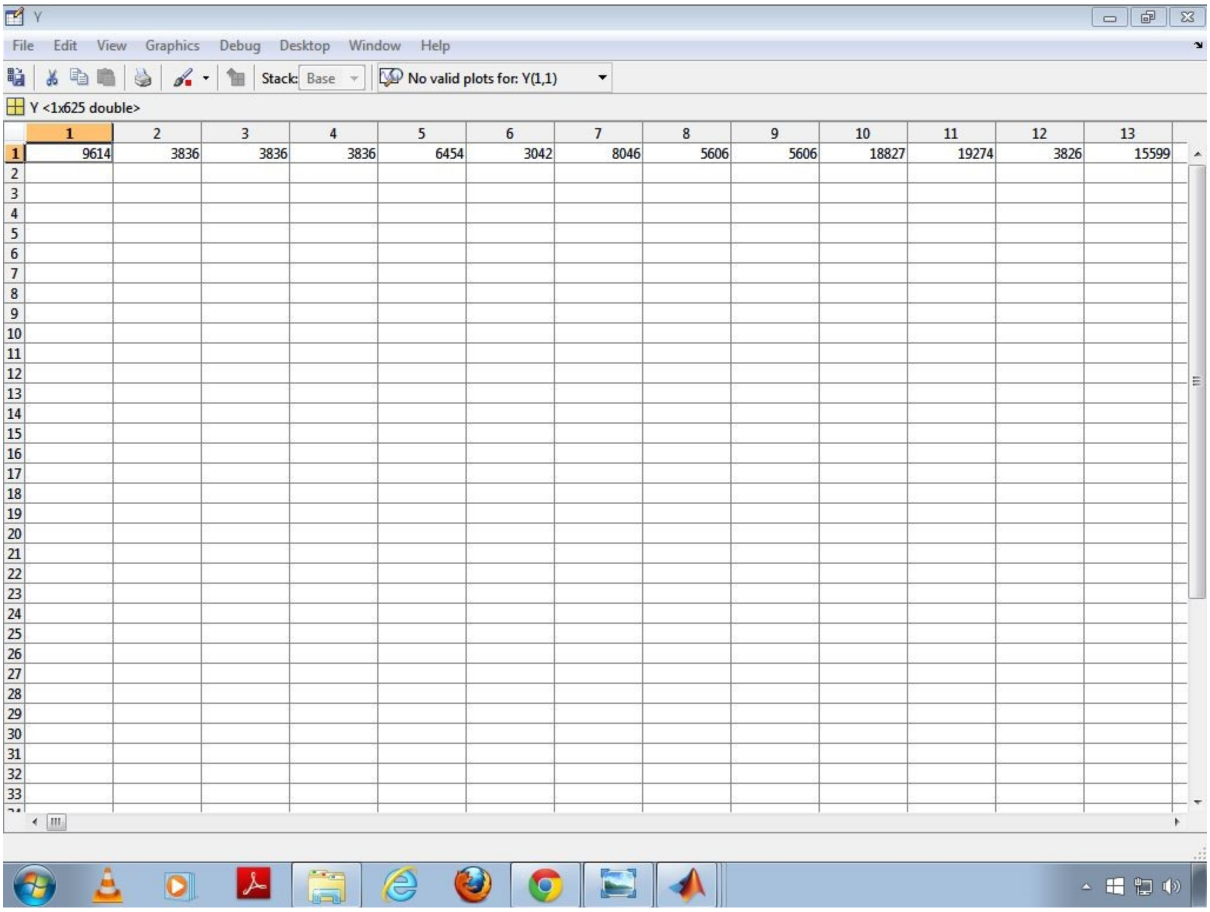
No valid plots for: X(1,1)

X <1x625 double>

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	28220424	26882009	26882009	26882009	26620104	26660589	26640292	26461823	26461823	26480045	26260100	44820168	66860351
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													
32													
33													

Figure 6.3 X co-ordinates matrix of Cipher Text

In figure 6.3, the values of X co-ordinates of cipher text are stored which are also x co-ordinate of elliptic curve points of " Y^2 mod 15554017 = (X^3+X+1) mod 15554017"


Y
File Edit View Graphics Debug Desktop Window Help
Stack: Base
No valid plots for: Y(1,1)
Y <1x625 double>

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	9614	3836	3836	3836	6454	3042	8046	5606	5606	18827	19274	3826	15599

Figure 6.4 Y co-ordinates matrix of Cipher Text

In figure 6.4, the values of Y co-ordinates of cipher text are stored which are also x co-ordinate of elliptic curve points of " Y^2 mod 15554017 = (X^3+X+1) mod 15554017"

Now the cipher text is both the matrices X matrix and Y matrix.

