B.Tech (Computer Science and Engineering) Syllabus for Admission Batch 2015-16 7th Semester

PCS7J001 Cryptography & Network Security 3-0-0

OBJECTIVES: The student should be made to:
• Understand OSI security architecture and classical encryption techniques.
• Acquire fundamental knowledge on the concepts of finite fields and number theory.
• Understand various block cipher and stream cipher models.
• Describe the principles of public key cryptosystems, hash functions and digital signature.

Module I : INTRODUCTION & NUMBER THEORY [10 hours]

Services, Mechanisms and attacks-the OSI security architecture-Network security model-Classical Encryption techniques (Symmetric cipher model, substitution techniques, transposition techniques, steganography).

FINITE FIELDS AND NUMBER THEORY: Groups, Rings, Fields-Modular arithmetic-Euclid‟s algorithm-Finite fields- Polynomial Arithmetic –Prime numbers-Fermat‟s and Euler‟s theorem-Testing for primality -The Chinese remainder theorem- Discrete logarithms.

Module II : BLOCK CIPHERS & PUBLIC KEY CRYPTOGRAPHY [10 hours]

Data Encryption Standard-Block cipher principles-block cipher modes of operation-Advanced Encryption Standard (AES)-Triple DES-Blowfish-RC5 algorithm. Public key cryptography: Principles of public key cryptosystems-The RSA algorithm-Key management – Diffie Hellman Key exchange-Elliptic curve arithmetic-Elliptic curve cryptography.

Module III : HASH FUNCTIONS AND DIGITAL SIGNATURES [10 hours]

Authentication requirement – Authentication function – MAC – Hash function – Security of hash function and MAC –MD5 – SHA – HMAC – CMAC – Digital signature and authentication protocols – DSS – EI Gamal – Schnorr.

SECURITY PRACTICE & SYSTEM SECURITY [8 hours]

Authentication applications – Kerberos – X.509 Authentication services – Internet Firewalls for Trusted System: Roles of Firewalls – Firewall related terminology- Types of Firewalls – Firewall designs – SET for E-Commerce Transactions. Intruder – Intrusion detection system – Virus and related threats – Countermeasures – Firewalls design principles – Trusted systems – Practical implementation of cryptography and security.

Module IV : E-MAIL, IP & WEB SECURITY [9 hours]

E-mail Security: Security Services for E-mail-attacks possible through E-mail – establishing keys privacy-authentication of the source-Message Integrity -Non-repudiation- Pretty Good Privacy-S/MIME. IP Security: Overview of IPSec – IP and IPv6-Authentication Header-Encapsulation Security Payload (ESP)-Internet Key Exchange (Phases of IKE, ISAKMP/IKE Encoding).

Web Security: SSL/TLS Basic Protocol-computing the keys- client authentication-PKI as deployed by SSL Attacks fixed in v3- Exportability-Encoding-Secure Electronic Transaction (SET).

OUTCOMES: Upon Completion of the course, the students should be able to:
• Compare various Cryptographic Techniques
• Design Secure applications
• Inject secure coding in the developed applications

TEXT BOOKS:

1. William Stallings, Cryptography and Network Security, 6th Edition, Pearson Education, March 2013. (UNIT I,II,III,IV).
2. Charlie Kaufman, Radia Perlman and Mike Speciner, "Network Security", Prentice Hall of India, 2002. (UNIT V).

REFERENCES:

1. Behrouz A. Ferouzan, "Cryptography & Network Security", Tata Mc Graw Hill, 2007.
2. Man Young Rhee, "Internet Security: Cryptographic Principles", "Algorithms and Protocols", Wiley Publications, 2003.
3. Charles Pfleeger, "Security in Computing", 4th Edition, Prentice Hall of India, 2006.
4. Ulysess Black, "Internet Security Protocols", Pearson Education Asia, 2000.
5. Charlie Kaufman and Radia Perlman, Mike Speciner, "Network Security, Second Edition, Private Communication in Public World", PHI 2002.
6. Bruce Schneier and Neils Ferguson, "Practical Cryptography", First Edition, Wiley Dreamtech India Pvt Ltd, 2003.
7. Douglas R Simson "Cryptography – Theory and practice", First Edition, CRC Press, 1995.
8. http://nptel.ac.in/.

UNIT 1
INTRODUCTION AND NUMBER THEORY

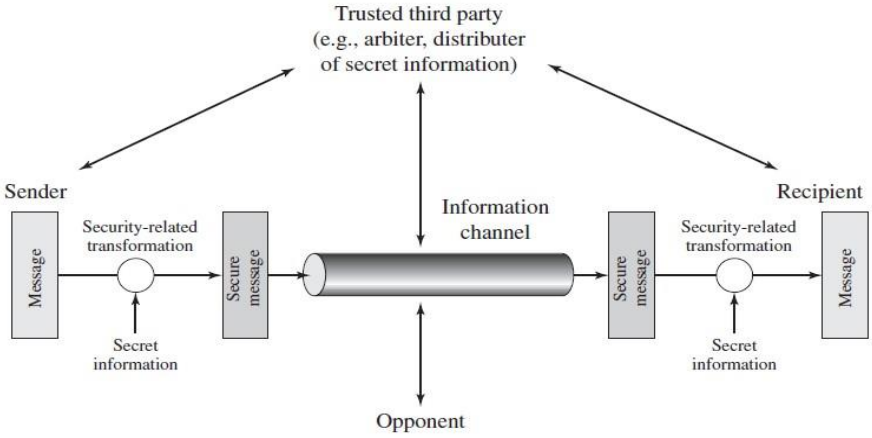
Lecture-1

1.1 INTRODUCTION

- Computer data often travels from one computer to another, leaving the safety of its protected physical surroundings. Once the data is out of hand, people with bad intention could modify or forge your data, either for amusement or for their own benefit.
- Cryptography can reformat and transform our data, making it safer on its trip between computers. The technology is based on the essentials of secret codes, augmented by modern mathematics that protects our data in powerful ways.
    - **Computer Security:** generic name for the collection of tools designed to protect data and to thwart hackers
    - **Network Security:** measures to protect data during their transmission
    - **Internet Security:** measures to protect data during their transmission over a collection of interconnected networks

1.2 THE OSI SECURITY ARCHITECTURE

- To assess effectively the security needs of an organization and to evaluate and choose various security products and policies, the manager responsible for security needs some systematic way of defining the requirements for security and characterizing the approaches to satisfying those requirements.



A MODEL FOR NETWORK SECURITY

- The OSI security architecture was developed in the context of the OSI protocol architecture.
- The OSI security architecture focuses on security attacks, mechanisms, and services. These can be defined briefly as follows

i. Threats and Attacks

- **Threat:** A potential for violation of security, which exists when there is a circumstance, capability, action, or event that could breach security and cause harm. That is, a threat is a possible danger that might exploit a vulnerability.

- **Attack:** An assault on system security that derives from an intelligent threat; that is, an intelligent act that is a deliberate attempt (especially in the sense of a method or technique) to evade security services and violate the security policy of a system.

ii. Security Attacks, Services And Mechanisms

To assess the security needs of an organization effectively, the manager responsible for security needs some systematic way of defining the requirements for security and characterization of approaches to satisfy those requirements. One approach is to consider three aspects of information security:

- **Security attack:** Any action that compromises the security of information owned by an organization.
- **Security mechanism:** A mechanism that is designed to detect, prevent or recover from a security attack.
- **Security service:** A service that enhances the security of the data processing systems and the information transfers of an organization. The services are intended to counter security attacks and they make use of one or more security mechanisms to provide the service.

1.2.1 SECURITY SERVICES

The classification of security services are as follows:

- **Confidentiality:** Ensures that the information in a computer system and transmitted information are accessible only for reading by authorized parties.
  Eg., printing, displaying and other forms of disclosure.
- **Authentication:** Ensures that the origin of a message or electronic document is correctly identified, with an assurance that the identity is not false.
- **Integrity:** Ensures that only authorized parties are able to modify computer system assets and transmitted information. Modification includes writing, changing status, deleting, creating and delaying or replaying of transmitted messages.
- **Non repudiation**: Requires that neither the sender nor the receiver of a message be able to deny the transmission.
- **Access control**: Requires that access to information resources may be controlled by or the target system.
- **Availability**: Requires that computer system assets be available to authorized parties when needed.

Lecture-02

Security Services (As per Standard X.800)

i. AUTHENTICATION: The assurance that the communicating entity is the one that it claims to be.

- **Peer Entity Authentication:** Used in association with a logical connection to provide confidence in the identity of the entities connected.
- **Data Origin Authentication:** In a connectionless transfer, provides assurance that the source of received data is as claimed.

ii. ACCESS CONTROL: The prevention of unauthorized use of a resource (i.e., this service controls who can have access to a resource, under what conditions access can occur, and what those accessing the resource are allowed to do).

iii. DATA CONFIDENTIALITY: The protection of data from unauthorized disclosure.

