Chapter 1. Introduction and Classical Ciphers

Prepared by:

Kobid Karkee

Himalaya College of Engineering

Cryptography

BScCSIT CSC316

Background

Information Security was provided, before digital age, in an organization by physical and administrative means

e.g. Filing cabinet with locking system, personnel screening at the time of recruitment etc.

With the introduction of computers, and development of shared systems, public telephone networks, data networks and the Internet, the term ***Computer Security*** was defined as "*A collection of tools designed to protect data and to thwart hackers*".

Distributed systems and network/communication facilities give rise to the need of security measures to protect data during their transmission, and hence the term ***Network security*** was introduced.

Nowadays, most organizations interconnect their data processing equipments with inter-connected networks (i.e. Internet). So, the term ***Internet security*** is used.

Computer Security : Definition

The NIST Computer Security Handbook [NIST95] defines: -

- "The protection afforded to an automated information system in order to attain the applicable objectives of preserving the integrity, availability, and confidentiality of information system resources (*includes hardware, software, firmware, information/data, and telecommunications*)."

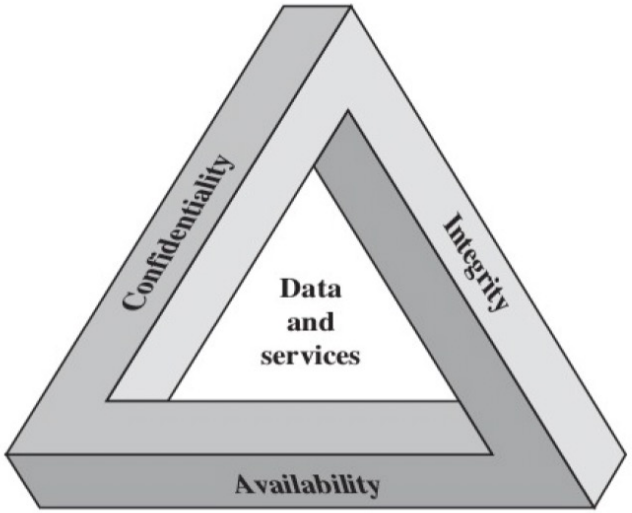
- This definition introduces three key concepts, collectively called CIA Triad: -
  - Confidentiality
  - Integrity
  - Availability

- Security characteristics have also been set on the basis of these objectives.

The CIA Triad: Objectives

The CIA triad embodies the fundamental security objectives for both data and for information and computing services.

- ***Confidentiality***: covers two 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.



- ***Integrity***: covers two concepts:
  - Data integrity:
    - Assures that information 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.
- ***Availability***:
  - Assures that systems work promptly and service is not denied to authorized users.

CIA Triad: Security Characteristics

(FIPS 199 provides characterization and the definition of a loss of security in each category.)

Confidentiality:

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

Integrity:

- Guarding against improper information modification or destruction
- includes ensuring information non-repudiation and authenticity.
- *A loss of integrity is the unauthorized modification or destruction of information.*

Availability:

- 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.

Additional concepts: -

- ***Authenticity***:
  - The property of being genuine and being able to be verified and trusted,
- ***Accountability***:
  - The security goal that generates the requirement for actions of an entity to be traced uniquely to that entity.

Impacts of Security Breach
(defined by FIPS)

Low: a ***limited effect*** on ***organizational operations***, organizational ***assets***, or ***individuals***

- (i) a minor degradation in mission capability to an extent and duration that the organization is able to perform its primary functions, but the effectiveness of the functions is noticeably reduced;
- (ii) minor damage to organizational assets;
- (iii) minor financial loss; or
- (iv) minor harm to individuals.

Moderate: a ***major effect*** on organizational operations, organizational assets, or individuals

- (i) a significant degradation in mission capability to an extent and duration that the organization is able to perform its primary functions, but the effectiveness of the functions is significantly reduced;
- (ii) significant damage to organizational assets;
- (iii) significant financial loss; or
- (iv) significant harm to individuals that does not involve loss of life or serious, life-threatening injuries.

High: a severe/catastrophic effect on organizational operations, organizational assets, or individuals

- (i) a severe degradation in or loss of mission capability to an extent and duration that the organization is not able to perform one or more of its primary functions;
- (ii) major damage to organizational assets;
- (iii) major financial loss; or
- (iv) severe or catastrophic harm to individuals involving loss of life or serious, life-threatening injuries.

Computer Security: Challenges

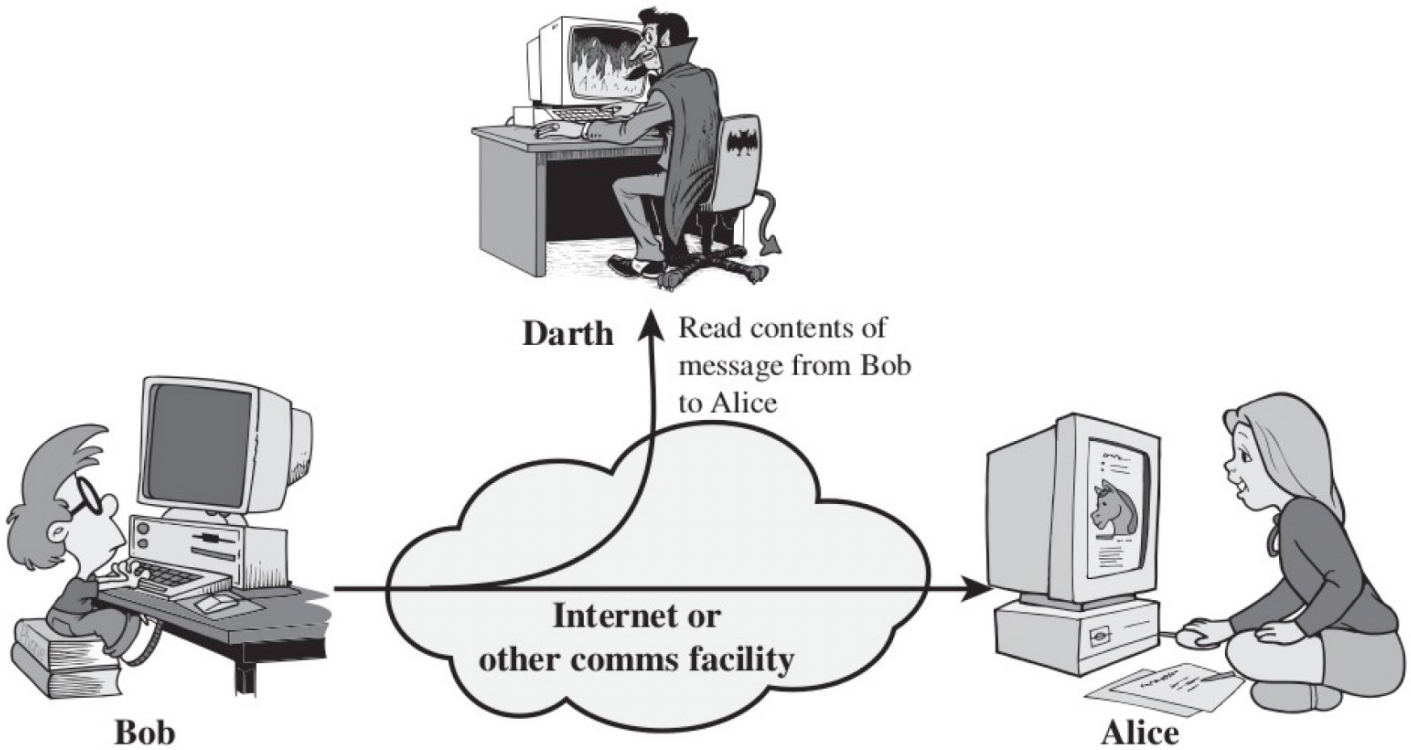
- Security requirements seem to be one-word labels (e.g. confidentiality, authentication, non-repudiation, or integrity etc.), but ***the mechanisms are complex***.
- To develop a security mechanism, all ***potential attacks on security features must be considered***.
- Procedures of a ***security mechanism is not commonly understood***, they make sense only when various aspects of threats are considered.
- It is necessary to decide, while designing security mechanism, ***where to use them*** - ***physical placement*** (e.g., at what points in a network) and a ***logical placement*** (e.g., at what layer(s) of an architecture.)
- Security mechanisms ***involve more than an algorithm/protocol*** such as possession of some secret information (e.g., an encryption key), reliance on communications protocols.

- An attacker just needs to find a single weakness but ***the designer must find and eliminate all weaknesses*** to achieve perfect security.
- Users and security managers are ***reluctant to invest in security*** unless a security failure occurs.
- Security ***requires regular, even constant, monitoring***, which is difficult in today's short-term, overloaded environment.
- Security is still ***incorporated into a system after the design is complete*** rather than being an integral part of the design process.
- A ***strong security is considered as an obstruction to efficient and user-friendly operation*** of an information system or use of information.

