Unit 1: Introduction and Classical Ciphers

- The word Cryptography comes from Greek words (kryptós = Secret,hidden ; graphein= Writing)
- Cryptography is the practice and study of techniques for secure communication in the presence of third parties called adversaries – Wikipedia
- Cryptography is a method of protecting information and communications through the use of code so that only those for whom the information is intended can read and process it.
- This field is very much associated with mathematics and computer science with application in many fields like computer security, electronic commerce, telecommunication, etc.
- In the ancient days, cryptography was mostly referred to as ***encryption*** – the mechanism to convert the readable ***plaintext*** into unreadable (incomprehensible) text i.e. ***ciphertext***, and ***decryption*** – the opposite process of encryption i.e. conversion of ciphertext back to the plaintext.
- Though the consideration of cryptography was on message confidentiality (encryption) in the past, nowadays cryptography considers the study and practices of authentication, digital signatures, integrity checking, and key management, etc.

Security

- Security, in information technology (IT), is the defense of digital information and IT assets against internal and external, malicious and accidental threats. This defense includes detection, prevention and response to threats through the use of security policies, software tools and IT services.

Computer Security

- It is a process and the collection of measures and controls that ensures the Confidentiality, Integrity and Availability (CIA) of the assets in computer systems. Computer Security protects you from both software and hardware part of a computer systems from getting compromised and be exploited.

Information Security

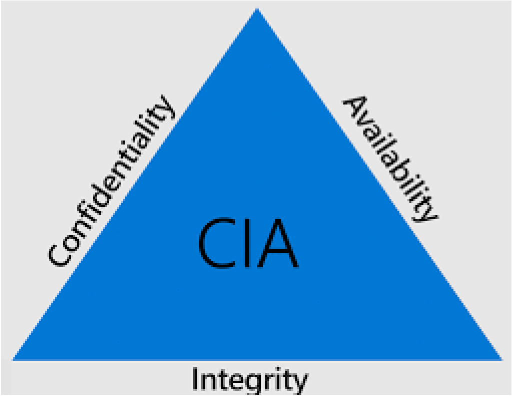
- Information security is primarily concerned with making sure that data in any form is kept secure in terms of preserving its confidentiality, integrity and availability.
- Information is a significant asset that can be stored in different ways such as digitally stored, printed, written on papers or in human memory. It can be communicated through different channels such as spoken languages, gestures or using digital channel such as email, SMS, social media, video, audio etc.
- Information security differs from cybersecurity such that information security aims to keep data in any form secure, whereas cybersecurity protects only digital data. Cybersecurity is the subset of information security.

Network Security:

- It is any activity designed to protect the usability and integrity of your network and data. It includes both hardware and software technologies.
- An effective network security manages access to the network. It targets a variety of threats and stop them from entering or spreading on your network.
- Network security, a subset of cybersecurity, aims to protect any data that is being sent through devices in your network to ensure that the information is not changed or intercepted.

CIA Triad:

- **Confidentiality**, **integrity** and **availability**, also known as the CIA triad, is a model designed to guide policies for information security within an organization.



- **Confidentiality**: Preserving authorized restrictions on information access and disclosure. This term covers two related concepts:
  - ✓ *Data confidentiality:* Assures that private or confidential information is not made available or disclosed to unauthorized individuals.
  - ✓ *Privacy:* Assures that individuals control or influence what information related to them may be collected and stored and by whom and to whom that information may be disclosed.

  Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information. A loss of confidentiality is the unauthorized disclosure of information.

  Confidentiality is the concealment of information or resources. Cryptography can be the better choice for maintaining the privacy of information, which traditionally is used to protect the secret messages. Similarly, privacy of resources, i.e. resource hiding can be maintained by using proper firewalls. Confidentiality is sometimes called **secrecy** or **privacy**.

- **Integrity:** Guarding against improper information modification or destruction. This term covers two related concepts:
  - ✓ *Data integrity:* Assures that information (both stored and in transmitted packets) and programs are changed only in a specified and authorized manner.
  - ✓ *System integrity:* Assures that a system performs its intended function in an unimpaired manner, free from deliberate or inadvertent unauthorized manipulation of the system.

  Guarding against improper information modification or destruction, including ensuring information nonrepudiation and authenticity. A loss of integrity is the unauthorized modification or destruction of information

Integrity ensures the correctness as well as trustworthiness of data or resources. For example, if we say that we have preserved the integrity of an item, we may mean that the item is: precise, accurate, unmodified, modified only in acceptable ways, modified only by authorized people, modified only by authorized processes, consistent, meaningful and usable

- **Availability:** Assures that systems work promptly and service is not denied to authorize users.

  Ensuring timely and reliable access to and use of information. A loss of availability is the disruption of access to or use of information or an information system.

  Availability refers to the ability to use the information or resource desired. An unavailable system is as bad as no system at all. An object or service is thought to be available if;

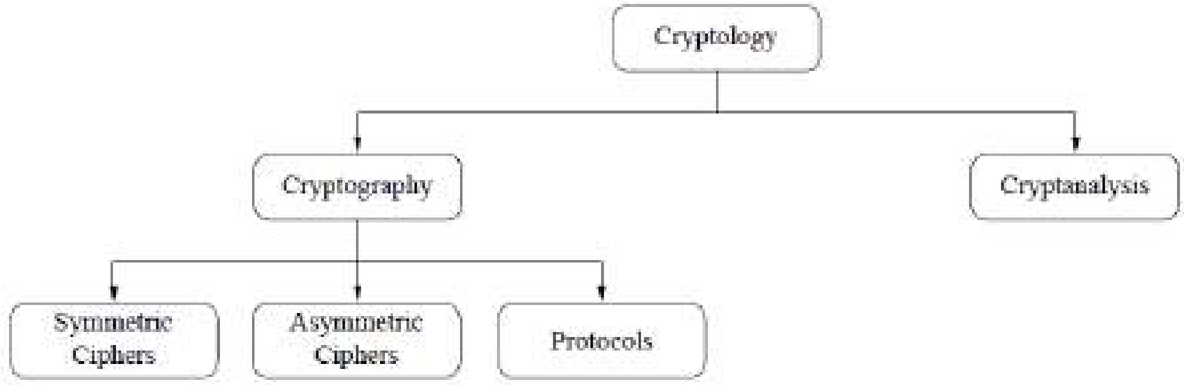
  - ✓ It is present in a usable form.
  - ✓ It has capacity enough to meet the service's needs.
  - ✓ It is making clear progress, and, if in wait mode, it has a bounded waiting time.
  - ✓ The service is completed in an acceptable period of time.

  Availability is usually defined in terms of "quality of service," in which authorized users are expected to receive a specific level of service. The aspect of availability that is relevant to security is that someone may intentionally arrange to deny access to data or to service by making it unavailable

Although the use of the CIA triad to define security objectives is well established, some in the security field feel that additional concepts are needed to present a complete picture. Some of them are:

- **Authenticity:** The property of being genuine and being able to be verified and trusted; confidence in the validity of a transmission, a message, or message originator. This means verifying that users are who they say they are and that each input arriving at the system came from a trusted source.
- **Accountability:** It means that every individual who works with an information system should have specific responsibilities for information assurance.
- **Access Control:** The prevention of unauthorized use of a resource.
- **Nonrepudiation:** Provides protection against denial by one of the entities involved in a communication of having participated in all or part of the communication

Overview of Cryptology



Cryptology: Study of techniques for ensuring the secrecy and\or authenticity of information

Cryptography: is the science of secret writing with the goal of hiding the meaning of a message.

Cryptanalysis: is the breaking of codes. Cryptanalysis encompasses all of the techniques to recover the plaintext and/or key from the ciphertext

The combined study of cryptography and cryptanalysis is known as ***cryptology***. Though most of the time we use cryptography and cryptology in the same way.

What are the meanings of the terms: Encryption, Decryption, Key and Cipher?

Encryption is the process of encoding a message so that its meaning is not obvious i.e. converting information from one form to some other unreadable form using some algorithm called ***cipher*** with the help of secret message called ***key***. The converting text is called is ***plaintext*** and the converted text is called ***ciphertext***

Decryption is the reverse process, transforming an encrypted message back into its normal, original form. In decryption process also the use of key is important.
Alternatively, the terms ***encode*** and ***decode*** or ***encipher*** and ***decipher*** are used instead of ***encrypt*** and ***decrypt***. That is, we say that we encode, encrypt, or encipher the original message to hide its meaning. Then, we decode, decrypt, or decipher it to reveal the original message.

