第二章 古典密码

基于字符的密码

- 代替密码（**substitution cipher)**：就是明文中的每一个字符被替换成密文中的另一个字符。接收者对密文做反向替换就可以恢复出明文。
- 置换密码**(permutation cipher)**，又称换位密码（**transposition cipher)**：明文的字母保持相同，但顺序被打乱了。

代替密码

- 简单代替密码（**simple substitution cipher),**又称单字母密码（**monoalphabetic cipher):**明文的一个字母用相应的一个密文字符代替。

- 多字母密码（**ployalphabetic cipher)**：明文空间中的字符映射到密文空间的字符。

注：关于代替密码，不做专门的讲解。仅讲它的特例，既代换密码。

字母代换密码

- 单表代换密码
  移位（**shift** ）密码、乘数（**multiplicative)**密码
- 复合代换密码
  仿射（**affine)** 密码**—HiLL**密码
  密钥短语（**Key Word)**密码**-Playfair**密码
- 多表代换密码
  维吉尼亚（**Vigenere)**密码
  博福特（**Beaufort**）密码
  滚动密钥**(running-key)**密码
  弗纳姆 **(Vernam)**密码
  转子机**(rotor machine)**

预备知识-模**q**的算术：

- 同余：
  给定任意整数**a**和**q**，以**q**除**a**，余数是**r**，则可以表示为**a=sq+r,0≤r<q**，其中**s=[a/q],**表示小于**a/q**的最大整数。定义**r**为**a mod q**的剩余，记为**r≡a mod q.**
  若整数**a**和**b**有**(a mod q)=(b mod q)**，则称**a**与**b**在**mod q**下同余。
  对于满足**{r}={a|a=sq+r,s∈Z}**的整数集称为同余类。
  模运算有下述性质：

（**1**）若**n|(a-b)**，则**a≡b mod q**

（**2**）**(a mod q)=(b mod q)**意味**a≡b mod q**

（**3**） **a≡b mod q**等价于**b≡a mod q**

（**4**）若**a≡b mod q**且**b≡c mod q**，则**a≡c mod q**

- 模算术（**Modular Arithmatic)**

在**mod q**的**q**个剩余类集**{0，1，2，…,q-1}**上可以定义加法和乘法运算如下**:**

加法**: (a mod q)+ (b mod q)= (a+b) mod q**

乘法**:(a mod q) × (b mod q)= (a × b) mod q**

单表代换密码:

移位密码

- 设P=C=K=Z/(26),对k∈ K,定义$e_k(x)=x+k \pmod{26}=y \in C$

  同时$d_k(y)=y-k \pmod{26}$

注1*：26个英文字母与模26剩余类集合{0,….,25}建立一一对应：

2*.当k=3时，为Caesar密码

abcdefghijklmnopqrstuvwxyz

DEFGHIJKLMNOPQRSTUVWXYZABC

例子: cipher => FLSKHU

实际算法为： $\forall x \in P$ 有 $e_3(x) = x + 3 \pmod{26} = y$

同时有，$d_3(y)=y-3 \pmod{26}$

移位密码分析：

- 给定加密的消息：

  PHHW PH DIWHU WKH WRJD SDUWB

由于（１）加解密算法已知

（２）可能尝试的密钥只有２**6**个

通过强力攻击得到明文：

meet me after the toga party.

移位密码很容易受到唯密文攻击。

复合密码算法

- 加密函数取形式为

  $e(x)=ax \pmod{26},\ a\in Z/(26)$

  要求唯一解的充要条件是gcd(a,26)=1它称之为乘数密码算法。该算法描述为：

  设$P=C=Z/(26)$， $K=\{a\in Z/(26)\ |\gcd(a,26)=1\}$,

  对$k=a\in K$,

  定义 $e_k(x)=ax \pmod{26}$和$d_k(y)=a^{-1}(y) \pmod{26}$

  $x,y\in Z/(26)$

  例子: a=9,

  ABCDEFGHIJKLMNOPQRSTUVWXYZ

  AJSBKTCLUDMVENWFOXGPYHQZIR

  明文 密文

  cipher => SUFLKX

对于乘数密码，当且仅当**a**与２６互素时，加密变换才是一一映射的，因此**a**的选择有１１种：

a=3,5,7,9,11,15,17,19,21,23,25

可能尝试的密钥只有**1**１个

- 加密函数取形式为

  $e(x)=ax+b \pmod{26}$, $a,b \in Z/(26)$，它称之为仿射加密算法；要求唯一解的充要条件是**$\gcd(a,26)=1$**。

