

Software Security and Dependability

ENGR5560G

Lecture 02

Cryptography Tools

Dr. Khalid A. Hafeez

Spring, 25



Lecture Outline

- Symmetric Cryptography
- Model of Symmetric Cryptography
- Classification of Cryptographic System
- Substitution Techniques
 - Caesar Cipher
 - Monoalphabetic Cipher
 - Playfair Cipher
 - Hill Cipher
 - Polyalphabetic Ciphers
 - Vigenère Cipher
 - Vernam Cipher
 - One-Time Pad
- Transposition Techniques
 - Rail Fence Cipher
 - Row Transposition Cipher





Cryptography

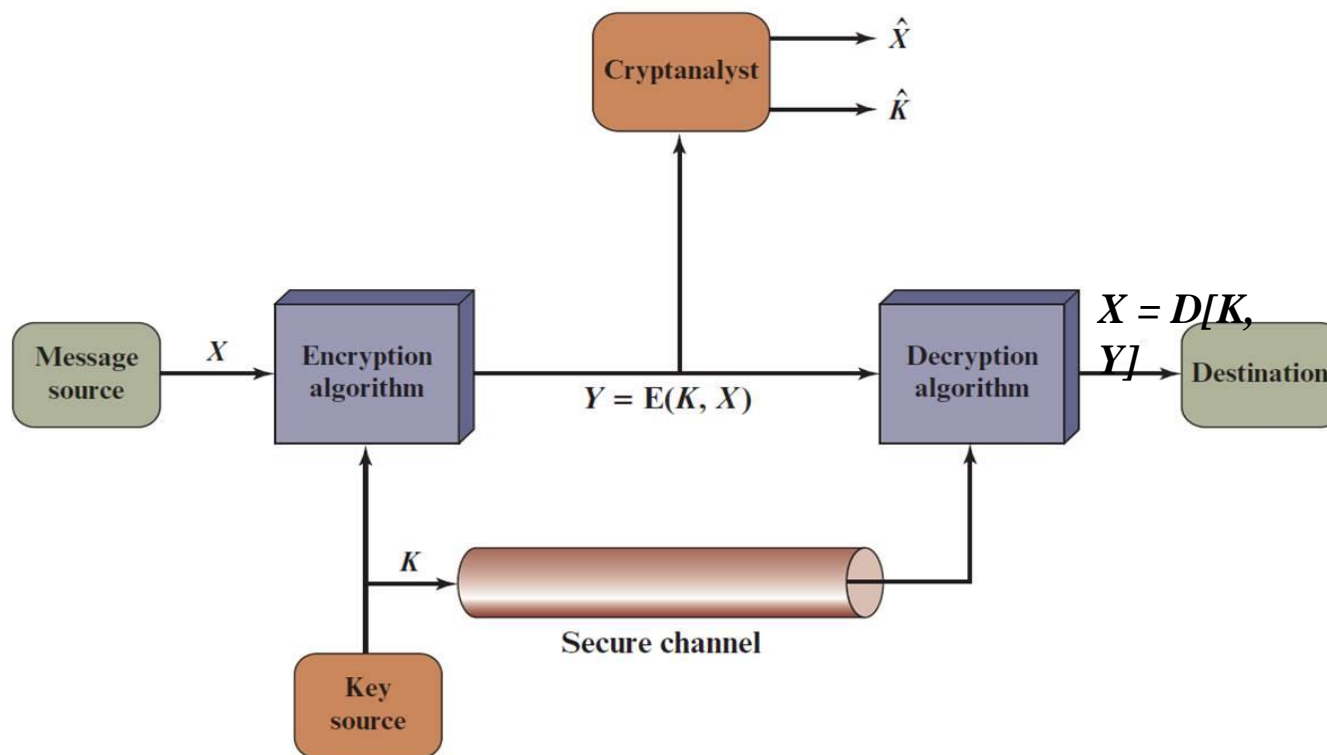
- **Cryptography:**
 - Is a science of secret writing with the goal of **hiding** the meaning of a message.
- **Cryptanalysis**
 - The science of breaking cryptosystems
- **Cryptology**
 - The areas of cryptography and cryptanalysis
- **Plaintext**
 - An original message
- **Ciphertext**
 - The coded message





Symmetric Cryptography

- It needs: Secret, Single, and Shared key
- The security of symmetric encryption depends on the secrecy of the key, not the secrecy of the algorithm
- **Two requirements for Secure use of Conventional Encryption:**
 - Need a strong encryption algorithm
 - Sender and receiver must have obtained copies of the secret key in a secure fashion and must keep the key secure
- It has five ingredients:





