Computer Security and Cryptography

CS381

2. Classical ciphers

来学嘉　　计算机科学与工程系　电院3-423室
34205440　1356 4100825 laix@sjtu.edu.cn
2016-03

Organization

- Week 1 to week 16　(2016-02-24 to 2016-06-08)
- 东上院502
- Monday 3-4节; week 9-16
- Wednesday 3-4节; week 1-16
- lecture 10 + exercise 40 + random tests 40 + other 10
- Ask questions **in** class – counted as points
- Turn ON your mobile phone　(after lecture)
- Slides and papers:
  - http://202.120.38.185/CS381
    - **computer-security**
  - http://202.120.38.185/references
- TA: '薛伟佳' icelikejia@qq.com，'黄格仕' <huang.ge.shi@foxmail.com>
- Send homework to: laix@sjtu.edu.cn and to TAs

Rule: do not disturb others!

2

Contents

- Introduction -- What is security?
- Cryptography
  - Classical ciphers
  - Today's ciphers
  - Public-key cryptography
  - Hash functions/MAC
  - Authentication protocols
- Applications
  - Digital certificates
  - Secure email
  - Internet security, e-banking

Network security
- SSL
- IPSEC
- Firewall
- VPN

Computer security
- Access control
- Malware
- DDos
- Intrusion

Examples
- Bitcoin
- Hardware
- Wireless

3

References

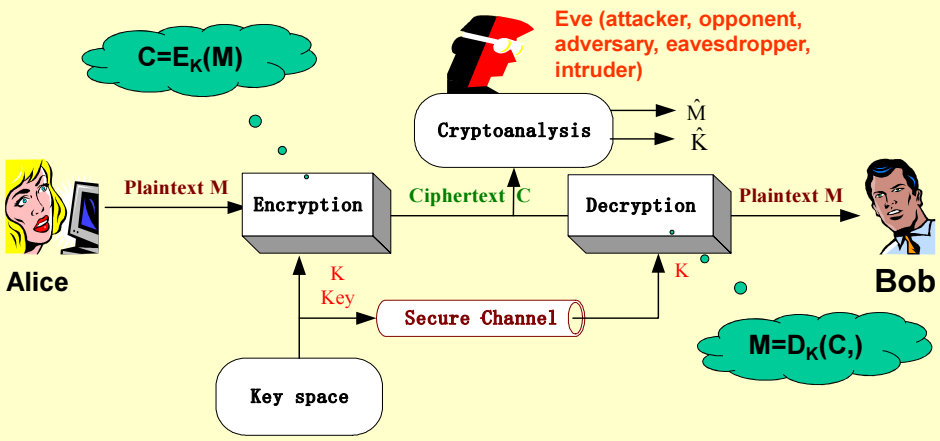
- W. Stallings, *Cryptography and network security - principles and practice*，Prentice Hall.
- W. Stallings, 密码学与网络安全：原理与实践（第4版），刘玉珍等译，电子工业出版社，2006
- Lidong Chen, Guang Gong, *Communication and System Security,* CRC Press, 2012.
- A.J. Menezes, P.C. van Oorschot and S.A. Vanstone, *Handbook of Applied Cryptography*. CRC Press, 1997, ISBN: 0-8493-8523-7, http://www.cacr.math.uwaterloo.ca/hac/index.html
- B. Schneier, *Applied cryptography*. John Wiley & Sons, 1995, 2nd edition.
- 裴定一,徐祥, 信息安全数学基础, ISBN 978-7-115-15662-4, 人民邮电出版社,2007.

4

Symmetric Encryption

- or conventional / private-key / single-key
- sender and recipient share a common key
- all classical encryption algorithms are private-key
- was the only type before invention of public-key in 1970's
- and by far most widely used

Shannon's model of secret communication



Five elements in a cipher system

$\{\mathcal{M}, C, \mathcal{K}, E_K, D_K\}$

- Plaintext (cleartext) M: the message to be sent to the receiver.
  - Plaintext space $\mathcal{M}$: the set of possible values of plaintext.
- Ciphertext C: an encrypted message.
  - Ciphertext space $\mathcal{C}$:the set of possible values of ciphertext.
- Key K: the secret information involves encryption and decryption.
  - Key space $\mathcal{K}$: the set of possible values of key.
- Encryption (encipher): the process of disguising a message in such way as to hide its substance. $C = E_K(M)$
- Decryption (decipher): The process of turning ciphertext back into plaintext.

