CHAPTER 3

第3章 古典密码

古典密码体制是加密和解密使用相同密钥的密码体制。虽然在现在看来，很多古典密码体制是很不安全的，但是古典密码的设计思想对现代密码的设计仍具有一定的借鉴作用。本章将介绍古典密码学的两大主要方法：置换(permutation)和代替(substitution)，并介绍几种著名的古典密码体制。

3.1 置换密码

在置换密码体制中，明文中的字或字母被重新排列，字或字母本身不变，但位置发生了改变，形成密文，又称为换位密码。从数学意义上看，置换密码中的密钥相当于一个置换函数。因为置换不会改变明文字母出现频率，所以通过比较密文中的字母频率和明文语言模型，可以检测出置换密码来。比如一条加密消息，如果单字母的出现频率与英语的一个模型匹配，而双字母频率不匹配，那么该密文就可能是使用置换密码加密的。

最简单的置换密码是采用明文倒置法，即将明文按字的顺序依次倒置，并截成固定长度的字母组，形成密文。例如，

明文：never accept failure no matter how often it visits you

密文：uoys tisi vtin etfo wohr etta mone ruli aftp ecca reve n

倒置法是简单的置换密码，经不起攻击。

攻击置换密码要求对密文字母重新排位，密码分析者通过重新编排字母，使得在密文中的字母以最高出现频率形成一些 n 字母组合。这个过程需要使用不同的 n，直到找到置换密码的换位模式。

3.2 代替密码

代替密码是把明文中的每一个字符替换成密文字母表中的另一个字符，并使用密钥 k 与之进行运算，得到密文。接收者对密文进行逆运算就可以恢复出明文。代替密码主要包括单表代替密码和多表代替密码。

3.2.1　单表代替密码

在单表代替密码中，只使用一个密文字母表，并且用密文字母表中的一个字母来代替明文字母表中的一个字母。设 A 和 B 分别为含 n 个字母的明文字母表和密文字母表：

$$A = \{a_0, a_1, \cdots, a_{n-1}\}$$

$$B = \{b_0, b_1, \cdots, b_{n-1}\}$$

单表代替密码定义了一个由 A 到 B 的一一映射 $f: A \rightarrow B$：$f(a_i)=b_i$。设明文为 $m=(m_0, m_1, \cdots, m_{n-1})$，则密文为 $c=(f(m_0), f(m_1), \cdots, f(m_{n-1}))$。下面介绍四种具体的单表代替密码体制。

1. 加法密码

加法密码的映射函数为

$$f(a_i) = b_i = a_j$$

$$j \equiv (i+k) \bmod n$$

其中，$a_i \in A$，k 是满足 $0<k<n$ 的正整数。

如果取消息空间$\mathcal{M}$、密文空间$\mathcal{C}$和密钥空间$\mathcal{K}$都为$\mathbb{Z}_q$。对任意消息 $m \in \mathcal{M}$和密钥 $k \in \mathcal{K}$，加法密码的加密算法可以表示为

$$c = E_k(m) \equiv (m+k) \bmod q$$

解密算法可以表示为

$$m = D_k(c) \equiv (c-k) \bmod q$$

如果取消息空间$\mathcal{M}$、密文空间$\mathcal{C}$和密钥空间$\mathcal{K}$都为$\mathbb{Z}_{26}$，则可以利用加法密码来加密普通英文句子，但首先需要建立英文字母与模 26 剩余之间的对应关系，如表 3.1 所示。

表 3.1　英文字母与模 26 剩余之间的对应关系

字母	a	b	c	d	e	f	g	h	i	j	k	l	m
数字	0	1	2	3	4	5	6	7	8	9	10	11	12
字母	n	o	p	q	r	s	t	u	v	w	x	y	z
数字	13	14	15	16	17	18	19	20	21	22	23	24	25

