Unit -4 Cryptography-I

Dr. Lokesh Chouhan

Associate Professor

गृह मंत्रालय
MINISTRY OF
HOME AFFAIRS
सत्यमेव जयते

राष्ट्रीय न्यायालयिक विज्ञान विश्वविद्यालय
(राष्ट्रीय महत्त्व का संस्थान, गृह मंत्रालय, भारत सरकार)
National Forensic Sciences University
(An Institution of National Importance under Ministry of Home Affairs, Government of India)

NFSU
विद्या अमृतं अश्नुते

E-Mail: Lokeshchouhan@gmail.com, Lokesh.chouhan_goa@nfsu.ac.in

Mob: +91-898924399, 9827235155

Symmetric Encryption

or **conventional / private-key / single-key**

sender and recipient **share a common key**

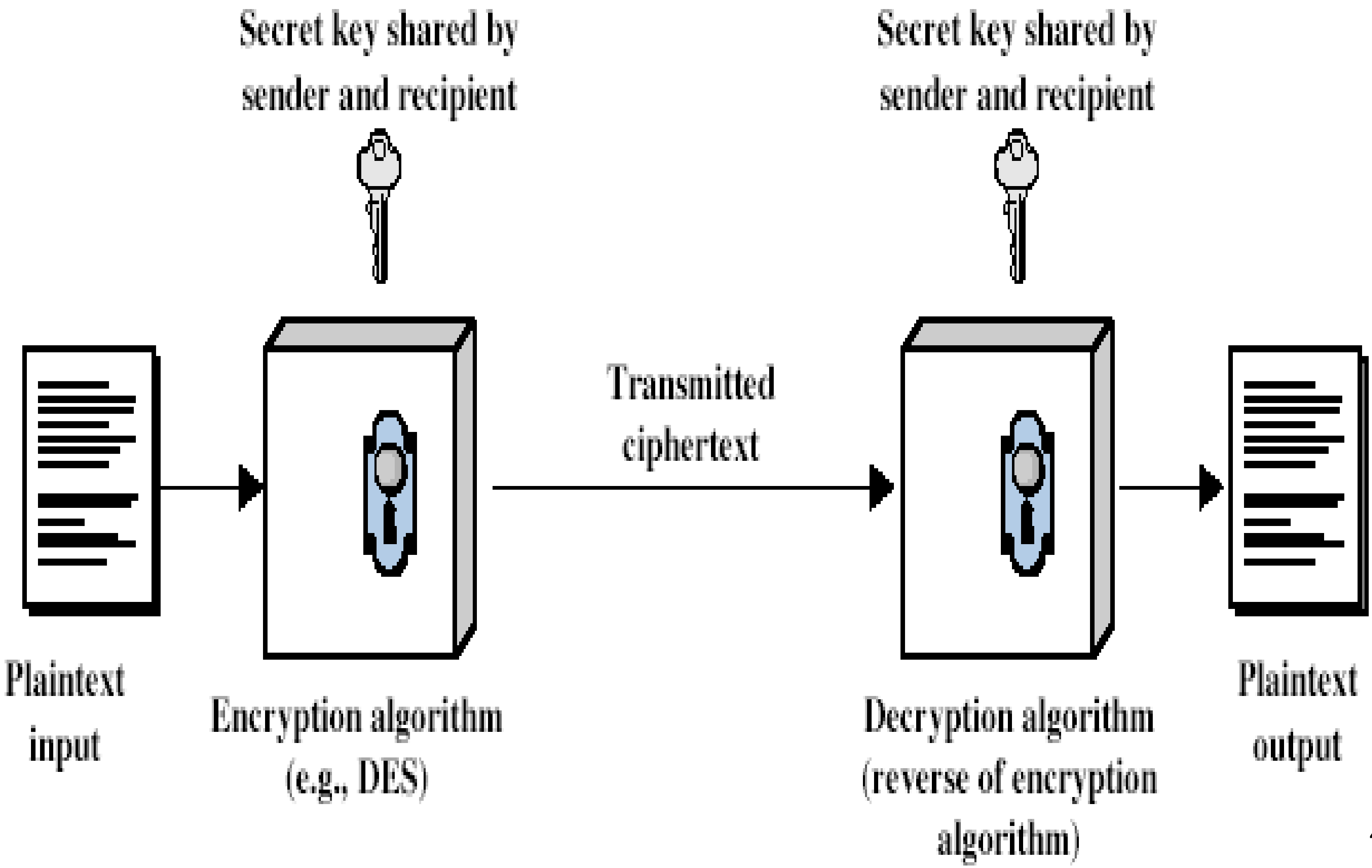
all classical encryption algorithms are **private-key**

