

Cryptography and Network Security

Chapter 2 – Classical Encryption Techniques

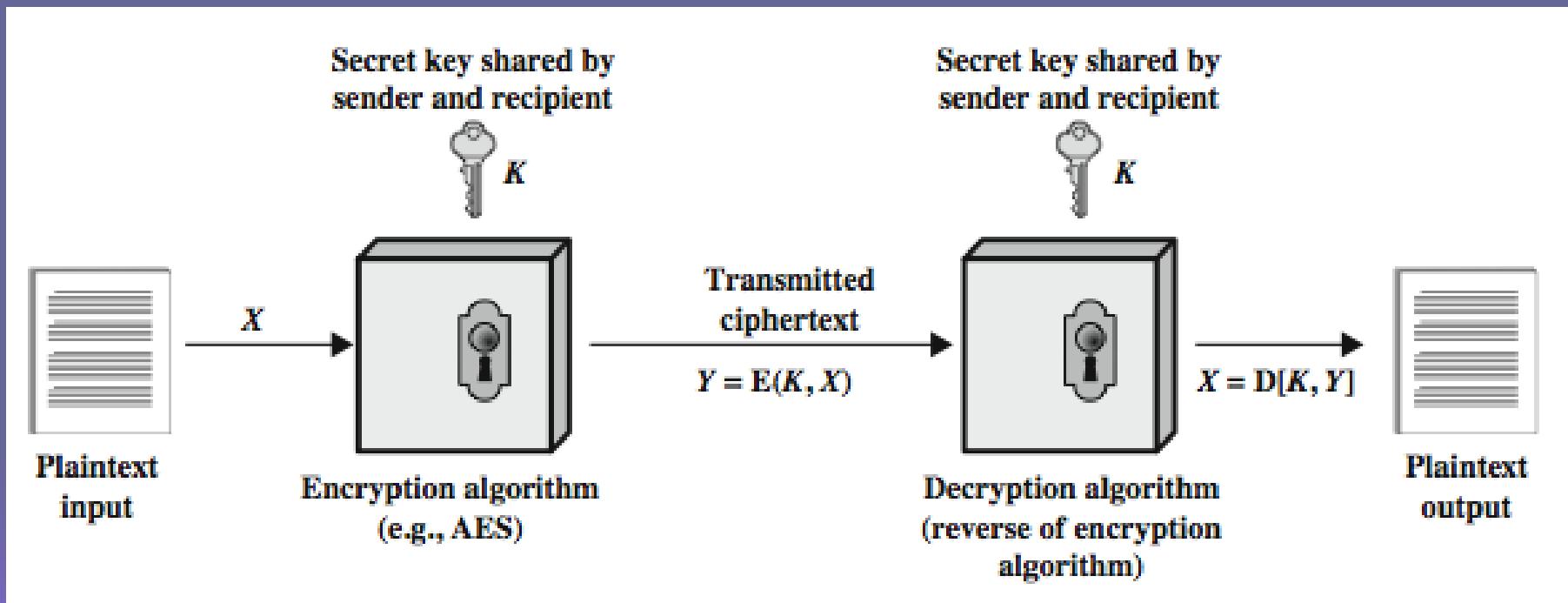
Symmetric Encryption

- 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
- and by far most widely used

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



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

- can characterize cryptographic system 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

Cryptanalysis

- objective to recover key not just message
- general approaches:
 - cryptanalytic attack
 - brute-force attack
- if either succeed all key use compromised

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

➤ **chosen text**

- select plaintext or ciphertext to en/decrypt

More Definitions

➤ **unconditional security**

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

More Definitions

- **The cost of breaking the cipher exceeds the value of the encrypted information**
- **The time required to break the cipher exceeds the useful lifetime of the information**
- **computational security**
 - An encryption scheme is said to
 - If either of the foregoing two criteria are met

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	$2^{32} = 4.3 \times 10^9$	$2^{31} \mu\text{s} = 35.8 \text{ minutes}$	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	$2^{55} \mu\text{s} = 1142 \text{ years}$	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	$2^{127} \mu\text{s} = 5.4 \times 10^{24} \text{ years}$	$5.4 \times 10^{18} \text{ years}$
168	$2^{168} = 3.7 \times 10^{50}$	$2^{167} \mu\text{s} = 5.9 \times 10^{36} \text{ years}$	$5.9 \times 10^{30} \text{ years}$
26 characters (permutation)	$26! = 4 \times 10^{26}$	$2 \times 10^{26} \mu\text{s} = 6.4 \times 10^{12} \text{ years}$	$6.4 \times 10^6 \text{ years}$

