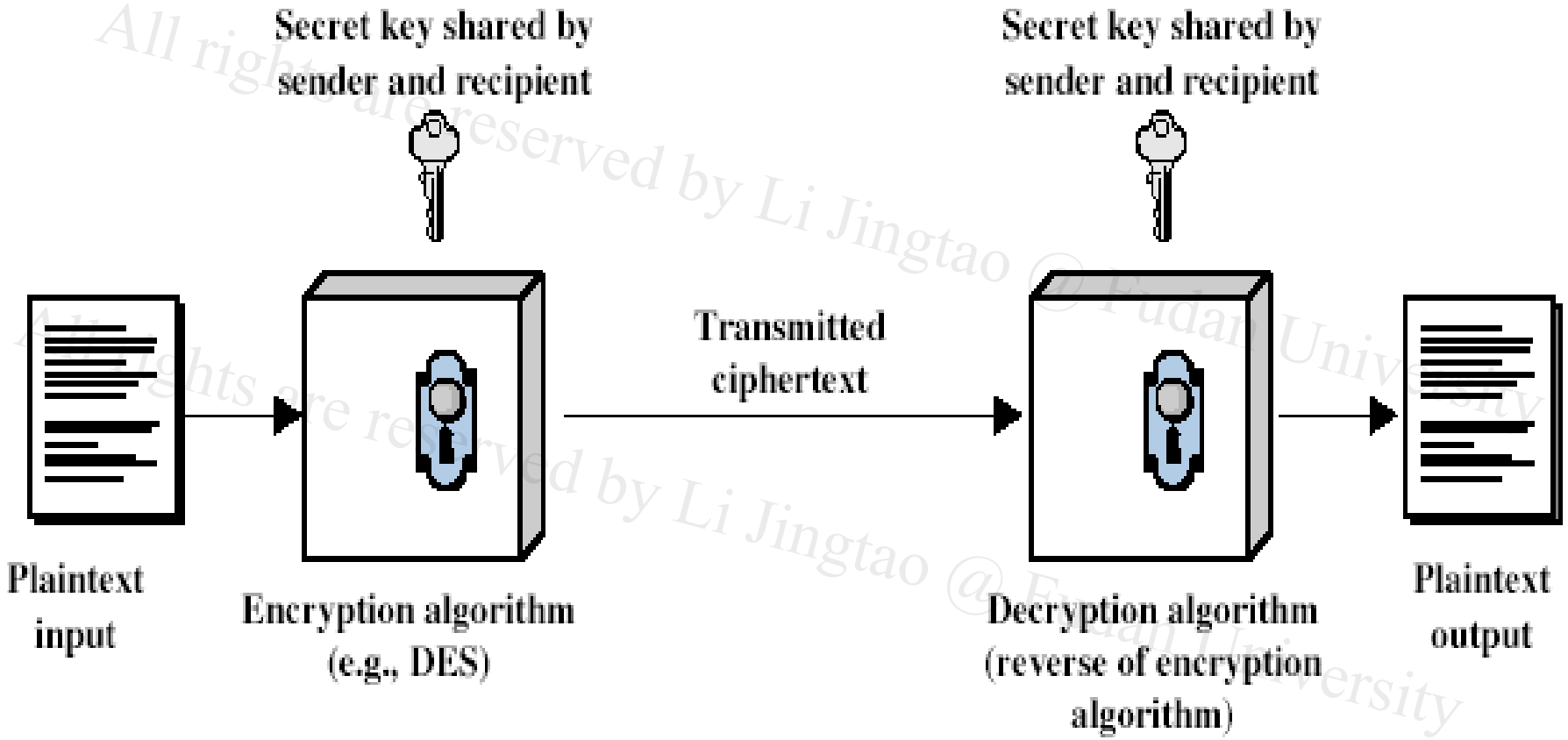
信息安全（02）

Introduction to Cryptography

-Classical Encryption Techniques (cont.)

Symmetric Cipher Model



Cryptanalytic Attacks

- 对于对手而言
  - 最坏情况下，仍有一种攻击方法可用

- **Brute Force Search**，穷举法

Brute Force Search

- always possible to simply try every key
- most basic attack, proportional to key size
- assume either know or recognise plaintext

Key Size (bits)	Number of Alternative Keys	Time required at 1 encryption/μs	Time required at 10^6 encryptions/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

More Definitions

- **unconditional security**
  - no matter how much computer power 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
- Unconditional security would be nice, but the only known such cipher is the **one-time pad** (later).
  - For all reasonable encryption algorithms, have to assume computational security where it either takes too long, or is too expensive, to bother breaking the cipher.