- **Connection Confidentiality:** The protection of all user data on a connection.
- **Connectionless Confidentiality:** The protection of all user data in a single data block.

- **Traffic Flow Confidentiality:** The protection of the information that might be derived from observation of traffic flows.

iv. INTEGRITY:

- **Connection Integrity with Recovery:** Provides for the integrity of all user data on a connection and detects any modification, insertion, deletion, or replay of any data within an entire data sequence, with recovery attempted.
- **Connection Integrity without Recovery:** Almost same as connection integrity without recovery, but provides only detection without recovery.
- **Selective-Field Connection Integrity:** Provides for the integrity of selected fields within the user data of a data block transferred over a connection and takes the form of determination of whether the selected fields have been modified, inserted, deleted, or replayed.
- **Connectionless Integrity:** Provides for the integrity of a single connectionless data block and may take the form of detection of data modification. Additionally, a limited form of replay detection may be provided.
- **Selective-Field Connectionless Integrity:** Provides for the integrity of selected fields within a single connectionless data block; takes the form of determination of whether the selected fields have been modified.

V. NONREPUDIATION:

Provides protection against denial by one of the entities involved in a communication of having participated in all or part of the communication.

- **Nonrepudiation, Origin:** Proof that the message was sent by the specified party.
- **Nonrepudiation, Destination:** Proof that the message was accepted by the specified party.

1.2.2 SECURITY MECHANISMS

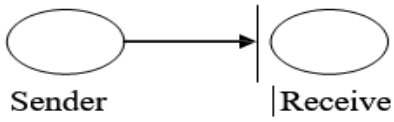
- One of the most specific security mechanisms in use is cryptographic techniques. Encryption or encryption-like transformations of information are the most common means of providing security. Some of the mechanisms are:
    - Encipherment: the process of making data unreadable to unauthorized entities by applying a cryptographic algorithm (an encryption algorithm). Decipherment (decryption) is the reverse operation by which the ciphertext is transformed to the plaintext.
    - Digital Signature: A digital signature is a mathematical scheme for verifying the authenticity of digital messages or documents. A valid digital signature, where the prerequisites are satisfied, gives a recipient very strong reason to believe that the message was created by a known sender (authentication), and that the message was not altered in transit (integrity). Digital signatures are a standard element of most cryptographic protocol suites, and are commonly used for software distribution, financial transactions, contract management software, and in other cases where it is important to detect forgery or tampering.
    - Access Control: This systems perform identification authentication and authorization of users and entities by evaluating required login credentials that can include passwords, personal identification numbers (PINs), biometric scans, security tokens or other authentication factors. Multifactor authentication, which requires two or more

authentication factors, is often an important part of layered defense to protect access control systems.

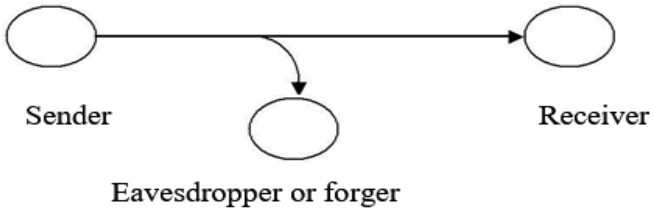
1.2.3 SECURITY ATTACKS

There are four general categories of attack which are listed below.

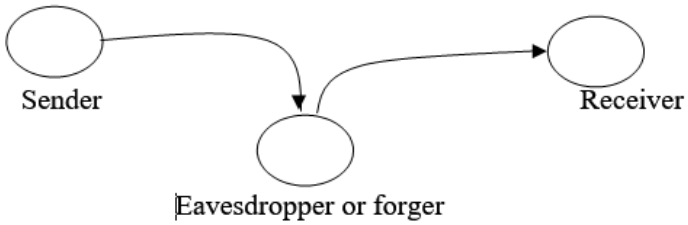
- **Interruption:** An asset of the system is destroyed or becomes unavailable or unusable. This is an attack on availability. e.g., destruction of piece of hardware, cutting of a communication line or disabling of file management system.



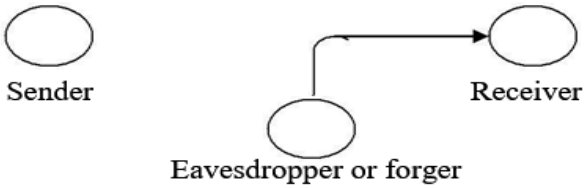
- **Interception:** An unauthorized party gains access to an asset. This is an attack on confidentiality. Unauthorized party could be a person, a program or a computer. e.g., wire tapping to capture data in the network, illicit copying of files



- **Modification:** An unauthorized party not only gains access to but tampers with an asset. This is an attack on integrity. e.g., changing values in data file, altering a program, modifying the contents of messages being transmitted in a network.



- **Fabrication:** An unauthorized party inserts counterfeit objects into the system. This is an attack on authenticity. e.g., insertion of spurious message in a network or addition of records to a file.



A useful categorization of these attacks is in terms of

i. Passive attacks
ii. Active attacks

Passive attack

Passive attacks are in the nature of eavesdropping on, or monitoring of, transmissions. Passive attacks are very difficult to detect because they do not involve any alteration of data. However, it is feasible to prevent the success of these attacks. The goal of the opponent is to obtain information that is being transmitted. Passive attacks are of two types:

- **Release of message contents:** A telephone conversation, an e-mail message and a transferred file may contain sensitive or confidential information. We would like to prevent the opponent from learning the contents of these transmissions.
- **Traffic analysis**: If we had encryption protection in place, an opponent might still be able to observe the pattern of the message. The opponent could determine the location and identity of communication hosts and could observe the frequency and length of messages being exchanged. This information might be useful in guessing the nature of communication that was taking place.

Active attacks

These attacks involve some modification of the data stream or the creation of a false stream. It is quite difficult to prevent active attacks absolutely, because to do so would require physical protection of all communication facilities and paths at all times. Instead, the goal is to detect them and to recover from any disruption or delays caused by them.
These attacks can be classified in to four categories:

- **Masquerade:** One entity pretends to be a different entity.
- **Replay:** involves passive capture of a data unit and its subsequent transmission to produce an unauthorized effect.
- **Modification of messages:** Some portion of message is altered or the messages are delayed or recorded, to produce an unauthorized effect.
- **Denial of service:** Prevents or inhibits the normal use or management of communication facilities. Another form of service denial is the disruption of an entire network, either by disabling the network or overloading it with messages so as to degrade performance.

LECTURE - 03

1.3 CLASSICAL CRYPTO SYSTEMS

1.3.1 CONVENTIONAL ENCRYPTION

- referred as conventional / private-key / single-key
- sender and recipient share a common key
- all classical encryption algorithms are private-key
- was only type prior to invention of public-key in 1970’s

Some basic terminologies used:

- **plaintext** - the original message
- **Ciphertext** - the coded message
- **cipher** - algorithm for transforming plaintext to ciphertext
- **key** - info used in cipher known only to sender/receiver
- **encipher (encrypt)** - converting plaintext to ciphertext

- **decipher (decrypt)** - recovering ciphertext from plaintext
- **cryptography** - study of encryption principles/methods
- **cryptanalysis (codebreaking)** - the study of principles/ methods of deciphering ciphertext *without* knowing key
- **cryptology** - the field of both cryptography and cryptanalysis

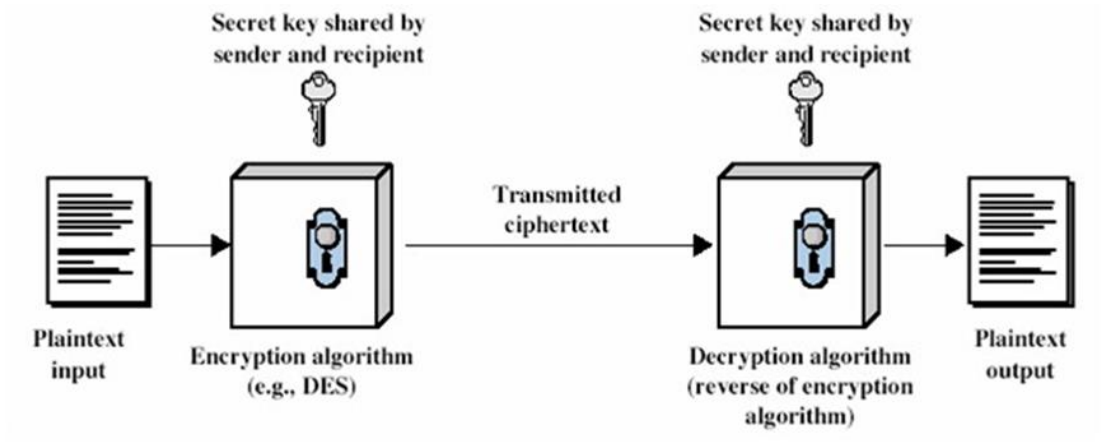


Figure: Conventional Encryption

- Here the original message, referred to as plaintext, is converted into apparently random nonsense, referred to as cipher text. The encryption process consists of an algorithm and a key. The key is a value independent of the plaintext. Changing the key changes, the output of the algorithm. Once the cipher text is produced, it may be transmitted. Upon reception, the cipher text can be transformed back to the original plaintext by using a decryption algorithm and the same key that was used for encryption.
- The security depends on several factors. First, the encryption algorithm must be powerful enough that it is impractical to decrypt a message on the basis of cipher text alone. Beyond that, the security depends on the secrecy of the key, not the secrecy of the algorithm.

