Traditional Symmetric-Key Ciphers

Dr. B C Dhara

Department of Information Technology

Jadavpur University

Objective of this chapter

- Define the terms and the concepts of symmetric key ciphers
- Emphasize the two categories of traditional ciphers: substitution and transposition ciphers
- Describe the categories of cryptanalysis used to break the symmetric ciphers
- Introduce the concepts of the stream ciphers and block ciphers
- Discuss some very dominant ciphers used in the past

Different terms of cryptography

- **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

Different terms of cryptography (contd…)

- **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

More Definitions

- **unconditional security**
  - no matter how much computer power or time is available, the cipher cannot be broken since the ciphertext provides insufficient information to uniquely determine the corresponding plaintext
  - A necessary condition for a symmetric-key encryption scheme to be unconditionally secure is that the key be at least as long as the message
  - The one-time pad is an example of an unconditionally secure encryption algorithm

Public-key encryption schemes cannot be unconditionally secure since, given a ciphertext c, the plaintext can in principle be recovered by encrypting all possible plaintexts until c is obtained.

computational security

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

Provable security

- A cryptographic method is said to be *provably secure if the difficulty of defeating it can be* shown to be essentially as difficult as solving a well-known and *supposedly difficult (typically* number-theoretic) problem, such as integer factorization or the computation of discrete logarithms

Brute Force Search

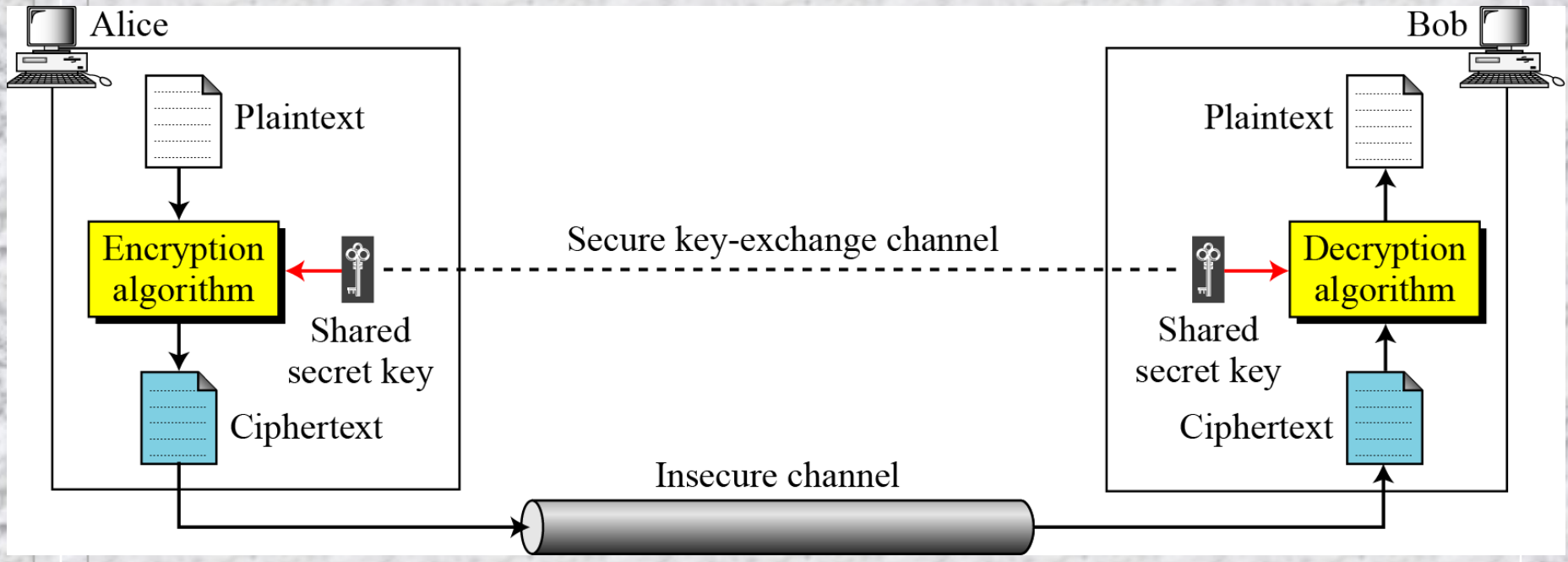
- Always possible to simply try every key
- Most basic attack, proportional to key size
- Assume either know / recognise plaintext

Key Size (bits)	Number of Alternative Keys	Time required at 1 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

Classification of cryptography

- Cryptography referred to encryption and decryption of messages using secret keys
  - Symmetric-key cryptography
  - Asymmetric-key cryptography

General idea of symmetric cipher



Intro to Cryptography

- 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

Intro to Cryptography (contd…)

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

Symmetric key …

Plain text P ➔ cipher text C= $E_k(P)$ [encryption process]

➔ $P=D_k(C)$ [decryption process]

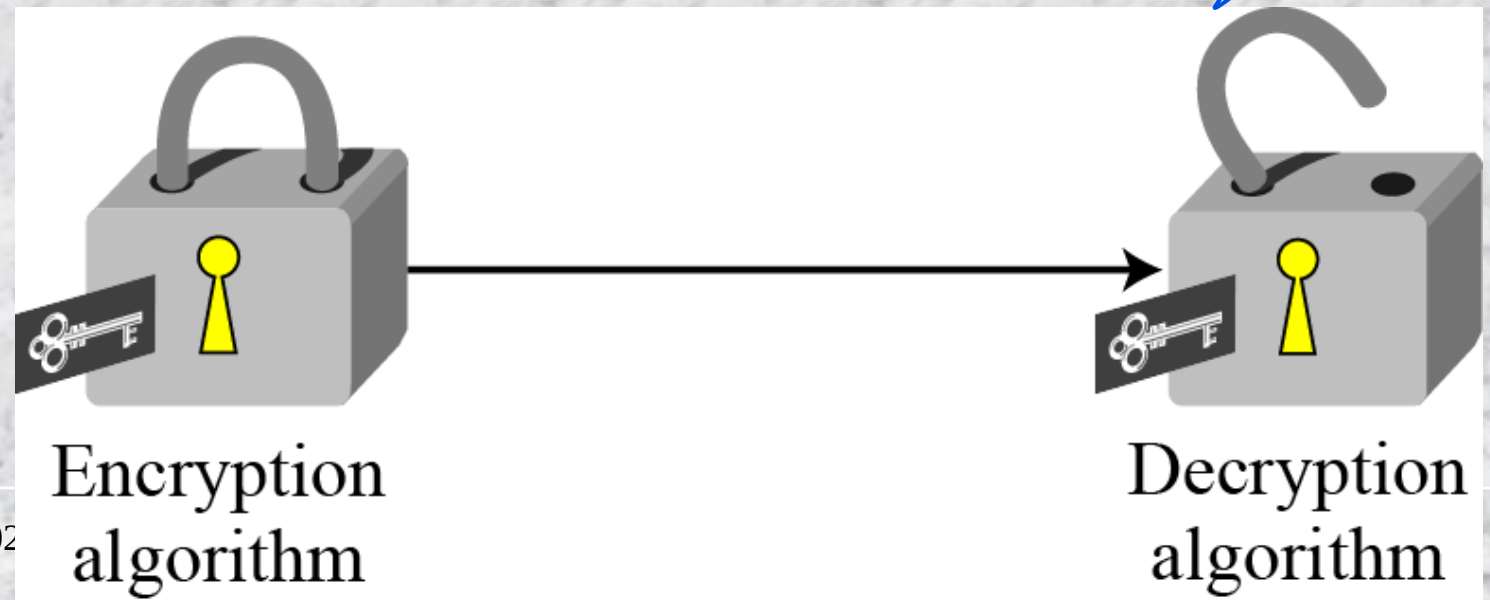
For any arbitrary message ‘x’

➔ $D_k(E_k(x)) = E_k(D_k(x)) = x$

- Suppose $P'= D_k(C) = D_k(E_k(P)) = P$

Symmetric key …

- Encryption can be thought as locking the message in a box, and
- Decryption can be thought as unlocking the box
  - In symmetric key system same key is used both for locking and unlocking



Kerckhoff's Principle

- A cipher text would be more secure if we hide both the encryption/decryption algorithm and the secret key
- According to Kerckhoff's principle:
  - one should always assume that the adversary, knows the encryption/decryption algorithm. The resistance of the cipher to attack must be based only on the secrecy of the key

Kerckhoff's Principle (contd…)

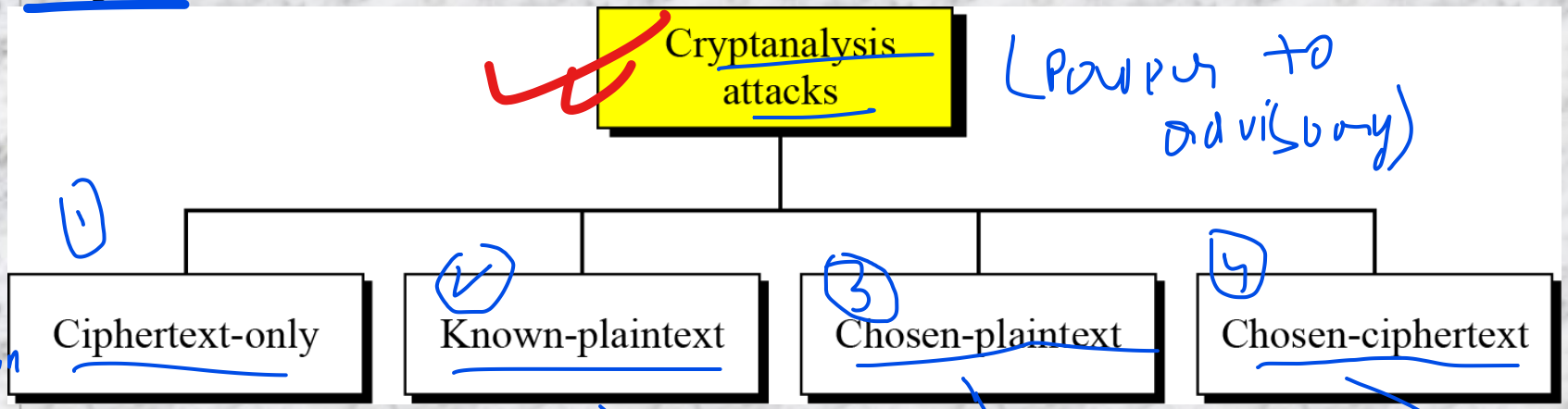
- Resistance to attack must be on the secrecy of the key
  - Guessing key should be so difficult that no need to hide the algorithms
  - Key domain should be so large that it makes difficult for an adversary to find the key

Cryptanalysis

- Cryptography is the science and art of creating secret codes, cryptanalysis is the science and art of breaking those codes
- objective to recover key not just message
  - general approaches:
    - cryptanalytic attack
    - brute-force attack
  - if either succeed all key use compromised

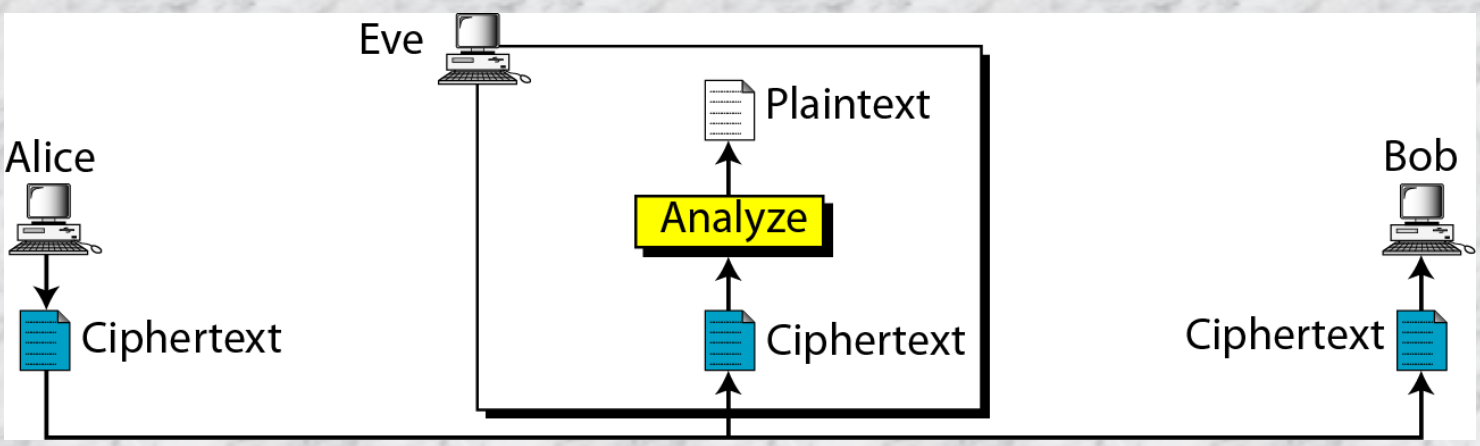
Cryptanalysis attacks

- **Attacks on encryption schemes**
- The objective of the following attacks is to systematically recover plaintext from ciphertext, or even more drastically, to deduce the decryption key.