例 3.1　设 $k=3$，明文为 alice。首先将明文中的字母对应于相应的整数，即 a 对应 0、l 对应 11、i 对应 8、c 对应 2、e 对应 4。然后对每一个整数执行加密运算 $c \equiv (m+3) \bmod 26$，得

$$(0+3) \bmod 26 \equiv 3 \bmod 26$$

$$(11+3) \bmod 26 \equiv 14 \bmod 26$$

$$(8+3) \bmod 26 \equiv 11 \bmod 26$$

$$(2+3) \bmod 26 \equiv 5 \bmod 26$$

$$(4+3) \bmod 26 \equiv 7 \bmod 26$$

最后再将加密后的整数转换为相应的字母，得到密文 dolfh。如果要对密文进行解密，首先将密文中的字母对应于相应的整数，即 d 对应 3、o 对应 14、l 对应 11、f 对应 5、h 对应 7。然后对每一个整数进行解密运算 $m \equiv (c-3) \bmod 26$，得

$$(3-3) \bmod 26 \equiv 0 \bmod 26$$

$$(14-3) \bmod 26 \equiv 11 \bmod 26$$
$$(11-3) \bmod 26 \equiv 8 \bmod 26$$
$$(5-3) \bmod 26 \equiv 2 \bmod 26$$
$$(7-3) \bmod 26 \equiv 4 \bmod 26$$

最后再将解密后的整数转换为相应的字母,得到明文 alice。这就是著名的凯撒(caesar)密码。

加法密码(模 q)是不安全的,可以利用密钥穷举攻击来破译,主要原因在于密钥空间太小,只有 q 种可能的情况。

2. 乘法密码

乘法密码的映射函数为

$$f(a_i) = b_i = a_j$$
$$j \equiv ik \bmod n$$

其中,k 与 n 互素。因为仅当$(k, n)=1$ 时,k 才存在乘法逆元,才能正确解密。

如果取消息空间$\mathcal{M}$和密文空间都为$\mathbb{Z}_q$,密钥空间$\mathcal{K}$为$\mathbb{Z}_q^*$。对任意消息 $m\in\mathcal{M}$和密钥 $k\in\mathcal{K}$,乘法密码的加密算法可以表示为

$$c = E_k(m) \equiv mk \bmod q$$

解密算法可以表示为

$$m = D_k(c) \equiv ck^{-1} \bmod q$$

乘法密码(模 q)也是不安全的,密钥空间也很小,只有 $\phi(q)$种可能的情况。

3. 仿射密码

乘法密码和加法密码相结合便构成仿射密码,其映射函数为

$$f(a_i) = b_i = a_j$$
$$j \equiv (k_1 + ik_2) \bmod n$$

其中 $0<k_1<n$ 且$(k_2, n)=1$。

如果取消息空间$\mathcal{M}$和密文空间都为$\mathbb{Z}_q$,密钥空间$\mathcal{K}$为$\mathbb{Z}_q\times\mathbb{Z}_q^*$。对任意消息 $m\in\mathcal{M}$和密钥$(k_1, k_2)\in\mathcal{K}$,仿射密码的加密算法可以表示为

$$c = E_k(m) \equiv (k_1 + mk_2) \bmod q$$

解密算法可以表示为

$$m = D_k(c) \equiv (c-k_1)k_2^{-1} \bmod q$$

显然,加法密码和乘法密码都是仿射密码的特例。仿射密码的密钥空间也不大,只有 $q\phi(q)$种可能的情况。

4. 密钥短语代替密码

这种密码选用一个英文短语或者单词串作为密钥,称为密钥字或密钥短语,例如 happy new year,去掉其中的重复字母,得到一个无重复字母的字母串,即 hapynewr,把它依次写在明文字母表之下,而后再将字母表中未在字母串中出现过的字母依次写于此短语之后,就可以构造一个字母替换表,如表 3.2 所示。