$M = D_K(C)$

Transposition: rail fence technique

- are—》rea

- plaintext ①②③④⑤⑥⑦⑧⑨⑩……
- ciphertext ①③⑤⑦⑨…②④⑥⑧⑩. . .
- plain： wait me at the gate

  encryption： w i m a t e a e

  a t e t h g t

  cipher： wimateaeatethgt

2016/3/1

8

Transposition, more complex

- Transposition
  - key : 3 4 2 1 5 6 7
    1 2 3 4 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

Put text in matrix form, read columns according to the key

- Decryption
  - Determine the number of letters in a column :length(ciphertext)/length(key)=4
  - Decryption key
    1 2 3 4 5 6 7
    3 4 2 1 5 6 7

T	A	T	A	C	K	P
T	P	S	O	O	N	E
N	T	U	D	I	L	T
A	M	O	W	X	Y	Z

⇒

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

Substitution cipher

- Caesar Cipher
  - 2000 years ago by Julius Caesar
  - Replace each letter by the letter that comes some fixed (3) distance before or after it in the alphabet.
- Math form: $E_3(x)=x-3 \bmod 26$

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
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

Omnia Gallia in tres partes divisa est

LJKF XDXI IFXF KQOB PMXO QBPA FSFP XBPQ

- Brute force attack: try every k in x-k mod 26, for k=0,1,…25

Monoalphabetic Cipher (Substitution)

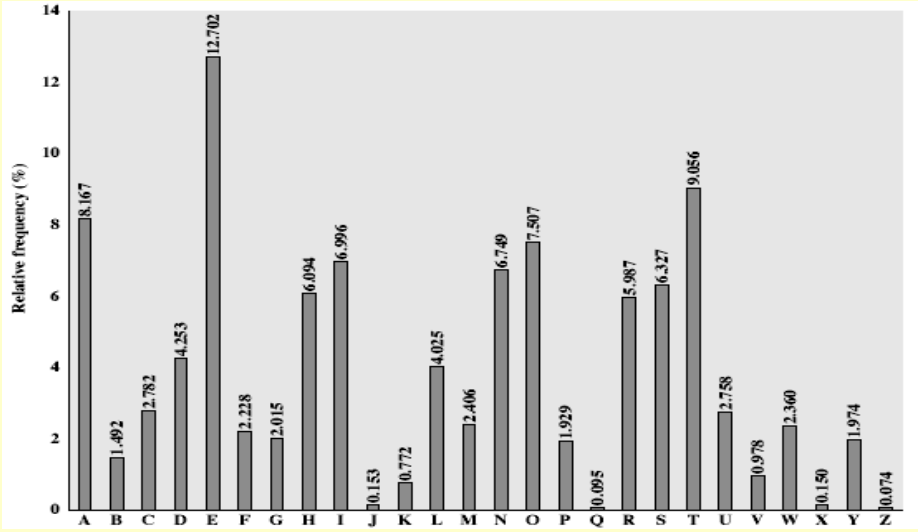
- K is a substitution table
  plain: **ABCDEFGHIJKLMNOPQRSTUVWXYZ**
  cipher：FDLGJKMSWNOPTIRQAYUHXEZBVC
- as long as it is a permutation of the 26 characters.
- Key space is large: $26! \approx 4.03 \times 10^{26}$
- Brute force does not work!

- Attack:
  – Languages have redundancy.
  – Different letters have different frequencies.

