CHAPTER 1

INTRODUCTION

Research analysis of information security attracts many researchers, because of its increasing practical importance in the growing field of electronic communications, e-business and other fields. Today, vast amount of sensitive information such as health and legal records, financial transactions, credit ratings and so on are routinely exchanged between computers via world wide public communication facilities and these valuable messages must be kept confidential and protected against manipulation. Cryptography ensures the strongest and the main tool for maintaining privacy, trust, access control in e-payments, e-voting, corporate security and many other applications. Cryptography is the science of applying strong Mathematics theory to increase the security of electronic transactions not only in military and diplomatic communications, but also to increase the level of privacy of individuals and groups.

From the earliest times, the art of cryptography has started becoming more and more popular and has advanced to a very high state of perfection. The history of security communications may be traced back to the use of tricks, spies in the disguise of ordinary people, carrier pigeons and other homing birds. Other techniques like redundancy in the number of messengers, hiding messages through mathematical and linguistic riddles, steganography and secret signalling mirrors between hilltop forts are also used to transfer secret messages in the early days. Cryptographic codes have become so complicated that, even during the times of the Second World War,
it took the enemies an average of two or three years in order to break each single code that was invented. But, the only relief that can be taken is that there were no modern computers at that time, and hence, most of the work had to be done using logic and manual mental labour.

The word cryptography is derived from the Greek words Kryptós, “hidden”, and gráphien, “to write”. It is the study of the principles and techniques through which information can be concealed and later revealed by legitimate users by means of the secret key, but in which it is either impossible or computationally infeasible for an unauthorized person to do so.

Research activities in cryptography have been conducted from the early 1970s. The National Bureau of Standards (NBS) project in computer security, identified a number of areas requiring research and development of standards. The International Business Machines Corporation (IBM) initiated a research program in cryptography because of the perceived need to protect electronic information during transmission between terminals and computers and between the computers, especially where the transmissions were to authorize the transfer or dispensing money. This development resulted in several publications, patents, cryptographic algorithms including the popular symmetric encryption Data Encryption Standard (DES).

Cryptography deals with the transformation of ordinary text (known as plaintext) into disguised form called cipher text by encryption and transformation of cipher text into original plaintext by decryption. Figure 1.1 shows the process of transforming intelligible message in to unintelligible form through encryption and decryption.
These transformations are parameterised by one or more keys and the motivation for encryption is to provide security in the transmissions over insecure channels. Encryption and decryption are easy when the key is known, but decryption should be virtually impossible without the use of correct key. Encryption protects communications and stored information from unauthorized access and disclosure. Other cryptographic techniques designed for authentication and digital signatures protect against spoofing and message forgeries.

Constructing and breaking cryptographic techniques are considered as intellectual challenge. Cryptanalysis is the process of attempting to find a shortcut method, not envisioned by the designer, for decrypting the cipher text when the key is unknown.

1.1 SECURITY ASPECTS OF CRYPTOGRAPHY

Cryptography is the study of Mathematical techniques related to the aspects of information security, confidentiality, authentication, data integrity and non repudiation.

1.1.1 Confidentiality

Confidentiality or privacy of information ensures that, only those with sufficient privileges and a demonstrated need may access certain information. It refers to the denial of access to information by unauthorized individuals. Confidentiality is breached, when unauthorized individuals or
system can view information. Thus, privacy of the system prevents the extraction of information by unauthorized parties from messages transmitted over a public and insecure channel, assuring the sender of the message that, it will be read, only by the intended receiver.

The cryptographic tools ensure confidentiality by concealing the private information through embedded encryption techniques. Encryption is accomplished through cryptographic algorithms to disguise the plaintext into the cipher text, for transmission. These cryptographic algorithms were based on the branches of Mathematics like Number theory, Statistics and so on. Confidentiality requires that, an intruder should not be able to determine the plain text corresponding to a given ciphertext and the secret keys involved. Figure 1.2 shows, the compromise of confidentiality in a communication between the legitimate users A and B.