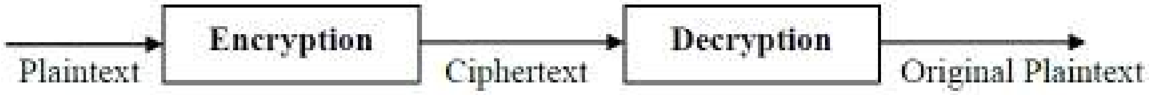


Fig: Encryption-Decryption

Key

A ***key*** is a parameter or a piece of information used to determine the output of cryptographic algorithm. While doing the encryption, key determines the transformation of plaintext to the cipher text and vice versa. Keys are also used in other cryptographic processes like message authentication codes and digital signatures. Most of the cryptographic systems depend upon the key and thus the secrecy of the key is very important and is one of the difficult problems in practice. Another important issue for the key is its length. Since key is the sole entity that defines the strength of the security (normally algorithm used is public) we need to select the key in a way such that attacker should take long enough to try all possibilities. To prevent the key from being guessed the choice of the key must be random.

Cipher

A ***cipher*** is an algorithm for performing encryption and decryption. The operation of cipher depends upon the special information called key. Without knowledge of the key, it should be difficult, if not nearly impossible, to decrypt the resulting cipher into readable plaintext. There are many types of encryption techniques that have advanced from history, however the distinction of encryption technique can be broadly categorized in terms of number of key used and way of converting plaintext to the ciphertext.

Cryptosystem

A cryptosystem is a five-tuple (P, C, K, E, D), where the following conditions are satisfied:

1. P is a finite set of possible plaintexts;
2. C is a finite set of possible ciphertexts;
3. K, the keyspace, is a finite set of possible keys;

4. For each k ∈ K, there is an encryption rule e_K ∈ E and a corresponding decryption rule d_K ∈ D. Each $e_K : P \rightarrow C$ and $d_K : C \rightarrow P$ are functions such that $d_K(e_K(x)) = x$ for every plaintext element x ∈ P.

In cryptography, a cryptosystem is a structure or scheme consisting of a set of algorithms needed to implement a particular security service, most commonly for achieving confidentiality.

- ✓ Typically, a cryptosystem consists of three algorithms: one for key generation, one for encryption, and one for decryption. The term cipher is often used to refer to a pair of algorithms, one for encryption and one for decryption.

Cryptanalysis

- ✓ Cryptanalysis is the science and sometimes art of breaking cryptosystems
- ✓ Cryptanalysis is the study of ciphertext, ciphers and cryptosystems with the aim of understanding how they work and finding techniques for defeating or weakening them.
- ✓ Cryptanalysis is used to breach cryptographic security systems and gain access to the contents of encrypted messages, even if the cryptographic key is unknown.
- ✓ Cryptanalysis is of central importance for modern cryptosystems: without people who try to break our crypto methods, we will never know whether they are really secure or not

Brute Force attack and Cryptoanalytic Attacks (Various attack models on Cryptosystem)

Typically, the objective of attacking an encryption system is to recover the key in use rather than simply to recover the plaintext of a single ciphertext. There are two general approaches to attacking a conventional encryption scheme:

- ⇒ **Brute-force attack**: The attacker tries every possible key on a piece of ciphertext until an intelligible translation into plaintext is obtained. On average, half of all possible keys must be tried to achieve success
- ⇒ **Cryptanalysis:** Cryptanalytic attacks rely on the nature of the algorithm plus perhaps some knowledge of the general characteristics of the plaintext

or even some sample plaintext–ciphertext pairs. This type of attack exploits the characteristics of the algorithm to attempt to deduce a specific plaintext or to deduce the key being used.

The attack model specifies the information available to the adversary when he mounts his attack. The most common types of attack models (cryptoanalytic attacks) are enumerated as follows:

Ciphertext-Only Attack

The opponent possesses a string of ciphertext, y.

Known Plaintext Attack

The opponent possesses a string of plaintext, x, and the corresponding ciphertext, y.

Chosen Plaintext Attack

The opponent has obtained temporary access to the encryption machinery. Hence he can choose a plaintext string, x, and construct the corresponding ciphertext string, y.

Chosen Ciphertext Attack

The opponent has obtained temporary access to the decryption machinery. Hence he can choose a ciphertext string, y, and construct the corresponding plaintext string, x.

Security Threats and Attacks

A threat to a computing system is a set of circumstances that has the potential to cause loss or harm. It is a potential violation of security, means that it is a possible danger that might exploit vulnerability. Attack is an assault on system security that derives from an intelligent threat, i.e. attack is an intelligent act that is an intentional attempt to evade security services and violate the security policy of a system

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.
- Threat can be anything that can take advantage of a vulnerability to breach security and negatively alter, erase, harm object or objects of interest. In simple words, a threat is a potential violation of security which might or might not occur.

.

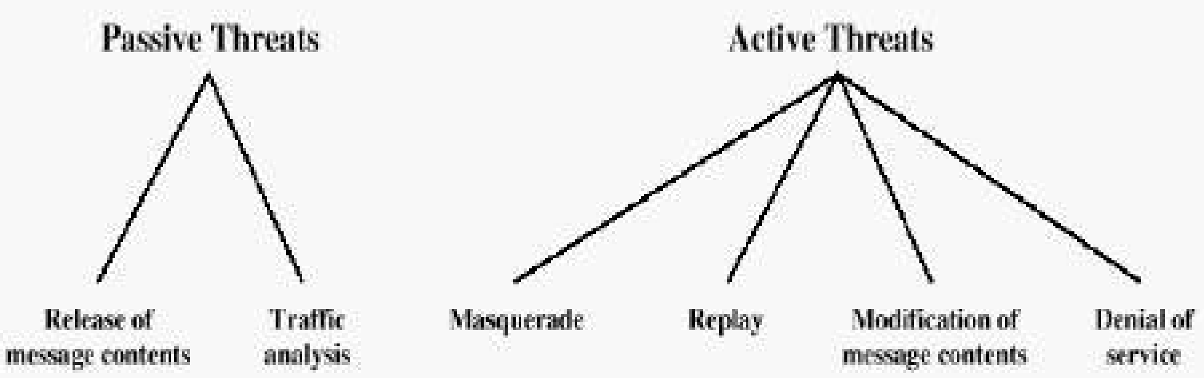
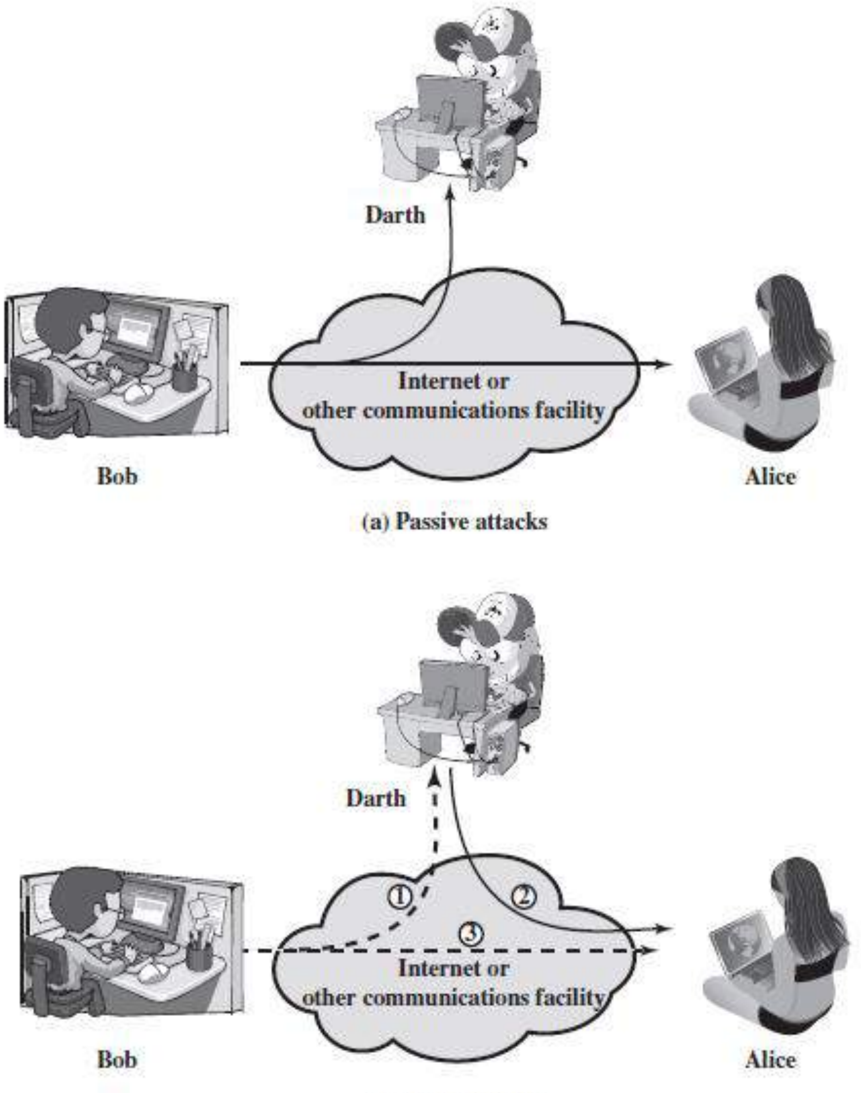


Figure 1.2 Active and Passive Security Threats

Security 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.
- Any action that compromises the security of information.
- Types of Security attacks :
    - **Passive attacks :** A passive attack attempts to learn or make use of information from the system but does not affect system resources
    - **Active attacks:** An active attack attempts to alter system resources or affect their operation



(a) Passive attacks

(b) Active attacks

Passive attacks

- A Passive attack attempts to learn or make use of information from the system but does not affect system resources.
- Passive Attacks are in the nature of eavesdropping on (act of secretly or stealthily listening to the private conversation or communications of others without their consent) or monitoring of transmission. The goal of the opponent is to obtain information is being transmitted
- Two types of passive attacks are: **the release of message contents** and **traffic analysis.**

The release of message contents: It is easily understood. A telephone conversation, an electronic mail message, and a transferred file may contain sensitive or confidential information. We would like to prevent an opponent from learning the contents of these transmissions.

