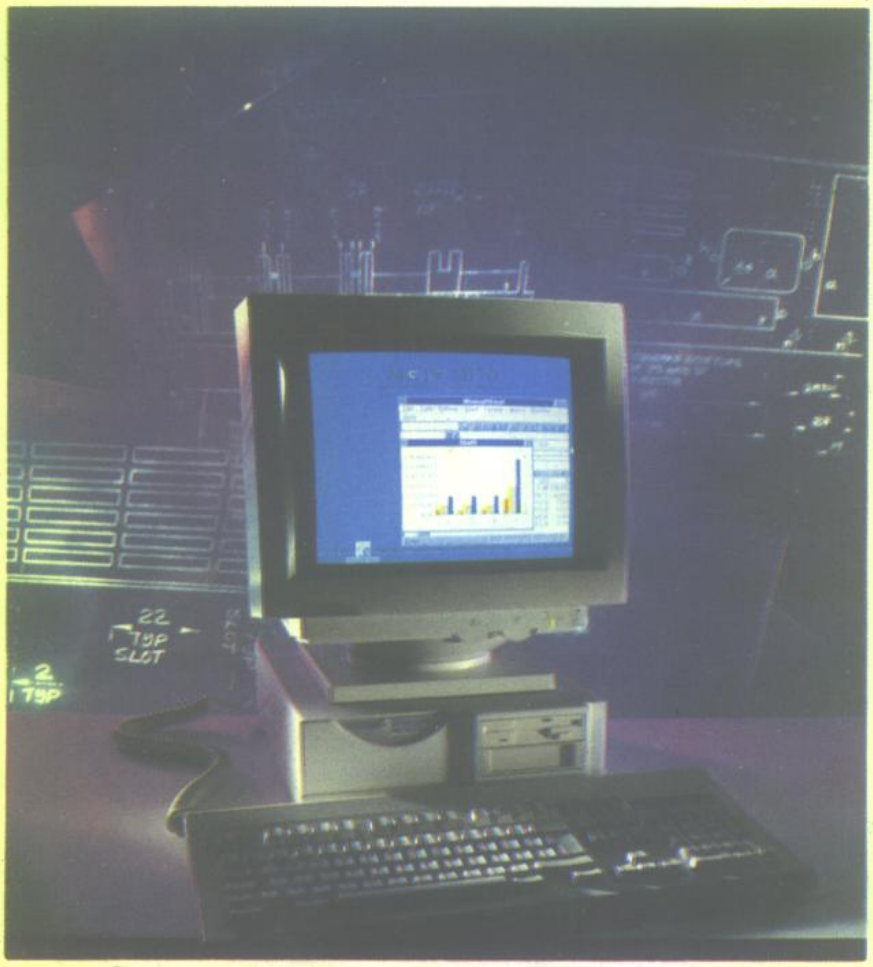
计算机基础教育系列教材

软件加密解密技术及应用

雷方桂
陈松乔 等 编著



中南工业大学出版社

软件加密解密技术及应用

雷方桂　陈松乔
吴耀斌　刘宇波　于　兵　编著

中南工业大学出版社

【湘】新登字 010 号

软件加密解密技术及应用

雷方桂　陈松乔等　编著

责任编辑：肖梓高

*

中南工业大学出版社出版发行
中南工业大学出版社印刷厂印装
新华书店总店北京发行所经销

*

开本：787×1092　1/16　印张：14　字数：340 千字
1995 年 10 月第 1 版　1995 年 10 月第 1 次印刷
印数：0001—6000

*

ISBN 7-81020-782-2/TP · 054

定价：13. 00 元

前 言

随着计算机的广泛应用和软件技术日新异月的发展，各种系统软件和应用系统如雨后春笋般地涌现并广泛应用于各个行业和部门。然而，软件保护等法规在今天却未能完全贯彻和执行，各种软件盗版仍然存在，因此，人们对于软件及数据保护的希望和要求越来越强烈。为了使各阶层计算机工作者能尽快掌握和熟悉软件加密解密技术的基本原理和方法，我们通过搜集大量资料和参考文献，结合工作中的一些体会和想法，编写了这本书。

该书从应用的角度，围绕加密解密技术讲述了磁盘工作原理，DOS 文件管理，数据、文件及磁盘的加密原理和方法，并介绍了一些典型的应用实例。该书具有一定的通用性和实用性，可作为计算机软件工作者和计算机爱好者的一本参考书和工具书，也是大中专院校相关专业学生有关课程的参考资料和手册。