Types of Cryptanalytic Attacks

- **ciphertext only**
  - Encryption algorithm
  - Ciphertext to be decoded
- **known plaintext**
  - Encryption algorithm
  - Ciphertext to be decoded
  - One or more plaintext-ciphertext pairs formed with the secret key
- **chosen plaintext**
  - Encryption algorithm
  - Ciphertext to be decoded
  - Plaintext message chosen by cryptanalyst, together with its corresponding ciphertext generated with the secret key

Types of Cryptanalytic Attacks

- **chosen ciphertext**
  - Encryption algorithm
  - Ciphertext to be decoded
  - Purported ciphertext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key
- **chosen text**
  - Encryption algorithm
  - Ciphertext to be decoded
  - Plaintext message chosen by cryptanalyst, together with its corresponding Ciphertext with the secret key
  - Purported cipehrtext chosen by cryptanalyst, together with its corresponding decrypted plaintext generated with the secret key.

Monoalphabetic Cipher Security

- now have a total of 26! = 4 x 10^{26} keys
- with so many keys, might think is secure
- but would be **!!!WRONG!!!**
- problem is language **characteristics**

An Improvement

- Homophone
- Assign each letter a number of different cipher symbols
- The number of symbols assigned to each letter is proportional to the relative frequency of that letter

提高单字母表密码安全性

- 两个角度
  - “多”对“一” → Playfair
  - “一”对“多” → **Vigenère**

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 密钥矩阵

- 5×5变换矩阵: I与J视为同一字符
- 基于关键字构造一个密钥矩阵
- eg. 使用关键字：cipher

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

加密方法

- Playfair:将明文中的双字母组合作为一个单元对待，并将这些单元转换为密文的双字母组合

- 加密规则:按双字母组合加密
  - 1) 相同的字母加分隔符(如x)
    balloon　　ba lx lo on
  - 2) 同行取右边: he -> EC
  - 3) 同列取下边: dm -> MT
  - 4) 其他取交叉: kt -> MQ, OD -> TR

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

Playfair举例

- 以前面的5×5变换矩阵(cipher)为例

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

- (1)balloon -> ba lx lo on
  -> db sp gs ug
- (2)book -> bo ok -> sr qg
- (3)fill -> fi lx lx -> ae sp sp

Playfair

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

- E, D, p, C, K?

Security of the Playfair Cipher

- security much improved over monoalphabetic
- since have 26 x 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. US & British military in WW1)
- it **can** be broken, given **a few hundred** letters
- since still has much of plaintext structure

Polyalphabetic Ciphers

- another approach to improving security is to use multiple cipher alphabets
- called polyalphabetic substitution ciphers
- makes 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 is the **Vigenère Cipher**
- effectively multiple caesar ciphers
- key is multiple letters long K = k1 k2 ... kd
- 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

Vigenère Cipher

Key \ Plaintext	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

- 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:
```

Example

- 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:
```

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

Security of Vigenère Ciphers

- have multiple ciphertext letters for each plaintext letter
- hence letter **frequencies** are **obscured**
- but **not totally lost**
- The ultimate defence against such a cryptanalysis is to choose a keyword that is as long as the plaintext and has no statistical relationship to it
  - AT&T, Vernam cipher

G. Vernam 1918

$$Ci = Pi \oplus Ki$$

$$Pi = Ci \oplus Ki$$

- Shannon在他的经典论文([Shannon 49]和[Shannon 51])中已经证明了一次一密所提供的绝对安全性

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
- have problem of safe distribution of key

One-Time Pad example

Plain-text: heilhitler
Key: wclnbtdefj
Cipher-text: DGTYIBWPJA

Message from spy

Cipher-text: DGTYIBWPJA
Key: wggsbtdefj
Plain-text: hanghiter

Lie of spy

Cipher text: DCYTIBWPJA
Key: wclnbtdefj
Plain-text: hanghitler

Cheat Spy

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 scheme
- 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

