

Cryptography and Network Security

Chapter 2

Fourth Edition
by William Stallings




Chapter 2 – Classical Encryption Techniques

Many savages at the present day regard their names as vital parts of themselves, and therefore take great pains to conceal their real names, lest these should give to evil-disposed persons a handle by which to injure their owners.

—The Golden Bough, Sir James George Frazer

The bottom of the slide features several decorative concentric circles in a lighter shade of purple, resembling ripples in water, positioned in the lower right and bottom center areas.

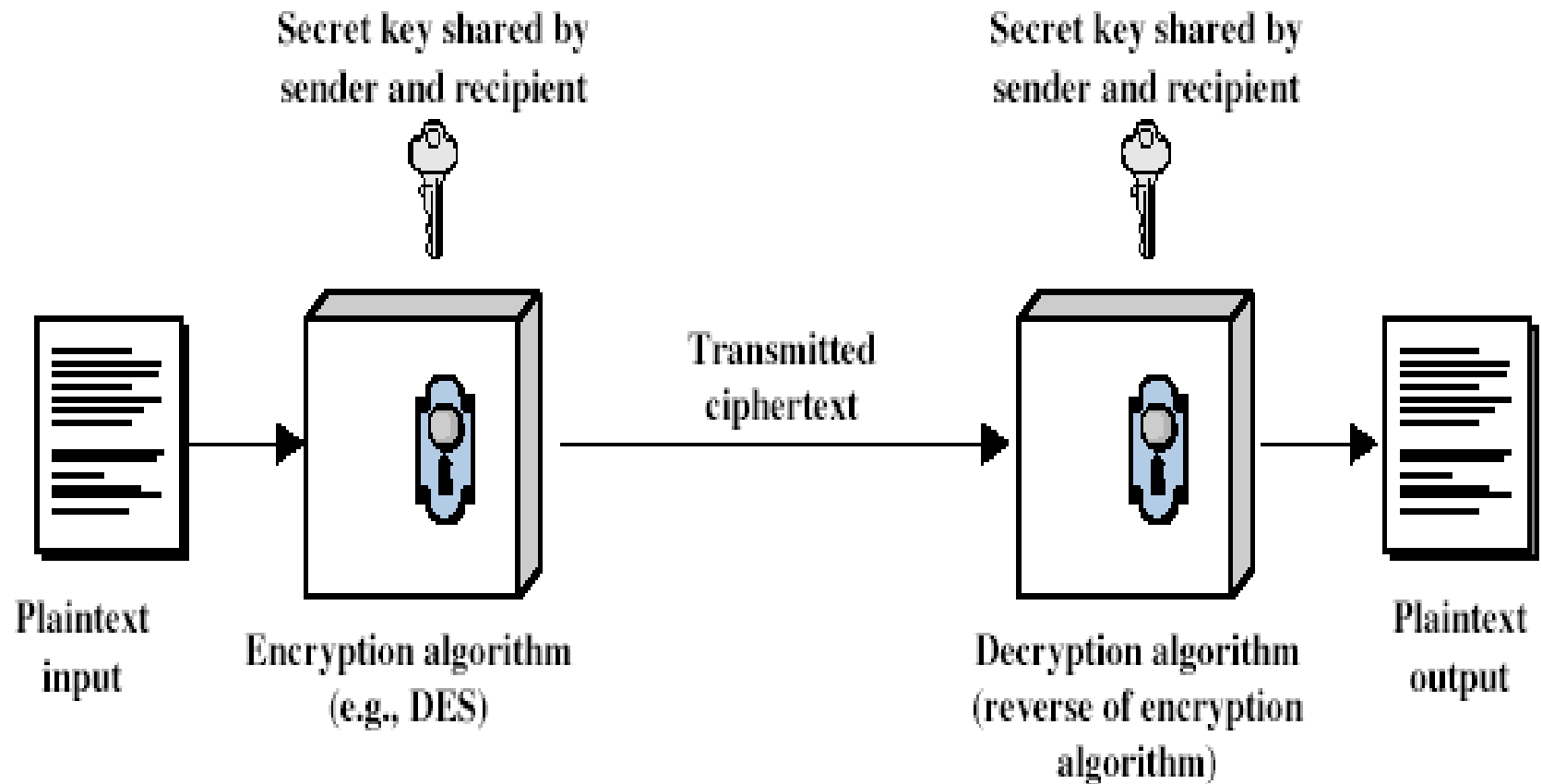
Symmetric Encryption

- 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
- 
- The bottom right corner of the slide features several decorative concentric circles in a lighter shade of purple, resembling ripples in water. There are three distinct sets of these circles, with the largest one being the most prominent.

