密码学简介

清华大学计算机系
于红波
2024年2月28日

什么是密码学？密码用在哪些地方？

什么是密码学?

- 经典定义：密码学是研究保密通信的一门科学。研究在不安全的环境中，如何把所要传输的信息在发给接收者之前进行秘密转换以防止第三者对信息的窃取

  保证信息的机密性
- 现代定义：密码学主要研究如何构建能够经受住任何滥用的安全方案，即：在任何恶意企图使它们偏离规定的情况下，该方案仍能维护它所设计的功能（From Foundations of Cryptography – O. Goldreich）

  保证信息的机密性和可认证性

提纲

- 密码学发展史：古典密码到现代密码
- 主要密码技术及密码常识

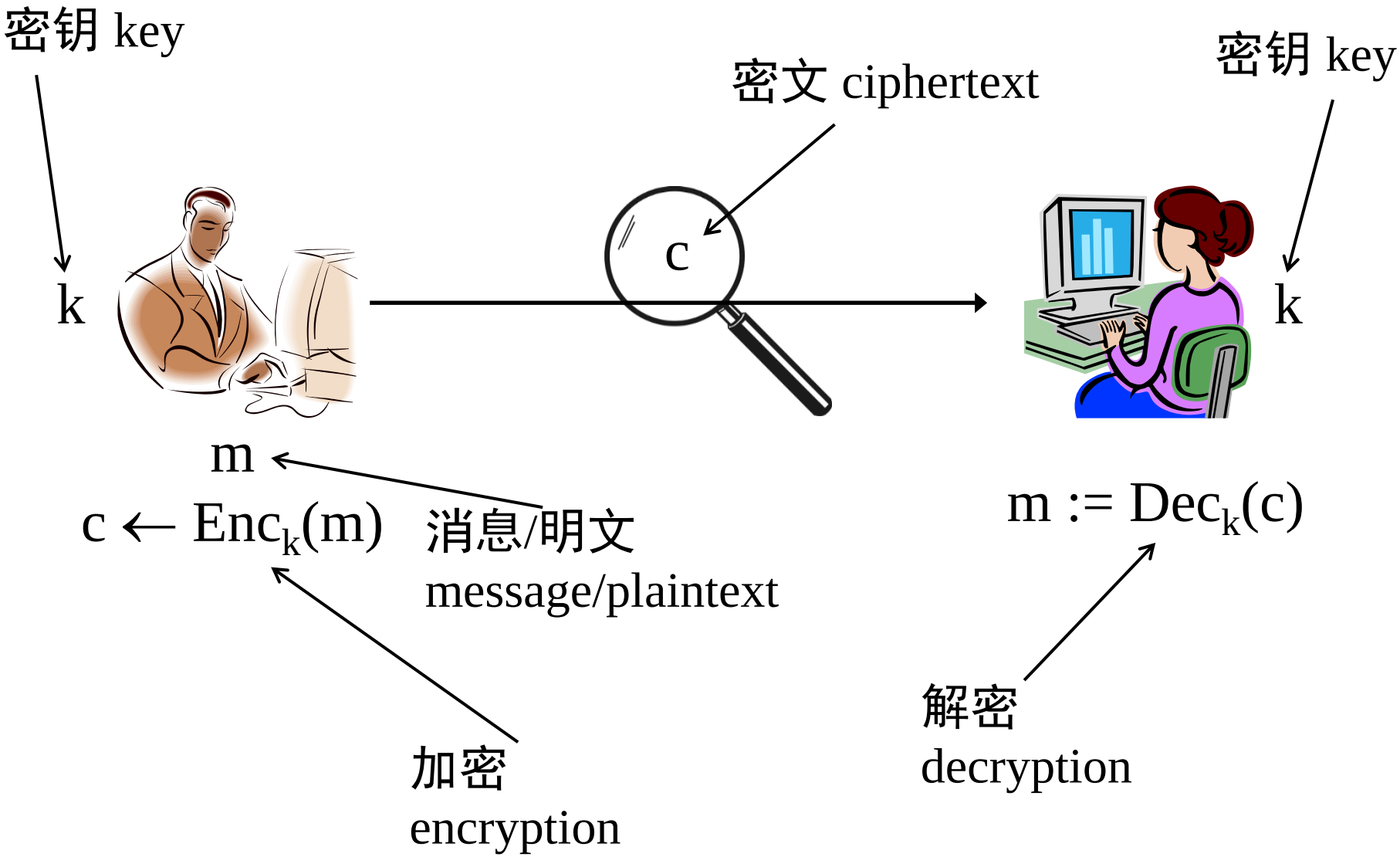
密码学发展史 —— 古典密码

- 动机
  - **逐步深入**：让我们能够“轻松地进入事物……”，引入符号
  - **揭示隐藏的挑战**：说明为什么事情比看起来更困难
  - **对严谨性的启示**：激发更严谨、更系统的科学方法

密码学发展史 —— 古典密码

- 直到 20 世纪 70 年代，只关注确保通信的**机密性**
  - 即，**加密**
- 直到 20 世纪 70 年代，它完全依赖于通信双方之间**预先共享**的秘密信息（*密钥*）
  - **私钥密码学**（Private-key cryptography），又名秘密密钥/共享密钥/**对称密钥密码学**（secret-key / shared-key / symmetric-key cryptography）

私钥加密（Private-key encryption）



私钥加密（Private-key encryption）

k

m

$c \leftarrow Enc_k(m)$

c

c

c

k

$m := Dec_k(c)$

Kerckhoffs's 原则

- 19世纪后期提出
  - 密码算法本身不需要保密，只需通信双方秘密密钥保密
  - 安全性只依赖密钥的安全性
- 优势
  - 维护密钥的安全性比维护密码算法的安全性代价小
  - 即使密钥泄露也不影响密码算法的安全性
  - 多人通信时，只需不同人维护相应的密钥
  - 能够公开对密码算法进行分析，利于发现问题，便于推广

密码学发展史-隐写术

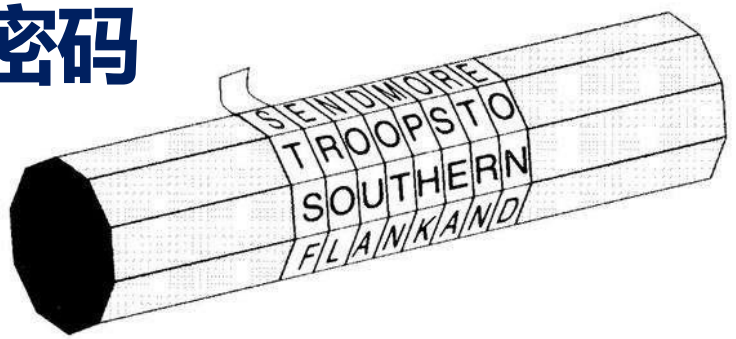
- 与计算设备有关
  - 笔、纸
    - Staganography（隐写术）：
      - Steganos（覆盖）
      - Graphein(写): 隐形墨水, 数字水印

芦花滩上有扁舟，
俊杰黄昏独自游。
义到尽头原是命，
反躬逃难必无忧。

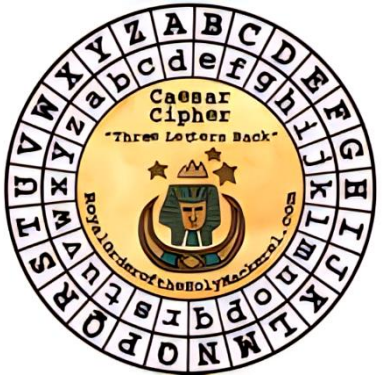
武则天时代　宰相裴炎给徐敬业、骆宾王传递信息“青鹅”

这是一棵树的照片，内含了隐蔽的图像。如果把每个色彩空间和数字3进行逻辑与运算，再把亮度增强85倍，得到下图。（来源于为维基百科）

密码学发展史 —— 古典密码



- 与计算设备有关 —— 笔、纸
  - Cryptography（密码术）：
  - 斯巴达密码：早在公元前5世纪，古希腊人发明了斯巴达密码，斯巴达人将一串信息附在在特定的木棍上，把木棍抽走，信息就会变成乱码。
  - 凯撒密码： 简单的代换密码(也称替换密码)
  - 反切码：我国抗倭明将戚继光发明



柳边求气低，波他争日时。

莺蒙语出喜，打掌与君知。

春花香，秋山开，嘉宾欢歌须金杯，孤灯光辉烧银缸。

之东郊，过西桥，鸡声催初天，奇梅歪遮沟。

前一首诗歌的20个字作为声母，依次编上号码1-20；

后一首诗歌的36个字作为韵母，依次编上号码1-36；

当时字音的八种声调，也依次编上号码1-8

5-25-2 敌

密码学的发展史 —— 古典密码

- 与计算设备有关 —— 笔、纸
  - Cryptography（密码术）
    - **Transposition（易位）**

      cow: COW, OCW, CWO, OWC, WCO, WOC

      For example, consider this short sentence.

      $\approx 10^{31.76} \approx 2^{105.51}$

    - **Substitution（替换、代换）**

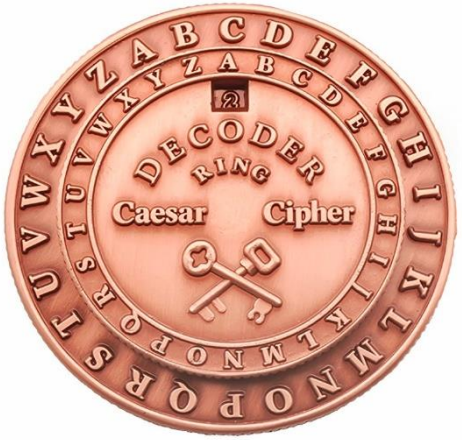
明码表	a	d	h	i	k	m	o	r	s	u	w	y	z
密码表	V	X	B	G	J	C	Q	L	N	E	F	P	T

      meet at midnight

      CUUZ VZ CGXSGIBZ

密码隐藏的是内容，隐写术隐藏的是消息本身。

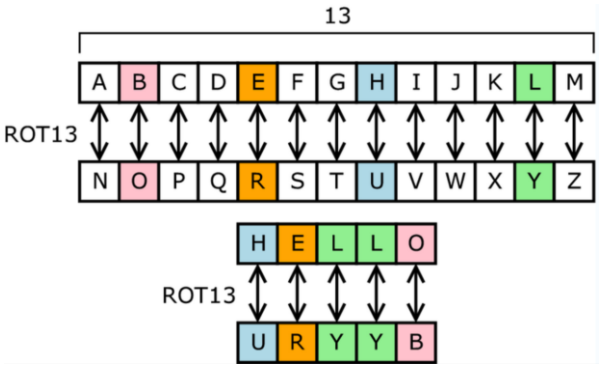
古典密码学 —— 移位密码

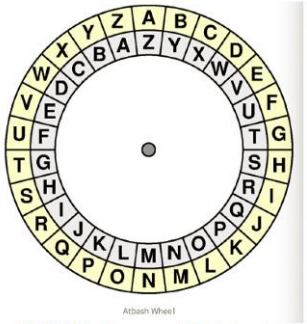


- 移位密码 (凯撒密码)
  - 比如，a由C表示，b由D表示，……，y由A表示，z由B表示

```
hello world
    ↓
JGNNQ YQTNF
```