was only type prior to invention of public-key in 1970's

Basic Terminology

plaintext	• the original message
ciphertext	• the coded message
cipher	• algorithm for transforming plaintext to ciphertext
key	• info used in cipher known only to sender/receiver
encipher (encrypt)	• converting plaintext to ciphertext
decipher (decrypt)	• recovering ciphertext from plaintext
cryptography	• study of encryption principles/methods
cryptanalysis (codebreaking)	• - the study of principles/ methods of deciphering ciphertext *without* knowing key
cryptology	• the field of both cryptography and cryptanalysis

Symmetric Cipher Model



Requirements

two requirements for secure use of symmetric encryption:

- a strong encryption algorithm
- a secret key known only to sender / receiver
  - $Y = \mathrm{E}_K(X)$
  - $X = \mathrm{D}_K(Y)$

assume encryption algorithm is known

implies a secure channel to distribute key

Cryptography

- **can characterize by:**
  - **type of encryption operations used**
    - substitution / transposition / product
  - **number of keys used**
    - single-key or private / two-key or public
  - **way in which plaintext is processed**
    - block / stream

Types of Cryptanalytic Attacks

ciphertext only

- only know algorithm / ciphertext, statistical, can identify plaintext

known plaintext

- know/suspect plaintext & ciphertext to attack cipher

chosen plaintext

- select plaintext and obtain ciphertext to attack cipher

chosen ciphertext

- select ciphertext and obtain plaintext to attack cipher

chosen text

- select either plaintext or ciphertext to en/decrypt to attack cipher

Brute Force Search

- always possible to **simply try every key**
- most basic attack, **proportional to key size**
- assume either **know / recognise plaintext**

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

More Definitions

unconditional security

- no matter how much computer power is available, the **cipher cannot be broken** since the ciphertext provides insufficient information to uniquely determine the corresponding plaintext

computational security

- given limited computing resources (eg time needed for calculations is greater than age of universe), the cipher cannot be broken

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 on
- **example:**

```
meet me after the toga party
PHHW PH DIWHU WKH WRJD SDUWB
```

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

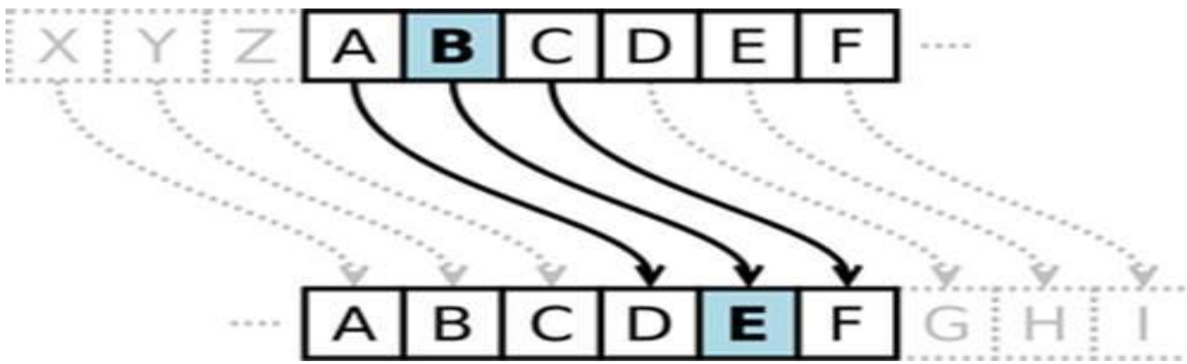
```
a b c d e f g h i j k  l  m
0 1 2 3 4 5 6 7 8 9 10 11 12
n  o  p  q  r  s  t  u  v  w  x  y  Z
13 14 15 16 17 18 19 20 21 22 23 24 25
```