Dimensions to Classify Cryptographic System

The type of **operations** used for transforming plaintext to ciphertext

Substitution

Each element in the plaintext is mapped into another element

Transposition

Elements in the plaintext are rearranged

The number of **keys** used

Symmetric, **single-key**, secret-key, conventional encryption

Asymmetric, **two-key**, or public-key encryption

The way in which the plaintext is **processed**

Block cipher

Process the input one block of elements at a time

Produce an output block for each input block

Stream cipher

Process the input elements continuously

Produce output one element at a time, as it goes along





Approaches to Attack Encryption

- **Cryptanalysis**

- Relies on the nature of the **algorithm** plus some knowledge of the general **characteristics** of the plaintext to **deduce a specific plaintext** or to deduce the **key** being used

- **Brute-force attack**

- Tries **every possible key** on a piece of ciphertext until an intelligible translation into plaintext is obtained
- On average, **half of all possible keys** must be tried to achieve success





Encryption Scheme Security

■ Unconditionally secure

- No matter how much time an opponent has, it is impossible for him or her to decrypt the ciphertext simply because the required information is not there
- There is no encryption algorithm that is unconditionally secure then **computationally** secure is the goal

■ Computationally secure

- 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

■ Properties that make an encryption algorithm strong are:

- Appropriate choice of cryptographic **algorithm**
- Use of sufficiently **long key lengths**
- Appropriate choice of **protocols**
- A well-engineered **implementation**
- Absence of deliberately introduced **hidden flaws**





Substitution Techniques





Substitution Techniques

- Is one in which the letters of plaintext are **replaced** by other letters or by numbers or symbols
- If the plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with ciphertext bit patterns





Caesar Cipher

- The **Simplest** and **earliest** known use of a substitution cipher
- Used by Julius Caesar
- **Replace** each letter of the alphabet with the letter standing **three** places further down the alphabet
 - 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
 - Alphabet is wrapped around so that the letter following Z is A

Example:

Plain: meet me after the toga party

Cipher: PHHW PH DIWHU WKH WRJD SDUWB





Caesar Cipher Algorithm

- 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

- Replace each letter of the alphabet with the letter standing **three places further down** the alphabet
- Algorithm can be expressed as:

$$c = E(3, p) = (p + 3) \bmod (26)$$

A shift may be of any amount, so that the general Caesar **Encryption** algorithm is:

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

Where k takes on a value in the range 1 to 25;

- The **decryption** algorithm is simply:

$$P = D(k, C) = (C - k) \bmod 26$$





- **Brute-Force Cryptanalysis of Caesar Cipher**

can be easily performed:
simply try all the 25 possible
keys

How to Increase key space

KEY	PHHW	PH	DIWHU	WKH	WRJD	SDUWB
1	oggv	og	chvgt	vjg	vqic	rctva
2	nffu	nf	bgufs	uif	uphb	qbsuz
3	meet	me	after	the	toga	party
4	ldds	ld	zesdq	sgd	snfz	ozqsx
5	kccr	kc	ydrpc	rfc	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	fqhjo
14	btti	bt	puitg	iwt	idvp	epgin
15	assh	as	othsf	hvs	hcuo	dofhm
16	zrrg	zr	nsgre	gur	gbtn	cnegl
17	yqqf	yq	mrfqd	ftq	fasm	bmdfk
18	xppe	xp	lqepc	esp	ezrl	alcej
19	wood	wo	kpdob	dro	dyqk	zkbdj
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



Monoalphabetic Cipher

- If the “cipher” line can be **any permutation of the 26 alphabetic characters**, then there are $26!$ or greater than 4×10^{26} possible keys
 - Approach is referred to as a ***monoalphabetic substitution*** cipher because a **single cipher alphabet is used per message**