Traffic analysis: A second type of passive attack, traffic analysis, is subtler. Suppose that we had a way of masking the contents of messages or other information traffic so that opponents, even if they captured the message, could not extract the information from the message. The common technique for masking contents is encryption. If we had encryption protection in place, an opponent might still be able to observe the pattern of these messages. The opponent could determine the location and identity of communicating hosts and could observe the frequency and length of messages being exchanged. This information might be useful in guessing the nature of the communication that was taking place

- Passive attacks are very difficult to detect, because they do not involve any alteration of the data. Typically, the message traffic is sent and received in an apparently normal fashion, and neither the sender nor receiver is aware that a third party has read the messages or observed the traffic pattern.
- However, it is feasible to prevent the success of these attacks, usually by means of encryption. Thus, the emphasis in dealing with passive attacks is on prevention rather than detection.

Active Attacks

- An Active attack attempts to alter system resources or effect their operations. Active attack involve some modification of the data stream or creation of false statement.
- Active attacks can be subdivided into four categories: **masquerade, replay, modification of messages, and denial of service**

Masquerade: Masquerade attack takes place when one entity pretends to be different entity. A masquerade attack usually includes one of the other forms of active attack

Replay: Replay involves the passive capture of a data unit and its subsequent retransmission to produce an unauthorized effect.

Modification of messages: Modification of messages simply means that some portion of a legitimate message is altered, or that messages are delayed or reordered, to produce an unauthorized effect

Denial of service (DOS): The denial of service prevents or inhibits the normal use or management of communications facilities. This attack may have a specific target; for example, an entity may suppress all messages directed to a particular destination (e.g., the security audit service). Another form of service denial is the disruption of an entire network, either by disabling the network or by overloading it with messages so as to degrade performance.

- Active attacks present the opposite characteristics of passive attacks. Whereas passive attacks are difficult to detect, measures are available to prevent their success. On the other hand, it is quite difficult to prevent active attacks absolutely because of the wide variety of potential physical, software, and network vulnerabilities. Instead, the goal is to detect active attacks and to recover from any disruption or delays caused by them. If the detection has a deterrent effect, it may also contribute to prevention

<u>Some forms of attacks</u>

Passive attack: unauthorized reading of a message or a file.

Active attack: modification of messages or files, and denial of service.

Interruption: This is an attack on availability

Interception: This is an attack on confidentiality

Modification: This is an attack on integrity

Fabrication: This is an attack on authenticity

Snooping: It is the unauthorized interception of information and disclosure. Passively listening (or reading) to communications or browsing through files or system information.

Modification or Alteration: Unauthorized change of information. If modified data controls the operation of the system, threats of failure may arise.

Masquerading or Spoofing: One entity pretends to be a different entity.

Repudiation of origin: A false denial that an entity sent or created something.

Denial of receipt: A false denial that an entity received some information or message.

Delay: Usually delivery of a message or service requires some time t. If an attacker can force the deliver to take more than time t, the attacker has successfully delayed delivery

Security Service

A service that enhances the security of data processing systems and information transfers. A security service makes use of one or more security mechanisms.

- ✓ enhance security of data processing systems and information transfers of an organization
- ✓ intended to counter security attacks
- ✓ using one or more security mechanisms

- Security services implement security policies and are implemented by security mechanisms

Definitions of Security Service According to the Standards X.800 and RFC 4949:

X.800 :

"a service that is provided by a protocol layer of communicating open systems and that ensures adequate security of the systems or of data transfers"

RFC 4949:

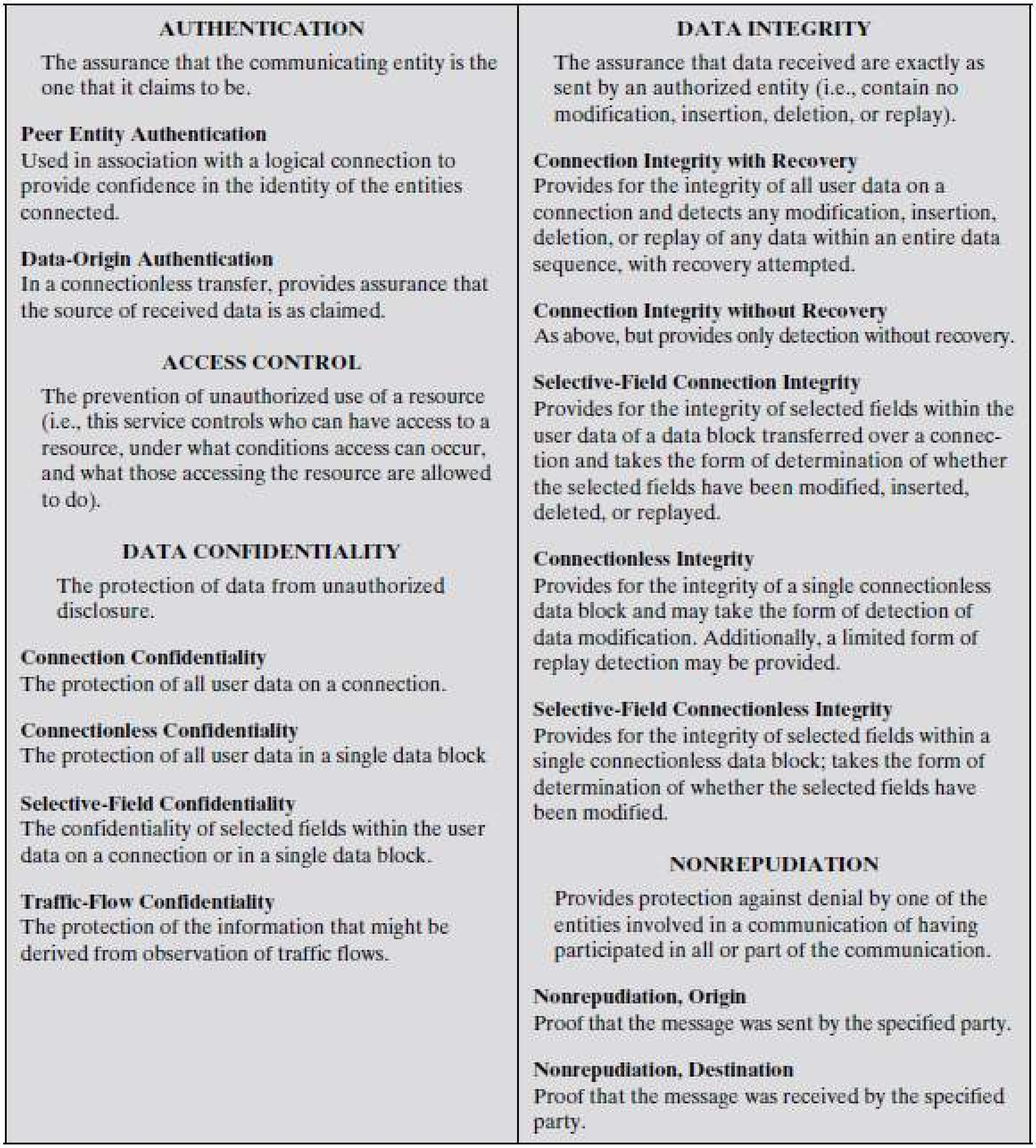
"*a processing or communication service that is provided by a system to give a specific kind of protection to system resources*"

X.800 divides **Security Services** into following five categories:

1. **Authentication** - The assurance that the communicating entity is the one that it claims to be
2. **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)

3. **Data Confidentiality** – The protection of data from unauthorized disclosure.
4. **Data Integrity** - The assurance that data received are exactly as sent by an authorized entity (i.e., contain no modification, insertion, deletion, or replay).
5. **Non-Repudiation** - Provides protection against denial by one of the entities involved in a communication of having participated in all or part of the communication.