Ciphertext-Only Attack

- Attacker has access to some ciphertext
  - Tries to find the corresponding key and the plaintext
  - This is most probable attack as only ciphertext is needed
  - Any encryption scheme vulnerable to this type of attack is considered to be completely insecure.



Common approaches in ciphertext only attack

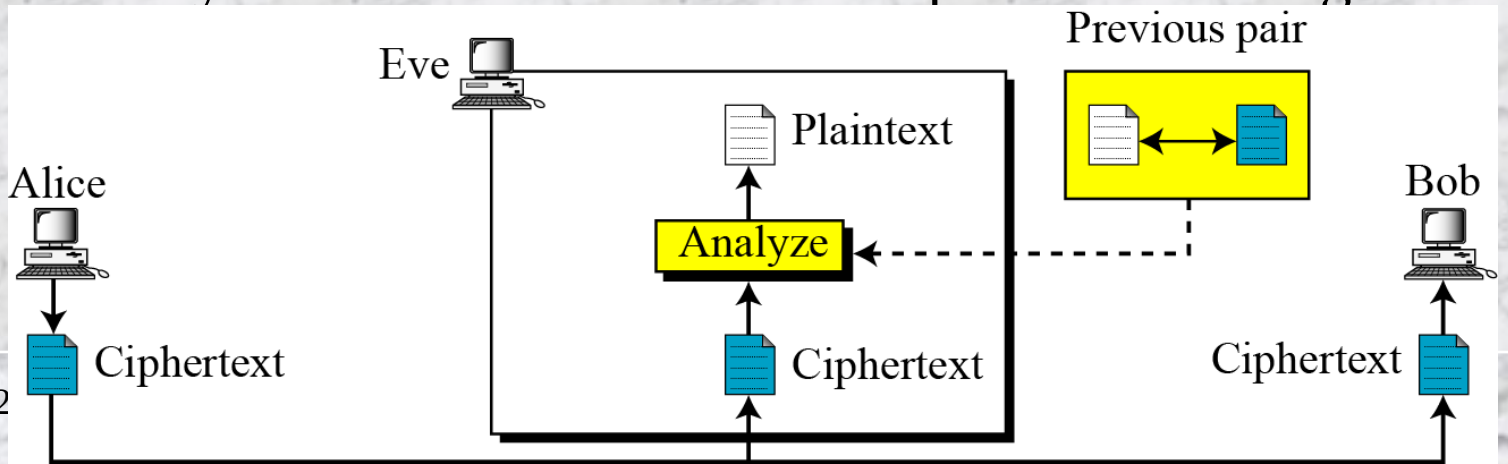
- Brute-force attack
  - Decrypt the ciphertext with every possible key until the plaintext makes sense
  - To prevent this attack, number of possible keys should be very large
- Statistical attack
  - Some characteristics of the plaintext language help in statistical attack
  - In English, ‘E’ is most frequently used character
    - Mostly available character in ciphertext may give ‘E’

Common approaches in ciphertext only attack

- Pattern attack
  - Some patterns available in ciphertext may help to break the ciphertext
  - Try produce ciphertext as random as possible

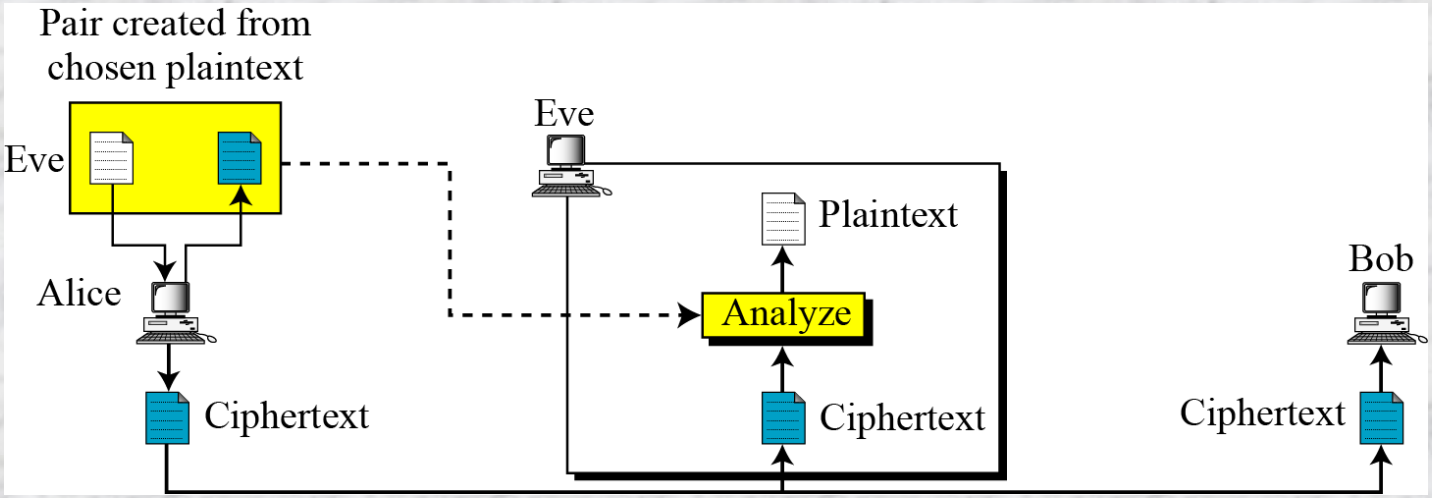
Known-Plaintext Attack

- Attacker has access some previous plaintext/ciphertext pairs
  - Assumption: No change of key
  - Assumption: receiver made the ciphertext public
  - Easier to implement
  - Less likely to happen:
    - Key may be changed
    - May not disclose the content of previous messages



Chosen-Plaintext Attack

- Plaintext/ciphertext pairs have been selected by attacker
  - Attacker can access the machine of sender

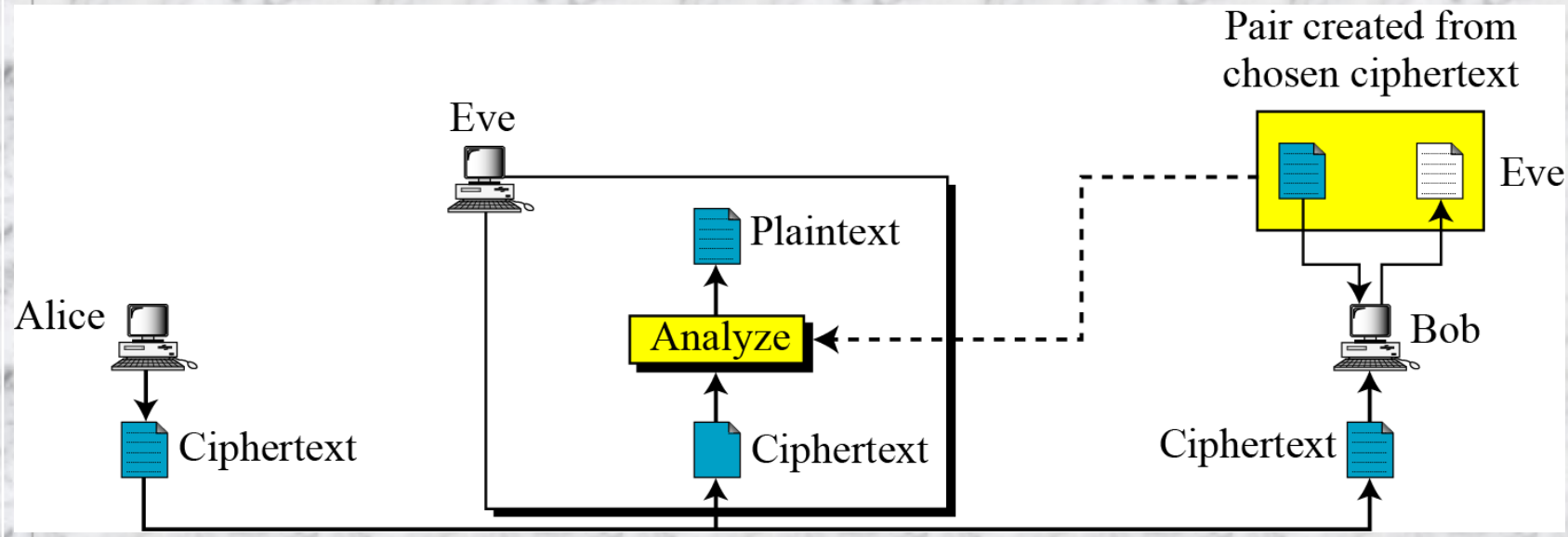


- It is less likely to happen

An *adaptive chosen-plaintext attack is a chosen-plaintext attack wherein the choice* of plaintext may depend on the ciphertext received from previous requests.

Chosen-Ciphertext Attack

- Attacker selects ciphertext/plaintext pairs
  - Attacker can access the machine of receiver



An *adaptive chosen-ciphertext attack is a chosen-ciphertext attack where the choice* of ciphertext may depend on the plaintext received from previous requests.

Traditional cipher techniques

- Traditional symmetric-key ciphers are broadly categorized into two groups:
  - Substitution ciphers
  - Transposition ciphers

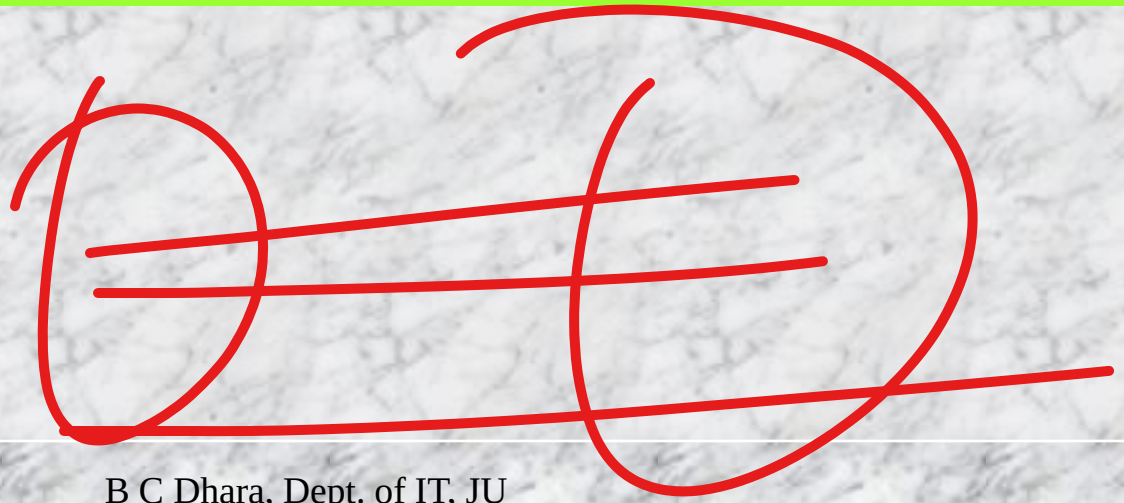
SUBSTITUTION CIPHERS

- A substitution cipher replaces one symbol with another.
  - Substitution ciphers can be categorized as either monoalphabetic ciphers or polyalphabetic ciphers.