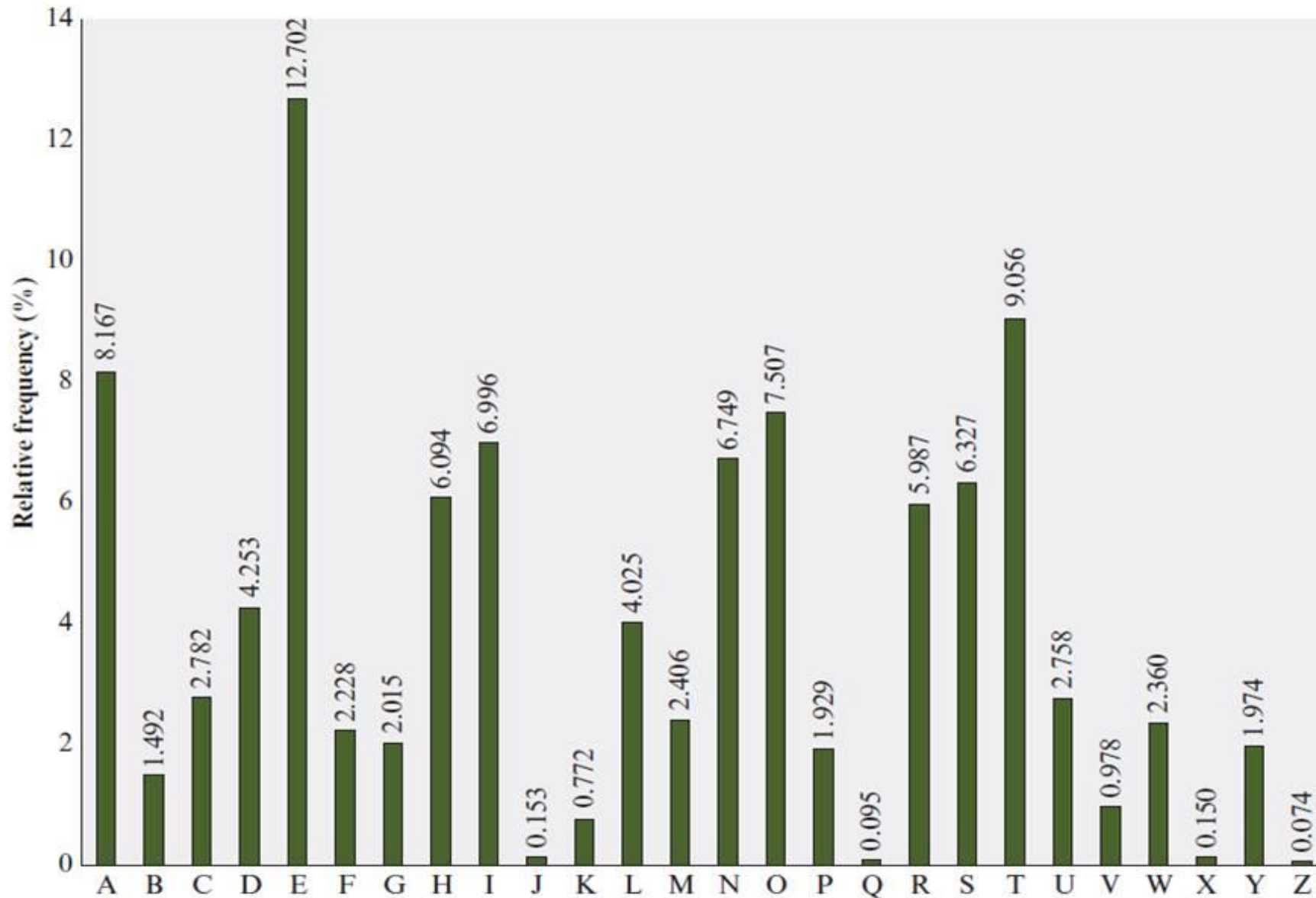
Example:

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





Cryptanalysis: Relative Frequency of Letters in English Text





Monoalphabetic Ciphers

- Easy to break because they reflect the frequency data of the original alphabet

Cypher text:

UZQSOVUOHXMOPVGPOZ**P**EVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQU**ZW**YMXUZUHSX
EPYEPOPDZSZUFPOMB**ZWP**FUPZHMDJUDTMOHMQ

Relative frequencies of the letters in the 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				

- Countermeasure is to provide multiple substitutes (homophones) for a single letter
 - Most common is *e*





Monoalphabetic Ciphers

- Easy to break because they **reflect the frequency data of the original alphabet**

Cypher text:

UZQSOVUOHXMOPVGPOZ**P**EVSG**ZW**SZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQU**ZW**YMXUZUHSX
EPYEPOPDZSZUFPOMB**ZWP**FUPZHMDJUDTMOHMQ

- **Digram**

- Two-letter combination
- Most common is **th**
- **ZW** ← in the ciphertext

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
t a e e te a that e e a
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
e t ta t ha e ee a e th t a
EPYEPOPDZSZUFPOMBZWPFPUPZHMDJUDTMOHMQ
e e e tat e the t

