2 Classical Encryption Techniques

ch3 in textbook

Yanwei Yu

E-mail: ywyu@ustc.edu.cn

- 可汗学院公开课 古典密码学：
- http://open.163.com/newview/movie/courseintro?newurl=%2Fspecial%2FKhan%2Fancientcryptography.html

课程列表

【第1集】 什么是密码学？ 译

【第2集】 概率空间 译

【第3集】 凯撒密码 译

【第4集】 多表密码 译

【第5集】 一次一密 译

【第6集】 频率稳定性 译

【第7集】 Enigma加密机 译

【第8集】 完全保密性 译

【第9集】 伪随机数生成器 译

- 可汗学院公开课现代密码学：
- http://open.163.com/newview/movie/courseintro?newurl=%2Fspecial%2FKhan%2Fmoderncryptography.html

课程列表

【第1集】 算术基本定理 译

【第2集】 公开密匙密码学 译

【第3集】 离散对数问题 译

【第4集】 迪菲·赫尔曼密钥交换 译

【第5集】 RSA加密：第一步 译

【第6集】 RSA加密：第二步 译

【第7集】 RSA加密：第三步 译

【第8集】 欧拉函数 译

【第9集】 RSA加密：第四步 译

【第10集】 后面应该学什么 译

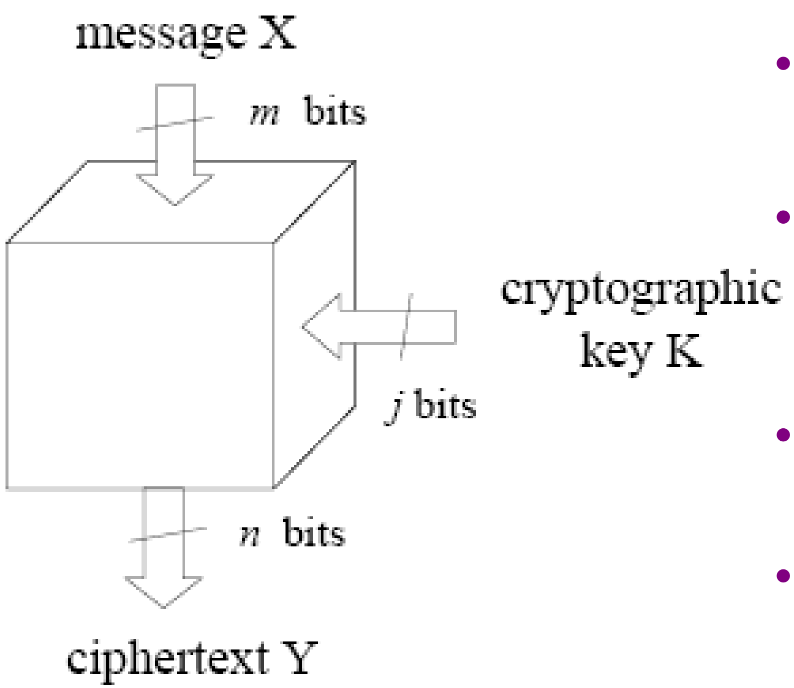
Outline

- Basic Terminology
- Why (not) to Classical Ciphers?
- Evolution of Cryptography
- Experiences and Lessons

1 Basic Terminology

- Cryptology=Crypto(secret)-log(word):
  - field of both cryptography and cryptanalysis
- Cryptography=Crypto(secret)-graph(write):
  - study of encryption principles/methods
- Cryptanalysis (codebreaking) :
  - study of principles/ methods of decrypting ciphertext *without* knowing key

1.1 Five Basic Elements

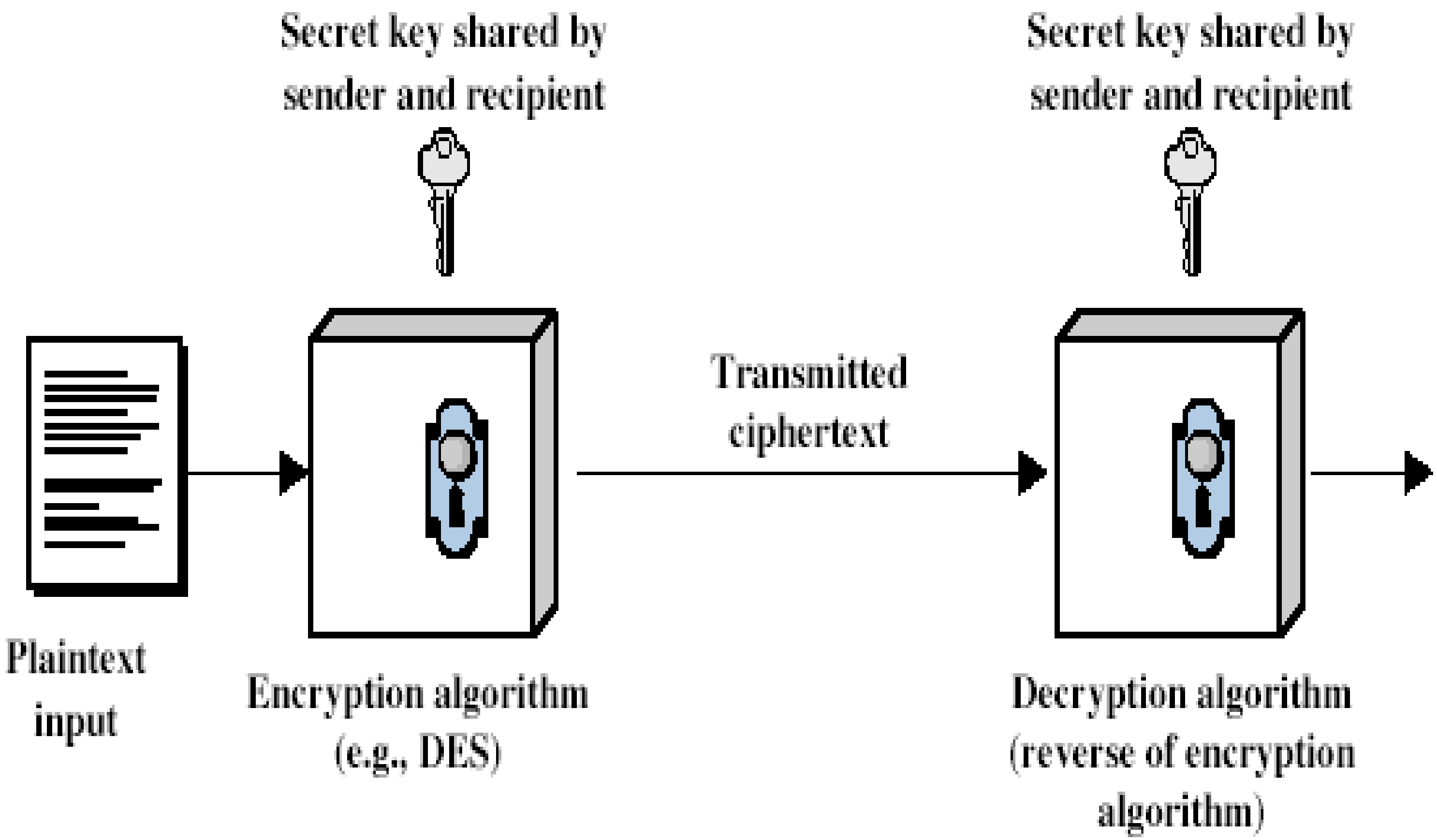


- **plaintext** - original message
- **ciphertext** - coded message
- **key** - info used in cipher known only to sender/receiver
- **encipher (encrypt)** - converting plaintext to ciphertext
- **decipher (decrypt)** - recovering ciphertext from plaintext