全书共分八章。第一章介绍了计算机密码学的基本知识，加密解密技术的发展以及基本的加密手段。第二章讲述了磁盘驱动器的基本工作原理和测试项目及测试方法；校正磁盘的工作原理及校正磁盘用于磁盘检测和校验等方面的应用。第三章讲解了 MS-DOS 系统的文件调用功能；如何建立文件 FCB 以及有关的磁盘文件管理功能。第四章阐述了数据加密的原理和经典算法：分组密码法、DES 算法、公开密钥码和传统密码法、序列密码法，同时介绍了一种数据硬件加密／解密器（DRU）Intel 8294A 的工作原理和加密解密方法。第五章阐明了文件的加密原理和方法，包括程序和文件名加密方法，同时介绍了一种反动态跟踪程序的破译方法。第六章磁盘结构和磁盘数据组织；说明了几种典型的磁盘加密方法，包括利用非标准格式化磁盘加密、利用 CRC 校验值加密、激光加密和硬件加密等，说明了硬加密技术和软加密技术。第七章介绍了一些加密解密技术的应用实例与技巧。第八章介绍了数据的保护方法和技巧。

全书在讲述基本原理的同时，介绍了一些经典的实例和处理方法。本书内容组织由浅入深，力求通俗易懂，适合于各个层次的计算机工作者使用和参考。

由于作者水平和时间有限，书中的错误和不足之处，敬请广大读者和同仁批评指正。

作者

1995 年 3 月

目 录

1　计算机密码学

1.1　计算机与现代密码技术

在数据处理中，为了加密起见，可设法将数据进行适当的交换，使之成为“面目全非”的密码信息，这个过程通常称为数据加密，这是数据加密的一种重要手段。有关加密的各种方法不断地出现，特别是计算机用于军事、经济、信贷、储蓄、管理等各个领域后，人们对数据信息的加密要求越来越重视。与之相应地就有一个如何将加密后的信息还原成原来的数据的“本来面目”的过程，这个过程称为信息解密。很明显，如果一个加密后的信息很容易被别人找到解密的办法，那么这就是一个价值不大的加密系统，所以，研究加密系统使得加密后的信息很难被别人找到解密的办法，就形成一门新的学科——密码学。

将加密到解密的过程用图 1-1 表示如下。

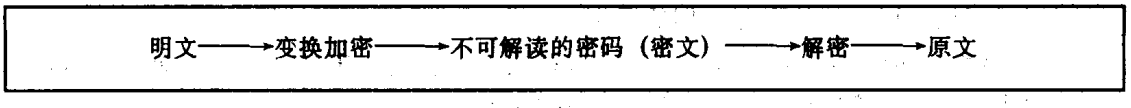


图 1-1　加密到解密的过程

上述过程所研究的原理、手段和方法，就是密码学的基本内容。

(一) 密码体制概述

1、加密的一般方法　加密的方法有很多种，以下是两种比较常用的方法。

(1) 代码法。设明文中所用到的所有词所组成的集合为 A，密码字符或数字的集合为 B。可以在 A、B 之间确定一种一一对应的关系。

$$f:\ A<=>B$$

对于每一个 $a \in A$，有：

$$f\ (a)\ =b \qquad b \in B$$

且：

$$f^{-1}\ (b)\ =a$$

加密和解密就是依靠着这种对应关系所编制的密码词典。显然，用这种方法进行加密，机动性很差，工作效率不高，如果利用电子计算机，那就要占用大量的存储空间以存放该密码词典。因此，这在实际应用中被认为是不适宜用于计算机的加密方法。

(2) 密码法。密码法通常有代替法、置换法和乘积密码法等多种。

①代替法。将明文中的每一个字母用别的字母或符号代替，从而达到加密的一种方法。

②置换法。将明文中的字母的排列顺序进行改变，从而达到加密的一种方法。

③乘积密码法。通过混合采用代替法和置换法，使明文变换成难于破译的密文的一种方法。乘积密码法是一种较可靠的加密方法。

乘积密码法是混合地使用代替法和置换法所得到的一种比较好的加密方法。用不同的混

合法所得的加密方法就形成不同的加密体制。

一般地说，加密体制就是用各种加密方法组合起来所形成的一种算法，并且通过该体制的用户所选定的单独密钥的作用来决定算法的具体实现。在这种情况下，如果整个算法也能加密的话，这对密码的用户来说，当然是再好也没有了，然而，密码体制是由电子线路、计算机程序来实现的，因此，就密码体制来说，在一般情况下是难以加密的。可见，密码体制的设计必须把注意力集中在这样一点上：使企图破译密码的人，在不知道密钥的情况下，难以实现对密文的破译。这样，保护整个数据的机密，就变成保护密钥的机密就可以了。

2、密码体制的准则 按照 C·F·Shanon 的提法，密码体制应该具备以下五个准则。

①破译密码需要极大的工作量。

②密钥的长度很小。

③加密和解密所进行的操作比较简单。

④即使产生错误，错误的扩散也很小。

⑤信息被加密后并不改变原信息的长度。

由于现代计算机的发展，算法可以组合到硬件中去，还可以用微编码和微程序，因此，在不降低容量的范围内，算法可以搞得复杂些。计算机都是二进制运算的，所以有关二进制信息的加密方法，很自然地引起人们的重视。

3、陷门函数 密码体制的要害是具有难以被攻击的特征，因此，一个密码体制总应该有某种关键所在。设计陷门函数并在密码体制中应用陷门函数，这就是一个很重要的手段。

所谓陷门函数是指具有以下性质的函数 F:

对于定义域 X，值域 Y，若:

①计算函数 F 的算法存在，即对于 $x_i \in X$，$F(x_1x_2\cdots x_n)= y$，$y \in Y$，是容易实现的。

②对于几乎所有的 $y \in Y$，要求出 x_i 所需要花费的计算时间或占用的存储空间，是不可能或很难实现的。

例如，一个大的合整数 n，分解它为素数的问题。在每次运算操作时间为 1us 的机器上，当 n= 30 位十进制数时，约花费 3.9 小时；当 n= 100 位时，需 74 年；当 n= 200 位时，要用 3.8×10^9 年。

因此，若 X 是由 100 位以下的所有的素数对组成的集合，则对于$(x_1, x_2)\in X$，若定义函数 $F(x_1, x_2)= x_1 \cdot x_2$，由于一个大的合数分解成两个素数的乘积是一个要花费很多时间的问题，所以，这里所定义的 F 便可看作是一个陷门函数。

(二) 几个具体的加密体制

加密体制有很多，这里就几个具有代表性的加密体制作些介绍。

1、DES 体制 DES 是 Data Eucryption Standard 的缩写，是美国国家标准局的数据加密标准。这个体制于 1972 初由 IBM 公司的 W·Tuchman 和 C·Meyers 研究成功，经过几年的研究和讨论于 1977 年 1 月批准为美国国家标准局的数据加密标准。后来，花了将近 17 人年的讨论也没有找到简捷的方法来攻击。如果用穷举法来攻击，即使一个微秒穷举一个密钥，共有 2^{56} 个密钥，要花费约 2283 年的时间。