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



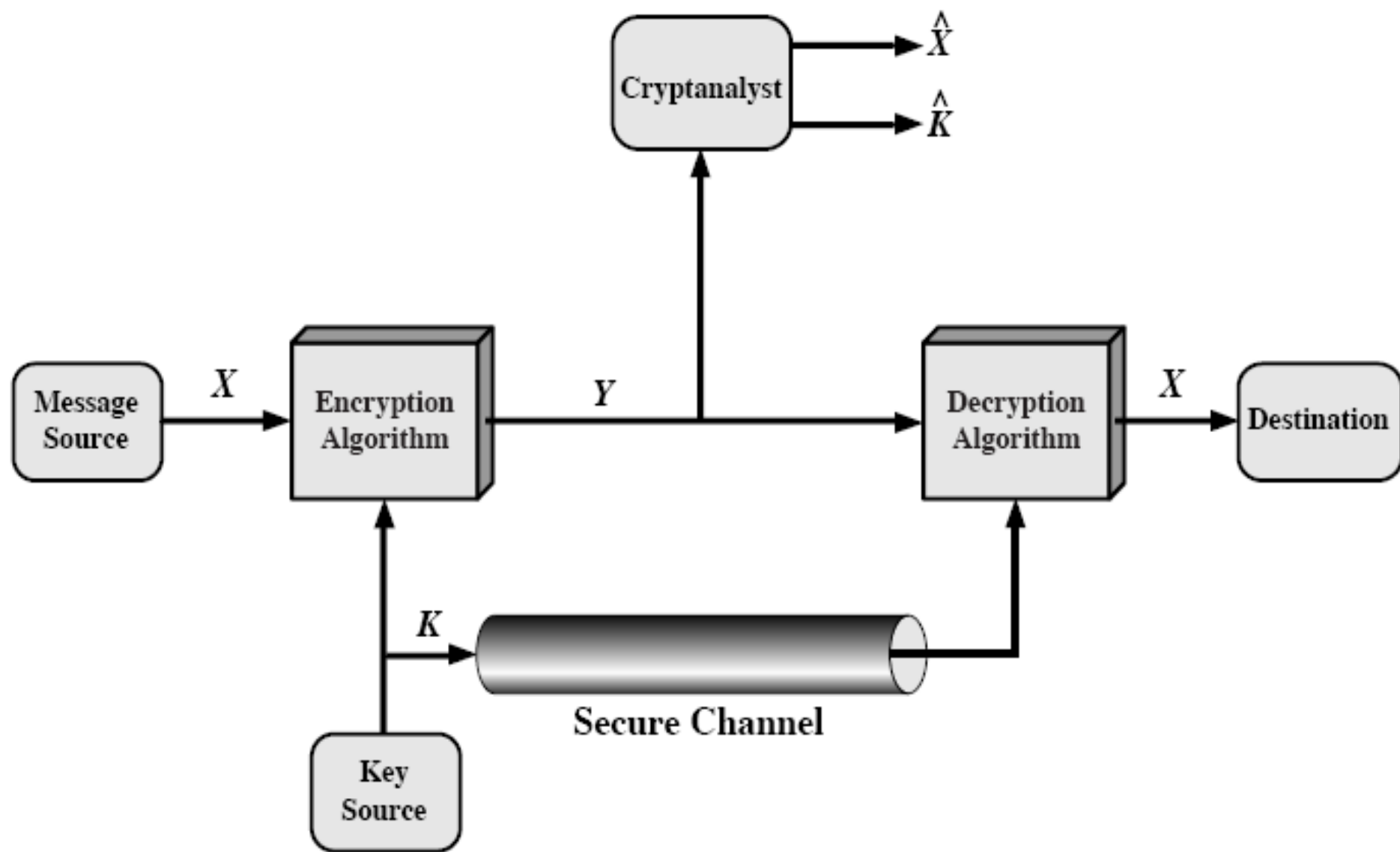


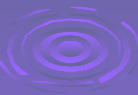
Figure 2.2 Model of Symmetric Cryptosystem