1.2 Symmetric Encryption

- **or conventional(传统的) / private-key / single-key**
- **sender and receiver 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**

Symmetric Cipher Model

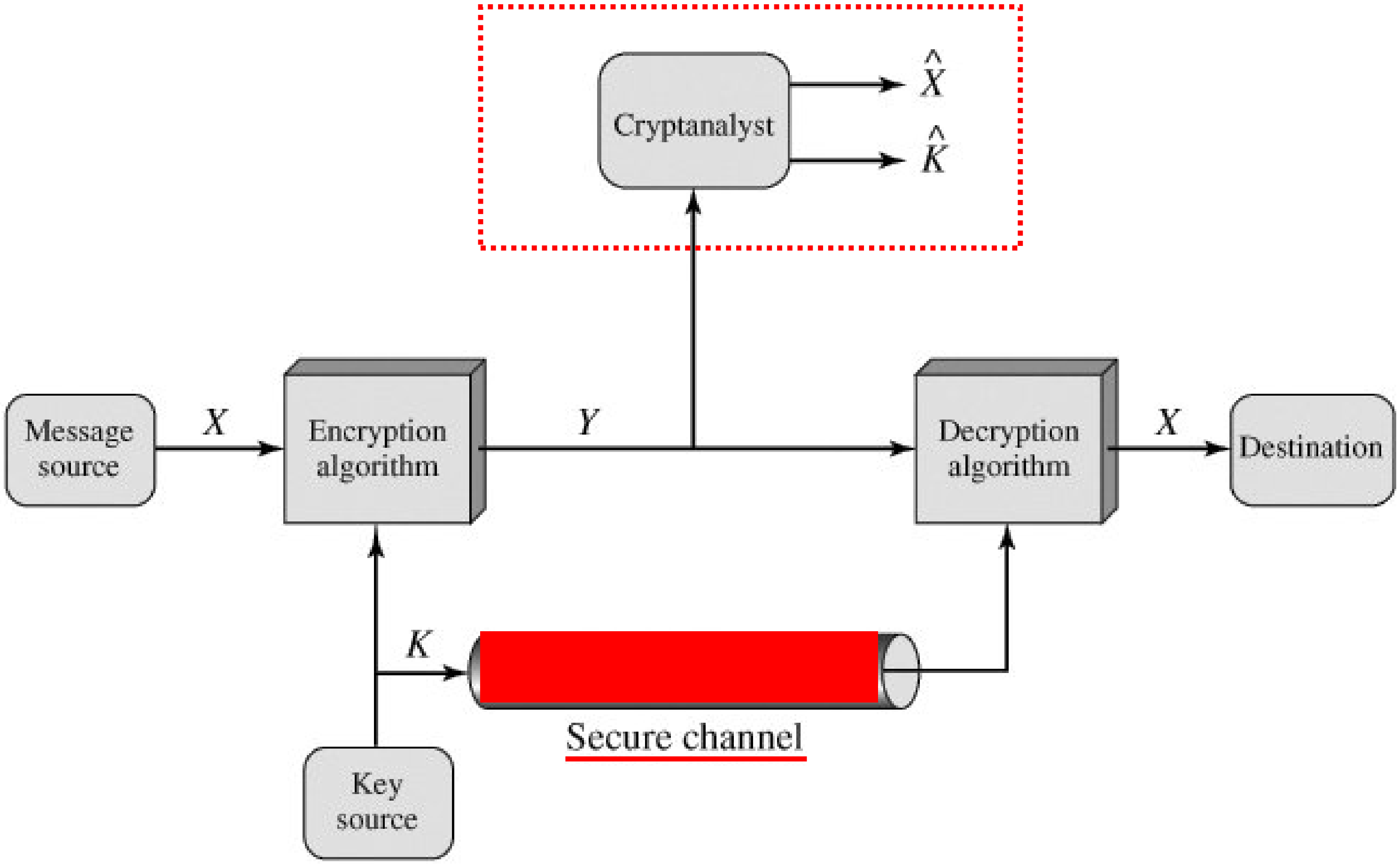


Requirements

- two requirements for secure use of symmetric encryption:
  - a strong encryption algorithm
  - a secret key known only to sender / receiver
- ***Kerckhoffs' principle***: assume encryption algorithm is known
- implies a secure channel to distribute key
- mathematically have:

$$Y = E_K(X)$$

$$X = D_K(Y)$$


Cryptanalyst
$\hat{X}$
$\hat{K}$
Message source
X
Encryption algorithm
Y
Decryption algorithm
X
Destination
K
Secure channel
Key source

1.3 Cryptography

- characterize cryptographic system by:
  - type of encryption operations used
    - substitution ;
    - transposition/permutation ;
    - product(乘积): Muliti-substitution/transposition
  - number of keys used
    - single-key or private / two-key or public
  - way in which plaintext is processed
    - block / stream

1.4 Cryptanalysis

- Objective: to recover key not just message
- general approaches:
  - – cryptanalytic attack
  - – brute-force(穷举) attack

1.4.1 Brute Force Search

- **always possible to simply try every key**
- **most basic attack, proportional(成比例) 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**

1.4.2 Cryptanalytic Attacks

- ciphertext only
  - only know algorithm & ciphertext, or can identify plaintext type
- known plaintext
  - know/suspect some plaintext-ciphertext pairs
- 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 ciphertext provides insufficient information to uniquely determine the corresponding plaintext**
- computational security
  - **given limited computing resources, the cipher cannot be broken**
    - **cost needed for calculations exceeds ciphertext value**
    - **time needed for calculations exceeds valid lifetime of ciphertext**

2 Why (not) to study Classical ciphers?

AGAINST

- Not similar to modern ciphers
- Long abandoned

FOR

- Basic components became a part of modern ciphers
- Under special circumstances modern ciphers reduce to historical ciphers
- Influence on world events
- The only ciphers you can break!

2.1 Evolution of Cryptology

- **Classical Cryptology**
  - **By hand, using electromechnical machine**

- **Modern Cryptology (after 1976)**
  - **Using computer**

Evolution of Cryptology(1)

Cryptography

Cryptanalysis

100 B.C. **(置换)Shift ciphers(加密法)**

Frequency analysis
al-Kindi, Baghdad — IX c.

(单表代换)Monoalphabetic substitution cipher

1586 Invention of the Vigenère Cipher
(维吉尼亚)

Black chambers — XVIII c.