A substitution cipher replaces one symbol with another.

Monoalphabetic Ciphers

In monoalphabetic substitution, the relationship between a symbol in the plaintext to a symbol in the ciphertext is always one-to-one.



Monoalphabetic Ciphers (contd…)

- **Example:** The following shows a plaintext and its corresponding ciphertext. The cipher is probably monoalphabetic because both l's (els) are encrypted as O's
  - Plaintext: hello Ciphertext: KHOOR
- **Example:** The following shows a plaintext and its corresponding ciphertext. The cipher is not monoalphabetic because each l (el) is encrypted by a different character
  - Plaintext: hello Ciphertext: ABNZF

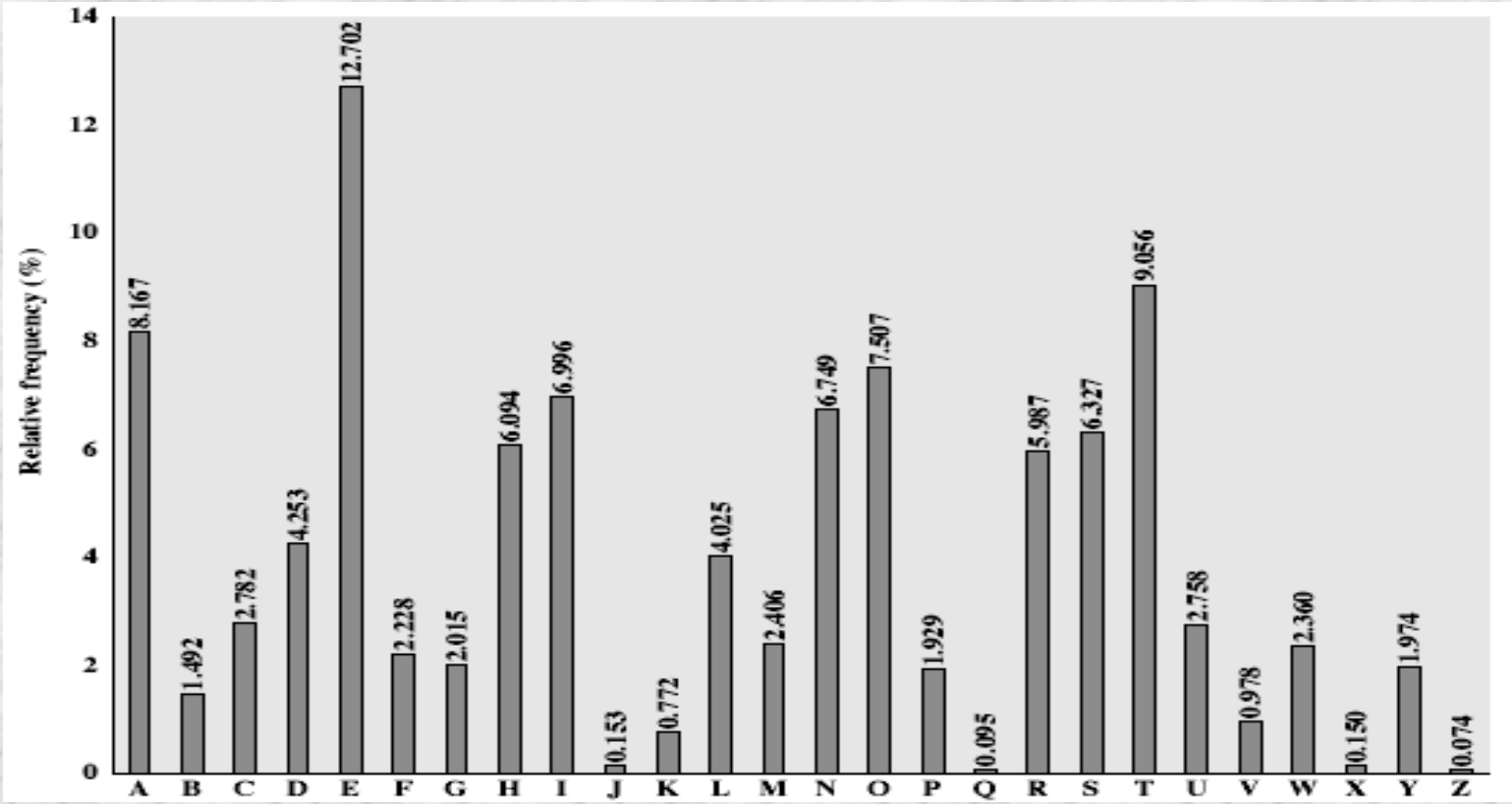
Monoalphabetic Cipher Security

- A total of 26! = 4 x 10^{26} keys
- So many keys, might think is secure
- but would be **!!!WRONG!!!**
- problem is language characteristics

Language Redundancy and Cryptanalysis

- human languages are **redundant**
- letters are not equally 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 ave tables of single, double & triple letter frequencies for various languages

English Letter Frequencies

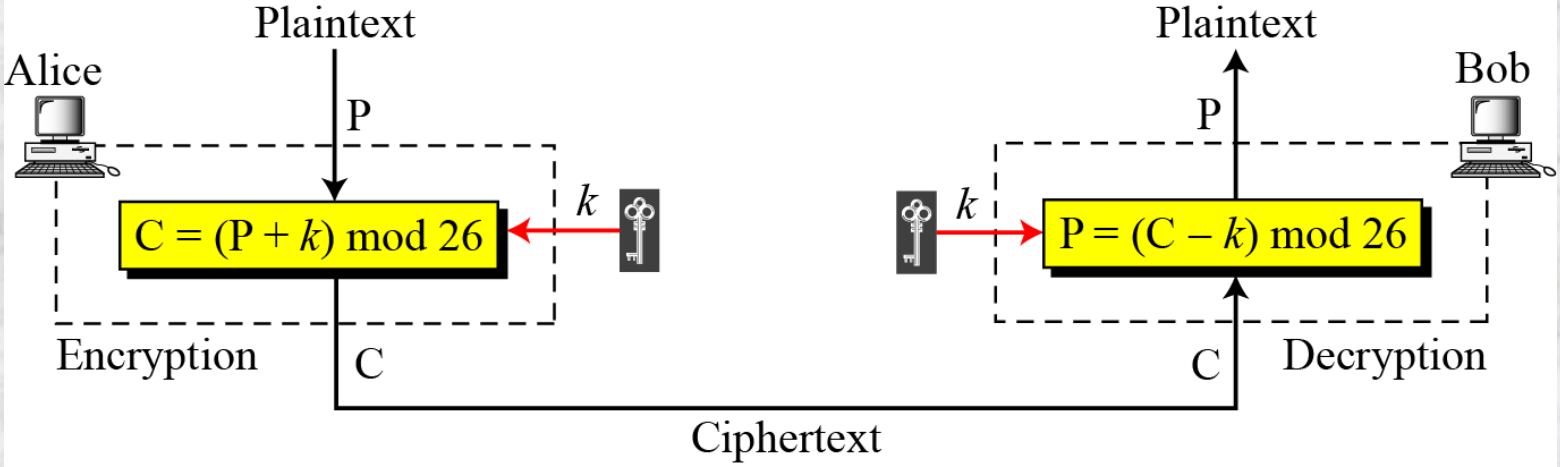


Additive cipher

- The simplest monoalphabetic cipher is the additive cipher
  - This cipher is sometimes called a shift cipher and sometimes a Caesar cipher, but the term additive cipher better reveals its mathematical nature.

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
Ciphertext →	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
Value →	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Additive cipher (contd…)



When the cipher is additive, the plaintext, ciphertext, and key are integers in Z_{26}.

C= (P+k) mod 26

P1=(C-k) mos 26 =(P+k-k) mod 26 = P

Additive cipher (contd…)

- Use the additive cipher with $k = 15$ for "hello" ⟷ "WTAAD".

Plaintext: h → 07	Encryption: (07 + 15) mod 26	Ciphertext: 22 → W
Plaintext: e → 04	Encryption: (04 + 15) mod 26	Ciphertext: 19 → T
Plaintext: l → 11	Encryption: (11 + 15) mod 26	Ciphertext: 00 → A
Plaintext: l → 11	Encryption: (11 + 15) mod 26	Ciphertext: 00 → A
Plaintext: o → 14	Encryption: (14 + 15) mod 26	Ciphertext: 03 → D

Ciphertext: W → 22	Decryption: (22 – 15) mod 26	Plaintext: 07 → h
Ciphertext: T → 19	Decryption: (19 – 15) mod 26	Plaintext: 04 → e
Ciphertext: A → 00	Decryption: (00 – 15) mod 26	Plaintext: 11 → l
Ciphertext: A → 00	Decryption: (00 – 15) mod 26	Plaintext: 11 → l
Ciphertext: D → 03	Decryption: (03 – 15) mod 26	Plaintext: 14 → o

- It is important to make sure that modulo operation will return non-negative integer.

Additive cipher (contd…)

- Historically, additive ciphers are called shift ciphers
  - Julius Caesar used an additive cipher to communicate with his officers
  - For this reason, additive ciphers are sometimes referred to as the Caesar cipher
    - Caesar used a key of 3 for his communications.

Additive ciphers are sometimes referred to as shift ciphers or Caesar cipher.

Attack to additive cipher

- Additive cipher is vulnerable to the ciphertext-only attack
  - Since key space is very small, exhaustive search (brute force attack)
  - K=0 has no effect, so only 25 possible keys

Ciphertext: UVACLYFZLJBYL

Attacker has intercepted the ciphertext "UVACLYFZLJBYL". Show how she can use a brute-force attack to break the cipher.

K = 1 → **Plaintext:** tuzbkxeykiaxk
K = 2 → **Plaintext:** styajwdxjhzwj
K = 3 → **Plaintext:** rsxzivcwigyvi
K = 4 → **Plaintext:** qrwyhubvhfxuh
K = 5 → **Plaintext:** pqvxgtaugewtg
K = 6 → **Plaintext:** opuwfsztfdvsf
K = 7 → **Plaintext:** notverysecure

Statistical attack on additive cipher

- Long sequence of ciphertext is available
- Frequency of individual characters are used
- Sometimes characters pattern (i.e., a group of consecutive characters) is also important
  - Two-letter groups (diagrams), three-letter groups (trigrams)

Frequency tables

Letter	*Frequency*	*Letter*	*Frequency*	*Letter*	*Frequency*	*Letter*	*Frequency*
E	12.7	H	6.1	W	2.3	K	0.08
T	9.1	R	6.0	F	2.2	J	0.02
A	8.2	D	4.3	G	2.0	Q	0.01
O	7.5	L	4.0	Y	2.0	X	0.01
I	7.0	C	2.8	P	1.9	Z	0.01
N	6.7	U	2.8	B	1.5		
S	6.3	M	2.4	V	1.0		

Digram	TH, HE, IN, ER, AN, RE, ED, ON, ES, ST, EN, AT, TO, NT, HA, ND, OU, EA, NG, AS, OR, TI, IS, ET, IT, AR, TE, SE, HI, OF
Trigram	THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR, DTH

Statistical attack on additive cipher (cont…)