Requirements

- two requirements for secure use of symmetric encryption:
 - a strong encryption algorithm
 - a secret key known only to sender / receiver
- mathematically have:
$$Y = E_K(X) = E(K, X)$$
$$X = D_K(Y) = D(K, Y)$$
- assume encryption algorithm is known
- implies a secure channel to distribute key

Cryptography

- 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



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

Table 2.1 Types of Attacks on Encrypted Messages

Type of Attack	Known to Cryptanalyst
Ciphertext only	<ul style="list-style-type: none">•Encryption algorithm•Ciphertext
Known plaintext	<ul style="list-style-type: none">•Encryption algorithm•Ciphertext•One or more plaintext-ciphertext pairs formed with the secret key
Chosen plaintext	<ul style="list-style-type: none">•Encryption algorithm•Ciphertext•Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key
Chosen ciphertext	<ul style="list-style-type: none">•Encryption algorithm•Ciphertext•Purported ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key
Chosen text	<ul style="list-style-type: none">•Encryption algorithm•Ciphertext•Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key•Purported ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key

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

➤ **computational security**

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

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/ μ s	Time required at 10^6 decryptions/ μ s
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)	$26! = 4 \times 10^{26}$	$2 \times 10^{26} \mu\text{s} \times 10^{12} \text{ years}$	$6.4 \times 10^6 \text{ years}$

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

" PHHW PH DIWHU WKH WRJD SDUWB "

KEY		PHHW	PH	DIWHU	WKH	WRJD	SDUWB
	1	oggv	og	chvgt	vjg	vqic	rectva
	2	nffu	nf	bgufs	uif	uphb	qbsuz
	3	meet	me	after	the	toga	party
	4	ldds	ld	zesdq	sgd	snfz	ozqsx
	5	kccr	kc	ydrp	rfe	rmey	nyprw
	6	jbbq	jb	xcqbo	qeb	qldx	mxoqv
	7	iaap	ia	wbpan	pda	pkcw	lwnpu
	8	hzzo	hz	vaozm	ocz	objv	kvmot
	9	gyyn	gy	uznyl	nby	niau	julns
	10	fxxm	fx	tymxk	max	mhzt	itkmr
	11	ewwl	ew	sxlwj	lzw	lgys	hsjlg
	12	dvvk	dv	rwkvi	kyv	kfxr	grikp
	13	cuuj	cu	qvjuh	jxu	jewq	fghjo
	14	btiti	bt	puittg	iwt	idvp	epgin
	15	assh	as	othsf	hvs	hcuo	dofhm
	16	zrrg	zr	nsgr	gur	gbtn	cnegl
	17	yqqf	yq	mrfqd	ftq	fasm	bmdfk
	18	xppe	xp	lqepc	esp	ezrl	alcej
	19	wood	wo	kpdob	dro	dyqk	zkbdi
	20	vnnc	vn	jocna	cqn	cxpj	yjach
	21	ummb	um	inbmz	bpm	bwoi	xizbg
	22	tlla	tl	hmaly	aol	avnh	whyaf
	23	skkz	sk	glzqx	znk	zumg	vgxze
	24	rjjy	rj	fkyjw	ymj	ytlf	ufwyd
	25	qiix	qi	ejxiv	xli	xske	tevxc

Figure 2.3 Brute-Force Cryptanalysis of Caesar Cipher

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

The background of the slide features several sets of concentric circles in a light blue color, resembling ripples in water. These circles are positioned in the lower right and bottom center areas of the slide.

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/ μ s	Time required at 10^6 decryptions/ μ s
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)	$26! = 4 \times 10^{26}$	$2 \times 10^{26} \mu\text{s} \times 10^{12} \text{ years}$	$6.4 \times 10^6 \text{ years}$

