中国计算机学会
学术著作丛书

密码分析学

冯登国

清华大学出版社
http://www.tup.tsinghua.edu.cn
广西科学技术出版社

中国计算机学会学术著作丛书

密码分析学

冯登国

清华大学出版社
广西科学技术出版社

(京)新登字 158 号
(桂)新登字 06 号

内 容 提 要

本书系统地介绍了现有的分析密码算法和密码协议的典型方法。内容主要包括:古典密码分析方法,分组密码分析方法,序列密码分析方法,公钥密码分析方法,密码协议的分析方法等。

本书可作为从事信息安全研究的科研人员的参考书,可作为信息安全专业的研究生以及相关专业的大学高年级本科生的教科书,也可供从事密码破译工作的科研人员参考。

图书在版编目(CIP)数据

密码分析学/冯登国著. —北京:清华大学出版社,2000
ISBN 7-302-03976-3

Ⅰ. 密… Ⅱ. 冯… Ⅲ. 保密编码-分析 Ⅳ. TN918.3

中国版本图书馆 CIP 数据核字(2000)第 35837 号

出版者:清华大学出版社(北京清华大学学研大厦,邮编 100084)
http://www.tup.tsinghua.edu.cn
印刷者:北京市清华园胶印厂
发行者:新华书店总店北京发行所
开 本:787×1092 1/16 印张:8.5 字数:199 千字
版 次:2000 年 8 月第 1 版 2000 年 8 月第 1 次印刷
书 号:ISBN 7-302-03976-3/TP・2326
印 数:0001~4000
定 价:14.00 元

前　言

密码学包括两部分，即密码编码学和密码分析学，这两部分既对立又统一，正是由于其对立性才促进了密码学的发展。一个密码系统的安全性只有通过对该系统抵抗当前各类攻击能力的考查和全面分析才能作出定论。密码体制的安全性分析是一个相当复杂的问题，但有一点是清楚的，那就是掌握现有的分析方法并用这些方法对相应的体制进行分析以考察其安全强度。作者有幸在中国科技大学研究生院为研究生开设了三个学期的“密码学进展”课程，在这一门课程中，主要讲解了密码分析学方面的内容，本书就是在此基础上完成的。密码分析很难讲解，它不仅涉及的知识面宽而且带有一定的实验性和经验性，任何一种破译方法只有通过实践才能真正掌握。密码分析方法很多，因为不同的算法或协议有不同的分析方法，甚至同一算法或协议有很多分析方法。本书中主要介绍了一些典型的分析方法，尤其是对分组密码和序列密码的典型分析方法作了比较详尽的介绍。在介绍公钥密码的分析方法时，主要以 RSA 体制、ElGamal 体制和背包体制的分析方法为主进行介绍（当然，一些攻击是针对体制的应用过程的）。在介绍安全协议的分析方法时，主要介绍了两种形式化分析方法，即 BAN 形式化分析方法和 Kailar 形式化分析方法，同时列举了一些具体协议的分析方法。作者这样做的目的是希望本书能起到抛砖引玉的作用，给对密码分析感兴趣的读者提供一些线索，以便于进一步阅读与研究。由于时间和水平有限，书中不足之处一定不少，请读者多提宝贵意见。作者特别感谢清华大学出版社为出版本书所做的努力。在写作过程中还得到了蔡吉人院士和裴定一教授的指点，在此对他们表示衷心的感谢。另外，也特别感谢国家重点基础研究发展规化项目（项目编号：G1999035800）的支持。

作者

2000 年 5 月 1 日

目　录

第1章 绪 论

密码学用于保护军事和外交通信可追溯到几千年前。在今天的信息时代，大量的敏感信息，如病历、法庭记录、私人财产等，常常通过公共通信设施或计算机网络来进行交换，而这些信息的秘密性和真实性是人们迫切需要的。因此，现代密码学的应用已不再局限于军事、政治和外交，其商用价值和社会价值已得到了广泛的重视。

本章主要介绍密码学中的基本概念、密码系统的理论安全性与实际安全性、Kerckhoff 假设与攻击类型和古典密码分析方法。

1.1 密码学中的基本概念

密码学是研究密码系统或通信安全的一门科学。它主要包括两个分支，即密码编码学和密码分析学。密码编码学的主要目的是寻求保证消息保密性和可认证性的方法，密码分析学的主要目的是研究加密消息的破译和消息的伪造。