2016/3/1

11

English Letter Frequencies

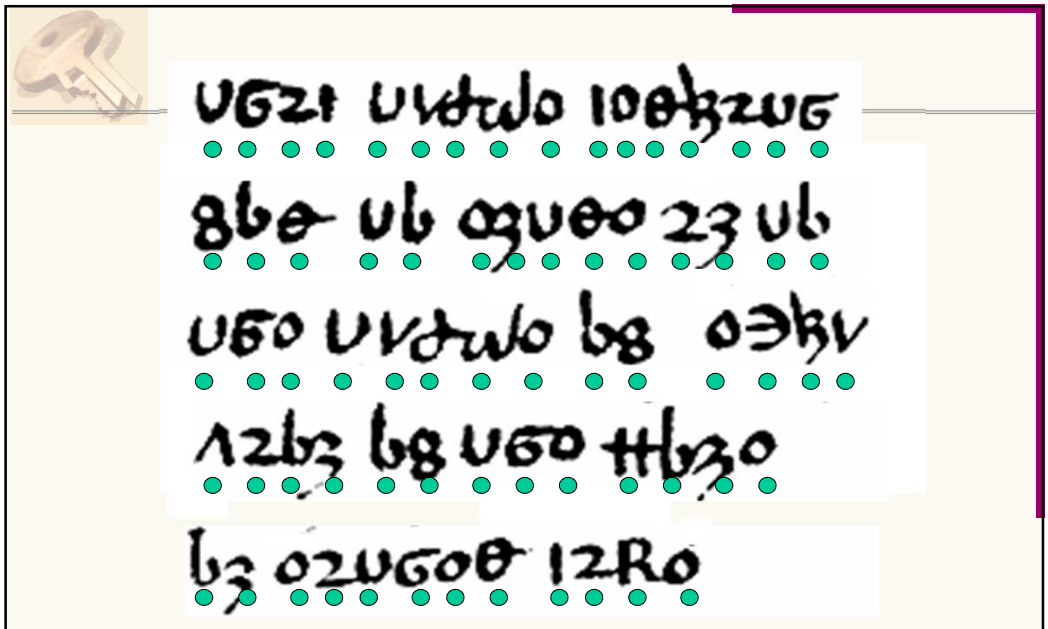


Frequency in English

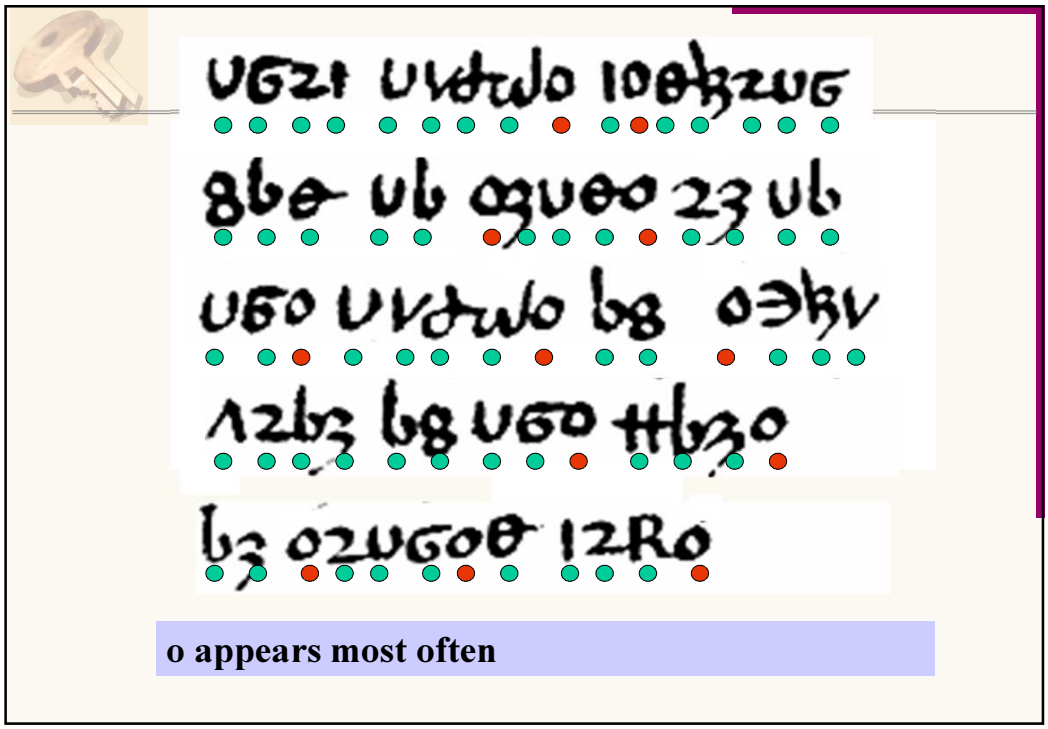
Single Letter	Double Letter	Triple Letter
E	TH	THE
T	HE	AND
R	IN	TIO
N	ER	ATI
I	RE	FOR
O	ON	THA
A	AN	TER
S	EN	RES