- **Trigram**

- Three-letter combination
- Most frequent is **the**
- **ZWP** ← in the ciphertext

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





Playfair Cipher

- Invented by British scientist Sir Charles Wheatstone in 1854
- Best-known **multiple-letter** encryption cipher
- Treats **digrams in the plaintext as single units** and translates these units into ciphertext digrams
- Based on the use of **a 5×5 matrix** of letters constructed using a keyword
- Used as the standard field system by the British Army in World War I and the U.S. Army and other Allied forces during World War II
- Fill in letters of **keyword** (minus duplicates) from **left to right** and from **top to bottom**, then fill in the remainder of the matrix with the remaining letters in alphabetic order
- **Example:**
 - 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



Playfair Cipher

Plaintext is encrypted two letters at a time, according to the following rules:

1. **Repeating plaintext** letters that are in the **same pair** are separated with a filler letter, such as **x**, so that **balloon** would be treated as **ba lx lo on**.
2. **Two plaintext letters** that fall in the **same row** of the matrix are each replaced by the letter to the right, with the first element of the row circularly following the last. For example, **ar** is encrypted as **RM**.
3. **Two plaintext letters** that fall in the **same column** are each replaced by the letter beneath, with the top element of the column circularly following the last. For example, **mu** is encrypted as **CM**.
4. Otherwise, each plaintext letter in a pair is replaced by the letter that lies in its **own row and the column** occupied by the **other plaintext** letter. Thus, **hs** becomes **BP** and **ea** becomes **IM** (or **JM**, as the encipherer wishes).

