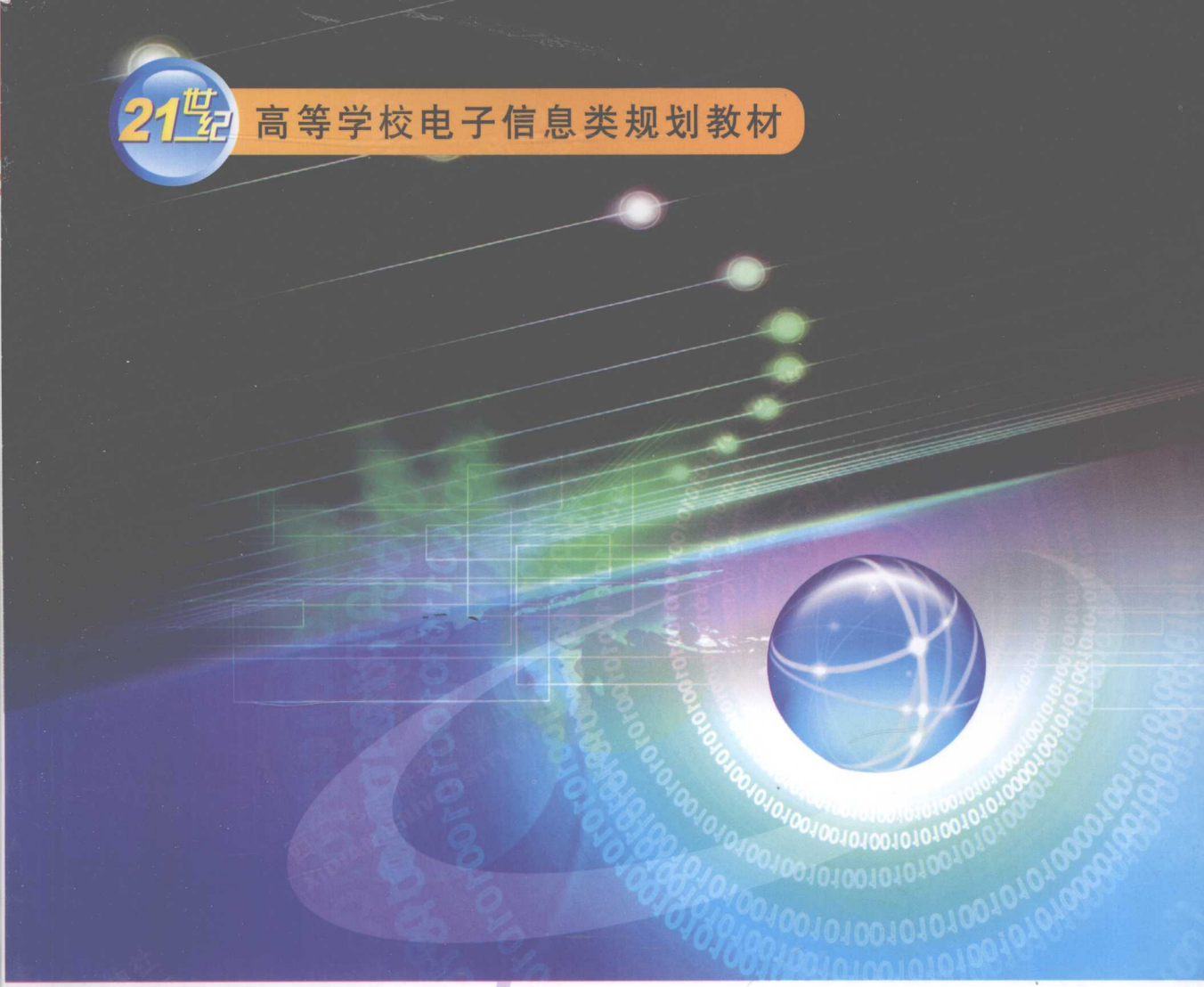

现代密码学原理与实践

于　工　牛秋娜
朱习军　刘　勇　编著
范九伦　主审

西安电子科技大学出版社
http://www.xduph.com

21 世纪高等学校电子信息类规划教材

现代密码学原理与实践

于 工 牛秋娜
朱习军 刘 勇 编著
范九伦 主审

西安电子科技大学出版社
2009

内容简介

本书全面、系统地论述了现代密码学的基本原理与方法，强调现代密码学在保密与认证两方面的功能，特别是在通信与网络安全中的重要作用。全书共七章，包括传统密码、序列密码、分组密码、公钥密码、签名与认证、密钥管理和密码协议、密码学在网络安全中的应用等内容。