采用密码技术可以隐蔽和保护需要保密的消息，使未授权者不能提取信息也不能窜改信息。被隐蔽的消息称作明文，隐蔽后的消息称作密文。将明文变换成密文的过程称作加密，其逆过程，即由密文恢复出原明文的过程称作解密。对明文进行加密操作的人员称作密码员。密码员对明文进行加密时所采用的一组规则称作加密算法，传送消息的意定对象称作接收者，他对密文进行解密时所采用的一组规则称作解密算法。加密和解密算法的操作通常都是在一组密钥控制下进行的，分别称为加密密钥和解密密钥。

根据密钥的特点，Simmons 将密码体制分为对称和非对称密码体制两种。对称密码体制又称单钥或私钥或传统密码体制，非对称密码体制又称双钥或公钥密码体制。在本书中，我们采用私钥和公钥密码体制这两个术语。在私钥密码体制中，加密密钥和解密密钥是一样的，或彼此之间容易相互确定。按加密方式又可将私钥密码体制分为序列(流)密码和分组密码两种。在序列密码中，将明文消息按字符逐位地加密。在分组密码中，将明文消息分组(每组含有多个字符)，逐组地进行加密。在公钥密码体制中，加密密钥和解密密钥不同，从一个难于推出另一个。现有的大多数公钥密码属于分组密码，只有概率加密体制属于序列密码。

在消息传输和处理系统中，除了意定的接收者外，还有非授权者。他们通过各种办法，如搭线窃听、电磁窃听、声音窃听等来窃取机密信息，称其为截收者。他们虽然不知道系统所用的密钥，但通过分析可能从截获的密文推断出原来的明文甚至密钥，这一过程称作密码分析。从事这一工作的人称作密码分析员或密码分析者。所谓一个密码是可破的，是指通过密文能够迅速地确定明文或密钥，或通过明文-密文对能够迅速地确定密钥。

在密码学中，“分析(analysis)”和“攻击(attack)”这两个术语含义相同，本书交替使用了这两个术语。

1.2 Kerckhoff 假设与攻击类型

通常假定密码分析者或敌手知道所使用的密码系统，这个假设称作 Kerckhoff 假设。当然，如果密码分析者或敌手不知道所使用的密码系统，那么破译密码更难，但是我们不应该把密码系统的安全性建立在敌手不知道所使用的密码系统这个前提之下，因此，在设计一个密码系统时，我们的目的是在 Kerckhoff 假设下达到安全性。

根据密码分析者破译时已具备的前提条件，通常人们将攻击类型分为下述四种：

（1）唯密文攻击：密码分析者有一个或多个密文。

（2）已知明文攻击：密码分析者有一些明文以及相应的密文。

（3）选择明文攻击：密码分析者有机会使用密码机，因此可选择一些明文，并产生密文。

（4）选择密文攻击：密码分析者有机会使用密码机，因此可选择一些密文，并产生明文。

上述每种攻击的目的是决定所使用的密钥。这四种攻击类型的强度按序递增，唯密文攻击是最弱的一种攻击，选择密文攻击是最强的一种攻击。如果一个密码系统能够抵抗选择密文攻击，那么它当然能够抵抗其余三种攻击。

对一个密码系统采取截获密文进行分析的这类攻击称作被动攻击。密码系统还可能遭受到的另一类攻击是主动攻击，非法入侵者主动向系统窜扰，采用删除、更改、增添、重放、伪造等手段向系统注入假消息。防止这种攻击的一种有效方法是使发送的消息具有可被验证的能力，使接收者或第三者能够识别和确认消息的真伪。实现这类功能的密码系统称作认证系统。消息的认证性和消息的保密性不同，保密性是使截获者在不知道密钥的条件下不能解读密文的内容，而认证性是使任何不知道密钥的人不能构造出一个密报，使意定的接收者解密成一个可理解的消息（合法的消息）。