1.3.2 Types of Cryptosystems

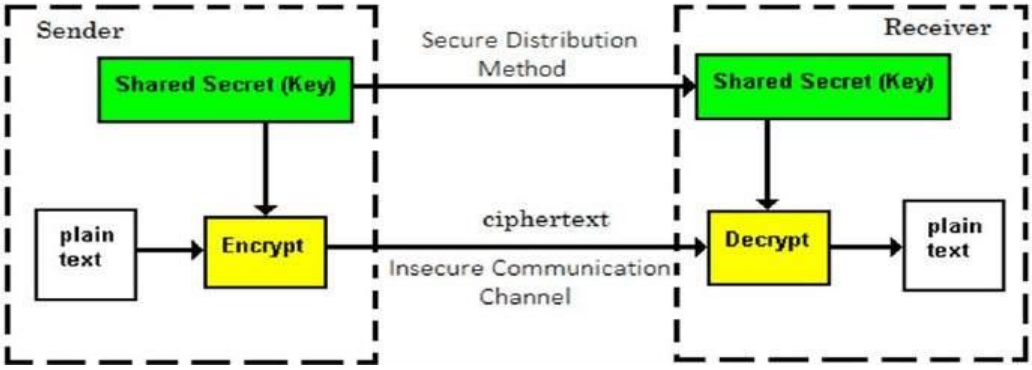
Fundamentally, there are two types of cryptosystems based on the manner in which encryption-decryption is carried out in the system −

- Symmetric Key Encryption
- Asymmetric Key Encryption

The main difference between these cryptosystems is the relationship between the encryption and the decryption key. Logically, in any cryptosystem, both the keys are closely associated. It is practically impossible to decrypt the ciphertext with the key that is unrelated to the encryption key.

(i) Symmetric Key Encryption:

- The encryption process where same keys are used for encrypting and decrypting the information is known as Symmetric Key Encryption.
- The study of symmetric cryptosystems is referred to as **symmetric cryptography**. Symmetric cryptosystems are also sometimes referred to as **secret key cryptosystems**.
- A few well-known examples of symmetric key encryption methods are − Digital Encryption Standard (DES), Triple-DES (3DES), IDEA, and BLOWFISH.



Prior to 1970, all cryptosystems employed symmetric key encryption. Even today, its relevance is very high and it is being used extensively in many cryptosystems. It is very unlikely that this encryption will fade away, as it has certain advantages over asymmetric key encryption.

The salient features of cryptosystem based on symmetric key encryption are −

- Persons using symmetric key encryption must share a common key prior to exchange of information.
- Keys are recommended to be changed regularly to prevent any attack on the system.
- A robust mechanism needs to exist to exchange the key between the communicating parties. As keys are required to be changed regularly, this mechanism becomes expensive and cumbersome.
- In a group of **n** people, to enable two-party communication between any two persons, the number of keys required for group is **n × (n – 1)/2**.
- Length of Key (number of bits) in this encryption is smaller and hence, process of encryption-decryption is faster than asymmetric key encryption.
- Processing power of computer system required to run symmetric algorithm is less.

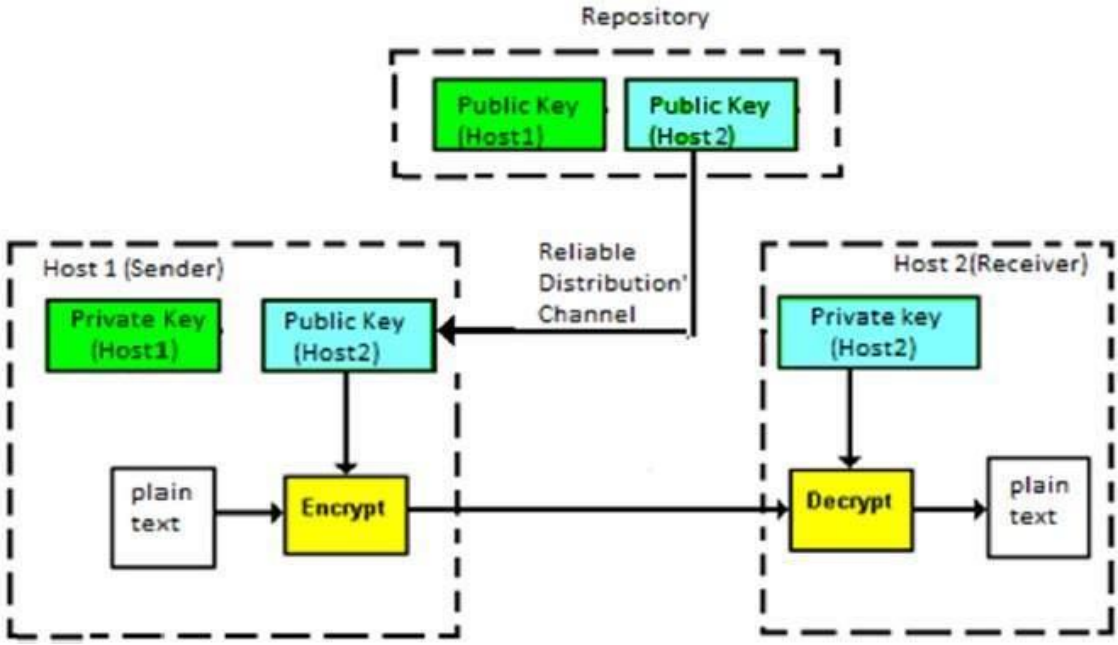
Challenge of Symmetric Key Cryptosystem

There are two restrictive challenges of employing symmetric key cryptography.

- **Key establishment** − Before any communication, both the sender and the receiver need to agree on a secret symmetric key. It requires a secure key establishment mechanism in place.
- **Trust Issue** − Since the sender and the receiver use the same symmetric key, there is an implicit requirement that the sender and the receiver 'trust' each other. For example, it may happen that the receiver has lost the key to an attacker and the sender is not informed.

These two challenges are highly restraining for modern day communication. Today, people need to exchange information with non-familiar and non-trusted parties. For example, a communication between online seller and customer. These limitations of symmetric key encryption gave rise to asymmetric key encryption schemes.

ii. Asymmetric Key Encryption: The encryption process where different keys are used for encrypting and decrypting the information is known as Asymmetric Key Encryption. Though the keys are different, they are mathematically related and hence, retrieving the plaintext by decrypting ciphertext is feasible. The process is depicted in the following illustration −



Asymmetric Key Encryption was invented in the 20th century to come over the necessity of pre-shared secret key between communicating persons.

The salient features of this encryption scheme are as follows −

- Every user in this system needs to have a pair of dissimilar keys, **private key** and **public key**. These keys are mathematically related − when one key is used for encryption, the other can decrypt the ciphertext back to the original plaintext.
- It requires to put the public key in public repository and the private key as a well-guarded secret. Hence, this scheme of encryption is also called **Public Key Encryption**.
- Though public and private keys of the user are related, it is computationally not feasible to find one from another. This is a strength of this scheme.
- When *Host1* needs to send data to *Host2*, he obtains the public key of *Host2* from repository, encrypts the data, and transmits.
- *Host2* uses his private key to extract the plaintext.
- Length of Keys (number of bits) in this encryption is large and hence, the process of encryption-decryption is slower than symmetric key encryption.
- Processing power of computer system required to run asymmetric algorithm is higher.

LECTURE-04

Challenge of Public Key Cryptosystem

- Public-key cryptosystems have one significant challenge − the user needs to trust that the public key that he is using in communications with a person really is the public key of that person and has not been spoofed by a malicious third party.
- This is usually accomplished through a Public Key Infrastructure (PKI) consisting a trusted third party. The third party securely manages and attests to the authenticity of public keys. When the third party is requested to provide the public key for any communicating person X, they are trusted to provide the correct public key.