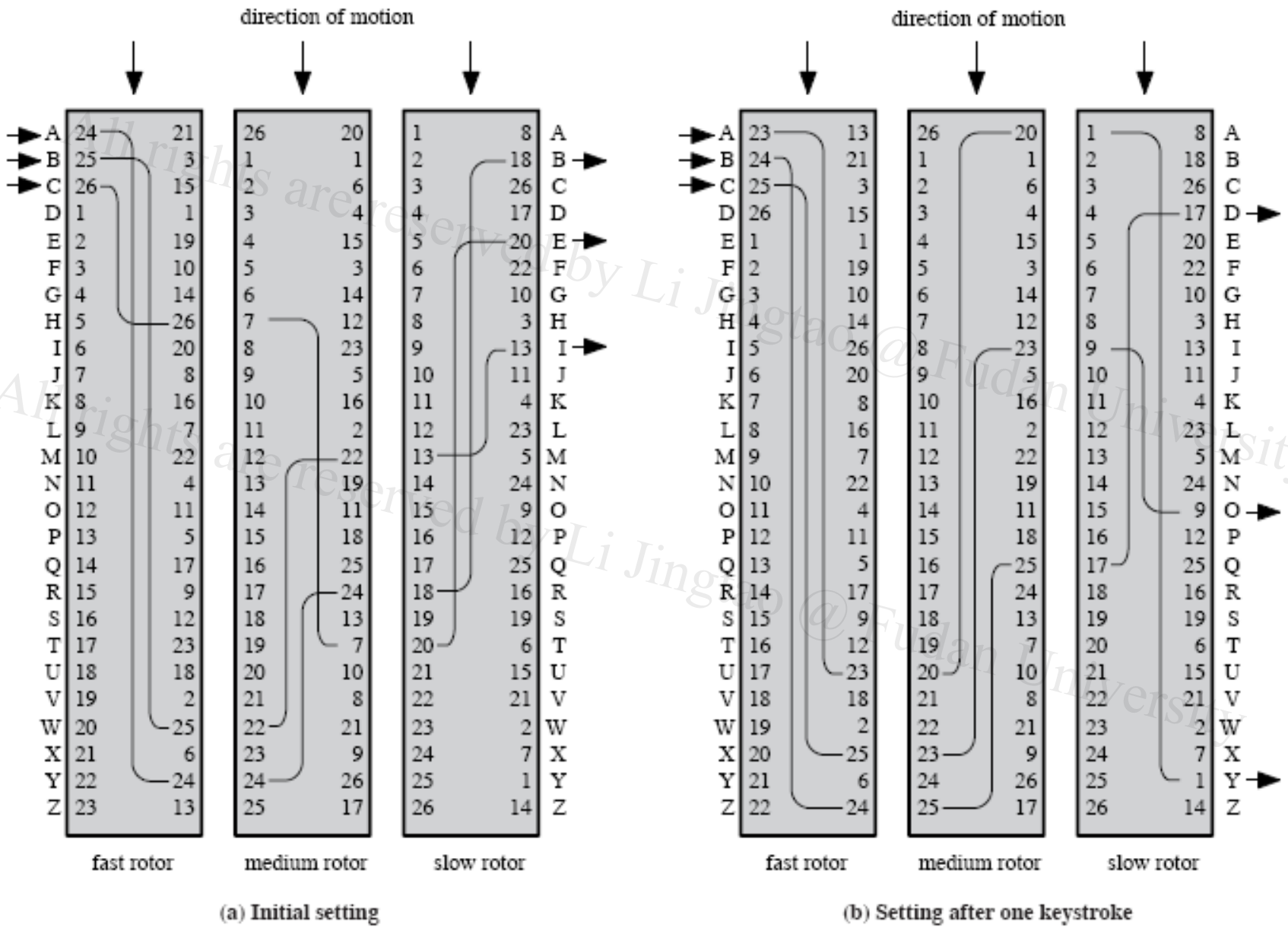
- ciphers using substitutions or transpositions are not secure because of language characteristics
- hence consider using several ciphers in succession to make harder, but:
  - two substitutions make a more complex substitution
  - two transpositions make more complex transposition
  - but a substitution followed by a transposition makes a new much harder cipher
- this is bridge from classical to modern ciphers