作为本科教材，本书理论深度适中，强调概念和思路，偏重编码方法与应用，书中所列的程序可使学生加深对知识的理解。另外，实验的程序代码与习题答案均在附录中提供，以方便教学与自学。

本书可作为高等院校信息工程、通信工程、计算机科学与技术以及电气工程与自动化等专业"信息安全"类课程的教材，也可供相关专业科技工作人员参考。

★本书配有电子教案，有需要的老师可与出版社联系，免费提供。

图书在版编目(CIP)数据

现代密码学原理与实践/于工等编著. —西安：西安电子科技大学出版社，2009.1
21世纪高等学校电子信息类规划教材
ISBN 978-7-5606-2130-2

Ⅰ.现… Ⅱ.于… Ⅲ.密码-理论-高等学校-教材 Ⅳ.TN918.1

中国版本图书馆CIP数据核字(2008)第177640号

策　　划　薛　媛
责任编辑　马晓娟　薛　媛
出版发行　西安电子科技大学出版社(西安市太白南路2号)
电　　话　(029)88242885　88201467　　邮　编　710071
网　　址　www.xduph.com　　电子邮箱：xdupfxb001@163.com
经　　销　新华书店
印刷单位　陕西华沐印刷科技有限责任公司
版　　次　2009年1月第1版　2009年1月第1次印刷
开　　本　787毫米×1092毫米　1/16　印张14.375
字　　数　339千字
印　　数　1～4000册
定　　价　20.00元
ISBN 978-7-5606-2130-2/TN·0463
XDUP 2422001-1

前　言

21 世纪是一个高度信息化的时代，信息安全问题引起了全世界的密切关注。为适应人才培养的需求，很多理工类大学都开设了信息安全及密码学方面的课程，因而广大学生和教师迫切需要一本密码学方面的简明教程，以满足教学的需要。

现代密码学涉及的知识面较宽，包括信息理论、通信技术、检错纠错编码、计算机网络等，它用到许多数学知识，如数论、群论等。没有系统学习过这些数学理论的人往往会望而生畏，就此止步。实际上，作为初学者，只需要对有关数学知识有一些初步了解，就能理解现代密码学的大多数基本观点和基本算法，能够从中受到新思想的启迪。在本书中，作者从不同角度对相关课题作了精辟的论述，这对于教师、科研人员及研究生的学习起到了积极的作用。

本书定位于本科生教材(也可作为相关专业研究生教材)，计划课时 48 学时左右。书中内容覆盖面比较广，且篇幅适中，论述浅显易懂，不苛求数学理论的严密性，偏重编码方法与应用，适合于初学者学习；概念和思路的讲述力求准确清晰，语言力求简练清楚，以达到教师好教、学生好用的目的。为了帮助初学者入门，书中将密码理论中所必需的初等数论与有限域的相关知识精练而通俗地集中于附录 A，供缺乏这方面基础的读者预先学习，为后续学习扫清障碍。为了加深学生对内容的理解，每章都安排有与理论教学相应的实践练习，通过程序设计与计算机演练，使学生能够理论联系实际，更好地掌握所学知识。另外，每章后均有习题，以便课后练习。

全书共分七章。第 1 章介绍了几个传统密码学和传统密码分析学的典型案例，以此作为入门的向导。第 2、3 章分别对序列密码和分组密码进行了论述，主要讨论现代对称密钥体系。第 4、5、6 章对现代非对称密钥体制、签名与认证以及密钥管理和密码协议进行分题讨论，着力反映现代密码学的新思想、新发展。第 7 章介绍密码学在无线通信网络与因特网中的应用。全书重点对现代密码基本理论和方法进行了论述，为了扩大知识面，书中还介绍了 20 世纪 90 年代以来密码学发展的新成果。附录 A 是数学补充知识，简单介绍了初等数论与有限域的相关理论；附录 B 是本书实践练习使用的一些计算机程序的源代码；附录 C 是习题参考答案。

由于编者水平有限，书中难免存在错误或不当之处，欢迎读者批评指正。

编　者

2008 年 3 月于青岛科技大学

目　录

第1章 传统密码

为了经济或军事的需要，古代就出现了各种各样的加密方法。经过人类多年的研究和改进，已经取得了大量的成果。现代密码学是在借鉴与发展传统密码学的基础上诞生的。在学习现代密码学之前，了解传统密码学的基本概念和成功经验是十分必要的。本章将介绍最典型的几种传统密码算法及其破译方法。