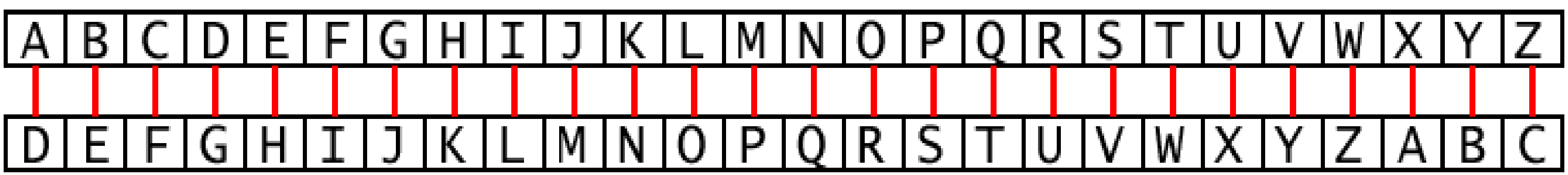
- then have **Caesar cipher** as:

$C = \mathrm{E}(p) = (p + k) \bmod (26)$

$p = \mathrm{D}(C) = (C - k) \bmod (26)$

Caesar Cipher





Alphabet shifted by 3 spaces.

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
- do need to recognize when have plaintext
- eg. break ciphertext "GCUA VQ DTGCM"

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
```

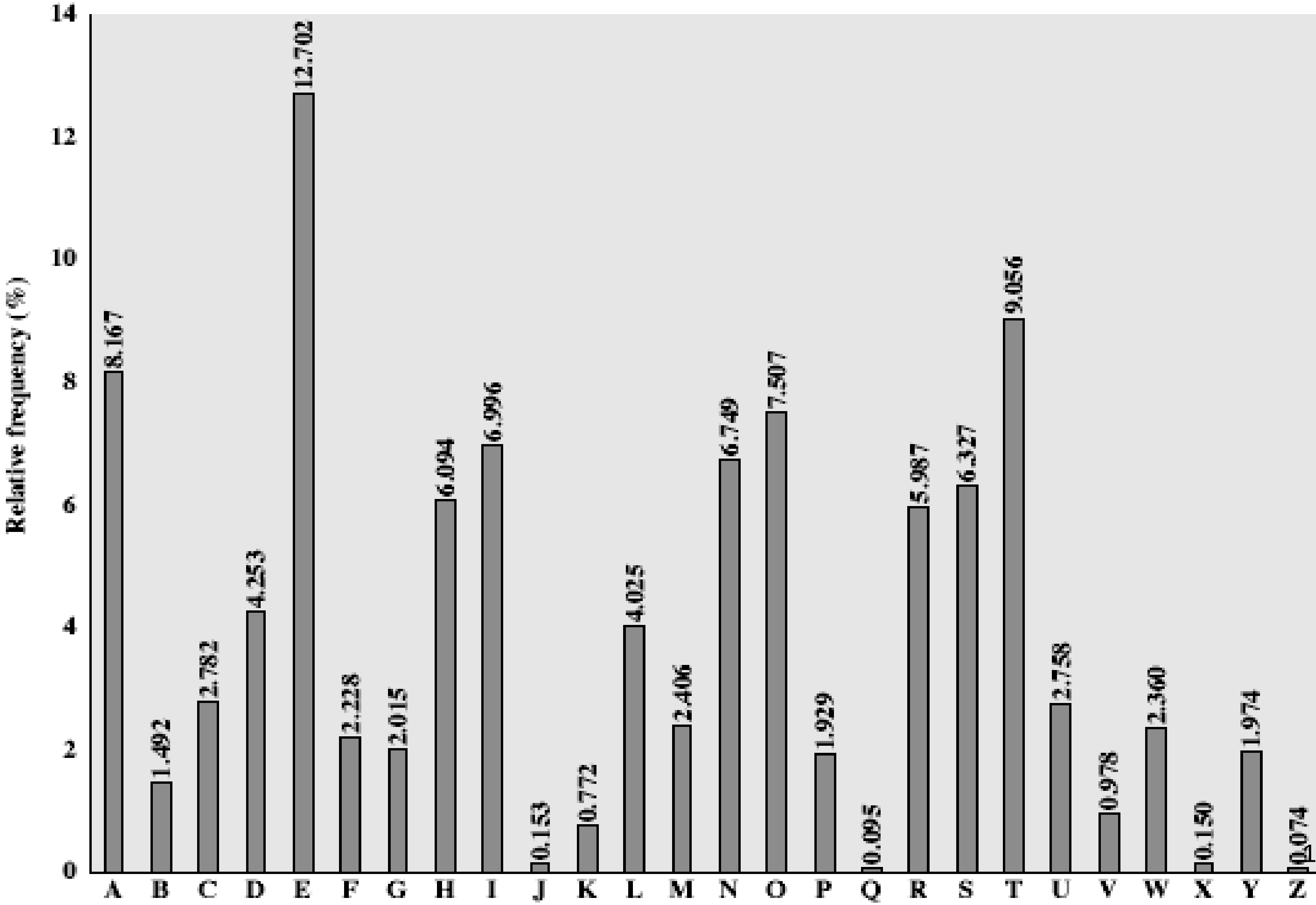
Monoalphabetic Cipher Security

- 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
- then **T,R,N,I,O,A,S**
- other letters are fairly rare
- cf. Z,J,K,Q,X
- have tables of single, double & triple letter frequencies

English Letter Frequencies



Use in Cryptanalysis

- **key concept** - monoalphabetic substitution ciphers do not change relative letter frequencies
- discovered by Arabian scientists in 9th century
- calculate letter frequencies for ciphertext
- compare counts/plots against known values
- if Caesar cipher look for common peaks/troughs
  - peaks at: A-E-I triple, NO pair, RST triple
  - troughs at: JK, X-Z
- for monoalphabetic must identify each letter
  - tables of common double/triple letters help

Example Cryptanalysis

- given ciphertext:

  UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
  VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
  EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

- count relative letter frequencies (see text)
- guess P & Z are **e** and **t**
- guess **ZW** is th and hence ZWP is 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

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 (sans duplicates)
- fill rest of matrix with other letters
- eg. using the keyword MONARCHY