- The third party satisfies itself about user identity by the process of attestation, notarization, or some other process − that X is the one and only, or globally unique, X. The most common method of making the verified public keys available is to embed them in a certificate which is digitally signed by the trusted third party.

Relation between Encryption Schemes:

A summary of basic key properties of two types of cryptosystems is given below −

	Symmetric Cryptosystems	**Public Key Cryptosystems**
Relation between Keys	Same	Different, but mathematically related
Encryption Key	Symmetric	Public
Decryption Key	Symmetric	Private

Due to the advantages and disadvantage of both the systems, symmetric key and public-key cryptosystems are often used together in the practical information security systems.

Kerckhoff's Principle for Cryptosystem

- In the 19th century, a Dutch cryptographer A. Kerckhoff furnished the requirements of a good cryptosystem. Kerckhoff stated that a cryptographic system should be secure even if everything about the system, except the key, is public knowledge.
- The six design principles defined by Kerckhoff for cryptosystem are −
  - The cryptosystem should be unbreakable practically, if not mathematically.
  - Falling of the cryptosystem in the hands of an intruder should not lead to any compromise of the system, preventing any inconvenience to the user.
  - The key should be easily communicable, memorable, and changeable.
  - The ciphertext should be transmissible by telegraph, an unsecure channel.
  - The encryption apparatus and documents should be portable and operable by a single person.
  - Finally, it is necessary that the system be easy to use, requiring neither mental strain nor the knowledge of a long series of rules to observe.
- The second rule is currently known as **Kerckhoff principle**. It is applied in virtually all the contemporary encryption algorithms such as DES, AES, etc. These public algorithms are considered to be thoroughly secure. The security of the encrypted message depends solely on the security of the secret encryption key.
- Keeping the algorithms secret may act as a significant barrier to cryptanalysis. However, keeping the algorithms secret is possible only when they are used in a strictly limited circle.
- In modern era, cryptography needs to protect the users' data who are connected to the Internet. In such cases, using a secret algorithm is not feasible, hence Kerckhoff principles became essential guidelines for designing algorithms in modern cryptography.

LECTURE-05

Cryptography:

Cryptographic systems are generally classified along 3 independent dimensions:

- **Type of operations used for transforming plain text to cipher text:** All the encryption algorithms are abased on two general principles: **substitution**, in which each element in the plaintext is mapped into another element, and **transposition**, in which elements in the plaintext are rearranged.
- **The number of keys used:**
  - If the sender and receiver uses same key then it is said to be **symmetric key (or) single key (or) conventional encryption**.
  - If the sender and receiver use different keys then it is said to be **public key encryption**.

The way in which the plain text is processed

- A **block cipher** processes the input and block of elements at a time, producing output block for each input block.
- A **stream cipher** processes the input elements continuously, producing output element one at a time, as it goes along.

Cryptanalysis:

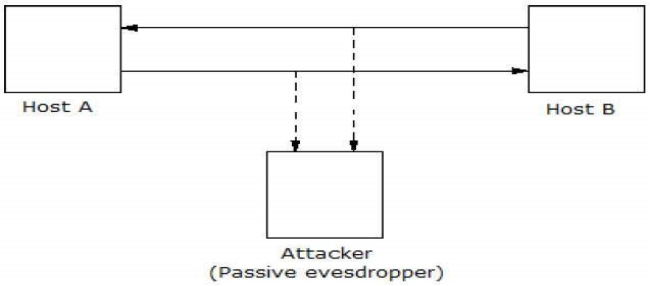
The process of attempting to discover **"Plaintext/original message"** or **"Key"** or both is known as cryptanalysis. The strategy used by the cryptanalysis depends on the nature of the encryption scheme and the information available to the cryptanalyst.

Attacks On Cryptosystems

- In the present era, almost all the aspects of human life are driven by information. Hence, it has become imperative to protect useful information from malicious activities such as attacks. Let us consider the types of attacks to which information is typically subjected to.
- Attacks are typically categorized based on the action performed by the attacker. An attack, thus, can be **passive** or **active**.

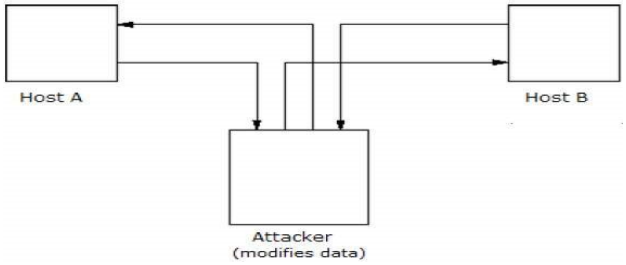
i. Passive Attacks:

- The main goal of a passive attack is to obtain unauthorized access to the information. For example, actions such as intercepting and eavesdropping on the communication channel can be regarded as passive attack.
- These actions are passive in nature, as they neither affect information nor disrupt the communication channel. A passive attack is often seen as *stealing* information. The only difference in stealing physical goods and stealing information is that theft of data still leaves the owner in possession of that data. Passive information attack is thus more dangerous than stealing of goods, as information theft may go unnoticed by the owner.



ii. Active Attacks: An active attack involves changing the information in some way by conducting some process on the information. For example,

- Modifying the information in an unauthorized manner.
- Initiating unintended or unauthorized transmission of information.
- Alteration of authentication data such as originator name or timestamp associated with information
- Unauthorized deletion of data.
- Denial of access to information for legitimate users (denial of service).



Cryptography provides many tools and techniques for implementing cryptosystems capable of preventing most of the attacks described above.

Assumptions of Attacker:

Let us see the prevailing environment around cryptosystems followed by the types of attacks employed to break these systems:

i. Environment around Cryptosystem:

- While considering possible attacks on the cryptosystem, it is necessary to know the cryptosystems environment. The attacker's assumptions and knowledge about the environment decides his capabilities.
- In cryptography, the following three assumptions are made about the security environment and attacker's capabilities.

a. Details of the Encryption Scheme: The design of a cryptosystem is based on the following two cryptography algorithms −

- **Public Algorithms:** With this option, all the details of the algorithm are in the public domain, known to everyone.
- **Proprietary algorithms:** The details of the algorithm are only known by the system designers and users.

b. Availability of Ciphertext: We know that once the plaintext is encrypted into ciphertext, it is put on unsecure public channel (say email) for transmission. Thus, the attacker can obviously assume that it has access to the ciphertext generated by the cryptosystem.

c. Availability of Plaintext and Ciphertext: This assumption is not as obvious as other. However, there may be situations where an attacker can have access to plaintext and corresponding ciphertext. Some such possible circumstances are −

- The attacker influences the sender to convert plaintext of his choice and obtains the ciphertext.
- The receiver may divulge the plaintext to the attacker inadvertently. The attacker has access to corresponding ciphertext gathered from open channel.
- In a public-key cryptosystem, the encryption key is in open domain and is known to any potential attacker. Using this key, he can generate pairs of corresponding plaintexts and ciphertexts.

LCTURE-06

Cryptographic Attacks:

- The basic intention of an attacker is to break a cryptosystem and to find the plaintext from the ciphertext. To obtain the plaintext, the attacker only needs to find out the secret decryption key, as the algorithm is already in public domain.
- Hence, he applies maximum effort towards finding out the secret key used in the cryptosystem. Once the attacker is able to determine the key, the attacked system is considered as *broken* or *compromised*.
- Based on the methodology used, attacks on cryptosystems are categorized as follows −

- **Ciphertext Only Attacks (COA):** In this method, the attacker has access to a set of ciphertext(s). He does not have access to corresponding plaintext. COA is said to be successful when the corresponding plaintext can be determined from a given set of ciphertext. Occasionally, the encryption key can be determined from this attack. Modern cryptosystems are guarded against ciphertext-only attacks.
- **Known Plaintext Attack (KPA):** In this method, the attacker knows the plaintext for some parts of the ciphertext. The task is to decrypt the rest of the ciphertext using this information. This may be done by determining the key or via some other method. The best example of this attack is *linear cryptanalysis* against block ciphers.
- **Chosen Plaintext Attack (CPA):** In this method, the attacker has the text of his choice encrypted. So he has the ciphertext-plaintext pair of his choice. This simplifies his task of determining the encryption key. An example of this attack is *differential cryptanalysis* applied against block ciphers as well as hash functions. A popular public key cryptosystem, RSA is also vulnerable to chosen-plaintext attacks.
- **Dictionary Attack:** This attack has many variants, all of which involve compiling a 'dictionary'. In simplest method of this attack, attacker builds a dictionary of ciphertexts and corresponding plaintexts that he has learnt over a period of time. In future, when an attacker gets the ciphertext, he refers the dictionary to find the corresponding plaintext.
- **Brute Force Attack (BFA):** In this method, the attacker tries to determine the key by attempting all possible keys. If the key is 8 bits long, then the number of possible keys is 2^8=256. The attacker knows the ciphertext and the algorithm, now he attempts all the 256 keys one by one for decryption. The time to complete the attack would be very high if the key is long.