(多表替换)Homophonic ciphers

Vigenère cipher
(Simple polyalphabetic substitution ciphers)

Kasiski's method — 1863

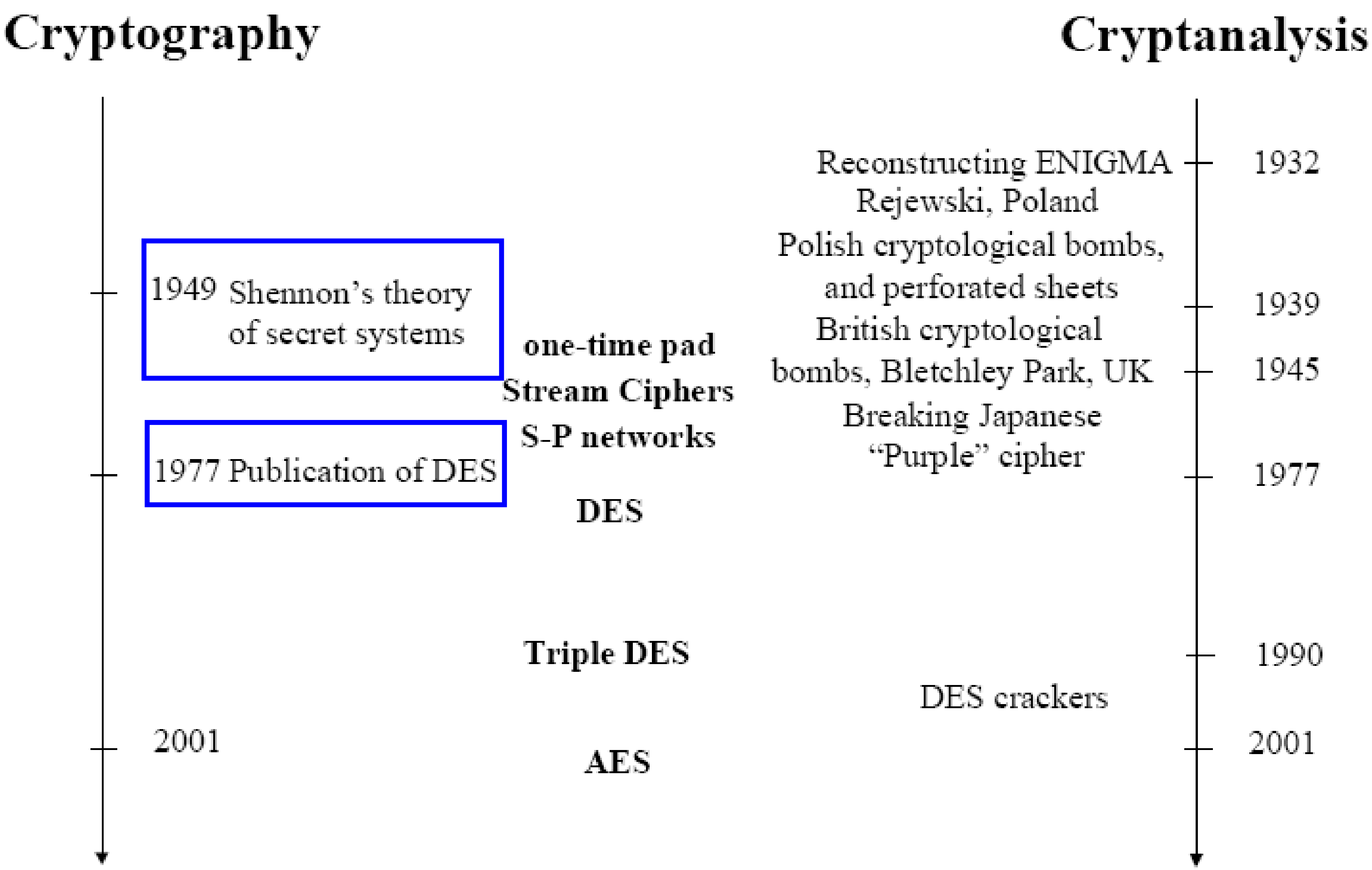
1919 Invention of rotor machines**(转轮机)**

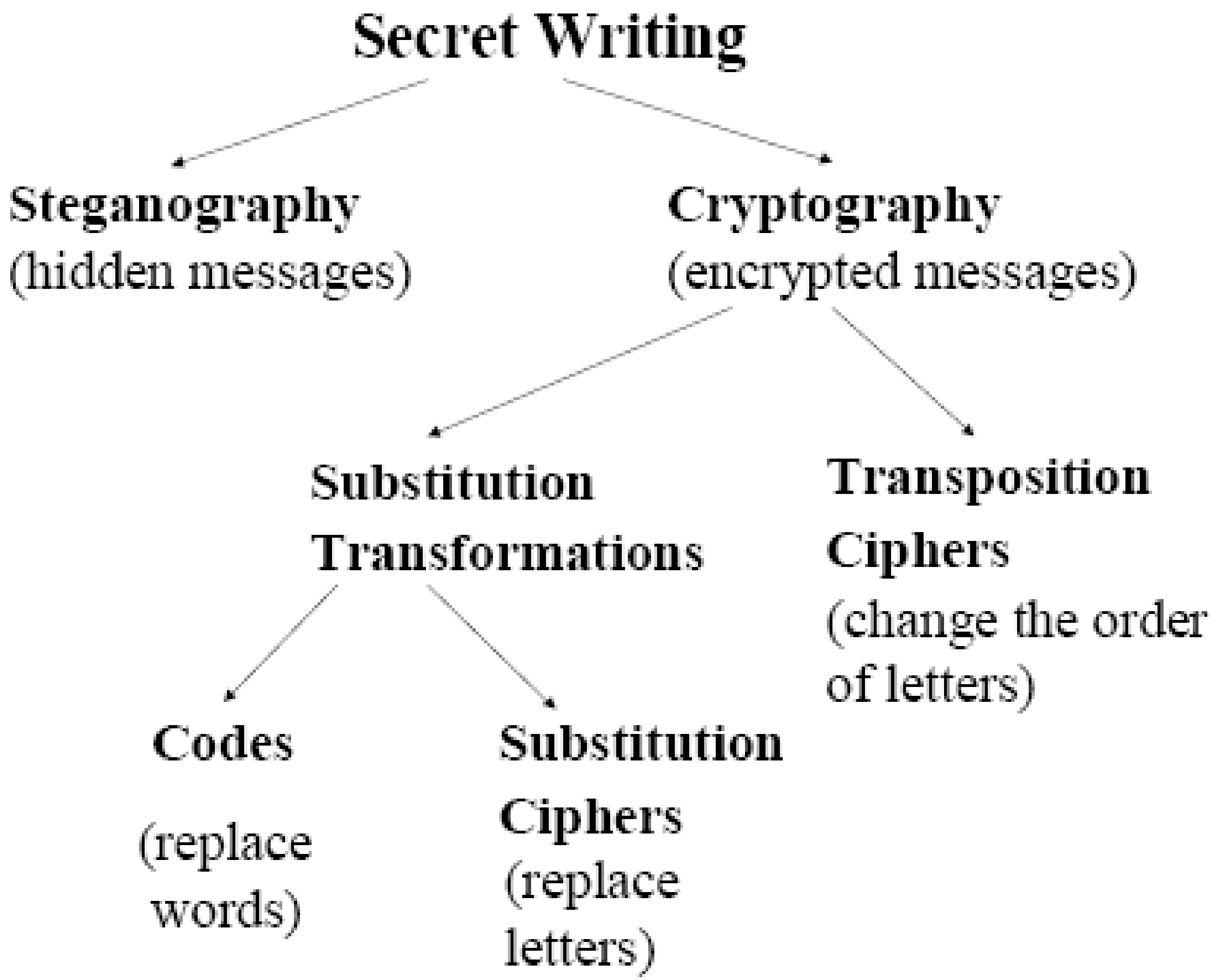
Index of coincidence
William Friedman — 1918