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 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:

$$c = E(k, p) = (p + k) \bmod 26$$

$$p = D(k, c) = (c - k) \bmod 26$$

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: DVQFIBJWPESCXHTMYAUOLRGZN

Plaintext: if we wish to replace letters

Ciphertext: WIRFRWAJUHYFTSDVFSUUUFYA

Monoalphabetic Cipher

Plaintext letter	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 letter	y	n	l	k	x	b	s	h	m	i	w	d	p	j	r	o	q	v	f	e	a	u	g	t	z	c

- To encrypt the message "meet me at nine", start by taking the first letter of the message, 'm', and look up the corresponding ciphertext letter 'p'.
- Repeat by looking up the next plaintext letter 'e', and noting it becomes 'x'. Continue this process for the rest of the message. Typically spaces, numbers, and punctuation are left alone.
- In this case "meet me at nine." would become "pxxe px ye jmjx."
- Decryption is similar. Start with the first ciphertext letter 'p', and look at the table to find the corresponding plaintext letter 'm'. Continue with the next letter 'x', and find it maps to 'e.'

Monoalphabetic Cipher

- Monoalphabetic encryption is very easy to break, for two main reasons.
- First, commonly used letters like 'e' show up very quickly as the 'x' in the example.
- Second, words with repeated letters like "meet" in the example show that repetition in the ciphertext.
- And indeed this is so weak that the daily cryptogram run by some newspapers is typically an monoalphabetic substitution.

Monoalphabetic Cipher Security

- now have a total of $26! = 4 \times 10^{26}$ 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