- **Birthday Attack:** This attack is a variant of brute-force technique. It is used against the cryptographic hash function. When students in a class are asked about their birthdays, the answer is one of the possible 365 dates. Similarly, if the hash function produces 64-bit hash values, the possible hash values are $1.8x10^{19}$. By repeatedly evaluating the function for different inputs, the same output is expected to be obtained after about $5.1x10^{9}$ random inputs. If the attacker is able to find two different inputs that give the same hash value, it is a **collision** and that hash function is said to be broken.
- **Man in Middle Attack (MIM):** The targets of this attack are mostly public key cryptosystems where key exchange is involved before communication takes place.
    - Host *A* wants to communicate to host *B*, hence requests public key of *B*.
    - An attacker intercepts this request and sends his public key instead.
    - Thus, whatever host *A* sends to host *B*, the attacker is able to read.
    - In order to maintain communication, the attacker re-encrypts the data after reading with his public key and sends to *B*.
    - The attacker sends his public key as *A*'s public key so that *B* takes it as if it is taking it from *A*.
- **Side Channel Attack (SCA):** This type of attack is not against any particular type of cryptosystem or algorithm. Instead, it is launched to exploit the weakness in physical implementation of the cryptosystem.
- **Timing Attacks:** They exploit the fact that different computations take different times to compute on processor. By measuring such timings, it is be possible to know about a particular computation the processor is carrying out. For example, if the encryption takes a longer time, it indicates that the secret key is long.
- **Power Analysis Attacks:** These attacks are similar to timing attacks except that the amount of power consumption is used to obtain information about the nature of the underlying computations.
- **Fault analysis Attacks:** In these attacks, errors are induced in the cryptosystem and the attacker studies the resulting output for useful information.

1.3.2 STEGANOGRAPHY

- A plaintext message may be hidden in any one of the two ways. The methods of steganography conceal the existence of the message, whereas the methods of cryptography render the message unintelligible to outsiders by various transformations of the text.
- A simple form of steganography, but one that is time consuming to construct is one in which an arrangement of words or letters within an apparently innocuous text spells out the real message.

Example:

(i) the sequence of first letters of each word of the overall message spells out the real (hidden) message.

(ii) Subset of the words of the overall message is used to convey the hidden message.

Various other techniques have been used historically, some of them are:

- **Character marking:** selected letters of printed or typewritten text are overwritten in pencil. The marks are ordinarily not visible unless the paper is held to an angle to bright light.
- **Invisible ink:** a number of substances can be used for writing but leave no visible trace until heat or some chemical is applied to the paper.
- **Pin punctures** – small pin punctures on selected letters are ordinarily not visible unless the paper is held in front of the light.

- **Typewritten correction ribbon** – used between the lines typed with a black ribbon, the results of typing with the correction tape are visible only under a strong light.

Drawbacks of steganography

- Requires a lot of overhead to hide a relatively few bits of information.
- Once the system is discovered, it becomes virtually worthless.

LECTURE - 07

1.4 CLASSICAL ENCRYPTION TECHNIQUES

There are two basic building blocks of all encryption techniques: substitution and transposition.

1.4.1 SUBSTITUTION TECHNIQUES

A substitution technique is one in which the letters of plaintext are replaced by other letters or by numbers or symbols. If the plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with cipher text bit patterns.

(i) **Caesar cipher (or) shift cipher**

The earliest known use of a substitution cipher and the simplest was by Julius Caesar. The Caesar cipher involves replacing each letter of the alphabet with the letter standing 3 places further down the alphabet.

e.g., Plain text: pay more money
Cipher text: SDB PRUH PRQHB

Note that the alphabet is wrapped around, so that letter following "z" is "a". For each plaintext letter p, substitute the cipher text letter c such that $C = E(p) = (p+3) \bmod 26$

A shift may be any amount, so that general Caesar algorithm is

- $C = E(p) = (p+k) \bmod 26$

Where k takes on a value in the range 1 to 25. The decryption algorithm is simply

- $P = D(C) = (C-k) \bmod 26$

(ii) Playfair cipher

- The best-known multiple letter encryption cipher is the playfair, which treats digrams in the plaintext as single units and translates these units into cipher text digrams. The playfair algorithm is based on the use of 5x5 matrix of letters constructed using a keyword. Let the keyword be „monarchy". The matrix is constructed by filling in the letters of the keyword (minus duplicates) from left to right and from top to bottom, and then filling in the remainder of the matrix with the remaining letters in alphabetical order.
- The letter "i" and "j" count as one letter. Plaintext is encrypted two letters at a time according to the following rules:
  - Repeating plaintext letters that would fall in the same pair are separated with a filler letter such as "x".
  - Plaintext letters that fall in the same row of the matrix are each replaced by the letter to the right, with the first element of the row following the last.
  - Plaintext letters that fall in the same column are replaced by the letter beneath, with the top element of the column following the last.
  - Otherwise, each plaintext letter is replaced by the letter that lies in its own row and the column occupied by the other plaintext letter.

M	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

Example: Plaintext = meet me at the school house
Splitting two letters as a unit => me et me at th es ch ox ol ho us ex
Corresponding cipher text => CL KL CL RS PD IL HY AV MP HF XL IU

Strength of playfair cipher

- Playfair cipher is a great advance over simple mono alphabetic ciphers.
- Since there are 26 letters, 26x26 = 676 diagrams are possible, so identification of individual digram is more difficult.
- Frequency analysis is much more difficult.

(iii)Polyalphabetic ciphers

- Another way to improve on the simple monoalphabetic technique is to use different monoalphabetic substitutions as one proceeds through the plaintext message. The general name for this approach is polyalphabetic cipher.
- All the techniques have the following features in common.
  - A set of related monoalphabetic substitution rules are used
  - A key determines which particular rule is chosen for a given transformation.

(iv) One Time Pad Cipher

- It is an unbreakable cryptosystem. It represents the message as a sequence of 0s and 1s. this can be accomplished by writing all numbers in binary, for example, or by using ASCII. The key is a random sequence of 0"s and 1"s of same length as the message.
- Once a key is used, it is discarded and never used again. The system can be expressed as follows:

$$C_i = P_i \oplus K_i$$

C_i - i^{th} binary digit of cipher text P_i - i^{th} binary digit of plaintext

K_i - i^{th} binary digit of key $\oplus$ – exclusive OR opearaiton

- Thus, the cipher text is generated by performing the bitwise XOR of the plaintext and the key. Decryption uses the same key. Because of the properties of XOR, decryption simply involves the same bitwise operation:

$$P_i = C_i \oplus K_i$$

e.g., plaintext = 0 0 1 0 1 0 0 1
Key = 1 0 1 0 1 1 0 0
---
ciphertext = 1 0 0 0 0 1 0 1

Advantage:

- Encryption method is completely unbreakable for a ciphertext only attack.

Disadvantages

- It requires a very long key which is expensive to produce and expensive to transmit.
- Once a key is used, it is dangerous to reuse it for a second message; any knowledge on the first message would give knowledge of the second.

LECTURE - 08

(v) Vigenere cipher

- In this scheme, the set of related monoalphabetic substitution rules consisting of 26 caesar ciphers with shifts of 0 through 25. Each cipher is denoted by a key letter.
- Ex: Caesar cipher with a shift of 3 is denoted by the key value "d" (since a=0, b=1, c=2 and so on).
- To aid in understanding the scheme, a matrix known as vigenere tableau is constructed.
- Each of the 26 ciphers is laid out horizontally, with the key letter for each cipher to its left. A normal alphabet for the plaintext runs across the top. The process of encryption is simple: Given a key letter X and a plaintext letter y, the cipher text is at the intersection of the row labeled x and the column labeled y; in this case, the ciphertext is V.
- To encrypt a message, a key is needed that is as long as the message. Usually, the key is a repeating keyword.

e.g., Key = d e c e p t i v e d e c e p t i v e d e c e p t i v e

PT = w e a r e d i s c o v e r e d s a v e y o u r s e l f
CT = ZICVTWQNGRZGVTWAVZHCQYGLMGJ

	PLAIN TEXT															
KEY LETTERS		a	B	c	D	E	F	g	h	i	j	k	…	X	y	Z
	A	A	B	C	D	E	F	G	H	I	J	K	…	X	Y	Z
	B	B	C	D	E	F	G	H	I	J	K	L	…	Y	Z	A
	C	C	D	E	F	G	H	I	J	K	L	M	…	Z	A	B
	D	D	E	F	G	H	I	J	K	L	M	N	…	A	B	C
	E	E	F	G	H	I	J	K	L	M	N	O	…	B	C	D
	F	F	G	H	I	J	K	L	M	N	O	P	…	C	D	E
	G	G	H	I	J	K	L	M	N	O	P	Q	…	D	E	F
	:	:	:	:	:	:	:	:	:	:	:	:	…	:	:	:
	X	X	Y	Z	A	B	C	D	E	F	G	H	…	U	V	W
	Y	Y	Z	A	B	C	D	E	F	G	H	I	…			X
	Z	Z	A	B	C	D	E	F	G	H	I	J	…			Y

1.4.2 TRANSPOSITION TECHNIQUES

All the techniques examined so far involve the substitution of a cipher text symbol for a plaintext symbol. A very different kind of mapping is achieved by performing some sort of permutation on the plaintext letters. This technique is referred to as a transposition cipher.