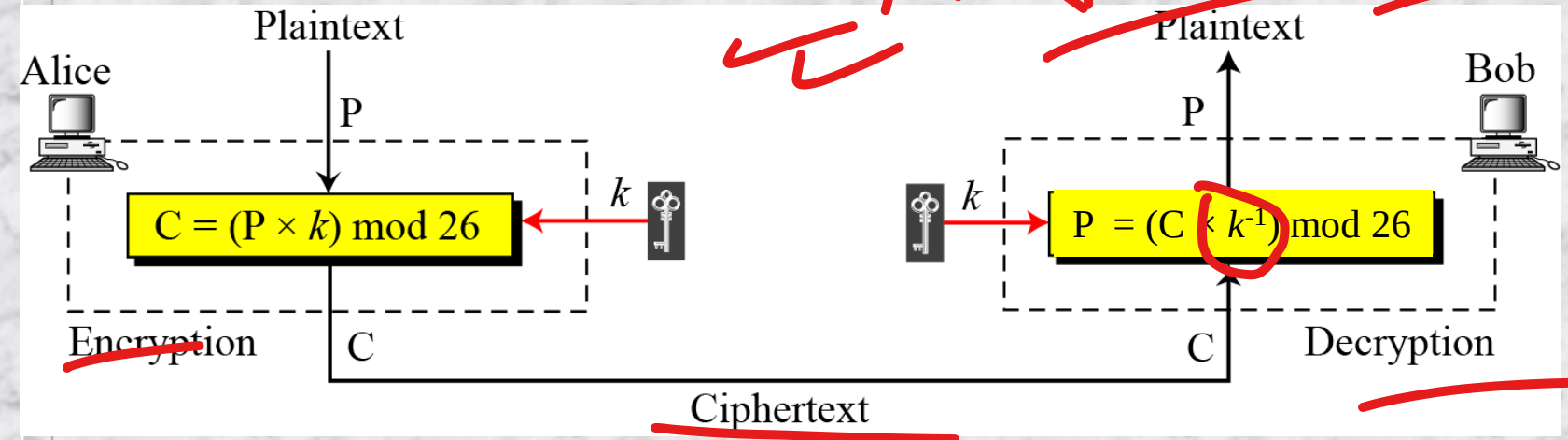
- Attacker has intercepted the following ciphertext. Using a statistical attack, find the plaintext

XLILSYWIMWRSAJSVWEPIJSVJSYVQMPPMSRHSPPEVWMXMWASVX-LQSVILY-VVCFIJSVIXLIWIPPIVVIGIMZIWQSVISJJIVW

- When tabulates the frequency of letters in this ciphertext,
  - The most common character is I with 14 occurrences. This means key = -4.

the house is now for sale for four million dollars it is worth more hurry before the seller receives more offers

Multiplicative Ciphers



- In a multiplicative cipher, the plaintext and ciphertext are integers in Z_{26};
  - the key is an integer in Z_{26}^*, due to the multiplicative inverse

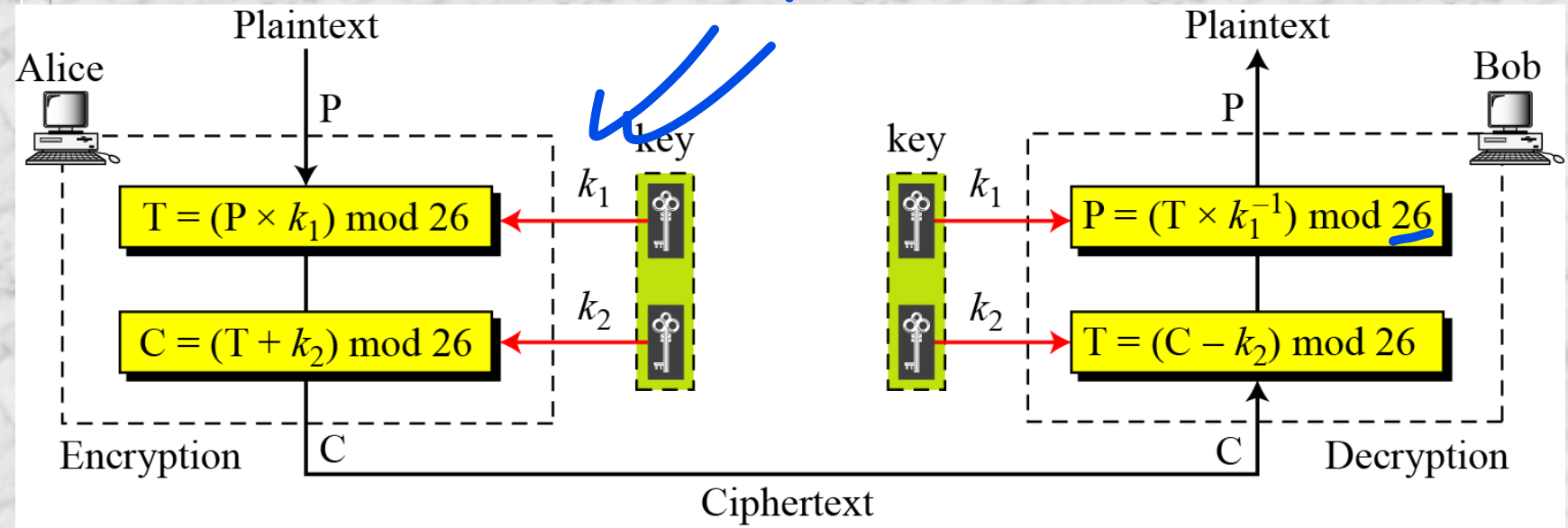
Multiplicative Ciphers (contd…)

- What is the key domain for any multiplicative cipher?
  - The key needs to be in Z_{26}^*. This set has only 12 members: {1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 25}
- We use a multiplicative cipher to encrypt the message "hello" with a key of 7. The ciphertext is "XCZZU"

Plaintext: h → 07	Encryption: (07 × 07) mod 26	ciphertext: 23 → X
Plaintext: e → 04	Encryption: (04 × 07) mod 26	ciphertext: 02 → C
Plaintext: l → 11	Encryption: (11 × 07) mod 26	ciphertext: 25 → Z
Plaintext: l → 11	Encryption: (11 × 07) mod 26	ciphertext: 25 → Z
Plaintext: o → 14	Encryption: (14 × 07) mod 26	ciphertext: 20 → U

Affine cipher

- A combination of additive and multiplicative ciphers



$$C = (P \times k_1 + k_2) \bmod 26 \qquad P = ((C - k_2) \times k_1^{-1}) \bmod 26$$

where k_1^{-1} is the multiplicative inverse of k_1 and $-k_2$ is the additive inverse of k_2

Affine cipher (cont…)

- The affine cipher uses a pair of keys in which the first key is from Z_{26}^* and the second is from Z_{26}. The size of the key domain is $26 \times 12 = 312$
- Use an affine cipher to encrypt the message "hello" with the key pair (7, 2)

P: h → 07	Encryption: $(07 \times 7 + 2) \bmod 26$	C: 25 → Z
P: e → 04	Encryption: $(04 \times 7 + 2) \bmod 26$	C: 04 → E
P: l → 11	Encryption: $(11 \times 7 + 2) \bmod 26$	C: 01 → B
P: l → 11	Encryption: $(11 \times 7 + 2) \bmod 26$	C: 01 → B
P: o → 14	Encryption: $(14 \times 7 + 2) \bmod 26$	C: 22 → W

Affine cipher (cont…)

- Use the affine cipher to decrypt the message "ZEBBW" with the key pair (7, 2) in modulus 26.

C: Z → 25	Decryption: $((25 - 2) \times 7^{-1}) \bmod 26$	P:07 → h
C: E → 04	Decryption: $((04 - 2) \times 7^{-1}) \bmod 26$	P:04 → e
C: B → 01	Decryption: $((01 - 2) \times 7^{-1}) \bmod 26$	P:11 → l
C: B → 01	Decryption: $((01 - 2) \times 7^{-1}) \bmod 26$	P:11 → l
C: W → 22	Decryption: $((22 - 2) \times 7^{-1}) \bmod 26$	P:14 → o

- The additive cipher is a special case of an affine cipher in which $k_1 = 1$
- The multiplicative cipher is a special case of affine cipher in which $k_2 = 0$

Cryptanalysis to Affine cipher

- Ciphertext-only attack like brute-force attack, statistical attack can be used
- Now, chosen-plaintext attack is considered

Ciphertext: PWUFFOGWCHFDIWEJOUUNJORSMDWRHVCMWJUPVCCG

Two-letter plaintext "et", two encryption algorithms (with different k_1 and k_2)

i) plaintext: et → ciphertext: WC

ii) plaintext: et → ciphertext: WF

Cryptanalysis to Affine cipher (cont…)

i) e → W ➔ 04 →22 and t →C ➔ 19 →02

$04 \times k_1 + k_2 \equiv 22 \pmod{26}$

$19 \times k_1 + k_2 \equiv 02 \pmod{26}$

$k_1 = 16$ and $k_2 = 10$, but $k_1 \notin Z_{26}^*$, not possible

Cryptanalysis to Affine cipher (cont…)

ii) e → W ➔ 04 →22 and t →F ➔ 19 →05

$04 \times k_1 + k_2 \equiv 22 \pmod{26}$

$19 \times k_1 + k_2 \equiv 05 \pmod{26}$

$k_1 = 11$ and $k_2 = 4$, this pair acceptable as $k_1 \in Z_{26}^*$

Encryption key pair is (11,4) and decryption key pair is (19, 22) and plaintext is

best time of the year is spring when flowers bloom

Monoalphabetic Substitution Cipher

- Because additive, multiplicative, and affine ciphers have small key domains, they are very vulnerable to brute-force attack
- A better solution is to create a mapping between each plaintext character and the corresponding ciphertext character
  - A table showing the mapping for each character

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

Monoalphabetic Substitution Cipher (cont…)

- Use the key in previous table to encrypt the message, *this message is easy to encrypt but hard to find the key*

- The ciphertext is

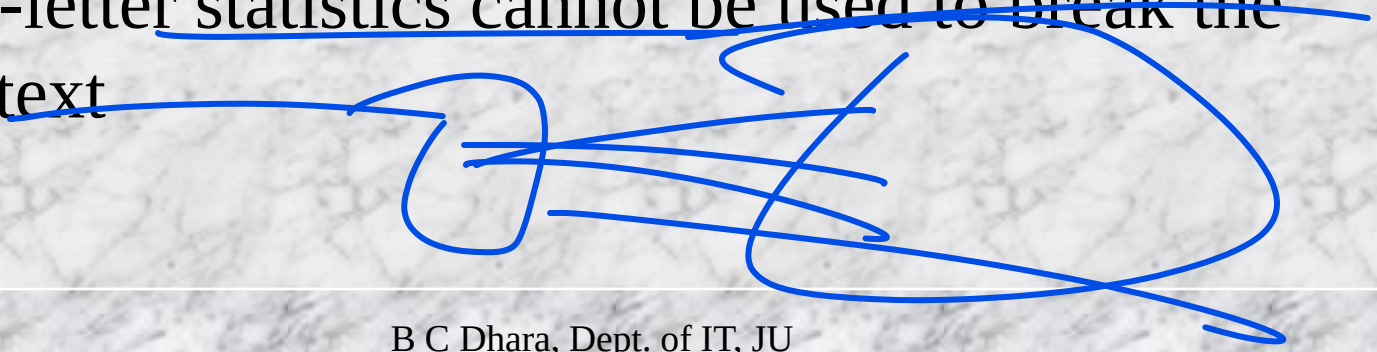
ICFVQRVVNEFVRNVSIYRGAHSLIOJICNHTIYBFGTICRXRS

Monoalphabetic Substitution Cipher (cont…)

- The size of key space for the monoalphabetic cipher is 26!
  - Brute-force attack is extremely difficult
- Monoalphabetic ciphers do not change the frequency of characters in the ciphertext
  - Which makes the ciphers vulnerable to statistical attack
- Affine cipher is a special case of substitution cipher with only 26 x 12 possible keys instead of 26! keys.