(机电)Electromechanical machine ciphers
(Complex polyalphabetic substitution ciphers)

1926 Vernam cipher (one-time pad)

Evolution of Cryptology(2)



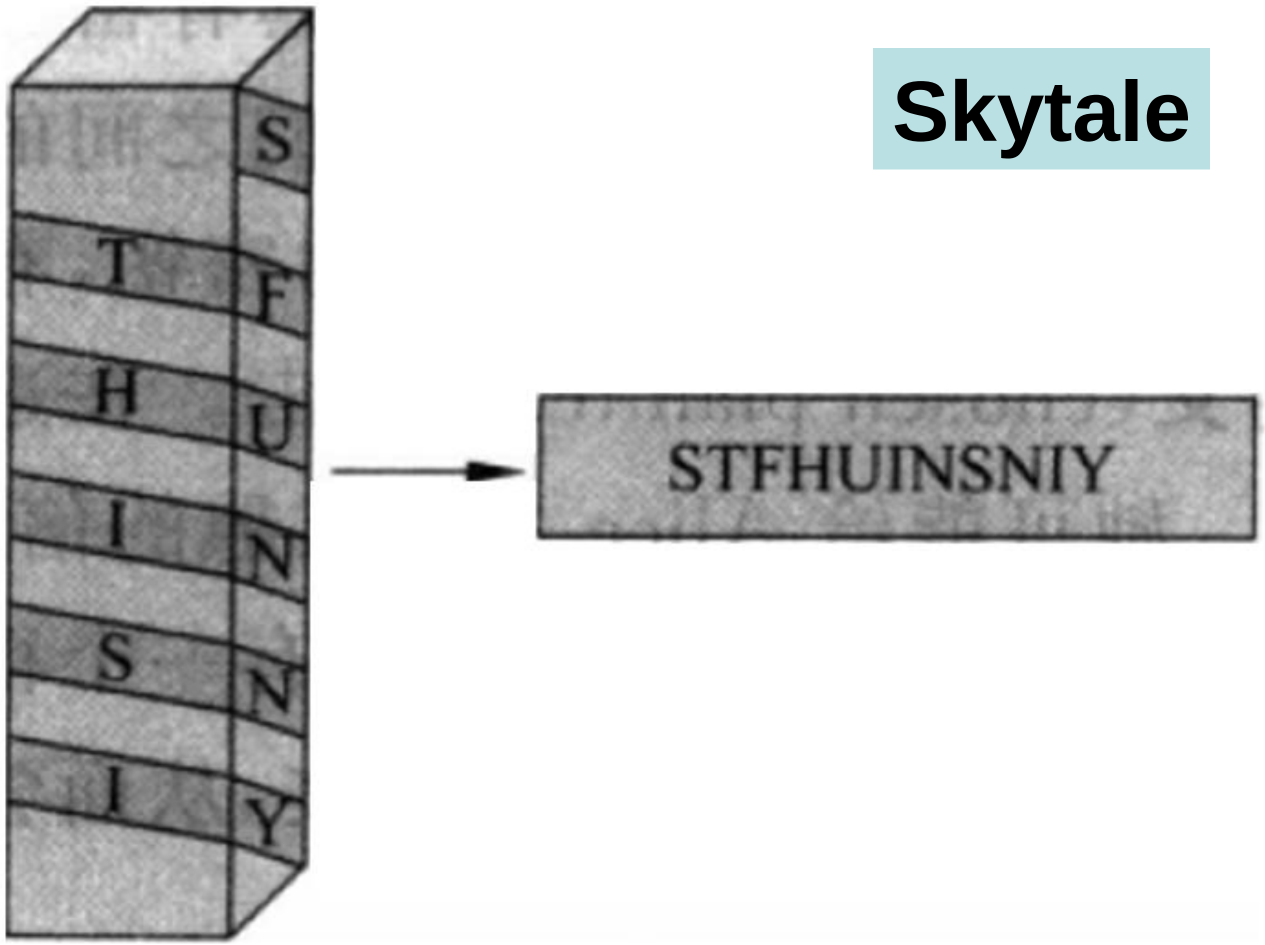

Secret Writing
Steganography
(hidden messages)
Cryptography
(encrypted messages)
Substitution
Transformations
Transposition
Ciphers
(change the order
of letters)
Codes
(replace
words)
Substitution
Ciphers
(replace
letters)

2.2 Classical Cryptology

- **Transposition /Permutation -> Substitution**
  - **No key -> have key**
  - **Substitution:**
    - **Mon-alphabetic Substitution -> Poly-alphabetic Substitution -> Combined with product**

2.2.1 Transposition /Permutation - Example 1

Skytale



2.2.1 Transposition /Permutation - Example 2

- **Plaintext**: MEET ME THURSDAY NIGHT
- **Ciphertext**:MESIETDGEHAHTUYTMRNX

M	E	E	T	M
E	T	H	U	R
S	D	A	Y	N
I	G	H	T	X

No Key!

2.2.1 Transposition /Permutation - Example 3

Plaintext: CRYPTANALYST

Key: KRIS

Encryption:

2	3	1	4
K	R	I	S
C	R	Y	P
T	A	N	A
L	Y	S	T

Ciphertext: YNSCTLRAYPAT

Class Exercise

- **Plaintext**: **MEET ME THURSDAY NIGHT**
- **Key**: **FRANK(25143)**
- **Ciphertext**:**EHAHMESIMRNXTUYTETDG**