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





Playfair Cipher

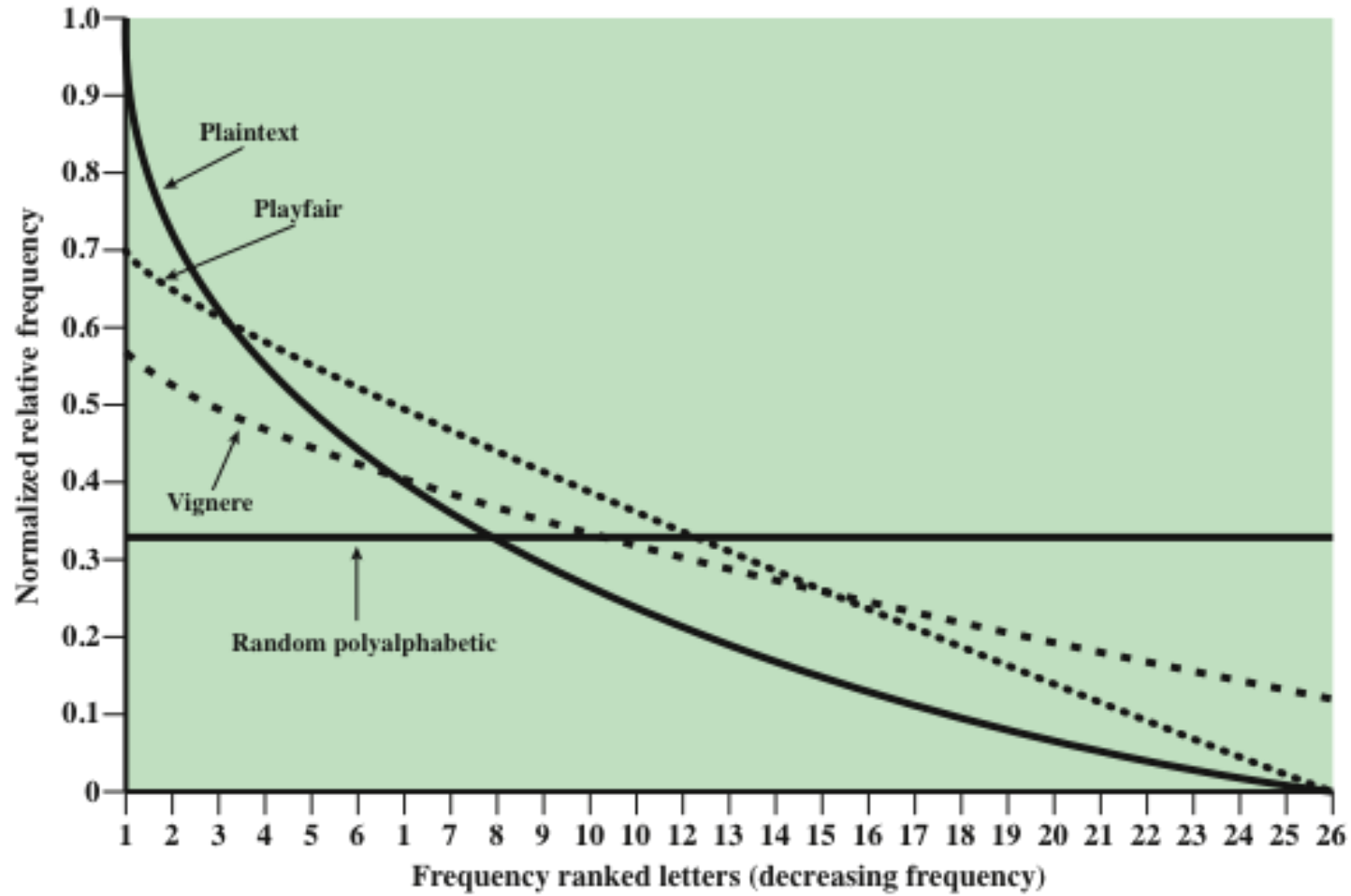


Figure 3.6 Relative Frequency of Occurrence of Letters





Hill Cipher

- Developed by the mathematician Lester Hill in 1929
- Strength is that it **completely hides single-letter frequencies**
 - The use of a larger matrix hides more frequency information
 - A 3 x 3 Hill cipher **hides not only single-letter but also two-letter frequency information**
- Strong against a ciphertext-only attack but easily broken with a known plaintext attack
- This 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 ($a = 0, b = 1, \dots, z = 25$).

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

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

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





Polyalphabetic Ciphers

- Polyalphabetic substitution cipher:
 - Improves on the simple monoalphabetic technique by using different monoalphabetic substitutions as one proceeds through the plaintext message
- All these techniques have the following features in common:
 - A set of related monoalphabetic substitution rules is used
 - A key determines which particular rule is chosen for a given transformation





Vigenère Cipher

- Best known and one of the simplest **polyalphabetic** substitution ciphers
- In this scheme the set of related monoalphabetic substitution rules consists of the 26 Caesar ciphers with shifts of 0 through 25
- Each cipher is denoted by a **key letter** which is the ciphertext letter that substitutes for the plaintext letter **a**
- Plaintext letters $P = p_0, p_1, p_2, \dots, p_{n-1}$
- Key consisting of the sequence of letters $K = k_0, k_1, k_2, \dots, k_{m-1}$, where typically $m < n$.
- Ciphertext letters $C = C_0, C_1, C_2, \dots, C_{n-1}$:

$$\begin{aligned} C &= C_0, C_1, C_2, \dots, C_{n-1} = E(K, P) = E[(k_0, k_1, k_2, \dots, k_{m-1}), (p_0, p_1, p_2, \dots, p_{n-1})] \\ &= (p_0 + k_0) \bmod 26, (p_1 + k_1) \bmod 26, \dots, (p_{m-1} + k_{m-1}) \bmod 26, \\ &\quad (p_m + k_0) \bmod 26, (p_{m+1} + k_1) \bmod 26, \dots, (p_{2m-1} + k_{m-1}) \bmod 26, \dots \end{aligned}$$





Vigenère Cipher

Example:

- To encrypt a message, a key is needed that is as long as the message
 - Usually, the key is a repeating keyword
- For example, if the keyword is *deceptive*, the message “*we are discovered save yourself*” is encrypted as:

Each character has a number from 0-25

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

key: *deceptive*deceptive*deceptive*

plaintext: wearediscoveredsaveyourself

ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ

Example: $(w+d) \bmod 26 = (22+3) \bmod 26 = 25 = Z$





Vigenère Autokey System

- The periodic nature of the keyword can be eliminated by using a nonrepeating keyword that is as long as the message itself.
- A **keyword is concatenated with the plaintext** itself to provide a running key
- Example:

key: deceptivewearediscoveredsav

plaintext: wearediscoveredsaveyourself

ciphertext: ZICVTWQNGKZEIIGASXSTSLVWLA

- Even this scheme is still vulnerable to cryptanalysis
 - Because the key and the plaintext share the same frequency distribution of letters, a statistical technique can be applied





Vernam Cipher

- AT&T engineer named Gilbert Vernam in 1918
- The ultimate defense against such a cryptanalysis is to **choose a keyword that is as long as the plaintext** and has no statistical relationship to it.