  该算法描述为：

  设**$P=C=Z/(26)$**

  $K=\{(a,b) \in Z/(26) \times Z/(26) \mid \gcd(a,26)=1\}$,

  对**$k=(a,b) \in K$,**

  定义 **$e_k(x)=ax+b \pmod{26}$**和**$d_k(y)=a^{-1}(y-b) \pmod{26}$**

  $x,y \in Z/(26)$

- **q=26**时，可能的密钥是**26×12=311**个。注意！

ϕ（26）=12。

- 例子，设**k＝（7，3），注意到7^{-1}(mod 26)=15,**加密函数是$e_k(x)=7x+3$,相应的解密函数是$d_k(y)=15(y-3)=15y-19$, 易见 $d_k(e_k(x))=d_k(7x+3)=15(7x+3)-19$

$$=x+45-19$$

$$=x \pmod{26}$$

若加密明文：**hot**，首先转换字母**h,o,t**成为数字**7,14,19,**

然后加密：

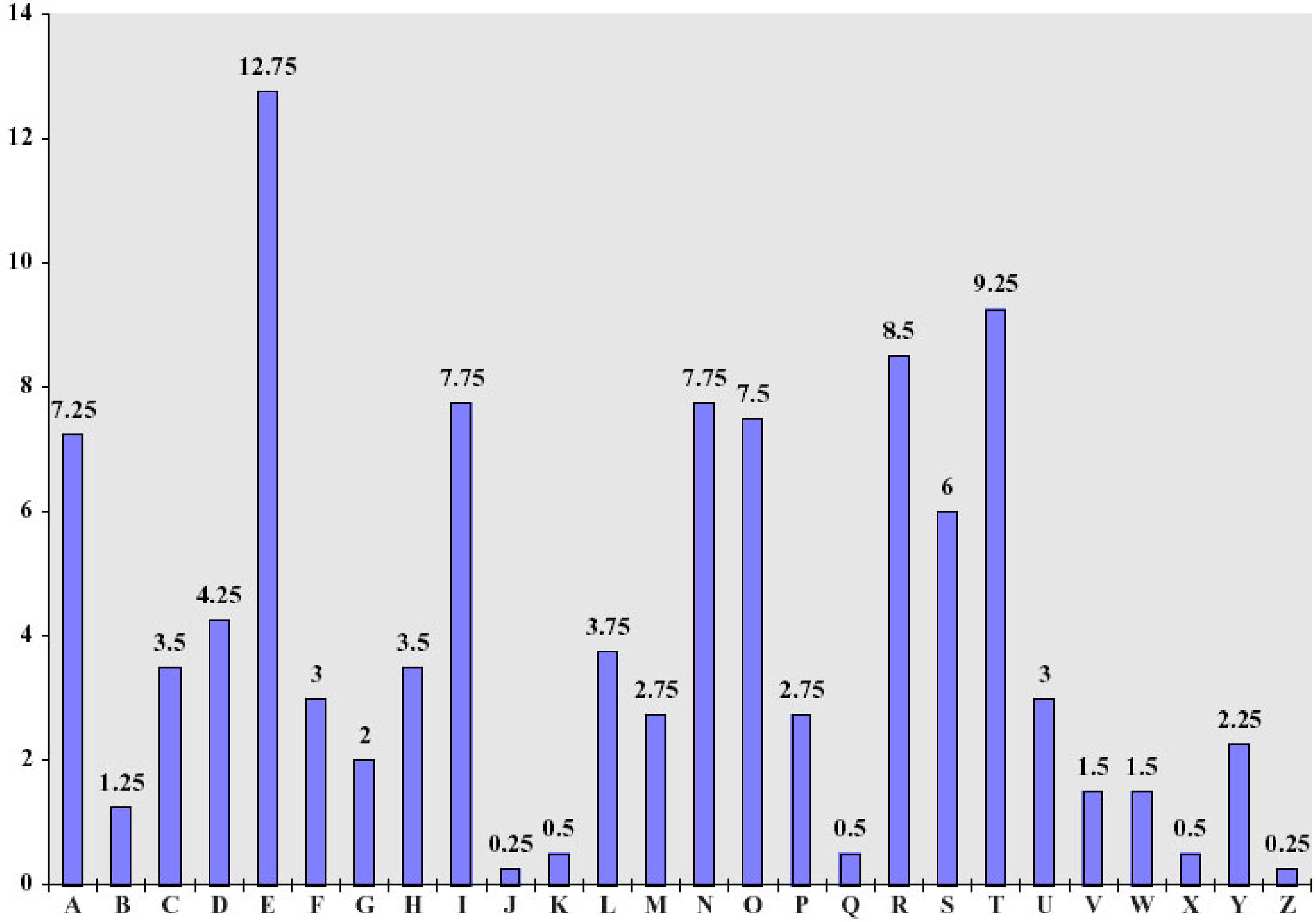
$$7\begin{bmatrix}7\\14\\19\end{bmatrix}+\begin{bmatrix}3\\3\\3\end{bmatrix}=\begin{bmatrix}0\\23\\6\end{bmatrix}=\begin{bmatrix}A\\X\\G\end{bmatrix}\pmod{26};$$

