Unit 2- Basics of Modern Cryptography

DR. VASUDHA ARORA

VASUDHA.ARORA@GDGU.ORG, VASUDHARORA6@GMAIL.COM

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

GD GOENKA UNIVERSITY, GURUGRAM

Some Basic Terminology

plaintext - original message

ciphertext - coded message

cipher - algorithm for transforming plaintext to ciphertext

key - info used in cipher known only to sender/receiver

encipher (encrypt) - converting plaintext to ciphertext

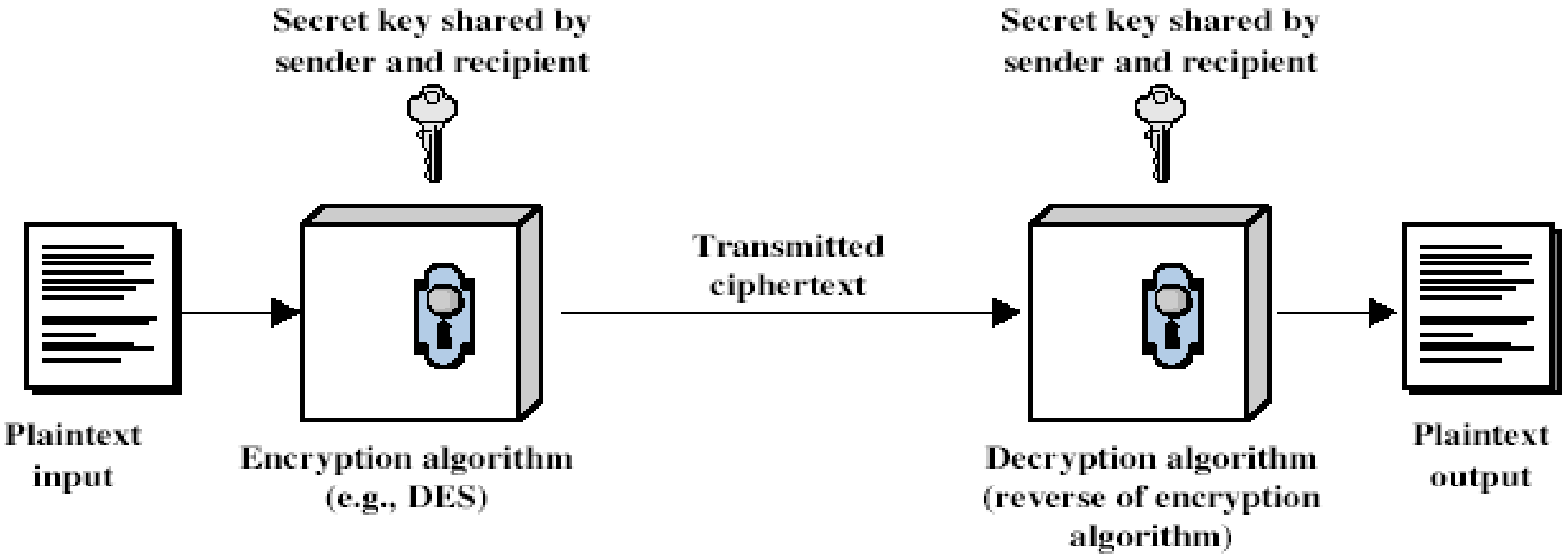
decipher (decrypt) - recovering ciphertext from plaintext

cryptography - study of encryption principles/methods

cryptanalysis (codebreaking) - study of principles/ methods of deciphering ciphertext *without* knowing key

cryptology - field of both cryptography and cryptanalysis

Symmetric Cipher Model/ Classical Cryptography



Requirements

two requirements for secure use of symmetric encryption:

- a strong encryption algorithm
- a secret key known only to sender / receiver

mathematically have:

$Y = E_K(X)$

$X = D_K(Y)$

- assume encryption algorithm is known
- implies a secure channel to distribute key

Cryptography

Cryptographic systems are characterized among three independent dimensions:

- **Type of operation used for transforming plaintext to ciphertext.**
  - substitution / transposition
- **The number of keys used**
  - single-key or private / two-key or public
- **The way in which plaintext is processed**.
  - block / stream

Type of operation used for transforming plaintext to ciphertext

All Encryption Algorithms are based on two general principles:

- **Substitution :** Each element in the plain text is mapped into another element.

e.g. Hello lkmmb

- **Transposition:** The elements in the plaintext are rearranged.

e.g. Hello lHole

The number of keys used

- If both sender and receiver use the same key (for both encryption and decryption), the system is referred to as **Symmetric key**, Secret key, **Private key** encryption system.