(i) Rail fence: It is simplest of such cipher, in which the plaintext is written down as a sequence of diagonals and then read off as a sequence of rows.

Ex: Plaintext = meet at the school house

To encipher this message with a rail fence of depth 2, we write the message as follows:

m e a t e c o l o s
e t t h s h o h u e

The encrypted message is: MEATECOLOSETTHSHOHUE

(ii) Row Transposition Ciphers-A more complex scheme is to write the message in a rectangle, row by row, and read the message off, column by column, but permute the order of the columns. The order of columns then becomes the key of the algorithm.
e.g., plaintext = meet at the school house

Key =	4	3	1	2	5	6	7
PT =	m	e	e	t	a	t	t
	h	e	s	c	h	o	o
	l	h	o	u	s	e	

CT = ESOTCUEEHMHLAHSTOETO

A pure transposition cipher is easily recognized because it has the same letter frequencies as the original plaintext. The transposition cipher can be made significantly more secure by performing more than one stage of transposition. The result is more complex permutation that is not easily reconstructed.

LECTURE - 09

Finite field and number theory

Group:

- A group G is a finite or infinite set of elements together with a binary operation (called the group operation) that together satisfy the four fundamental properties of closure, associativity, the identity property, and the inverse property. The operation with respect to which a group is defined is often called the "group operation," and a set is said to be a group "under" this operation. Elements A, B, C, ... with binary operation between A and B denoted AB form a group if
  - Closure: If A and B are two elements in G, then the product AB is also in G.
  - Associativity: The defined multiplication is associative, i.e., for all $A, B, C \in G$ $(AB)C = A(BC)$.
  - Identity: There is an identity element I such that $IA = AI = A$ for every element $A \in G$.
  - Inverse: There must be an inverse (reciprocal) of each element. Therefore, for each element A of G, the set contains an element $B = A^{-1}$ such that $AA^{-1} = A^{-1}A = I$.
- A group is a monoid each of whose elements is invertible.
- A group must contain at least one element, with the unique (up to isomorphism) single-element group known as the trivial group.
- The study of groups is known as group theory. If there are a finite number of elements, the group is called a finite group and the number of elements is called the group order of the group. A subset of a group that is closed under the group operation and the inverse operation is called a subgroup. Subgroups are also groups, and many commonly encountered groups are in fact special subgroups of some more general larger group.

Ring:

- In mathematics, a **ring** is one of the fundamental algebraic structures used in abstract algebra. It consists of a set equipped with two binary operations that generalize the arithmetic operations of addition and multiplication. Through this generalization, theorems from arithmetic are extended to non-numerical objects such as polynomials, series, matrices and functions.
- The most familiar example of a ring is the set of all integers, ”Z”, consisting of the numbers

… , −5, −4, −3, −2, −1, 0, 1, 2, 3, 4, 5, …

- A **ring** is a set R equipped with two binary operations + and · satisfying the following three sets of axioms, called the **ring axioms.**

1. R is an abelian group under addition, meaning that:
   - $(a + b) + c = a + (b + c)$ for all a, b, c in R (that is, + is associative).
   - $a + b = b + a$ for all a, b in R (that is, + is commutative).
   - There is an element 0 in R such that $a + 0 = a$ for all a in R (that is, 0 is the additive identity).
   - For each a in R there exists $-a$ in R such that $a + (-a) = 0$ (that is, $-a$ is the additive inverse of a).
2. R is a monoid under multiplication, meaning that:
   - $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for all a, b, c in R (that is, · is associative).
   - There is an element 1 in R such that $a \cdot 1 = a$ and $1 \cdot a = a$ for all a in R (that is, 1 is the multiplicative identity).[5]
3. Multiplication is distributive with respect to addition, meaning that:
   - $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ for all a, b, c in R (left distributivity).
   - $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$ for all a, b, c in R (right distributivity).

Modular Arithmetic:

In mathematics, **modular arithmetic** is a system of arithmetic for integers, where numbers "wrap around" when reaching a certain value, called the **modulus**.

Modular Arithmetic as Remainders:

- The easiest way to understand modular arithmetic is to think of it as finding the remainder of a number upon division by another number. For example, since both 15 and -9 leave the same remainder 3 when divided by 12, we say that: $15 \equiv -9 \pmod{12}$.
- This allows us to have a simple way of doing modular arithmetic: first perform the usual arithmetic, and then find the remainder. For example, to find $123+321 \pmod{11}$, we can take 123 + 321 = 444 and divide it by 11, which gives us $123+321 \equiv 4 \pmod{11}$.
- However, this could get messy when the numbers get larger. One approach that we could take is to first find the remainders of 123 and 321 when divided by 11 (the remainders are both 2), perform the usual arithmetic, and find the remainder again. In this example, since $123 \equiv 2 \pmod{11}$ and $321 \equiv 2 \pmod{11}$, we can conclude that

$$123+321 \equiv 2+2 \pmod{11} \equiv 4 \pmod{11}.$$

Properties:

The congruence relation satisfies all the conditions of an equivalence relation:

- Reflexivity: $a \equiv a \pmod{n}$
- Symmetry: $a \equiv b \pmod{n}$ if $b \equiv a \pmod{n}$ for all a, b, and n.
- Transitivity: If $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$

If $a_1 \equiv b_1 \pmod{n}$ and $a_2 \equiv b_2 \pmod{n}$, or if $a \equiv b \pmod{n}$, then:

- $a + k \equiv b + k \pmod{n}$ for any integer k (compatibility with translation)
- $k\,a \equiv k\,b \pmod{n}$ for any integer k (compatibility with scaling)
- $a_1 + a_2 \equiv b_1 + b_2 \pmod{n}$ (compatibility with addition)
- $a_1 - a_2 \equiv b_1 - b_2 \pmod{n}$ (compatibility with subtraction)

- $a_1\ a_2 \equiv b_1\ b_2 \pmod{n}$ (compatibility with multiplication
- $a^k \equiv b^k \pmod{n}$ for any non-negative integer k (compatibility with exponentiation)
- $p(a) \equiv p(b) \pmod{n}$, for any polynomial $p(x)$ with integer coefficients (compatibility with polynomial evaluation)
- If $a \equiv b \pmod{n}$, then it is false, in general, that $k^a \equiv k^b \pmod{n}$.
- If $c \equiv d \pmod{\varphi(n)}$, where φ is Euler's totient function,
  then $a^c \equiv a^d \pmod{n}$ provided a is coprime with n

For cancellation of common terms, we have the following rules:

- If $a + k \equiv b + k \pmod{n}$ for any integer k, then $a \equiv b \pmod{n}$
- If $k\ a \equiv k\ b \pmod{n}$ and k is coprime with n, then $a \equiv b \pmod{n}$

The modular multiplicative inverse is defined by the following rules:

- Existence: there exists an integer denoted a^{-1} such that $aa^{-1} \equiv 1 \pmod{n}$ if and only if a is coprime with n. This integer a^{-1} is called a *modular multiplicative inverse* of a modulo n.
- If $a \equiv b \pmod{n}$ and a^{-1} exists, then $a^{-1} \equiv b^{-1} \pmod{n}$ (compatibility with multiplicative inverse, and, if $a = b$, uniqueness modulo n)
- If $a\ x \equiv b \pmod{n}$ and a is coprime to n, the solution to this linear congruence is given by $x \equiv a^{-1}b \pmod{n}$

In particular, if p is a prime number, then a is coprime with p for every a such that $0 < a < p$; thus a multiplicative inverse exists for all a not congruent to zero modulo p.