1.1 基本概念

1.1.1 密码与密码学

密码是为解决信息安全而进行的编码。安全指通信系统抗御外来攻击的能力。外来攻击主要有两大类：一类是以截获或窃听通信内容为目标的被动攻击，攻击者截获他人信息、窃取密码、打探隐私、偷盗机密、危害民众；另一类是以篡改或伪造信息为手段的主动攻击，攻击者冒充合法发信人，发布信息，安置黑客，散布病毒，甚至破坏通信系统。针对被动攻击，密码可以使所传输信息具有保密功能，窃听者即使截获了一些信息，也会因不懂密文而不知所云。针对主动进攻，密码应具备“认证”功能，对发信人身份、消息来源以及消息完整性等加以认证，使非法发信人不得进入系统，使虚假消息能被识别，使篡改行为得以被发现。保密和认证是密码的两大基本功能。

为了军事、政治、司法等方面的需求，有时也需要破译对方的密码或者赚取对方的认证，由此便出现了与“密码编码学”原理相同但目的相反的另一分支，叫做“密码分析学”。这种保密与破译的斗争如同矛与盾的关系一样，魔高一尺，道高一丈，相依相存，相克相长，促进了二者共同的发展。近年来曾多次听到某种保密系统悬赏破译，某个防火墙产品欢迎投诉，其目的就是通过不断发现问题与解决问题，增强自己的抗攻击性能。为了设计出更加安全可靠的密码系统，设计者不仅要研究出更新更强的密码技术，还要研究对手有哪些可能的攻击手段。设计以分解大数复杂性为基础的 RSA 密码体制时，必定要讨论分解大数的技巧；设计以离散对数复杂性为基础的密码体制时，难免要讨论离散对数的计算方法，因此，聪明的密码设计者同时也应该是高明的密码分析者，二者统一于同一个目标。“密码学”则是“密码编码学”与“密码分析学”的总称。

1.1.2 传统密码学与现代密码学

密码技术的历史源远流长。很久以前，人们为了保存或传递经济、军事、外交上的重

要信息，就已经开始使用密码或类似的保密技术，发明了多种多样的加密和解密方法，其手段由手工加密发展到机械加密，直到近代的电子加密。密码技术的发展历史上，记载着不少精辟的成就，书写了丰富的经验，密码技术的成果也曾在二战中发挥了巨大作用。然而直到 20 世纪 40 年代以前，密码技术却一直是纯经验性的学问。1949 年，Shannon 发表了“保密系统的通信理论[1]”，用信息论的观点对密源、密钥、密文等概念进行了数学描述，对保密系统进行了定量化的分析，才使保密学建立在科学的理论基础之上。

我们把 20 世纪 70 年代以前的密码研究与应用称为传统密码学。那时，互联网、智能卡以及很多现代交互方式都还没有出现，传统密码学以研究秘密通信为目的，它关注的是怎样使所传输的信息不被第三者所窃取，发信人身份的认证问题并未引起足够的重视。

随着计算机互联网的出现与普及，一方面通信网络极大地方便了人们对信息的交换和共享，另一方面也同时给攻击者提供了更多的机会。如何更好地解决信息安全问题，已成为刻不容缓的任务。社会需求的推动与科研成果的积累，终于迎来了 20 世纪 70 年代以公开密钥体制为标志的密码学大发展阶段。古老的密码学重新焕发出蓬勃的生机，成为新的热门学科。密码学作为信息安全的卫士，已走出密码学家神秘的殿堂，成为广大民众和商界关注的学科。

一般认为现代密码学诞生于 20 世纪 70 年代，标志性的大事有两件[2]：一是 1977 年美国国家标准局正式公布实施了美国数据加密标准(DES)，并批准用于非机要部门和商业用途，传统密码学的神秘面纱被揭开；二是 1976 年 11 月，美国斯坦福大学电气工程系研究生 W. Diffie 和副教授 Helman 在 IEEE 上发表了题为“密码学新方向”的学术论文，公钥密码研究的序幕就此拉开。从此密码学以一种全新的理念和姿态迅速崛起，理论研究成果频出的同时，实际应用也大踏步地走进了人们的生活。

传统密码学只重视保密功能，而现代密码学除了重视保密功能之外，还对认证功能提出了多方面的要求，如发信人身份认证、信息完整性认证、不可抵赖性认证等，对于抵御主动攻击发挥了巨大的作用。数字签名是一种常见的认证手段。现代密码学还提出了公开算法的思想，开创了公开密钥体制的新思路，从而在安全观念与设计理念上明显地区别于传统密码学，独树一帜，成为当代信息安全的重要技术支柱。由于现代密码学的需要，过去某些纯数学理论的领域，如经典数论、椭圆曲线等，一下子找到了用武之地，变得火爆起来。