- If the sender and receiver use different keys (for encryption and decryption) the system is referred to as **asymmetric key**, two-key or **public key** encryption system.

The way in which plaintext is processed.

The plain text can be processed in two different ways to be converted into ciphertext:

- **Block Cipher:** A block cipher process the input one block of elements at a time, producing an output block for each input block.

- **Stream Cipher:** A stream cipher processes the input elements continuously, producing as an output one element at a time as it goes along.

Cryptanalysis

- objective to recover key not just message
- general approaches:
  - cryptanalytic attack
  - brute-force attack

Cryptanalytic Attacks

ciphertext only

- only know algorithm & ciphertext, is statistical, know or can identify plaintext

known plaintext

- know/suspect plaintext & ciphertext

chosen plaintext

- select plaintext and obtain ciphertext

chosen ciphertext

- select ciphertext and obtain plaintext

Brute Force Search

always possible to simply try every key

most basic attack, proportional to key size

Classical Substitution Ciphers

where letters of plaintext are replaced by other letters or by numbers or symbols

or

if plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with ciphertext bit patterns

Caesar Cipher

- Earliest known substitution cipher
- by Julius Caesar
- first attested use in military affairs
- replaces each letter by 3rd letter down the lane

example:

```
meet me after the class
PHHW PH DIWHU WKH FODVV
```

Caesar Cipher

- can define transformation as:

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
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

- mathematically give each letter a number

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

- then have Caesar cipher as:
  - $c = E(p) = (p + k) \bmod (26)$
  - $p = D(c) = (c - k) \bmod (26)$
- **Modified Caesar Cipher**

Cryptanalysis of Caesar Cipher

- only have 26 possible ciphers
  - A maps to A,B,..Z
- could simply try each in turn
- a **brute force search**
- given ciphertext, just try all shifts of letters

Monoalphabetic Cipher

- rather than just shifting the alphabet
- could shuffle (jumble) the letters arbitrarily
- each plaintext letter maps to a different random ciphertext letter
- hence key is 26 letters long

```
Plain:  abcdefghijklmnopqrstuvwxyz
Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN

Plaintext:  ifwewishtoreplaceletters
Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA
```

Monoalphabetic Cipher Security

- now have a total of 26! = 4 x 1026 keys
- with so many keys, might think is secure
- but would be **!!!WRONG!!!**
- problem is language characteristics

Language Redundancy and Cryptanalysis

- human languages are **redundant**
- eg "th lrd s m shphrd shll nt wnt"
- letters are not equally commonly 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
- have tables of single, double & triple letter frequencies for various languages

Playfair Cipher

- not even the large number of keys in a monoalphabetic cipher provides security
- one approach to improving security was to encrypt multiple letters
- the **Playfair Cipher** is an example
- invented by Charles Wheatstone in 1854, but named after his friend Baron Playfair

Playfair Key Matrix

- a 5X5 matrix of letters based on a keyword
- fill in letters of keyword (minus duplicates)
- fill rest of matrix with other letters in alphabetic order
- eg. using the keyword 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

Encrypting and Decrypting

plaintext is encrypted two letters at a time

- if a pair is a repeated letter, insert filler like ‘X’
- if both letters fall in the same row, replace each with letter to right (wrapping back to start from end)
- if both letters fall in the same column, replace each with the letter below it (again wrapping to top from bottom)
- otherwise each letter is replaced by the letter in the same row and in the column of the other letter of the pair

Security of Playfair Cipher

security much improved over monoalphabetic

since have 26 x 26 = 676 digrams

would need a 676 entry frequency table to analyse (verses 26 for a monoalphabetic)

and correspondingly more ciphertext

was widely used for many years

- eg. by US & British military in WW1

it **can** be broken, given a few hundred letters

since still has much of plaintext structure

Hill Cipher

Developed by a mathematician Lester Hill in 1929.

The Encryption Algorithm takes **m successive plaintext letters** and substitutes for them m ciphertext letters.