![Figure 1.2 Loss of confidentiality](image)

1.1.2 Authentication

Authenticity refers to validating the source of a message. It is a service, addressing the identification of the persons involved in the
communication. One of the requirements of a cryptosystem is to provide a service for the recipient of a message to ascertain its origin and an intruder should not be able to masquerade as a legitimate user of a system. The absence of authentication in a communication between legitimate users A and B is shown in Figure 1.3. The actual communication should be between A and B, but an intruder C impersonates as user A and communicates with B. But the user B is not aware of the fact that, he/she is communicating with an intruder and not with the user A.

![Figure 1.3 Absence of authentication](image)

### 1.1.3 Data Integrity

Integrity refers to the assurance that a message was not modified accidentally or deliberately in transit, by substitution, insertion or deletion. Data may lose its integrity by intentionally through viruses, worms, etc and accidentally through faulty programs, noise in the transmission channel or media and so on. To eradicate these external and internal threats to the integrity of information, systems should employ a variety of error control techniques including signature algorithms involving hash values, error-correcting techniques. Typically, integrity is provided by sending a
compressed form of the message known as message digest along with the full message as a check.

In a communication between the legitimate users A and B, an intruder C manages to access the message from A, changes its contents and send the changed message to B. Users A and B are unaware of such an attack called modification and it is shown in Figure 1.4.

![Figure 1.4 Loss of integrity](image)

**1.1.4 Non-repudiation**

Non-repudiation is a service, which prevents an entity from denying previous commitments or actions. A sender should not be able to deny later that he/she has sent a message. It should be possible for the recipient of a message to detect, whether it is a replay of previous transmission. When disputes arise due to an entity denying that certain actions were taken, a means to resolve the situations is necessary. A procedure involving a trusted third party is needed to resolve the dispute.
1.2 CRYPTOSYSTEM

A cryptosystem is a five-tuple Mathematical system \((M, C, K, E, D)\), where

- Message space \(M\) : a finite set of possible plain-texts
- Cipher text space \(C\) : a finite set of possible cipher-texts
- Key space \(K\) : a finite set of possible keys
- For each \(e \in K\), there is an encryption process \(E_e \in E\) and a corresponding decryption process \(D_d \in D\). Each \(E_e : M \rightarrow C\) and \(D_d : C \rightarrow M\) are bijective functions such that \(D_d(E_e(m)) = m\), for every plain text \(m \in M\).

1.2.1 Symmetric Key Cryptosystem

In a symmetric key cryptosystem, the same secret key \((e=d)\) is used for both encryption and decryption processes. It means that, some one who has enough information to encrypt messages automatically has enough information to decrypt messages as well. There are many popular symmetric key algorithms available in the literature; to name a few, Caesar ciphers, Data Encryption Standard (DES), Advanced Encryption Standard (AES) and Triple DES. Detailed studies of these symmetric key algorithms are given in the books authored by Simmons (1992), Bruce Schneier (1996), Alfred Menezes et al (2001), Stinson (2000), Hans Delfs and Helmut Knebl (2006) and Mollin (2006).

Symmetric key ciphers are designed to have high rates of data throughput and the computations are faster with relatively shorter keys. In end-end communication, the key must be kept secret by both the parties and
hence, the key to be frequently changed. If a key is compromised, then an attacker can decrypt all message traffic encrypted with that key and he/she can pretend to be one of the parties and produce false messages to forge the other party. In a large network with $n$ users, assuming a separate key is used for each pair of users, the total number of keys required is $\frac{n(n-1)}{2}$.

For example, 1000 users need 499500 keys to communicate with each other. Thus, key generation and management is a difficult task in symmetric key cryptosystem. Figure 1.5 depicts the secured communication between the hosts A and B by using the same key.

![Figure 1.5 Symmetric key cryptosystem](image)

1.2.2 Public Key Cryptosystem

In a public key cryptosystem or asymmetric algorithm, the key used for encryption ($e$) is different from the key used for decryption ($d$). Each person in the messaging cycle will have a private and a public key. Any message that is encrypted using the public key can only be decrypted using the private key, and vice versa. The algorithms are called “public key”, because every person knows everybody’s public key in the cycle. The private
key is kept secret. The encryption and decryption keys are constructed in such a way that, one key cannot be derived from the other with in a reasonable amount of time. It must be computationally infeasible to find the secret key from public key. The idea of a public key cryptosystem was due to the work of Diffie and Hellman (1976). Rivest, Shamir and Adleman (1978) proposed the first practical public key cryptosystem, now widely known as RSA cryptosystem.