DES 体制是将加密的信息按每 64 比特分成组，然后用密钥 K_i(i= 1，2，…，16)经过 16 次迭代加密运算，从而得到 64 比特的密文予以输出。

DES 体制的加密算法，如图 1–2 所示。几点说明。

①对于 f(R，K)，采用如图 1–3 所示的体制。

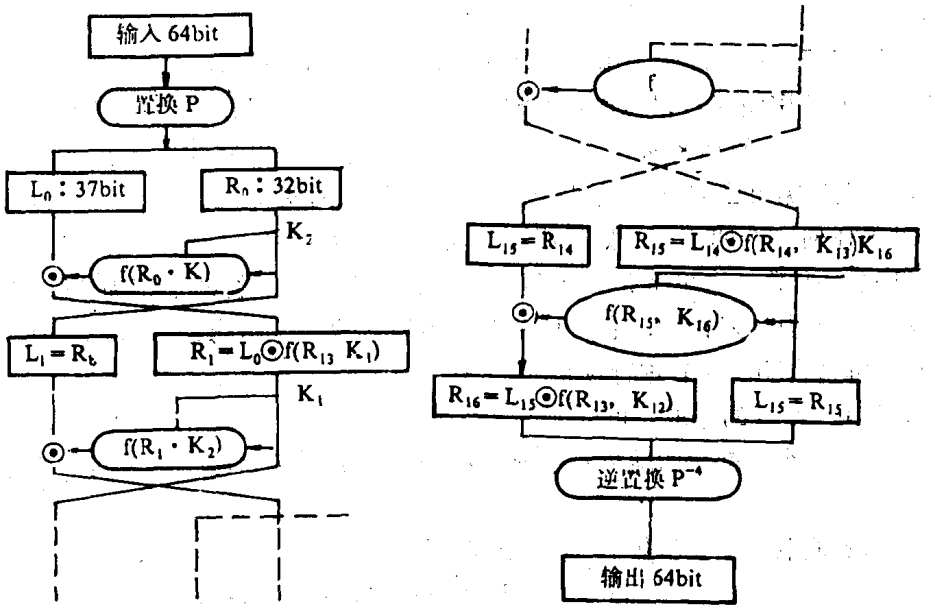


图 1–2　DES 体制加密算法示意图

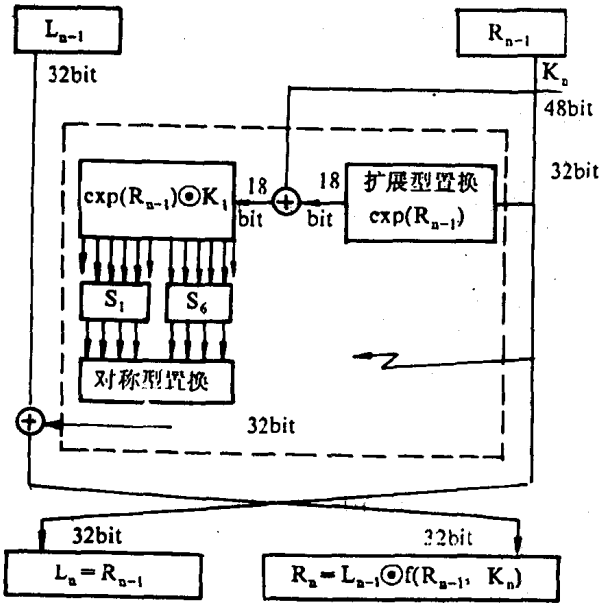


图 1–3　对于 f(R，K)采用体制示意图

②这个体制经过如此复杂的 16 轮操作，其目的就是使明文尽可能增大其混乱性和扩散性，使得输出不残存统计规律，以使破译者不能从反向推算出密钥。

③在算法中所使用的密钥 K_1—K_{16}，是用户从给定的 64bit 的主密钥经过一定的运算而产生的，并不是对用户直接给出 16 种密钥。

由于在主密钥的 64bit 中，有 8bit 是用作奇偶校验码，剩下 56bit 才作为有效密钥使用，因此，用户可选用的密钥有 2^{56} 种。

因此，即使在给定输入(明文)和输出(密文)的情况下，要想从中推出密钥 K_1，K_2，…，K_{16}，必须用穷举法去搜索 $P_{10}^{2^{56}}$ 种情况，这在实际应用中是难以完成的。尽管对这个系统有

不同看法，但是 W·Tuchman 和 C·Meyers 认为：DES 系统可以在较长的时间内保密。用计算机来破译 DES 密码，将花费极大的费用和力量，因而也将是不现实的。

2. RSA 体制 RSA 体制由 Riverst、Shamir 和 Adleman 三人提出的。它是一种公开密钥体制。

所谓"公开"就是加密密钥可以公开，而解密密钥则保密。这样，关于密钥分发问题就很方便了，公开密钥就可以像电话号码本那样地公开。譬如，A 发给 B，那么 A 只要查一下密钥本上有关 B 的加密密钥，然后，A 就可按这个公开密钥对明文进行加密后发送给 B，最后，B 就根据自己的解密密钥将密文变换成原来的明文。

对于任何一个明文 P，如果明文很长的话，可以将它分段，每一段信息 N 可以看成 0 到 n-1 的一个数，这里 n 是一个很大的合数，取数 r，将 N 加密为 M，使得 $N^r \equiv M \pmod n$，称(r，n)为加密密钥，它可以公开。可以选取一个适当的 s，使得 $M^s \equiv N \pmod n$，称(s，n)为解密密钥，将它进行保密。

对于 n，r，s 的选取，可以如下进行：

①先取两个非常大的随机素数 P、q，均为几百位且这两个数之间相差近一百，令：

$$n = P.q$$

②随机地取一个大整数 r，使它满足：

$$GCD(r, P-1) = 1, \quad GCD(r, q-1) = 1$$

③由 $s.r \equiv 1 \bmod (p-1)\cdot(q-1)$，求出整数 S。

上述办法选取的 n，r，s，由于 n 很大，在不经过 P，q 的情况下，要从 n 出发进行分解是很难的，因此，在公开(r，n)的情况下，要求得 s 是困难的。这就达到了加密的目的。