Decryption

Take kG point and multiply it by N_B which just equal to k G N_B and it is $kP_{B.}$

So $kGN_B =$ (4606449,2971607)

Subtract it by P_{m1} and find the value of P_m means

$P_m= P_{m1} - kGN_B$

$P_m=$ (3386389,10408953) - (4606449,2971607)

$P_m=$ (6997299,4930323)

Now to find the $D_{p\,,}$we subtract R_p from P_m

$D_p= P_m - R_p$

$D_p=$ (13333819,1520)

Now apply reverse koblitz's method to find the mapped point which was represented by m

so m= greatest integer of ((x-1)/k)

m= greatest integer of ((13333819-1)/3000)

m= 4444

now we can just mapped m to mapping table and then find the intensity value is 255.

Store it for all values into a matrix and this is our original image.

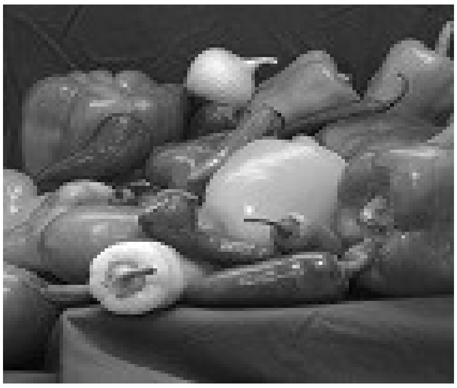


Figure 6.5 Decrypted Image

6.2 ANALYSIS

Draw the histogram of both the images which are original image and decrypted image.