- hello world 一共有多少种不同形式的密文？为什么？





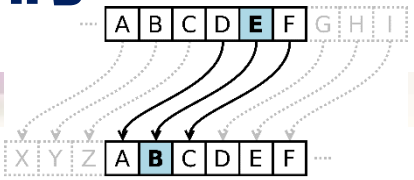
古典密码学 —— 移位密码

- 移位密码的数学描述
  - 考虑加密的是英文文本，去掉除字母之外的字符
  - 假定0表示字母‘a’，1表示‘b’，……，25表示‘z’
  - **密钥**的集合为 $k \in \mathcal{K} = \{0, 1, 2, \ldots\ldots, 25\}$

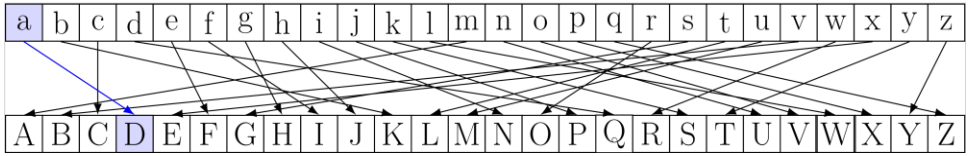
古典密码学 —— 移位密码

- 移位密码安全吗？只有 26 种可能的密钥！
- 移位密码的暴力搜索
  - 密文c:　　URYYBJBEYQ
    移位1　vszzckcfzr
    移位2　wtaadldgas
    ……
    axeehphkew
    ……
    移位13 helloworld
    ……
  - 获得一个有意义的明文，同时可以验证不存在其他有意义明文
- 移位密码的启示：*sufficient key-space principle*
  - 为了防止暴力搜索，密钥空间必须足够大

古典密码学 —— 单字母替换密码



- 单字母替换密码
  - 不仅仅是移位变换



  - 每个字母可以用其它任何一个字母替换(不能重复)
  - 每个字母可以随机的映射到其它字母

```
原字母：abcdefghijklmnopqrstuvwxyz
替换后：DKVQFIBJWPESCXHTMYAUOLRGZN
m: if   we   wish   to   replace   letters
c: WI   RF   RWAJ   UH   YFTSDVF   SFUUFYA
```

  - 尝试给出一种替换方法?
  - 假定被加密的信息为英文文本，该方法是否安全?

古典密码学 —— 单字母替换密码

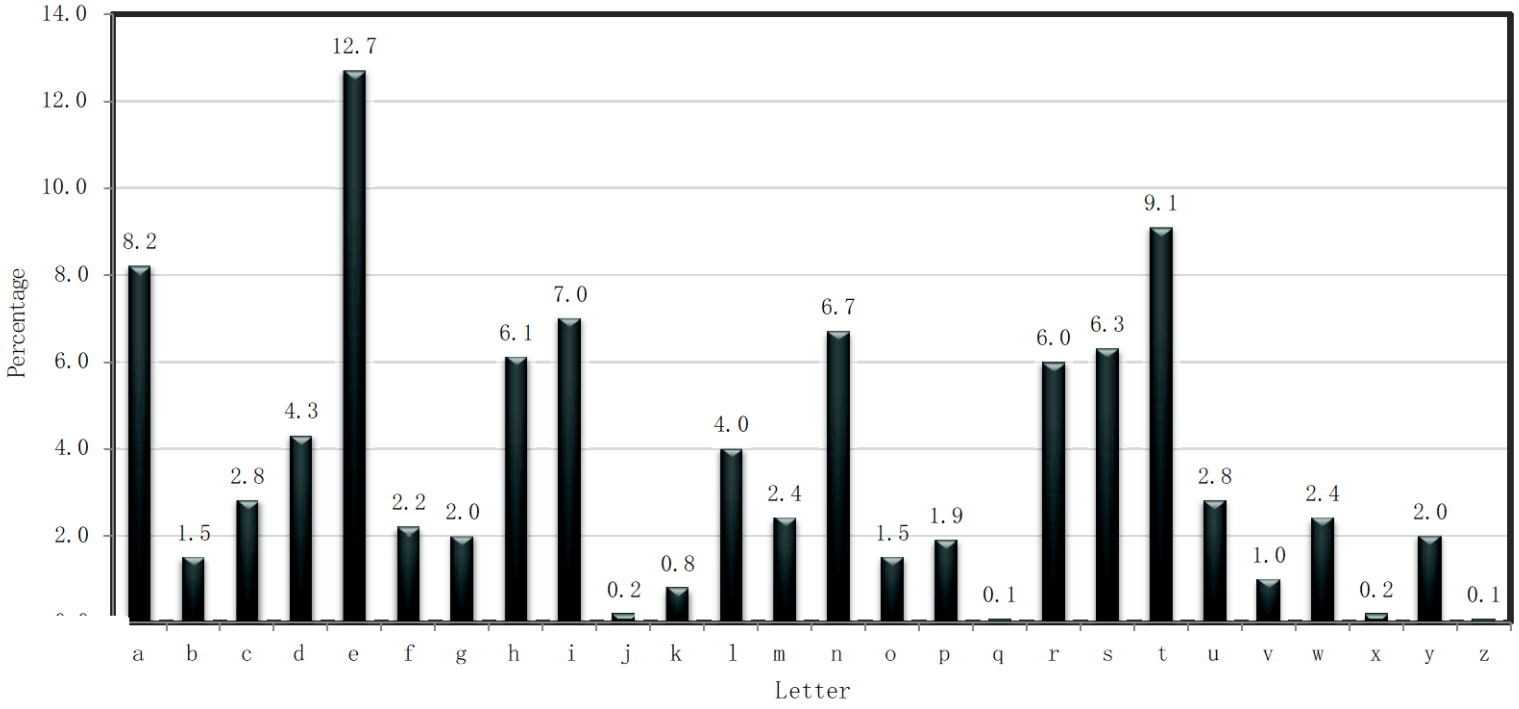
- 单字母替换密码的替换方法
  - 随机选取一段英文文本，挑选不重复的字母

    The problems of cryptography and secrecy systems furnish an interesting application of communication theory. In this paper a theory of secrecy systems is developed. The approach is on a theoretical level and is intended to complement the treatment found in standard works on cryptography.

    THEPROBLMSFCYGANDUIVW……
    - 字符 J K Q X Z 在普通文本中出现频率过低
  - 逐项暴力法（Entry-by-entry brute force method）：依次从0到25之间**随机**取一个之前没有出现过的数值，排序即可获得一个替换
  - 费舍尔-耶茨洗牌（Fisher-Yates shuffles）：从任意一个置换（permutation）开始，从头到尾遍历数组，对于每个位置 i，将其与从位置 i 到数组末尾之间**随机**选择的元素进行**交换**。

古典密码学 —— 单字母替换密码

- 单字母替换密码的分析方法
  - 英文文本的单字母概率分布已知
  - 频率统计方法



古典密码学 —— 单字母替换密码

- 单字母替换密码
  - 单字母替换密码的密钥空间大小为

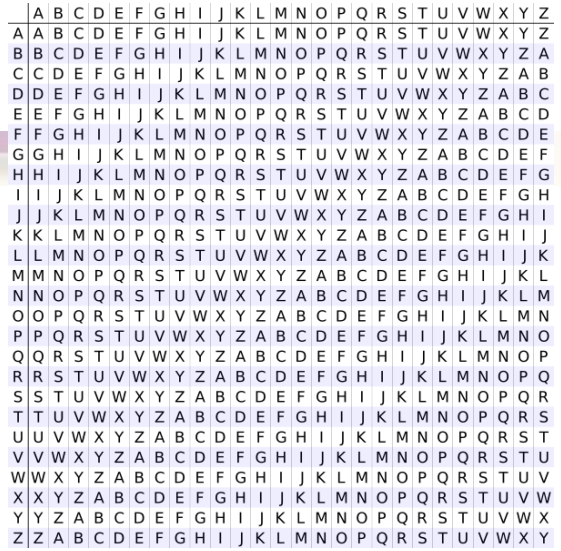
$$26! = 26 \times 25 \times 24 \times \cdots \times 2 \times 1 \approx 2^{88}$$

  - 虽然密钥空间很大，但是不能有效抵抗频率统计方法
- 单字母替换密码的启示
  - 加解密算法不能暴露明文的统计特征：比如语言字母的频率、固有的特性和模式。

古典密码学—单字母替换（单表）替换

- 问题
  - 明文中字母发生的频率没有被随机化，每个字母被加密成唯一的另外的一个字母
- 如何掩盖加密后密文的统计规律
  - 多表替换（Polyalphabetic substitution）
    - Vigenere密码
    - http://en.wikipedia.org/wiki/Vigenere_cipher
  - 加密多字母（Polygraphic substitution）
    - Playfair 密码
    - http://en.wikipedia.org/wiki/Playfair_cipher
    - Hill密码

古典密码学 —— 多表替换密码



- Vigenère密码
  - Blaise de Vigenère 发明
  - 使用多个单字母替换表
  - 相同的字母可以被多个字母替换
  - 依次使用多个移位加密，当密钥字符用完时折回

```
m:  thisprocesscanalsobeexpressed
k:  CIPHERCIPHERCIPHERCIPHERCIPHE
c:  VPXZTIQKTZWTCVPSWFDMTETIGAHLH
```

  - 假定密钥字长度为 10，密钥空间大小?
- Vigenère密码是否安全?