SJTU

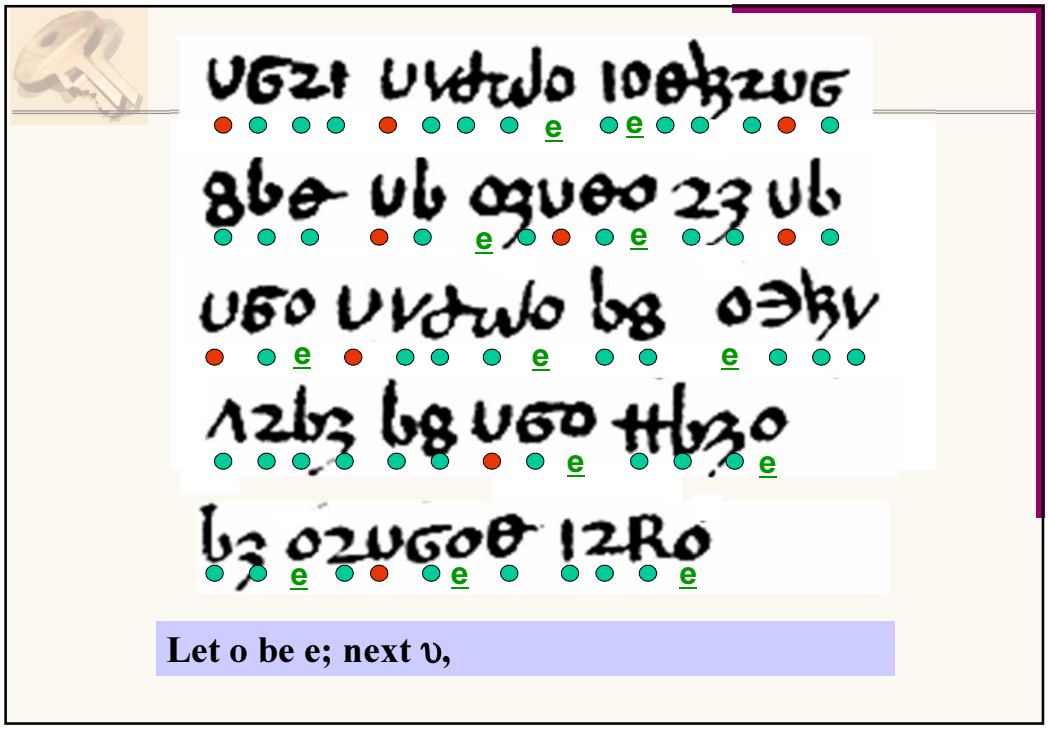
http://cis.sjtu.edu.cn



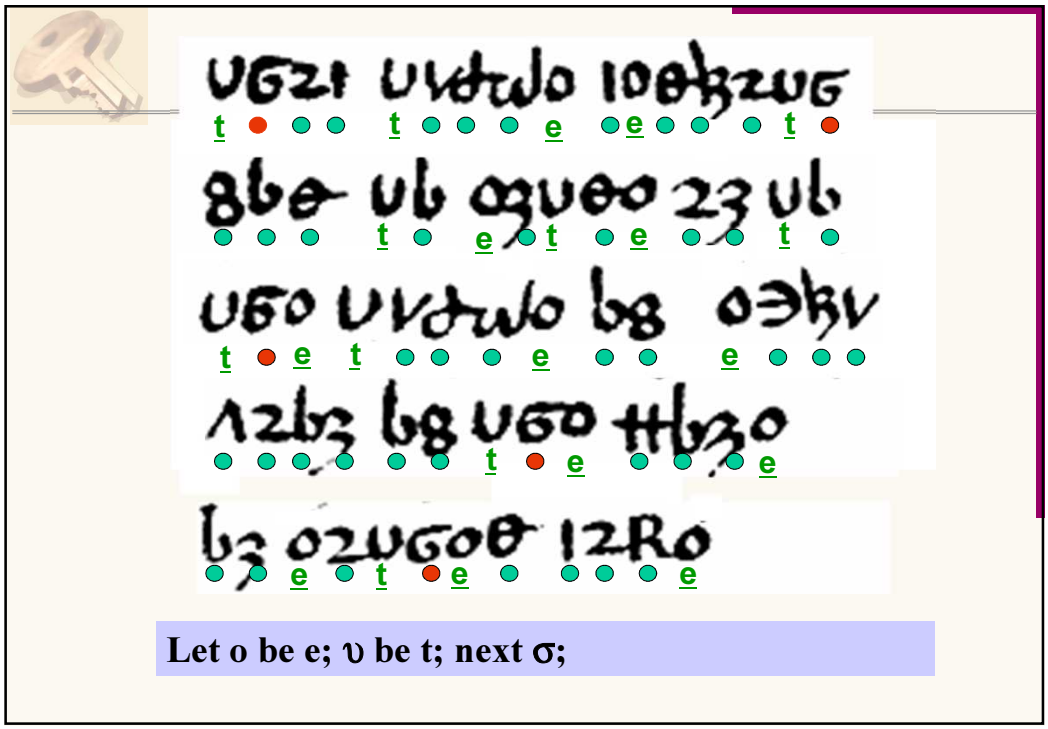
An english text encrypted by Caesar cipher with unknown key