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

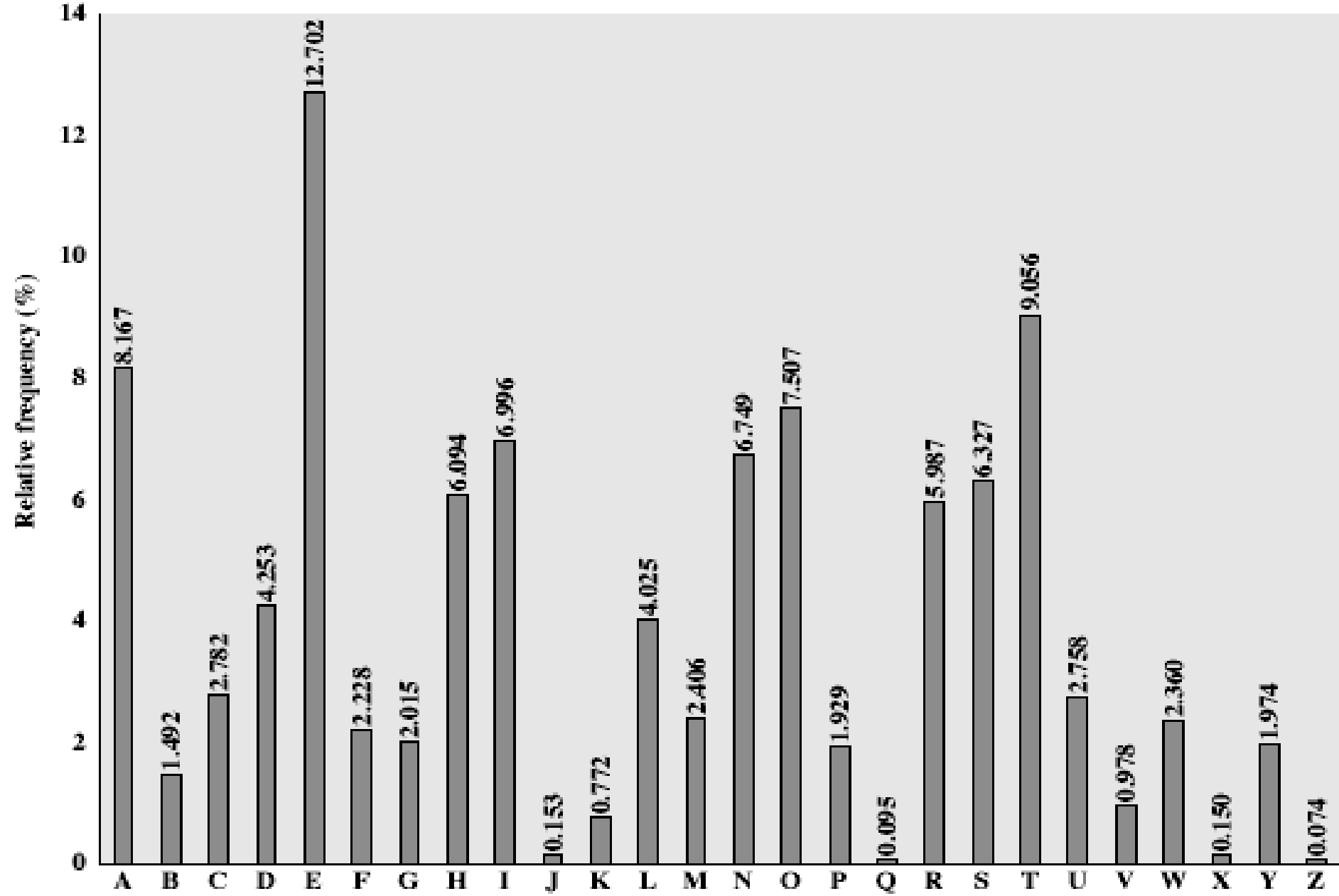


Language Redundancy and Cryptanalysis

- 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



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

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 (again 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