2	5	1	4	3
M	E	E	T	M
E	T	H	U	R
S	D	A	Y	N
I	G	H	T	X

Class Exercise

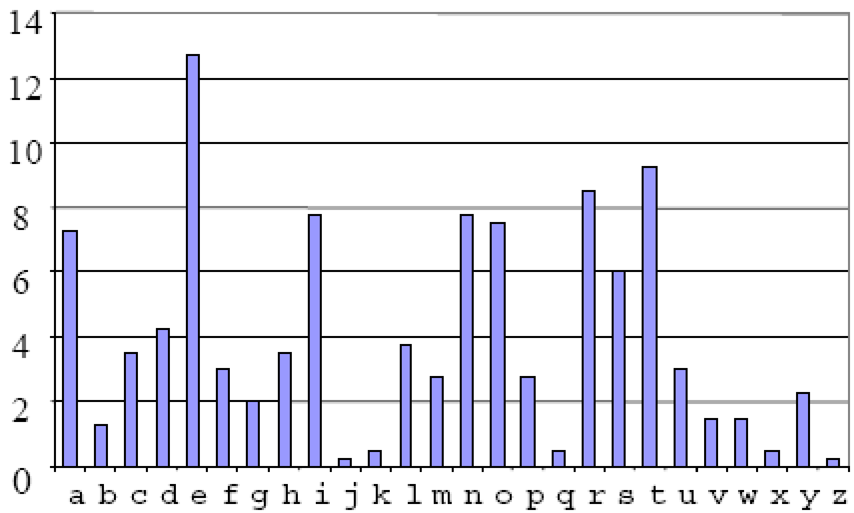
- **Ciphertext**:MESITUYTMRNXEHAHETDG
- **Key**: Alice (15423)
- **Plaintext**:
  - MEET ME THURSDAY NIGHT

1	5	4	2	3
M	E	E	T	M
E	T	H	U	R
S	D	A	Y	N
I	G	H	T	X

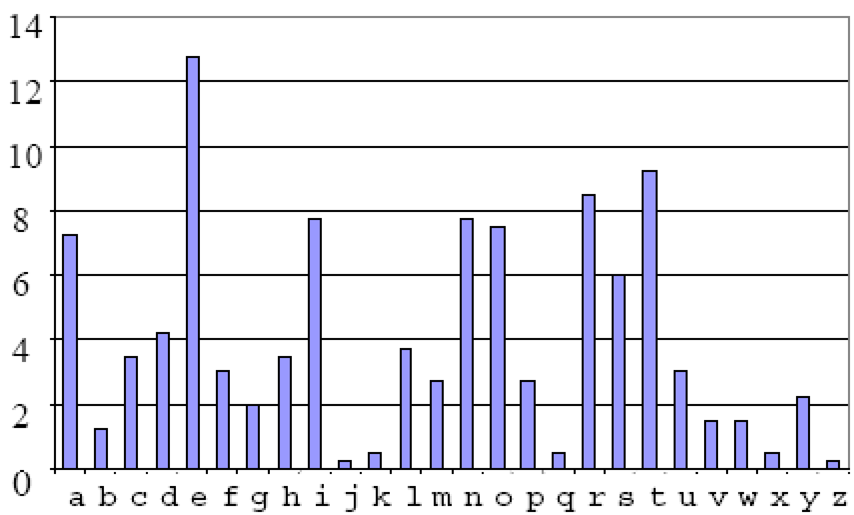
2.2.1 Transposition/Permutation Ciphers

$$
\begin{array}{lcccccc}
M = & m_1 & m_2 & m_3 & m_4 & \ldots\ldots & m_N \\
C = & m_{f(1)} & m_{f(2)} & m_{f(3)} & m_{f(4)} & \ldots\ldots & m_{f(N)}
\end{array}
$$

Letters of the plaintext are rearranged without changing them



Character frequency in a long English plaintext



Character frequency in the corresponding ciphertext for a transposition cipher

2.2.2 Monalphabetic Substitution(单表替换) Ciphers

- Caesar Cipher
- Shift Cipher
- general Monalphabetic substitution ciphers

2.2.2 Monalphabetic Substitution (单表替换) Ciphers - Caesar Cipher

- **Caesar Cipher**
  - **Coding Characters into Numbers**
  - **Using mathematical compution**

$c_i = f(m_i) = m_i + 3 \bmod 26$

$m_i = f^{-1}(c_i) = c_i - 3 \bmod 26$

No key

A	⇔	0	N	⇔	13
B	⇔	1	O	⇔	14
C	⇔	2	P	⇔	15
D	⇔	3	Q	⇔	16
E	⇔	4	R	⇔	17
F	⇔	5	S	⇔	18
G	⇔	6	T	⇔	19
H	⇔	7	U	⇔	20
I	⇔	8	V	⇔	21
J	⇔	9	W	⇔	22
K	⇔	10	X	⇔	23
L	⇔	11	Y	⇔	24
M	⇔	12	Z	⇔	25

Casear Cipher: Example

Plaintext:

11 5 3 15 7 11 21 3 25 11 5 17 16 19 23 7 20 7 6

Ciphertext: L F D P H L V D Z L F R Q T X H U H G

2.2.2 Monalphabetic Substitution (单表替换) Ciphers - Shift Cipher

- **Shift Cipher**

  $c_i = f(m_i) = m_i + k \bmod 26$

  $m_i = f^{-1}(c_i) = c_i - k \bmod 26$

  Key = k

  Number of keys = 26

Exercise in class

- Try to break ciphertext "GCUA VQ DTGCM"