Polyalphabetic Ciphers

- In polyalphabetic substitution, each occurrence of a character may have a different substitute.
  - The relationship between a character in the plaintext to a character in the ciphertext is one-to-many
  - This technique hides the letter frequency of the underlying language
  - Single-letter statistics cannot be used to break the ciphertext



Polyalphabetic Ciphers (cont…)

- Ciphertext character depends on the corresponding plaintext and position of the plaintext character in the message
- A key stream k=(k_1, k_2, ….) in which k_i is used to encipher the i^{th} character in the plaintext to create i^{th} character in the ciphertext

Autokey cipher

$P = P_1P_2P_3 \ldots$ $C = C_1C_2C_3\ldots$ $k = (k_1, P_1, P_2, \ldots)$

Encryption: $C_i = (P_i + k_i) \bmod 26$ Decryption: $P_i = (C_i - k_i) \bmod 26$

- Consider an autokey cipher with initial key value k1 = 12. Encipher the message "Attack is today"

Plaintext:	a	t	t	a	c	k	i	s	t	o	d	a	y
P's Values:	00	19	19	00	02	10	08	18	19	14	03	00	24
Key stream:	*12*	*00*	*19*	*19*	*00*	*02*	*10*	*08*	*18*	*19*	*14*	*03*	*00*
C's Values:	12	19	12	19	02	12	18	00	11	7	17	03	24
Ciphertext:	**M**	**T**	**M**	**T**	**C**	**M**	**S**	**A**	**L**	**H**	**R**	**D**	**Y**

Cryptanalysis on Autokey cipher

- It is vulnerable to the brute-force attack
  - since first subkey can be only one of the 25 values (1 to 25)
  - Ciphers should have large key domains

Playflair cipher

- Used during World War I
- Secret key is of 25 alphabets stored in 5x5 matrix
- Different arrangements of the letters create many secret keys

Secret Key =

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

Playflair cipher (contd…)

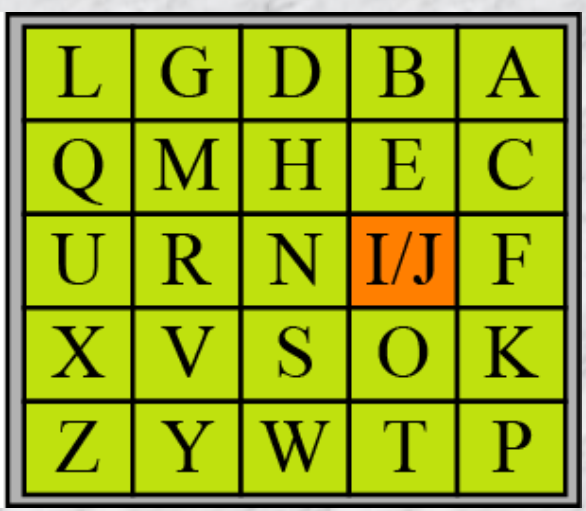
- A message string is encrypted using the following logic:
- Consider two letters at a time
- If a pair with same two letters are not considered
  - If appears, an auxiliary symbol is inserted between them
- If length of the string is odd, one auxiliary symbol is inserted at the end to make the string of length even

Playflair cipher (contd…)

- Consider group of two letters from the pre-processed message string
- If both letters appear on the same row of the secret key matrix
  - Cipher letter for each letter is the next letter to the right in the same row with wrapping to beginning of the row
- If both letters appear on the same column of the secret key matrix
  - Cipher letter for each letter is the next letter to the beneath in the column with wrapping to beginning of the column
- If two pair letters are not in the same row or column
  - Cipher letters are at the row-column intersection points (with row priority)

Example of Playflair cipher

Secret Key =



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

Let us encrypt the plaintext “hello” using the above key matrix

he → EC lx → QZ lo → BX

Plaintext: hello Ciphertext: ECQZBX

Notes on Playflair cipher

- Auxiliary letter should be something from {26,27,28,29,30,31} as letters are represented by {0,1,2,…,25}
- Auxiliary symbols are of two types:
  - i) as the separator for ‘xx’, and
  - ii) separator for ‘IJ’ or ‘JI’

Secret Key =

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

Cryptanalysis of Playfair cipher

- Brute-force attacke is very difficult
- Encipherment hides the signle-letter frequency
- Frequencies of diagrams are preserved (some change due to filler insertion)
  - An attacker can use ciphertext-only attack based on diagram frequency test to find the key

Vigenere Cipher

$P = P_1P_2P_3 \ldots$ $C = C_1C_2C_3 \ldots$ $K = [(k_1, k_2, \ldots, k_m), (k_1, k_2, \ldots, k_m), \ldots]$

Encryption: $C_i = P_i + k_i$ Decryption: $P_i = C_i - k_i$

We can encrypt the message "She is listening" using the 6-character keyword "PASCAL".

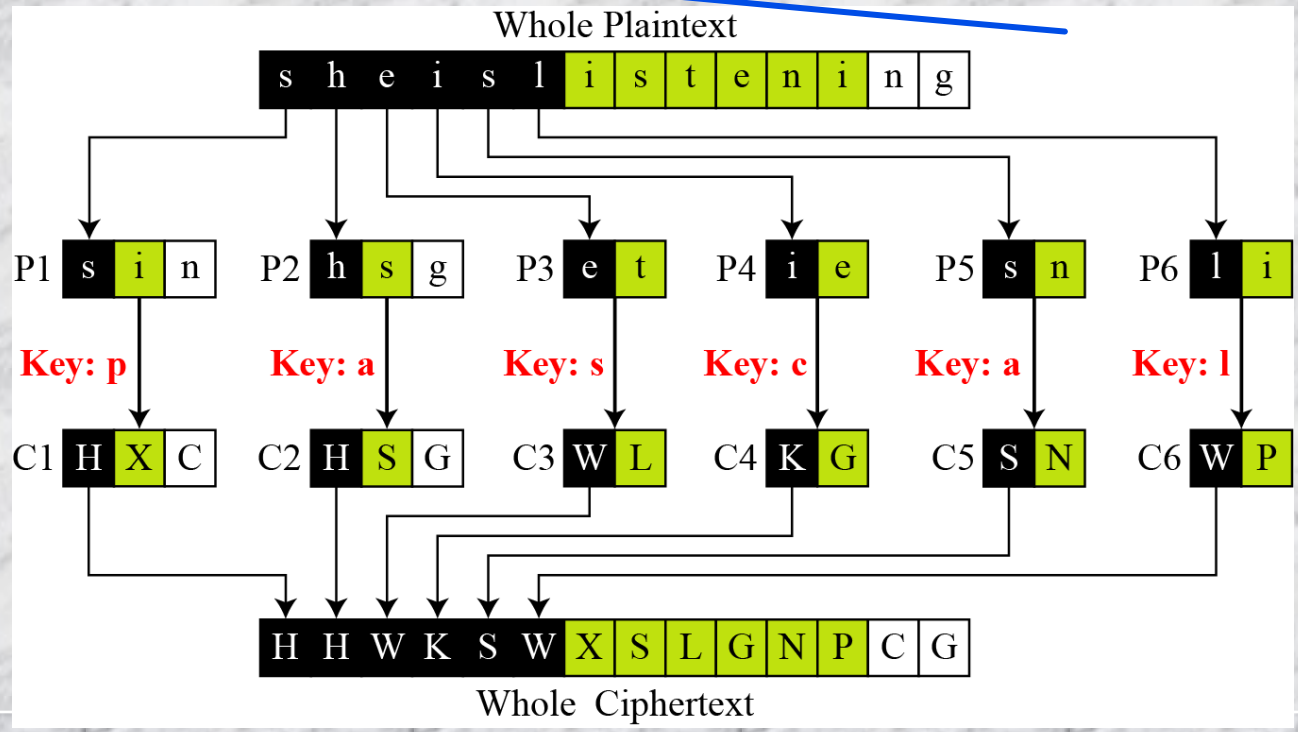
Plaintext:	s	h	e	i	s	l	i	s	t	e	n	i	n	g
P's values:	18	07	04	08	18	11	08	18	19	04	13	08	13	06
Key stream:	***15***	***00***	***18***	***02***	***00***	***11***	***15***	***00***	***18***	***02***	***00***	***11***	***15***	***00***
C's values:	**07**	**07**	**22**	**10**	**18**	**22**	**23**	**18**	**11**	**6**	**13**	**19**	**02**	**06**
Ciphertext:	**H**	**H**	**W**	**K**	**S**	**W**	**X**	**S**	**L**	**G**	**N**	**T**	**C**	**G**

Vigenere Cipher (contd …)

- Vigenere key stream does not depend on the plaintext characters
- It depends only on the position of the character in the plaintext
  - Key stream can be created without knowing the plaintext

Vigenere Cipher (contd …)

- Vigenere cipher can be seen as combinations of m additive ciphers.



Vigenere Cipher (contd …)

- Additive cipher is a special case of Vigenere cipher in which m = 1

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

Cryptanalysis of Vigenere cipher

- Vigenere cipher does not preserve the frequency of characters
- Cryptanalysis consists of two parts:
  - Finding the key length
  - Finding the key itself

Cryptanalysis of Vigenere cipher (contd…)

- Kasiski test
  - Find repeated text segment, of length at least three, in the ciphertext
  - Suppose, two segments found and distance between them is d
    - d| m, m the length of the key
  - If more repeated segments found with distances d_1, d_2,…,d_n , then gcd(d_1, d_2,…,d_n) | m
- When length of key is found,
  - Divides the ciphertext into m pieces and …
  - If fails then divides the message into 2m pieces, …

Vigenere Cipher (Crypanalysis)

Let us assume we have intercepted the following ciphertext:

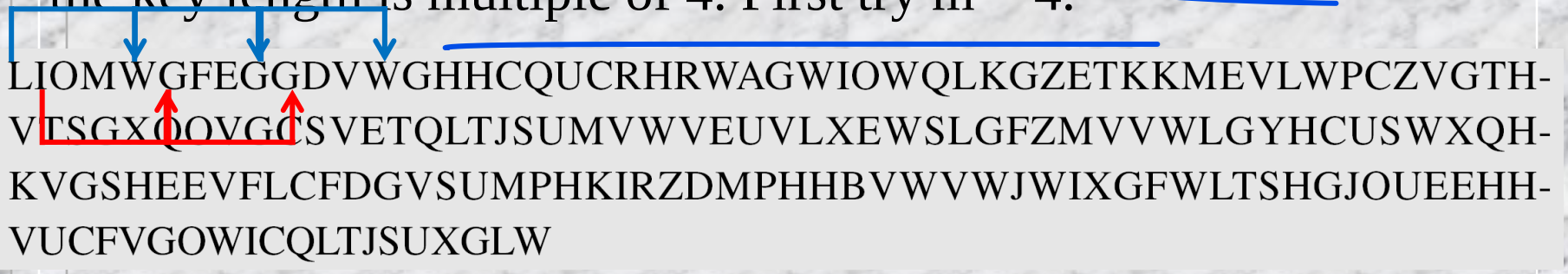
LIOMWGFEGGDVWGHHCQUCRHRWAGWIOWQLKGZETKKMEVLWPCZVGTH-
VTSGXQOVGCSVETQLTJSUMVWVEUVLXEWSLGFZMVVWLGYHCUSWXQH-
KVGSHEEVFLCFDGVSUMPHKIRZDMPHHBVWVWJWIXGFWLTSHGJOUEEHH-
VUCFVGOWICQLTJSUXGLW

The Kasiski test for repetition of three-character segments yields the results shown in the following Table

String	*First Index*	*Second Index*	*Difference*
JSU	68	168	100
SUM	69	117	48
VWV	72	132	60
MPH	119	127	8

Vigenere Cipher (Crypanalysis) (contd…)

- The greatest common divisor of differences is 4, which means that the key length is multiple of 4. First try m = 4.



C1: **LWGWCRAOKTEPGTQCTJVUEGVGUQGECVPRPVJGTJEUGCJG**

P1: *jueuapymircneroarhtsthihytrahcieixsthcarrehe*

C2: **IGGGQHGWGKVCTSOSQSWVWFVYSHSVFSHZHWWFSOHCOQSL**

P2: *usssctsiswhofeaeceihcetesoecatnpntherhctecex*

C3: **OFDHURWQZKLZHGVVLUVLSZWHWKHFDUKDHVIWHUHFWLUW**

P3: *lcaerotnwhiwedssirsiirhketehretltiideatrairt*

C4: **MEVHCWILEMWVVXGETMEXLMLCXVELGMIMBWXLGEVVITX**

P4: *iardysehaisrrtcapiafpwtethecarhaesfterectpt*

Vigenere Cipher (Crypanalysis) (contd…)

C1: **LWGWCRAOKTEPGTQCTJVUEGVGUQGECVPRPVJGTJEUGCJG**

P1: *jueuapymircneroarhtsthihytrahcieixsthcarrehe*

C2: **IGGGQHGWGKVCTSOSQSWVWFVYSHSVFSHZHWWFSOHCOQSL**

P2: *usssctsiswhofeaeceihcetesoecatnpntherhctecex*

C3: **OFDHURWQZKLZHGVVLUVLSZWHWKHFDUKDHVIWHUHFWLUW**

P3: *lcaerotnwhiwedssirsiirhketehretltiideatrairt*

C4: **MEVHCWILEMWVVXGETMEXLMLCXVELGMIMBWXLGEVVITX**

P4: *iardysehaisrrtcapiafpwtethecarhaesfterectpt*

In this case, the plaintext makes sense.