1.1.3 对称密钥与非对称密钥

就系统使用的密钥来分类，有两大类密钥体制。一类是对称密钥体制，加密与解密的密钥相同，掌握加密密钥的人，必然也能解密，这种密码体制也叫做单密钥体制。另一类是非对称密钥体制，加密密钥与解密密钥是不相同的，因而可以将其中一把密钥公开，只需秘密保管另一把，这种密码体制也叫做公开密钥体制或双钥体制。传统密码系统全都属于对称密钥体制，而现代密码系统中两类密钥体制都存在。比如 DES 属于对称密钥体制，RSA 则属于非对称密钥体制。

1.1.4 分组密码与序列密码

就系统加密与解密方式来分类，有分组密码与序列密码两类。分组密码是把欲加密的

文本分成一些较短的段落，每次使用相同的密钥与算法处理一段，然后将处理后的各个小段连接起来。解密过程也是按原来的分段一段段解译，然后再连接起来。序列密码方式则不进行分段，直接把整个文章顺序送入加密器，密钥也应当源源不断地输入加密器，加密器通常只是完成某种很简单的算法，比如模二加法。这种看似简单的一字一密加密方式已被 Shannon 从理论上证明是无法破译的，当然它要求的密钥是无限长的随机序列。

1.1.5 密码学基本术语[3]

保密通信过程的流程图如图 1.1 所示。

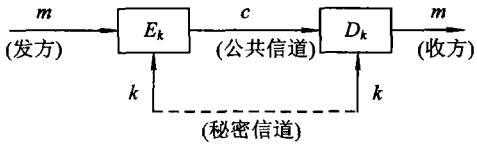

图 1.1 保密通信的流程示意

保密通信的一些基本术语对于传统密码学与现代密码学都同样有用。如：

· 明文(plaintext)：发送方(sender)未经过加密处理的信息，其内容是容易理解的。

· 密文(ciphertext)：发送方经过加密处理后的信息，文字被改变，其内容是难以理解的。

· 密钥(key)：发送方加密处理时所用的秘密信息。在传统密码学中，解密也使用同样的密钥。

· 加密(encryption，encipher)：发送方把明文加工成密文的变换，即

$$c = E_k(m) \tag{1-1}$$

· 解密(decryption，decipher)：接收方(receiver)把密文解译成明文的变换，即

$$m = D_k(c) \tag{1-2}$$

为了顺利进行上述保密通信过程，通信之前还必须完成以下准备工作：

(1) 协议(protocol)：约定通信双方的通信步骤和各个技术细节。

(2) 密钥交换(key exchange)：通信双方必须设法取得所约定的密钥(指对称密钥)。用公共信道传输密钥是不安全的，除非利用某种专门用于传送密钥的秘密信道来传输，然而其代价可能是昂贵的，或是很不方便的。可见，密钥交换问题是传统密码体制的一大难题。

1.2 传统密码举例[2]

1.2.1 逆序密码

设：明文为

m=Thousands of years ago, cryptography had been used to keep secrets of military operations or treasure.

加密算法 E_k 为将原文忽略空格并不计字母大小写，每 5 字符一组，各组取逆序，连接后得到密文：

c=suohtosdnaraeyfcogasotpyrhpargbdahysuneekotdeespeesterclimfoyratiareposno-itertroerusa

这里，k=5 是密钥。

1.2.2 棋盘密码

将 26 个字母写入 5×5 方阵中，i 和 j 占同一格，使每个字符都由一对行、列脚标表示之，如表 1.1 所示。例如，明文为

m=There are all kinds of encryption schemes in classic cryptography.

则将原文忽略空格，用表 1.1 的行、列脚标编码，可写出如表 1.2 所示的密文。

表 1.1 棋盘密码的字母排列

	1	2	3	4	5
1	a	b	c	d	e
2	f	g	h	i, j	k
3	l	m	n	o	p
4	q	r	s	t	u
5	v	w	x	y	z

表 1.2 棋盘密码对明文的加密结果

44	23	15	42	15	11	42	15	11	31
31	25	14	33	14	43	34	21	15	33
13	42	54	35	44	24	34	33	43	13
23	15	32	15	43	24	33	13	31	11
43	43	24	13	13	42	54	35	44	34
44	34	22	42	11	35	23	54		