```
MONAR
CHYBD
EFGIK
LPQST
UVWXZ
```

Encrypting and Decrypting

- plaintext encrypted two letters at a time:
  1. if a pair is a repeated letter, insert a filler like 'X', eg. "balloon" encrypts as "ba lx lo on"
  2. if both letters fall in the same row, replace each with letter to right (wrapping back to start from end), eg. “ar" encrypts as "RM"
  3. if both letters fall in the same column, replace each with the letter below it (again wrapping to top from bottom), eg. “mu" encrypts to "CM"
  4. otherwise each letter is replaced by the one in its row in the column of the other letter of the pair, eg. “hs" encrypts to "BP", and “ea" to "IM" or "JM" (as desired)

PLAYFAIR EXAMPLE

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

- Encrypting the message
  - "Hide the gold in the tree stump" (note the null "X" used to separate the repeated "E"s) :
- HI DE TH EG OL DI NT HE TR EX ES TU MP

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

HI

Shape: Rectangle

Rule: Pick Same Rows, Opposite Corners

BM

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

DE

Shape: Column

Rule: Pick Items Below Each Letter, Wrap to Top if Needed

OD

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

TH

Shape: Rectangle
Rule: Pick Same Rows, Opposite Corners

ZB

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

EG

Shape: Rectangle
Rule: Pick Same Rows,
Opposite Corners

XD

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

OL

Shape: Rectangle

Rule: Pick Same Rows, Opposite Corners

NA

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

EX

Shape: Row

Rule: Pick Items to Right of Each Letter, Wrap to Left if Needed

XM

More Examples

- Key:

Example

- Plaintext:

we will meet at the exit

Keyword Matrix

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

- Plaintext Pairs

we	wi	lx	lm	ex	et	at	th	ex	ex	it

<table>
<tr><th>Ciphertext Digraph</th><th colspan="5">Square</th><th>Rule</th><th>Plaintext Digraph</th></tr>
<tr><td rowspan="5">UA</td><td>E</td><td>K</td><td>A</td><td>M</td><td>P</td><td rowspan="5">Rule 4: Rectangle</td><td rowspan="5">we</td></tr>
<tr><td>L</td><td>B</td><td>C</td><td>D</td><td>F</td></tr>
<tr><td>G</td><td>H</td><td>I</td><td>K</td><td>N</td></tr>