表 3.2　密钥短语代替密码

明文	a	b	c	d	e	f	g	h	i	j	k	l	m
密文	h	a	p	y	n	e	w	r	b	c	d	f	g
明文	n	o	p	q	r	s	t	u	v	w	x	y	z
密文	i	j	k	l	m	o	q	s	t	u	v	x	z

当选择密钥短语代替密码和上面的密钥进行加密时，若明文为 hello，则密文为 rnffj。不同的密钥字可以得到不同的替换表，对于明文为英文单词时，密钥短语密码最多可能有 $26!=4\times10^{26}$ 个不同的替换表。

单表代替密码除了密钥空间较小外，另外一个弱点是没有将明文字母出现的频率隐藏起来，这给破译工作提供了极大的方便。因为无论是英文字母还是中文汉字，每个字母或者单字的出现频率是不同的，当统计范围足够大时，可以发现，每个字母或单字的出现频率也是比较稳定的。表 3.3 给出了英文字母出现的频率。

表 3.3　英文字母的出现频率

字母	频率	字母	频率
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

例 3.2　已知利用仿射密码加密后的密文为

```
fmdlrhrskfprhhfxrkviviizrslezykdvsprkavo
```

这些密文的频率分析见表 3.4。

表 3.4　密文中出现字母的频率

字母	频率	字母	频率
a	1	g	0
b	0	h	3
c	0	i	3
d	2	j	0
e	1	k	4
f	3	l	2

续表

字母	频率	字母	频率
m	1	t	0
n	0	u	0
o	1	v	4
p	2	w	0
q	0	x	1
r	6	y	1
s	3	z	2

虽然这里只有 40 个字母,但它足以分析仿射密码。从表 3.4 可以看出,r 出现了 6 次,k 和 v 出现了 5 次,f、h、i 和 s 出现了 3 次。根据表 3.3 知道,e 和 t 是两个出现频率最高的字母,因此可以首先猜测 r 是 e 的密文,k 是 t 的密文。根据仿射密码的加密算法得

$$17 \equiv (k_1 + 4k_2) \bmod 26$$

$$10 \equiv (k_1 + 19k_2) \bmod 26$$

这个同余式组有唯一解 $k_1=5$, $k_2=3$,从而得到了加密算法为

$$c \equiv (5 + m \times 3) \bmod 26$$

解密算法为

$$m \equiv (c-5) \times 3^{-1} \equiv (c-5) \times 9 \bmod 26$$

利用所得解密算法解密上述密文得

```
alicesentamessagetobobbyencryptionmethod
```

当然,有时候可能不会有这么幸运,需要猜测多次才能得到正确的明文。

3.2.2 多表代换密码

多表代换密码首先将明文 m 分为 n 个字母构成的分组 $\boldsymbol{m}_1, \boldsymbol{m}_2, \cdots, \boldsymbol{m}_j$,加密算法可以表示为

$$\boldsymbol{c}_i = (\boldsymbol{A}\boldsymbol{m}_i + \boldsymbol{B}) \bmod q, \quad i = 1, 2, \cdots, j$$

其中$(\boldsymbol{A}, \boldsymbol{B})$是密钥,$\boldsymbol{A}$ 是$\mathbb{Z}_q$ 上的 $n \times n$ 可逆矩阵,满足 $\gcd(|\boldsymbol{A}|, N)=1$($|\boldsymbol{A}|$是行列式)。$\boldsymbol{B}=(b_1, b_2, \cdots, b_n) \in \mathbb{Z}_q^n$,$\boldsymbol{c}_i=(y_1, y_2, \cdots, y_n) \in \mathbb{Z}_q^n$,$\boldsymbol{m}_i=(x_1, x_2, \cdots, x_n) \in \mathbb{Z}_q^n$。解密算法可以表示为

$$\boldsymbol{m}_i = \boldsymbol{A}^{-1}(\boldsymbol{c}_i - \boldsymbol{B}) \bmod q, \quad i = 1, 2, \cdots, j$$