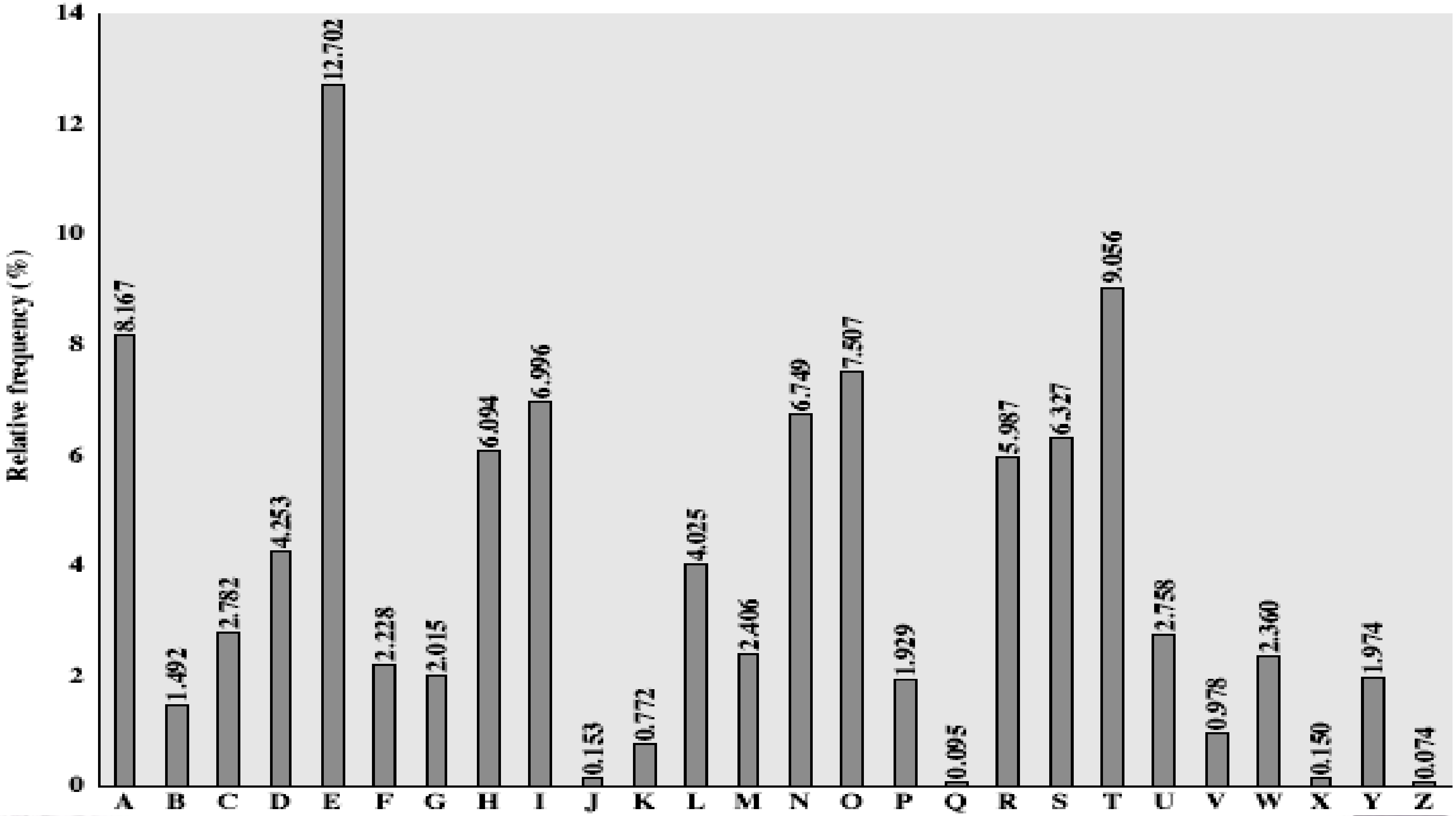
EASY TO BREAK

- **Cryptanalysis of Shift 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**

2.2.2 Monalphabetic Substitution (单表替换) Ciphers - General Ciphers

- General Monalphabetic Substitution Ciphers Security
  - now have a total of 26! = 4 x 1026 keys
    - A maps to A' from {A,B,..Z}
    - B maps to B' from {A,B,...Z} excluding A'
    - ……
  - with so many keys, might think is secure
  - but would be !!!WRONG!!!
  - problem is language characteristics

English Letter Frequencies



Average Frequency of Single Letter

- **Average fequency in a random string of letters:**

$$\frac{1}{26} = 0.038 = 3.8\%$$

- **Average fequency in a long english text:**

E	—	13%
T, N, R, I, O, A, S	—	6%-9%
D, H, L	—	3.5%-4.5%
C, F, P, U, M, Y, G, W, V	—	1.5%-3%
B, X, K, Q, J, Z	—	< 1%

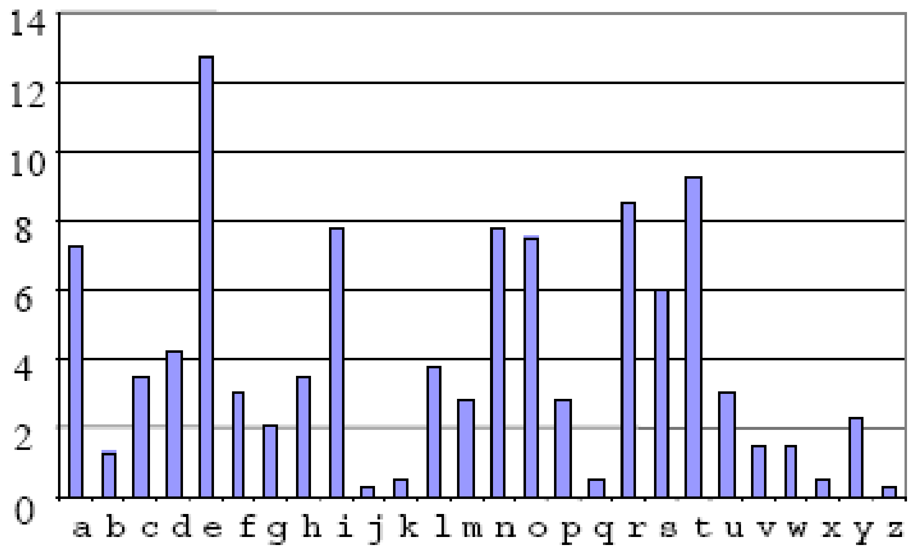
Most Frequent digrams and trigrams

- **Digrams:**

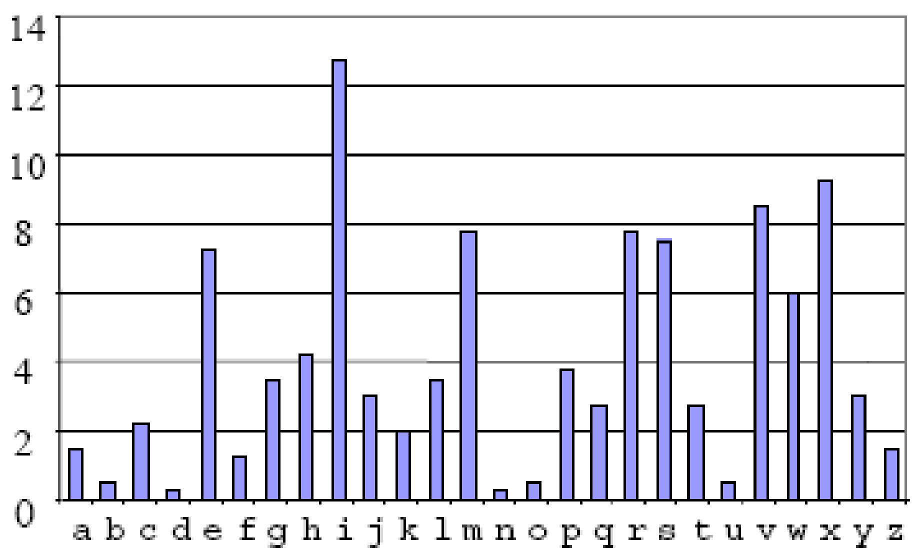
TH, HE, IN, ER, RE, AN, ON, EN, AT

- **Trigrams:**

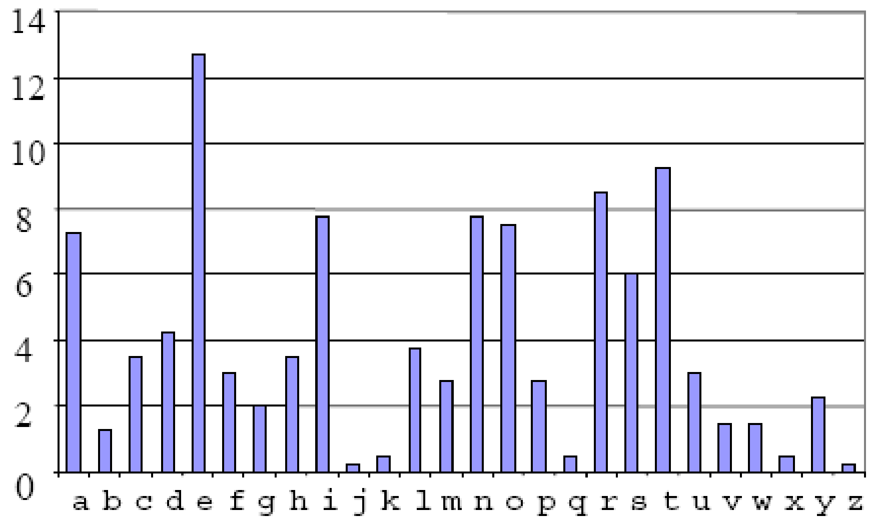
THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR, DTH