Some of the more advanced properties of congruence relations are the following:

- Fermat's little theorem: If p is prime and does not divide a, then $a^{p-1} \equiv 1 \pmod{p}$.
- Euler's theorem: If a and n are coprime, then $a^{\varphi(n)} \equiv 1 \pmod{n}$, where φ is Euler's totient function
- A simple consequence of Fermat's little theorem is that if p is prime, then $a^{-1} \equiv a^{p-2} \pmod{p}$ is the multiplicative inverse of $0 < a < p$. More generally, from Euler's theorem, if a and n are coprime, then $a^{-1} \equiv a^{\varphi(n)-1} \pmod{n}$.
- Another simple consequence is that if $a \equiv b \pmod{\varphi(n)}$, where φ is Euler's totient function, then $k^a \equiv k^b \pmod{n}$ provided k is coprime with n.
- Wilson's theorem: p is prime if and only if $(p-1)! \equiv -1 \pmod{p}$.
- Chinese remainder theorem: For any a, b and coprime m, n, there exists a unique $x \pmod{mn}$ such that $x \equiv a \pmod{m}$ and $x \equiv b \pmod{n}$.
  - In fact, $x \equiv b\ \ m_n^{-1}\ m + a\ \ n_m^{-1}\ n \pmod{mn}$ where m_n^{-1} is the inverse of m modulo n and n_m^{-1} is the inverse of n modulo m.
- Primitive root modulo n: A number g is a primitive root modulo n if, for every integer a coprime to n, there is an integer k such that $g^k \equiv a \pmod{n}$. A primitive root modulo n exists if and only if n is equal to 2, 4, p^k or $2p^k$, where p is an odd prime number and k is a positive integer. If a primitive root modulo n exists, then there are exactly $\varphi(\varphi(n))$ such primitive roots, where φ is the Euler's totient function.
- Quadratic residue: An integer a is a quadratic residue modulo n, if there exists an integer x such that $x^2 \equiv a \pmod{n}$. Euler's criterion asserts that, if p is an odd prime, and a is not a multiple of p, then a is a quadratic residue modulo p if and only if

$$a^{(p-1)/2} \equiv 1 \pmod{p}.$$

LECTURE-10

Finite Fields:

- Finite Fields, also known as Galois Fields, are very much useful in cryptography applications.
- A field can be defined as a set of numbers that we can add, subtract, multiply and divide together and end up with a result that exists in our set of numbers. This is particularly useful for crypto as we can deal with a limited set of extremely large numbers.
- To have a finite field we need the following properties (the dot symbol · denotes the remainder after multiplying/adding two elements):
  - **Closed:** any operation performed with elements from the set returns an element contained in the original set.
  - **Associative:** it states that: **(a· b) ·c = a · (b ·c)**
  - **Identity:** there exists a neutral element (usually 1) such that: **a · 1 = a**
  - **Inverse:** within the set there's another element such that **a · (a)^-1= 1**
  - **Commutative:** it states that: **a · b = b · a**
- The most important property of a finite field is that it has **p ^ m** elements where p is a prime number and m is any number.
- For example: A finite field with 11 elements can be defined as GF (11^1). A finite field with 256 elements would be written as GF (2^8). But we can't have finite field with 12 elements, although we can write 12 as 2^2 * 3 it breaks the convention of p ^ m.
- With our notation of GF(p^m):
  - If m = 1 then we get prime fields
  - If m > 1 then we get extension fields.

A) Prime Field Arithmetic

- The notation GF(p) means we have a finite field with the integers {0, … , p-1}. Suppose we have GF (5), our initial set will be {0, 1, 2, 3, 4}. Any operations we do on 5 should return 0, 1, 2, 3 or 4 (closure property).

Addition:
(3 + 4) mod 5 = 2
(1 + 4) mod 5 = 0
(1 + 2) mod 5 = 3
Subtraction:
(4 - 0) mod 5 = 4
(4 - 2) mod 5 = 2
(3 - 0) mod 5 = 1

Multiplication:
(0 * 4) mod 5 = 0
(2 * 4) mod 5 = 3
(3 * 4) mod 5 = 2
Division/Inversion:
(4 * 4) mod 5 = 1
(3 * 2) mod 5 = 1
(2 * 3) mod 5 = 1
(1 * 1) mod 5 = 1

B) Extension Fields

- Unlike finite fields, whose elements are integers, extension fields' elements are polynomials. Extension fields = GF(2^m) where m > 1
- These polynomials take the form of:

$$a_{m-1}X^{m-1} + \ldots + a_1X^1 + a_0$$

- To make it less cryptic, let's use the example of GF(2^3) which will result in the equation form:

$$A(x) = a_2X^2 + a_1X^1 + a_0$$

- GF (2^3) = GF (8) which means there'll be a total of 8 elements in this set.
- If we write the elements, they'll have the following values for (a2, a1, a0) where a1, a2 or a0 can only ever be 0 or 1.
  - (a2, a1, a0) (0, 0, 0) = 0
  - (0, 0, 1) = 1
  - (0, 1, 0) = x
  - (1, 0, 0) = x^2
  - (0, 1, 1) = x+1
  - (1, 1, 0) = x^2+x
  - (1, 0, 1) = x^2+1
  - (1, 1, 1) = x^2+x+1
- Putting it all together, GF (2^3) = {0, 1, x, x^2, x+1, x^2+x, x^2+1, x^2+x+1}
- We can perform addition, subtraction, multiplication, and division and return an element in our set.
- For example if we add: x^2+1 andx^2+x+1 the resultant value will be: (1+1) x^2+x+(1+1). Since 1 + 1 mod 2 = 0, equation simplifies to x which is contained in the original set.
- Let's try another example
- A · B = (x^2+1)(x^2+x+1) = x^4+x^3+x^2+x^2+1 = x^4+x^3+(1+1)x^2+1= x^4+x^3+1
- Here, x^4+x^3+1 doesn't exist in our finite field. These are called irreducible polynomials which are defined as polynomials which can't be broken down into smaller polynomials (from the power of original polynomial).

Euclid's algorithms:

GCD of two numbers is the largest number that divides both of them. A simple way to find GCD is to factorize both numbers and multiply common factors.

```
36  = 2 x 2 x 3 x 3
60  = 2 x 2 x 3 x 5

GCD = Multiplication of common factors
    = 2 x 2 x 3
    = 12
```

Basic Euclidean Algorithm for GCD

The algorithm is based on below facts.

- If we subtract smaller number from larger (we reduce larger number), GCD doesn't change. So if we keep subtracting repeatedly the larger of two, then we end up with GCD.
- Now instead of subtraction, if we divide smaller number, the algorithm stops when we find remainder 0.

The Algorithm

The Euclidean Algorithm for finding GCD(A,B) is as follows:

- If A = 0 then GCD(A,B)=B, since the GCD(0,B)=B, and we can stop.
- If B = 0 then GCD(A,B)=A, since the GCD(A,0)=A, and we can stop.
- Write A in quotient remainder form (A = B·Q + R)
- Find GCD(B,R) using the Euclidean Algorithm since GCD(A,B) = GCD(B,R)

Example:

Find the GCD of 270 and 192

- A=270, B=192
- A ≠0, B ≠0
- Use long division to find that 270/192 = 1 with a remainder of 78.
- We can write this as: 270 = 192 * 1 +78

Find GCD (192,78), since GCD (270, 192) =GCD (192,78)

- A=192, B=78
- A ≠0, B ≠0
- Use long division to find that 192/78 = 2 with a remainder of 36.
- We can write this as: 192 = 78 * 2 + 36

Find GCD (78,36), since GCD (192,78) =GCD (78,36)

- A=78, B=36
- A ≠0, B ≠0
- Use long division to find that 78/36 = 2 with a remainder of 6.
- We can write this as: 78 = 36 * 2 + 6

Find GCD (36,6), since GCD (78,36) =GCD (36,6)

- A=36, B=6
- A ≠0, B ≠0
- Use long division to find that 36/6 = 6 with a remainder of 0.
- We can write this as: 36 = 6 * 6 + 0

Find GCD (6,0), since GCD (36,6) =GCD (6,0)

- A=6, B=0
- A ≠0, B =0, GCD(6,0)=6

Hence,

- GCD (270,192) = GCD (192,78) = GCD (78,36) = GCD (36,6) = GCD (6,0) = 6
- GCD (270,192) = 6

LECTURE-12

prime numbers:

- prime numbers only have divisors of 1 and self
- They cannot be written as a product of other numbers

NOTE: 1 is prime, but is generally not of interest eg. 2,3,5,7 are prime, 4,6,8,9,10 are not prime numbers are central to number theory

- List of prime number less than 200 is:

2	3	5	7	11	13	17	19	23	29	31	37	41	43	47	53	59	61	67	
71	73	79	83	89	97	101	103	107	109	113	127	131	137	139					
149	151	157	163	167	173	179	181	191	193	197	199								

Prime Factorization

- To factor a number n is to write it as a product of other numbers: n = a x b x c
- The prime factorization of a number n is when its written as a product of primes
- Example: 91 = 7 x 13 ; 3600 = 24 x 32 x 52