Example

Plain Text

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

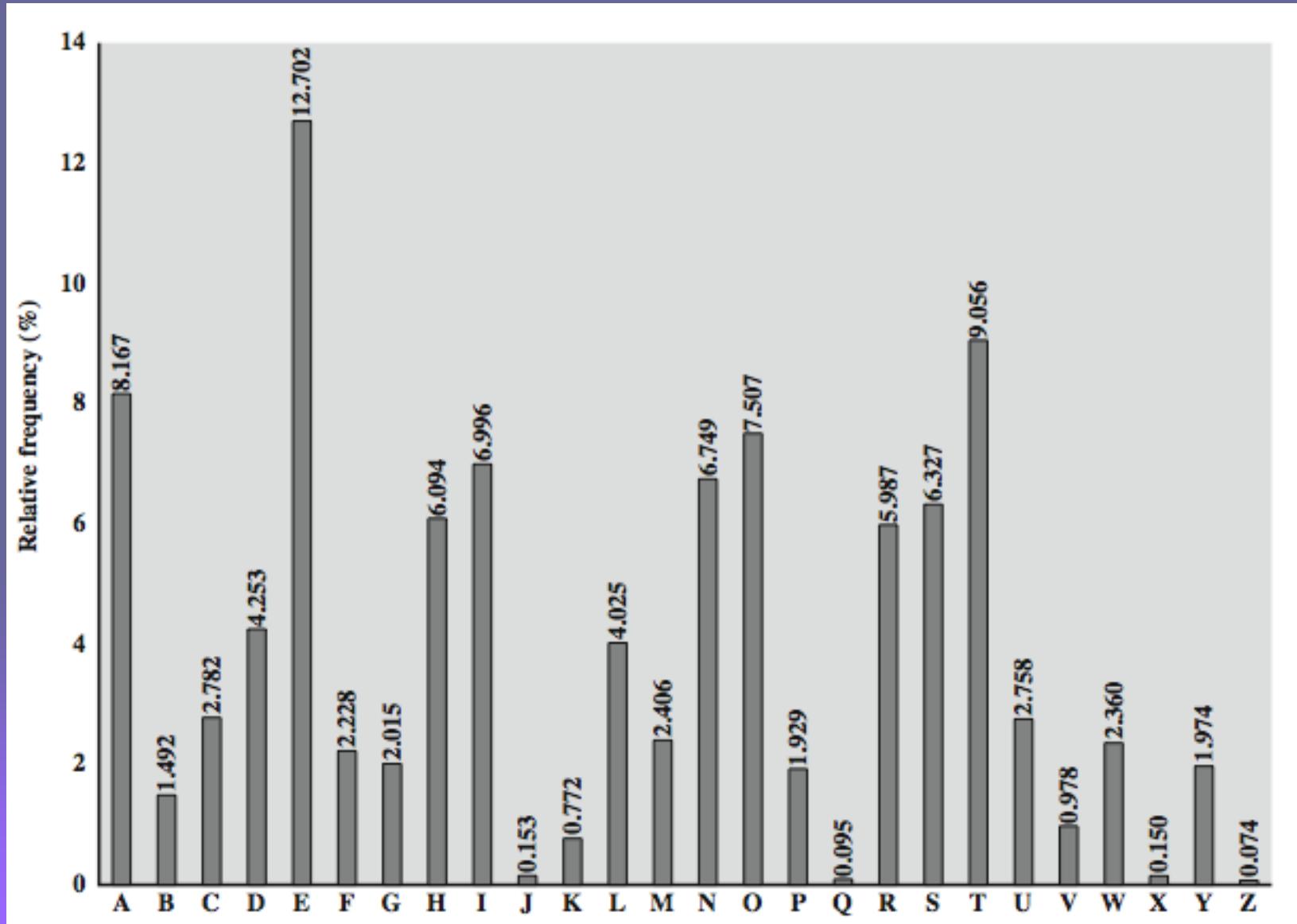
Cipher Text

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFAPPDTSPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

Frequency Table

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				

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
VUEPHZHMDZSHZOWSFAPPDTSPQUZWYMXUZUHSX
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

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 letter to right (wrapping back to start from end)
 3. if both letters fall in the same column, replace each with the letter below it (wrapping to top from bottom)
 4. 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 $26 \times 26 = 676$ 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 the mathematician Lester Hill in 1929. The encryption algorithm takes
- m successive plaintext letters and substitutes for them m ciphertext letters. The substitution is determined by m linear equations in which
- each character is assigned a numerical value ($a = 0, b = 1 \dots z = 25$).

Hill Cipher

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} \text{ mod } 26$$

- $\mathbf{C} = \mathbf{KP}$ mod 26
- where \mathbf{C} and \mathbf{P} are column vectors of length 3, representing the plaintext and ciphertext, and \mathbf{K} is a 3×3 matrix, representing the encryption key. Operations are performed mod 26.
- For example, consider the plaintext "paymoremoney" and use the encryption key

Hill Cipher

$$\mathbf{K} = \begin{pmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 19 \end{pmatrix}$$

$$\begin{pmatrix} 15 \\ 0 \\ 24 \end{pmatrix}. \text{ Then } \mathbf{K} \begin{pmatrix} 15 \\ 0 \\ 24 \end{pmatrix} = \begin{pmatrix} 375 \\ 819 \\ 486 \end{pmatrix} \bmod 26 = \begin{pmatrix} 11 \\ 13 \\ 18 \end{pmatrix} = \text{LNS. Continuing in this fashion,}$$

the ciphertext for the entire plaintext is LNSHDLEWMTRW.

Hill Cipher

Decryption requires using the inverse of the matrix \mathbf{K} . We can compute

$$\mathbf{K}^{-1} = \begin{pmatrix} 4 & 9 & 15 \\ 15 & 17 & 6 \\ 24 & 0 & 17 \end{pmatrix}$$

$$\mathbf{C} = \mathbf{E}(\mathbf{K}, \mathbf{P}) = \mathbf{PK} \bmod 26$$

$$\mathbf{P} = \mathbf{D}(\mathbf{K}, \mathbf{C}) = \mathbf{CK}^{-1} \bmod 26 = \mathbf{PKK}^{-1} = \mathbf{P}$$

Polyalphabetic Ciphers

- **polyalphabetic substitution ciphers**
- improve security using multiple cipher alphabets
- make 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
- effectively multiple caesar ciphers
- key is multiple letters long $K = k_1 \ k_2 \ \dots \ k_d$
- 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

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

Example of Vigenère Cipher

- 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*
key: deceptivedeceptivedeceptive
plaintext: wearediscoveredsaveyourself
ciphertext:ZICVTWQNGRZGVTWAVZHCQYGLMGJ