$$c_i = p_i \oplus k_i$$

where

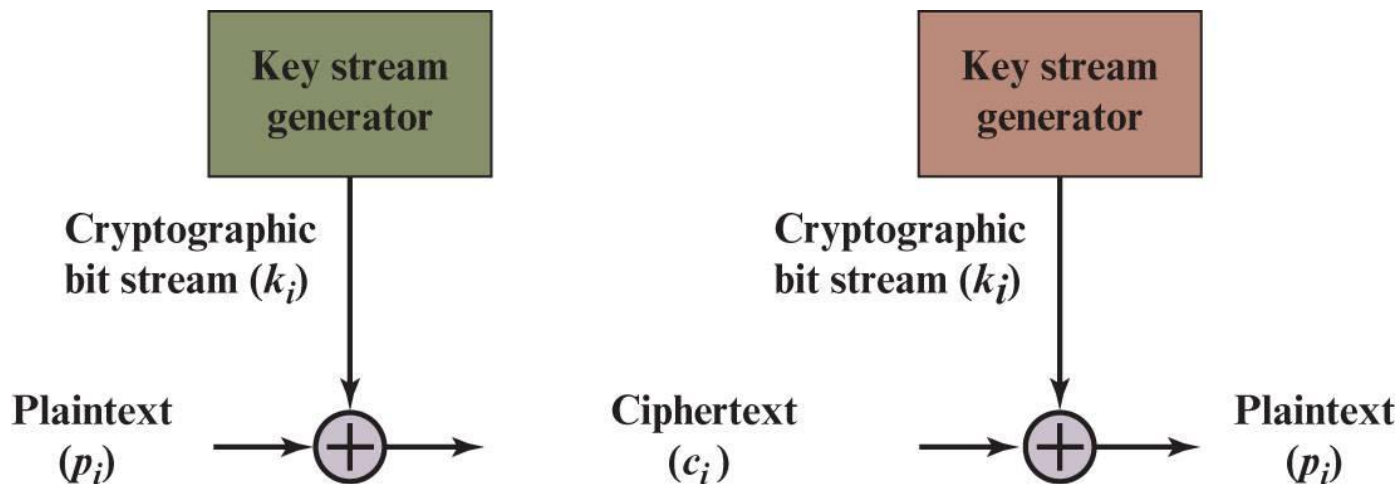
p_i = i^{th} binary digit of plaintext

k_i = i^{th} binary digit of key

c_i = i^{th} binary digit of ciphertext

\oplus = exclusive-or (XOR) operation

$$p_i = c_i \oplus k_i$$





One-Time Pad

- Improvement to Vernam cipher proposed by an Army officer, Joseph Mauborgne
- Use a **random key that is as long as the message** so that the key need not be repeated
- Key is used to encrypt and decrypt a **single message and then is discarded**
- Each **new message requires a new key** of the same length as the new message
- This Scheme is **unbreakable**
 - Produces random output that bears no statistical relationship to the plaintext
 - Because the ciphertext contains no information whatsoever about the plaintext, there is simply no way to break the code





One-Time Pad: Difficulties

- The one-time pad **offers complete security** but, in practice, has **two fundamental difficulties**:
 1. There is a practical problem of **making** large quantities of random keys
 - Any heavily used system might require millions of random characters on a regular basis
 2. Mammoth key **distribution** problem
 - For every message to be sent, a key of equal length is needed by both sender and receiver
- Because of these difficulties, the one-time pad is of limited utility
 - Useful primarily for low-bandwidth channels requiring very high security
- The one-time pad is the **only cryptosystem that exhibits *perfect secrecy***





Transposition Techniques





Rail Fence Cipher

- Simplest transposition cipher
- Plaintext is written down as a sequence of diagonals and then read off as a sequence of rows
- To encipher the message “meet me after the toga party” with a rail fence of depth 2, we would write:

m e m a t r h t g p r y
 ↘ ↙ e t e f e t e o a a t

Encrypted message is:

MEMATRHTGPRYETEFETEOAAT






Transposition Cipher

- The positions of plaintext letters

Example:

Plaintext: LAST NITE WAS HEAVEN PLEASE MARRY ME



L	A	S	T	N	I
T	E	W	A	S	H
E	A	V	E	N	P
L	E	A	S	E	M
A	R	R	Y	M	E

Cypertext: L T E L A A E A E R S W V A R T A E S Y N S N E M I H P M E





Row Transposition Cipher

- a more complex transposition
- Write the message in a rectangle, row by row, and read the message off, column by column, but **permute the order of the columns**
- The order of the columns then becomes the key to the algorithm

Message: attack postponed until two am xyz

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

- Transposition cipher can be made significantly more secure by performing **more than one stage of transposition**.