一种密码攻击的复杂度可以分为两部分，即数据复杂度和处理复杂度。数据复杂度是实施该攻击所需输入的数据量；而处理复杂度是处理这些数据所需的计算量。这两部分的主要部分通常被用来刻划该攻击的复杂度。例如，在穷尽密钥搜索攻击中，所需要的数据量与计算量相比是微不足道的，因此，穷尽密钥搜索攻击的复杂度实际是处理复杂度。在 Biham 和 Shamir 的差分密码分析中，实施攻击所需的计算量相对于所需的明密文对的数量来说是比较小的，因此，差分密码分析的复杂度实际是数据复杂度。

1.3 密码系统的理论安全性与实际安全性

衡量一个密码系统的安全性有两种基本的方法，一种是实际安全性；另一种是无条件安全性，又称理论安全性。实际安全性是根据破译密码系统所需的计算量来评价其安全性的。实际安全性又分为计算安全性和可证明安全性两种。如果破译一个系统在原理上是可能的，但用所有已知的算法和现有的计算工具不可能完成所要求的计算量，就称其为计算上安全的。如果能够证明破译某体制的困难性等价于解决某个数学难题，就称其为可证

明安全的。这两种安全性虽都是从计算量来考虑，但不尽相同，计算安全要算出或估计出破译它的计算量下限，而可证明安全则要从理论上证明破译它的计算量不低于解已知难题的计算量。人们说一个密码系统是“实际上安全的”，意指利用已有的最好的方法破译该系统所需要的努力超过了敌手的破译能力（诸如时间、空间和资金等资源），或破译该系统的难度等价于解数学上的某个已知难题。当然，这只是提供了系统是计算上安全的一些证据，并没有真正证明系统是计算上安全的。理论安全性与敌手的计算能力或时间无关，也就是说敌手破译体制所做的任何努力都不会优于随机地选择来碰运气。说一个密码系统是“理论上安全的”，则具有无限计算资源（诸如时间、空间、设备和资金等）的密码分析者也无法破译该系统。

1.4 古典密码分析方法

简单的单表代换密码，如移位密码，极易破译。仅统计标出最高频度字母再与明文字母表字母对应决定出移位量，就差不多可以得到正确解了。其他如乘法密码、一般的仿射密码要复杂些，但多考虑几个密文字母统计表与明文字母统计表的匹配关系也不难解出。另外单表代换密码，如移位密码也很容易用穷举密钥搜索来破译，因为密钥量仅为 N。可见，密码系统是安全的一个必要条件是密钥空间必须足够的大，使得穷举密钥搜索破译是不可行的，但这不是一个密码系统安全的充分条件。

多表代换密码的破译要比单表代换密码的破译难得多。因为在单表代换下，字母的频度、重复字母模式、字母结合方式等统计特性除了字母名称改变以外，都未发生变化，依靠这些不变的统计特性就能破译单表代换；而在多表代换下，原来明文中的这些特性通过多个表的平均作用而被隐蔽了起来。已有事实表明，用唯密文攻击法分析单表和多表代换密码是可行的，但用唯密文攻击法分析多字母代换密码如 Hill 密码是比较困难的。分析多字母代换密码多用已知明文攻击法。本节我们以 Vigenère 密码为例来说明多表代换密码的一些分析方法。

Vigenère 密码是由法国密码学家 Blaise de Vigenère 于 1858 年提出的一种密码，它是一种以移位代换为基础的周期代换密码。d 个代换表 $f=(f_1,f_2,\cdots,f_d)$ 由 d 个字母序列给定的密钥 $k=(k_1,k_2,\cdots,k_d)\in Z_N^d$ 决定，其中 $k_i(i=1,2,\cdots,d)$ 确定明文的第 $i+td$ 个字母（t 为正整数）的移位次数，即加密公式为

$$c_{i+td}=E_{k_i}(m_{i+td})=(m_{i+td}+k_i)\bmod N \tag{1.4.1}$$

从而解密公式为

$$m_{i+td}=D_{k_i}(c_{i+td})=E_{N-k_i}(c_{i+td})=(N-k_i+m_{i+td}+k_i)\bmod N=m_{i+td} \tag{1.4.2}$$

称 k 为用户密钥（user key）或密钥字（key word）。密钥量为 N^d，当 N 与 d 较大时，密钥量是很大的。将用户密钥 k 周期地延伸就给出了整个明文加密所需的工作密钥（working key）。