Security Attacks/Threats

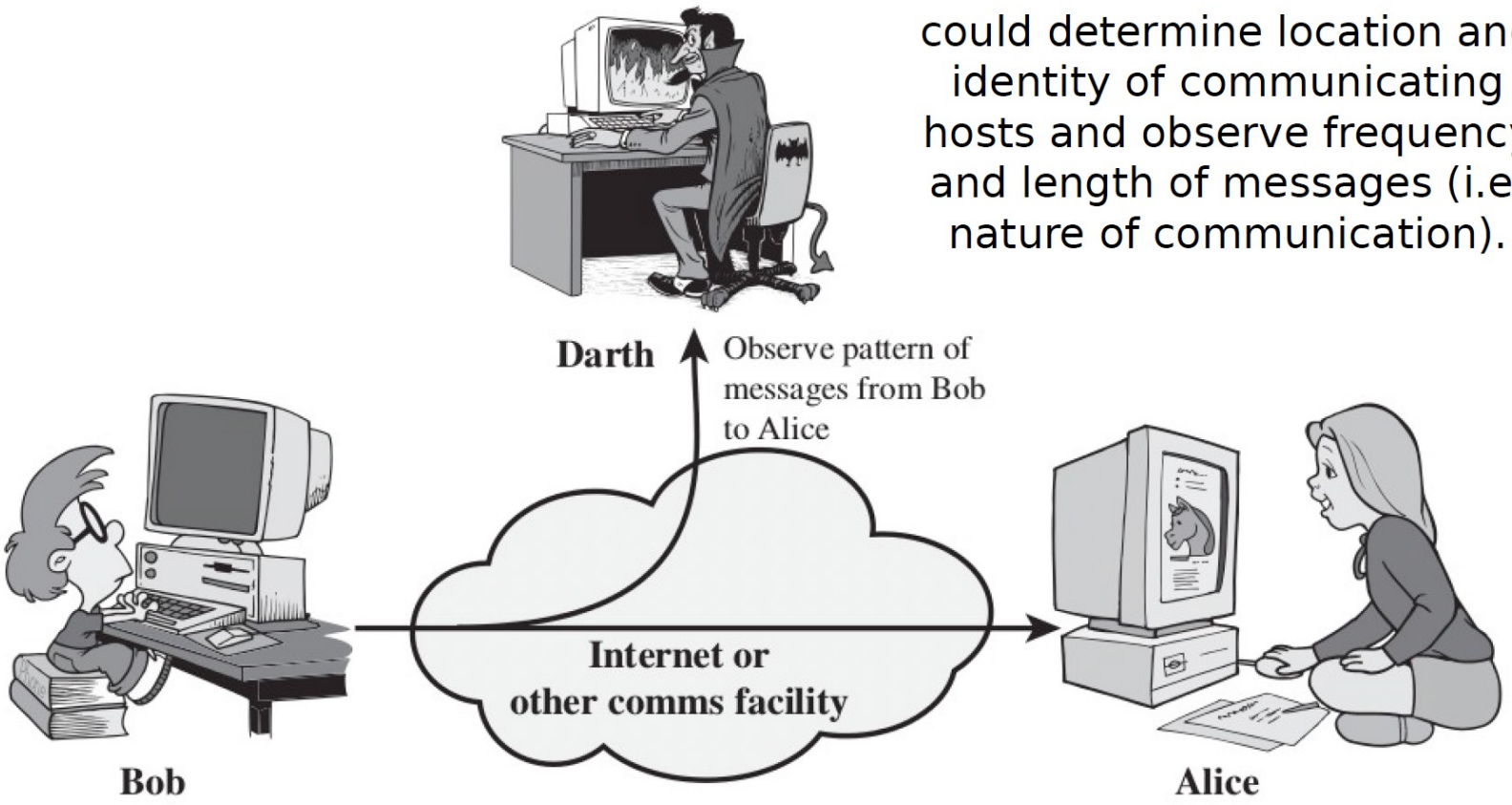
- ***Passive attacks***:
  - attempt to learn or make use of information from the system but ***does not affect system resources***
  - are in the nature of ***eavesdropping*** (listening secretly to the private conversation) on, or ***monitoring*** of, ***transmissions***.
  - goal of the opponent is ***to obtain information*** that is being transmitted.
  - Very ***difficult to detect***, because they do not involve any alteration of the data. 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
  - e.g. ***release of message contents*** and ***traffic analysis*** etc.
- ***Active attacks***:
  - attempt to ***alter system resources*** or ***affect their operation***.
  - involve some ***modification of the data stream*** or the ***creation of a false stream***
  - Active attacks are quite ***difficult to be absolutely prevented*** because of the wide variety of potential physical, software, and network vulnerabilities.
  - The goal is ***to detect active attacks and to recover from any disruption or delays*** caused by them
  - e.g. ***masquerade***, ***replay***, ***modification of messages***, ***denial of service*** etc.

Passive attack: Release of Message Contents



Telephone conversation, email messages, file transfer etc.

Passive attack: Traffic Analysis

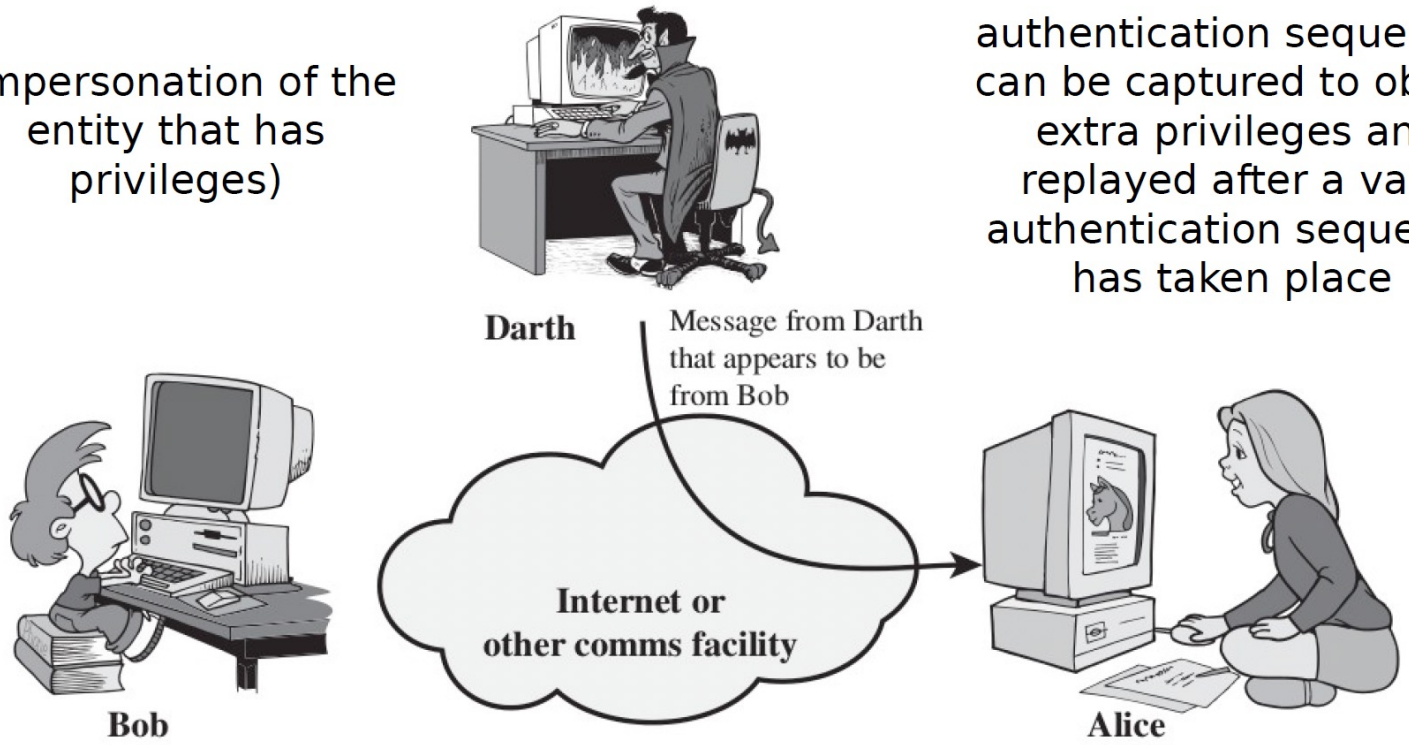


could determine location and identity of communicating hosts and observe frequency and length of messages (i.e. nature of communication).