In an asymmetric key cryptosystem, only the private key is to be kept confidential and the keys are used as long run keys. Public key cryptosystems are based on one way trap door functions, in which, it is easy to compute one direction (encryption) and hard to compute the other direction (decryption). However, with the secret key (trap door), decryption can be done easily. In a network with n-users, the number of keys required for the users to communicate with each other is n and thus, the asymmetric cryptography is scalable to a large population and does not have difficult key distribution problems. Throughput rates for the most popular encryption algorithms are numerous orders of magnitude and relatively slower than the symmetric key schemes. Asymmetric encryption expands the cipher text. Computational performance of public key encryption is inferior to that of
symmetric key encryption. Key sizes are typically much larger than those required for symmetric key techniques. The security of most effective public key algorithms is based on the apparent difficulty of a set of number theoretic computational problems. Figure 1.6 explains the message transmission from host A to host B, using public key cryptosystem.

![Public key cryptosystem diagram]

**Figure 1.6 Public key cryptosystem**

### 1.2.3 Hybrid Cryptosystem

A hybrid cryptosystem is a combination of symmetric and public key cryptosystem. At the commencement of the communication, a public key cryptosystem is used for the distribution of keys and in later stages, the symmetric key algorithm is used for the bulk encryption of the data. Hybrid cryptosystem is used in Pretty Good Privacy (PGP).

### 1.3 COMPUTATIONAL HARD PROBLEMS

The security of many public key cryptosystems depends on the apparent intractability of the computational problems. A computational problem is said to be easy or tractable if it can be solved in polynomial time, that is, in expected time, at least for a non-negligible fraction of all possible

1.3.1 Integer Factorization Problem

Given a large positive integer $n$, find its prime factorization: that is, writing $n = p_1^{e_1} p_2^{e_2} \ldots p_k^{e_k}$, where $p_i$’s are pair wise distinct primes and $e_i \geq 1$ is hard.

1.3.2 Discrete Logarithm Problem

Given a prime $p$, a generator $\alpha$ of $\mathbb{Z}_p^*$ and an element $b \in \mathbb{Z}_p^*$, find the integer $a$, $0 < a \leq p-2$, such that $\alpha^a \equiv b \pmod{p}$ is computationally infeasible.

1.3.3 Diffie-Hellman Problem

Given a prime $p$, a generator $\alpha$ of $\mathbb{Z}_p^*$, elements $\alpha^a \pmod{p}$ and $\alpha^b \pmod{p}$, computing $\alpha^{ab} \pmod{p}$ with in reasonable amount of time is infeasible.

1.3.4 Elliptic Curve Discrete Logarithm Problem

Given an elliptic curve $E(F_p)$ over a prime field $F_p$ and a point $P$ on elliptic curve, it is relatively easy to calculate $Q = xP$ if $x$ and $P$ are known, but it is very hard to determine $x$ from known $Q$ and $P$. 
1.3.5 Polynomial Congruences Problem

Given a prime number $p$, and a set of $m$ integer polynomials $f_1, f_2, \ldots, f_m$ in $x_1, x_2, \ldots, x_n$ of degree bounded by $d$, solving the following system of congruences is hard

$$f_1(x_1, x_2, \ldots, x_n) \equiv 0 \mod p$$
$$f_2(x_1, x_2, \ldots, x_n) \equiv 0 \mod p$$
$$\ldots$$
$$\ldots$$
$$f_m(x_1, x_2, \ldots, x_n) \equiv 0 \mod p$$

1.4 PRELIMINARIES

In this section, the mathematical preliminaries used in the dissertation are discussed.