解密：

$$15\begin{bmatrix}0\\23\\6\end{bmatrix}-\begin{bmatrix}19\\19\\19\end{bmatrix}=\begin{bmatrix}7\\14\\19\end{bmatrix}$$

单表替换密码的破译

- 通过字母的使用频率破译



对抗频率分析的办法

- 多表代换密码
- 多字母代替密码

多对应替代密码

- 与简单代替密码类似，只是映射是一对多的，每个明文字母可以加密成多个密文字母。

例如，**A**可能对应于**5**、**13**、**25**

B可能对应于**7**、**9**、**31**、**42**

- 当对字母的赋值个数与字母出现频率成比例时。这是因为密文符号的相关分布会近似于平的，可以挫败频率分析。
- 然而，若明文字母的其它统计信息在密文中仍很明显时，那么同音代替密码仍然是可破译的。

多表代换密码

- 多表代换密码： 是以一系列（两个以上）代换表依此对明文消息的字母进行代换的方法。

➢ 非周期多表代换密码：

代换表是非周期的无限序列 →

一次一密密码（**one time padding):**对每个明文每次采用不同的代换表。

➢ 周期多表代换密码：代换表个数有限，重复使用。

Vigenére cipher (1858)

- 是一种多表移位代换密码

设**d**为一固定的正整数，**d**个移位代换表π=（ π_1, π_2 ,… π_d) 由密钥序列**K**＝（ $k_1,k_2,\ldots,k_d$)给定 ，第**i+td**个明文字母由表π_i决定，即密钥k_i决定

$e_k(x_{i+td})=(x_{i+td}+k_i) \bmod q = y$

$d_k(y_{i+td})= (x_{i+td}-k_i) \bmod q = x$

例子：**q=26, x=polyalphabetic cipher, K=RADIO**

明文 **x=p o l y a l p ha b e t i c c i p h e r**

密钥 **k=RAD I ORADI ORADI ORADIO**

密文 **y=GOOGOCPKTP NTLKQZPKMF**

Vigenére cipher-破译

- 依然保留了字符频率某些统计信息
- 重码分析法：间距是密钥长度整数倍的相同子串有相同密文，反过来，密文中两个相同的子串对应的密文相同的可能性很大。

```
a b c  d e f  g h i  j k l  m
000102 030405 060708 091011 12
n o p  q r s  t u v  w x y  z
131415 161718 192021 222324 25
```

密钥: *cryptographycryptographycr*

明文: **yourpackagereadyroomathree**

密文: **AFSGIOI   PG   PG**

滚动密钥密码

- 对于周期代换密码，当密钥的长度**d**和明文一样长时，就成为滚动密钥密码。

  Vigen**é**re本人建议密钥与明文一样长

若干古典密码算法的例子

随机序列密钥算法

- 1918年，Gillbert Vernam建议密钥与明文一样长并且没有统计关系的密钥内容，他采用的是二进制数据：
  加密：$C_i = P_i \oplus K_i$
  解密 $P_i = C_i \oplus K_i$