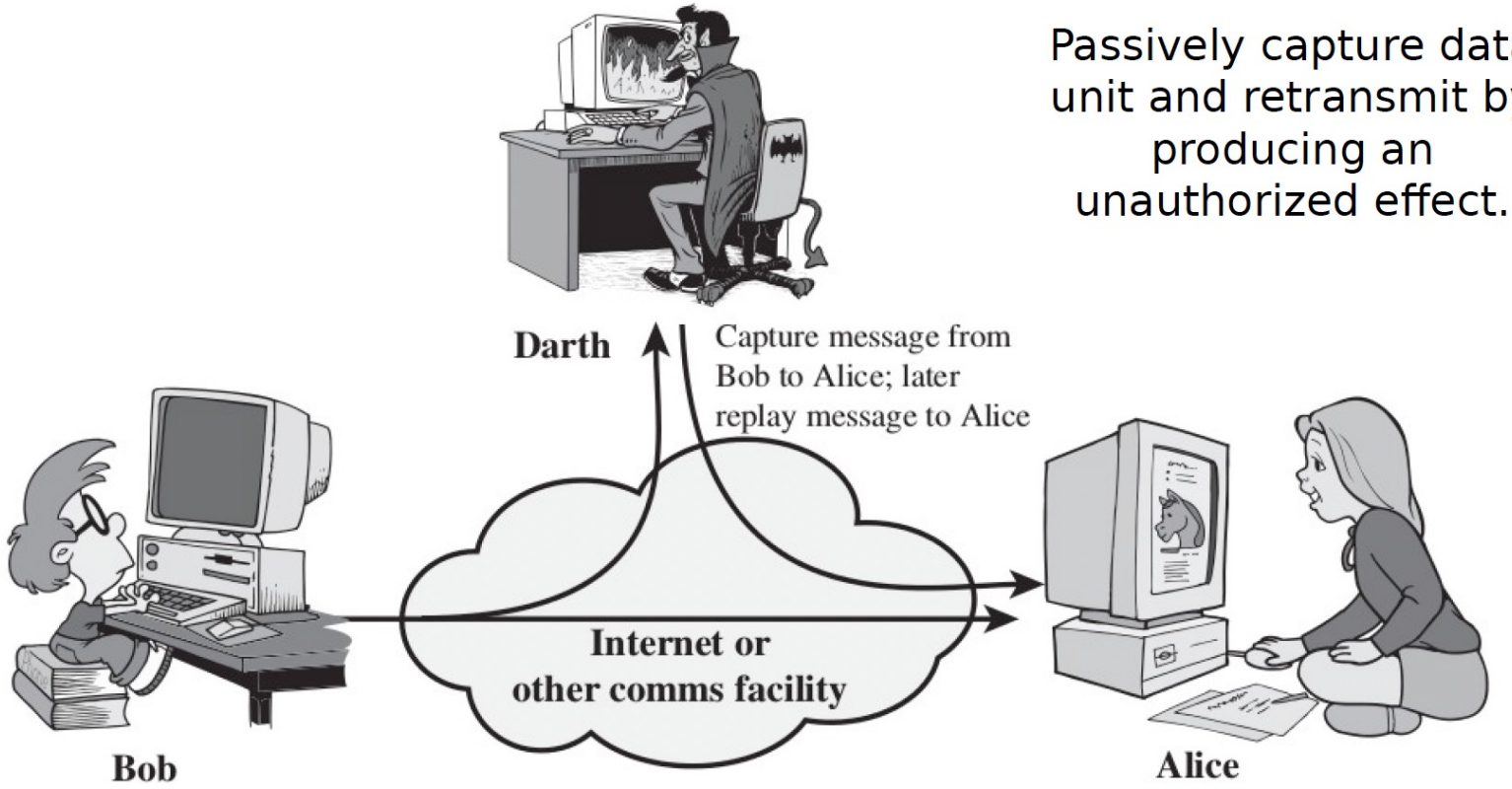
Active attack: Masquerade

(Impersonation of the entity that has privileges)

authentication sequences can be captured to obtain extra privileges and replayed after a valid authentication sequence has taken place

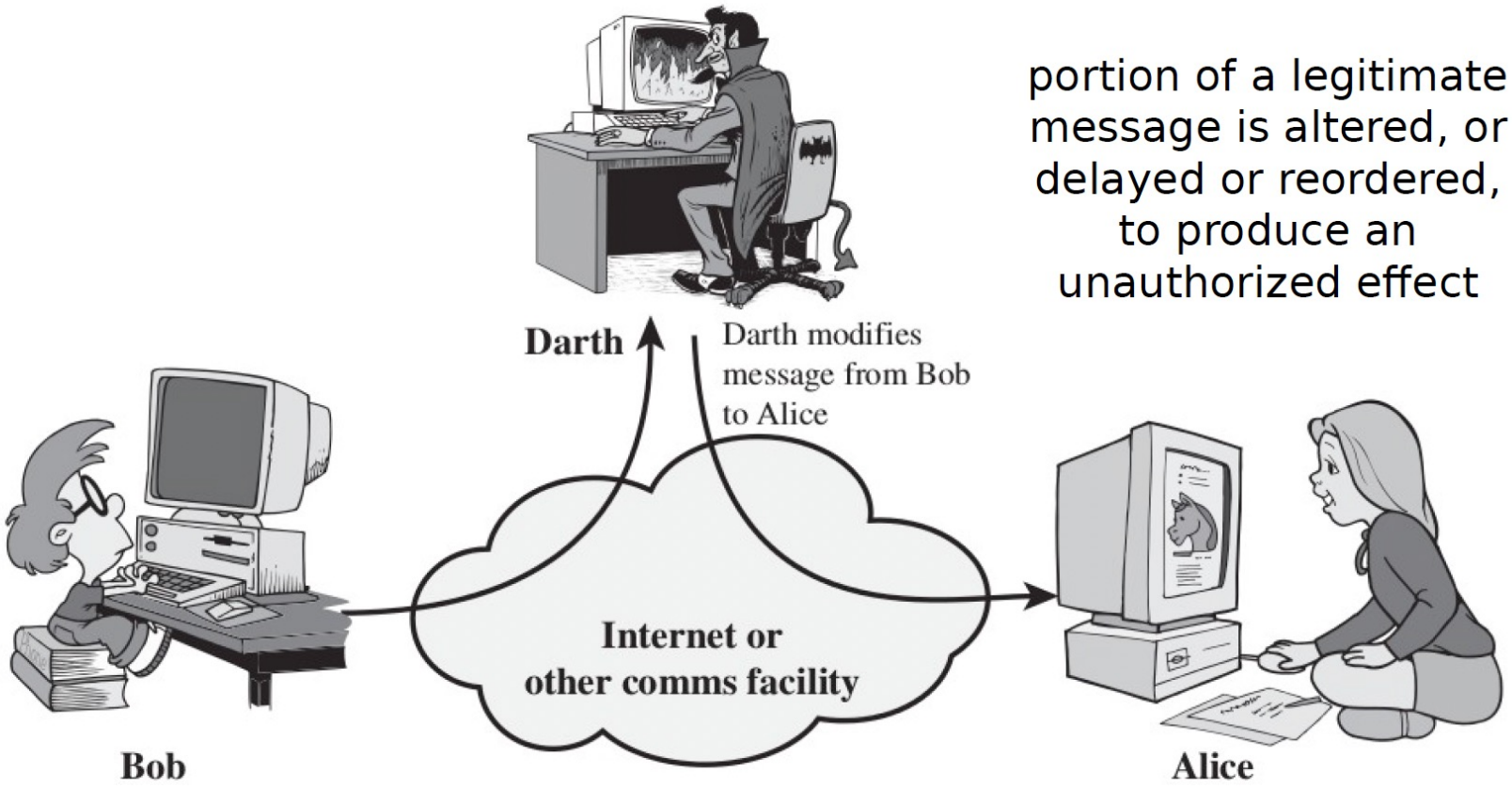


Active attack: Replay



Passively capture data unit and retransmit by producing an unauthorized effect.

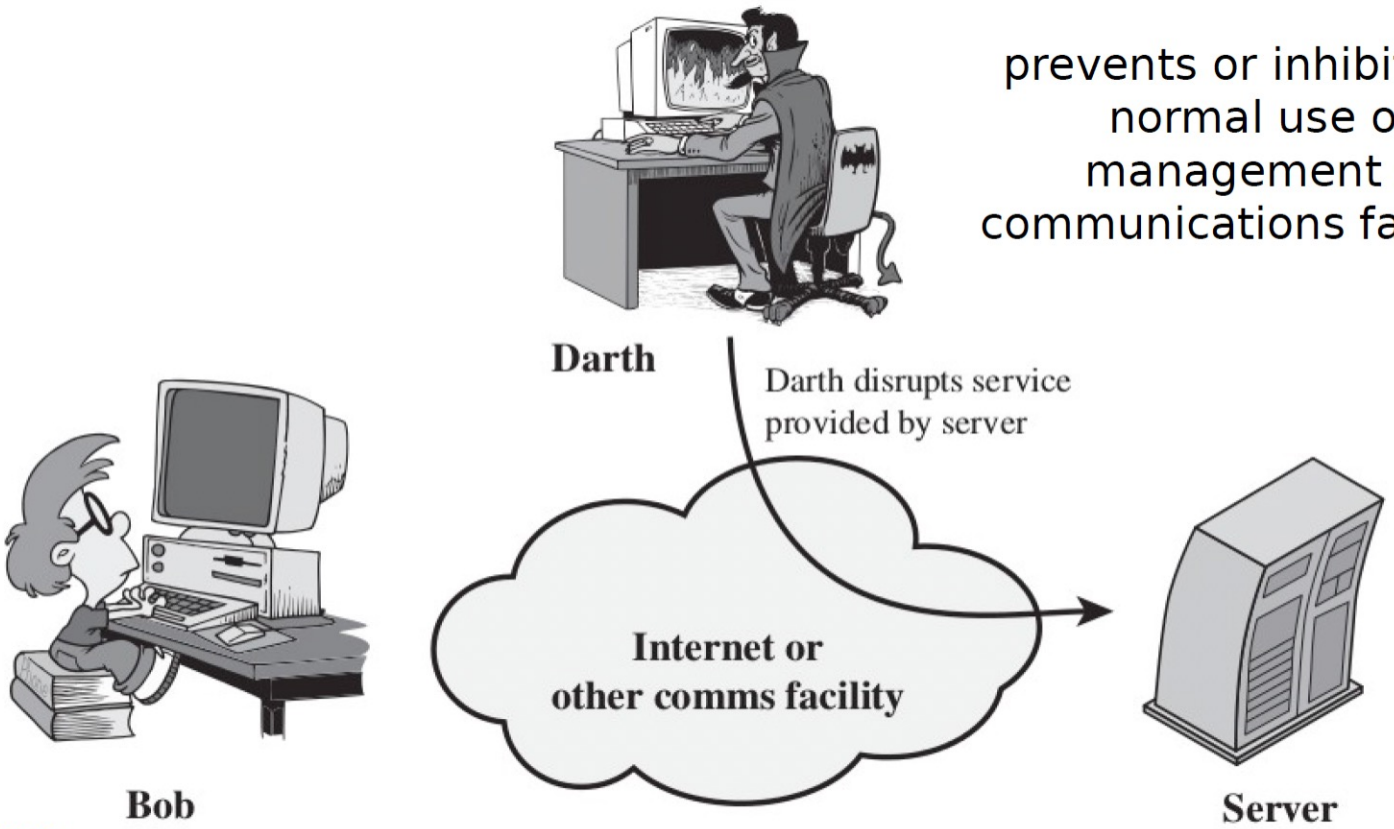
Active attack: Modification of Messages



portion of a legitimate message is altered, or delayed or reordered, to produce an unauthorized effect