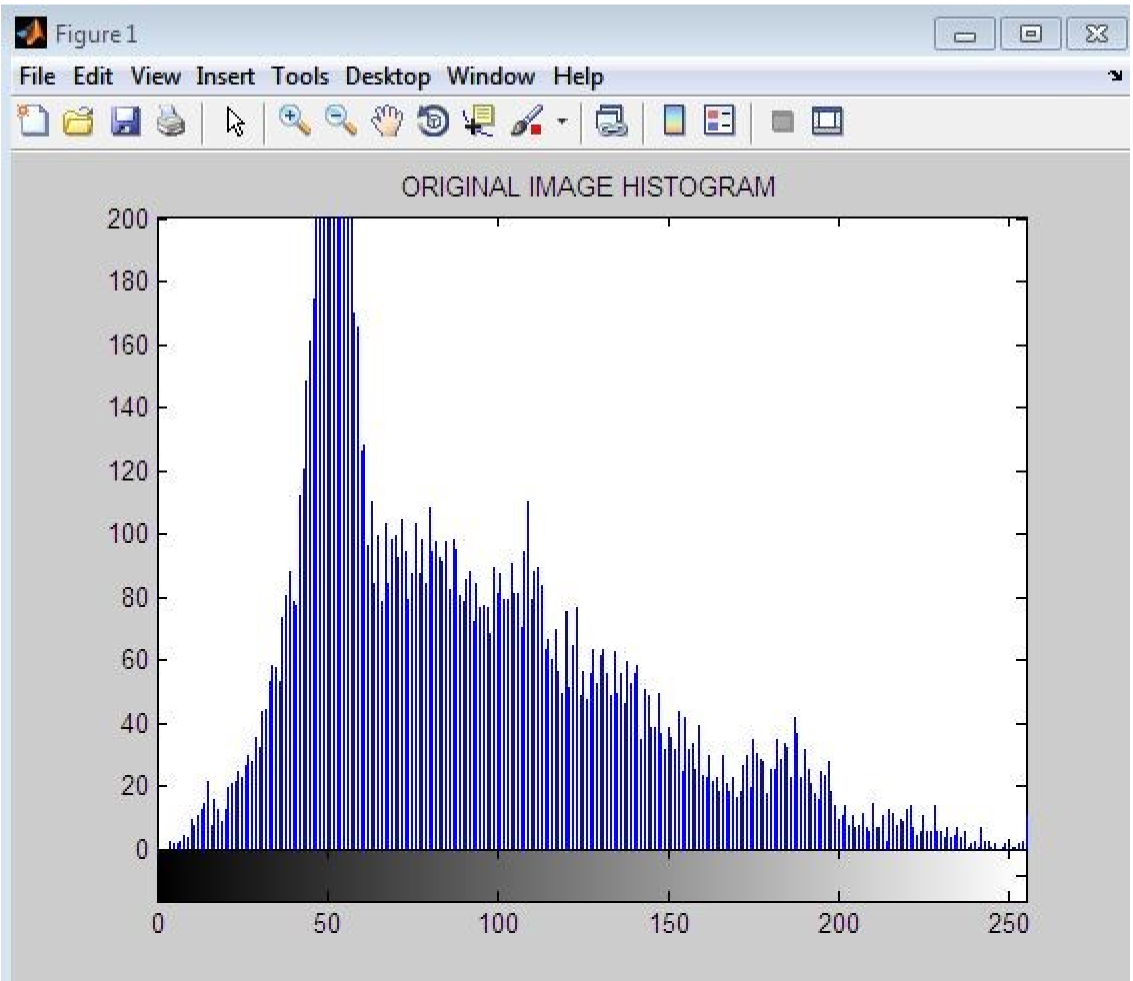


Figure 6.6 Histogram of original image

Figure 6.6 shows the histogram of original image which is 6.1.

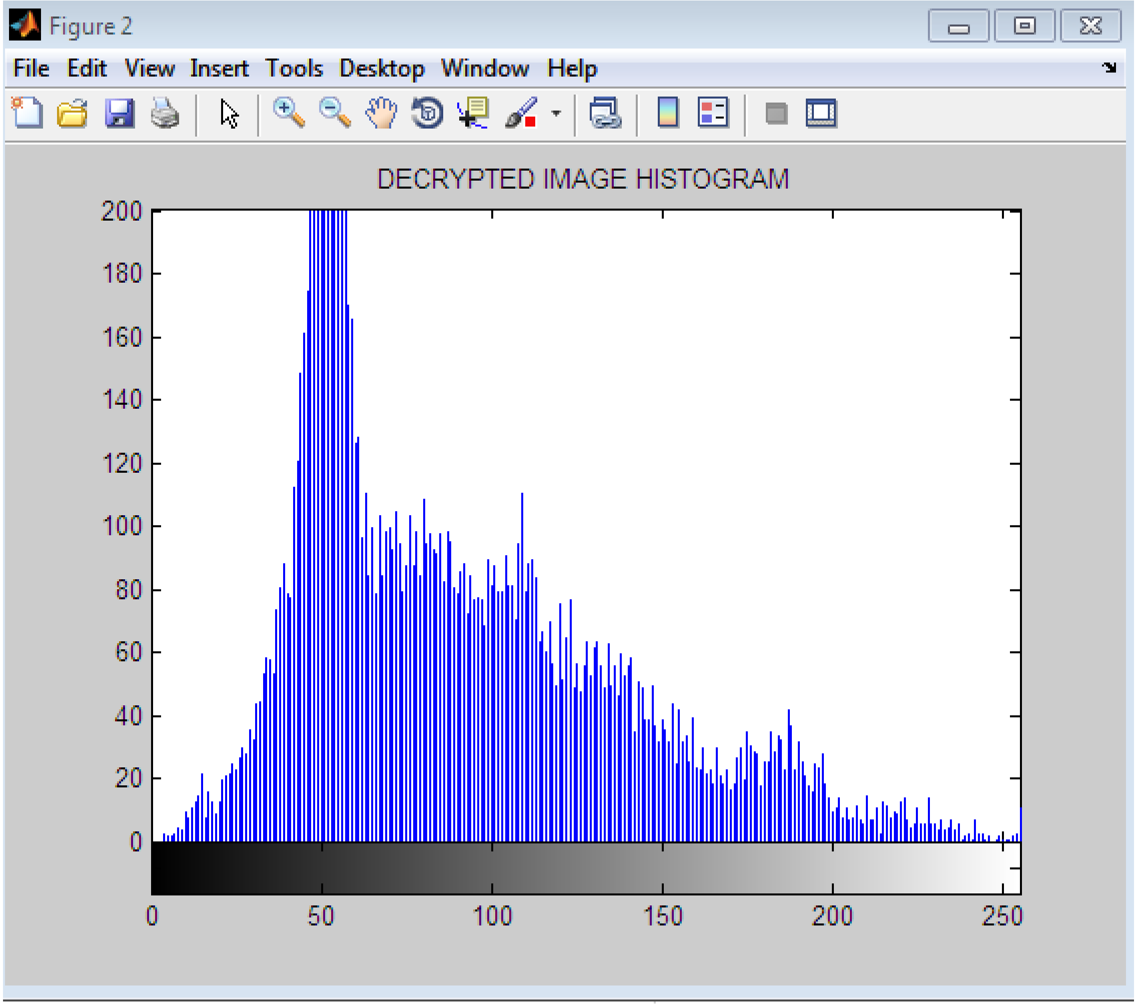


Figure 2
File Edit View Insert Tools Desktop Window Help
DECRYPTED IMAGE HISTOGRAM
200
180
160
140
120
100
80
60
40
20
0
0
50
100
150
200
250

Figure 6.7 Histogram of Decrypted Image

Figure 6.7 shows the histogram of Decrypted image which is 6.5.