例 3.3 设 $n=3$, $q=26$

$$\boldsymbol{A} = \begin{bmatrix} 11 & 2 & 19 \\ 5 & 23 & 25 \\ 20 & 7 & 17 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

明文为

```
your pin no is four one two six
```

将明文分成 3 个字母组成的分组

you rpi nno isf our one two six

得

$$\boldsymbol{m}_1=\begin{bmatrix}24\\14\\20\end{bmatrix},\quad \boldsymbol{m}_2=\begin{bmatrix}17\\15\\8\end{bmatrix},\quad \boldsymbol{m}_3=\begin{bmatrix}13\\13\\14\end{bmatrix},\quad \boldsymbol{m}_4=\begin{bmatrix}8\\18\\5\end{bmatrix}$$

$$\boldsymbol{m}_5=\begin{bmatrix}14\\20\\17\end{bmatrix},\quad \boldsymbol{m}_6=\begin{bmatrix}14\\13\\4\end{bmatrix},\quad \boldsymbol{m}_7=\begin{bmatrix}19\\22\\14\end{bmatrix},\quad \boldsymbol{m}_8=\begin{bmatrix}18\\8\\23\end{bmatrix}$$

执行加密运算得

$$\boldsymbol{c}_1=\boldsymbol{A}\begin{bmatrix}24\\14\\20\end{bmatrix}+\begin{bmatrix}0\\1\\2\end{bmatrix}=\begin{bmatrix}22\\7\\10\end{bmatrix},\quad \boldsymbol{c}_2=\boldsymbol{A}\begin{bmatrix}17\\15\\8\end{bmatrix}+\begin{bmatrix}0\\1\\2\end{bmatrix}=\begin{bmatrix}5\\7\\11\end{bmatrix},\quad \boldsymbol{c}_3=\boldsymbol{A}\begin{bmatrix}13\\13\\14\end{bmatrix}+\begin{bmatrix}0\\1\\2\end{bmatrix}=\begin{bmatrix}19\\13\\19\end{bmatrix}$$

$$\boldsymbol{c}_4=\boldsymbol{A}\begin{bmatrix}8\\18\\5\end{bmatrix}+\begin{bmatrix}0\\1\\2\end{bmatrix}=\begin{bmatrix}11\\8\\9\end{bmatrix},\quad \boldsymbol{c}_5=\boldsymbol{A}\begin{bmatrix}14\\20\\17\end{bmatrix}+\begin{bmatrix}0\\1\\2\end{bmatrix}=\begin{bmatrix}23\\20\\9\end{bmatrix},\quad \boldsymbol{c}_6=\boldsymbol{A}\begin{bmatrix}14\\13\\4\end{bmatrix}+\begin{bmatrix}0\\1\\2\end{bmatrix}=\begin{bmatrix}22\\2\\25\end{bmatrix}$$

$$\boldsymbol{c}_7=\boldsymbol{A}\begin{bmatrix}19\\22\\14\end{bmatrix}+\begin{bmatrix}0\\1\\2\end{bmatrix}=\begin{bmatrix}25\\16\\20\end{bmatrix},\quad \boldsymbol{c}_8=\boldsymbol{A}\begin{bmatrix}18\\8\\23\end{bmatrix}+\begin{bmatrix}0\\1\\2\end{bmatrix}=\begin{bmatrix}1\\18\\3\end{bmatrix}$$

密文为

whk fhl tnt lij xuj wcz zqu bsd

解密时,先求出

$$\boldsymbol{A}^{-1}=\begin{bmatrix}11&2&19\\5&23&25\\20&7&17\end{bmatrix}^{-1}=\begin{bmatrix}10&23&7\\15&9&22\\5&9&21\end{bmatrix}$$