Active attack: Denial of Service

prevents or inhibits the normal use or management of communications facilities



Examples: -

(1) an entity may suppress all messages directed to a particular destination

(2) disable the whole network

(3) overload the network with messages so as to degrade performance

Cryptography

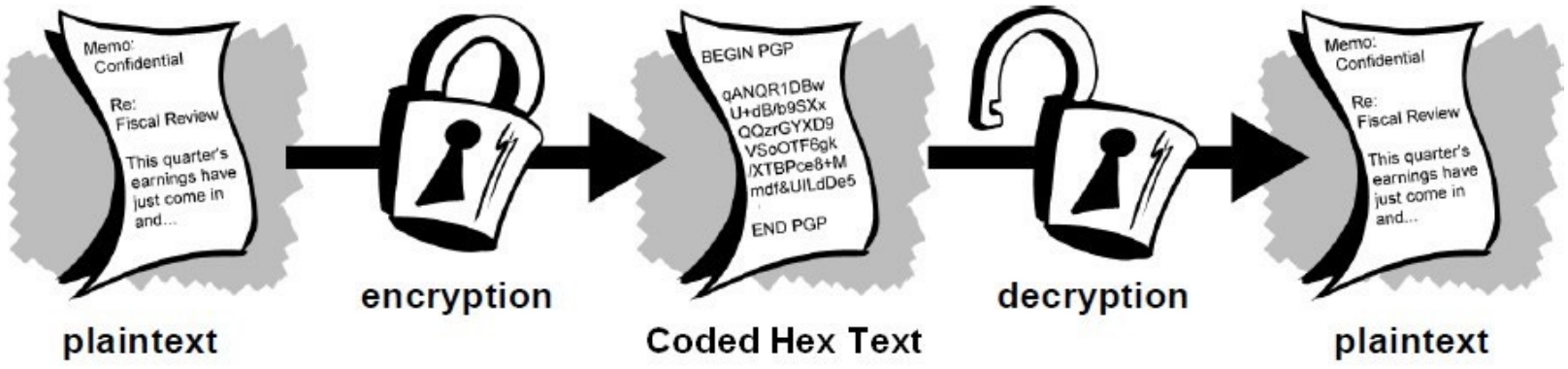
- Cryptography is a method of protecting information and communications through the use of codes so that only those for whom the information is intended can read and process it.
- Cryptography is the science of using mathematics to encrypt and decrypt data. :Phil Zimmermann
- Cryptography is the art and science of keeping messages secure.:Bruce Schneier

Cryptography

- A message is plaintext (sometimes called cleartext).
- The process of disguising a message in such a way as to hide its substance is encryption.
- An encrypted message is ciphertext.
- The process of turning ciphertext back into plaintext is decryption.

Cryptography

- A cipher (or cypher) is an algorithm for performing encryption or decryption—a series of well-defined steps that can be followed as a procedure.



Cryptography

- Goal: The primary goal of cryptography is to secure important data on the hard disk or as it passes through a medium that may not be secure itself. Usually, that medium is a computer network.
- Services: Cryptography can provide the following services:
  - Confidentiality (secrecy)
  - Integrity (anti-tampering)
  - Authentication
  - Non-repudiation.

Cryptography

Confidentiality (secrecy)

- Ensuring that no one can read the message except the intended receiver
- Data is kept secret from those without the proper credentials, even if that data travels through an insecure medium

Integrity (anti-tampering)

- Assuring the receiver that the received message has not been altered in any way from the original.

Cryptography

Authentication

- The process of proving one's identity is called authentication.
- Cryptography can help establish identity for authentication purposes.

Non-repudiation

- A mechanism to prove that the sender really sent this message

Types of Cryptography

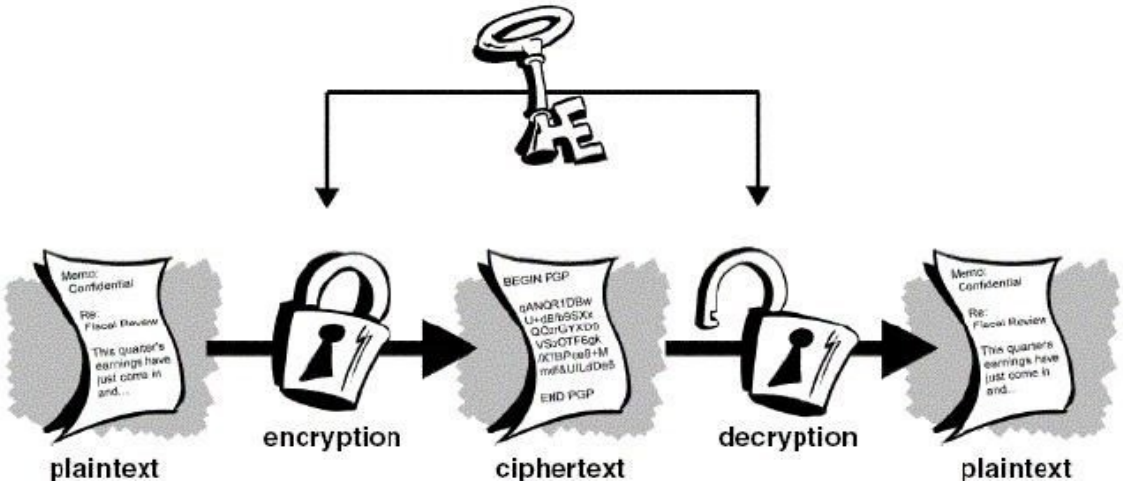
1. Symmetric Key Cryptography
2. Asymmetric Key Cryptography
3. Hash Functions

Symmetric Key Cryptography

- Also known as Secret Key Cryptography or Conventional Cryptography
- Symmetric Key Cryptography is an encryption system in which the sender and receiver of a message share a single, common key that is used to encrypt and decrypt the message.
- A key is a piece of information (a parameter) that determines the functional output of a cryptographic algorithm or cipher.

Symmetric Key Cryptography

- The key for encrypting and decrypting the file had to be known to all the recipients. Else, the message could not be decrypted by conventional means.



Symmetric Key Cryptography

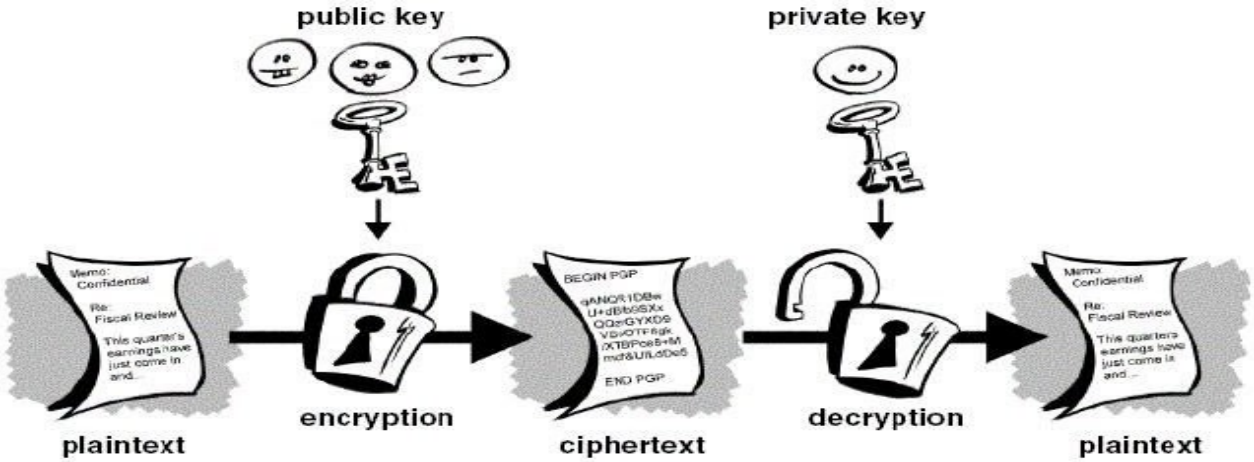
Problems with Conventional Cryptography

Key Management

- Symmetric-key systems are simpler and faster; their main drawback is that the two parties must somehow exchange the key in a secure way and keep it secure after that.
- Key Management caused nightmare for the parties using the symmetric key cryptography. They were worried about how to get the keys safely and securely across to all users so that the decryption of the message would be possible. This gave the chance for third parties to intercept the keys in transit to decode the top-secret messages. Thus, if the key was compromised, the entire coding system was compromised and a “Secret” would no longer remain a “Secret”.
- This is why the “Public Key Cryptography” came into existence.

Asymmetric Key Cryptography

- Asymmetric cryptography , also known as Public-key cryptography, refers to a cryptographic algorithm which requires two separate keys, one of which is private and one of which is public.
- The public key is used to encrypt the message and the private one is used to decrypt the message.