As shown in figure 6.5 and figure 6.6, the both images are same which are original image and decrypted image.

For the color images this approach also works ,Firstly to convert it into R,G,B component and then apply this proposed approach on these component and after decryption user will get the R,G,B component then join it .This will same as original image

Chapter 7

CONCLUSION AND FUTURE WORK

7.1 CONCLUSION

Cryptography deals with securely transferring information from one end to another end over insecure channels. It ensures that input data or the sensitive information being transferred will only be interpreted by the one for whom it is targeted and will not be interpreted by anyone else present in the network. In this era of internet where highly sensitive information is transmitted over it, cryptographic techniques must ensure high data security.

RSA is one of the public key cryptographic technique used for this purpose .RSA ensures security but key sizes used are very large. This increases complexity of the algorithm. This shortcoming of RSA is removed by Elliptic curve cryptography which uses keys of smaller sizes smaller key sizes leads to less computational overhead. Thus overhead of managing keys is reduced to certain extent.

In this thesis we have applied ECC on image data. Intensity value of each pixel is mapped to a random DNA sequence which provided extra level of security in our scheme. In ECC, input data is mapped to points on elliptic curve. Different encoding techniques are used for this purpose. One of the best suited techniques is Koblitz's method. Our scheme has also used Koblitz's method for converting intensity of each pixel of the original image into points of elliptic curve. This is the pre-processing step of ECC. On the other hand, during decryption every point on elliptic curve is converted or deciphered into intensity value of original image. This is post processing step of ECC. Pre and post processing steps are performed using Koblitz's method.

After mapping points on elliptic curve, each obtained point is encrypted using ECC encryption and in decryption phase every encrypted point is brought back to its original value. These coordinates of elliptic curve are decoded back to intensity values by using decryption algorithm of koblitz's method.

Thus using ECC along with DNA sequence ensures less computational overhead and more security. Therefore we can say that our scheme is better than existing traditional approaches because it is providing more security and decreasing computational overhead unlike other approaches which are less secure and requires more computational power.

Future work includes modification in encoding and decoding technique used. More computational resources are used in encoding and decoding phase. If this technique can be replaced by another one with uses less resources overall computational overhead of the scheme can be reduced to certain extent.

7.2 FUTURE WORK

Future work includes extending the implementation to support other formats of audio files as well as to work with other data types such as video. Continuing research is being carried out to incorporate improvements to improve the encryption and decryption speeds.

It has proven that ECC is a successful Public key generation and encryption technique. In this age when people is becoming more dependent on internet for communication and transferring files over public network such as internet then a such technology is necessary which is secure and unbreakable from network attacks i.e. ECC. In future this is a very wide area for research because many secrets of ECC can be disclosed that are unrevealed still and that day is not so far on which ECC replaces RSA and other public key encryption techniques.

REFERENCES

[1] N.S. Raghava, Ashish Kumar "Image Encryption Using Henon Chaotic Map With Byte Sequence", International Journal of Computer Science Engineeringand Information Technology Research (IJCSEITR)ISSN(P): 2249-6831; ISSN(E): 2249-7943Vol. 3, Issue 5, Dec 2013

[2] Ray C. C. Cheng, Nicolas Jean-baptiste, Wayne Luk, and Peter Y. K Cheung,"Customizable Elliptic Curve Cryptosystems" , IEEE Trans. On VLSI Systems, vol. 13, no. 9, pp. 1048-1059, Sep. 2005.

[3] Alessandro Cilardo, Luigi Coppolino, Nicola Mazzocca, and Luigi Romano,"Elliptic Curve Cryptography Engineering", Proceedings of the IEEE, Vol. 94, no. 2, pp. 395 - 406, Feb. 2006.

[4] N.Koblitz, Elliptic Curve Cryptosystems, Mathematics of Computation, volA8, 1987, pp.203 -209.

[5] Saenger, Wolfram (1984). Principles of Nucleic Acid Structure. New York: SpringerVerlag .

[6] Watson JD, Crick FH (1953). "A Structure for Deoxyribose Nucleic Acid" (PDF). Nature 171 (4356): 737–738.