$26^{10} \approx 2^{47}$

古典密码学 —— Vigenère密码的分析

- 分析方法
  - 密钥空间大小为 26^t，若 t 很大，使用计算机穷尽密钥搜索也需要很长时间
  - 寻找密钥长度，将问题变成简单的移位密码
- 如何寻找 t？
  - 两种方法
    - Kasiski测试法，1863 （查尔斯·巴贝奇 Charles Babbage, 1846）
    - 重合指数法分析，1920

古典密码学 —— Vigenère密码的分析

- Kasiski test（卡西斯基测试）
  - 基于以下事实
    - 两段相同的明文段将被加密成相同的密文段，则他们的位置间距为 t 的倍数
  - 算法
    - 搜索长度至少为 3 的相同的密文段
    - 记下离起始密码段的距离 $d_1,d_2,d_3...$
    - 则 t 整除 $\gcd(d_1,d_2,d_3...)$

古典密码学 —— Vigenère密码的分析

- Kasiski test
  - 例

明文:

cryptoisshortforcryptography

密钥:

ABCDABCDABCDABCDABCDABCDABCDABCD

密文:

CSASTPKVSIQUTGQUCSASTPIUAQJB

距离为16， 则 t 可能为 $4, 8, 16$

古典密码学 —— Vigenère密码的分析

- 重合指数法（Index of coincidence）
  - 定义：设 $\boldsymbol{x} = x_1x_2x_3 \ldots, x_n$ 是一条长度为 $\boldsymbol{n}$ 的串，$\boldsymbol{x}$ 的重合指数 $\boldsymbol{I_c(x)}$ 定义为 $\boldsymbol{x}$ 中两个随机元素相同的概率
    - 设 $\boldsymbol{f_0, f_1, \ldots f_{25}}$ 分别表示A, B, …, Z在串 $\boldsymbol{x}$ 中出现的次数，则

$$\boldsymbol{I_c(x) = \frac{\sum_{i=0}^{25}\binom{f_i}{2}}{\binom{n}{2}} = \frac{\sum_{i=0}^{25} f_i(f_i-1)}{n(n-1)}}$$

    - 若 $\boldsymbol{x}$ 是英语文本串，则 $\boldsymbol{I_c(x) = \sum_{i=0}^{25} p_i^2 = 0.065}$
    - 若 $\boldsymbol{x}$ 是一个完全的随机串，则

$$\boldsymbol{I_c(x) = \frac{1}{26} = 0.038}$$

古典密码学 —— Vigenère密码的分析

- $y = y_1 y_2 y_3 \cdots y_n$, 将 y 分割为 t 个长度相等的子串

$$\boldsymbol{y_1} = y_1\, y_{1+t} y_{1+2t} y_{1+3t} \cdots$$
$$\boldsymbol{y_2} = y_2\, y_{2+t} y_{2+2t} y_{2+3t} \cdots$$
$$\cdots \qquad\qquad \cdots$$
$$\boldsymbol{y_t} = y_t\, y_{t+t} y_{t+2t} y_{t+3t} \cdots$$

古典密码学 —— Vigenère密码的分析

- 重合指数法
  - 若 t 猜对，则每一条字符串的重合指数接近于0.065
  - 若 t 猜错，则每一条字符串的重合指数接近于0.038
- Kasiski 测试

CHREEVOAHMAERATBIAXXWTNXBEEOPHBSBQMQEQERBW
RVXUOAKXAOSXXWEAHBWGJMMQMNKGRFVGXWTRZXWIAK
LXFPSKAUTEMNDCMGTSXMXBTUIADNGMGPSRELXNJELX
VRVPRTULHDNQWTWDTYGBPHXTFALJHASVBFXNGLLCHR
ZBWELEKMSJIKNBHWRJGNMGJSGLXFEYPHAGNRBIEQJT
AMRVLCRREMNDGLXRRIMGNSNRWCHRQHAEYEVTAQEBBI
PEEWEVKAKOEWADREMXMTBHHCHRTKDNVRZCHRCLQOHP
WQAIIWXNRMGWOIIFKEE

古典密码学 —— Vigenère密码的分析

- 重合指数
  - $t = 1,\ \ I_c = 0.045$
  - $t = 2,\ \ I_c = 0.046, 0.041$
  - $t = 3,\ \ I_c = 0.045, 0.050, 0.047$
  - $t = 4,\ \ I_c = 0.042, 0.039, 0.045, 0.040$
  - $t = 5,\ \ I_c = 0.063, 0.068, 0.069, 0.061, 0.072$
- Kasiski 测试
  - CHR 出现5次，位置 $1, 166, 236, 276, 286,$
  - $t = \gcd(165, 235, 275, 285) = 5$

Playfair Cipher

- 产生密码表
  - 5*5=25的密钥表
    - I和J看成一个字母
    - 第一行（列）是密钥, 密钥是一个单词或词组，去掉重复字母。例如：Playfair example
    - 其余按照字母顺序

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

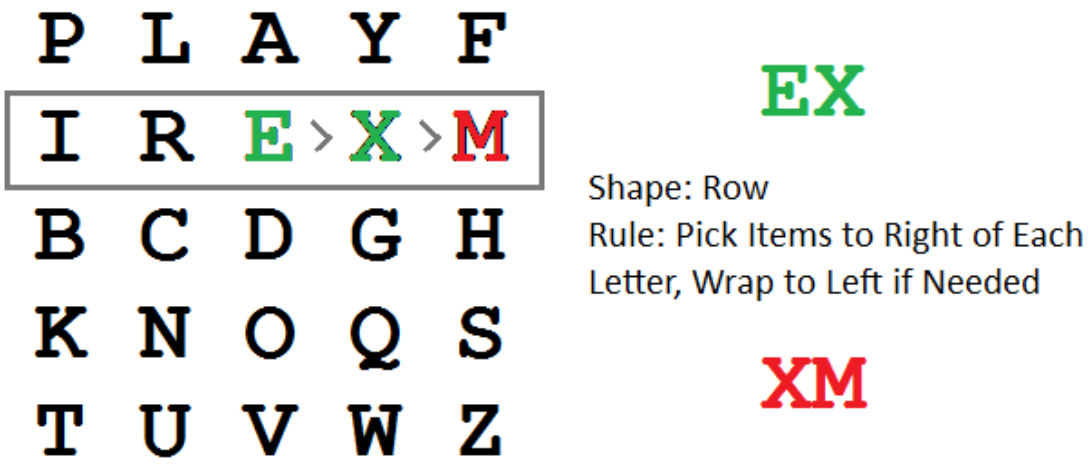
P	L	A	Y	F
I	R	E	X	M
B	C	D	G	H
K	N	O	Q	S
T	U	V	W	Z

Playfair Cipher

- 消息分组
  - 将消息分成两字母一组。如果成对后有两个相同字母紧挨或最后一个字母是单个的，就插入一个字母X
    - 例： communist，应成为co,mx,mu,ni,st

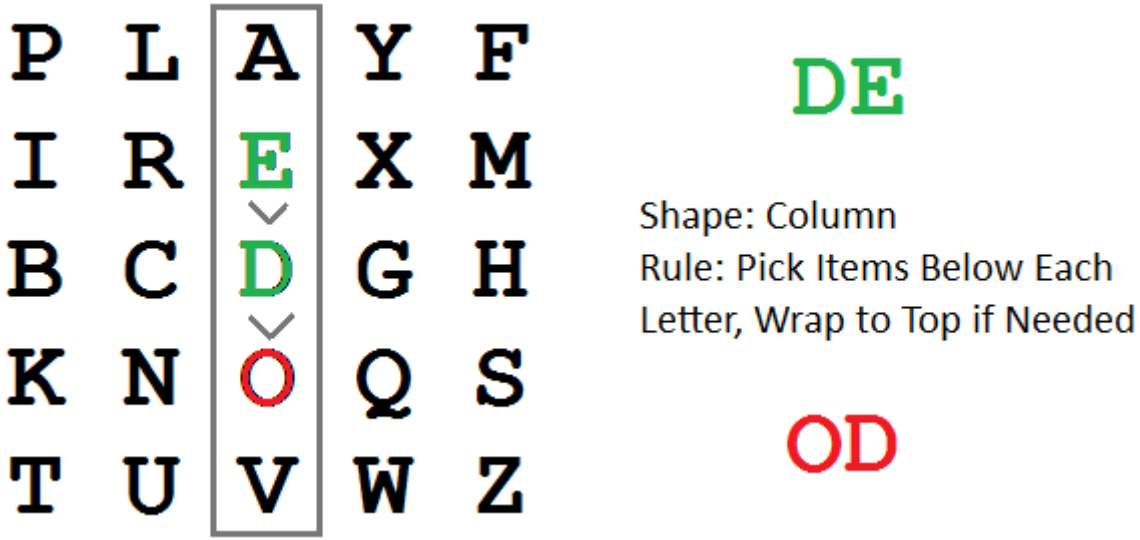
Playfair Cipher

- 加密消息
  - 若明文p_1p_2在同一行，对应密文c_1c_2分别是紧靠p_1p_2右端的字母。其中第一列被看做是最后一列的右方。
    - 如，按照前表，EX→XM



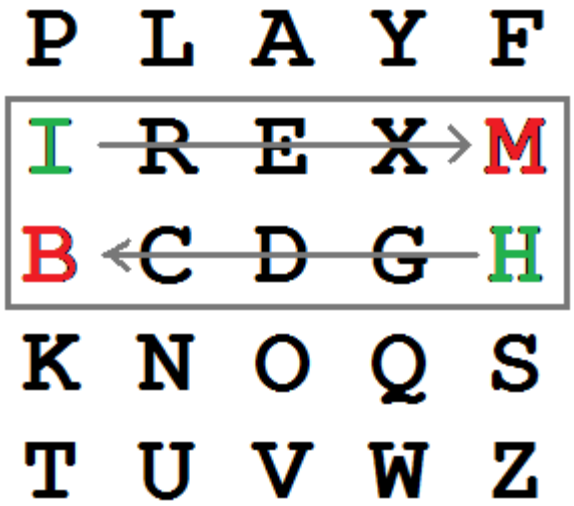
Playfair Cipher

- 加密消息
  - 若p_1 p_2在同一列，对应密文c_1 c_2分别是紧靠p_1 p_2 下方的字母。其中第一行被看做是最后一行的下方。
    - 例如 DE→OD