Asymmetric Key Cryptography

Step 1: Give your public key to the sender

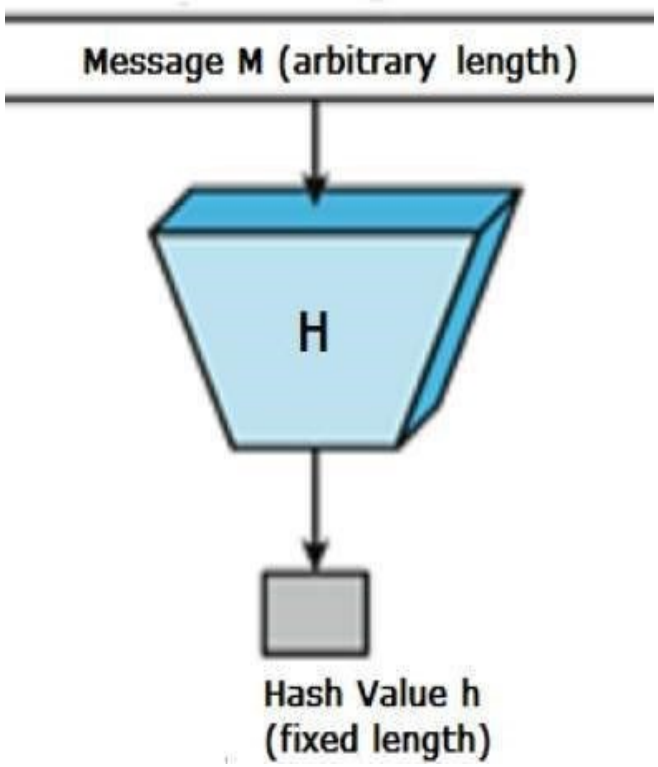
Step 2: Sender uses your public key to encrypt the plaintext

Step 3: Sender gives the ciphertext to you

Step 4: Use your private key (and passphrase) to decrypt the ciphertext

Hash Function

- A cryptographic hash function is a hash function that takes an arbitrary block of data and returns a fixed-size bit string, the cryptographic hash value, such that any (accidental or intentional) change to the data will (with very high probability) change the hash value.
- The data to be encoded are often called the message, and the hash value is sometimes called the message digest or simply digest.



Cryptosystem

- A cryptosystem is an implementation of cryptographic techniques and their accompanying infrastructure to provide information security services.
- A cryptosystem is also referred to as a cipher system.
- The various components of a basic cryptosystem are as follows −
  - Plaintext
  - Encryption Algorithm
  - Ciphertext
  - Decryption Algorithm
  - Encryption Key
  - Decryption Key

Cryptanalysis

- 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.

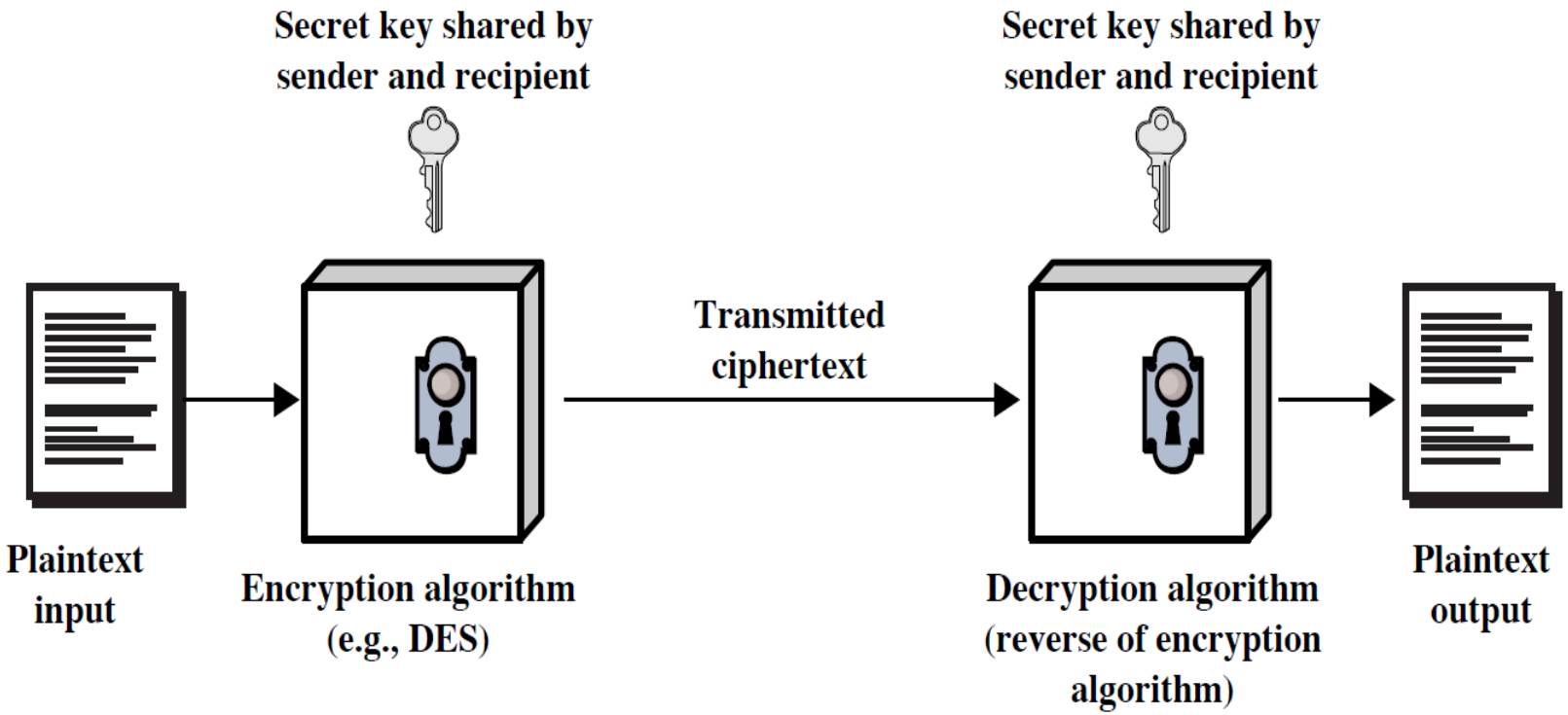
Classical encryption techniques

- As opposed to modern cryptography
- Goals:
  - to introduce basic concepts & terminology of encryption
  - to prepare us for studying modern cryptography

Basic terminology

- Plaintext: original message to be encrypted
- Ciphertext: the encrypted message
- Enciphering or encryption: the process of converting plaintext into ciphertext
- Encryption algorithm: performs encryption
  - Two inputs: a plaintext and a secret key

Symmetric Cipher Model



- Deciphering or decryption: recovering plaintext from ciphertext
- Decryption algorithm: performs decryption
  - Two inputs: ciphertext and secret key
- Secret key: same key used for encryption and decryption
  - Also referred to as a symmetric key

- Cipher or cryptographic system : a scheme for encryption and decryption
- Cryptography: science of studying ciphers
- Cryptanalysis: science of studying attacks against cryptographic systems
- Cryptology: cryptography + cryptanalysis

Ciphers

- Symmetric cipher: same key used for encryption and decryption
  - Block cipher: encrypts a block of plaintext at a time (typically 64 or 128 bits)
  - Stream cipher: encrypts data one bit or one byte at a time
- Asymmetric cipher: different keys used for encryption and decryption

Symmetric Encryption

- or conventional / secret-key / single-key
- sender and recipient share a common key
- all classical encryption algorithms are symmetric
- The only type of ciphers prior to the invention of asymmetric-key ciphers in 1970's
- by far most widely used

Symmetric Encryption

- Mathematically:

  $Y = E_K(X)$ or $Y = E(K, X)$

  $X = D_K(Y)$ or $X = D(K, Y)$
- X = plaintext
- Y = ciphertext
- K = secret key
- E = encryption algorithm
- D = decryption algorithm
- Both E and D are known to public

Cryptanalysis

- Objective: to recover the plaintext of a ciphertext or, more typically, to recover the secret key.
- Kerkhoff's principle: the adversary knows all details about a cryptosystem except the secret key.
- Two general approaches:
  - brute-force attack
  - non-brute-force attack (cryptanalytic attack)

Brute-Force Attack

- Try every key to decipher the ciphertext.
- On average, need to try half of all possible keys
- Time needed proportional to size of key space

Key Size (bits)	Number of Alternative Keys	Time required at 1 decryption/μs	Time required at 10^6 decryptions/μs
32	$2^{32} = 4.3 \times 10^9$	2^{31} μs = 35.8 minutes	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	2^{55} μs = 1142 years	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	2^{127} μs = 5.4×10^{24} years	5.4×10^{18} years
168	$2^{168} = 3.7 \times 10^{50}$	2^{167} μs = 5.9×10^{36} years	5.9×10^{30} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	2×10^{26} μs = 6.4×10^{12} years	6.4×10^6 years

Cryptanalytic Attacks

- Types are based on the amount of information known to the cryptanalyst

Types of Attack	***Known to Cryptanalyst***
Ciphertext Only	• Encryption algorithm • Ciphertext
Known Plaintext	• Encryption algorithm • Ciphertext • One or more plaintext–ciphertext pairs formed with the secret key
Chosen Plaintext	• Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key
Chosen Ciphertext	• Encryption algorithm • Ciphertext • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key
Chosen Text	• Encryption algorithm • Ciphertext • Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key • Ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key

Example: chosen-plaintext attack

- In 1942, US Navy cryptanalysts discovered that Japan was planning an attack on “AF”.
- They believed that “AF” means Midway island.
- Pentagon didn’t think so.
- US forces in Midway sent a plain message that their freshwater supplies were low.
- Shortly, US intercepted a Japanese ciphertext saying that “AF” was low on water.
- This proved that “AF” is Midway.