[8] Clausen-Schaumann H, Rief M, Tolksdorf C, Gaub HE (2000). "Mechanical stability of single DNA molecules". Biophys J 78 (4): 1997–2007.

[9] Crick, Francis (1988). "Chapter 8: The genetic code". What mad pursuit: a personal view of scientific discovery. New York: Basic Books. pp. 89–101.

[10] Kang Ning(2009),"A Psuedo DNA Cryptography Method", Cornell University Library.

[11] Debnath Bhattacharyya, Samir Kumar Bandyopadhyay," Hiding Secret Data in DNA Sequences", International Journal of Scientific & Engineering Research Volume 4(2013).

[12] J. Solinas, "Efficient Arithmetic on Koblitz Curves," Designs, Codes and Cryptography, vol. 19, pp. 195-249, 2000.

[13] N. Koblitz, "CM-Curves with Good Crytographic Properties," Proc.CRYPTO 91, pp. 279-287, Springer, 1992.

[14] V. Dimitrov, K. Ja ¨rvinen, M. Jacobson, W. Chan, and Z. Huang, "Provably Sublinear Point Multiplication on Koblitz Curves and Its Hardware

Implementation," IEEE Trans. Computers, vol. 57, no. 11, pp. 1469-1481, Nov. 2008.

[15] V. Dimitrov, L. Imbert, and P. Mishra, "Efficient and Secure Elliptic Curve Point Multiplication Using Double-Based Chains," Proc. ASIACRYPT 2005, pp. 59-78, Springer, 2005.

[16] D. Hankerson, A. Menezes, and S. Vanstone, Guide to Elliptic Curve Cryptography. Springer, 2004.

[17] I. Blake, G. Seroussi, and N. Smart, Elliptic Curves in Cryptography. Cambridge Univ. Press, 1999.

[18] Kevin M. Finnigin, Barry E. Mullins, Richard A. Raines, Henry B.Potoczny, "Cryptanalysis of an elliptic curve cryptosystem for wireless sensor networks," International journal of security and networks*,* Vol. 2, No.3/4, pp. 260- 271,2006.

[19] Standard specifications for public key cryptography, IEEE standard*,* pI363,2000.

[20] Williams Stallings, Cryptography and Network Security, Prentice Hall*,* 4 th Edition, 2006.

[21] R.V.Kurja, Kirti Joshi, N.Mohan Kumar, Kapil H Raranape,A.Ramanathan, T.N.Shorey, R.R.Simha, and V.Srinivas, Elliptic Curves, *International Distribution by American Mathematical Society,* 2006.

[22] Sangook Moon, "A Binary Redundant Scalar Point Multiplication In Secure Elliptic Curve Cryptosystems," International journal of network security, Vol.3, No.2, PP.132-137, Sept. 2006.

[23] Brijesh Kumar Patel, N.S.Raghava, "A Novel Approach for Multimedia Encryption Based on Confusion and Chaotic Logistic Map" International Journal of computer science Engineering and Information Technology Research (IJCSEITR) ISSN(P): 2249-6831;ISSN(E): 2249-7934 Vol. 4, Issue 4,Aug 2014, 47-58

[24] Miles E. Smld And Dennis K. Branstad, "The Data Encryption Standard: Past and Future, " PROCEEDINGS OF THE IEEE, VOL. 76, NO. 5, MAY 1988

[25] István Zsolt BERTA∗ and Zoltán Ádám MANN ."IMPLEMENTING ELLIPTIC CURVE CRYPTOGRAPHY ON PC AND SMART CARD" .PERIODICA POLYTECHNICA SER. EL. ENG. VOL. 46, NO. 1–2, PP. 47–73 (2002)

[26] Whitfield Diffie, Martin E Hellaman, "New Directions In Cryptography", IEEE Transactions On Information Theory,Vol.It-22,No.6,November 1976.

[27] H. J. Shiu, K. L. Ng, J. F. Fang, R. C. T. Lee and C. H. Huang, "Data hiding methods based upon DNA sequences", Information of Science, vol.180, no.11, pp.2196-2208, 2010.

[28] Qiang Zhang *, Ling Guo, Xianglian Xue, Xiaopeng Wei,"An Image Encryption Algorithm Based on DNA Sequence Addition Operation",Key Laboratory of Advanced Design and Intelligent Computing ,2011

[29] Jin-Shiuh Taur1, Heng-Yi Lin1, Hsin-Lun Lee1 and Chin-Wang Tao, "Data Hiding in DNA Sequences Based On Table Lookup Substitution", International Journal of Innovative Computing, Information and Control, Volume 8, Number 10(A), October 2012