1.4.1 Number Theoretic Transforms (NTT)

Let $\{h_n\}$ be a sequence of $m$ integers (m-point integer sequence) and $\beta$ and $m$ are integers mutually prime with a composite number $N$. Then the Number Theoretic Transforms (NTT) of $\{h_n\}$ is defined as

$$H_k = \sum_{n=0}^{m-1} h_n \beta^{nk} \mod N, \quad k = 0, 1, 2, \ldots, m - 1$$  \hfill (1.1)$$

and the Inverse number theoretic transforms (INTT) of $\{H_n\}$ is defined as

$$h_k = m^{-1} \sum_{n=0}^{m-1} H_n \beta^{-nk} \mod N, \quad k = 0, 1, 2, \ldots, m - 1$$  \hfill (1.2)$$
where \( mm^{-1} \equiv 1 \mod N \), \( \beta^m \equiv 1 \mod N \) and \( \sum_{k=0}^{m-1} \beta^{uk} = 0 \mod N \), for every \( u \) such that \( \frac{m}{u} \) is a prime.

1.4.2 Euler Totient Function

For an integer \( N \geq 1 \), the Euler totient function \( \phi(N) \) is defined to be the number of positive integers not exceeding \( N \), which are relatively prime to \( N \).

1.4.3 Hash Function

A hash function \( h : M \rightarrow D \) with \( |M| > |D| \), take a message as input strings of arbitrary finite length from \( M \) and produce an output of strings of fixed length in \( D \) referred as a hash value. The function is many-to-one implying the collisions.

A hash function is pre-image resistant or one way if for any \( d \in D \), it is computationally infeasible to find the pre-image \( m \in M \) such that \( d = h(m) \).

A hash function is collision resistant, if given the description of \( h \), it is computationally infeasible to find two distinct messages \( m_1, m_2 \in M \) such that \( h(m_1) = h(m_2) \).

1.4.4 Digital Signatures

A signature scheme is a method of signing a message stored in electronic form to provide data integrity in message communication. It consists of two components, namely
(i) signing algorithm
(ii) verification algorithm

A signature scheme is a five-tuple \((M, A, K, S, V)\), where the following conditions are satisfied:

- \(M\) is a set of possible messages
- \(A\) is a set of possible signatures
- \(K\), the key space is a finite set of possible keys
- For each \(k \in K\), there is a signing algorithm \(\text{sig}_k : M \rightarrow A\) in \(S\) and a verification algorithm \(\text{ver}_k : M \times A \rightarrow \{\text{true}, \text{false}\}\) in \(V\), such that for every \((x, y) \in M \times A:\)

\[
\text{ver}(x, y) = \begin{cases} 
\text{true} & \text{if } y = \text{sig}(x) \\
\text{false} & \text{if } y \neq \text{sig}(x) 
\end{cases}
\]

The fast public cryptographic hash functions are used to produce message digest of arbitrary length message and then the message digest is signed instead of the original long message to minimise the cost and time.

### 1.4.5 Group Signature Schemes

Group signatures allow a group of members to anonymously sign an arbitrary message on behalf of the group with or with out the assistance of a trusted authority.

### 1.4.6 Primitive Root

An integer \(g\) is called the primitive root \(\mod N\), if \(\phi(N)\) is the smallest positive integer satisfying \(g^{\phi(N)} \equiv 1 \mod N\).
1.4.7 Symmetric Functions

Let \( P(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1x + a_0 \) be a polynomial of degree \( n \) with roots \( x_i \) for \( i = 1, 2, \ldots, n \), then the symmetric functions of the roots in terms of the coefficients are given by

\[
\sum_{i=1}^{n} x_i = -a_{n-1}, \quad \sum_{1 \leq i < j \leq n} x_ix_j = a_{n-2}, \ldots, \quad \prod_{i=1}^{n} x_i = (-1)^n a_0
\]

1.4.8 Chinese Remainder Theorem

If \( m_1, m_2, \ldots, m_r \) are pairwise relatively prime integers greater than 1 and \( b_1, b_2, \ldots, b_r \) are any integers, then the system of linear congruences \( x \equiv b_1 \pmod{m_1}, \ x \equiv b_2 \pmod{m_2}, \ldots, \ x \equiv b_r \pmod{m_r} \) has a unique solution \( x = \sum_{k=1}^{r} b_k M_k M'_k \pmod{M} \), where \( M = \prod_{i=1}^{n} m_i \), \( M_k = \frac{M}{m_k} \) and \( M'_k = M_k^{-1} \pmod{m_k} \).