Following table includes all five categories and their sub categories



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.

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).

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

Selective-Field Confidentiality
The confidentiality of selected fields within the user data on a connection or in a single data block.

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

DATA INTEGRITY

The assurance that data received are exactly as sent by an authorized entity (i.e., contain no modification, insertion, deletion, or replay).

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
As above, 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.

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 received by the specified party.

Availability Service

Both X.800 and RFC 4949 define availability to be the property of a system or a system resource being accessible and usable upon demand by an authorized system entity, according to performance specifications for the system (i.e., a system is available if it provides services according to the system design whenever users request them)

Security Mechanism

- ✓ A mechanism that is designed to detect, prevent, or recover from a security attack.
- ✓ The specific means of implementing one or more security services
- ✓ No single mechanism that will support all services required
- ✓ However one particular element underlies many of the security mechanisms in use: **cryptographic techniques**

X.800 defines following Specific and Pervasive Security Mechanisms:

- **specific security mechanisms:**
  - encipherment, digital signatures, access controls, data integrity, authentication exchange, traffic padding, routing control, notarization
- **pervasive security mechanisms:**
  - trusted functionality, security labels, event detection, security audit trails, security recovery

SPECIFIC SECURITY MECHANISMS	PERVASIVE SECURITY MECHANISMS
May be incorporated into the appropriate protocol layer in order to provide some of the OSI security services. **Encipherment** The use of mathematical algorithms to transform data into a form that is not readily intelligible. The transformation and subsequent recovery of the data depend on an algorithm and zero or more encryption keys. **Digital Signature** Data appended to, or a cryptographic transformation of, a data unit that allows a recipient of the data unit to prove the source and integrity of the data unit and protect against forgery (e.g., by the recipient). **Access Control** A variety of mechanisms that inforce access rights to resources. **Data Integrity** A variety of mechanisms used to assure the integrity of a data unit or stream of data units **Authentication Exchange** A mechanism intended to ensure the identity of an entity by means of information exchange. **Traffic Padding** The insertion of bits into gaps in a data stream to frustrate traffic analysis attempts. **Routing Control** Enables selection of particular physically secure routes for certain data and allows routing changes, especially when a breach of security is suspected. **Notarization** The use of a trusted third party to assure certain properties of a data exchange.	Mechanisms that are not specific to any particular OSI security service or protocol layer. **Trusted Functionality** That which is perceived to be correct with respect to some criteria (e.g., as established by a security Policy). **Security Label** The marking bound to a resource (which may be a data unit) that names or designates the security attributes of that resource. **Event Detection** Detection of security-relevant events. **Security Audit Trail** Data collected and potentially used to facilitate a security audit, which is an independent review and examination of system records and activities. **Security Recovery** Deals with requests from mechanisms, such as event handling and management functions, and takes recovery actions.

Classical Cryptosystems

Classical cryptosystems (also called single-key or symmetric cryptosystems) are cryptosystems that use the same key for encipherment and decipherment.

Symmetric encryption, also referred to as conventional encryption or single-key encryption, was the only type of encryption in use prior to the development of public-key encryption in the 1970s. It remains by far the most widely used of the two types of encryption. *All traditional schemes are **symmetric / single key / private-key** encryption algorithms, with a **single key**, used for both encryption and decryption. Since both sender and receiver are equivalent, either can encrypt or decrypt messages using that common key.*

- An original message is known as the **plaintext**, while the coded message is called the **ciphertext**.
- The process of converting from plaintext to ciphertext is known as **enciphering** or **encryption**; restoring the plaintext from the ciphertext is **deciphering** or **decryption.**
- The many schemes used for encryption constitute the area of study known as **cryptography**. Such a scheme is known as a **cryptographic system** or a **cipher**.
- Techniques used for deciphering a message without any knowledge of the enciphering details fall into the area of **cryptanalysis**. Cryptanalysis is what the layperson calls “breaking the code.”
- The areas of cryptography and cryptanalysis together are called **cryptology**

Symmetric Cipher Model

A symmetric encryption scheme has five ingredients as in following figures

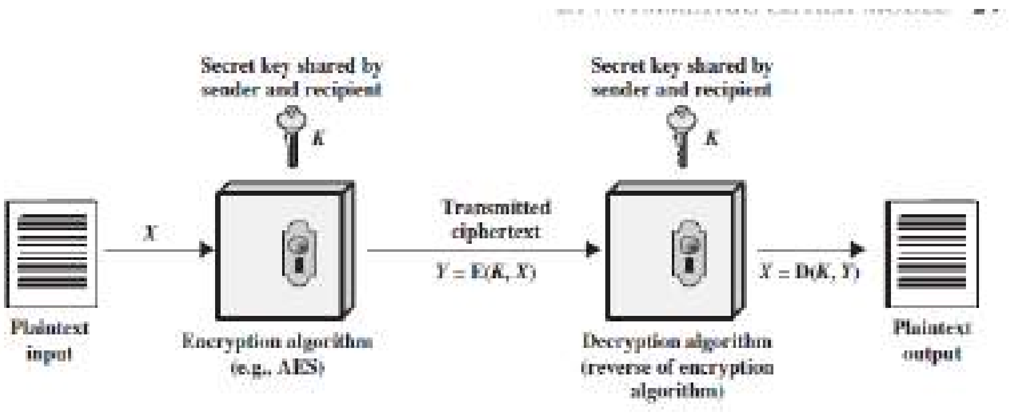

Secret key shared by sender and recipient
K
Secret key shared by sender and recipient
K
X
Transmitted ciphertext
Y = E(K, X)
X = D(K, Y)
Plaintext input
Encryption algorithm (e.g., AES)
Decryption algorithm (reverse of encryption algorithm)
Plaintext output

Figure: Simplified Model of Symmetric Encryption

- **Encryption algorithm:** The encryption algorithm performs various substitutions and transformations on the plaintext.
- **Secret key:** The secret key is also input to the encryption algorithm. The key is a value independent of the plaintext and of the algorithm. The algorithm will produce a different output depending on the specific key being used at the time. The exact substitutions and transformations performed by the algorithm depend on the key.
- **Ciphertext:** This is the scrambled message produced as output. It depends on the plaintext and the secret key. For a given message, two different keys will produce two different ciphertexts. The ciphertext is an apparently random stream of data and, as it stands, is unintelligible.
- **Decryption algorithm:** This is essentially the encryption algorithm run in reverse. It takes the ciphertext and the secret key and produces the original plaintext.

There are two requirements for secure use of conventional encryption:

1. ***We need a strong encryption algorithm***. At a minimum, we would like the algorithm to be such that an opponent who knows the algorithm and has access to one or more ciphertexts would be unable to decipher the ciphertext or figure out the key.
2. ***Sender and receiver must have obtained copies of the secret key in a secure***

fashion and must keep the key secure. If someone can discover the key and knows the algorithm, all communication using this key is readable.

There are two basic types of classical ciphers:

1. Substitution ciphers.
2. Transposition ciphers

Substitution Techniques

A substitution cipher changes characters in the plaintext to produce the ciphertext.

- A substitution technique is one in which the letters or symbols of plaintext are replaced by other letters or symbols. If the plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with ciphertext bit patterns
- In substitution ciphers the letters are systematically replaced by other letters or symbols. Eg. **Caesar Cipher**, **Vigenere Cipher**.

Caesar cipher:

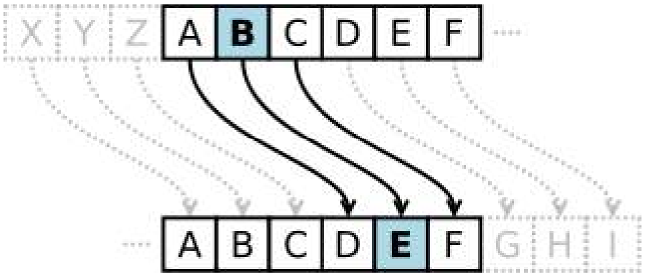
The Caesar cipher involves replacing each letter of the alphabet with the letter standing three places further down the alphabet.

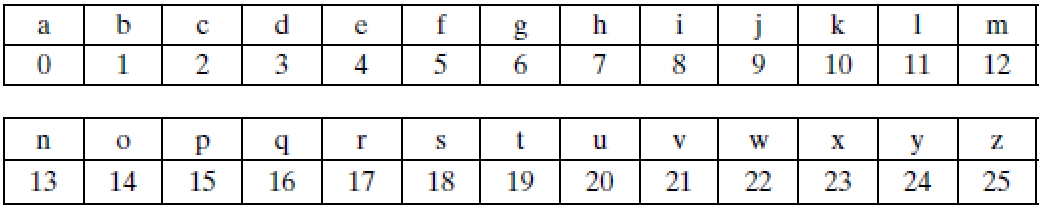
Note that the alphabet is wrapped around, so that the letter following Z is A.