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

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*

Vernam Cipher

- ultimate defense is to use a key as long as the plaintext
- with no statistical relationship to it
- invented by AT&T engineer Gilbert Vernam in 1918
- originally proposed using a very long but eventually repeating key

Vernam Cipher

- The system can be expressed succinctly as follows:
- $c_i = p_i \text{ XOR } k_i$
- Where p_i = i th binary digit of plaintext
- k_i = ith binary digit of key
- c_i = ith binary digit of ciphertext
- = exclusive-or (XOR) operation
- Thus, the ciphertext is generated by performing the bitwise XOR of the plaintext and the key. Because of the properties of the XOR,
- decryption simply involves the same bitwise operation:
- $p_i = c_i \text{ XOR } k_i$

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
- problems in generation & 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
- Plain text message: meetmeafterthetogaparty
- 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

- is a more complex transposition
- 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
- Plain text message: attack postponed until two am

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:	TTNAAPMTSUOAODWCOIXKNLYPETZ

Row Transposition Ciphers

Key: 4 3 1 2 5 6 7

Input: t t n a a p t

m t s u o a o

d w c o i x k

n l y p e t z

Output: NSCYAUOPTTWLTMDNAOIEPAXTTOKZ

Original message of 28 letters: After first encryption

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 26 27 28

After the first transposition we have,

03 10 17 24 04 11 18 25 02 09 16 23 01 08

15 22 05 12 19 26 06 13 20 27 07 14 21 28

Row Transposition Ciphers

After second transposition we have,

17	09	05	27	24	16	12	07	10	02	22	20	03	25
15	13	04	23	19	14	11	01	26	21	18	08	06	28