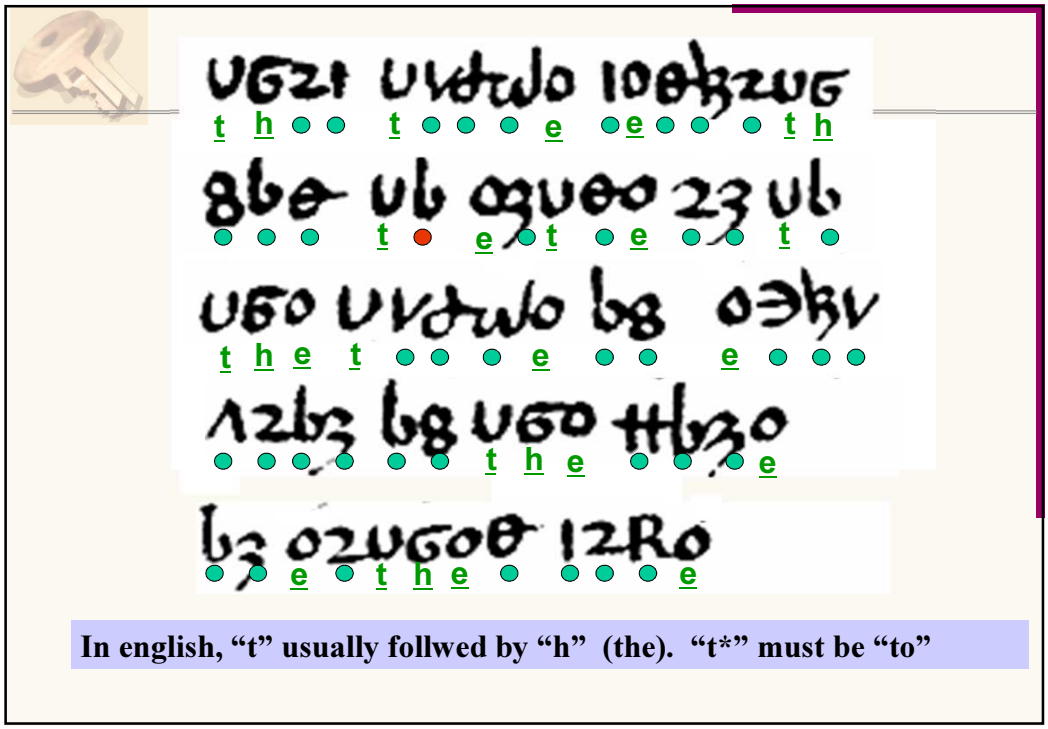
o appears most often



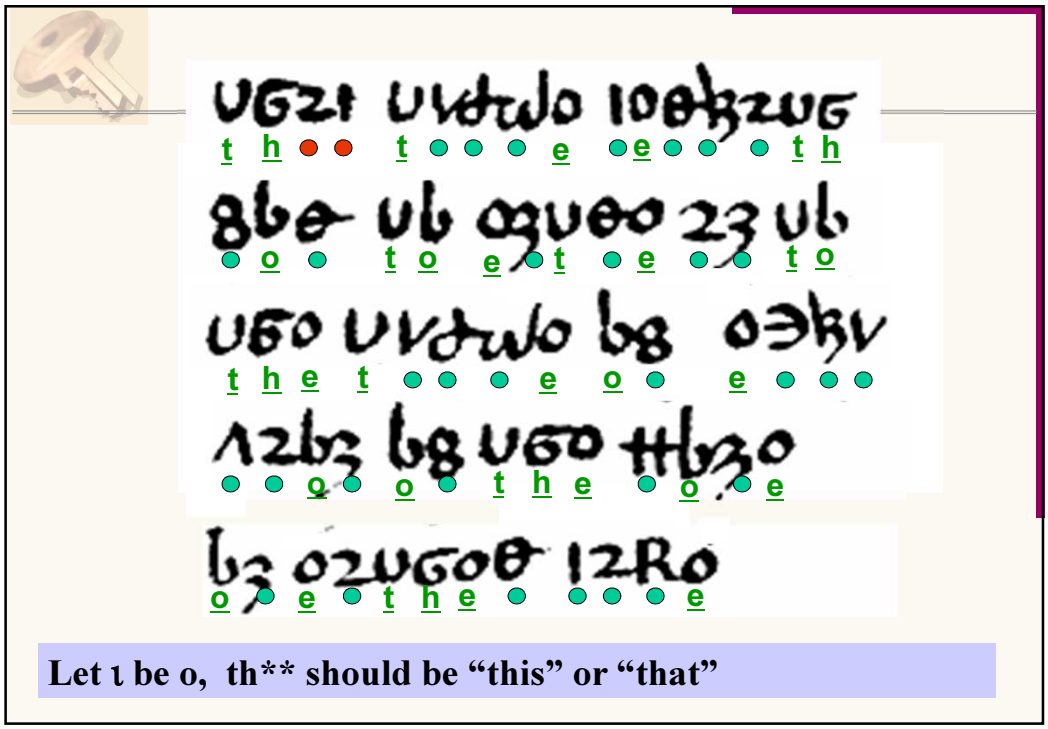
Let o be e; next ʊ,



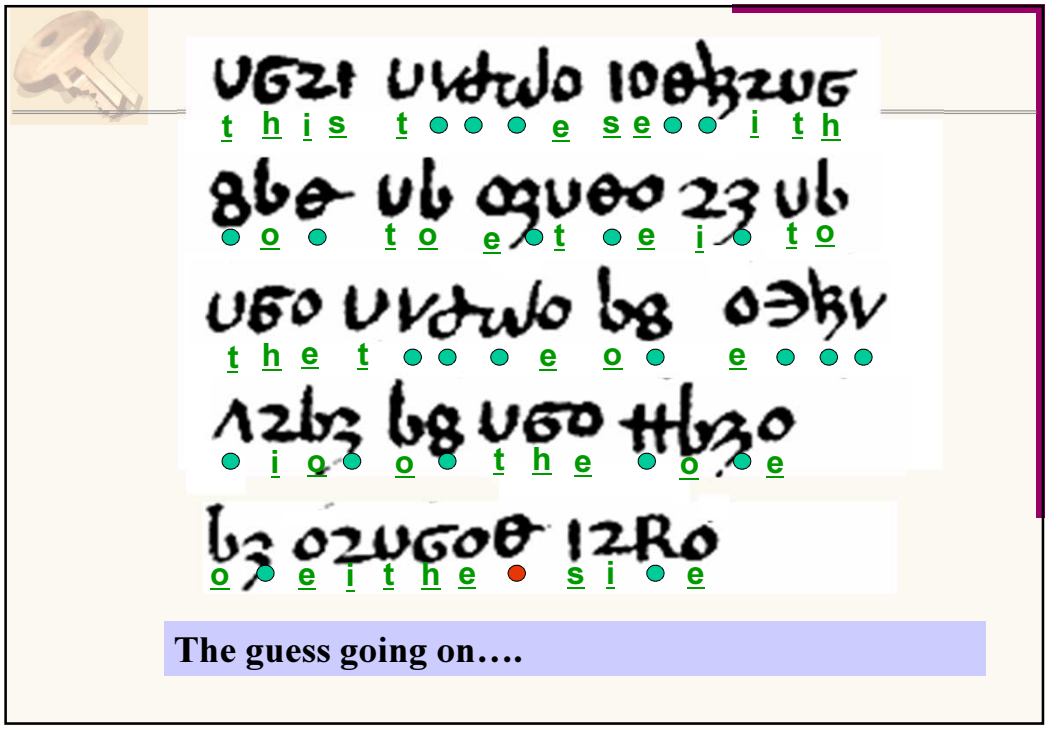
Let o be e; υ be t; next σ;