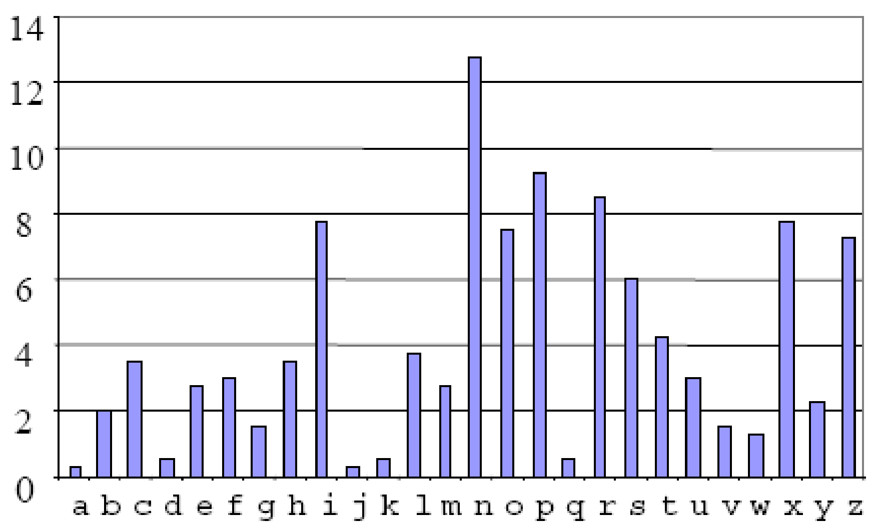
Character frequency in a long English plaintext



Character frequency in the corresponding ciphertext for a shift cipher

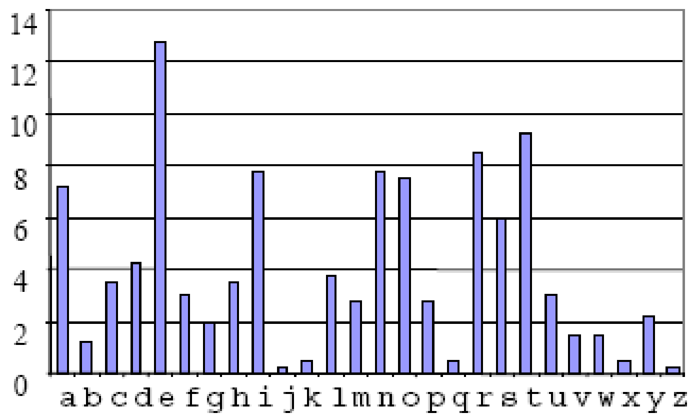


Character frequency in a long English plaintext

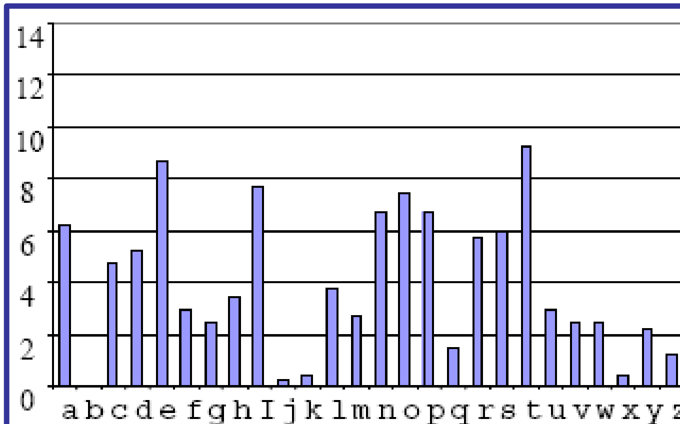


Character frequency in the corresponding ciphertext for a general monoalphabetic substitution cipher

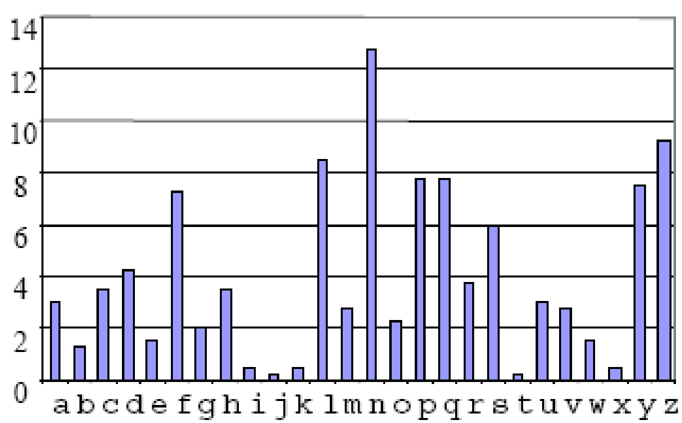
Frequency Analysis Attack: Relevant frequencies



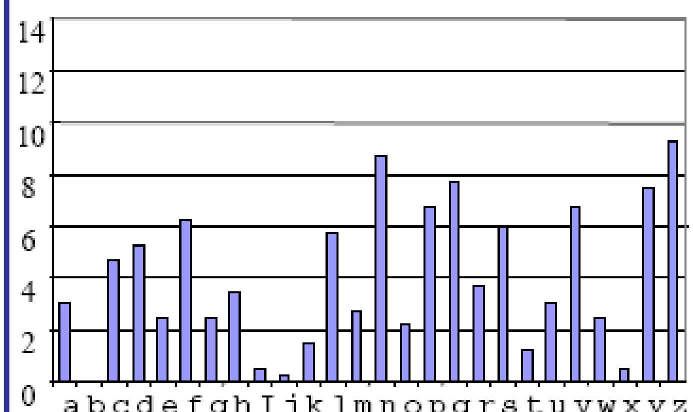
Long English text T



Short English message M



Ciphertext of the long English text T



Ciphertext of the short English message M

Frequency Analysis Example

- **Ciphertext:**

```
FMXVE DKAPH FERBN DKRXR SREFM ORUDS
DKDVS HVUFE DKAPR KDLYE VLRHH RH
```

- **Plaintext =?**

2.2.3 Polyalphabetic Substitution (多表替换) 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
  - Vigenere cipher
  - Rotor machines
  - One-time pad