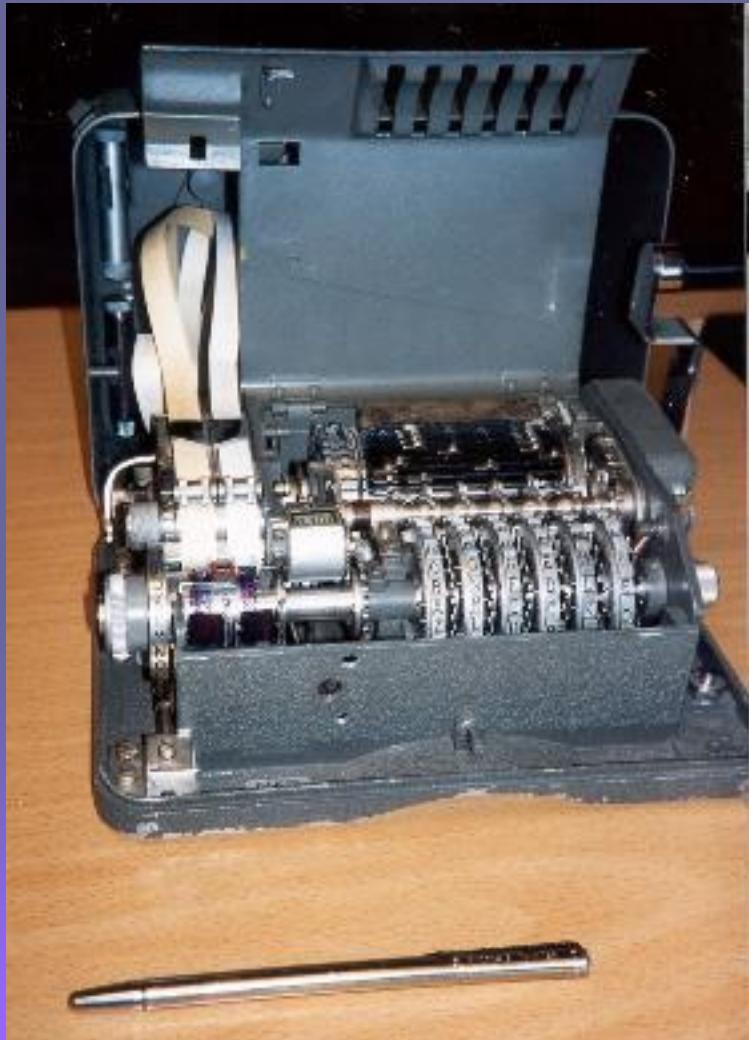
Product Ciphers

- 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

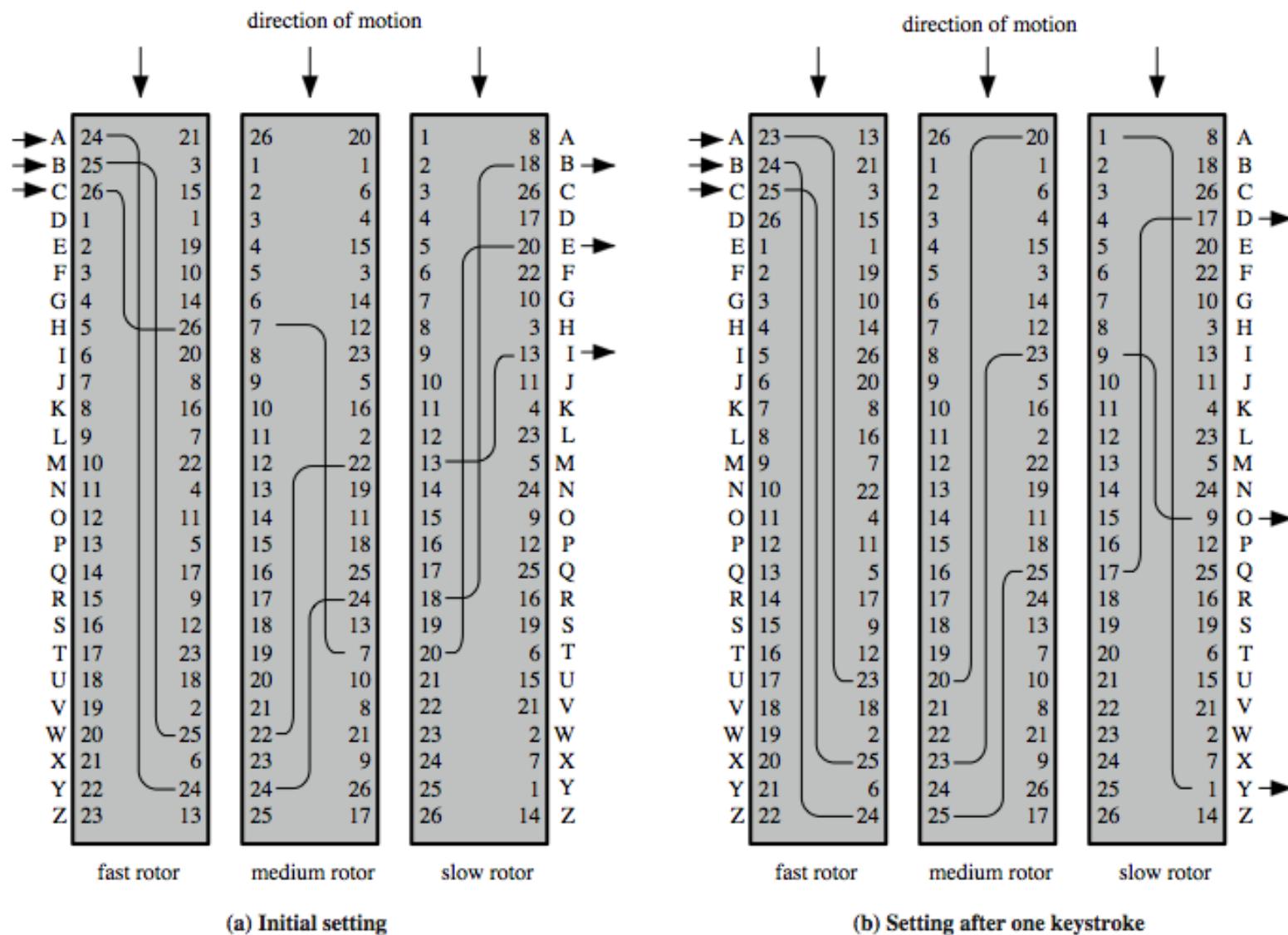
Rotor Machines

- before modern ciphers, rotor machines were most common complex ciphers in use
- 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

Hagelin Rotor Machine



Rotor Machine Principles



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
- advantage is can obscure encryption use

Summary

➤ have considered:

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