RSA 体制的具体算法描述如下。

对于给定的 g，w，u。

(1) 随机地选择奇正整数 a，使得：

$$g - 1 < \log_2 a < g + \frac{w}{2} - 1$$

(2) 对于每个质数 r< u，作：

$$GCD(r, a),\ GCD(r, 2a+1),\ GCD(a, (u-1)/2)$$

如果这三者中有一个不是 1，则重作到 (1)。

(3) 判定 a 与 2a+1 是否同时为质数，若不是则重作 (1)；

(4) 随机地选择奇正整数 b，使得：

$$g + w - 1 < \log_2 b < g + \frac{3}{2}w - 1$$

(5) 对于每个质数 r< u，作：

$$GCD(r, b),\ GCD(r, 2b+1),\ GCD(b, (u-1)/2)$$

如果这三者中有一个不是 1，则重作到 (4)。

(6) 判定 b 与 2b+1 是否同时为质数，若不是则再重做 (4)。

(7) 作如下操作。

$$GCD(a, b),\ GCD(a, 2b+1),\ GCD(2a+1, b),\ GCD(2a+1, 2b+1)$$

若其中一个不等于 1，则返回到 (1)。

(8) 解以下六对同余式。

① $A\equiv 0 \pmod{2a+1}$

$A\equiv 1 \pmod{2b+1}$

② $B\equiv 1 \pmod{2a+1}$

$B\equiv 0 \pmod{2b+1}$

③ $C\equiv 0 \pmod{2a+1}$

$C\equiv -1 \pmod{2b+1}$

④ $D\equiv -1 \pmod{2a+1}$

$D\equiv 0 \pmod{2b+1}$

⑤ $E\equiv 1 \pmod{2a+1}$

$E\equiv -1 \pmod{2b=1}$

⑥ $F\equiv -1 \pmod{2a+1}$

$F\equiv 1 \pmod{2b+1}$

当这六对同余方程的解都不是发文信息时，就转向（9），否则就转向（1）。

（9）计算。

$$P=2a=1,\quad q=2b+1,\quad m=q\cdot P,\quad v=2ab$$

（10）由 $u\cdot d\equiv 1 \pmod{v}$求出最小正整数 d。

然后，以（u，m）为加密密钥，（d，m）为解密密钥，那么很明显有：

$$(X^n)^d=X^{nd}=X^{k\cdot 2ab+1}$$

故有：

$$(X^n)d\equiv X \pmod{m}$$

为了使得所取的 m 较大，且 P，q 相差也较大，所以要求使 m 满足以下的关系式：

$$2g<\log_2 m<2g=2w$$

如果要求选取的 P、q 范围为：

$$2^g<P,\qquad q<10\times 2^g$$

则有：

$$2g<\log_2(P\cdot q)<2\log_2 10+2g$$

即：

$$2g<\log_2 m<2\log_2 10+2g$$

按上述范围取 P 和 q，那么 $w=\log_2 10=3.321\cdots$，称 w 为宽度。

如果取 P、q 为两个十进制的 100 位的质数，那么，以上述范围可求得 $g=328.870\cdots$，称 g 为规格。

为了保证加密密钥中的 u 能使下面的不等式成立：

$$X^u>m \qquad \text{(对于每一个正整数 X)}$$

只须取 $u>2g+2^w$ 即可，这是因为：

$$X^a>X^{2g+2w}>X^{\log_2 m}\geqslant 2^{\log_2 m}=m$$

因此，如果取 g=350，w=5，那么，u 应取为 $u>2g+2w=2\times 350+2\times 5=710$。

随着计算机网络的发展，公开密钥密码体制较好地解决了通信用户数量极其庞大时，密钥的分发问题，可以与数量不确定的大量用户进行加密通信，这显然是它的一大优点。然而，在具体的实施中，对于公开密钥表的使用和管理能否有可信赖的公正第三者，这也是一

个实际问题，加密通信总是在互相认识的双方，为了保守信息的机密而采用的通信方式。在这种情况下，通信双方之间或几方之间当然是经过周密而又慎重的考虑，才确定加密和解密的办法。因此，有关密钥的分发数量也是相当有限并且受到严格限制的。有关这些问题，只能在实际使用的过程中获得满意的解答。

3、渐缩体制 对于给定的正整数 a_1，a_2，…，a_n 和 S，求使得：

$$S=a_1x_1+a_2x_2+\cdots+a_nx_n$$

成立的{0，1}解(即 $x_i\in\{0,1\}$)，这就称为渐缩问题。这个问题当 n 相当大时，解的选择方式用穷举法，共有 2^n 种不同的方式。因此，对于一般的 a_i(i=1，2，…，n)来说，渐缩问题是一个 NP 问题。

当 a_1，a_2，…，a_n 满足以下的关系式时：

$$a_2>a_1$$

$$a_3>a_1+a_2$$

$$a_4>a_1+a_2+a_3$$

……

$$a_k>a_1+a_2+\cdots+a_{k-1} \qquad (1-1)$$

渐缩问题的解是容易得到的。

例如：$a_1=1$；$a_2=4$；$a_3=6$；$a_4=12$，S=19

因为：19>12，故应取 $x_4=1$；

19−12=7；7>6，故应取 $x_3=1$；

7−6=1；1<4，故应取 $x_2=0$；

1=1，故应取 $x_1=1$；

1−1=0，正好结束。

所以，解得 $x_1=1$，$x_2=0$，$x_3=1$，$x_7=1$。

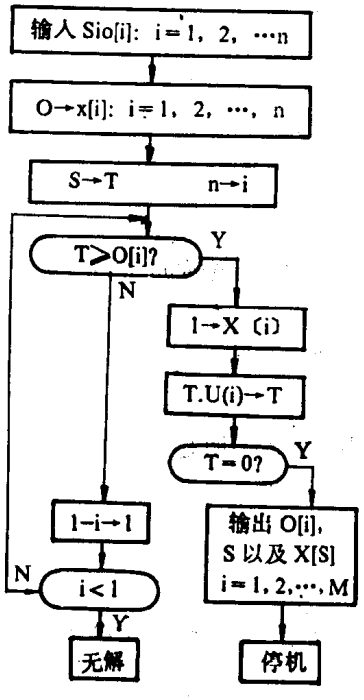


图 1-4 渐缩问题算法示意图

对于满足式 (1−1) 的渐缩问题，可以有如图 1−4 所示的算法。

在介绍过渐缩问题后，再来建立渐缩体制。

设已知 a_i=(i=1，2，…，n)满足条件式(1)，任取正整数 m 满足：

$$m>\sum_{i=1}^{n}a_i$$

取得 w，使得(w，m)=1，即 w 与 m 互质。

由 $w'\cdot w\equiv 1 \pmod m$ 解出 w'。

作：

$$a_i'\equiv a_iw_i \pmod m$$

对于已知的{0，1}信息 x_1，x_2，…，x_n，利用 a1 来进行加密，得到密文 S：

$$S\equiv\sum_{i=1}^{n}a_i'x_i \pmod m$$

接收者收到 S 后，可用以下的方法来解密：

$$S_w'\equiv S' \pmod m$$

因为：

$$S' \equiv S_w' \equiv \sum_{i=1}^{n} a_i' w' x_i \equiv \sum_{i=1}^{n} a_i' w w' x_i \equiv \sum_{i=1}^{n} a_i' x_i \pmod{m}$$