Julius Caesar used a cryptosystem in his wars, which is now referred to as Caesar cipher. It is an additive cipher with the key set to three. Each character in the plaintext is shifted three characters to create ciphertext.

Hill cipher

- Divides the plaintext into equal-size blocks
- Blocks are encrypted one at a time in such a way that each character in the block contributes to the encryption of other character in the block
  - This is block cipher
- Key is a square matrix of size m × m, size of block is m

Hill cipher (contd…)

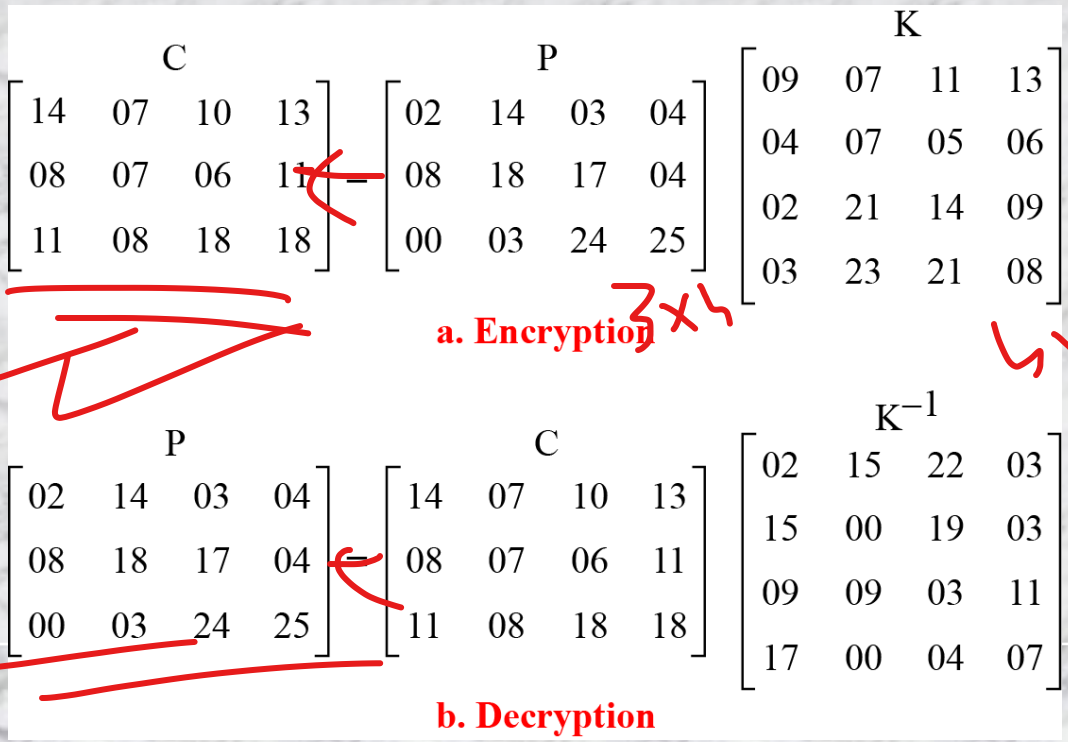
$$K = \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1m} \\ k_{21} & k_{22} & \cdots & k_{2m} \\ \vdots & \vdots & & \vdots \\ k_{m1} & k_{m2} & \cdots & k_{mm} \end{bmatrix}$$

$$C_1 = P_1 k_{11} + P_2 k_{21} + \cdots + P_m k_{m1}$$
$$C_2 = P_1 k_{12} + P_2 k_{22} + \cdots + P_m k_{m2}$$
$$\cdots$$
$$C_m = P_1 k_{1m} + P_2 k_{2m} + \cdots + P_m k_{mm}$$

The key matrix in the Hill cipher needs to have a multiplicative inverse.

Hill cipher (contd…)

- For example, the plaintext "code is ready" can make a 3 × 4 matrix when adding extra bogus character "z" to the last block and removing the spaces. The ciphertext is "OHKNIHGKLISS"



Cryptanalysis of Hill cipher

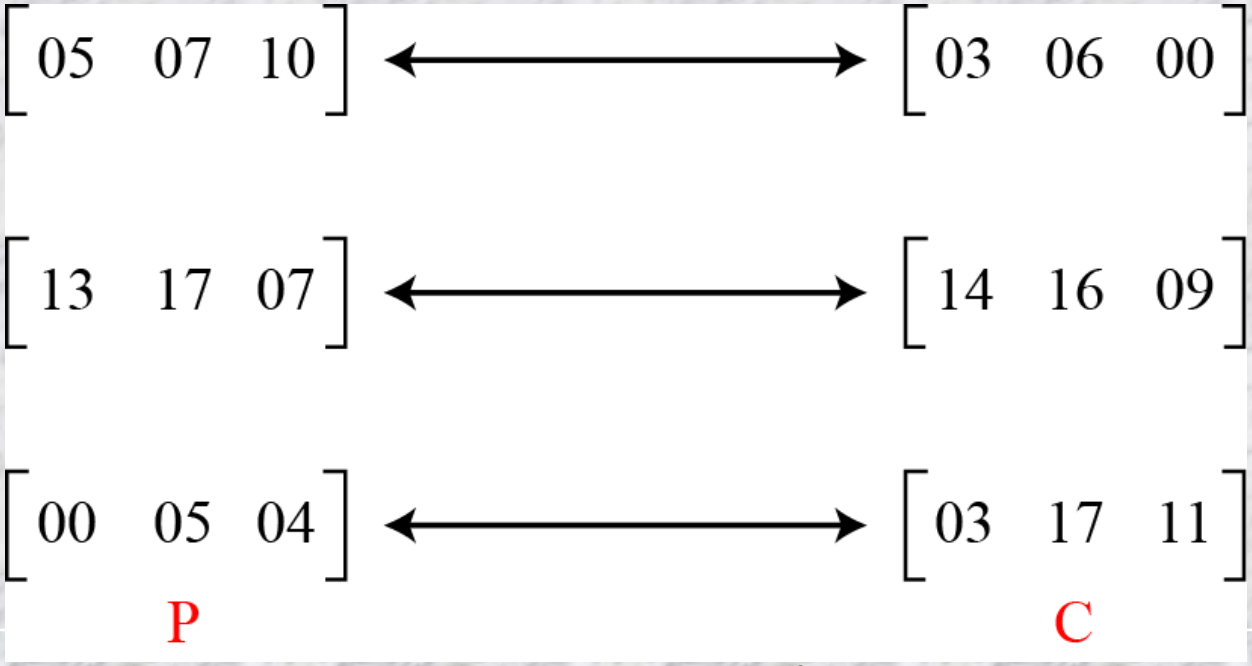
- Ciphertext-only attack is difficult as:
  - Brute-force attack is extremely difficult
    - Since size of key matrix is m × m and in each location possibilities is 26
    - Size of key space is $26^{m \times m}$
    - Non-invertible matrices cannot be key
  - This cipher does not preserve the statistics of the plaintext
    - Frequency analysis on single letter, diagrams, trigrams will not work
    - Frequency analysis on words of length m might work that plaintext has many words of length m that are same

Cryptanalysis of Hill cipher (contd…)

- Attacker can execute known-plaintext attack if
  - m is known
  - At least m plaintext/ciphertext pairs are known (may be from same/different messages)
  - Blocks must be distinct
    - Two matrices of size m × m, P (plaintext) and C (ciphertext) and key is computed as $K=CP^{-1}$
    - If P is not invertible, new set of m plaintext/ciphertext pairs are needed

Cryptanalysis of Hill cipher: known-plaintext attack

- Assume that attacker knows that m = 3 and intercepted three plaintext/ciphertext pair blocks (not necessarily from the same message)



Cryptanalysis of Hill cipher: known-plaintext attack (contd…)

- Matrices P and C are constructed from these pairs
  - Because P is invertible, then inverts the P matrix and multiplies it by C to get the K matrix

$$\begin{bmatrix} 02 & 03 & 07 \\ 05 & 07 & 09 \\ 01 & 02 & 11 \end{bmatrix} = \begin{bmatrix} 21 & 14 & 01 \\ 00 & 08 & 25 \\ 13 & 03 & 08 \end{bmatrix} \begin{bmatrix} 03 & 06 & 00 \\ 14 & 16 & 09 \\ 03 & 17 & 11 \end{bmatrix}$$

$$K \qquad P^{-1} \qquad C$$

Now the key and can break any ciphertext encrypted with that key

One-Time Pad

- One of the goals of cryptography is perfect secrecy
- A study by Shannon has shown that perfect secrecy can be achieved if each plaintext symbol is encrypted with a key randomly chosen from a key domain
- This idea is used in a cipher called one-time pad, invented by Vernam

One-Time Pad *(contd…)*

- In additive cipher,
  - To encrypt each symbol a key randomly selected from {0,1,2,..,25} instead of a fixed key
  - Ciphertext-only attack is impossible
- One-time pad cipher, key has the same length as the plaintext
- One-time pad is perfect
  - But, almost impossible to implement it becaue
    - How the keys will be communicated between sender and receiver?

TRANSPOSITION CIPHERS

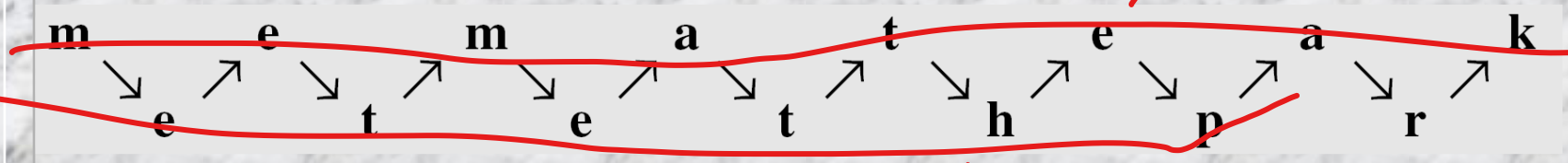
- Consider classical **transposition** or **permutation** ciphers
- A transposition cipher does not substitute one symbol for another, instead it changes the location of the symbols
  - Hide the message by rearranging the letter order without altering the actual letters used
- Keyless Transposition Ciphers
- Keyed Transposition Ciphers
- Combining Two Approaches

Keyless Transposition Ciphers

- Simple transposition ciphers, which were used in the past, are keyless.

Rail fence cipher

- Write message letters out diagonally over a number of rows
- Then read off cipher row by row
- For example, the message “Meet me at the park” is encipher as follows



- The ciphertext is “MEMATEAKETETHPR”.