(1) Vigenère Cipher *1568*

$$c_i = f_{i \bmod d}(m_i) = m_i + k_{i \bmod d} \bmod 26$$

$$m_i = f^{-1}_{i \bmod d}(m_i) = m_i - k_{i \bmod d} \bmod 26$$

$$\text{Key} = k_0, k_1, \ldots, k_{d-1}$$

Number of keys for a given period d = $(26)^d$

Vigenère Cipher: example

- **Plaintext**:ILOVEYOU
- **Key**: MATH
- **Ciphertext**:ULHCQYHB

	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

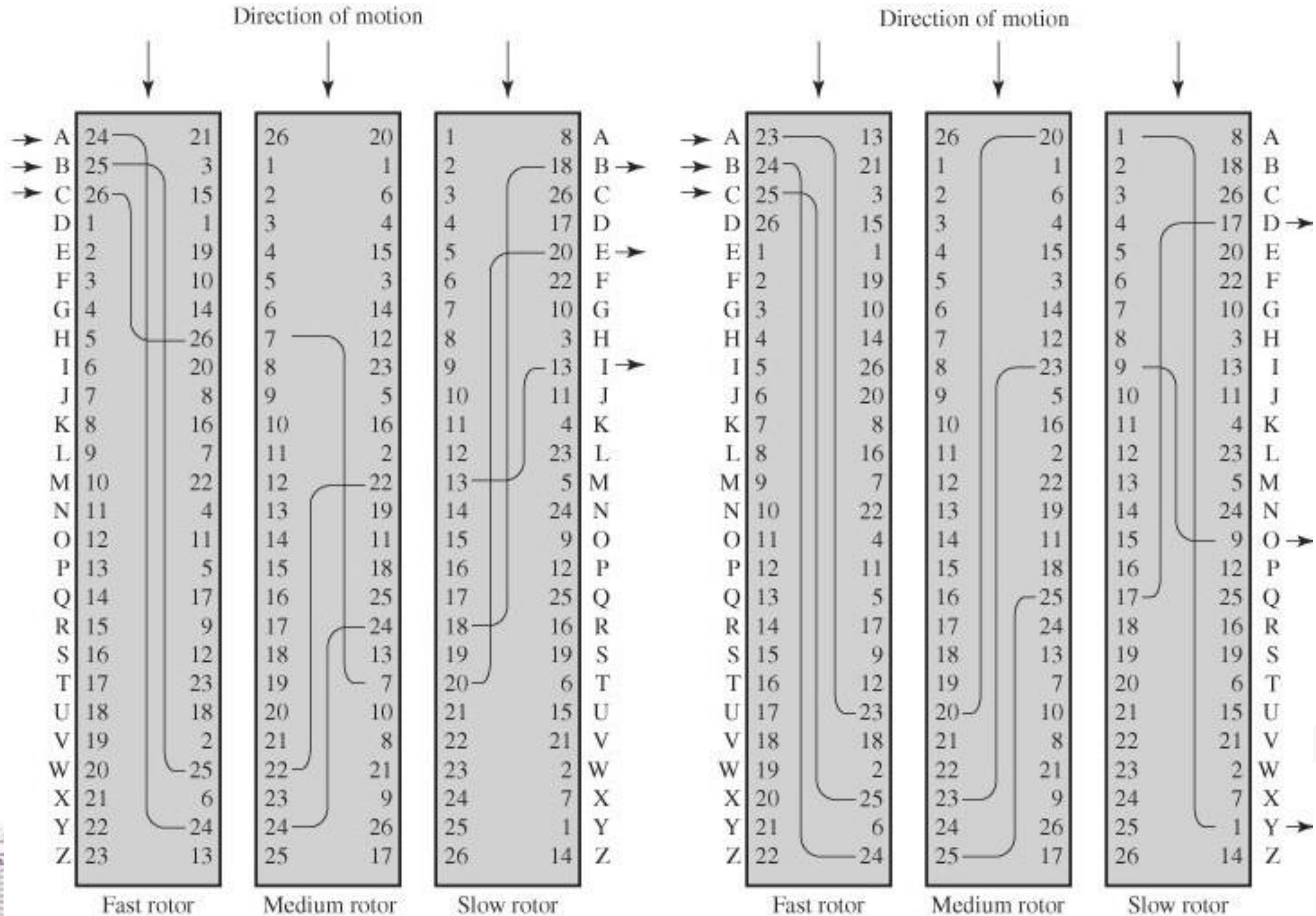
Security of Vigenère Ciphers

- **letter frequencies are obscured**
  - **have multiple ciphertext letters for each plaintext letter**
- **letter frequencies do 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 attack each separately** (Kasiski Method)

(2) Rotor Machines

- **used before and during the WWII**

Country	Machine	Period
Germany:	Enigma	$d=26\cdot25\cdot26 = 16{,}900$
U.S.A.:	M-325, Hagelin M-209	
Japan:	“Purple”	
UK:	Typex	$d=26\cdot(26-k)\cdot26$, $k=5, 7, 9$
Poland:	Lacida	$d=24\cdot31\cdot35 = 26{,}040$



(a) Initial setting

(b) Setting after one keystroke

(按键)

Hagelin Rotor Machine

亚瑟·谢尔比乌斯
(Arthur Scherbius)

马里安·雷杰夫斯基
（Marian Rejewski ）

阿兰·图灵(Alan Turing)
(1912,6-1954,6)

(3) One-Time Pad

Gilbert Vernam, AT&T
Major Joseph Mauborgne

1926

$$c_i = m_i \oplus k_i$$

$$
\begin{array}{ll}
m_i & \texttt{01110110101001010110101} \\
k_i & \texttt{11011101110110101110110} \\
\hline
c_i & \texttt{10101011011111111000011}
\end{array}
$$

All bits of the key must be chosen at random and never reused

$$c_i = m_i + k_i \bmod 26$$

m_i	TO	BE	OR	NOT	TO	BE
k_i	AX	TC	VI	URD	WM	OF
c_i	TL	UG	JZ	HFW	PK	PJ

All letters of the key must be chosen at random and never reused

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(额外开支) to hide relatively few info bits
- Usually hide encrypted message

Experiences and Lessons

- language can be characterized by frequency of letters
- Two basic processing: transposition and substitution
- Security depends on key security, not on keeping cipher secret
- need high key space

Review Questions

3.1 Describe the five main requirements for the secure use of symmetric encryption.
3.2 What are the two basic functions used in encryption algorithms?
3.4 What is the difference between a block cipher and a stream cipher?
3.5 What are the two general approaches to attacking a cipher?
3.6 List and briefly define types of cryptanalytic attacks based on what is known to the attacker.
3.7 What is the difference between an unconditionally secure cipher and a computationally secure cipher?

Review Questions (Cont.)

3.8 Why is the Caesar cipher substitution technique vulnerable to a brute-force cryptanalysis?
3.9 How much key space is available when a monoalphabetic substitution cipher is used to replace plaintext with ciphertext?
3.10 What is the drawback of a Playfair cipher?
3.11 What is the difference between a monoalphabetic cipher and a polyalphabetic cipher?
3.12 What are two problems with the one-time pad?
3.13 What is a transposition cipher?
3.14 What are the drawbacks of Steganography?

Thanks!