所以，就可以解出 x_1，x_2，…，x_n。

在这个体制中，(a′，m)可以公开，w 作为解密密钥。

1.2 加密系统模型

加密系统做为一种信息处理系统，它遵循控制论的“黑箱”的原理，如图 1–5 所示。

输入数据——→<黑箱>——→输出数据

图 1–5 “黑箱”原理示意图

在加密系统中，输入数据是明文信息，输出数据是密文信息。反过来，对于解密系统，输入数据是密文信息，而输出数据则是明文信息。

“黑箱”就是加密控制过程或解密控制过程。显而易见，“黑箱”是加密系统的加密关键。图 1–6 给出了最一般的加密系统方框图。

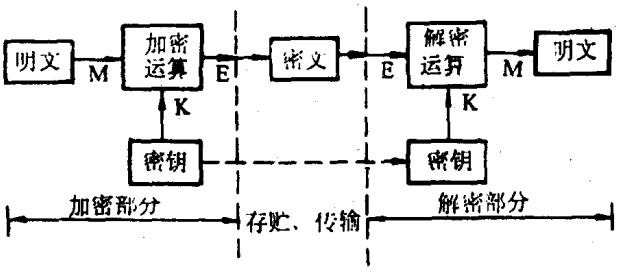


图 1–6 加密系统基本原理方框图

加密和解密系统都有两个类似的信息来源:

①明文或密文;

②密钥。系统可靠性的前提是所用的具体密钥不被解密者获得。

对明文的加密方法是用密钥 K 对明文 M 进行加密处理，得出加密信息后，再存贮或传输。密文可以说是公开的信息。但是，由于解密者不知加密所用的具体密钥 K_i 及加密算法。因而在一定时期内，解密者无法知道信息的内容，信息的收方知道具体的密钥 K_i，故可通过解密处理迅速地译出密文，得到明文。

令 M 代表明文，K 代表密钥，E 代表密文，这三者之间的关系写成一般的数学方程式为:

$$E = f\ (M,\ K) \tag{1–2}$$

即 E 是 M 和 K 的函数。

函数 f 也可能是个变换族。当 E 是一个变换族时，(1–2) 式写成:

$$E = T_i M \tag{1–3}$$

变换族 T 可以认为是密钥 K 的数学表示，变换 T_i 与明文 M 的运算得到密文 E。注角 i

表示使用的密钥的结构号数。一般说来，可供选择的密钥有有限多个。因此，密钥做为一个信息来源，对应于相应的变换 T_1，T_2，…，T_m。

当知道了 E 和 K 时，收方应能恢复密文，因此，对于变换 T_i，反变换 T_i^{-1} 必须是唯一的，也就是：

$$T_iT^{-1}=1 \tag{1-4}$$

式中，1 表示收发完全相同的变换。

这样，收端解密过程可表示为：

$$M=T_x^{-1}E \tag{1-5}$$

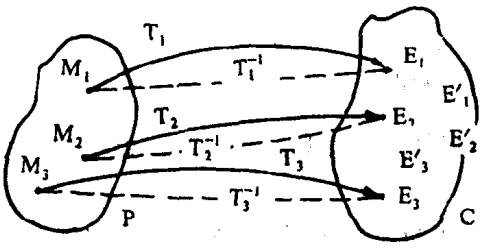


图 1-7　明文与密文的一一对应关系

对任何情况，反变换 T_j^{-1} 应能实现，并且对每一个 E 都是单值的，即明文与密文都是一一对应的关系。图 1-7 表示了这种一一对应关系。

从数学概念上看，明文构成了一个集合 P，密文构成了一个集合 C。

以上讨论可以用函数的概念，以数学术语来描述，其定义如下：

①集合 P 叫做函数的定义域；

②集合 C 叫做函数的值域；

③变换 T 将 P 中的每一个元素 M_i 与 C 中的一个元素 E_i 联系起来。

因而，如果说（P，C，T）是个函数，则 T 是从 P 到 C 的函数。如果 M_i 是 P 中的元素，并且 E_i 是 C 中在变换 T 下与之相对应的元素，则可写成为：

$$T\ (M_i)\ =E_i \tag{1-6}$$

P 中全部 T（M_i）的集，也可以用 T(P)表示，定义 T 的值域。因而值域 T(P)是 C 的子集。在图 1-7 中，T 的定义域是$\{M_1,\ M_2,\ M_3\}$，值域是元素集$\{E_1,\ E_3,\ E_4\}$。

注意，函数 f：P→C 称为一对一的函数，则要求 P 中两个不同的元素不能对应于 C 中的同一元素，这就是说，无论何时，对于 P 中的 M_i 与 M_j，当 $M_i\neq M_j$ 时，就意味着 $f(M_i)\neq f(M_j)$。与之等价的命题是：若 f 是一对一的函数，如果 $f(M_i)=f(M_j)$，则有 $M_i=M_j$。由于只有在每一个密文元素对应一个，而且仅对应一个明文元素时，明文才能被正确还原。否则，在解密时还原出来的明文可能多于一个，这样就会在解密过程中引起二义性。