Keyless method: second method

- The sender and receiver can agree on the number of columns and write the plaintext, row by row, in a table
- Finally, ciphertext is constructed considering column wise letters

m	**e**	**e**	**t**
m	**e**	**a**	**t**
t	**h**	**e**	**p**
a	**r**	**k**	

the ciphertext is "MMTAEEHREAEKTTP"

Keyless method: second method (contd…)

The cipher is achieved by permutation of each character in the plaintext into the ciphertext based on the positions

$\mathbf{m}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{t}_4$
$\mathbf{m}_5$	**e**	**a**	**t**
t	**h**	**e**	**p**
a	**r**	**k**	

01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
01	05	09	13	02	06	10	13	03	07	11	15	04	08	12

Keyed Transposition Ciphers

- The keyless ciphers permute the characters by using writing plaintext in one way and reading it in another way
- The permutation is done on the whole plaintext to create the whole ciphertext
- Another method is to divide the plaintext into groups of predetermined size, called blocks, and then use a key to permute the characters in each block separately

Keyed Transposition Ciphers (cont…)

- Send the message “Enemy attacks tonight”
- The key used for encryption and decryption is a permutation key, which shows how the character are permuted

Encryption ↓

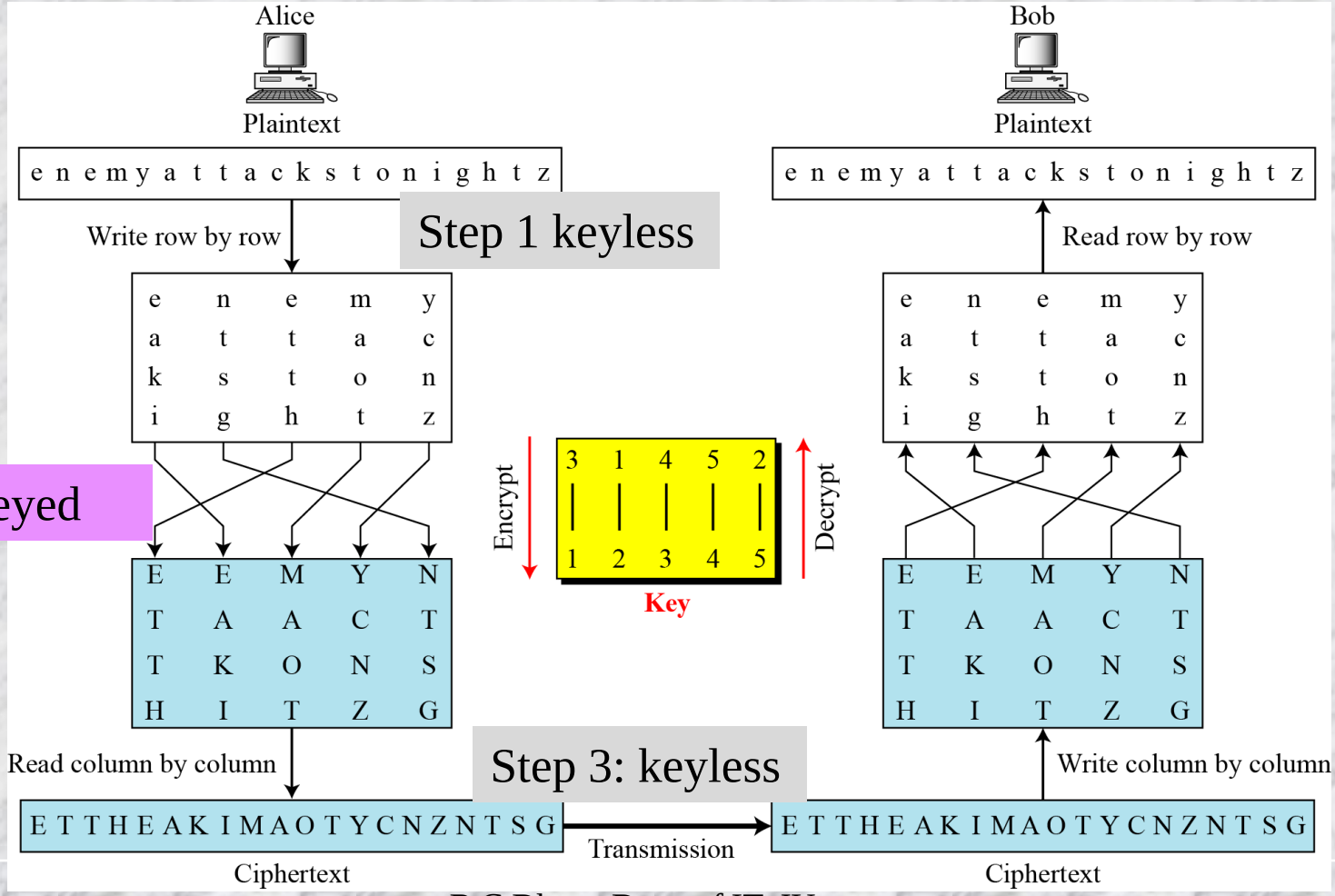
3	1	4	5	2
1	2	3	4	5

↑ Decryption

- The permutation yields

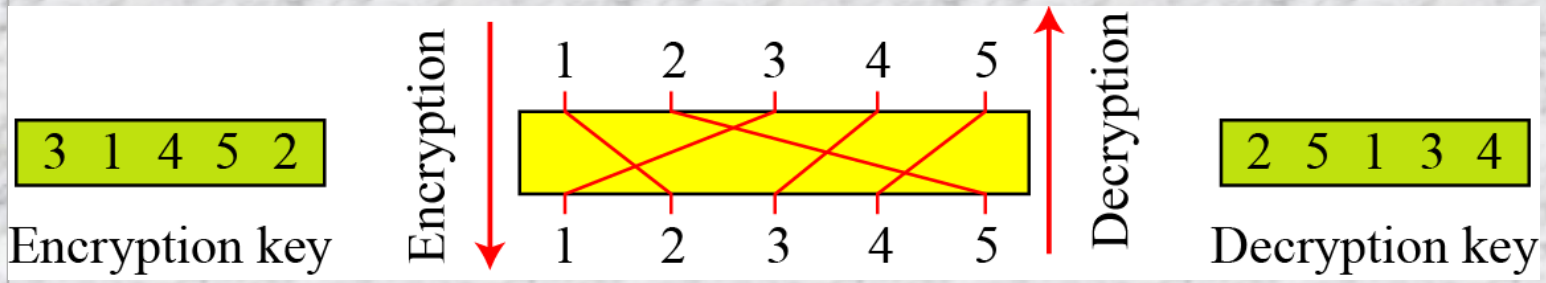
E E M Y N T A A C T T K O N S H I T Z G

Combining Two Approaches

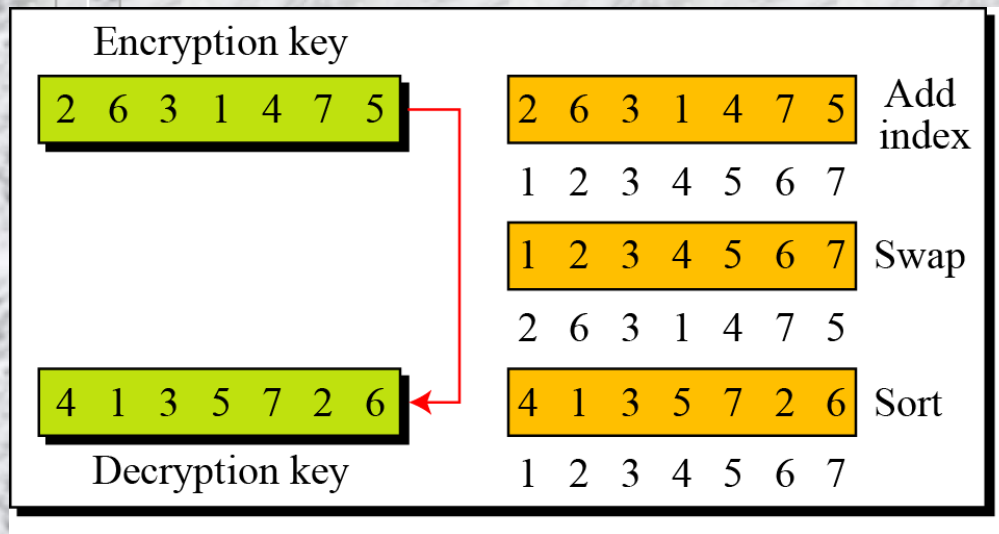


Representation of permutation

- Representation of permutation in encryption and decryption



How the decryption key can be computed, if encryption key given?



a. Manual process

```
Given: EncKey [index]
index ← 1
while (index ≤ Column)
  {
    DecKey[EncKey[index]] ← index
    index ← index + 1
  }
Return : DecKey [index]
```

b. Algorithm

Use of matrix in transposition

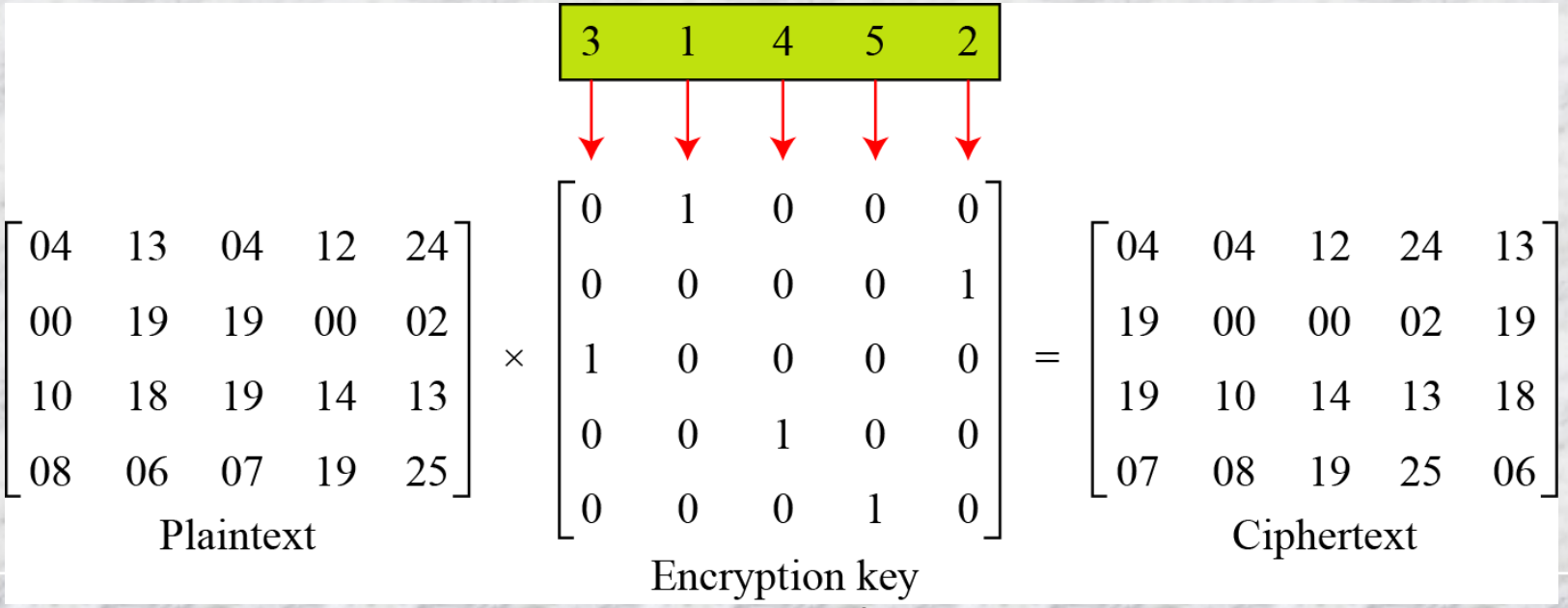
- The encryption/decryption process for a transposition cipher can be accomplished using matrix multiplications
  - The permutation of rows/columns can be executed by multiplication with permutation matrix
  - Plaintext/ciphertext matrix are n×m and permutation matrix is m×m

Permutation matrix

- Obtained from an identity matrix, by swapping rows or columns
- In each row/cloumn, exactly one 1 and rest are 0s
- Determinant is ± 1, nonsingular
- $AA^t = I$
  - A^t is the inverse of A

Use of matrix in transposition: example

- The encryption process
  - Multiplying the 4 × 5 plaintext matrix by the 5 × 5 encryption key gives the 4 × 5 ciphertext matrix



Cryptanalysis of Transposition ciphers

- Transposition ciphers are vulnerable to ciphertext-only attacks
- **Statistical attacks**
  - No change in frequency of letters in ciphertext
  - Single letter frequency analysis is useful if long length ciphertext is available
  - Does not preserve the frequency of diagram and tridiagram