Playfair Cipher

- 加密消息
  - 若p_1 p_2不在同一行，不在同一列，则c_1 c_2是由p_1 p_2确定的矩形的其他两角的字母
    - 如：HI→BM

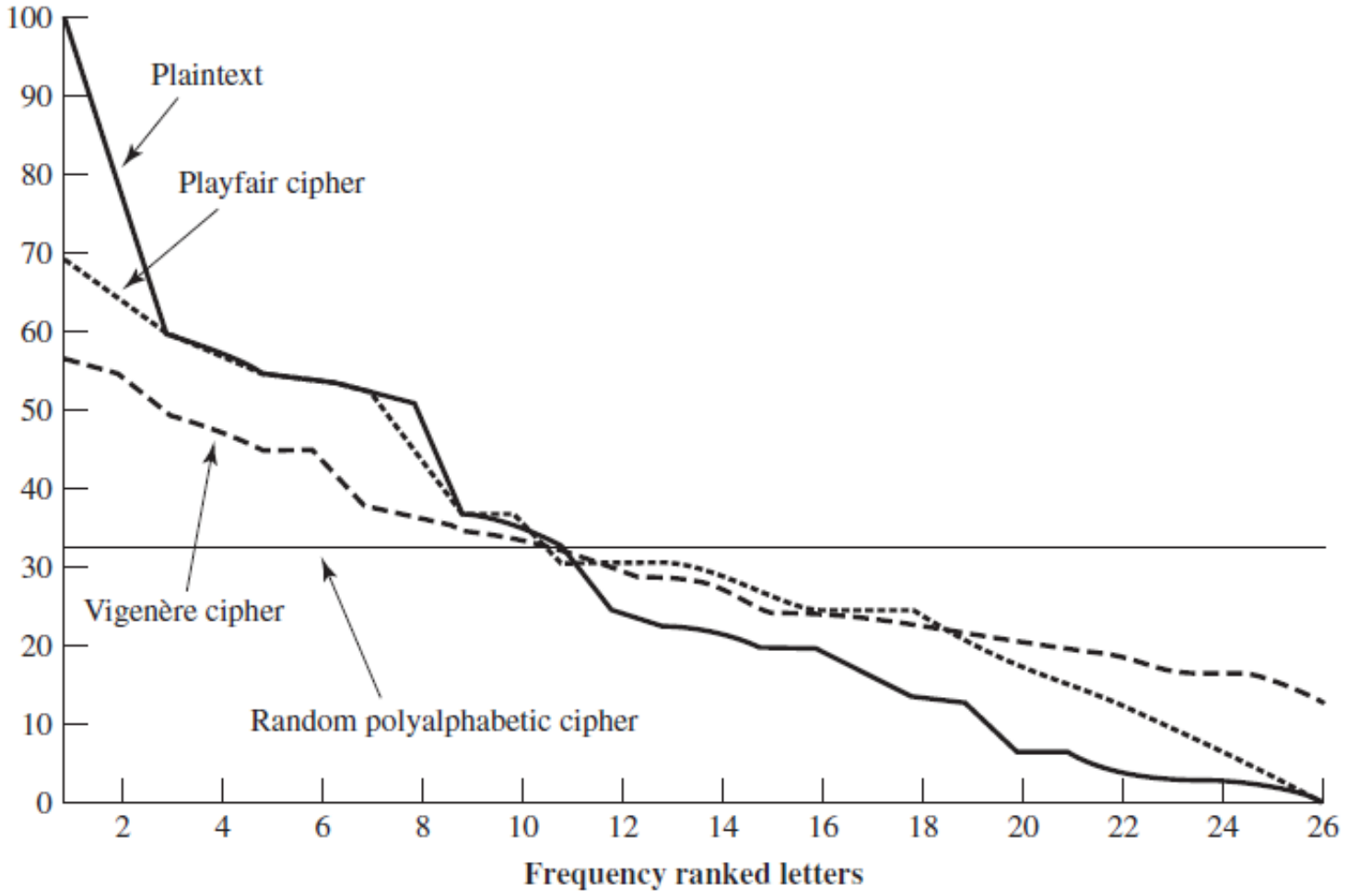


HI

Shape: Rectangle
Rule: Pick Same Rows,
Opposite Corners

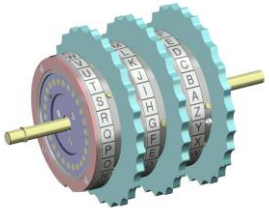
BM

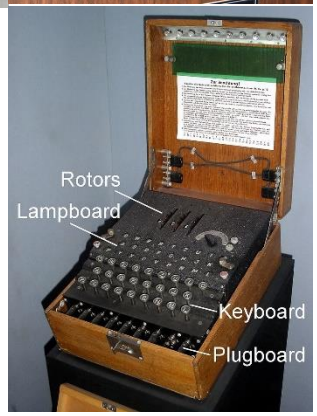
频率特征


100
90
80
70
60
50
40
30
20
10
0
Plaintext
Playfair cipher
Vigenère cipher
Random polyalphabetic cipher
2
4
6
8
10
12
14
16
18
20
22
24
26
Frequency ranked letters

密码学发展史 —— 古典密码

- 与计算设备有关
  - 密码盘：15世纪-20世纪
    - 简化了加密过程
    - 单表代换、多表代换
  - 机械或机电装置
    - 轮转机（Roter Machines: 1920s-1960)
      - Enigma密码机(德国)
      - Sigaba（美国）
      - Typex（英国）
      - Lorenz SZ 40/42（德国，盟军代号“金枪鱼”）
      - Siemens and Halske T52（德国，盟军代号“鲟鱼”）
      - 紫密码机（日本，Type B Cipher Machine, codenamed Purple）

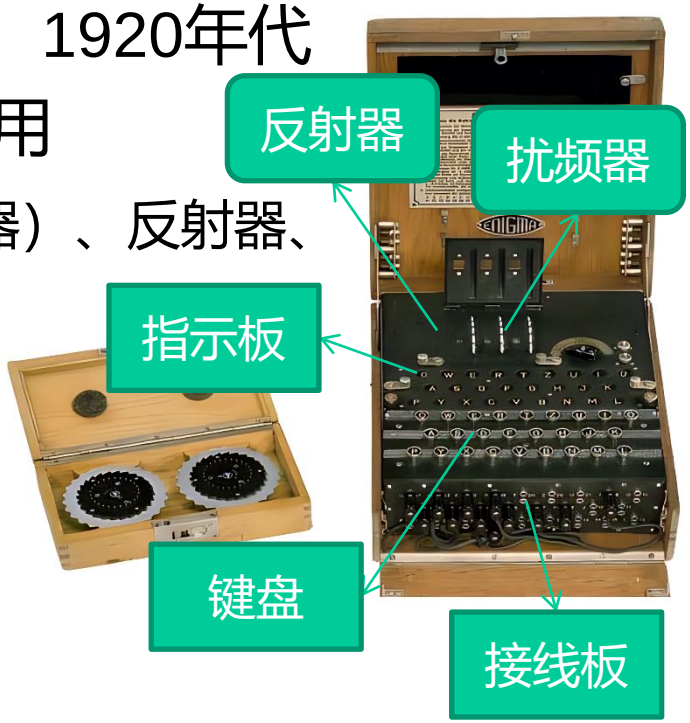


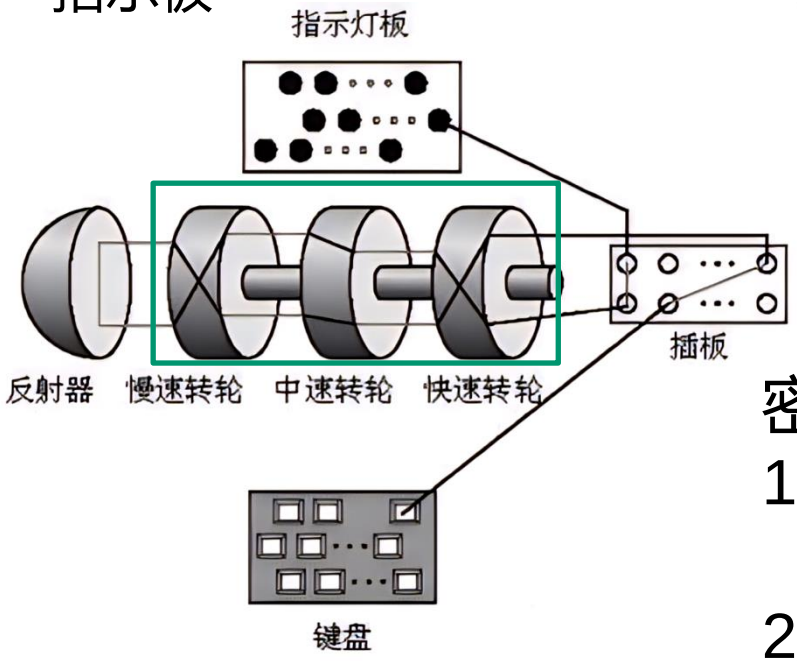


Enigma machine

密码学发展史 —— 二战时期

- Enigma密码机由德国Scherbius发明，1920年代被用于商业，在二战中被军事广泛使用
  - 键盘、接线板（插板）、扰频器（轮转器）、反射器、指示板



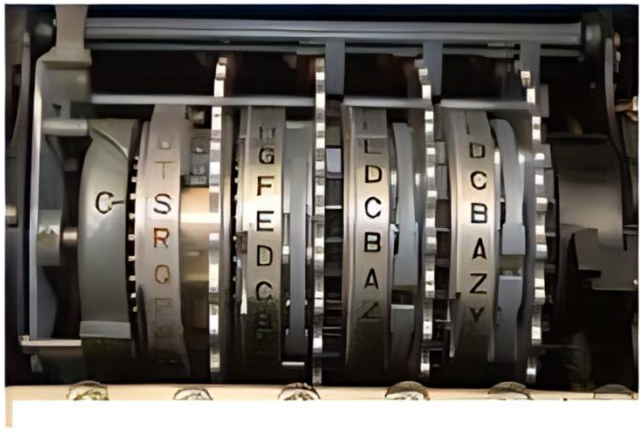


密钥：

1. 接线板设置： A/L, P/R, T/D, B/W, K/F, O/Y
2. 扰频器排列：2-3-1
3. 扰频器定位：Q-C-W

密码学发展史 —— 二战时期

- Enigma密码机
  - 3个扰频器：
    - $26 \times 26 \times 26 = 17576$
  - 接线板：26个字母中任意交换6对
    - 100391791500
  - 密钥总的数目
    - $17576 \times 6 \times 100391791500 = 10^{16} = 2^{53.15}$