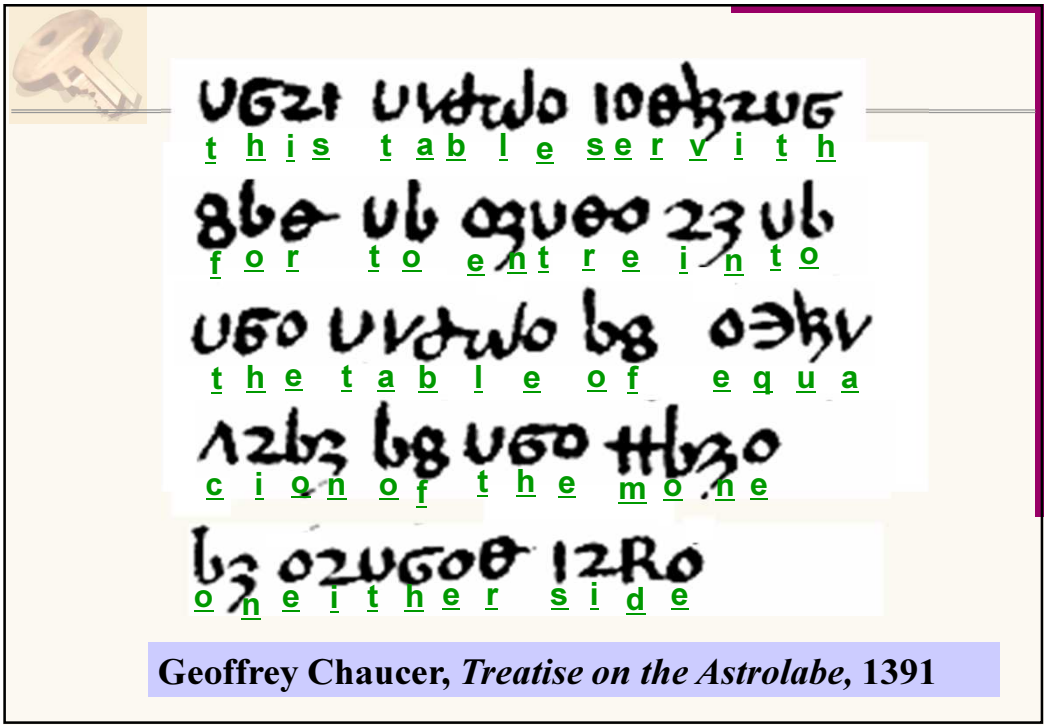
In english, “t” usually follwed by “h” (the). “t*” must be “to”



Let ι be o, th** should be “this” or “that”



The guess going on….



Geoffrey Chaucer, *Treatise on the Astrolabe*, 1391

frequency analysis

- Substitution cipher is easy to break by frequency analysis
- To make it a stronger, one can use multiple substitutions
- For example, the Hill cipher and the Vigenère Cipher.

Hill cipher (1929)

- *m* successive plaintext letters are substituted by *m* successive ciphertext.
- Substitution is determined by *m* linear equations.
- m=3
  - encryption （a=0,b=1,…z=25)
- C = KP

$$C_1 = (k_{11}p_1 + k_{12}p_2 + k_{13}p_3) \bmod 26$$

$$C_2 = (k_{21}p_1 + k_{22}p_2 + k_{23}p_3) \bmod 26$$

$$C_3 = (k_{31}p_1 + k_{32}p_2 + k_{33}p_3) \bmod 26$$

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

Example: Hill cipher

- Plaintext: pay more money
- Encryption key is

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

  - p=15,a=0,y=24
  - $K(15,0,24)^T \bmod 26 = (11,13,18)^T$

Cryptanalysis of Hill Cipher

$(\underline{C_1}\,\underline{C_2}\ldots\underline{C_r}) = K\,(\underline{m_1}\,\underline{m_2}\ldots\underline{m_r})$

As long as matrix $(\underline{m_1}\,\underline{m_2}\ldots\underline{m_r})$ is non-singular, K can be solved.

Weakness of Hill cipher: linearity.

Polyalphabetic cipher

- Key determines several different mono-alphabetic substitution: $\pi_1, \pi_2, \ldots, \pi_r$
- Simple example: Vigenère Cipher
  - Use of several mono-alphabetic substitutions, so one letter can be replaced by different letters.
  - A key letter determines a Caesar cipher;
  - The length of key determines the number of Caesar ciphers.
  - The key letters are used periodically.

Vigenère Cipher

	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
⇨	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
⇨	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
⇨	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
⇨	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
⇨	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
⇨	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
⇨	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
⇨	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

Fight fiercely, Harvard! Fight! Fight! Fight!