The substitution

The *key* for a hill cipher is a matrix e.g.

$$\begin{bmatrix} 2 & 4 & 5 \\ 9 & 2 & 1 \\ 3 & 17 & 7 \end{bmatrix}$$

Hill Cipher

Each letter is represented by a number modulo 26. Often the simple scheme A = 0, B = 1, …, Z = 25 is used

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).

Hill Cipher

For m = 3, the system is described as

$$\begin{pmatrix} C1 \\ C2 \\ C3 \end{pmatrix} = \begin{pmatrix} K11 & K12 & K13 \\ K21 & K22 & K23 \\ K31 & K32 & K33 \end{pmatrix} \begin{pmatrix} P1 \\ P2 \\ P3 \end{pmatrix} \quad \text{mod } 26$$

Or

C = K P mod 26

Where, C and P are column vectors of length 3 representing ciphertext and plaintext respectively.

K is a 3x3 matrix representing Encryption key

Hill Cipher

Example:

Input : Plaintext: ACTING IS AN ART

Key: GYBNQKURP

Output : Ciphertext: POH

For m=3, we create a 3x3 matrix

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

Hill Cipher

The first 3 letters of the message ‘ACT’ is written as vector:

$$\begin{bmatrix} 0 \\ 2 \\ 19 \end{bmatrix}$$

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 of ‘POH’

Hill Cipher

Continuing in same fashion for entire plaintext we get the complete ciphertext.

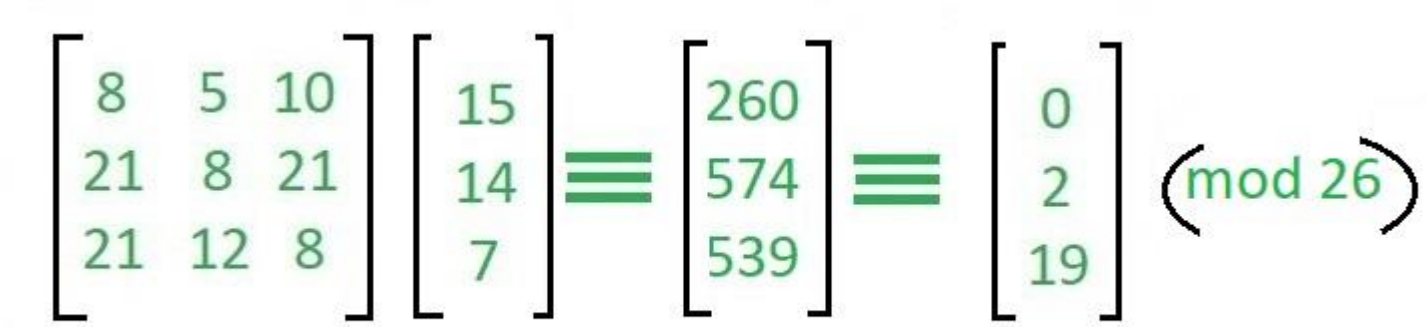
Decryption is done as $P = K^{-1} \bmod 26$

To decrypt the message, we turn the ciphertext back into a vector, then simply multiply by the inverse matrix of the key matrix

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

Hill Cipher

For the previous Ciphertext ‘POH



which gives us back ‘ACT’.

Polyalphabetic substitution ciphers

- It improves security using multiple cipher alphabets
- The development of Polyalphabetic Substitution Ciphers was the cryptographers answer to Frequency Analysis.
- uses a key to select which alphabet is used for each letter of the message
- use each alphabet in turn
- repeat from start after end of key is reached

Vigenère Cipher

- simplest polyalphabetic substitution cipher
- effectively multiple caesar ciphers
- key is multiple letters long $K = k_1\ k_2\ \ldots\ k_d$
- i^{th} letter specifies i^{th} alphabet to use
- use each alphabet in turn
- repeat from start after d letters in message
- decryption simply works in reverse

Vigenère Cipher

eg. given key *deceptive*

```
key:        deceptivedeceptivedeceptive
plaintext:  wearediscoveredsaveyourself
ciphertext:ZICVTWQNGKZEIIGASXSTSLVVWLA
```

mathematically give each letter a number

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Security of Vigenère Ciphers