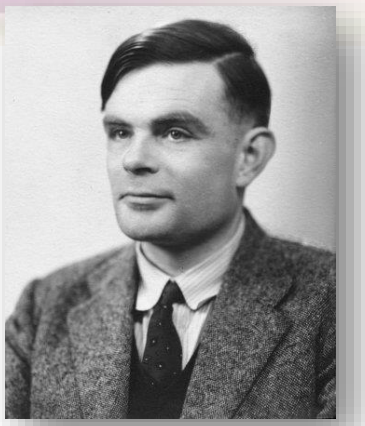


密码学发展史 —— ENIGMA密码机的破解

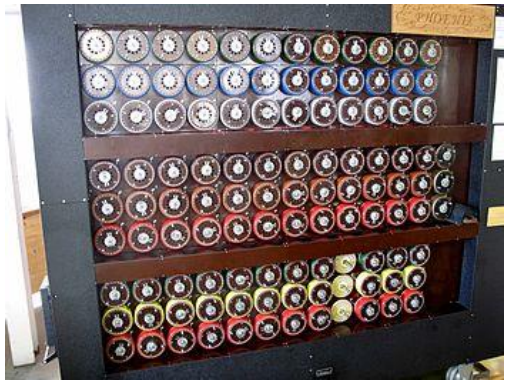
- 1932年，波兰密码学家Rejewski，Zygalski, 和Rozycki设计Bomba破译德军使用的ENIGMA (3个扰频器)
  - Schmidt出卖军用密码机扰频器文件给法国
  - 密钥设定： 每一条信息传递新密钥(消息密钥)
  - 每条消息密钥重发送两次
    - 原始密钥（扰频器定位）QCW, 消息密钥PGH
    - PGHPGH加密成KIVBJE
- 1938年12月德国增加了ENIGMA的安全性(5个扰频器)，Rejewski的技术破解受限
- 1939年6月波兰将Bomba技术提供给法国和英国

密码学发展史 —— ENIGMA密码机的破解

- 自1939年英国的科学家和数学家在Bletchley（布莱切里）庄园开始了新的ENIGMA的破译工作
- 图灵设计了the British bombe，利用已知的明密文对破译ENIGMA
- 得到德军法国边境集结的详细计划：盟军法国诺曼底登陆，德军损失四十万
- 美国弗里德曼小组经过
  两年的努力破解了日本紫密，
  从而击落了日本舰队总司令
  山本五十六座机
  Enigma的破解使得二战至少提前一年结束



阿兰. 图灵



the British bombe:图灵机模型

密码学发展史 —— 一次一密（one-time pad）

- 一次一密乱码本：一个大的不重复的真随机密钥字母集被写在几张纸上并粘在一起成为一个乱码本。
- 发方：用乱码本中的每一密钥字母加密一个明文字符(明文与密钥模26加)，每个密钥仅对一个消息使用一次，加密后销毁乱码本中用过的部分。
- 收方：有一个同样的乱码本，并依次使用每个密钥去解密密文的每个字符。收方解密后也同样销毁乱码本中用过的部分。
- 新的消息用乱码本新的密钥加密，不能重复使用。
- 所以叫做“one-time pad”。
- 对现代密码的影响：
  - 流密码、分组密码CTR工作模式

One-time Pad(OTP)

- 例：

明文 H E L L O　　L A T E R

密钥 X M C K L　　T Q U R A

密文 E Q N V Z　　E Q N V 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 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

H→7, X→23, 7+23(mod 26)=4(E)

香农简介

- 1949年《保密系统的通信理论》
  "Communication Theory of Secrecy Systems"
- 利用信息熵在理论上**证明了一次一密是无法被破译的，同时证明了一个无法被破译的密码系统必须具备与一次一密相同的条件**，即密钥必须有以下特征：
  - 完全随机
  - 不能重复使用
  - 保密
  - 和明文一样长
- 使保密通信由艺术变成科学
- 密码设计的新思想，对现代密码体制设计非常重要。

OTP的完善保密性

- 一个密码体制的完善保密性是指已知的密文不会泄露明文的任何信息
- 定义：一个密码体制具有完善保密性， 如果对任意的明文p和任意的密文c, 都有

$$Pr(P=p|C=c)=Pr(P=p)$$

  - Pr(P=p|C=c)是已知密文c时明文p的后验概率
  - Pr(P=p)是明文p的先验概率
- 即使知道密文后，攻击者也不能以更高的概率猜测出明文

OTP的完善保密性

- One-time pad
  - $P=C=K=\{0,1\}^n$
  - K随机产生
    - $Pr(K=k)=1/2^n$
  - 证明$Pr(P=p|C=c)=P(P=p)$
- 证明

$$\Pr[C=c \mid P=p]=\Pr[K=p\oplus c]=1/2^n$$

$$\Pr[C=c]=\sum_{p\in P}\Pr[P=p]\Pr[C=c \mid P=p]$$

$$=1/2^n\sum_{p\in P}\Pr[P=p]=1/2^n$$

$$\therefore \Pr(P=p \mid C=c)=\frac{\Pr[C=c \mid P=p]\Pr[P=p]}{\Pr[C=c]}=\Pr[P=p]$$

密码学发展史 —— 古典密码

David Kahn: “The Codebreakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet”, edition Rev Sub, 1996, Scribner. ISBN 978-0-684-83130-5

Simon Singh: “The Code Book: The Secret History of Codes and Code-breaking”, 2000, New edition, Fourth Estate, ISBN-13: 978-1857028898

V • T • E	Classical cryptography	
Ciphers by family	Polyalphabetic	Alberti • Enigma • Trithemius • Vigenère
	Polybius square	ADFGVX • Bifid • Nihilist • Tap code • Trifid • VIC cipher
	Square	Playfair • Two-square • Four-square
	Substitution	Affine • Atbash • Autokey • Beaufort • **Caesar** • Chaocipher • Great • Hill • Pigpen • ROT13 • Running key
	Transposition	Columnar • Double • Myszkowski • Rail fence • Route
	Other	BATCO • DRYAD • Kama Sutra • One-time pad • Rasterschlüssel 44 • Reihenschieber • Reservehandverfahren • Slidex • Solitaire
Codes	Book • Code talker • Poem	
Steganography	Bacon • Grille • Null	
Cryptanalysis	Cryptogram • Frequency analysis • Index of coincidence (Units: Ban and Nat) • Information leakage • Kasiski examination	
V • T • E	**Cryptography**	[show]

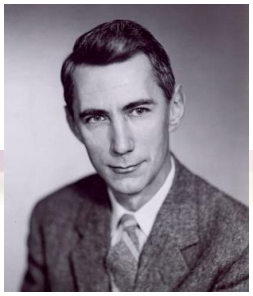
Categories: Classical ciphers | Group theory | Julius Caesar - *https://en.wikipedia.org/wiki/*

密码体制的安全性

- 无条件安全性(unconditional security)即完善保密性(perfect security)
  - 即使攻击者有无限的计算资源也不可能攻破密码体制，则该密码体制是无条件安全的
  - OTP是无条件安全的

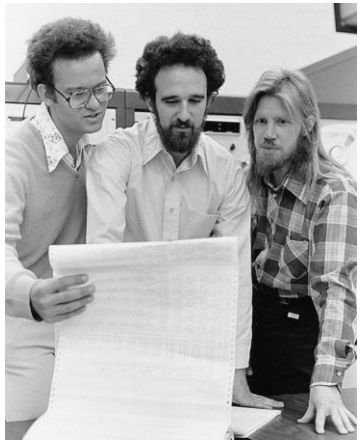
- 计算安全性(computational security)
  - 破译一个密码体制所做的计算上的努力
  - 如果使用最好的算法破译一个密码体制至少需要N次操作（N是一个特定的非常大的数字），定义该密码体制是计算安全的
- 可证明安全性(provable security)
  - 通过规约的方式为安全性提供证据
  - 如果可以破译密码体制A，则就可以解决一个数学难题B（分解因子问题，离散对数问题）

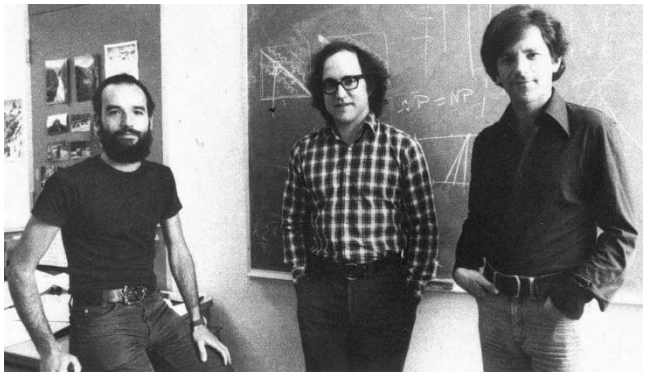
密码学发展史 —— 现代密码学



- Claude Shannon

- 与计算设备有关 — 电子计算机
  - 1948~1949年，Claude Shannon先后发表了《通信的数学理论》和《保密系统的信息理与计算设备有关论》，提出信息熵的概念，建立了信息论，为密码学和其他领域提供了理论基础，使密码学从艺术成为科学。
  - 私钥密码体制（对称密码体制）：
    - DES（1976年），AES（2001年）
  - 密钥分发问题
  - 公钥密码体制： RSA, ECC
    - Whitfield Diffie，Martin Hellman
    - Ralph Merkle
    - Ron. Rivest, Adi Shamir, Leonard Adleman.