<tr><td>O</td><td>Q</td><td>R</td><td>S</td><td>T</td></tr>
<tr><td>U</td><td>V</td><td>W</td><td>Y</td><td>Z</td></tr>
</table>

<table>
<tr><th>Ciphertext Digraph</th><th colspan="5">Square</th><th>Rule</th><th>Plaintext Digraph</th></tr>
<tr><td rowspan="5">AR</td><td>E</td><td>X</td><td>A</td><td>M</td><td>P</td><td rowspan="5">Rule 3: Same Column</td><td rowspan="5">wi</td></tr>
<tr><td>L</td><td>B</td><td>C</td><td>D</td><td>F</td></tr>
<tr><td>G</td><td>H</td><td>I</td><td>K</td><td>N</td></tr>
<tr><td>O</td><td>Q</td><td>R</td><td>S</td><td>T</td></tr>
<tr><td>U</td><td>V</td><td>W</td><td>Y</td><td>Z</td></tr>
</table>

Ciphertext Digraph	Square	Rule	Plaintext Digraph
BE	E X A M P L B C D F G H I K N O Q R S T U V W Y Z	Rule 4: Rectangle	lx

Ciphertext Digraph	Square	Rule	Plaintext Digraph
DE	E X A M P L B C D F G H I K N O Q R S T U V W Y Z	Rule 4: Rectangle	lm

<table>
<tr><th>Ciphertext Digraph</th><th>Square</th><th>Rule</th><th>Plaintext Digraph</th></tr>
<tr><td>XA</td><td><table>
<tr><td>E</td><td>X</td><td>A</td><td>M</td><td>P</td></tr>
<tr><td>L</td><td>B</td><td>C</td><td>D</td><td>F</td></tr>
<tr><td>G</td><td>H</td><td>I</td><td>K</td><td>N</td></tr>
<tr><td>O</td><td>Q</td><td>R</td><td>S</td><td>T</td></tr>
<tr><td>U</td><td>V</td><td>W</td><td>Y</td><td>Z</td></tr>
</table></td><td>Rule 2: Same Row</td><td>ex</td></tr>
</table>

Ciphertext Digraph	Square	Rule	Plaintext Digraph
PO	E X A M P L B C D F G H I K N O Q R S T U V W Y Z	Rule 4: Rectangle	et

<table>
<tr><th>Ciphertext Digraph</th><th colspan="5">Square</th><th>Rule</th><th>Plaintext Digraph</th></tr>
<tr><td rowspan="5">PR</td><td>E</td><td>X</td><td>A</td><td>M</td><td>P</td><td rowspan="5">Rule 4: Rectangle</td><td rowspan="5">at</td></tr>
<tr><td>L</td><td>B</td><td>C</td><td>D</td><td>F</td></tr>
<tr><td>G</td><td>H</td><td>I</td><td>K</td><td>N</td></tr>
<tr><td>O</td><td>Q</td><td>R</td><td>S</td><td>T</td></tr>
<tr><td>U</td><td>V</td><td>W</td><td>Y</td><td>Z</td></tr>
</table>

Ciphertext Digraph	Square	Rule	Plaintext Digraph
QN	E X A M P L B C D F G H I K N O Q R S T U V W Y Z	Rule 4: Rectangle	th

<table>
<tr><th>Ciphertext Digraph</th><th colspan="5">Square</th><th>Rule</th><th>Plaintext Digraph</th></tr>