我们将通过建立英文字母和模 26 的剩余之间的对应关系来使用 Vigenère 密码加密普通的英文消息（见表 1.4.1）。

表 1.4.1 英文字母和模 26 的剩余之间的对应关系

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

例 1.4.1 假定我们仍使用表 1.4.1，$d=6$，$k=$cipher，明文串是 this cryptosystem is not secure。

首先将 k 及明文串转化为数字串：$k=(2,8,15,7,4,17)$，$m=(19,7,8,18,2,17,24,15,19,14,18,24,18,19,4,12,8,18,13,14,19,18,4,2,20,17,4)$。其次模 26“加”密钥字 $k=(2,8,15,7,4,17)$得

19	7	8	18	2	17
2	8	15	7	4	17
21	15	23	25	6	8

24	15	19	14	18	24
2	8	15	7	4	17
0	23	8	21	22	15

18	19	4	12	8	18
2	8	15	7	4	17
20	1	19	19	12	9

13	14	19	18	4	2
2	8	15	7	4	17
15	22	8	25	8	19

20	17	4
2	8	15
22	25	19

最后将所得的密文数字串利用表 1.41 转化成密文字母串

VPXZGIAXIVWPUBTTMJPWIZITWZT

解密过程与加密过程类似，不同的只是进行模 26 减，而不是模 26 加。

分析 Vigenère 密码的第一步是确定密钥字的长度 d。确定密钥字的长度 d 的方法最常用的有两种，即 Kasiski 测试法和重合指数(index of coincidence)法。

Kasiski 测试法是由普鲁士军官 Kasiski 在 1863 年提出的一种重码分析法。这种方法的基本原理是：若用给定的 d 个密钥表周期地对明文字母加密，则当明文中有两个相同字母组在明文序列中间隔的字母数为 d 的倍数时，这两个明文字母组对应的密文字母组必相同。但反过来，若密文中出现两个相同的字母组，它们所对应的明文字母组未必相同，但相同的可能性很大。如果我们将密文中相同字母组找出来，并对其相间字母数综合研究，找出它们的相同字母数的最大公因子，就有可能提取出有关密钥字的长度 d 的信息。

下面我们来介绍确定 Vigenère 密码的密钥字的长度 d 的另一种方法——重合指数法。

定义 1.4.1 设 $x=x_1x_2\cdots x_n$ 是 n 个字母的一个串，x 的重合指数定义为 x 中的两个随机元素相同的概率，记为 $\mathrm{I}_c(x)$。

假定 $f_0,f_1,\cdots,f_{25}$分别表示 x 中字母 A,B,C,…,Z 出现的频率。我们能以 C_n^2 种方法选择 x 中的两个元素，对每一个 i，$0\leqslant i\leqslant 25$，x 中的两个元素都被选择为 i 的方法有 $C_{f_i}^2$ 种，因此，

$$\mathrm{I}_c(x)=\frac{\sum_{i=0}^{25}C_{f_i}^2}{C_n^2}=\frac{\sum_{i=0}^{25}f_i(f_i-1)}{n(n-1)} \tag{1.4.3}$$

现在假定 x 是一个英文明文串。记字母 A,B,…,Z 出现的期望概率分别为 $p_0, p_1, \cdots, p_{25}$。通过对大量的小说、杂志、报纸等的汇编统计，人们已经获得了英文的 26 个字母的概率分布的一个估计，参见表 1.4.2。

表 1.4.2　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

我们期望 $I_c(x) \approx \sum_{i=0}^{25} p_i^2 = 0.065$，这是因为两个随机元素都是 A 的概率为 p_0^2，两个随机元素都是 B 的概率为 p_1^2，等等。如果 x 是利用任何多表代换密码获得的一个密文，那么在这种情况下，各个概率只是被作了一个置换，但量 $\sum_{i=0}^{25} p_i^2$ 是不变的。因此对用多表代换密码获得的一个密文 x，也有 $I_c(x) \approx \sum_{i=0}^{25} p_i^2 = 0.065$。