因此，在集合论原理中，可以确定如下：

从明文元素集(P)到密文函数集(C)可能的一对一函数的数目：第一个明文元素可以变换成|C|个元素中的任一个；第二个明文元素可以变换成|C|−1 个元素中的任一个；依此类推，最后一个明文元素就可以映射到（|C|−|P|+1）个元素中的任一个。所以，一对一函数的总数等于 C 中可用于每一个明文元素相对应元素的数目的乘积，即：

$$|C|\cdot\ (|C|-1)\ \cdot\cdots\cdot\ (|C|-|P|+1=\frac{|C|!}{(|C|-|P|)\ !}) \tag{1-7}$$

其中，|C|表示 C 中元素的总数。$n!=1\cdot2\cdot3\cdot\cdots\cdot n$（称为 n 阶乘）。

在图 1-7 所示的例子中，|P|=3，|C|=6，可能的一对一函数有：

$$\frac{|C|!}{(|C|-|P|)\ !}=\frac{6!}{(6-3)\ !}=120\text{个}$$

图中只示出了其中的一个函数。

如果集合 S 表示从 P 到 C 可能的一对一函数集合，则有|S|个这样的函数，其中的任一个都作为密钥算法的侯选者，选定一个密钥就相当于选出一个这样的函数，密钥可以看成一个密码变量，用符号 K_i 表示一个密钥，符号 V 表示一个密钥集合，因此，可能的密钥总数 r=|V|，集合 V 可表示为:

$$V=\{K_1,\ K_2,\ \cdots,\ K_r\} \tag{1-8}$$

设:

$$T=\{T_{v1},\ T_{v2},\ \cdots,\ T_{vr}\} \tag{1-9}$$

指定为定义这种加密过程所对应的函数的集合。并设:

$$T^{-1}=\{T_{v1}^{-1},\ T_{v2}^{-1},\ \cdots,\ T_{vr}^{-1}\} \tag{1-10}$$

指定为定义这种解密过程所对应的函数的集合。

算法就由加密 T 和解密 T^{-1} 所组成。这里 T 代表全部可能的加密函数集(或称为变换)。而 T^{-1} 代表全部可能的解密函数集，它实际是 T 的逆变换，满足(4)式。

如果能独立指定的密钥数超过或等于一对一函数的数目，即|V|≥|S|，那么即使使用不同的密钥也会出现明文与密文完全对应的情况(即使 $K_i\neq K_i$，也有 $T_{ki}=T_{ki}$)。这样的密钥叫做等值密钥，对于密钥数少于一对一函数的数目|V|<|S|等值密钥也有可能存在。在实际算法中，要证实或否定等值密钥的存在性是非常困难的。

一个能减少等值密钥可能性的好的设计原则是：保证可能性密钥的数目远小于可能的一对一函数的数目，即满足条件:

$$|V|<<|S| \tag{1-11}$$

从以上的讨论中，可以得出一个结论：破译的最好办法就是尽快找出加密方法使用的具体密钥号，然后通过密文与密钥的运算，还原明文。

设破译者每试验一个密钥，"对"与"否"各有 1/2 的可能性，如果已经找到了全部可能的密钥数目为 R，这样，破译 R 个独立密钥的平均时间 T_p 为:

$$T_p=(1/2)(R/V) \tag{1-12}$$

式中，V 为试破密钥或更换密钥的速度，当 V 的单位为次/秒，T_p 的单位为小时或年时，上式可改写成:

$$T_p=1.39*10^{-4}(R/V)\quad (小时) \tag{1-13}$$

$$T_p=1.59*10^{-8}(R/V)\quad (年) \tag{1-14}$$

当 R 大到一定程度时，即使使用高速电子计算机破译，也要花很长时间，这将使破译完全失去意义。

例如，当 $R=10^{20}$，$V=10^8$（即破译速度为 1 亿次/秒的电子计算机）时，平均破译时间 $T_p=1.6*10^4$ 年（即 16000 年）。那么一年内有无破译的可能性呢？根据概率公式，可以得到破译概率公式为:

$$P_c=\frac{1}{2}\cdot\frac{1}{T_p}\times C \tag{1-15}$$

式中，C 为实际破译的时间，单位与 T_p 相同。

因此，对于这个例子可以算出一年的破译率:

$$P_1=\frac{1}{2\times 1.6\times 10^{10}}=3\times 10^{-5}$$

三年的破译率:

$$P_3 = \frac{1}{2 \times 1.6 \times 10^{10}} \times 3 = 9 \times 10^{-5}$$

计算结果表明，一年内破译的概率为三十万分之一，三年内破译的概率也不过为十万分之一。

对于计算机密码学，还有一类型值得注意的加密原理，称为没有密钥的加密系统。图1-8给出了实现的方框图。这一类型在微机加密，特别是硬件加密中经常运用，加密的关键在于保证加密及解密运算方法的安全。