Relatively Prime Numbers & GCD

- Two numbers a, b are relatively prime (coprime) if they have no common divisors apart from 1
- For example, 8 and 15 are relatively prime since factors of 8 are 1,2,4,8 and of 15 are 1,3,5,15 and 1 is the only common factor
- conversely, we can determine the greatest common divisor by comparing their prime factorizations and using least powers
- For example: 300 = 21x31x52 ; 18 = 21x32 hence GCD(18,300) = 21x31x50 = 6

Euler Totient Function ø(n):

- When performing arithmetic modulo n complete set of residues is: 0 …. n-1
- The reduced set of residues is those numbers (residues) which are relatively prime to n
- For example: for n = 10,
  - complete set of residues is {0,1,2,3,4,5,6,7,8,9}
  - reduced set of residues is {1,3,7,9}
- Number of elements in reduced set of residues is called the Euler Totient Function ø(n)
- To compute ø(n) need to count number of residues to be excluded
- In general, it needs prime factorization, but
  - for p (p prime): ø(p)=p-1
  - for p.q (p,q prime): ø(p.q)=(p-1)x(q-1)
- For example:

  ø(37) = 36
  ø(21) = (3–1)x(7–1) = 2x6 = 12

Euler's Theorem:

- It states that a^ø(n) = 1 (mod n) for any a, n where GCD(a, n) = 1
- For example:
  - Let a = 3 ; n = 10 ; ø(10) = 4; hence 3^4 = 81 = 1 mod 10
  - Again if a = 2 ; n = 11 ; ø(11) = 10; hence 210 = 1024 = 1 mod 11
- Also have: a ^ (ø(n)+1) = a (mod n)

PRIMALITY TEST:

A **primality test** is an algorithm for determining whether an input number is prime. Among other fields of mathematics, it is used for cryptography. Unlike integer factorization, primality tests do not generally give prime factors, only stating whether the input number is prime or not. Factorization is thought to be a computationally difficult problem, whereas primality testing is comparatively easy (its running time is polynomial in the size of the input). Some primality tests *prove* that a number is prime, while others like Miller–Rabin prove that a number is composite. Therefore, the latter might more accurately be called *compositeness tests* instead of primality tests.

Fermat Method

- Given a number n, check if it is prime or not.
- Fermat's method is a probabilistic method and is based on below Fermat's Little Theorem.

Fermat's Little Theorem:

- If n is a prime number, then for every a, 1 < a < n-1,

  $\mathbf{a^{n-1} \equiv 1 \pmod{n}}$ **or** $\mathbf{a^{n-1} \% n = 1}$

Example:

- Since 5 is prime, $2^4 \equiv 1 \pmod 5$ [or $2^4\%5 = 1$], $3^4 \equiv 1 \pmod 5$ and $4^4 \equiv 1 \pmod 5$
- Since 7 is prime, $2^6 \equiv 1 \pmod 7$, $3^6 \equiv 1 \pmod 7$, $4^6 \equiv 1 \pmod 7$, $5^6 \equiv 1 \pmod 7$ and $6^6 \equiv 1 \pmod 7$

Proofs

- If a given number is prime, then this method always returns true. If given number is composite (or non-prime), then it may return true or false, but the probability of producing incorrect result for composite is low and can be reduced by doing more iterations.

Algorithm:

// Higher value of k indicates probability of correct
// results for composite inputs become higher. For prime
// inputs, result is always correct
1) Repeat following k times:
   a) Pick a randomly in the range [2, n - 2]
   b) If gcd (a, n) $\neq$ 1, then return false
   c) If a^{n-1} != 1 (mod n), then return false
2) Return true [probably prime].

Chinese remainder theorem

- In number theory, the **Chinese remainder theorem** states that if one knows the remainders of the Euclidean division of an integer *n* by several integers, then one can determine uniquely the remainder of the division of *n* by the product of these integers, under the condition that the divisors are pairwise coprime.
- The earliest known statement of the theorem is by the Chinese mathematician Sun-tzu in the *Sun-tzu Suan-ching* in the 3rd century AD.
- The Chinese remainder theorem is widely used for computing with large integers, as it allows replacing a computation for which one knows a bound on the size of the result by several similar computations on small integers.

Statement

- Let n_1, ..., n_k be integers greater than 1, which are often called moduli or divisors. Let us denote by N the product of the n_i.
- The Chinese remainder theorem asserts that if the n_i are pairwise coprime, and if a_1, ..., a_k are integers such that $0 \leq a_i < n_i$ for every i, then there is one and only one integer x, such that $0 \leq x < N$ and the remainder of the Euclidean division of x by n_i is a_i for every i.
- This may be restated as follows in term of congruences: If the n_i are pairwise coprime, and if a_1, ..., a_k are any integers, then there exist integers x such that

$$\begin{aligned} x &\equiv a_1 \pmod{n_1} \\ &\vdots \\ x &\equiv a_k \pmod{n_k}, \end{aligned}$$

- and any two solutions, say x_1 and x_2, are congruent modulo N, that is, $x_1 \equiv x_2 \pmod N$.
- In abstract algebra, the theorem is often restated as: if the n_i are pairwise coprime, the map defines a ring isomorphism

$$x \bmod N \;\mapsto\; (x \bmod n_1, \ldots, x \bmod n_k)$$

Implementation:

We are given two arrays num[0..k-1] and rem[0..k-1].
In num[0..k-1], every pair is coprime (gcd for every pair is 1).
We need to find minimum positive number x such that:

x % num[0] = rem[0],
x % num[1] = rem[1],
.....
x % num[k-1] = rem[k-1]

Basically, we are given k numbers which are pairwise coprime, and given remainders of these numbers when an unknown number x is divided by them. We need to find the minimum possible value of x that produces given remainders.

Example 1

Input: num[] = {5, 7}, rem[] = {1, 3}
Output: 31
Explanation:
31 is the smallest number such that:
(1) When we divide it by 5, we get remainder 1.
(2) When we divide it by 7, we get remainder 3.

Example 2

Input: num[] = {3, 4, 5}, rem[] = {2, 3, 1}
Output: 11
Explanation:
11 is the smallest number such that:
(1) When we divide it by 3, we get remainder 2.
(2) When we divide it by 4, we get remainder 3.
(3) When we divide it by 5, we get remainder 1.

Discrete Logarithm

- In the mathematics of the real numbers, the logarithm $\log_b a$ is a number x such that $b^x = a$, for given numbers a and b. Analogously, in any group G, powers b^k can be defined for all integers k, and the **discrete logarithm** $\log_b a$ is an integer k such that $b^k = a$. In number theory, the more commonly used term is **index**: we can write $x = \operatorname{ind}_r a \pmod m$ (read the index of a to the base r modulo m) for $r^x \equiv a \pmod m$ if r is a primitive root of m and $\gcd(a,m) = 1$.

Definition

- Let G be any group. Denote its group operation by multiplication and its identity element by 1. Let b be any element of G. For any positive integer k, the expression b^k denotes the product of b with itself k times:

$$b^k = \underbrace{b \cdot b \cdots b}_{k \text{ factors}}.$$

- Similarly, let b^{-k} denote the product of b^{-1} with itself k times. For $k = 0$ and $b \neq 0$, the kth power is the identity: $b^0 = 1$.
- Let a also be an element of G. An integer k that solves the equation $b^k = a$ is termed a **discrete logarithm** (or simply **logarithm**) of a to the base b. One writes $k = \log_b a$.

Examples: Powers of 10

- The powers of 10 form an infinite subset $G = \{\ldots, 0.001, 0.01, 0.1, 1, 10, 100, 1000, \ldots\}$ of the rational numbers. This set G is a cyclic group under multiplication, and 10 is a generator.
- For any element a of the group, we can compute $\log_{10} a$. For example, $\log_{10} 10000 = 4$, and $\log_{10} 0.001 = -3$. These are instances of the discrete logarithm problem.
- Other base-10 logarithms in the real numbers are not instances of the discrete logarithm problem, because they involve non-integer exponents. For example, the equation $\log_{10} 53 = 1.724276\ldots$ means that $10^{1.724276\ldots} = 53$. While integer exponents can be defined in any group using products and inverses, arbitrary real exponents in the real numbers require other concepts such as the exponential function.

Powers of a fixed real number

- A similar example holds for any non-zero real number b. The powers form a multiplicative subgroup $G = \{\ldots, b^{-3}, b^{-2}, b^{-1}, 1, b^1, b^2, b^3, \ldots\}$ of the non-zero real numbers. For any element a of G, one can compute $\log_b a$.
- Exponentiation is relatively easy, where finding discrete logarithms is generally a hard problem. So, it is preferred for cryptography.