Rotor Machines

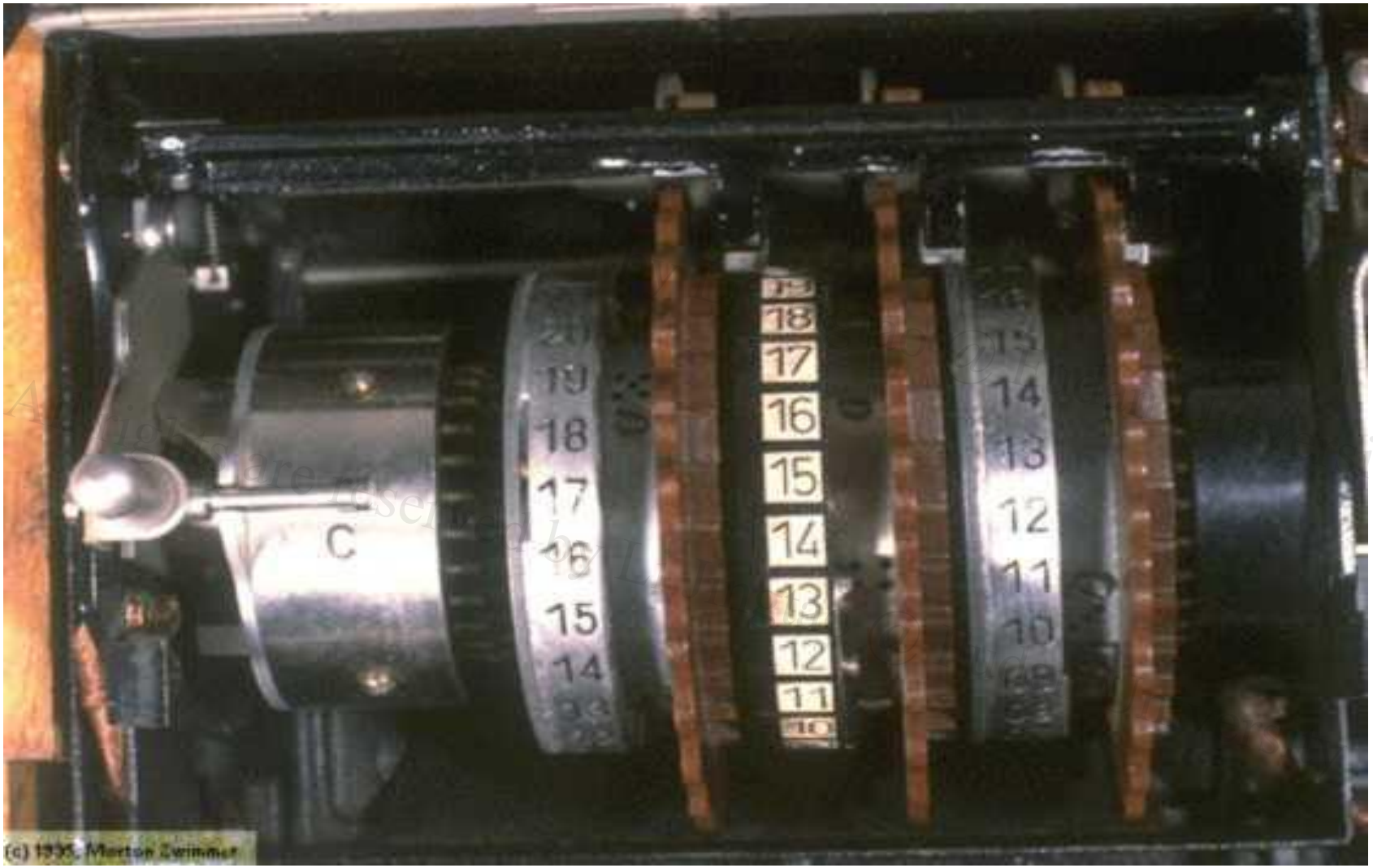
- before modern ciphers, rotor machines were most common product cipher
- were widely used in WW2
  - German Enigma, Allied Hagelin, Japanese Purple
- implemented a very complex, varying substitution **product** cipher
- used a series of cylinders, each giving one substitution, which rotated and changed after each letter was encrypted
- with 3 cylinders have 26^3=17576 alphabets

Enigma

Enigma



Enigma-Rotors



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 noise in graphic image or sound file
- has drawbacks
  - high overhead to hide relatively few info bits

现代隐写术的变迁

- 一方面，网上传输的大量多媒体信息，如图像、声音、视频，甚至文本信息，对于人类的视觉、听觉感知系统，都或多或少存在着一些冗余空间，而利用这些冗余空间，就可以进行信息的秘密传递，同时不影响载体的视觉或听觉效果，因此就可以实现信息的隐蔽传递。

现代隐写术的变迁

- 另一方面，数字产品的无失真复制，越来越多的数字产品在网上传播，如电影、音乐等；
- 造成了版权保护方面存在很大的漏洞。
- 如何既能够充分利用互联网的优势，实现信息共享，同时又不损害数字产品所有者的利益，因此就出现了数字水印的技术。

信息隐藏技术

- 伪装式保密通信
- 数字水印

伪装式保密通信

- 目前在这一研究领域中主要研究在图像、视频、声音以及文本中隐藏信息。如：
- 在一幅普通图像中隐藏一幅机密图像。
- 在一段普通谈话中隐藏一段机密谈话或各种数据。
- 在一段视频流中隐藏各种信息等。
- 文本中的冗余空间比较小，但利用文本的一些特点也可以隐藏一些信息。

数字水印

- 目前存在两种基本的数字版权标记手段，数字水印和数字指纹。

- 数字水印是嵌入在数字作品中的一个版权信息，它可以给出作品的作者、所有者、发行者以及授权使用者等等版权信息。
- 数字指纹可以作为数字作品的序列码，用于跟踪盗版者。

Summary

- **Cryptography** is a good tool to ensure the confidentiality of sensitive message
- **Cryptography** has two basic command: encrypt/encipher, decrypt/decipher
- Classical Cryptography include:
  - Julius Caesar
  - Playfair
  - Vigenère
  - Transposition Ciphers
  - One-time Padding