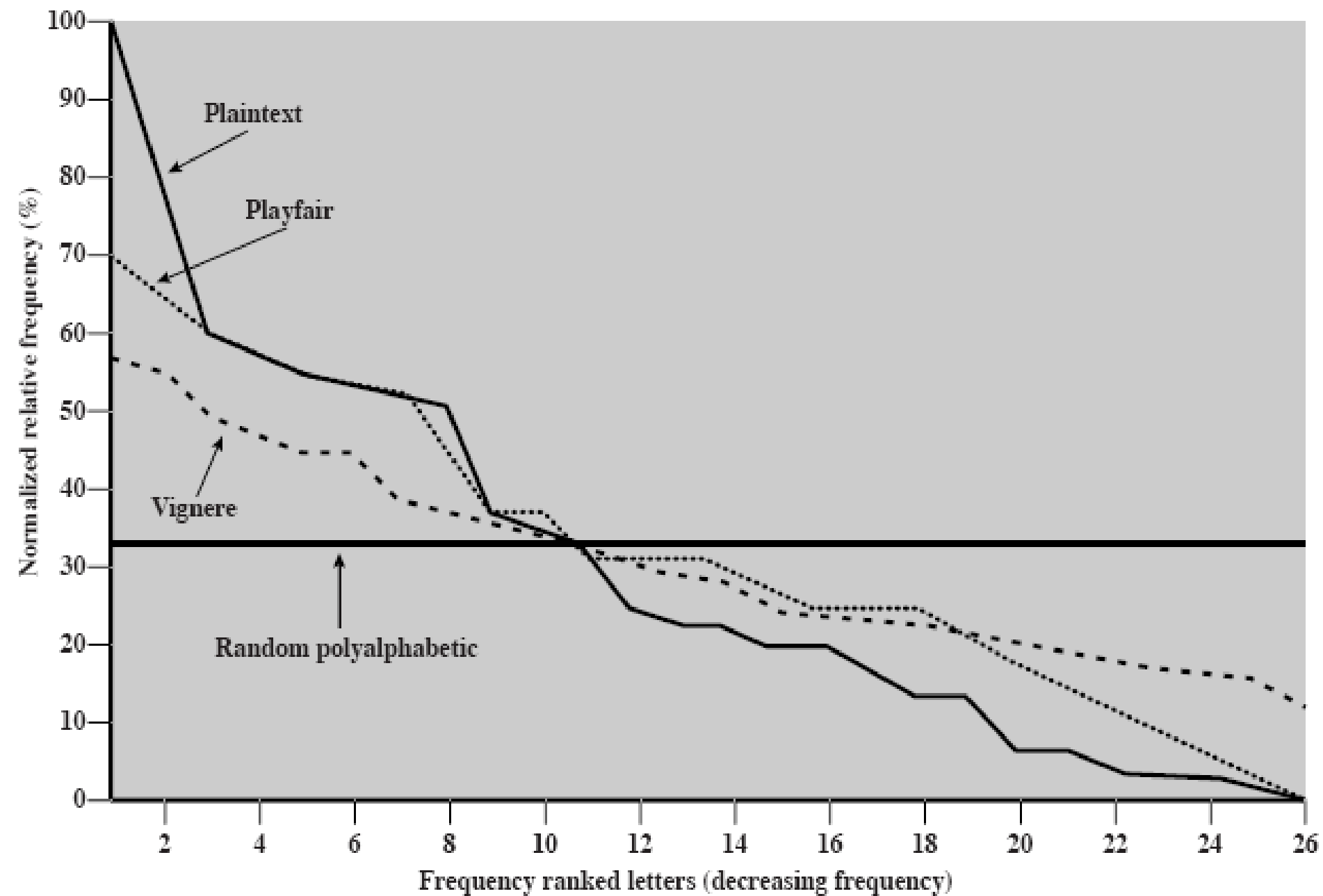


Figure 2.6 Relative Frequency of Occurrence of Letters

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

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

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**

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: Z ICV TWQNGKZE I IGASXSTSLVVWLA

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

HILL CIPHER

$$C1 = [k11 \ p1 + k12 \ p2 + k13 \ p3] \bmod 26$$

$$C2 = [k21 \ p1 + k22 \ p2 + k23 \ p3] \bmod 26$$

$$C3 = [k31 \ p1 + k32 \ p2 + k33 \ p3] \bmod 26$$

$$C = K P \bmod 26$$

Example: Plaintext "PAY"

$$K = \begin{bmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 9 \end{bmatrix}$$

$$\begin{bmatrix} C1 \\ C2 \\ C3 \end{bmatrix} = \begin{bmatrix} 17 & 17 & 5 \\ 21 & 18 & 21 \\ 2 & 2 & 9 \end{bmatrix} \times \begin{bmatrix} 15 \\ 0 \\ 24 \end{bmatrix} \bmod 26$$

$$= \begin{bmatrix} 11 \\ 13 \\ 18 \end{bmatrix} = \begin{bmatrix} L \\ N \\ S \end{bmatrix}$$

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
- 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