1.4.9 Elliptic Curve over a Prime Field \( \mathbb{F}_p \)

The elliptic curve \( E: y^2 = x^3 + ax + b \) over the prime field \( \mathbb{F}_p \), \( p > 3 \), is the set of points \( (x, y) \in \mathbb{F}_p \times \mathbb{F}_p \) satisfying \( y^2 = x^3 + ax + b \pmod{p} \), where \( a, b \in \mathbb{F}_p \) are constants such that \( 4a^3 + 27b^2 \neq 0 \pmod{p} \) together with a special point \( O_E \) called the point at infinity.
1.4.10 Addition of Points on Elliptic Curve

The addition $P \oplus Q$ of two distinct points $P$ and $Q$ on the elliptic curve $E: y^2 = x^3 + ax + b$ is geometrically represented in Figure 1.7(a). A chord (non-vertical) joining $P$ and $Q$ intersects the elliptic curve at a third point $R$, then $P \oplus Q$ is the reflection of $R$ in the $X$ axis. The addition of the same point $P$ is geometrically represented in Figure 1.7 (b). A tangent is drawn at $P$ and the tangent intersects the elliptic curve at a third point $R$, then $P \oplus P = 2P$ is the reflection of $R$ on in the $X$ axis.

![Graph of Addition of Points on Elliptic Curve](image)

(a) Distinct Points
(b) Same Point

Figure 1.7 Addition of points on elliptic curve

If $P = (x_1, y_1), Q = (x_2, y_2)$, then $P \oplus Q$ is given algebraically by,

$$P + Q = \begin{cases} O_E, & \text{if } x_1 = x_2 \text{ & } y_1 = -y_2 \\ (x_3, y_3), & \text{otherwise} \end{cases}$$
where, \((x_3,y_3) = (\lambda^2 - x_1 - x_2, \lambda(x_1 - x_3) - y_1)\) and 
\[
\lambda = \begin{cases} 
\frac{3x_1^2 + a}{2y_1} & \text{if } P = Q \\
\frac{y_2 - y_1}{x_2 - x_1} & \text{otherwise}
\end{cases}
\]

The notation \(nP\) stands for the addition of the point \(P \oplus P \oplus \ldots n\) times.

A point \(P\) on the elliptic curve is said to be of order \(n\), if \(n\) is the smallest positive integer such that \(nP = O_E\).

### 1.4.11 Analogue of the Elliptic Curve ElGamal Encryption

The communication between Alice and Bob is done as follows:

1. Alice and Bob choose an elliptic curve \(E(F_p)\) and a random base point \(P\) of order \(q\).
2. Alice and Bob chooses secret keys as the random integers \(r_a\) and \(r_b\). The public keys of Alice and Bob are \(r_aP\) and \(r_bP\).
3. To send a message point \(m\) to Bob, Alice chooses a random integer \(k\) and sends the pair of points \((kP, m + k(r_bP))\).
4. To read \(m\), Bob computes \(m + k(r_bP) - r_b(kP) = m\).
1.4.12 Homomorphic Encryption

If \((c_1, c_2) = (\alpha P, m + \alpha(bP))\) and \((c_1', c_2') = (\alpha' P, m' + \alpha'(bP))\) are encryptions of \(m\) and \(m'\), then \((c_1, c_2) + (c_1', c_2')\) is an encryption for \(m + m'\), since \((c_1, c_2) + (c_1', c_2') = (c_1 + c_1', c_2 + c_2') = (\alpha P + \alpha' P, m + \alpha(bP) + m' + \alpha'(bP)) = ((\alpha + \alpha') P, (m + m') + (\alpha + \alpha') (bP))\)

Thus, Elgamal encryption is homomorphic.

1.4.13 \((t, n)\) Secret Sharing Scheme

Let \(t, n\) be positive integers, \(t \leq n\). A \((t, n)\) threshold scheme is a method of sharing a key \(K\) among a set of \(n\) participants, in such a way that any \(t\) or more participants can compute the value of \(K\), but not less than \(t\) participants.