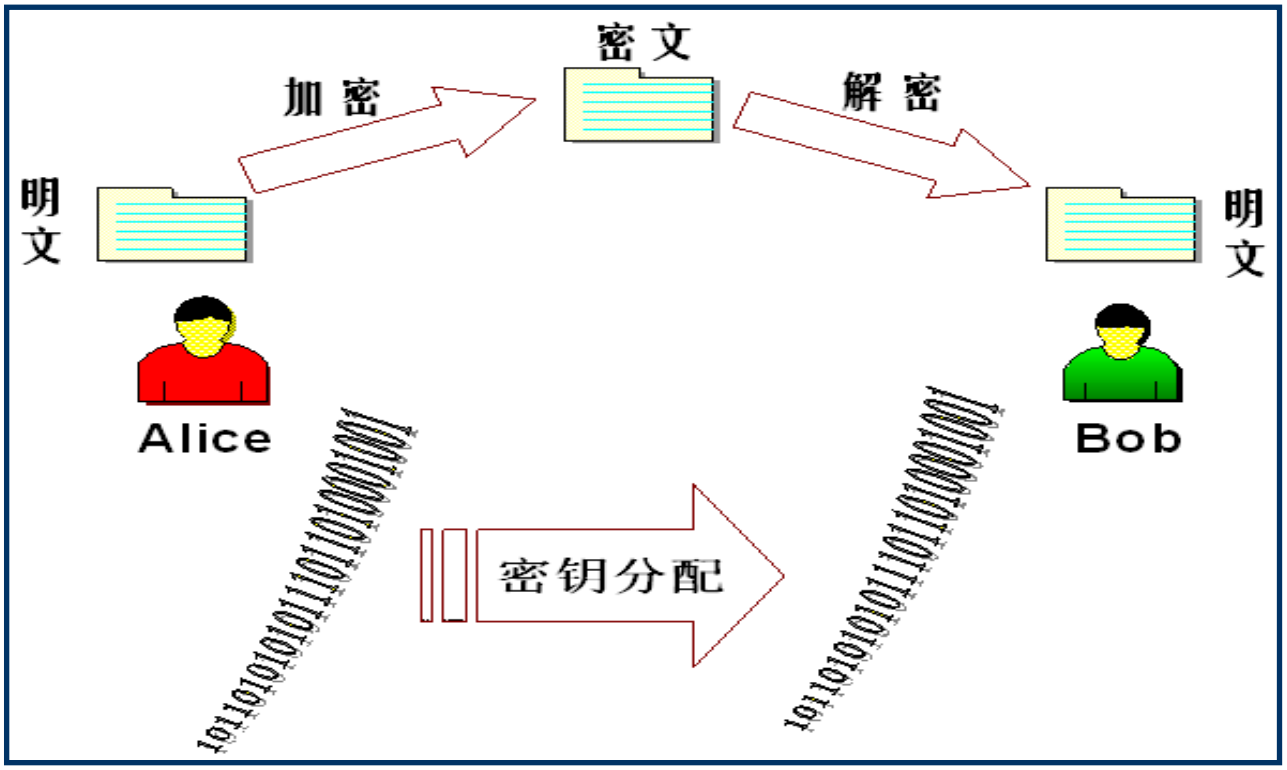




- Ron Rivest, Adi Shamir and Leonard Adleman

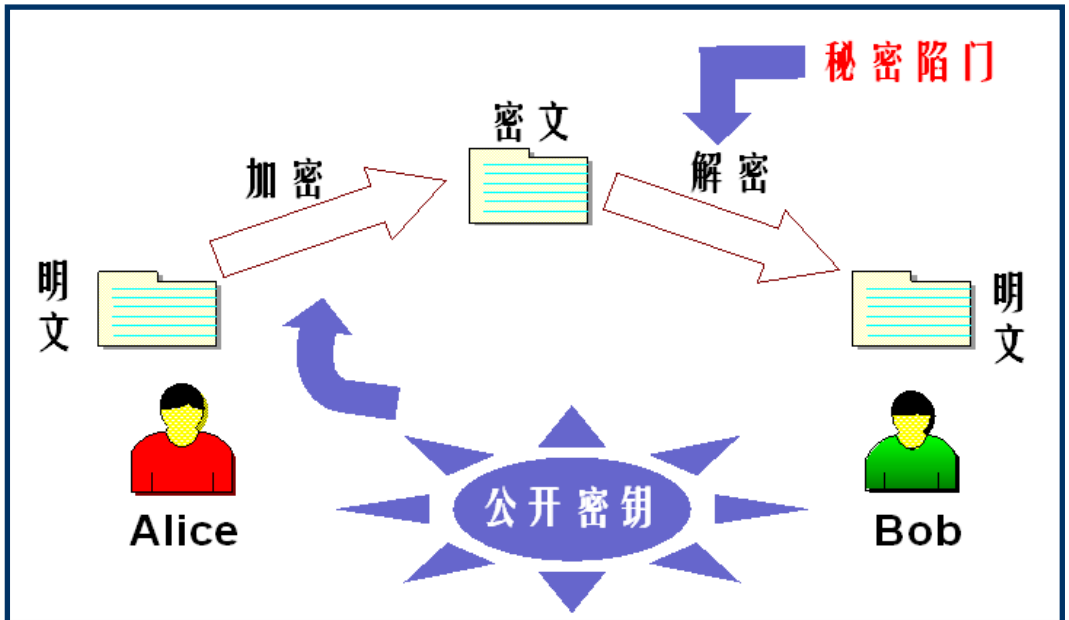
现行密码体制

- 对称密码体制：基于代换、置换、移位寄存器等
  - 优点：加密速度快，适合批量加密数据
  - 缺点：密钥分配、密钥管理、没有签名功能



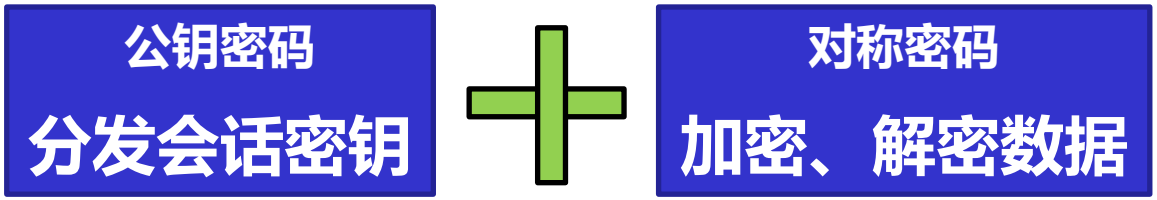
现行密码体制

- 对称密码体制：基于代换、置换、移位寄存器等
  - 优点：加密速度快，适合批量加密数据
  - 缺点：密钥分配、密钥管理、没有签名功能
- 公钥密码体制：基于大数分解等数学难题
  - 优点：可解决密钥分配、管理问题，可用于签名
  - 缺点：加密速度慢



现行密码体制

- 对称密码体制：基于代换、置换、移位寄存器等
  - 优点：加密速度快，适合批量加密数据
  - 缺点：密钥分配、密钥管理、没有签名功能
- 公钥密码体制：基于大数分解等数学难题
  - 优点：可解决密钥分配、管理问题，可用于签名
  - 缺点：加密速度慢
- 实际使用：混合密码体制

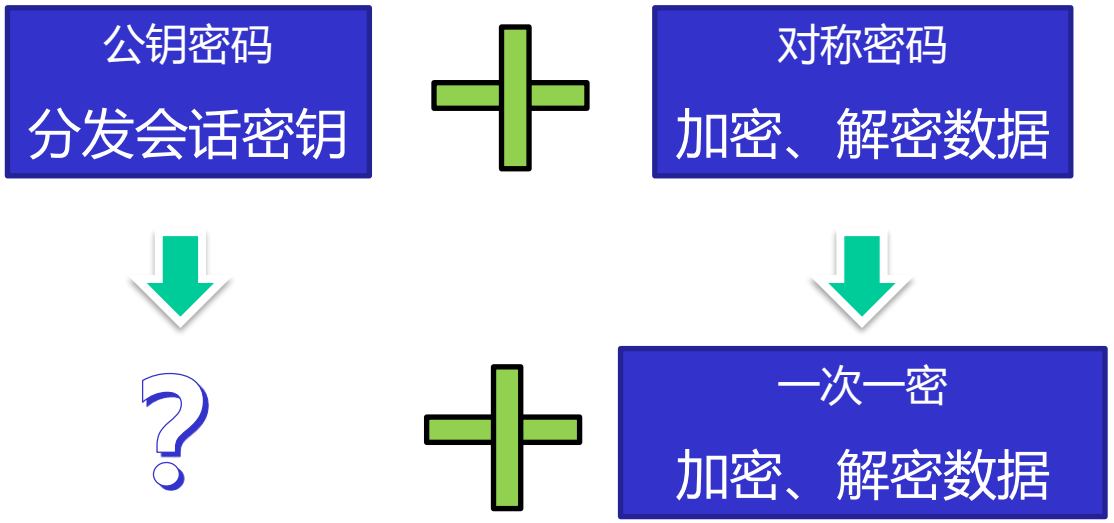


量子计算的挑战

- Shor算法（1994）：多项式时间内解决大数分解问题
  - 受影响密码体制：RSA等大多数公钥密码
  - 影响程度：完全攻破
- Grover算法（1996）：快速穷搜索密钥
  - 受影响密码体制：DES，AES等对称密码
  - 影响程度：相当于密钥长度被缩短了一半

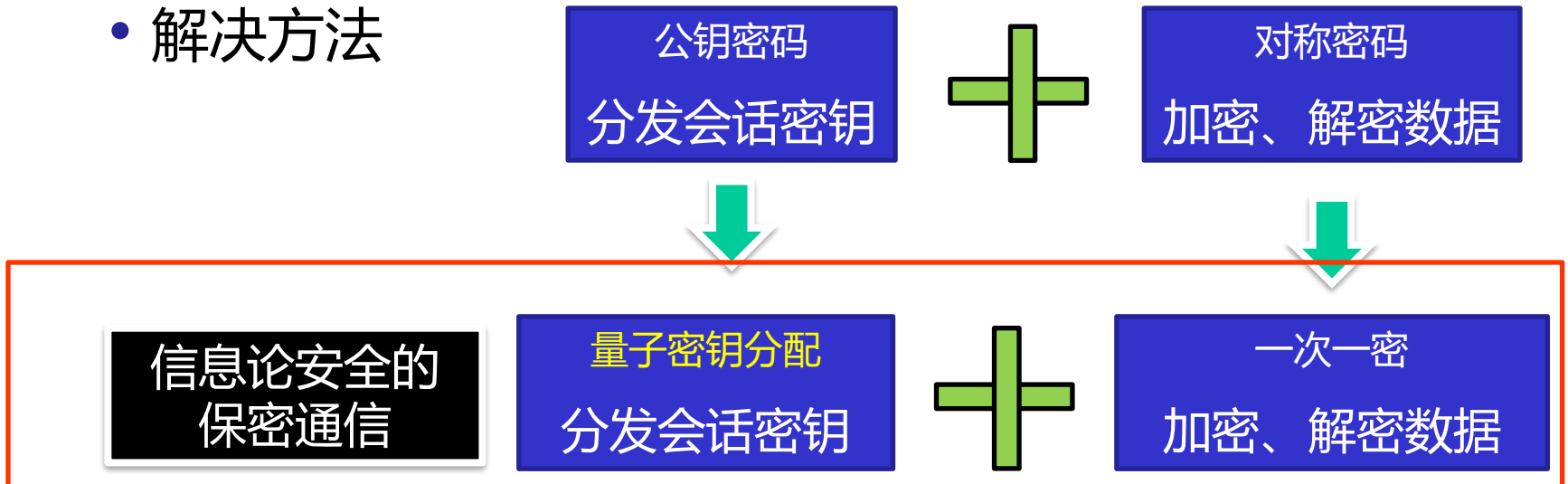
量子计算的挑战

- Shor算法（1994）：多项式时间内解决大数分解问题
  - 受影响密码体制：RSA等大多数公钥密码
  - 影响程度：完全攻破
- Grover算法（1996）：快速穷搜索密钥
  - 受影响密码体制：DES，AES等对称密码
  - 影响程度：相当于密钥长度被缩短了一半
- 解决方法