假定 $y = y_1 y_2 \cdots y_n$ 是通过 Vigenère 密码所获得的一个密文串。把 y 分成 d 个长为 n/d 的子串，记为 $Y_1, Y_2, \cdots, Y_d$。如果 d 的确是密钥字长度，那么每个 $I_c(Y_i)(1 \leqslant i \leqslant d)$ 都将大概等于 0.065。如果 d 不是密钥字的长度，那么子串 Y_i 将看起来更随机些，因为它们将是采用不同的密钥移位加密获得的。而对一个完全随机的串 x，$I_c(x) \approx 26\left(\frac{1}{26}\right)^2 = \frac{1}{26} = 0.038$。因为两个值 0.065 和 0.038 间隔的充分远，所以我们通常能确定正确的密钥字的长度。

例 1.4.2　假定有一段来自用 Vigenère 密码加密的密文：

CHREEVOAHMAERATBIAXXWTNXBEEOPHBSBQMQEQERBWRVXUOAKXAOS
XXWEAHBWQIMMQMNKGRFVGXWTRZXWIAXLXFPSKAUTEMNDCMGTSXM
XBTUIADNGMGPSRELXNJELXVRVPTULHDNQWTWDTYGBPHXTFALJHASVB

FXNGLLCHRZBWELEKMSJIKNBHWRJGNMGJSGLXFEYPHAGNRBIEQJTAMRV
LCRREMNDGLXRRIMGNSNRWCHRQHAEYEVTAQEBBIPEEWEVKAKOEWADR
EMXMTBHHCHRTKDNVRICHRCLQOHPWQAIIWXNRMGWOIIFKEE。

首先我们用 Kasiski 测试法确定密钥字的长度。密文串 CHR 在密文中的四处出现，起始位置在 1,166,236 和 286。从第 1 个到其他 3 个的距离分别为 165,235 和 285,这三个数的最大公因子是 5,所以 5 很可能是密钥字的长度。

其次我们计算一下重合指数，看是否与 Kasiski 测试法所得的结果一致。当 $d=1$ 时，重合指数为 0.045;当 $d=2$ 时,两个重合指数分别为 0.046 和 0.041;当 $d=3$ 时,三个重合指数分别为 0.043,0.050 和 0.047;当 $d=4$ 时,四个重合指数分别为 0.042,0.039,0.046 和0.040;当 $d=5$ 时,五个重合指数分别为 0.063,0.068,0.069,0.061 和 0.072,这也为密钥字的长度是 5 提供了有力的证据。

分析 Vigenère 密码的第二步是确定密钥字(密钥字的长度已由第一步确定)。通常采用重合互指数(mutual index of coincidence)法来确定密钥字,重合互指数的精确定义如下。

定义 1.4.2 假定 $x=x_1x_2\cdots x_n$ 和 $y=y_1y_2\cdots y_{n'}$ 分别是长为 n 和 n' 的字母串。x 和 y 的重合互指数定义为 x 的一个随机元素等于 y 的一个随机元素的概率,记为 $\mathrm{MI}_c(x,y)$。

如果我们将 x 和 y 中的字母 A,B,C,…,Z 出现的频率分别表示为 $f_0,f_1,\cdots,f_{25}$ 和 $f'_0,f'_1,\cdots,f'_{25}$,那么

$$\mathrm{MI}_c(x,y)=\sum_{i=0}^{25}\frac{f_i}{n}\cdot\frac{f'_i}{n'}=\frac{\sum_{i=0}^{25}f_if'_i}{nn'} \tag{1.4.4}$$

现在假定我们已经确定了 d 的值,子串 Y_i 是通过明文的移位加密获得的,并假定 $k=(k_1,k_2,\cdots,k_d)$ 是密钥字。我们来估计 $\mathrm{MI}_c(Y_i,Y_j)$。考虑 Y_i 中的一个随机字母和 Y_j 中的一个随机字母,两个字母都是 A 的概率是 $p_{-k_i}p_{-k_j}$,两个字母都是 B 的概率是 $p_{1-k_i}p_{1-k_j}$,等等(注:所有的下标都是模 26 运算)。因此,我们可估计

$$\mathrm{MI}_c(Y_i,Y_j)\approx\sum_{h=0}^{25}p_{h-k_i}p_{h-k_j}=\sum_{h=0}^{25}p_hp_{h+k_i-k_j} \tag{1.4.5}$$