<tr><td rowspan="5">XA</td><td>E</td><td>X</td><td>A</td><td>M</td><td>P</td><td rowspan="5">Rule 2: Same Row</td><td rowspan="5">ex</td></tr>
<tr><td>L</td><td>B</td><td>C</td><td>D</td><td>F</td></tr>
<tr><td>G</td><td>H</td><td>I</td><td>K</td><td>N</td></tr>
<tr><td>O</td><td>Q</td><td>R</td><td>S</td><td>T</td></tr>
<tr><td>U</td><td>V</td><td>W</td><td>Y</td><td>Z</td></tr>
</table>

<table>
<tr><th>Ciphertext Digraph</th><th colspan="5">Square</th><th>Rule</th><th>Plaintext Digraph</th></tr>
<tr><td rowspan="5">XA</td><td>E</td><td>X</td><td>A</td><td>M</td><td>P</td><td rowspan="5">Rule 2: Same Row</td><td rowspan="5">ex</td></tr>
<tr><td>L</td><td>B</td><td>C</td><td>D</td><td>F</td></tr>
<tr><td>G</td><td>H</td><td>I</td><td>K</td><td>N</td></tr>
<tr><td>O</td><td>Q</td><td>R</td><td>S</td><td>T</td></tr>
<tr><td>U</td><td>V</td><td>W</td><td>Y</td><td>Z</td></tr>
</table>

Ciphertext Digraph	Square	Rule	Plaintext Digraph
NR	E X A M P L B C D F G H I K N O Q R S T U V W Y Z	Rule 4: Rectangle	it

Example

- **Keyword: adjoin**
- **Message: Jammu and Kashmir becomes two UTs**
- **Find out the cipher text?**

Security of the 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. US & British military in WW1)
- it **can** be broken, given a few hundred letters
- since still has much of plaintext structure

Polyalphabetic Ciphers

- another approach to improving security is to use **multiple cipher alphabets**
- called **polyalphabetic substitution ciphers**
- makes cryptanalysis harder with more alphabets to guess and flatter frequency distribution
- use 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 is the **Vigenère Cipher**
- effectively **multiple caesar ciphers**
- key is multiple letters long **K = k1 k2 ... kd**
- 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

Example

- write the plaintext out
- write the keyword repeated above it
- use each key letter as a caesar cipher key
- encrypt the corresponding plaintext letter
- eg using keyword ***deceptive***
- **Message:** ***We are discovered save yourself***

key: deceptivedeceptivedeceptive

plaintext: wearediscoveredsaveyourself

ciphertext:ZICVTWQNGRZGVTWAVZHCQYGLMGJ

Aids

- simple aids can assist with en/decryption
- a **Saint-Cyr Slide** is a simple manual aid