- 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

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

- 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



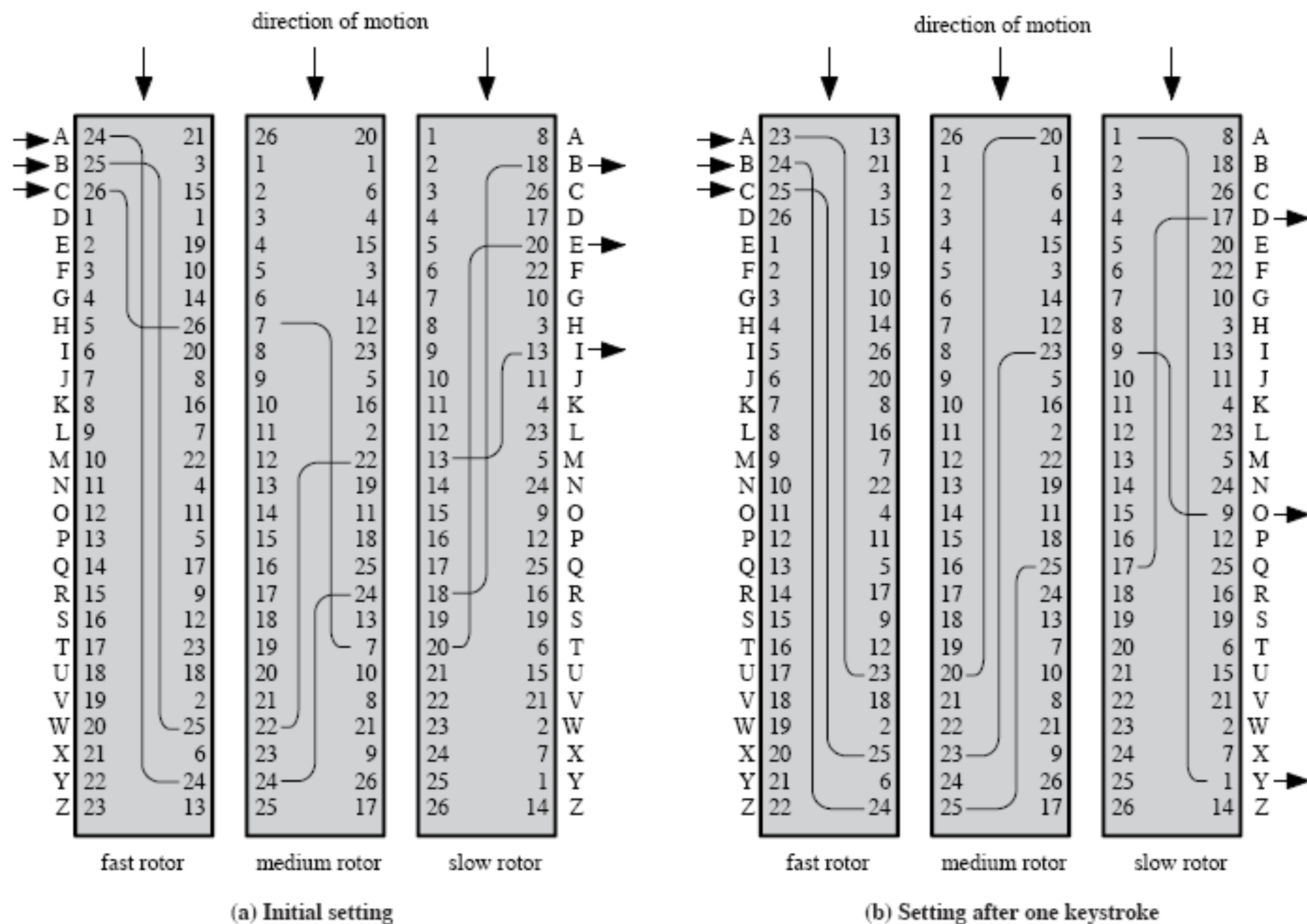


Figure 2.7 Three-Rotor Machine With Wiring Represented by Numbered Contacts

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

3rd March

Dear George,

Greetings to all at Oxford. Many thanks for your letter and for the Summer examination package. All Entry Forms and Fees Forms should be ready for final despatch to the Syndicate by Friday 20th or at the very latest, I'm told, by the 21st. Admin has improved here, though there's room for improvement still; just give us all two or three more years and we'll really show you! Please don't let these wretched 16+ proposals destroy your basic O and A pattern. Certainly this sort of change, if implemented immediately, would bring chaos.

Sincerely yours,

Figure 2.8 A Puzzle for Inspector Morse
(from *The Silent World of Nicholas Quinn*, by Colin Dexter)

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