易知,$\mathrm{MI}_c(Y_i,Y_j)$的估计值只依赖于差$(k_i-k_j)\bmod 26$,我们将这个差称为 Y_i 和 Y_j 的相对移位(relative shift)。又 $\sum_{h=0}^{25}p_hp_{h+l}=\sum_{h=0}^{25}p_hp_{h-l}$,所以 MI_c 关于相对移位 l 和 $26-l$ 的估计值是一样的。

由表 1.4.3 知,当相对移位不是 0 时,这些估计值均在 0.031 到 0.045 之间,而相对移位是 0 时,估计值是0.065。我们能使用这个表来推测 Y_i 和 Y_j 的相对移位 $l=(k_i-k_j)\bmod 26$。具体做法如下:假定我们固定 Y_i,考虑由长为 d 的密钥字 $e_0=(0,0,\cdots,0)$,$e_1=(1,1,\cdots,1)$,$e_2=(2,2,\cdots,2)$,…,加密 Y_j 的影响,将加密结果分别记为 $Y_j^0,Y_j^1,Y_j^2,\cdots$。利用公式 $\mathrm{MI}_c(x,y^g)=\frac{\sum_{i=0}^{25}f_if'_{i-g}}{nn'}$ 容易计算 $\mathrm{MI}_c(Y_i,Y_j^g)$,$0\leqslant g\leqslant 25$。当 $g=l$ 时,MI_c 将接近

0.065，因为 Y_i 和 Y_j' 的相对移位是 0。然而对 $g \neq l$ 的值，MI_c 的值将在 0.031 和 0.045 之间。通过使用重合互指数这种技术，我们能获得子串 $Y_1, Y_2, \cdots$ 中的任何两个的相对移位。

表 1.4.3　期望的重合指数

相对移位	MI_c 的期望值
0	0.065
1(25)	0.039
2(24)	0.032
3(23)	0.034
4(22)	0.044
5(21)	0.033
6(20)	0.036
7(19)	0.039
8(18)	0.034
9(17)	0.034
10(16)	0.038
11(15)	0.045
12(14)	0.039
13	0.043

例 1.4.2(续)　我们已经假定了密钥字的长度是 5。现在极力计算相对移位。利用计算机不难算出 260 个值 $MI_c(Y_i, Y_j^g)$，$1 \leqslant i < j \leqslant 5$，$0 \leqslant g \leqslant 25$。这些值参见表 1.4.4。

对每对(i, j)，找出接近于 0.065 的 $MI_c(Y_i, Y_j^g)$的值。如果对一个给定的对(i, j)有唯一的一个这样的值，我们就把它视作相对移位的值(也可能不是)。在表 1.4.4 中，用方框已划出了六个这样的值。它们提供了如下的信息：Y_1 和 Y_2 的相对移位可能是 9；Y_1 和 Y_5 的相对移位可能是 16；Y_2 和 Y_3 的相对移位可能是 13；Y_2 和 Y_5 的相对移位可能是 7；Y_3 和 Y_5 的相对移位可能是 20；Y_4 和 Y_5 的相对移位可能是 11。这些关系给出了关于未知量 k_1, k_2, k_3, k_4, k_5 的如下关系式：

$$k_1 - k_2 = 9$$
$$k_1 - k_5 = 16$$
$$k_2 - k_3 = 13$$
$$k_2 - k_5 = 7$$
$$k_3 - k_5 = 20$$
$$k_4 - k_5 = 11$$

这样 $k_2 = k_1 + 17$，$k_3 = k_1 + 4$，$k_4 = k_1 + 21$，$k_5 = k_1 + 10$。所以密钥字的可能形式为$(k_1, k_1 + 17, k_1 + 4, k_1 + 21, k_1 + 10)$，$k_1 \in Z_{26}$。不难确定密钥字是 JANET(最多穷搜索 26 种可能)。