  - a slide with repeated alphabet
  - line up plaintext 'A' with key letter, eg 'C'
  - then read off any mapping for key letter
- can bend round into a **cipher disk**
- or expand into a **Vigenère Tableau** (see text Table 2.3)

Security of Vigenère Ciphers

- have multiple ciphertext letters for each plaintext letter
- hence letter frequencies are obscured
- but not totally lost
- start with letter frequencies
  - see if look monoalphabetic or not
- if not, then need to determine number of alphabets, since then can attach each

Kasiski Method

- method developed by Babbage / Kasiski
- repetitions in ciphertext give clues to period
- so find same plaintext an exact period apart
- which results in the same ciphertext
- of course, could also be random fluke
- eg repeated "VTW" in previous example
- suggests size of 3 or 9
- then attack each monoalphabetic cipher individually using same techniques as before

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

- 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
- have problem of safe distribution of key

Transposition Ciphers

- now consider classical **transposition** or **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
- Message: meet me after toga party
- eg. write message out as:

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

- giving ciphertext

MEMATRHTGPRYETEFETEOAAT

Rail Fence cipher

- eg. write message out as:

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

- Ciphertext (redaing row one by one):
- **m e m a t r h t g p r y e t e f e t e o a a t**

- giving ciphertext

MEMATRHTGPRYETEFETEOAAT

Row Transposition Ciphers

- a more **complex scheme**
- write letters of message out in rows over a specified number of columns
- then reorder the columns according to some key before reading off the rows

Ex.

Message: attack postponed until two am

Key: **3 4 2 1 5 6 7**

Row Transposition Ciphers

Key:	3	4	2	1	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	x	y	z

attack postponed until two am

Row Transposition Ciphers

Key:	3	4	2	1	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	x	y	z

attack postponed until two am

Ciphertext: **TTNA APTM TSUO AODW COIX KNLY PETZ**

TTNA	APTM	TSUO	AODW	COIX	KNLY	PETZ