Cryptanalysis of Transposition ciphers (contd…)

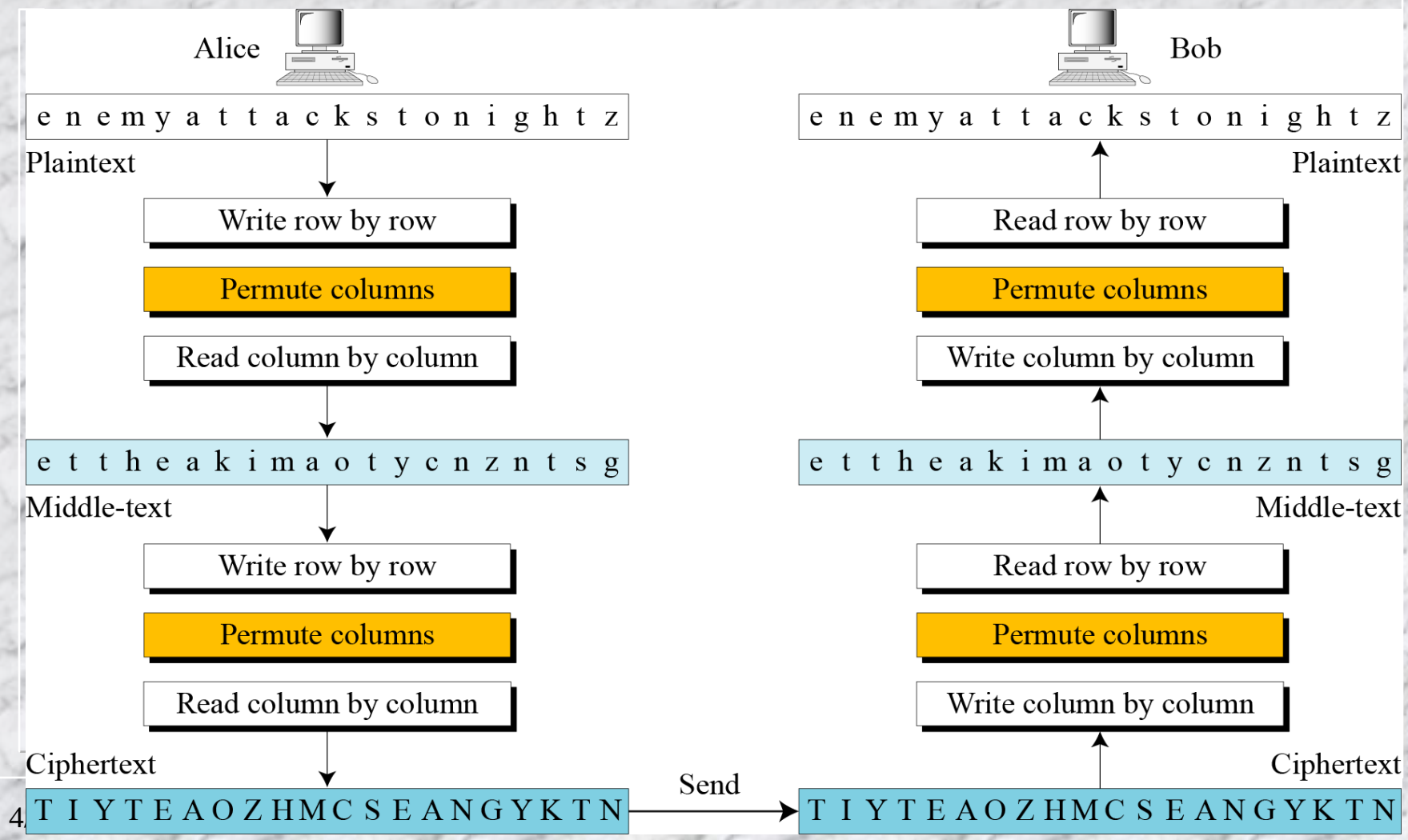
- **Brute-Force attacks**
  - message length L
  - Partition the message into columns
    - Number of columns divides L
    - E.g., if message length L=20, then possible number of columns: 1, 2, 4, 5, 10, 20

      only one column is not useful

      try with all possible columns and within which all possible permutations

Double Transposition Ciphers

To make cryptanalysis task difficult: apply the transposition twice, with different keys (normally same key is used)



STREAM AND BLOCK CIPHERS

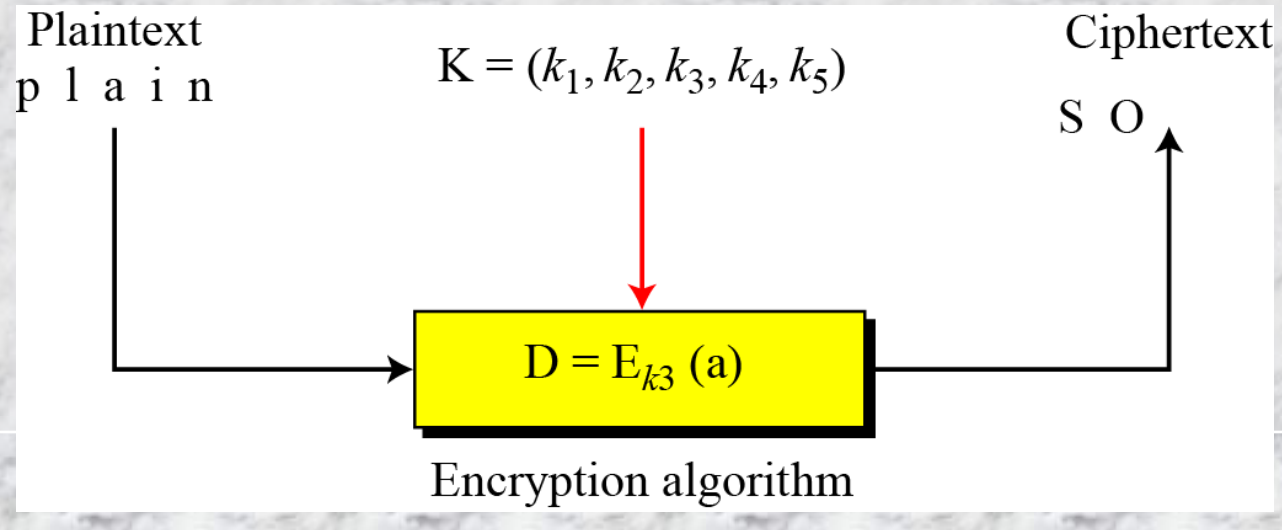
- The literature divides the symmetric ciphers into two broad categories: stream ciphers and block ciphers
  - Stream Ciphers
  - Block Ciphers
  - Combination

Stream Ciphers

- Call the plaintext stream P, the ciphertext stream C, and the key stream K.

$$P = P_1P_2P_3, \ldots \qquad C = C_1C_2C_3, \ldots \qquad K = (k_1, k_2, k_3, \ldots)$$

$$C_1 = E_{k1}(P_1) \qquad C_2 = E_{k2}(P_2) \qquad C_3 = E_{k3}(P_3) \ldots$$



Stream Ciphers (contd…)

- Key stream:
  - It may be a stream of predefined values
  - May be created one value at a time using an algorithm
  - Values may depend on the plaintext or ciphertext characters
  - Values may aldo depend on the previous values

Stream Ciphers examples

Additive ciphers can be categorized as stream ciphers in which the key stream is the repeated value of the key. In other words, the key stream is considered as a predetermined stream of keys or K = (k, k, …, k). In this cipher, however, each character in the ciphertext depends only on the corresponding character in the plaintext, because the key stream is generated independently.

The monoalphabetic substitution ciphers discussed in this chapter are also stream ciphers. However, each value of the key stream in this case is the mapping of the current plaintext character to the corresponding ciphertext character in the mapping table.

Stream Ciphers examples (contd…)

Vigenere ciphers are also stream ciphers according to the definition. In this case, the key stream is a repetition of m values, where m is the size of the keyword. In other words,

$$K = (k_1, k_2, \ldots k_m, k_1, k_2, \ldots k_m, \ldots)$$

We can establish a criterion to divide stream ciphers based on their key streams. We can say that a stream cipher is a monoalphabetic cipher if the value of k_i does not depend on the position of the plaintext character in the plaintext stream; otherwise, the cipher is polyalphabetic.

Stream Ciphers examples (contd…)

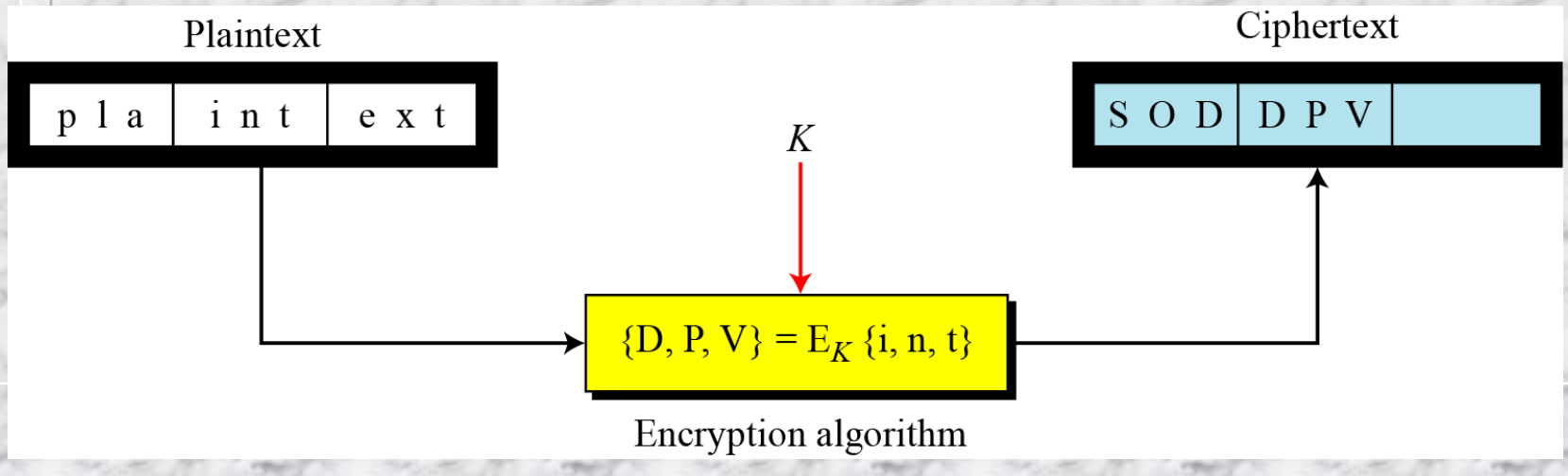
- Additive ciphers are definitely monoalphabetic because k_i in the key stream is fixed; it does not depend on the position of the character in the plaintext.

- Monoalphabetic substitution ciphers are monoalphabetic because k_i does not depend on the position of the corresponding character in the plaintext stream; it depends only on the value of the plaintext character.

- Vigenere ciphers are polyalphabetic ciphers because k_i definitely depends on the position of the plaintext character. However, the dependency is cyclic. The key is the same for two characters m positions apart.

Block cipher

- In a block cipher, a group of plaintext symbols of size m (m > 1) are encrypted together creating a group of ciphertext of the same size.
  - A single key is used to encrypt the whole block even if the key is made of multiple values.



Block cipher example

Playfair ciphers are block ciphers. The size of the block is m = 2. Two characters are encrypted together.

Hill ciphers are block ciphers. A block of plaintext, of size 2 or more is encrypted together using a single key (a matrix). In these ciphers, the value of each character in the ciphertext depends on all the values of the characters in the plaintext. Although the key is made of m × m values, it is considered as a single key.

From the definition of the block cipher, it is clear that every block cipher is a polyalphabetic cipher because each character in a ciphertext block depends on all characters in the plaintext block.

Combination of block/stream ciphers

In practice, blocks of plaintext are encrypted individually, but they use a stream of keys to encrypt the whole message block by block. In other words, the cipher is a block cipher when looking at the individual blocks, but it is a stream cipher when looking at the whole message considering each block as a single unit.