Classical Encryption Techniques

4CS401

M. K. Chavan,
Asst. Professor, CSE Department,
Walchand College of Engineering, Sangli

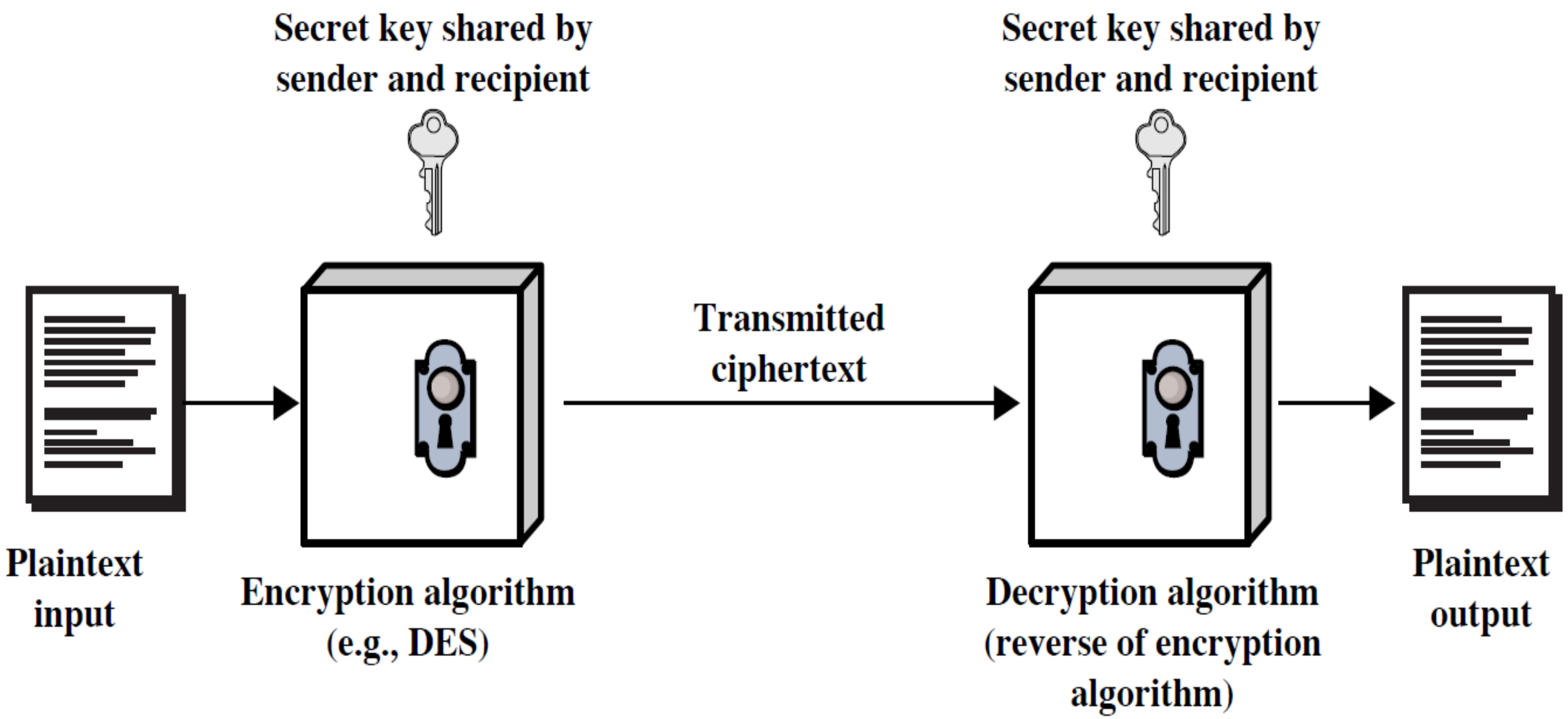
Classical encryption techniques

- 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.

- 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

Classical Cryptography

- Sender, receiver share common key
  - Keys may be the same, or trivial to derive from one another
  - Sometimes called *symmetric cryptography*
- Two basic types
  - Transposition ciphers
  - Substitution ciphers
  - Combinations are called *product ciphers*

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

Transposition Cipher

- Rearrange letters in plaintext to produce ciphertext
- Example (Rail-Fence Cipher or 2-columnar transposition)
  - Plaintext is `HELLO WORLD`
  - `HE`
`LL`
`OW`
`OR`
`LD`
  - Ciphertext is `HLOOL ELWRD`