- have multiple ciphertext letters for each plaintext letter
- hence letter frequencies are obscured but not totally lost

Autokey Cipher

- ideally want a key as long as the message
- Vigenère proposed the **autokey** cipher
- with keyword is prefixed to message as key
- knowing keyword can recover the first few letters
- use these in turn on the rest of the message
- but still have frequency characteristics to attack

eg. given key *deceptive*

```
key:        deceptivewearediscoveredsav
plaintext:  wearediscoveredsaveyourself
ciphertext:ZICVTWQNGKZEIIGASXSTSLVVWLA
```

One-Time Pad (OTP)

- if a truly random key as long as the message is used, the cipher will be secure
- called a One-Time pad
- is unbreakable since ciphertext bears no statistical relationship to the plaintext
- since for **any plaintext** & **any ciphertext** there exists a key mapping one to other
- can only use the key **once** though
- problems in generation & safe distribution of key

eg. given key *any random key*

```
Random key:dxcabnmolprtjhdfjhdufdfdjfi
plaintext: wearediscoveredsaveyourself
```

Transposition Ciphers

- Also Known as **permutation** ciphers
- these hide the message by rearranging the letter order
- without altering the actual letters used
- can recognise these since have the same frequency distribution as the original text

Rail Fence cipher

write message letters out diagonally over a number of rows

then read off cipher row by row

eg. `meet me after the party`

write message out as:

```
m e m a t r h p r y
  e t e f e t e a t
```

giving ciphertext

```
MEMATRHPRYETEFETEAT
```

Rail Fence cipher

Here, no. of rows over which text is distributed is the key for encryption

eg. `meet me after the party is over`

write message out in 3 rows as:

```
m  t  a  e  h  a  y  o  r
 e  m  f  r  e  r  i  v
  e  e  t  t  p  t  s  e
```

giving ciphertext

`MTAEHAYOREMFRERIVEETTPTSE`

Simple Columnar Transposition Technique

- Write the message row-by-row in a rectangle of pre defined size
- Read the message off column by column, but permute the order of the columns.
- The order of the columns then become the key to the algorithm.
- E.g. attack postpone until two am

```
Key:        3 4 1 2 5 6 7
Plaintext: 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
Ciphertext: TTNAAPTMAODWTSUOCOIKNLPET
```

Columnar Transposition with Multiple Rounds

Transposition cipher can be made more secure by performing more than one stage of transpositioning

```
Ciphertext: TTNAAPTMAODWTSUOCOIKNLPET
```

Obtained after Round 1

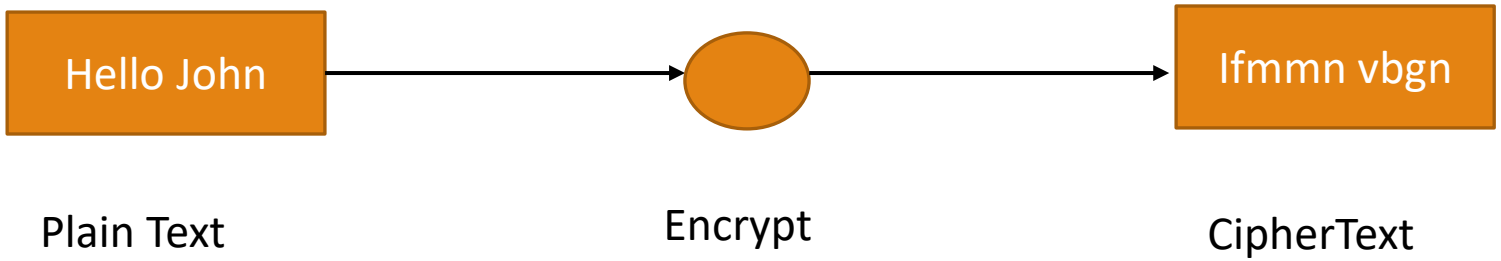
Key for Round 2 is same as Round 1.

```
Key:        3 4 1 2 5 6 7
Plaintext: t t n a a p t
           m a o d w t s
           u o c o I k n
           l p e t
Ciphertext: NOCEADOTTMULTAOPAWIPTKTSN
```