We can define the transformation by listing all possibilities, as follows:

plain: a b c d e f g h i j k l m n o p q r s t u v w x y z

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





a	b	c	d	e	f	g	h	i	j	k	l	m
0	1	2	3	4	5	6	7	8	9	10	11	12

n	o	p	q	r	s	t	u	v	w	x	y	z
13	14	15	16	17	18	19	20	21	22	23	24	25

For example,

plain: meet me after the toga party

cipher: PHHW PH DIWHU WKH WRJD SDUWB

Algorithm:

For each plain text letter p, substitute the ciphertext letter C.

$C = E(3, p) = (p + 3) \bmod 26$

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

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

where k takes on a value in the range 1 to 25.

The decryption algorithm is simply $p = D(k, C) = (C - k) \bmod 26$

Note: This general form of Caesar Cipher is also known as **shift cipher**

Monoalphabetic Ciphers

- Monoalphabetic cipher is a substitution cipher in which for a given key, the cipher alphabet for each plain alphabet is fixed throughout the encryption process. For example, if ‘A’ is encrypted as ‘D’, for any number of occurrences in that plaintext, ‘A’ will always get encrypted to ‘D’.

- Caesar Cipher is an example of Monoalphabetic Cipher.
- With only 25 possible keys, the Caesar cipher is far from secure. An increase in the key space can be achieved by allowing an arbitrary substitution, which can improve the security.
- In Caesar cipher,
  plain: a b c d e f g h i j k l m n o p q r s t u v w x y z
  cipher: D E F G H I J K L M N O P Q R S T U V W X Y Z A B C

- If, instead, the "cipher" line can be any permutation of the 26 alphabetic characters, then there are 26! or greater than 4 * 10^26 possible keys.

- Let $p=C=Z_{26}$ and K consists of all possible permutations of the 26 alphabets.
  For each permutation π ε K, define:

$$e_\pi (x)= \pi (x)$$
$$d_\pi (y)= \pi^{-|}(y)$$

  where $\pi^{-|}$ is the inverse permutation. (possible keys= 26!)

For example,

Let π be:

a b c d e f g h i j k l m n o p q r s t u v w x y z

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

Plain Text: hello

Cipher text: PGDDL

Plain text: meet me after the toga party

Cipher text: UGGV UG RSVGY VPG VLKR NRYVI

Polyalphabetic cipher

- A polyalphabetic cipher is any cipher based on substitution, using multiple substitution alphabets.
- Same plaintext alphabet can be substituted by more than one (many) ciphertext alphabets.
- Example : Playfair Cipher, Hill Cipher, etc.

Playfair Cipher

- Invented by British scientist Sir Charles Wheatstone in 1854, but it bears the name of his friend Baron Playfair of St. Andrews, who championed the cipher at the British foreign office
- The best-known multiple-letter encryption cipher is the Playfair, which treats digrams in the plaintext as single units and translates these units into ciphertext digrams.
- Playfair Cipher is based on the use of a 5 * 5 matrix of letters constructed using a keyword.

Algorithm:

1. **Generate the Key Square**
   - 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 alphabetic order.
   - The key square is a 5×5 grid of alphabets that acts as the key for encrypting the plaintext. Each of the 25 alphabets must be unique and one letter of the alphabet (usually J) is omitted from the table (as the table can hold only 25 alphabets). If the plaintext contains J, then it replaces I (i.e. The letters I and J count as one letter.)

Example: (Key: MONARCHY)

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

In this example, the key is "MONARCHY". Thus, the initial entries are ‘M', 'O', 'N', 'A', 'R', 'C', 'H', 'Y’ followed by remaining characters of A-Z (except 'J') in that order

2. **Encrypt the plain text:**

The plaintext is split into pairs of two letters (digrams/digraphs). If there is an odd number of letters, a filler letter (any letter such as z) is added to the last letter.

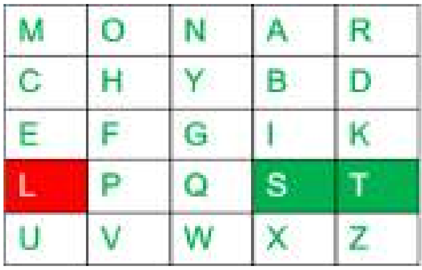
For example:

PlainText: "instruments"
After Split: 'in' 'st' 'ru' 'me' 'nt' 'sz'

Plaintext is encrypted two letters at a time, according to the following rules:

– Repeating plaintext letters that are in the same pair are separated with a filler letter, such as x, so that *balloon* would be treated as *ba lx lo on*.

– If both the letters are in the same row: Take the letter to the right of each one (going back to the leftmost if at the rightmost position).



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

✓ For example, digraph “st “ is encrypted as TL, “ar” is encrypted as RM

Diagraph: "st"
Encryption: s → T
t → L

Encrypted Text: TL

– If both the letters are in the same column: Take the letter below each one (going back to the top if at the bottom).

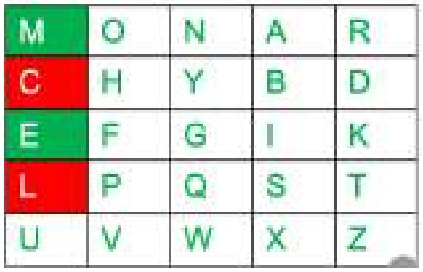
For example:

Diagraph: "me"

Encryption:

m → C
e → L

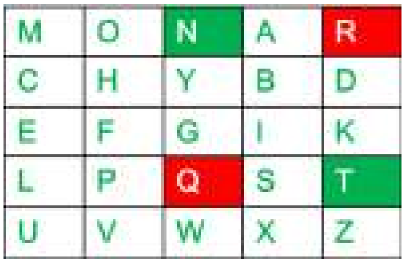
Encrypted Text: CL



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

– If neither of the above rules is true: Form a rectangle with the two letters and take the letters on the horizontal opposite corner of the rectangle.

For example:



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

Diagraph: "nt"

Encryption:

n -> R

t -> Q

Encrypted Text: RQ

Example:

Plain Text: "instruments"

After Split: 'in' 'st' 'ru' 'me' 'nt' 'sz'

Encryption:

i -> g

n -> a

s -> t

t -> l

r -> m

u -> z

m -> c

e -> l

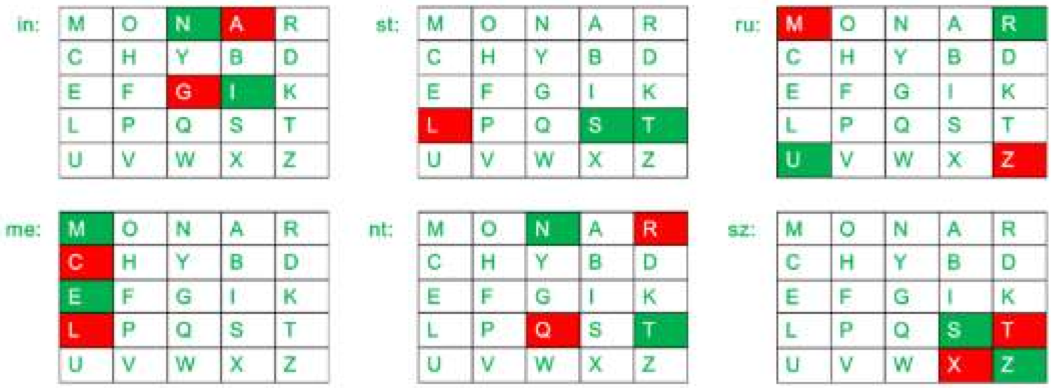
n -> r

t -> q

s -> t

z -> x

Encrypted Text: GATLMZCLRQTX



in:

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

st:

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

ru:

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

me:

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

nt:

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

sz:

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

Hill Cipher

- Hill cipher is a multi-lettered substitution cipher based on linear algebra. the **Hill** Invented by Lester S. Hill in 1929.
- Each letter is represented by a number modulo 26. Often the simple scheme A = 0, B = 1, …, Z = 25 is used, but this is not an essential feature of the cipher.
- To encrypt a message, each block of n letters (considered as an n-component vector) is multiplied by an invertible n × n matrix, against modulus 26.
- To decrypt the message, each block is multiplied by the inverse of the matrix used for encryption.
- The matrix used for encryption is the cipher key, and it should be chosen randomly from the set of invertible n × n matrices (modulo 26).
- This encryption algorithm takes m successive plaintext letters and substitutes for them m ciphertext letters. The substitution is determined by m linear equations in which each character is assigned a numerical value (a = 0, b = 1, c, z = 25).
  For m = 3, the system can be described as

$$c1 = (k11p1 + k21p2 + k31p3) \bmod 26$$

$$c2 = (k12p1 + k22p2 + k32p3) \bmod 26$$

$$c3 = (k13p1 + k23p2 + k33p3) \bmod 26$$

This can be expressed in terms of row vectors and matrices

$$(c_1\ c_2\ c_3) = (p_1\ p_2\ p_3)\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \bmod 26$$