Transposition Cipher

- Generalize to n-columnar transpositions
- Example 3-columnar
  - ```
    HEL
    LOW
    ORL
    DXX
    ```
  - `HLODEORXLWLX`
    ```

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 “DWWDFN“

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

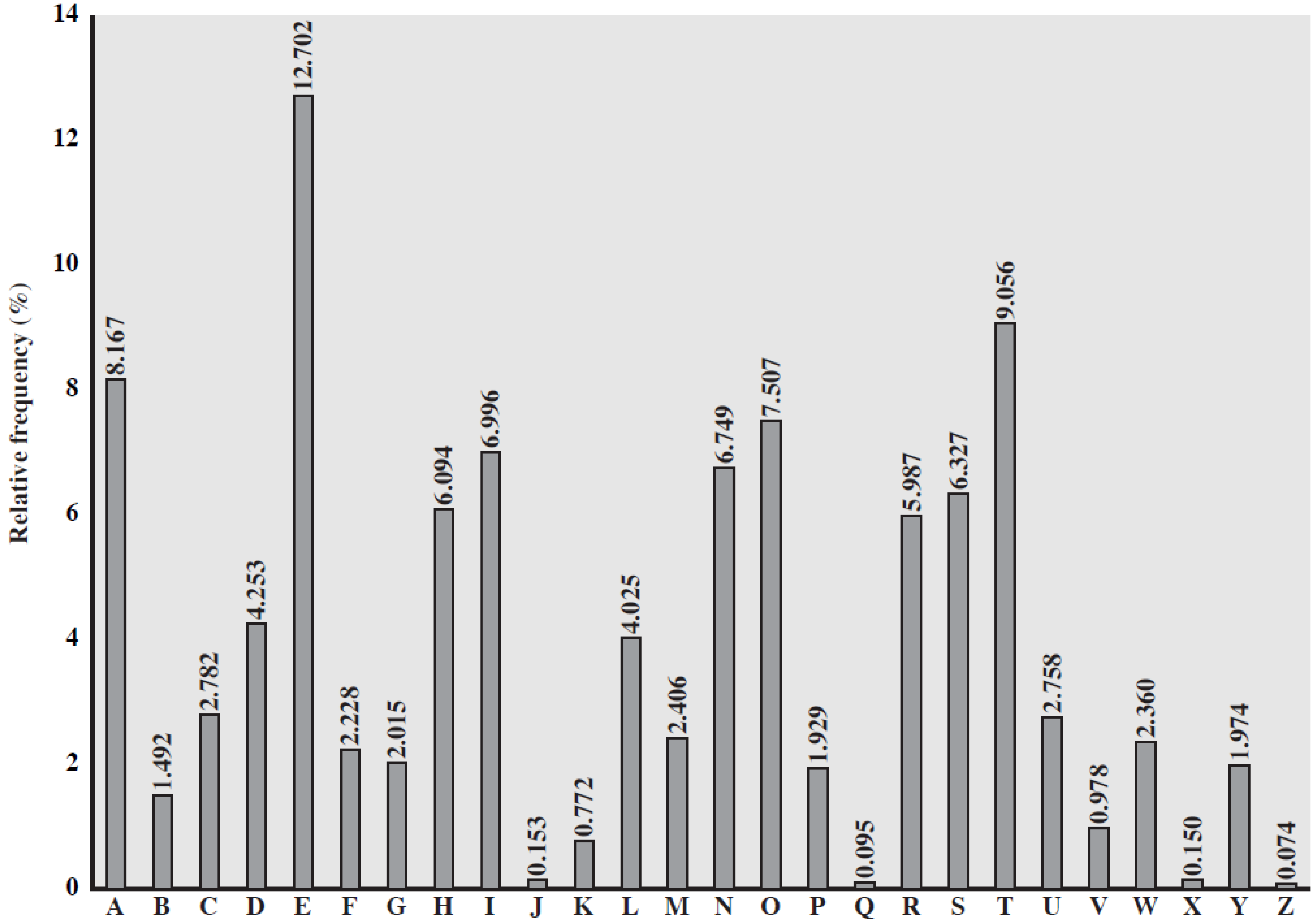
Monoalphabetic Cipher Security

- 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				

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

- 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.

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.

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.

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.

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.

Example of Vigenère Cipher

- Keyword: *deceptive*

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

C=(p+k) mod 26

Transposition Ciphers

- Also called **permutation** ciphers.
- Shuffle the plaintext, without altering the actual letters used.
- Example: Row Transposition Ciphers

Rail Fence cipher

- write message letters out diagonally over a number of rows
- then read off cipher row by row
- 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
```

Row Transposition Ciphers

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

Key: 4 3 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	x	y	z

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.

An unconditionally Secure Cipher

Vernam's one-time pad cipher

- Key = $k_1 k_2 k_3 k_4 \ldots$ (random, used one-time only)
- Plaintext = $m_1 m_2 m_3 m_4 \ldots$
- Ciphertext = $c_1 c_2 c_3 c_4 \ldots$

  where $c_i = m_i \oplus k_i$
- Can be proved to be unconditionally secure.

Summary

- Have considered:
  - classical cipher techniques and terminology
  - monoalphabetic substitution ciphers
  - cryptanalysis using letter frequencies
  - Playfair cipher
  - polyalphabetic ciphers
  - transposition ciphers
  - product ciphers