表 1.4.4

i	j	$MI_c(Y_i,Y_j')$的值								
1	2	0.028	0.027	0.028	0.034	0.039	0.037	0.026	0.025	0.052
		0.068	0.044	0.026	0.037	0.043	0.037	0.043	0.037	0.028
		0.041	0.041	0.034	0.037	0.051	0.045	0.042	0.036	
1	3	0.039	0.033	0.040	0.034	0.028	0.053	0.048	0.033	0.029
		0.056	0.050	0.045	0.039	0.040	0.036	0.037	0.032	0.027
		0.037	0.036	0.031	0.037	0.055	0.029	0.024	0.037	
1	4	0.034	0.043	0.025	0.027	0.038	0.049	0.040	0.032	0.029
		0.034	0.039	0.044	0.044	0.034	0.039	0.045	0.044	0.027
		0.055	0.047	0.032	0.027	0.039	0.037	0.039	0.035	
1	5	0.043	0.033	0.028	0.046	0.043	0.044	0.039	0.031	0.026
		0.030	0.036	0.040	0.041	0.024	0.029	0.048	0.070	0.044
		0.028	0.038	0.044	0.043	0.047	0.033	0.026	0.046	
2	3	0.046	0.048	0.041	0.032	0.036	0.035	0.036	0.030	0.024
		0.039	0.034	0.029	0.040	0.067	0.041	0.033	0.037	0.045
		0.033	0.033	0.027	0.033	0.045	0.052	0.042	0.030	
2	4	0.046	0.034	0.043	0.044	0.034	0.031	0.040	0.045	0.040
		0.048	0.044	0.033	0.024	0.028	0.042	0.039	0.026	0.034
		0.050	0.035	0.032	0.040	0.056	0.043	0.028	0.028	
2	5	0.033	0.033	0.036	0.046	0.026	0.018	0.043	0.080	0.050
		0.029	0.031	0.045	0.039	0.037	0.027	0.026	0.031	0.039
		0.040	0.037	0.041	0.032	0.051	0.032	0.034	0.030	
3	4	0.038	0.036	0.040	0.033	0.036	0.060	0.035	0.041	0.029
		0.058	0.035	0.035	0.034	0.053	0.030	0.032	0.035	0.036
		0.036	0.028	0.046	0.032	0.051	0.032	0.034	0.030	
3	5	0.035	0.034	0.034	0.036	0.030	0.043	0.043	0.050	0.025
		0.041	0.051	0.050	0.035	0.032	0.033	0.033	0.052	0.031
		0.027	0.030	0.072	0.035	0.034	0.032	0.043	0.027	
4	5	0.052	0.038	0.033	0.038	0.041	0.043	0.037	0.048	0.028
		0.028	0.036	0.061	0.033	0.033	0.032	0.052	0.034	0.027
		0.39	0.043	0.033	0.027	0.030	0.039	0..048	0.035	

将密文可解密为：The almond tree was in tentative blossom. The days were longer, often ending with magnificent evenings of corrugated pink skies. The hunting season was over, with hounds and guns put away for six months. The vineyards were busy again as the well-organized farmers treated their vines and the more lackadaisical neighbors hurried to be the pruning they should have done in November.

第 2 章　分组密码的分析方法

任何一种密码都可能有很多种分析方法，不同密码的分析方法也不尽相同。本章主要针对一些典型的分组密码介绍一些典型的攻击方法，比如强力攻击方法、差分密码分析方法及其推广、线性密码分析方法及其推广、差分-线性密码分析方法、插值攻击方法和密钥相关攻击方法等。

2.1　强力攻击

强力攻击可用于任何分组密码，且攻击的复杂度只依赖于分组长度和密钥长度，严格地讲攻击所需的时间复杂度依赖于分组密码的工作效率（包括加解密速度，密钥扩展速度以及存储空间等）。

2.1.1　穷尽密钥搜索攻击

设 k 是密钥长度（以比特为单位），在唯密文攻击下，攻击者依次试用密钥空间中所有 2^k 个密钥解密一个或多个截获的密文，直至得到一个或多个有意义的明文块。在已知（选择）明文攻击下，攻击者试用密钥空间中的所有 2^k 个密钥对一个已知明文加密，将加密结果同该明文相对应的已知密文比较，直至二者相等，然后再用其他几个已知明密文对来验证该密钥的正确性。

穷尽密钥搜索的复杂度平均为 2^{k-1} 次加密，实际上这种攻击方法适用于任何密码体制。

2.1.2　字典攻击

攻击者搜集明密文对，并把它们编排成一个“字典”。攻击者看见密文时，检查这个密文是否在字典里，如果在，他就获得了该密文相对应的明文。如果 n 是分组长度，那么字典攻击需要 2^n 个明密文对才能使攻击者在不知道密钥的情况下加解密任何消息。