Defending Attacks! How easy?

- Ciphertext-only attack:
  - the ***easiest to defend*** against because the opponent has the least amount of information to work with.
  - Only relatively weak algorithms fail to withstand a ciphertext-only attack
- Known plaintext:
  - analyst may be able to capture one or more plaintext messages as well as their encryptions. Or the ***analyst may know certain plaintext patterns*** in a message
  - the analyst ***may be able to deduce the key*** on the basis of the way in which the known plaintext is transformed
  - Generally, an encryption algorithm is designed to withstand a known-plaintext attack
- Probable-word attack: (similar to known plaintext)
  - If the opponent is working with the encryption of some general prose message, he or she ***may have little knowledge of what is in the message***. However, if the opponent is after some very specific information, then ***parts of the message may be known.***
- Chosen-plaintext attack:
  - the analyst is able somehow to get the source system to insert into the system a message chosen by the analyst
  - if the analyst is able to choose the messages to encrypt, the analyst ***may pick patterns*** that can be expected to ***reveal the structure of the key***.
- Chosen ciphertext and Chosen text:
  - These are less commonly employed as cryptanalytic techniques but are nevertheless possible avenues of attack

Classical Ciphers

- Plaintext is viewed as a sequence of elements (e.g., bits or characters)
- Substitution cipher: replacing each element of the plaintext with another element.
- Transposition (or permutation) cipher: rearranging the order of the elements of the plaintext.
- Product cipher: using multiple stages of substitutions and transpositions

Caesar Cipher

- Earliest known substitution cipher
- Invented by Julius Caesar
- Each letter is replaced by the letter three positions further down the alphabet.
- 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
- Example: ohio state → RKLR VWDWH

Caesar Cipher

- Mathematically, map letters to numbers:

  a, b, c, ..., x, y, z

  0, 1, 2, ..., 23, 24, 25

- Then the general Caesar cipher is:

  $c = E_K(p) = (p + k) \bmod 26$

  $p = D_K(c) = (c - k) \bmod 26$

- Can be generalized with any alphabet.

Cryptanalysis of Caesar Cipher

- Key space: {0, 1, ..., 25}
- Vulnerable to brute-force attacks.
- E.g., break ciphertext "UNOU YZGZK“

- Need to recognize it when have the plaintext
- What if the plaintext is written in Swahili?

Monoalphabetic Substitution Cipher

- Shuffle the letters and map each plaintext letter to a different random ciphertext letter:

  Plain letters: abcdefghijklmnopqrstuvwxyz

  Cipher letters: DKVQFIBJWPESCXHTMYAUOLRGZN

  Plaintext: ifwewishtoreplaceletters

  Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA

- What does a key look like?

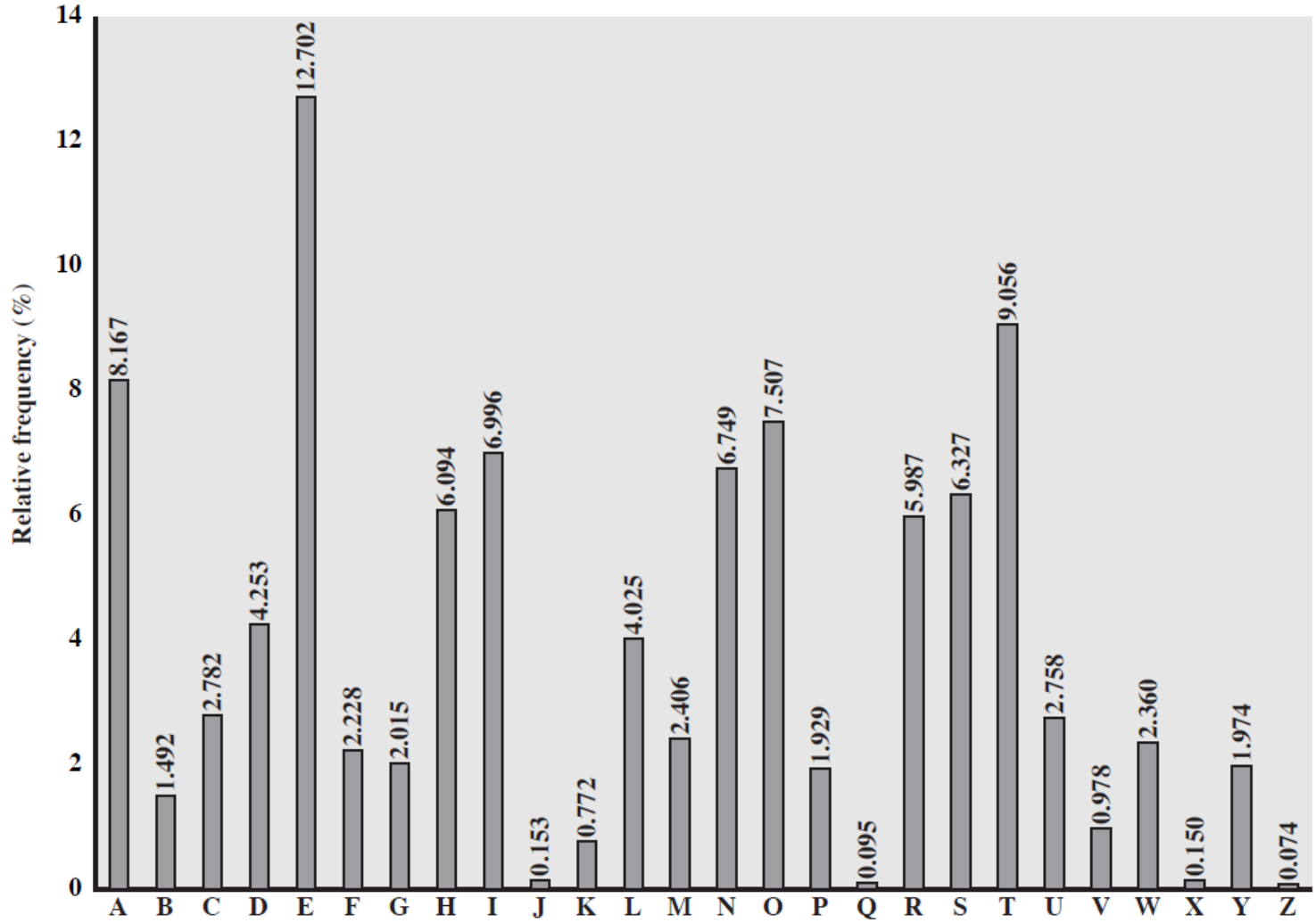
Monoalphabetic Cipher Security

- Now we have a total of 26! = 4 x 10^{26} keys.
- With so many keys, it is secure against brute-force attacks.
- But not secure against some cryptanalytic attacks.
- Problem is language characteristics.

Language Statistics and Cryptanalysis

- Human languages are not random.
- Letters are not equally frequently used.
- In English, E is by far the most common letter, followed by T, R, N, I, O, A, S.
- Other letters like Z, J, K, Q, X are fairly rare.
- There are tables of single, double & triple letter frequencies for various languages

English Letter Frequencies



Statistics for double & triple letters

- In decreasing order of frequency
- Double letters:

  th he an in er re es on, …
- Triple letters:

  the and ent ion tio for nde, …

Use in Cryptanalysis

- Key concept: monoalphabetic substitution does not change relative letter frequencies

- To attack, we
  - calculate letter frequencies for ciphertext
  - compare this distribution against the known one

Example Cryptanalysis

- Given ciphertext:

  UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
  VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
  EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

- Count relative letter frequencies (see next page)
- Guess {P, Z} = {e, t}
- Of double letters, ZW has highest frequency, so guess ZW = th and hence ZWP = the
- Proceeding with trial and error finally get:

  it was disclosed yesterday that several informal but direct contacts have been made with political representatives of the viet cong in moscow

Letter frequencies in ciphertext

P 13.33	H 5.83	F 3.33	B 1.67	C 0.00
Z 11.67	D 5.00	W 3.33	G 1.67	K 0.00
S 8.33	E 5.00	Q 2.50	Y 1.67	L 0.00
U 8.33	V 4.17	T 2.50	I 0.83	N 0.00
O 7.50	X 4.17	A 1.67	J 0.83	R 0.00
M 6.67				

Polyalphabetic Substitution Ciphers

- A sequence of monoalphabetic ciphers (M_1, M_2, M_3, ..., M_k) is used in turn to encrypt letters.
- A key determines which sequence of ciphers to use.
- Each plaintext letter has multiple corresponding ciphertext letters.
- This makes cryptanalysis harder since the letter frequency distribution will be flatter.

Playfair Cipher

- Not even the large number of keys in a monoalphabetic cipher provides security.
- One approach to improving security is to encrypt multiple letters at a time.
- The **Playfair Cipher** is the best known such cipher.
- Invented by Charles Wheatstone in 1854, but named after his friend Baron Playfair.

Playfair Key Matrix

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

- Unlike traditional cipher we encrypt a pair of alphabets(digraphs) instead of a single alphabet.
- Use a 5 x 5 matrix.
- Fill in letters of the key (w/o duplicates).
- Fill the rest of matrix with other letters.
- E.g., key = MONARCHY.