量子计算的挑战

- Shor算法（1994）：多项式时间内解决大数分解问题
  - 受影响密码体制：RSA等大多数公钥密码
  - 影响程度：完全攻破
- Grover算法（1996）：快速穷搜索密钥
  - 受影响密码体制：DES，AES等对称密码
  - 影响程度：相当于密钥长度被缩短了一半
- 解决方法

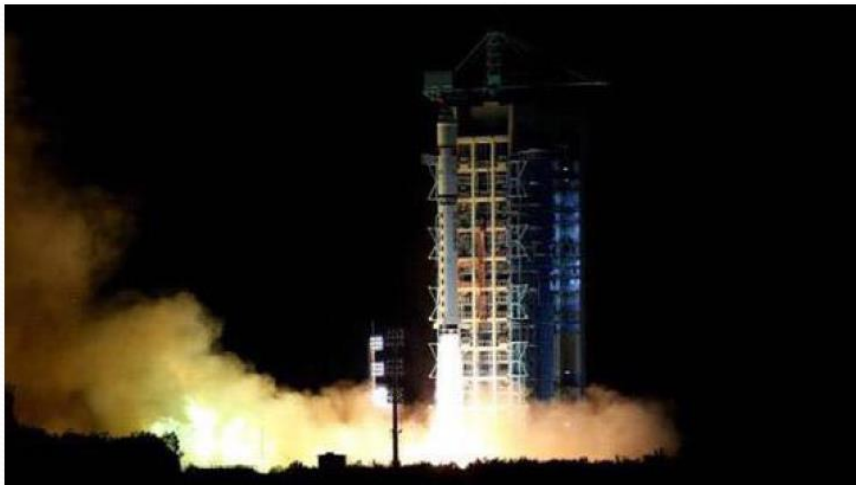


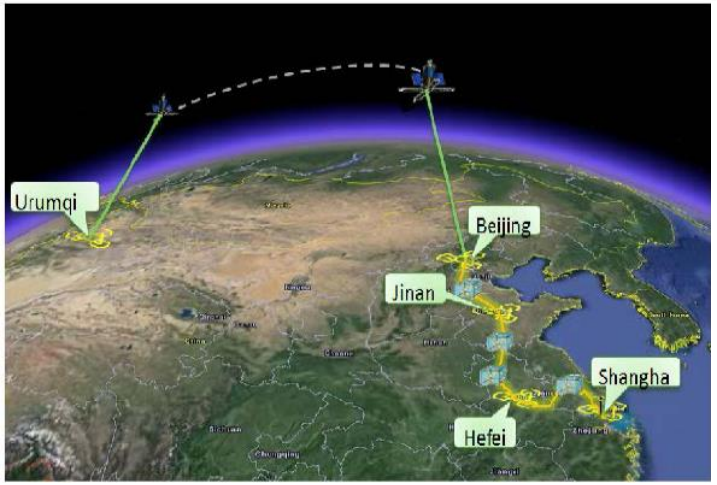
量子密码

量子密钥分发

京沪干线

墨子号量子卫星

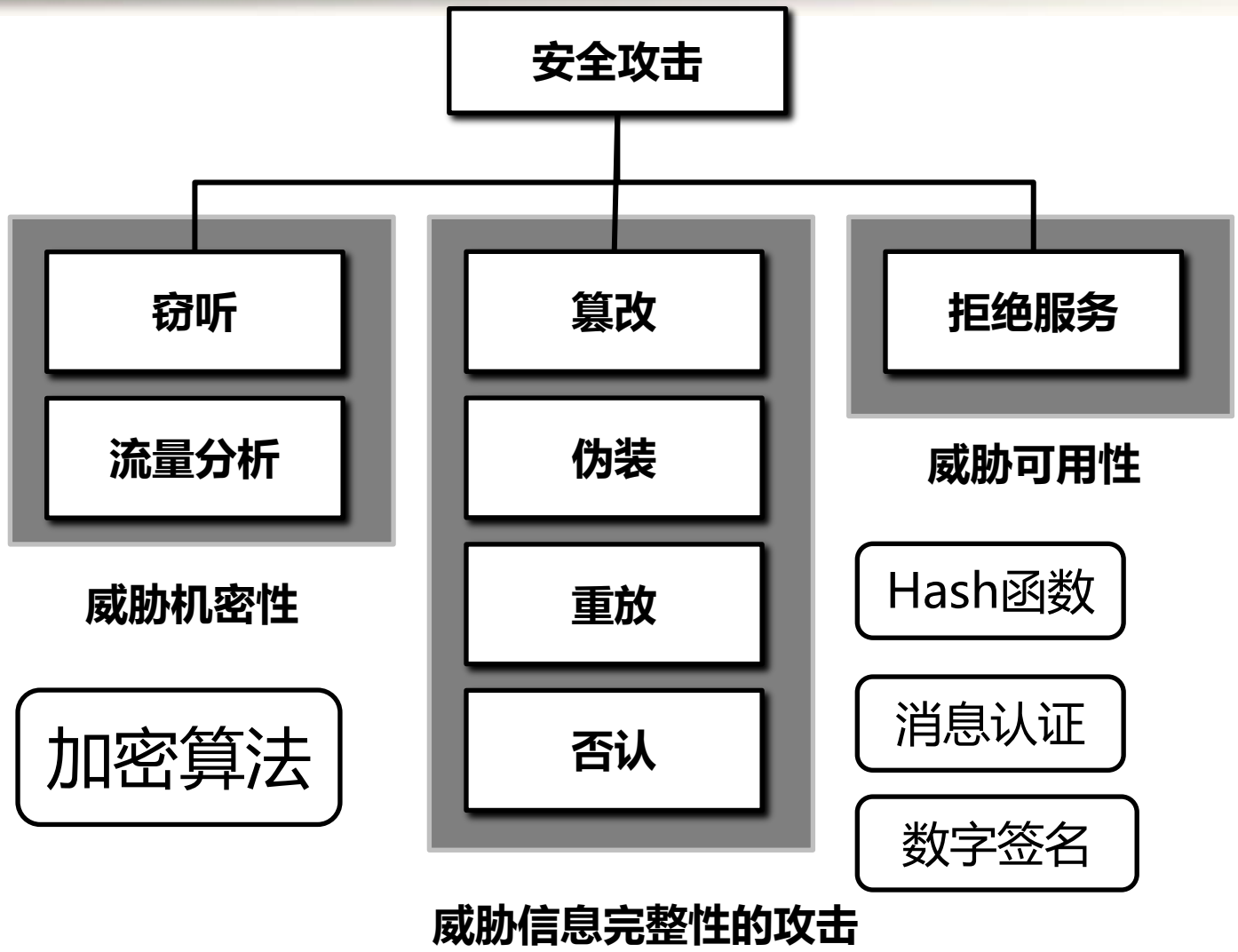




安全目标： CIA

- Confidentiality：保密性、机密性
  - 保护信息内容不会被泄露给未授权的实体
  - 保密场景：业务数据、网络拓扑、流量
- Integrity：完整性
  - 保证信息不被未授权地修改，或者可以检测出非授权修改
  - 攻击示例：篡改、插入、重放
- Availability：可用性
  - 保证资源的授权用户能够访问到应得资源或服务
  - 攻击示例：拒绝服务（网络、系统、硬件、人）

信息安全面临的威胁


安全攻击
窃听
流量分析
威胁机密性
篡改
伪装
重放
否认
威胁信息完整性的攻击
拒绝服务
威胁可用性
加密算法
Hash函数
消息认证
数字签名

几条密码常识

- **不要使用保密的密码算法**
  - “由公司开发一种密码算法， 并将这种算法保密，这样就能保证安全” ？
  - 使用公开的，被公认为强度较高的密码算法
    - 密码算法早晚会被公诸于世， 如RC4
    - 开发高强度的密码算法是非常困难的
- **使用低强度的密码比不进行任何加密更危险**
  - 获得一种错误的安全感， 如16世纪苏格兰玛丽女王

几条密码常识

- 任何密码总有一天都会被破解
  - 一次一密 完美密码
  - 量子密码 可能完美的密码
- 密码只是信息安全的一部分
  - 安全是“系统安全”，系统的强度取决于其中最脆弱的环节的强度

谢　谢！

课后预习

- 二战Enigma密码
  - 密码机的原理
  - 加密过程
  - 秘钥的产生和传递（日密码，通信密码）
  - Enigma的破译

Enigma密码阅读

https://www.ciphermachinesandcryptology.com/index.htm

Enigma密码破译：电影《模仿游戏》

代换密码频率分析

- 公元9世纪，阿拉伯科学家al-Kindi 发明
- 主要思想
  - 代换密码没有掩藏密文字母出现的频率
  - 计算密文中字母出现的频率，与（英文）字母统计表相比较，很容易确定代换表（密码表）。

代换密码频率分析

- 26个英文字母出现的概率

字母	概率	字母	概率
A	0.082	N	0.067
B	0.015	O	0.075
C	0.028	P	0.019
D	0.043	Q	0.001
E	0.127	R	0.060
F	0.022	S	0.063
G	0.020	T	0.091
H	0.061	U	0.028
I	0.070	V	0.010
J	0.002	W	0.023
K	0.008	X	0.001
L	0.040	Y	0.020
M	0.024	Z	0.001

代换密码频率分析

- 两个连续的字母（diagrams）出现的概率

th 1.52%	en 0.55%	ng 0.18%
he 1.28%	ed 0.53%	of 0.16%
in 0.94%	to 0.52%	al 0.09%
er 0.94%	it 0.50%	de 0.09%
an 0.82%	ou 0.50%	se 0.08%
re 0.68%	ea 0.47%	le 0.08%
nd 0.63%	hi 0.46%	sa 0.06%
at 0.59%	is 0.46%	si 0.05%
on 0.57%	or 0.43%	ar 0.04%
nt 0.56%	ti 0.34%	ve 0.04%
ha 0.56%	as 0.33%	ra 0.04%
es 0.56%	te 0.27%	ld 0.02%
st 0.55%	et 0.19%	ur 0.02%