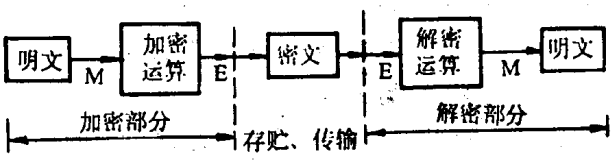


图 1-8　没有密钥的加密系统方框图

1.3　基本加密手段

所有的密码从根本上说都是代替密码。因为在任何一种密码中，明文信息总是通过给定的密钥，唯一地变换成对应的密文信息，实际上，这就构成了明文与密文之间的代替。然而，密码通常可以划分为三类：代替密码、换位密码和乘积密码。下面按照上一节所述的加密基本方程式来研究这三种典型的加密算法。

(一) 代替密码

所谓代替密码就是用一个或数个字母、数字、符号或比特来代替明文码元，并保持明文码元原来的次序，这些用来置换明文码元的字母、数字、符号或比特就构成了密文。代替密码也称为置换密码。

设明文信息 M 的表达式为:

$$M = m_1 m_2 m_3 m_4 \cdots$$

式中，$m_1 m_2 m_3 m_4 \cdots$为明文码元，变换成密文 E 为:

$$E = f(m_1)f(m_2)f(m_3)f(m_4)\cdots = e_1 e_2 e_3 e_4 \cdots$$

这里函数 $f(m_i)$存在唯一的反函数。密钥就是字符的重排方法。

最早出现的代替密码叫做"凯撒"密码。加密时，规定每个字母用其后的第三个字母置换。即明文 A 用 D 代替，B 用 E 代替，依此类推，解密时，按照加密的逆变换进行。即 D 用 A 替换，E 用 B 替换，依此类推。

(二) 换位密码

换位密码是由改变明文码元的位置而形成的，换位后的密文元仍使用明文中的码元，但其排列次序按照密钥进行换位，一般是将密文进行分组，每组再根据密钥进行换位。例如，

取分组长度为五个字符，按照 23514 进行换位。解密实际上采用了与 23514 相对应的逆变换 41253。换位运算可以两次或多次进行，这种算法称为组合换位。

（三）密码强度与解密

代替密码的算法所产生的密文，一方面保持了明文的顺序；另一方面明文中的某一码元不论其出现的位置如何，都用另外的一个特定码元进行了置换。因此，从统计的角度而言，对于明文中码元的出现概率，在密文中并没有得到改变，所以根据报文所具有的字母出现概率特性，可以相应地推断密文码元的明文值。在英文资料中，就统计的角度而言，英文单个字母出现的概率从大到小具有不同的分布次序。

因此，在有大量密文（从理论上可以算出，大约需要 100，000 个密文字符）的情况下，可以类似地统计密文中各个字母出现的概率，从统计和概率的意义上，可以认为：在密文中单个字母出现概率最大的码元就可能是 E，其次为 T，以下类推，通过这种统计推算就有可能分析出代替密码的明文码元与密文码元的对应关系，也就是可以找出代替密码的密钥。

实际上，通过统计双字母或三字母在英文材料中出现的概率，可以更快地找出代替密码的密钥（双字母统计分析只需要 300 个密文字符）。因此可以说，代替密码是一种强度不高的密码算法。

换位密码相对于代替密码，其密码强度要高一些。对换位密码的解密从原理上说，一般只有采用密钥穷举法进行。

所谓密钥穷举法，就是将一份已知的明文用试验密钥来加密并将其输出结果与已知的密文进行比较。如果是一致的，那么这个试验密钥就是未知密钥的候选者，这种密码攻击是在假定破密者已知密码算法以及他掌握着可以用于密码分析的明文及其对应的密文条件下进行的。从理论上讲，通过反复试验总是可以找到正确的密钥的，但在实际攻击中，往往由于计算量和数据存贮量过大，而无法进行下去.

针对破译者对代替密码和换位密码的攻击。提出以下两个有效对策：

①要使有效密钥长度足够长，对于换位密码要求分组的数据足够长。

②在明文信息中加入冗余字符，也就是加入“噪声”，或者叫“干扰”。目的是使输出的密文具有随机性。

注意，有一种加密算法可以被称之为“不可破译”的密码，叫做连续加密算法。

连续加密算法是将每个明文符号单独地变换成一个密文符号，其加密运算公式为：

$$L_i = m_i + K_i \quad \text{(对英文字母和间隔模 27)} \qquad (1\text{-}16)$$

式中，m_i 是明文字母序号，K_i 是随机密钥的数字号码($K_i < 26$)，L_i 是密文字母序号。

连续加密法系统的加密可靠性取决于密钥序列，如果这种密钥序列是完全随机的，那么采用这种算法的加密系统的可靠性能达到任意高。这就是所谓理论上“绝对安全”的加密系统。但是，这种算法的最大问题在于密钥序列的长度等于明文字母的长度。因此，对数据量很大的明文序列，这种加密方法是不可取的。

（四）乘积密码

将代替密码和换位密码方法结合在一起，交替地连接使用(使用一次称为一级)，就构成了多级的乘积密码。