  核心：构造和消息一样长的随机密钥

Playfair密码算法

- **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。
- 5×5变换矩阵：I与J视为同一字符

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

- 加密规则:按成对字母加密

1) 相同对中的字母加分隔符(如x)
2) balloon → ba lx lo on
3) 同行取右边：he → EC
4) 同列取下边：dm → MT
5) 其他取交叉：kt → MQ　OD → TR

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(cipher)

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密码分析

- Playfair有26×26=676种字母对组合
- 字符出现概率一定程度上被均匀化
- 基于字母频率的攻击比较困难
- 依然保留了相当的结构信息

Hill密码（1929）

- 基于矩阵的线性变换:
- K是一个m*m矩阵, 在Z/(26)上可逆, 即存在K^{-1}使得:

  $KK^{-1} = I$ （在Z/(26)）

  对每一个$k\in K$,定义$e_k(x)=xK \pmod{26}$

  和 $d_k(y)=yK^{-1} \pmod{26}$

注：明文与密文都是 m元的向量（$x_1, x_2 \ldots, x_m$);($y_1, y_2,\ldots,y_m$)，Z/(26)为同余类环。在这个环上的可逆矩阵A_{mxm},是指行列式$\det A_{mxm}$的值$\in Z^*/(26)$,它为Z/(26)中全体可逆元的集合。$Z^*/(26)= \{a \in Z/(26)|(a,26)=1\}$，

$Z^*/(26)=\{1,3,5,7,9,11,15,17,19,21,23,25\}$

Hill密码的例子-i

例子：当 **m=2**时，明文元素$x=(x_1,x_2)$,密文元素$y=(y_1,y_2)$

$$(y_1,y_2)=(x_1,x_2)\begin{pmatrix}11 & 8\\ 3 & 7\end{pmatrix} \qquad K$$

若$K=\begin{pmatrix}11 & 8\\ 3 & 7\end{pmatrix}$,可得$K^{-1}=\begin{pmatrix}7 & 18\\ 23 & 11\end{pmatrix}$

若对明文**july**加密，它分成**2**个元素（**j,u),(l,y),**分别对应于（**9,20**），（**11,24**）**,**有

$$(9,20)\begin{pmatrix}11 & 8\\ 3 & 7\end{pmatrix}=(99+60,72+140)=(3,4)$$

且

$$(11,24)\begin{pmatrix}11 & 8\\ 3 & 7\end{pmatrix}=(121+72,88+168)=(11,22)$$

于是对 **july**加密的结果为**DELW**。

Hill密码的例子-ii

为了解密，**Bob**计算

$$(3,4)\begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix} = (9,20)$$

且

$$(11,22)\begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix} = (11,24)$$

因此，得到了正确的明文**“july”**

Hill密码分析

- 完全隐藏了字符(对)的频率信息
- 线性变换的安全性很脆弱，易被已知明文攻击击破。
- 对于一个mxm的hill密码，假定有m个明文-密文对，明文和密文的长度都是m. 可以把明文和密文对记为：
  $P_j=(p_{1j}, p_{2j}, ..\, p_{mj})$和$C_j=(C_{1j}, C_{2j}, \ldots, C_{mj})$，
  $C_j=KP_j,\ 1\le j \le m$
  定义mxm的方阵X=（P_{ij}） Y=(C_{ij})，得到Y=KX，$K=YX^{-1}$

例子：friday ⟹ PQCFKU

K(5 17)=(15 16) K(8 3)=(2 5) K(0 24)=(10 20)

$$\begin{pmatrix} 15 & 16 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 17 \\ 8 & 3 \end{pmatrix}\mathbf{K} \qquad \mathbf{X}^{-1} = \begin{pmatrix} 5 & 17 \\ 8 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 9 & 1 \\ 2 & 15 \end{pmatrix}$$

因此，

$$\mathbf{K} = \begin{pmatrix} 9 & 1 \\ 2 & 15 \end{pmatrix}\begin{pmatrix} 15 & 16 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 7 & 19 \\ 8 & 3 \end{pmatrix}$$