XWTNU NZHJQRRZ PRUNOEJ GQXKL TVMIB WLYVG

security

- Vigenère Cipher is a multi-alphabetic substitution
- It is stronger than monoalphabetic substitution, but still can be broken by frequency analysis.
- Period– the length of the keyword (songbird), the same substitution is used periodically.
- With large enough ciphertext, the frequency analysis still works.
- Kasiski (1805 – 1881) Method
  - repetitions in ciphertext give clues to period
  - attack each monoalphabetic cipher individually
- The Hagelin Machine (long period)

28

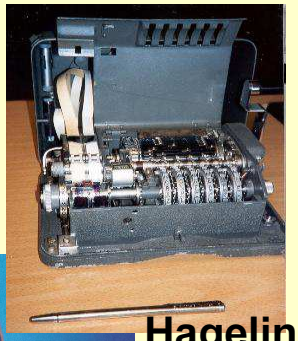
Rotor Machine

widely used in WW2

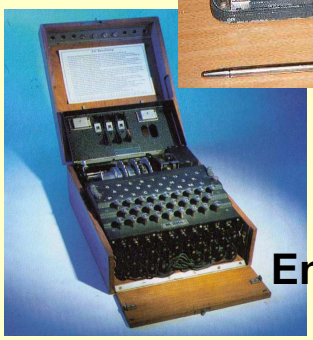
German Enigma, Allied

Hagelin, Japanese Purple

- a series of cylinders,
- each giving one substitution,
- rotate and change after each letter was encrypted
- with 3 cylinders
  - 26^3=17576 alphabets



Hagelin



Enigma

construction

- each rotor giving one substitution

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

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

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

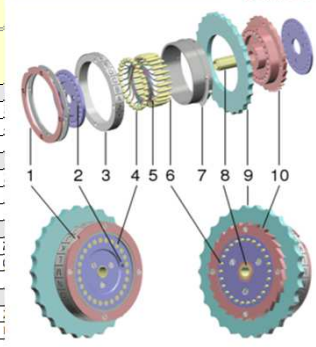
Rotor IV																									
In	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
Out	E	S	O	V	P	Z	J	A	Y	Q	U	I	R	H	X	L	N	F	T	G	K	D	C	M	W

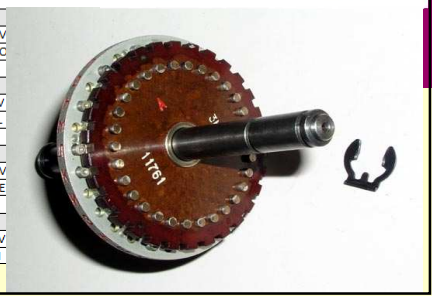
Rotor V																									
In	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V			
Out	V	Z	B	R	G	I	T	Y	U	P	S	D	N	H	L	X	A	W	M	J	Q	O			

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

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

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





3-rotor

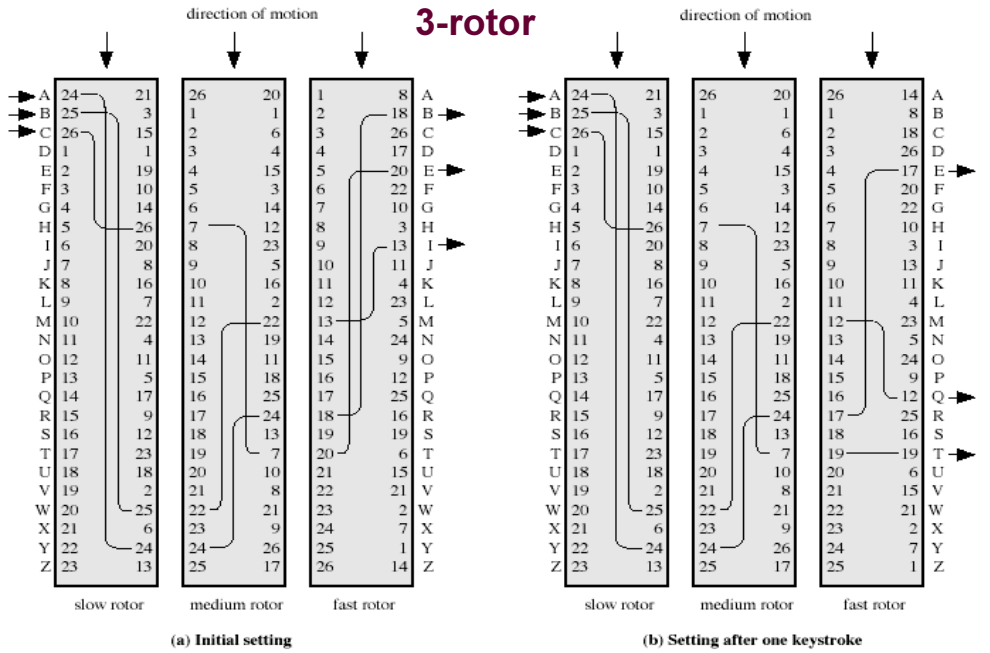
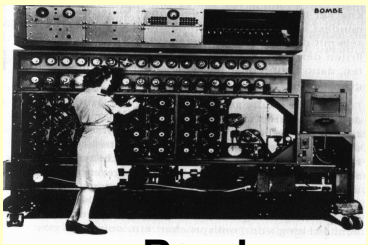


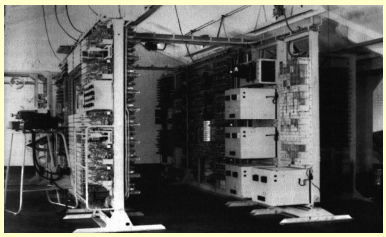
Figure 2.7 Three-Rotor Machine With Wiring Represented by Numbered Contacts

Rotor Machines

- i cylinders have 26^i substitution tables (period).
- However, it is still vulnerable to statistical attacks
- The 1st generation of computer were used to attack it.
- The break of German Enigma cipher played an important role in WWII.