1.4.14 Shamir \((t, n)\) Secret Sharing Scheme

Suppose there are \(n\) participants involved in a communication then a secret key \(K\) can be shared among the participants as \((x_i, y_i = f(x_i))\) for \(i = 1, 2, \ldots, n\), where \(x_i\)'s are distinct and \(f(x)\) is a polynomial given by \(f(x) = K + a_1 x + a_2 x^2 + \ldots + a_{t-1} x^{t-1} \mod p\) with \(p\) to be a prime and \(t \leq n\). If, at least \(t\) participants pool their shares, then the secret polynomial \(f(x)\) can be reconstructed using Lagrange interpolation formula as \(f(x) = \sum_{j=1}^{t} y_j \prod_{1 \leq k \leq t, k \neq j} (x - x_k) \mod p\) and in turn the secret \(K\) is
obtained as \( K = f(0) = \sum_{j=1}^{t} y_j \prod_{1 \leq k \leq t, k \neq j} \frac{(0-x_k)}{(x_j-x_k)} \mod p \). The secret key is recovered only if at least \( t \)-participants join together and not less than \( t \).

### 1.4.15 Division Algorithm

Given integers \( a \) and \( b \) with \( b > 0 \), there exists a unique pair of integers \( q \) and \( r \), such that \( a = bq + r \), with \( 0 \leq r < b \).

### 1.4.16 B-smooth

An integer \( n \) is said to be B-smooth with respect to a bound \( B \), if all its prime factors are less than or equal to \( B \).

### 1.4.17 Factor Base

A set consisting of first \( t \) primes, \( \{ p_1, p_2, ..., p_t \} \), which are used to find pairs of integers \( (a_i, b_i) \) satisfying \( a_i^2 \equiv b_i \pmod{n} \) and \( b_i \) is \( p_i \)-smooth, where \( n \) is a composite number.

### 1.5 OBJECTIVE OF THE THESIS

The objective of this dissertation is

- To analyse the cryptographic goals and to develop a public key cryptosystem based on number theoretic transforms satisfying the basic security aspects in end to end communication and group communication.

- To develop signature schemes, which provide data integrity in group oriented communication.
To offer an interdisciplinary presentation on number theory and probability theory uses in the security of cryptographic algorithms and to discuss the practical application of cryptographic protocols in electronic voting using elliptic curves.

To develop an efficient special purpose deterministic factorization algorithm with analytical justification, since the security of many popular cryptosystems is based on the computationally hard integer factorization problem.

1.6 ORGANIZATION OF THE THESIS

The thesis is organized as follows.

In Chapter 2, a new public key cryptosystem based on number theoretic transforms satisfying the security goals confidentiality, authentication and data integrity is designed. The security of the developed cryptosystem is based on computational hard problems, integer factorization, discrete logarithm and Diffie-Hellman problem. The proposed public key cryptosystem is also extended to a group oriented communication environment to achieve the security goal privacy.

Chapter 3 analyzes the group oriented signature schemes, which provide data integrity in the transmission of messages in a group oriented communication. Four different group oriented signature schemes based on primitive roots, symmetric functions, elliptic curves and the Chinese remainder theorem are developed. In all the proposed schemes, the group signature is constructed with the assistance of a trusted authority.

Chapter 4 is devoted to develop suitable cryptographic protocols for electronic voting schemes, based on elliptic curves. The various voting
schemes, namely, single authority two way electronic voting scheme, multi authority two way electronic voting scheme and multi authority multi way electronic voting schemes, based on elliptic curves were discussed. Elliptic curve cryptography is a suitable algorithm for wireless devices such as handhelds and personal digital assistants (PDA), which have limited bandwidth and processing power. Elliptic curve cryptography uses keys of smaller size but offers the same level of security and improved performance over other public key cryptosystems like RSA, DSA, etc. Thus, elliptic curve cryptography needs smaller memory and processor requirements.

In Chapter 5, a special purpose factorization algorithm based on recursion having polynomial running time is developed to factorize an integer, which is a product of two primes. The analytical justification is provided for the proposed algorithm. The successes of popular public key cryptosystems RSA, DSA, ElGamal etc led to significant developments in the areas like multiplying, factoring and finding prime numbers. Factorization is the leading one among the various attacks on the popular cryptosystems and this is the motivation for the designing the algorithm.

Chapter 6 concludes the thesis by presenting overall work and their scope for the future enhancement.