执行解密运算得

$$\boldsymbol{m}_1=\boldsymbol{A}^{-1}\left(\begin{bmatrix}22\\7\\10\end{bmatrix}-\begin{bmatrix}0\\1\\2\end{bmatrix}\right)=\begin{bmatrix}24\\14\\20\end{bmatrix},\quad \boldsymbol{m}_2=\boldsymbol{A}^{-1}\left(\begin{bmatrix}5\\7\\11\end{bmatrix}-\begin{bmatrix}0\\1\\2\end{bmatrix}\right)=\begin{bmatrix}17\\15\\8\end{bmatrix}$$

$$\boldsymbol{m}_3=\boldsymbol{A}^{-1}\left(\begin{bmatrix}19\\13\\19\end{bmatrix}-\begin{bmatrix}0\\1\\2\end{bmatrix}\right)=\begin{bmatrix}13\\13\\14\end{bmatrix},\quad \boldsymbol{m}_4=\boldsymbol{A}^{-1}\left(\begin{bmatrix}11\\8\\9\end{bmatrix}-\begin{bmatrix}0\\1\\2\end{bmatrix}\right)=\begin{bmatrix}8\\18\\5\end{bmatrix}$$

$$\boldsymbol{m}_5=\boldsymbol{A}^{-1}\left(\begin{bmatrix}23\\20\\9\end{bmatrix}-\begin{bmatrix}0\\1\\2\end{bmatrix}\right)=\begin{bmatrix}14\\20\\17\end{bmatrix},\quad \boldsymbol{m}_6=A^{-1}\left(\begin{bmatrix}22\\2\\25\end{bmatrix}-\begin{bmatrix}0\\1\\2\end{bmatrix}\right)=\begin{bmatrix}14\\13\\4\end{bmatrix}$$

$$\boldsymbol{m}_7=\boldsymbol{A}^{-1}\left(\begin{bmatrix}25\\16\\20\end{bmatrix}-\begin{bmatrix}0\\1\\2\end{bmatrix}\right)=\begin{bmatrix}19\\22\\14\end{bmatrix},\quad \boldsymbol{m}_8=\boldsymbol{A}^{-1}\left(\begin{bmatrix}1\\18\\3\end{bmatrix}-\begin{bmatrix}0\\1\\2\end{bmatrix}\right)=\begin{bmatrix}18\\8\\23\end{bmatrix}$$

也就是说,明文为

you rpi nno isf our one two six

习题

1. 加法密码的加密算法为

$$c \equiv (m+5) \bmod 26$$

试对明文 data 加密,并使用解密算法

$$m \equiv (c-5) \bmod 26$$

验证加密结果。

2. 仿射密码的加密算法为

$$c \equiv (5m+7) \bmod 26$$

试对明文 uestc 加密,并使用解密算法

$$m \equiv 5^{-1}(c-7) \bmod 26$$

验证加密结果。

3. 设由仿射密码对一个明文加密得到的密文为

stqdylwqhkejyxychletqcwcqytcygtnhycftexeukejcfqtg

又已知明文的前两个字符是 un。试对该密文进行解密。

4. 在多表代换密码中

$$\boldsymbol{A} = \begin{bmatrix} 3 & 13 & 21 & 9 \\ 15 & 10 & 6 & 25 \\ 10 & 17 & 4 & 8 \\ 1 & 23 & 7 & 2 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 1 \\ 21 \\ 8 \\ 17 \end{bmatrix}$$

加密算法为

$$\boldsymbol{c}_i \equiv (\boldsymbol{A}\boldsymbol{m}_i + \boldsymbol{B}) \bmod 26$$

试对明文

cryptography is the core technology of information security

加密,并使用

$$\boldsymbol{m}_i \equiv \boldsymbol{A}^{-1}(\boldsymbol{c}_i - \boldsymbol{B}) \bmod 26$$

验证加密结果,其中

$$\boldsymbol{A}^{-1} = \begin{bmatrix} 26 & 13 & 20 & 5 \\ 0 & 10 & 11 & 0 \\ 9 & 11 & 15 & 22 \\ 9 & 22 & 6 & 25 \end{bmatrix}$$