Bombe



Colossus

Bletchley Park

2015*奥斯卡提名*《模仿游戏》
The Imitation Game

33

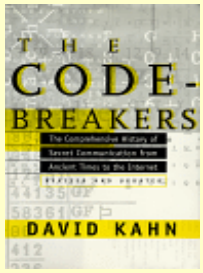
Alan Turing's Office

- Chief-Scientist of >20,000 workers
- Attacking algorithm was found by 4 Polish mathematicians, sold to French -> UK

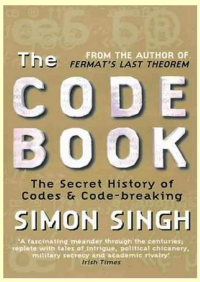
34

Literature

- **History of Cryptography**
  - David Kahn, "The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet", 1181 pages, 1996, Scribner Book Company, ISBN 0-684-83130-9



- **The Code Book**
  - Simon Singh, "The Code Book : The Science of Secrecy from Ancient Egypt to Quantum Cryptography", 402 pages, 2000, Fourth Estate, ISBN 1-857-02889-9



"unbreakable"

if a cipher is provably secure,
then it is probably breakable
--L.Knudson

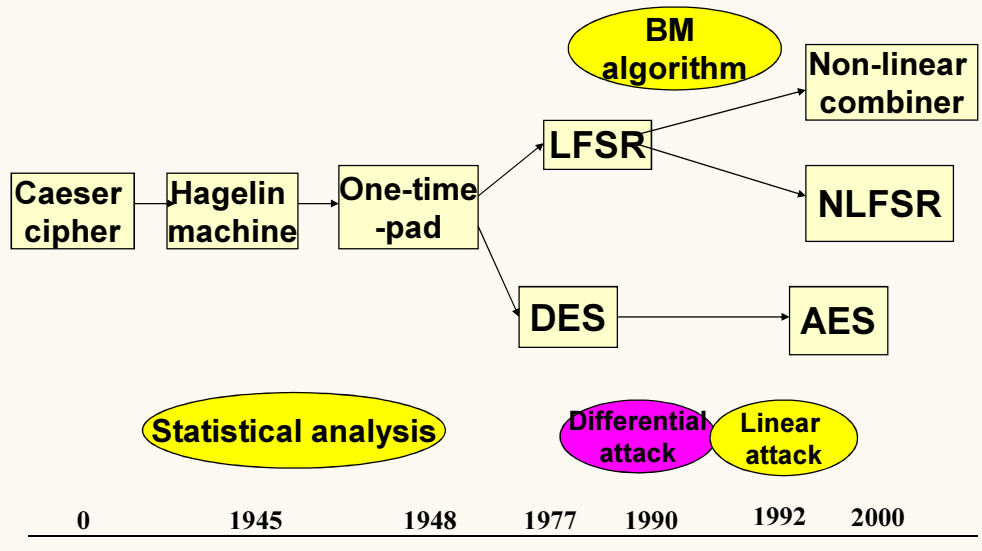
One-Time Pad

- Vernam cipher（Gilbert Vernam，1917）
  - Binary texts，$c_i = p_i \oplus k_i$, i=1,2,3,…
  - Key is statistically independent of plaintext, same length.
- Joseph Mauborgne (1881–1971): one-time-pad
  - Key is random, and used only once
- Theorem (Shannon 1949):

if k (key) is uniformly random, independent of p and used only once, then the ciphertext is statistically independent of plaintext for Vernam cipher (perfect secrecy)

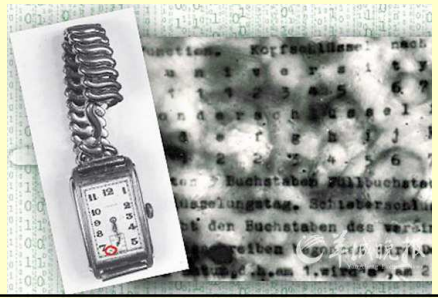
- This is indeed unbreakable, but impractical.

Ciphers and attacks



Steganography

- an alternative to encryption
- Hiding the 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



The red circle contains text

在二战期间，间谍所使用的手表，其中的红色小圈被放大后，显示的是几行德国文字

summary

- model of secret communication
- Five elements in a cipher system
- classical cipher
  - Transposition
  - Substitution
  - *Casear* cipher
  - *Vigenère* cipher
  - Rotor machines
  - One time pad

Next part: unconditional security

http://cis.sjtu.edu.cn

Exercise 2

1. Which security services are required to do secure online shopping?
2. Decrypt the ciphertext of TZUYH of Vigenère Cipher with keyword OR.
3. Is one-time-pad cipher practical, and why?

Deadline: 1 day before next lecture\

Subject: CS381-EX#-name

41