3	4	2	1	5	6	7

Product Ciphers

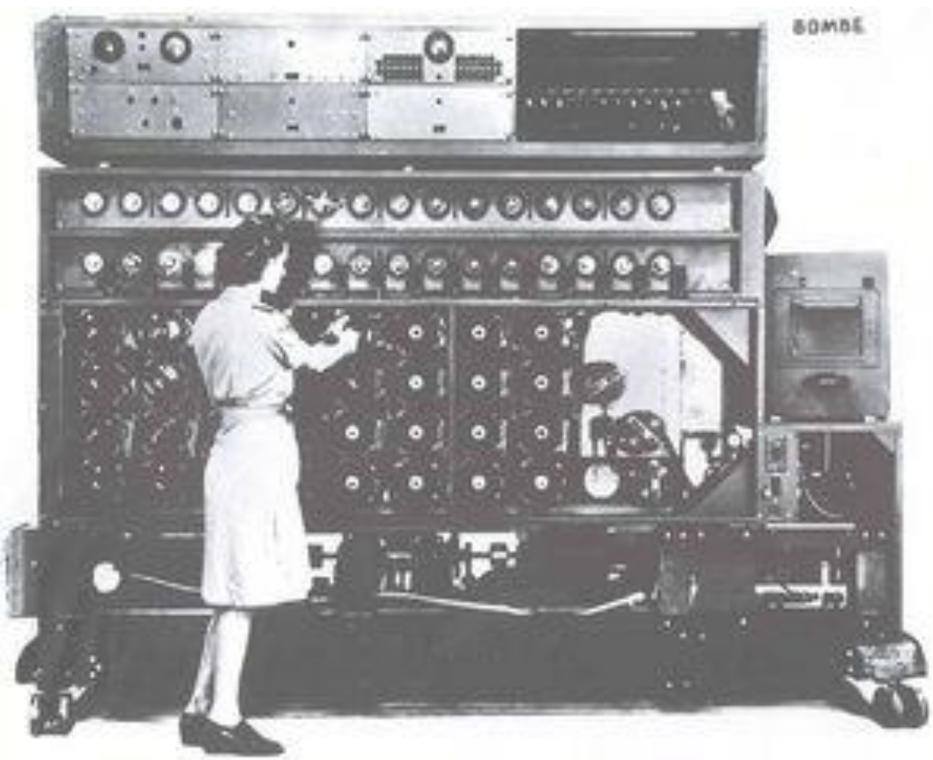
- ciphers using substitutions or transpositions are not secure because of language characteristics
- hence consider using several ciphers in succession to make harder, but:
  - two substitutions make a more complex substitution
  - two transpositions make more complex transposition
  - but a substitution followed by a transposition makes a new much harder cipher
- this is bridge from classical to modern ciphers

Rotor Machines

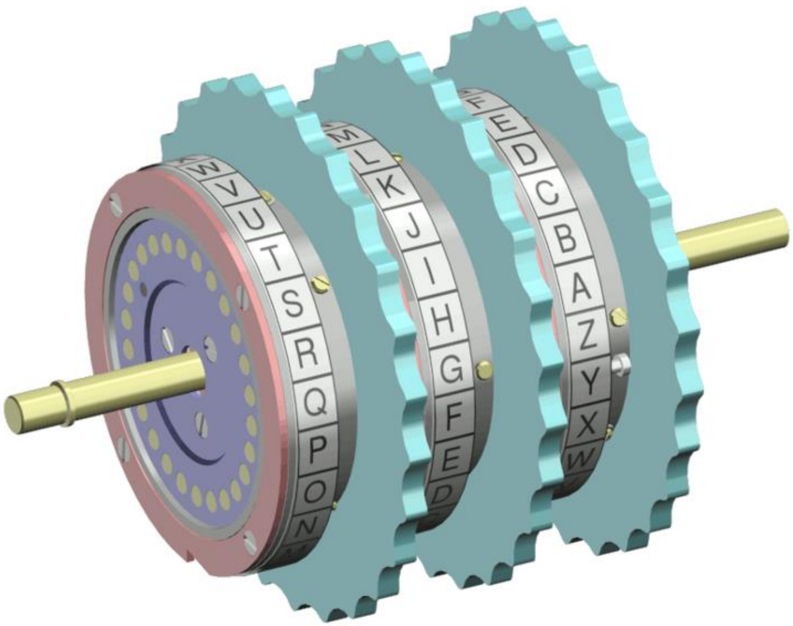
- before modern ciphers, rotor machines were most common product cipher
- were widely used in **WW2**
  - German Enigma, Allied Hagelin, Japanese Purple
- implemented a very complex, varying substitution cipher
- used a series of cylinders, each giving one substitution, which rotated and changed after each letter was encrypted
- with 3 cylinders have 26^3=17576 alphabets

Four-Rotor Enigma coding machine

- A **Bombe electromechanical codebreaking machine** built at the Naval Computing Machine Laboratory in the National Cash Register plant, Dayton, Ohio.
- The Bombe shown was in operation in **1943** at the Navy Communications Supplementary Activity on Nebraska Ave. in Washington, D.C.
- Bombes simulated the **rotor movents of the four-rotor Enigma coding machine** carried by German submarines.
- As the war progressed the Bombes were supplanted by much faster all-electronic machines with electronic ring counters simulating the rotors.

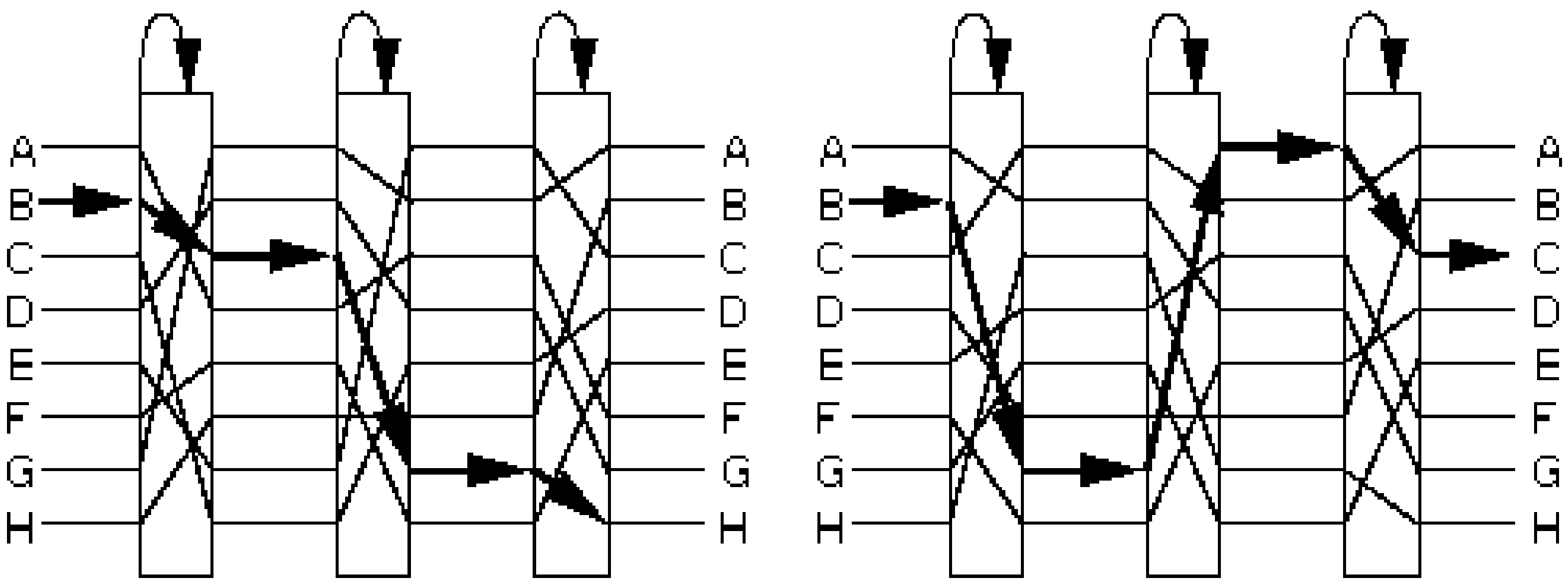


Rotor Machines



Modern Rotor Machines 2002

Rotor Machines



Steganography

- an alternative to encryption
- hides existence of message
  - using only a subset of letters/words in a longer message marked in some way
  - using invisible ink
  - hiding in LSB in graphic image or sound file
- has drawbacks
  - high overhead to hide relatively few info bits

Steganography

Latest Chanrayaan3 Image

1.

- 2

Summary

- have considered:
  - classical cipher techniques and terminology
  - monoalphabetic substitution ciphers
  - cryptanalysis using letter frequencies
  - Playfair ciphers
  - polyalphabetic ciphers
  - transposition ciphers
  - product ciphers and rotor machines
  - stenography