Or,

C = PK mod 26

- Some cryptography books express the plaintext and ciphertext as column vectors, so that the column vector is placed after the matrix rather than the row vector placed before the matrix

- In general Hill System can be expressed as

$$\mathbf{C = E(K, P) = PK \bmod 26}$$
$$\mathbf{P = D(K, C) = CK^{-1} \bmod 26 = PKK^{-1} = P}$$

Example:

For example:

Let K= $\begin{bmatrix} 11 & 8 \\ 3 & 7 \end{bmatrix}$

Plaintext: july

ju= (9 20) & ly= (11 24)

(9 20) $\begin{bmatrix} 11 & 8 \\ 3 & 7 \end{bmatrix}$ mod 26

= (3 4)

i.e DE

So, (11 24) $\begin{bmatrix} 11 & 8 \\ 3 & 7 \end{bmatrix}$ mod 26

= (11 22)

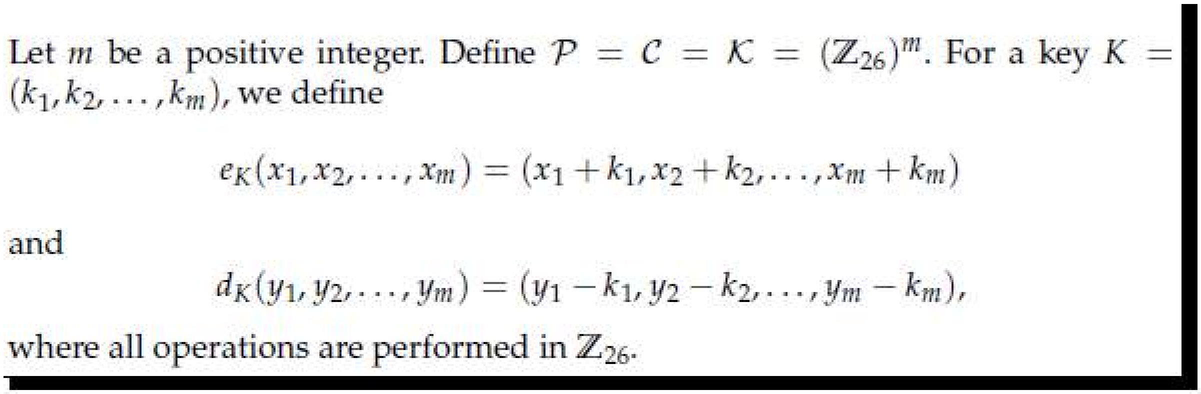
i.e. LW

Hence, the ciphertext is DELW.

For decryption, find $k^{-|}$ and multiply with the ciphertext in the form of matrix.

Vigenère Cipher

- → Polyalphabetic ciphers
- → This cipher is named after Blaise de Vigen`ere, who lived in the sixteenth century.
- → Using the correspondence A = 0, B = 1, . . . , Z = 25 described earlier, we can associate each key K with an alphabetic string of length **m**, called a keyword. The Vigenère Cipher encrypts **m** alphabetic characters at a time: each plaintext element is equivalent to **m** alphabetic characters.
- → This scheme of cipher uses a text string (say, a word) as a key, which is then used for doing a number of shifts on the plaintext



Let m be a positive integer. Define $\mathcal{P} = \mathcal{C} = \mathcal{K} = (\mathbb{Z}_{26})^m$. For a key $K = (k_1, k_2, \ldots, k_m)$, we define

$$e_K(x_1, x_2, \ldots, x_m) = (x_1 + k_1, x_2 + k_2, \ldots, x_m + k_m)$$

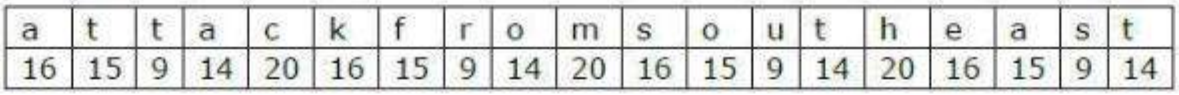
and

$$d_K(y_1, y_2, \ldots, y_m) = (y_1 - k_1, y_2 - k_2, \ldots, y_m - k_m),$$

where all operations are performed in $\mathbb{Z}_{26}$.

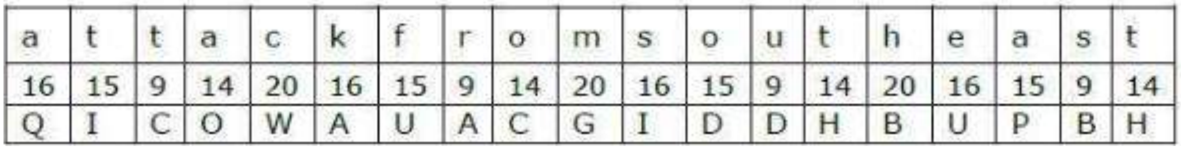
Process of Vigenere Cipher

- → The sender and the receiver decide on a key. Say **'point'** is the key. Numeric representation of this key is '16 15 9 14 20'.
- → The sender wants to encrypt the message, say **'attack from south east'**. He will arrange plaintext and numeric key as follows:



a	t	t	a	c	k	f	r	o	m	s	o	u	t	h	e	a	s	t
16	15	9	14	20	16	15	9	14	20	16	15	9	14	20	16	15	9	14

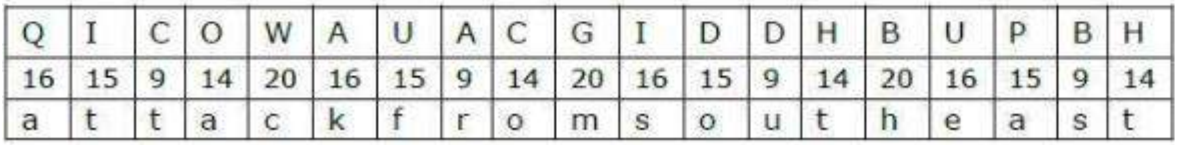
- → The sender now shifts each plaintext alphabet by the number written below it to create ciphertext as shown below:



a	t	t	a	c	k	f	r	o	m	s	o	u	t	h	e	a	s	t
16	15	9	14	20	16	15	9	14	20	16	15	9	14	20	16	15	9	14
Q	I	C	O	W	A	U	A	C	G	I	D	D	H	B	U	P	B	H

Here, each plaintext character has been shifted by a different amount and that amount is determined by the key. The key must be less than or equal to the size of the message.

→ For decryption, the receiver uses the same key and shifts received ciphertext in reverse order to obtain the plaintext.



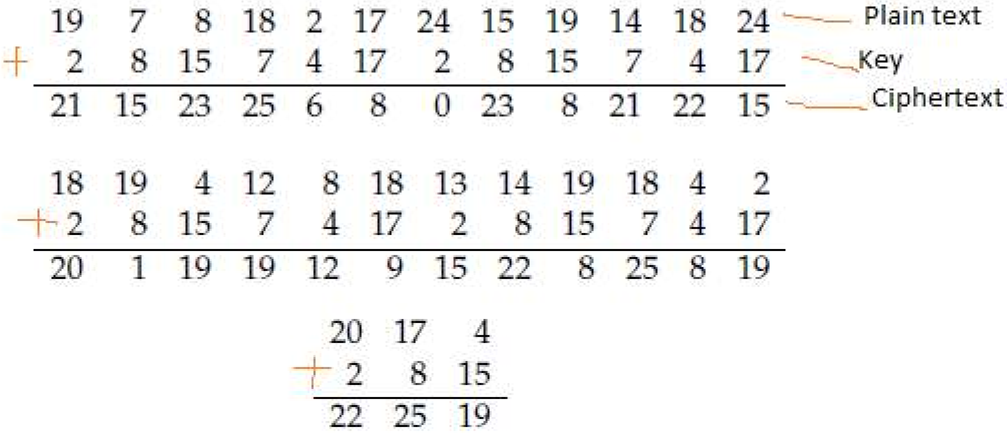
Q	I	C	O	W	A	U	A	C	G	I	D	D	H	B	U	P	B	H
16	15	9	14	20	16	15	9	14	20	16	15	9	14	20	16	15	9	14
a	t	t	a	c	k	f	r	o	m	s	o	u	t	h	e	a	s	t

Suppose the key is CIPHER. Compute the ciphertext of

“this crypto system is not secure”

➔ Here the key length (m) = 6 and numeric representation of key is
K = (2, 8, 15, 7, 4, 17).