代换密码频率分析

- 英文中最常用的3字母

  the, ing, and, her, ere, ent,

  tha, nth, was, eth, for, dht

代换密码频率分析

- 例：利用代换密码获得如下密文，如何恢复明文？

```
YIFQFMZRWQFYVECFMDZPCVMRZWNMDZVEJBTXCDDUMJ
NDIFEFMDZCDMQZKCEYFCJMYRNCWJCSZREXCHZUNMXZ
NZUCDRJXYYSMRTMEYIFZWDYVZVYFZUMRZCRWNZDZJJ
XZWGCHSMRNMDHNCMFQCHZJMXJZWIEJYUCFWDJNZDIR
```

在下面的分析中，明文用小写字母表示，密文用大写字母表示

代换密码频率分析

Step1：将密文字母出现的频率 与英文字母表中字母出现的频率作比较

字母	频数	字母	频数
A	0	N	9
B	1	O	0
C	15	P	1
D	13	Q	4
E	7	R	10
F	11	S	3
G	1	T	2
H	4	U	5
I	5	V	5
J	11	W	8
K	1	X	6
L	0	Y	10
M	16	Z	20

字母	概率	字母	概率
A	0.082	N	0.067
B	0.015	O	0.075
C	0.028	P	0.019
D	0.043	Q	0.001
E	0.127	R	0.060
F	0.022	S	0.063
G	0.020	T	0.091
H	0.061	U	0.028
I	0.070	V	0.010
J	0.002	W	0.023
K	0.008	X	0.001
L	0.040	Y	0.020
M	0.024	Z	0.001

Z出现次数最多，可能 S(e)=Z

代换密码频率分

Step2:考虑双字出现频率 假设S(e)=Z

字母	概率	字母	概率
A	0.082	N	0.067
B	0.015	O	0.075
C	0.028	P	0.019
D	0.043	Q	0.001
E	0.127	R	0.060
F	0.022	S	0.063
G	0.020	T	0.091
H	0.061	U	0.028
I	0.070	V	0.010
J	0.002	W	0.023
K	0.008	X	0.001
L	0.040	Y	0.020
M	0.024	Z	0.001

th 1.52%	en 0.55%	ng 0.18%
he 1.28%	ed 0.53%	of 0.16%
in 0.94%	to 0.52%	al 0.09%
er 0.94%	it 0.50%	de 0.09%
an 0.82%	ou 0.50%	se 0.08%
re 0.68%	ea 0.47%	le 0.08%
nd 0.63%	hi 0.46%	sa 0.06%
at 0.59%	is 0.46%	si 0.05%
on 0.57%	or 0.43%	ar 0.04%
nt 0.56%	ti 0.34%	ve 0.04%
ha 0.56%	as 0.33%	ra 0.04%
es 0.56%	te 0.27%	ld 0.02%
st 0.55%	et 0.19%	ur 0.02%

1） ZW 出现4次

W可能是r,s,n,d,a

2) WZ没有出现

W 不是 r

3) W出现8次（0.047）

所以 S(d)=W

代换密码频率分析

Step3：考虑双字

- 假设S(d)=W
- RW出现2次
  - R可能是e,n
    e的密文是Z
  - R可能是n
    S(n)=R

th 1.52%	en 0.55%	ng 0.18%
he 1.28%	ed 0.53%	of 0.16%
in 0.94%	to 0.52%	al 0.09%
er 0.94%	it 0.50%	de 0.09%
an 0.82%	ou 0.50%	se 0.08%
re 0.68%	ea 0.47%	le 0.08%
nd 0.63%	hi 0.46%	sa 0.06%
at 0.59%	is 0.46%	si 0.05%
on 0.57%	or 0.43%	ar 0.04%
nt 0.56%	ti 0.34%	ve 0.04%
ha 0.56%	as 0.33%	ra 0.04%
es 0.56%	te 0.27%	id 0.02%
st 0.55%	et 0.19%	ur 0.02%

代换密码频率分析

- 通过以上三步的分析

```
------end---------e----ned---e------------
YIFQFMZRWQFYVECFMDZPCVMRZWNMDZVEJBTXCDDUMJ

--------e----e---------n--d---en----e----e
NDIFEFMDZCDMQZKCEYFCJMYRNCWJCSZREXCHZUNMXZ

-e---n------n------ed---e---e--ne-nd-e-e--
NZUCDRJXYYSMRTMEYIFZWDYVZVYFZUMRZCRWNZDZJJ

-ed-----n----------e----ed-------d---e--n
XZWGCHSMRNMDHNCMFQCHZJMXJZWIEJYUCFWDJNZDIR
```

代换密码频率分析

Step4:双字母

假设S(e)=Z

1)NZ出现3次

N可能是h,r,t

2)ZN没有出现

N可能不是r,t

故 可能 S(h)=N

th 1.52%	en 0.55%	ng 0.18%
he 1.28%	ed 0.53%	of 0.16%
in 0.94%	to 0.52%	al 0.09%
er 0.94%	it 0.50%	de 0.09%
an 0.82%	ou 0.50%	se 0.08%
re 0.68%	ea 0.47%	le 0.08%
nd 0.63%	hi 0.46%	sa 0.06%
at 0.59%	is 0.46%	si 0.05%
on 0.57%	or 0.43%	ar 0.04%
nt 0.56%	ti 0.34%	ve 0.04%
ha 0.56%	as 0.33%	ra 0.04%
es 0.56%	te 0.27%	ld 0.02%
st 0.55%	et 0.19%	ur 0.02%

代换密码频率分析

- 考虑3个字母
  - 明文中有 ne-ndhe, 其中 ‘–’ 对应密文C
- 从3个字母组的分布看 and 出现的频率较高
  - 猜测S(a)=C

代换密码频率分析

- 进一步

```
------end-----a---e-a--nedh--e------a-----
YIFQFMZRWQFYVECFMDZPCVMRZWNMDZVEJBTXCDDUMJ

h-------ea---e-a---a---nhad-a-en--a-e-h--e
NDIFEFMDZCDMQZKCEYFCJMYRNCWJCSZREXCHZUNMXZ

he-a-n------n------ed---e---e--neandhe-e--
NZUCDRJXYYSMRTMEYIFZWDYVZVYFZUMRZCRWNZDZJJ

-ed-a---nh---ha---a-e----ed-----a-d--he--n
XZWGCHSMRNMDHNCMFQCHZJMXJZWIEJYUCFWDJNZDIR
```

代换密码频率分析

- 确定M
  - M是密文中出现频率次高的字母
    - M可能是 t, a, o , i, n, s, h, r
  - NRM解密成nh-
    - h-可能是一个词的开头
    - M应该是一个元音 o,i
  - CM出现在密文中
    - ai, ao
    - 猜测S(i)=M

字母	概率	字母	概率
A	0.082	N	0.067
B	0.015	O	0.075
C	0.028	P	0.019
D	0.043	Q	0.001
E	0.127	R	0.060
F	0.022	S	0.063
G	0.020	T	0.091
H	0.061	U	0.028
I	0.070	V	0.010
J	0.002	W	0.023
K	0.008	X	0.001
L	0.040	Y	0.020
M	0.024	Z	0.001

代换密码频率分析

```
-----iend-----a-i-e-a-inedhi-e------a---i-
YIFQFMZRWQFYVECFMDZPCVMRZWNMDZVEJBTXCDDUMJ

h-----i-ea-i-e-a---a-i-nhad-a-en--a-e-hi-e
NDIFEFMDZCDMQZKCEYFCJMYRNCWJCSZREXCHZUNMXZ

he-a-n-----in-i----ed---e---e-ineandhe-e--
NZUCDRJXYYSMRTMEYIFZWDYVZVYFZUMRZCRWNZDZJJ

-ed-a--inhi--hai--a-e-i--ed-----a-d--he--n
XZWGCHSMRNMDHNCMFQCHZJMXJZWIEJYUCFWDJNZDIR
```

代换密码频率分析

- 确定J
  - 密文中JN出现两次，对应的明文-
    - th出现频率高
    - 猜测S(t)=J
- 确定Y
  - 密文中JY出现一次，对应明文t-
  - 猜测S(o)=Y

th 1.52%	en 0.55%	ng 0.18%
he 1.28%	ed 0.53%	of 0.16%
in 0.94%	to 0.52%	al 0.09%
er 0.94%	it 0.50%	de 0.09%
an 0.82%	ou 0.50%	se 0.08%
re 0.68%	ea 0.47%	le 0.08%
nd 0.63%	hi 0.46%	sa 0.06%
at 0.59%	is 0.46%	si 0.05%
on 0.57%	or 0.43%	ar 0.04%
nt 0.56%	ti 0.34%	ve 0.04%
ha 0.56%	as 0.33%	ra 0.04%
es 0.56%	te 0.27%	ld 0.02%
st 0.55%	et 0.19%	ur 0.02%

代换密码频率分析

- 确定D
  - MD出现4次，i-
  - D可能对应n,t,s
  - 猜测S（s)=D
- 确定H
  - 密文HNCMF
  - 明文chai-
  - 猜测S(r)=F

th 1.52%	en 0.55%	ng 0.18%
he 1.28%	ed 0.53%	of 0.16%
in 0.94%	to 0.52%	al 0.09%
er 0.94%	it 0.50%	de 0.09%
an 0.82%	ou 0.50%	se 0.08%
re 0.68%	ea 0.47%	le 0.08%
nd 0.63%	hi 0.46%	sa 0.06%
at 0.59%	is 0.46%	si 0.05%
on 0.57%	or 0.43%	ar 0.04%
nt 0.56%	ti 0.34%	ve 0.04%
ha 0.56%	as 0.33%	ra 0.04%
es 0.56%	te 0.27%	ld 0.02%
st 0.55%	et 0.19%	ur 0.02%

代换密码频率分析

```
o-r-riend-ro--arise-a-inedhise--t---ass-it
YIFQFMZRWQFYVECFMDZPCVMRZWNMDZVEJBTXCDDUMJ

hs-r-riseasi-e-a-orationhadta-en--ace-hi-e
NDIFEFMDZCDMQZKCEYFCJMYRNCWJCSZREXCHZUNMXZ

he-asnt-oo-in-i-o-redso-e-ore-ineandhesett
NZUCDRJXYYSMRTMEYIFZWDYVZVYFZUMRZCRWNZDZJJ

-ed-ac-inhischair-aceti-ted--to-ardsthes-n
XZWGCHSMRNMDHNCMFQCHZJMXJZWIEJYUCFWDJNZDIR
```

代换密码频率分析

- 进一步猜测，得到密文

Our friend from Paris examined his empty glass with surprise, as if evaporation had taken place while he wasn't looking. I poured some more wine and he settled back in his chair, face tilted up towards the sun.