Encrypting and Decrypting

Plaintext is encrypted two letters at a time.

1. If a pair is a repeated letter, insert filler like 'X’.
2. If both letters fall in the same row, replace each with the letter to its right (circularly).
3. If both letters fall in the same column, replace each with the the letter below it (circularly).
4. Otherwise, each letter is replaced by the letter in the same row but in the column of the other letter of the pair.

Playfair Example

- Message = Move forward
- Plaintext = mo ve fo rw ar dx
- Here x is just a filler, message is padded and segmented
- **mo ->ON;**

  ve->UF;

  fo ->PH and so on.
- Ciphertext = **ON UF PH NZ RM BZ**

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**

Security of Playfair Cipher

- Equivalent to a monoalphabetic cipher with an alphabet of 26 x 26 = 676 characters.
- Security is much improved over the simple monoalphabetic cipher.
- Was widely used for many decades
  - eg. by US & British military in WW1 and early WW2
- Once thought to be unbreakable.
- Actually, it **can** be broken, because it still leaves some structure of plaintext intact.

Hill Cipher

- Hill cipher is a polygraphic substitution cipher based on linear algebra.
- 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.

Hill 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.

Hill Cipher

- 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).

  Input : Plaintext: ACT

  Key: GYBNQKURP

  Output : Ciphertext: POH

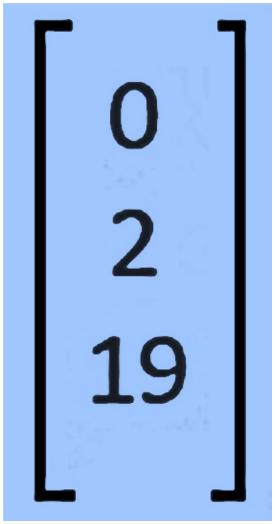
Hill Cipher

- We have to encrypt the message ‘ACT’ (n=3).
- The key is ‘GYBNQKURP’ which can be written as the nxn matrix:

$$\begin{bmatrix} 6 & 24 & 1 \\ 13 & 16 & 10 \\ 20 & 17 & 15 \end{bmatrix}$$

Hill Cipher

- The message ‘ACT’ is written as vector:



Hill Cipher

- The enciphered vector is given as:

$$\begin{bmatrix} 6 & 24 & 1 \\ 13 & 16 & 10 \\ 20 & 17 & 15 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 19 \end{bmatrix} = \begin{bmatrix} 67 \\ 222 \\ 319 \end{bmatrix} \equiv \begin{bmatrix} 15 \\ 14 \\ 7 \end{bmatrix} \pmod{26}$$

- Which corresponds to ciphertext POH

Note: Mod 26 division means number%26 i.e. remainder after divide by 26

Hill Cipher

- To decrypt the message, we turn the ciphertext back into a vector, then simply multiply by the inverse matrix of the key matrix (IFKVIVVMI in letters).
- The inverse of the matrix used in the previous example is:

$$\begin{bmatrix} 6 & 24 & 1 \\ 13 & 16 & 10 \\ 20 & 17 & 15 \end{bmatrix}^{-1} \equiv \begin{bmatrix} 8 & 5 & 10 \\ 21 & 8 & 21 \\ 21 & 12 & 8 \end{bmatrix} \pmod{26}$$

Hill Cipher

- For the previous Ciphertext ‘POH’:

$$\begin{bmatrix} 8 & 5 & 10 \\ 21 & 8 & 21 \\ 21 & 12 & 8 \end{bmatrix} \begin{bmatrix} 15 \\ 14 \\ 7 \end{bmatrix} \equiv \begin{bmatrix} 260 \\ 574 \\ 539 \end{bmatrix} \equiv \begin{bmatrix} 0 \\ 2 \\ 19 \end{bmatrix} \pmod{26}$$

- Which gives back our plaintext ‘ACT’.

Vigenère Cipher

- Simplest polyalphabetic substitution cipher
- Consider the set of all Caesar ciphers:

  $\{ C_a, C_b, C_c, ..., C_z \}$
- Key: e.g. security
- Encrypt each letter using C_s, C_e, C_c, C_u, C_r, C_i, C_t, C_y in turn.
- Repeat from start after C_y.
- Decryption simply works in reverse.

Vigenère Cipher

- **Encryption**

  The plaintext(P) and key(K) are added modulo 26.

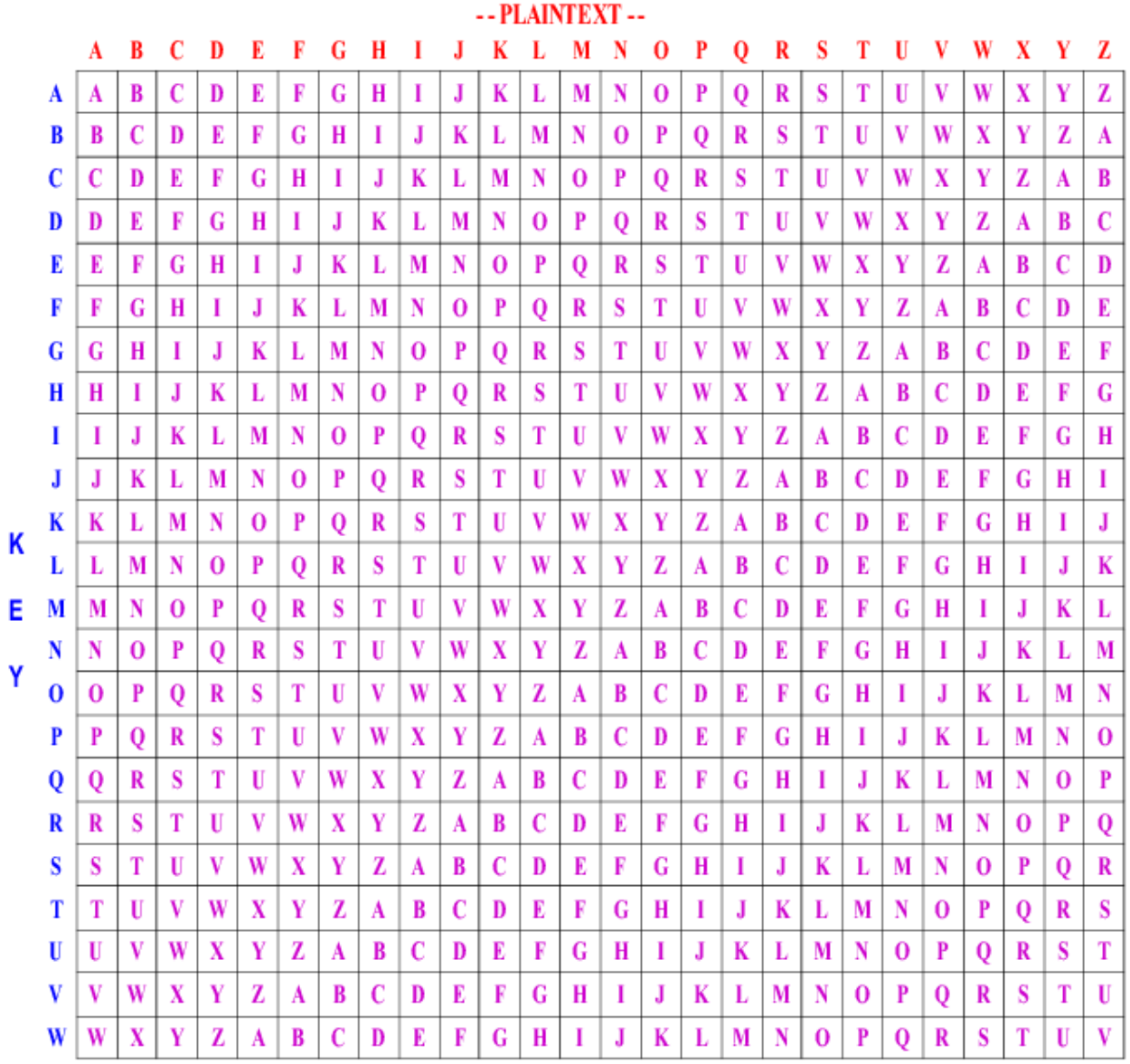
  $E_i = (P_i + K_i) \bmod 26$

- **Decryption**

  $D_i = (E_i - K_i + 26) \bmod 26$

Encryption Table

Vigenère Cipher



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

Example of Vigenère Cipher

- Keyword: *deceptive*

```
key:        deceptivedeceptivedeceptive
plaintext:  wearediscoveredsaveyourself
ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```

Expressed numerically, we have the following result.

key	3	4	2	4	15	19	8	21	4	3	4	2	4	15
plaintext	22	4	0	17	4	3	8	18	2	14	21	4	17	4
ciphertext	25	8	2	21	19	22	16	13	6	17	25	6	21	19

key	19	8	21	4	3	4	2	4	15	19	8	21	4
plaintext	3	18	0	21	4	24	14	20	17	18	4	11	5
ciphertext	22	0	21	25	7	2	16	24	6	11	12	6	9

Security of Vigenère Ciphers

- There are multiple (how many?) ciphertext letters corresponding to each plaintext letter.
- So, letter frequencies are obscured but not totally lost.
- To break Vigenere cipher:

1. Try to guess the key length. How? (See next slide)
2. If key length is N, the cipher consists of N Caesar ciphers. Plaintext letters at positions k, N+k, 2N+k, 3N+k, etc., are encoded by the same cipher.
3. Attack each individual cipher as before.

Guessing the Key Length

- Main idea: Plaintext words separated by multiples of the key length are encoded in the same way.
- In our example, if plaintext = "…thexxxxxxthe…" then "the" will be encrypted to the same ciphertext words.
- So look at the ciphertext for repeated patterns.
- E.g. repeated "VTW" in the previous example suggests a key length of 3 or 9:

```
ciphertext:   ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```

- Of course, the repetition could be a random fluke.