We convert the plaintext elements to residues modulo 26, write them in groups of six, and then “add” the keyword modulo 26, as follows:



$$\begin{array}{rrrrrrrrrrrrrl} & 19 & 7 & 8 & 18 & 2 & 17 & 24 & 15 & 19 & 14 & 18 & 24 & \text{Plain text} \\ + & 2 & 8 & 15 & 7 & 4 & 17 & 2 & 8 & 15 & 7 & 4 & 17 & \text{Key} \\ \hline & 21 & 15 & 23 & 25 & 6 & 8 & 0 & 23 & 8 & 21 & 22 & 15 & \text{Ciphertext} \end{array}$$

$$\begin{array}{rrrrrrrrrrrrr} & 18 & 19 & 4 & 12 & 8 & 18 & 13 & 14 & 19 & 18 & 4 & 2 \\ + & 2 & 8 & 15 & 7 & 4 & 17 & 2 & 8 & 15 & 7 & 4 & 17 \\ \hline & 20 & 1 & 19 & 19 & 12 & 9 & 15 & 22 & 8 & 25 & 8 & 19 \end{array}$$

$$\begin{array}{rrrr} & 20 & 17 & 4 \\ + & 2 & 8 & 15 \\ \hline & 22 & 25 & 19 \end{array}$$

The alphabetic equivalent of the ciphertext string would thus be:

VPX ZGIAXIVWPUBT TM JPW IZITWZT

There are two special cases of Vigenere cipher:

- The *keyword length is same as plaintext message*. This case is called **Vernam Cipher**. It is more secure than typical Vigenere cipher.
- Vigenere cipher becomes a cryptosystem with *perfect secrecy*, which is called *One-Time pad.*

  Perfect secrecy is the concept that given a ciphertext from a perfectly secure encryption system, absolutely nothing will be revealed about the plaintext by the ciphertext.

One-Time Pad

- An Army Signal Corp officer, Joseph Mauborgne, proposed an improvement to the Vernam cipher that yields the ultimate in security
- Random key is used which is as long as the message, so that the key need not be repeated
- The key is to be used to encrypt and decrypt a single message, and then is discarded.
- Each new message requires a new key of the same length as the new message.
- Such a scheme, known as a one-time pad, is unbreakable. It produces random output that bears *no statistical relationship to the plaintext*. Because the ciphertext contains no information whatsoever about the plaintext, there is simply *no way to break the code*.
- Messages encrypted with keys based on randomness have the advantage that there is theoretically no way to "break the code" by analyzing the messages.
- Each encryption is unique and bears no relation to the next encryption, so patterns between the messages cannot be detected.
- When a message is to be sent, the sender uses the secret key to encrypt each character, one at a time.
- With a one-time pad, however, the decrypting party must have access to the same key used to encrypt the message and this raises the problem of how to get the key to the decrypting party safely.
- The key used in a one-time pad is called a secret key because if it is revealed, the messages encrypted with it can easily be deciphered

The security of the one-time pad is entirely due to the randomness of the key. If the stream of characters that constitute the key is truly random, then the stream of characters that constitute the ciphertext will be truly random. Thus, there are no patterns or regularities that a cryptanalyst can use to attack the ciphertext.

In theory, we need look no further for a cipher. The one-time pad offers complete security but, in practice, has two fundamental difficulties:

1. There is **the practical problem of making large quantities of random keys**. Any heavily used system might require millions of random characters on a regular basis. Supplying truly random characters in this volume is a significant task.
2. Even more daunting is **the problem of key distribution and protection**. For every message to be sent, a key of equal length is needed by both sender and receiver. Thus, a mammoth key distribution problem exists.

➔ Because of these difficulties, the one-time pad is of limited utility and is useful primarily for low-bandwidth channels requiring very high security.

➔ The one-time pad **is the only cryptosystem that exhibits what is referred to as *perfect secrecy.***

Transposition Techniques

- In transposition technique, the identity of the characters remains unchanged, but their positions are changed to create the ciphertext.
- Transposition Techniques are based on the permutation of the plain-text instead of substitution.
- In a transposition cipher, the order of the alphabets is re-arranged to obtain the ciphertext
- All Ciphertext letters will be the letters from plain text.
- Example: Rail Fence Cipher

Rail Fence Cipher

- Also called a zigzag cipher
- Rail fence is the simplest transposition cipher technique in which the plaintext is written down as a sequence of diagonals and then read off as a sequence of rows. i.e. first, we write the message in a zigzag manner then read it out direct row-wise to change it to cipher-text
- The number of lines (rails) used in a Rail Fence Cipher is the key

Encryption process:

- In the rail fence cipher, the plain-text is written downwards and diagonally on successive rails of an imaginary fence.
- When we reach the bottom rail, we traverse upwards moving diagonally, after reaching the top rail, the direction is changed again. Thus the alphabets of the message are written in a zig-zag manner.
- After each alphabet has been written, the individual rows are combined to obtain the cipher-text.

Note: Number of rows = number of rails = depth of Rail fence = Key

Number of columns= Number alphabets in plain text

For example, (Rail = 3i.e. Key = 3)

Plaintext: “meet me after the toga party”

m				m				t				h				g				r		
	e		t		e		f		e		t		e		o		a		a		t	
		e				a				r				t				p				y

Ciphertext = MMTHGRTETEFETEOAATEARTPY

Decryption

- The number of columns in rail fence cipher remains equal to the length of plain-text message. And the key corresponds to the number of rails.
- Rail matrix can be constructed accordingly. Once we’ve got the matrix we can figure-out the spots where texts should be placed (using the same way of moving diagonally up and down alternatively).
- Then, we fill the cipher-text row wise. After filling it, we traverse the matrix in zig-zag manner to obtain the original text.

Example :

Ciphertext = MMTHGRTETEFETEOAATEARTPY

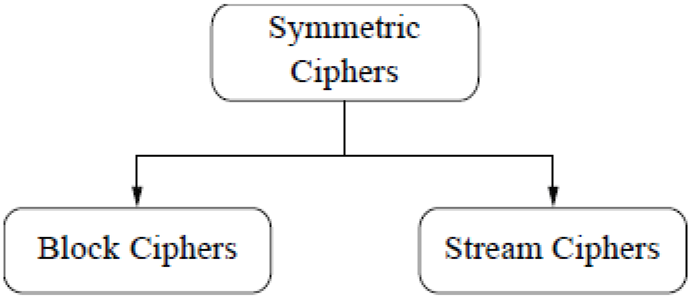
Key = 3

M				M				T				H				G				R		
	E		T		E		F		E		T		E		O		A		A		T	
		E				A				R				T				P				Y

Reading diagonally plain text is obtained.

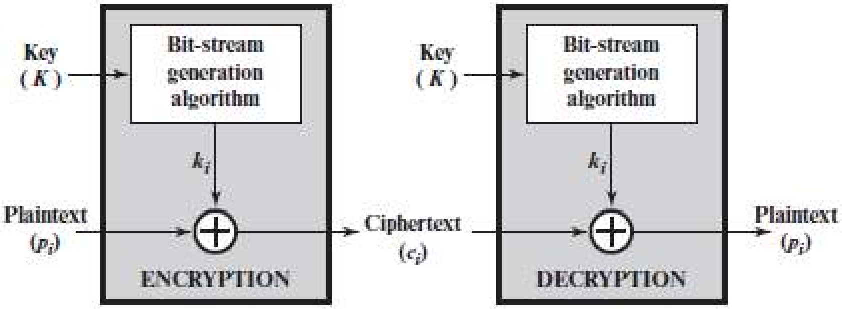
Modern Ciphers

- In Modern ciphers, digital data is represented in strings of binary digits (bits) unlike alphabets.
- Modern cryptography is the cornerstone of computer and communications security. Its foundation is based on various concepts of mathematics such as number theory, computational-complexity theory, and probability theory.
- Modern cryptosystems need to process these binary strings to convert into another binary string. Based on how these binary strings are processed, a symmetric encryption scheme can be classified into stream cipher and block cipher



Stream cipher

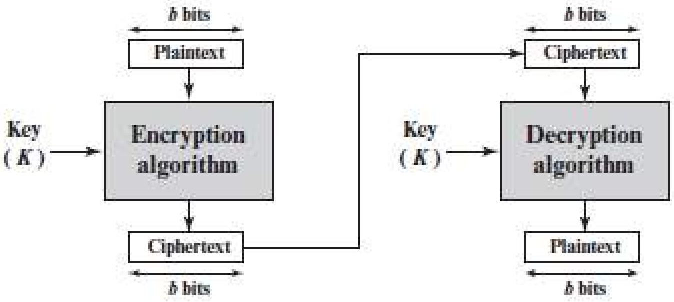
- A stream cipher is the mechanism that encrypts a digital data stream one bit or one byte at a time.
- In this scheme, the plaintext is processed one bit at a time i.e. one bit of plaintext is taken, and a series of operations are performed on it to generate one bit of ciphertext.
- For practical reasons, the bit-stream generator must be implemented as an algorithmic procedure, so that the cryptographic bit stream can be produced by both users.
- In this approach, the bit-stream generator is a key-controlled algorithm and must produce a bit stream that is cryptographically strong. That is, it must be computationally impractical to predict future portions of the bit stream based on previous portions of the bit stream. The two users need only share the generating key, and each can produce the keystream.