2.1.3　查表攻击

设 k 是密钥长度，查表法采用选择明文攻击，其基本观点是：对一个给定的明文 x，用所有 2^k 个密钥 K（记其全体为 $\mathbf{K}$），预计算密文 $y_K = E_K(x)$。构造一张有序对表 $\{(y_K, K)\}_{K\in\mathbf{K}}$，以 y_K 给出 K 的标号。因此，对于给定的密文，攻击者只需从存储空间中找出相对应的密钥 K 即可。

2.1.4　时间-存储权衡攻击

时间-存储权衡（time-memory trade-off）攻击是一种选择明文攻击方法，它由穷尽密

钥搜索攻击和查表攻击两种方法混合而成，它在选择明文攻击中以时间换取空间。它比穷尽密钥搜索攻击的时间复杂度小，比查表攻击的空间复杂度小。我们以 DES 为例介绍时间-存储权衡攻击方法。

2.1.4.1 Feistel 型密码和 DES

Feistel 型密码是一种迭代密码，它发明于 1974 年，人们已利用这种结构设计了许多密码。现在我们来描述 Feistel 型密码的加、解密过程。

设 m 是一个长为 $n=2t$ 比特的明文组，记 $m=m_0m_1$，$m_i(i=0,1)$的长度为 t 比特。给定一个密钥 k，用它生成 r 个子密钥 $k_1,k_2,\cdots,k_r$，每一轮使用一个，共 r 轮。加密过程如下：

for $i=2$ to $r+1$ do

$m_i=m_{i-2}\oplus f(m_{i-1},k_{i-1})$；

密文为 $m_{r+1}m_r$。其中 $f(\cdot,\cdot)$是轮函数。

解密过程与加密过程类似，其解密过程如下：

for $i=r$ to 1 do

$m_{i-1}=m_{i+1}\oplus f(m_i,k_i)$；

最后获得明文 m_0m_1。

可见，在 Feistel 型密码的设计中，关键是轮函数 f 的设计。DES 就是一种 Feistel 型密码，在 DES 中，密钥长度为 56 比特，分组长度为 $n=64$ 比特，$t=32$。其加密工作程序如下：

(1) 给定一个明文 x，通过一个固定的初始置换 IP 置换 x 的比特获得 x_0，记 $x_0=\mathrm{IP}(x)=L_0R_0$，这里 L_0 是 x_0 的前 32 比特，R_0 是 x_0 的后 32 比特。

(2) 然后进行 16 轮完全相同的运算，在这里数据与密钥结合。我们根据下列规则计算 L_iR_i，$1\leqslant i\leqslant 16$：

$$L_i = R_{i-1}$$
$$R_i = L_{i-1}\oplus f(R_{i-1},k_i)$$

这里⊕表示两个比特串的异或，f 是一个函数(f 将在下面描述)，$k_1,k_2,\cdots,k_{16}$都是密钥 k 的函数，长度均为 48 比特(实际上，每一个 k_i 是来自密钥 k 的比特的一个置换选择)，$k_1,k_2,\cdots,k_{16}$构成了密钥方案。

(3) 对比特串 $R_{16}L_{16}$ 应用初始置换 IP 的逆置换 IP^{-1}，获得密文 y，即 $y=\mathrm{IP}^{-1}(R_{16}L_{16})$。注意最后一次迭代后，左边和右边未交换，而将 $R_{16}L_{16}$作为 IP^{-1}的输入，目的是为了使算法可同时用于加密和解密。

函数 $f(A,J)$的第一个变量 A 是一个长度为 32 的比特串，第二个变量 J 是一个长度为 48 的比特串，输出是一个长度为 32 的比特串，f 的计算过程如下：

(1) 将 f 的第一个变量 A 根据一个固定的扩展函数 E 扩展成一个长度为 48 的比特串。

(2) 计算 $E(A)\oplus J$，并将所得结果分成 8 个长度为 6 的比特串，记为 $B=B_1B_2B_3B_4B_5B_6B_7B_8$。