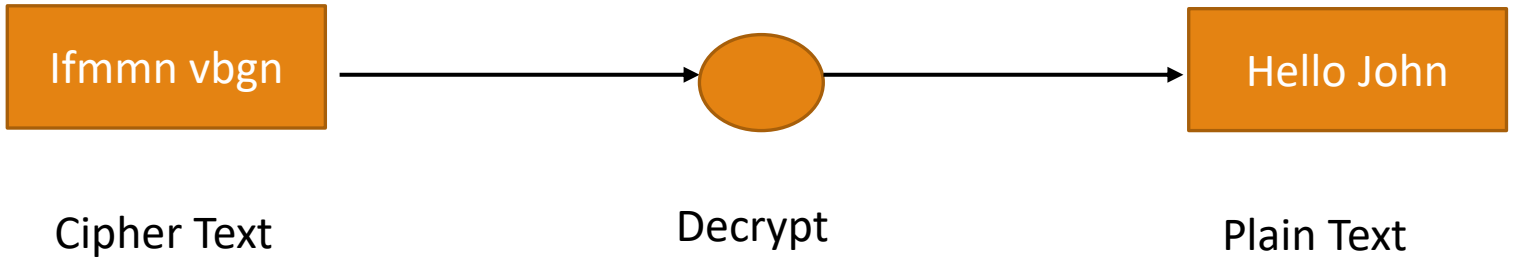
AND SO ON…….

Encryption and Decryption

- T The process of Encoding plain text messages into ciphertext messages is known as Encryption



- Decryption is exactly opposite of encryption. Decryption transforms a cipher text message back into a plain text message.



Encryption and Decryption

Every Encryption and decryption process has two aspects:

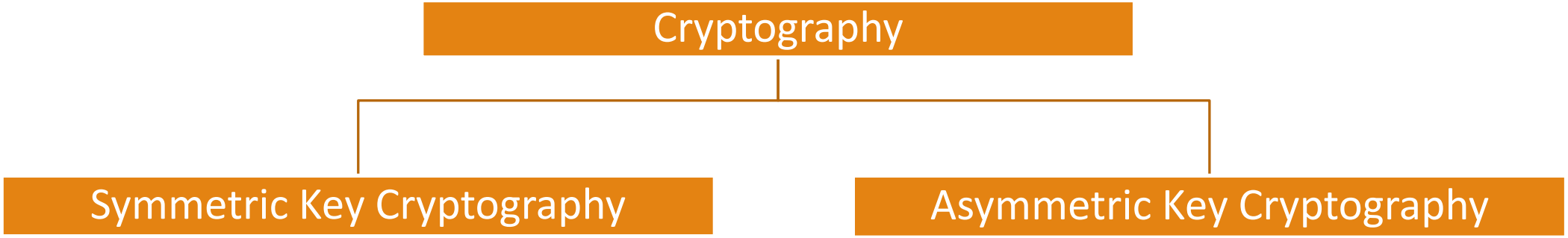
Algorithm

- In general Algorithm for Encryption or decryption is known to Everybody

Key

- The used for the Encryption and decryption makes the process of cryptography secure

Cryptography Techniques



➢If the same key is used for Encryption and decryption we call the mechanism as **Symmetric key cryptography.**

➢ If two different keys are used in a cryptographic mechanism, wherein one key is used for encryption and another different key is used for decryption, we call that mechanism as **Asymmetric key cryptography.**

Advantages of Symmetric Key Cryptography

- **Extremely Secure**
- **Relatively Fast**

Problems with Symmetric Key Cryptography

- **For n parties to communicate requires n(n-1)/2 keys.**

For example, if ten parties want to communicate with each other securely they would need 45 different key pairs: 10(10-1)/2 = 45. This would increase to 4,950 if there were 100 communicating parties!

- **Problem of key distribution: i.e. Sharing the Key between sender and receiver**

it was necessary for either the sender or the recipient to create a key and then send it to the other party. While the key was in transit, it could be stolen or copied by a third party who would then be able to decrypt any ciphertexts encrypted with that key.

Video

Advantages of Asymmetric Key Cryptography

ADVANTAGES

- In asymmetric or public key, cryptography there is no need for exchanging keys, thus eliminating the key distribution problem.
- Can provide digital signatures.

DISADVANTAGES

- A disadvantage of using public-key cryptography for encryption is speed: there are popular secret-key encryption methods which are significantly faster than any currently available public-key encryption method.