(a) Stream cipher using algorithmic bit-stream generator

Block Cipher

- A block cipher is the mechanism in which a block of plaintext is treated as a whole and used to produce a ciphertext block of equal length.
- The number of bits in a block is fixed. Typically, a block size of 64 or 128 bits is used.
- As with a stream cipher, the two users share a symmetric encryption key



(b) Block cipher

- In general, block ciphers seem applicable to a broader range of applications than stream ciphers

Symmetric vs. Asymmetric Ciphers

Symmetric Cryptography/Cryptosystem/Encryption/Cipher

- Symmetric encryption is a technique which allows the use of **only one key for performing both the encryption and the decryption of the message** shared over the internet. It is also known as the conventional method used for encryption.
- In symmetric encryption, the plaintext is encrypted and is converted to the ciphertext using a key and an encryption algorithm. While the cipher text is converted back to plain text using the same key that was used for encryption, and the decryption algorithm.
- In symmetric ciphers, the same key encrypts and decrypts the data. So, both sender and receiver needs to have the shared key.
- Symmetric cryptography is also called **private key cryptography.**

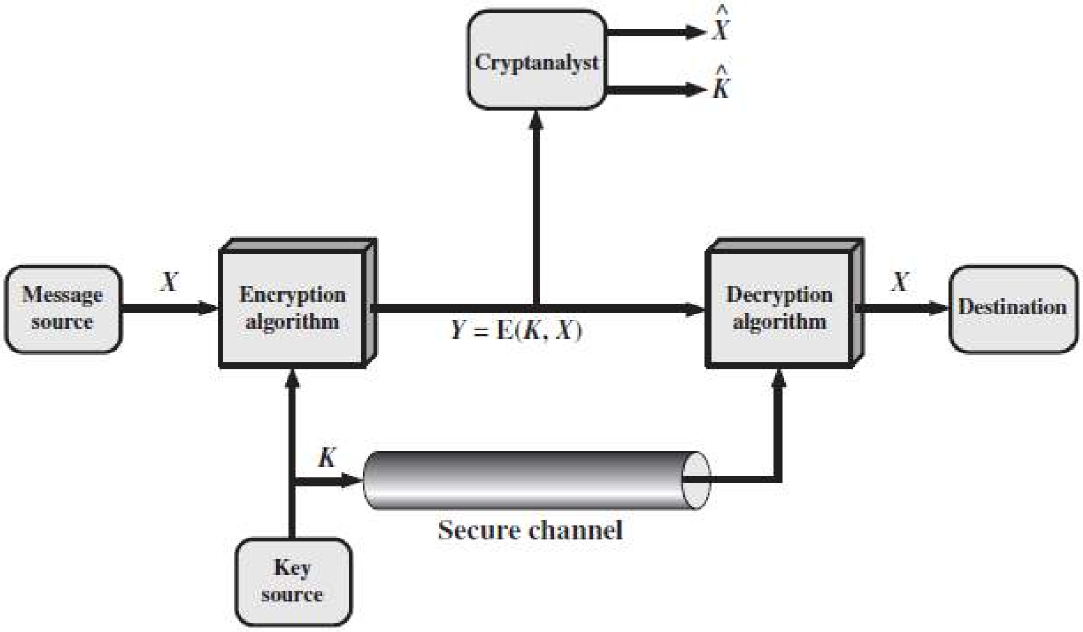


Figure: Model of Symmetric Cryptosystem

Asymmetric Cryptography/Cryptosystem/Encryption/Cipher

- Asymmetric encryption is an encryption technique that uses a pair of key (private key and public key) for encryption and decryption
- Asymmetric algorithms rely on one key for encryption and a different but related key for decryption
- Asymmetric algorithms rely on one key for encryption and a different but related key for decryption
- Asymmetric encryption uses the **one key for the encryption** of the message and the **other key for the decryption of the message**.
- The **public key is freely available** to anyone who is interested in sending the message. The **private key is kept secret** with the receiver of the message.
- Also known as **public key cryptosystem**.
- In this approach, it is computationally infeasible to determine the decryption key given only knowledge of the cryptographic algorithm and the encryption key.
- Either of the two related keys can be used for encryption, with the other used for decryption.
    - ⇒ Data encrypted by the public key can only be decrypted by the private key. (useful to ensure confidentiality)
    - ⇒ Data encrypted by the private key can only be decrypted by the public key. (useful in system like digital signature)

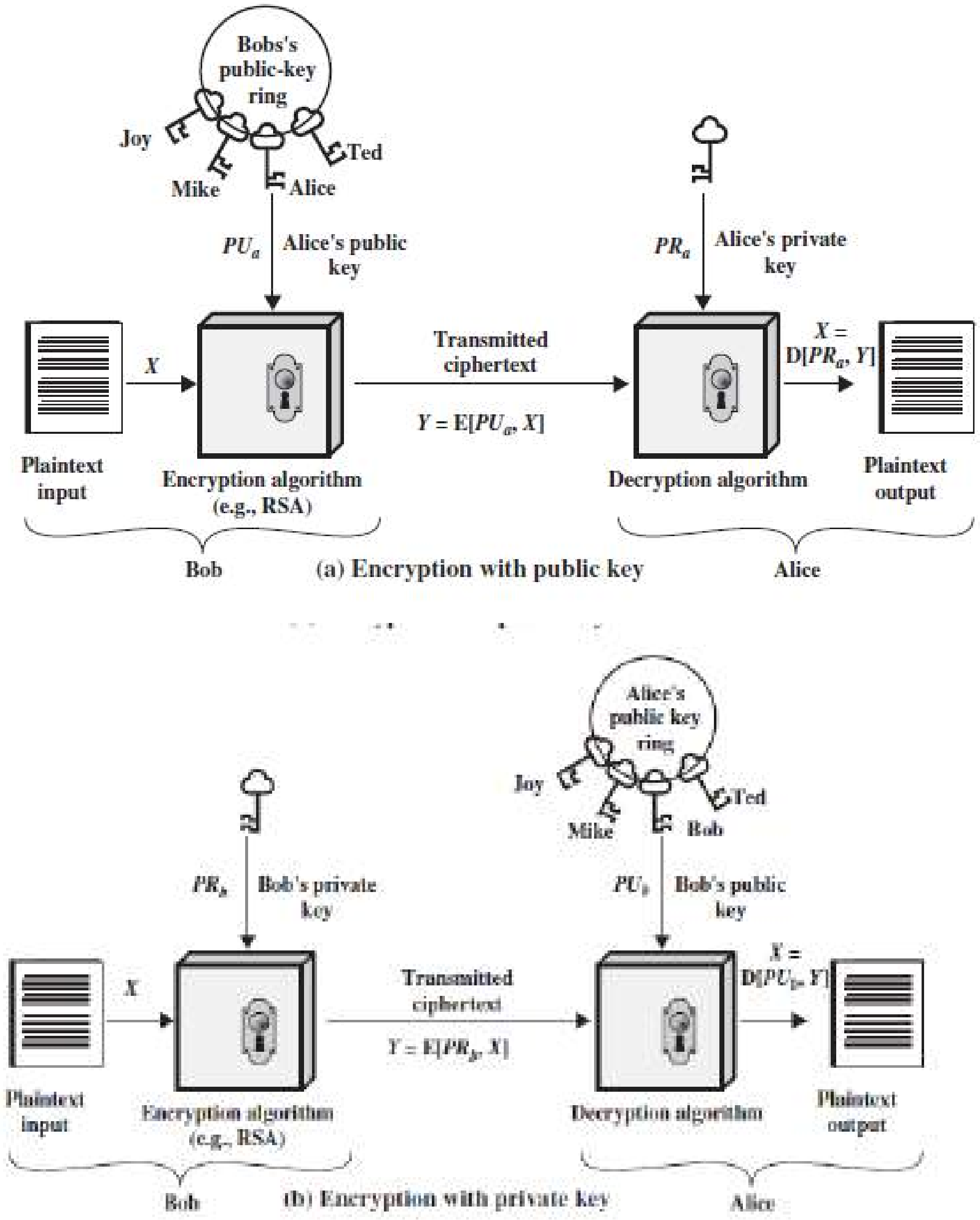


Figure: Asymmetric Cryptosystem / Public key Cryptography

Kerckhoffs's principle

- Kerckhoffs's principle is one of the basic principles of modern cryptography. It was formulated in the end of the nineteenth century (in 1883) by Dutch cryptographer Auguste Kerckhoffs. The principle is as follows:

  A cryptographic system should be secure even if everything about the system, except the key, is public knowledge.