Vernam Cipher

- The ultimate defense against such a cryptanalysis is to choose a keyword that is as long as the plaintext and has no statistical relationship to it.
- Such a system was introduced by an AT&T engineer named Gilbert Vernam in 1918.
- His system works on binary data rather than letters.
- The system can be expressed succinctly as follows:
- $c_i = p_i \text{ XOR } k_i = p_i \oplus k_i$

Vernam Cipher

- P_i = ith binary digit of plaintext
- K_i = ith binary digit of key
- C_i = ith binary digit of ciphertext
- Thus, the ciphertext is generated by performing the bitwise XOR of the plaintext and the key. Because of the properties of the XOR, decryption simply involves the same bitwise operation:
- $p_i = c_i$ XOR k_i

Vernam Cipher

- **Encryption Algorithm:**

1. Assign a number to each character of the plain-text and the key according to alphabetical order.
2. Bitwise XOR both the number (Corresponding plain-text character number and Key character number).
3. Subtract the number from 26 if the resulting number is greater than or equal to 26, if it isn't then leave it.

Vernam Cipher

- **Plain-Text:** O A K
- **Key:** S O N
- **O ==>** 14 = 0 1 1 1 0
- **S ==>** 18 = 1 0 0 1 0
- **Bitwise XOR Result: 1 1 1 0 0 = 28**
- Since the resulting number is greater than 26, subtract 26 from it. Then convert the Cipher-Text character number to the Cipher-Text character.
- **28 - 26 = 2 ==> C CIPHER-TEXT: C**

Vernam Cipher

- Similarly, do the same for the other corresponding characters,
- **PT:** O A K
  NO: 14 00 10
- **KEY:** S O N
  NO: 18 14 13
- New Cipher-Text is after getting the corresponding character from the resulting number.
- **CT-NO:** 02 14 07
  CT: C O H

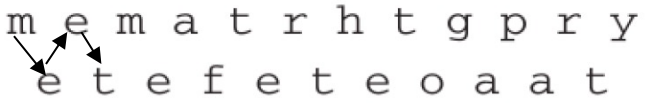
Transposition Ciphers

- Also called **permutation** ciphers.
- Shuffle the plaintext, without altering the actual letters used.
- Example: Rail Fence Cipher, Row Transposition Ciphers (Rectangular)

Rail Fence Cipher

A very different kind of mapping is achieved by performing some sort of ***permutation on the plaintext*** letters: called Transposition Techniques

In Rail Fence: The plaintext is written down as a ***sequence of diagonals*** and then read off as a sequence of rows



Plain text:
meetmeafterthetogaparty

The encrypted message is
MEMATRHTGPRYETEFETEOAAT

- Trivial to cryptanalyze

Row Transposition Ciphers (Rectangular Scheme)

- Plaintext is written row by row in a rectangle.
- Ciphertext: write out the columns in an order specified by a key.

Key: 3 4 2 1 5 6 7

a	t	t	a	c	k	p
o	s	t	p	o	n	e
d	u	n	t	i	l	t
w	o	a	m	x	y	z

Plaintext:

Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ

Product Ciphers

- Uses a sequence of substitutions and transpositions
  - Harder to break than just substitutions or transpositions
- This is a bridge from classical to modern ciphers.

Unconditional & Computational Security

- A cipher is unconditionally secure if it is secure no matter how much resources (time, space) the attacker has.
- A cipher is computationally secure if the best algorithm for breaking it will require so much resources (e.g., 1000 years) that practically the cryptosystem is secure.
- All the ciphers we have examined are not unconditionally secure.

One-Time Pad

- An Army Signal Corp officer, Joseph Mauborgne, proposed an improvement to the Vernam cipher.
- Mauborgne suggested using a random key that is as long as the message, so that the key need not be repeated.
- In addition, 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, and hence called one-time pad.

One-Time Pad

- It produces random output that bears no statistical relationship to the plaintext.
- Considered “Unbreakable” - As the ciphertext contains no information whatsoever about the plaintext, there is simply no way to break the code.

One-Time Pad

Suppose that we are using a ***Vigenère scheme with 27 characters*** in which the twenty-seventh character is the space character, but with a one-time key that is as long as the message

Consider the ciphertext :

ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS

Case 1:

ciphertext: ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS

key: pxlmvmsydofuyrvzwc tnlebnecvgdupahfzzlmnyih

plaintext: mr mustard with the candlestick in the hall

Case 2:

ciphertext: ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS

key: mfugpmiydgaxgoufhkllmhsqdqogtewbqfgyovuhwt

plaintext: miss scarlet with the knife in the library

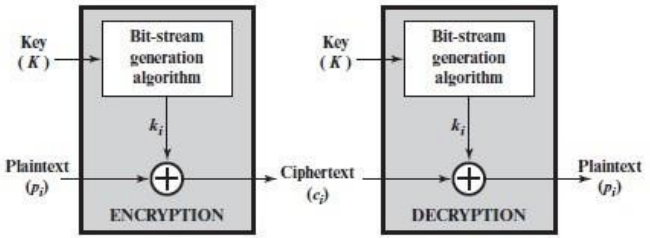
Question: which is the correct key? (*No answer*)

Modern Ciphers

- In Modern ciphers, digital data is represented in strings of binary digits (bits) unlike alphabets.
- 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.

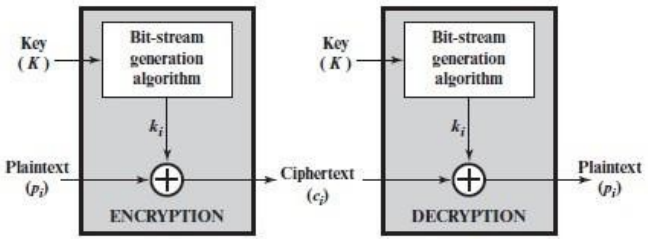
Stream Ciphers

- 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.



(a) Stream cipher using algorithmic bit-stream generator

Stream Ciphers

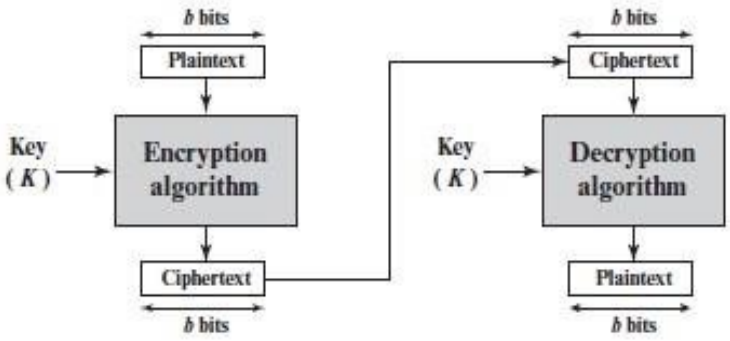


(a) Stream cipher using algorithmic bit-stream generator

- 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.

Block Ciphers

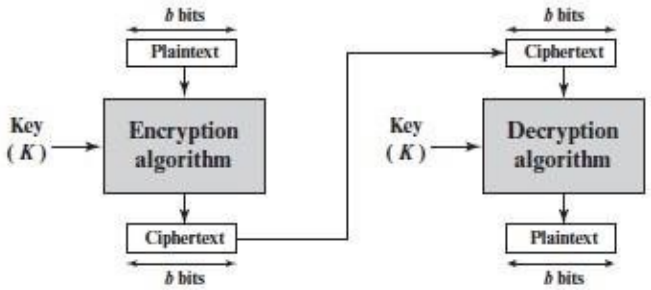
- A block cipher is one in which a block of plaintext is treated as a whole and used to produce a ciphertext block of equal length.
- Many block ciphers have a Feistel structure. Such a structure consists of a number of identical rounds of processing.
- Typically, a block size of 64 or 128 bits is used. As with a stream cipher, the two users share a symmetric encryption key.



(b) Block cipher

Block Ciphers

- A block cipher operates on a plaintext block of n bits to produce a ciphertext block of n bits.
- There are 2^n possible different plaintext blocks.
- Using some of the modes of operation (e.g. Cipher feedback mode or Output feedback mode) , a block cipher can also be used to achieve the same effect as a stream cipher.



(b) Block cipher

Block vs Stream Cipher

S.N.	BLOCK CIPHER	STREAM CIPHER
1	Block Cipher Converts the plain text into cipher text by taking plain text's block at a time.	Stream Cipher Converts the plaint text into cipher text by taking 1 byte (8 bits) of plain text at a time.
2	Block cipher uses either 64 bits or more than 64 bits.	Stream cipher uses 8 bits.
3	Simple design	Complex comparatively
4	Reversing encrypted text is hard.	Reversing encrypted text is comparatively easy.

Symmetric Key vs Assymetric Key Cipher

Characteristic	Symmetric Cryptography	Asymmetric Cryptography
Key used for encryption/decryption	Same key is used	One key is used for encryption and another for decryption
Speed of encryption/decryption	Very fast	Slower
Size of resulting encrypted text	Usually same as or less than the original plaintext size	More than the original plaintext size
Known keys	Both parties should know the key in symmetric key encryption	One of the keys is known by the two parties in public key encryption
Usage	Confidentiality	Confidentiality, Digital signature