1.2.3 凯撒密码

将 26 个英文字母按字母表序号 0～25 编号(有时也可以按 1～26 编号)，如表 1.3 所示。加密编码的方法是把明文的每个字母用向后循环移动 k 位后的字符代替：

$$c = E_k(m) = m + k \bmod 26 \tag{1-3}$$

移位距离 $k(0<k<26)$是密钥。

表 1.3 英文字母按字母表序号 0～25 编号

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
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

例如明文是：

m=Julius Caesar(朱丽叶斯·凯撒)

则不计空格与大小写，明文各字符的序号为

9 20 11 8 20 18 2 0 4 18 0 17

假设平移 k=11，则变为

20 31 22 19 31 29 13 11 15 29 11 28

循环位移意味着：

31=5 mod 26，29=3 mod 26，28=2 mod 26

于是，得到平移后的序号：

20 5 22 19 5 3 13 11 15 3 11 2

再按字母表写出，则密文为

c=UFWTFDNLPDLC

更一般的方法是不仅平移而且作“仿射变换”：

$$c = k_1 m + k_2 \bmod 26 \tag{1-4}$$

式中，k_1 和 k_2 是两个密钥，k_1 要求与 26 互素。

例如，m=information 对应的字母序号是：

8 13 5 14 17 12 0 19 8 14 13

若选 $k_1=7$，$k_2=10$，则变换后的数字是：

14 23 19 4 25 16 10 13 14 4 23

对应的密文是

c=OXTEZQKNOEX

1.2.4 单表置换密码

实际上，除了移位，还可以置换，甚至任意调换，26 个字母的各种排列共 26! 种，每指定一种调换方式，都可以写出一张与原字母表对应的置换对照表，称为单表置换。在置换表中引入密钥的方法有许多种。例如可按下述方式编制置换对照表：设密钥为 University of Science and Technology，略去重复字符，得到 UNIVERSTYOFCADHLG，排在置换对照表的最前面，后面顺序接上字母表中尚未出现过的字符，就构成了下面的列置换表：

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

1.2.5 多表置换——维吉利亚(Vigenere)密码

单表置换虽然改变了明文的模样，但密文中同一个字符总来自明文的同一字母，这样就很容易从概率分析上破译(找到概率最大的字母，就可能是 e 变来的)。为此引入多表变换，原文中的同一字母出现在不同位置时，会变换成密文的不同字符。

设密钥 k 是长度为 n 的字符串 $(k_1 k_2 \cdots k_n)$，这里 k_j 是密码的第 j 个字符在字母表的序号。明文 m 也被分成许多长为 n 的小段，第 i 段 $M_i = m_1 m_2 \cdots m_n$；那么该段明文对应的密文 $C_i = c_1 c_2 \cdots c_n$。则有

$$c_j = (m_j + k_j) \bmod 26 \qquad j = 1, 2, \cdots, n \tag{1-5}$$

例如：密钥 k=shift，其字母序号为 18，7，8，5，19；密钥长度 n=5。将明文 m=encode algorithm 忽略空格，按 n=5 分段后，其字母序号是：

4，13，2，14，3； 4，0，11，6，14； 17，8，19，7，12；

则第 1 段 5 个字符分别移位 18，7，8，5，19 后序号变为 22，20，10，19，22，即 WUKTW；第 2 段 5 个字符再分别移位 18，7，8，5，19 后序号变为 22，7，19，11，7，即 WHTLH；第 3 段 5 个字符又分别移位 18，7，8，5，19 后序号变为 9，15，1，12，5，即 JPBMF。最终的密文是 c=WUKTWWHTLHJPBMF。

为了容易地查到不同密钥字符的变换结果，可以列出一张“维吉利亚”方阵表，第一行 26 个字母按序排列。第二行是第一行左移一位的结果，B 开头，字母顺序不变，直到 Z，最后是 A 作结尾。第三行是循环左移二位 ……直到第 26 行以 Z 开头，后接 AB……XY。当

密钥字符为 a 时，就查第一行；当密钥字符为 b 时，就查第二行…… 当密钥字符为 z 时，就查最后一行。表中的列用来查询每个明文字母应当对应什么密文字母。这种方式称为多表置换。

后来，比欧福特(Beaufort)将“维吉利亚”方阵表的各行按逆序排列得到了“比欧福特方阵”，使破译更加困难。

1.2.6 维尔南(Vernam)加密算法

当明文、密文、密钥均为二元代码(0，1)时，维吉利亚密码就成了维尔南密码。维尔南密钥在计算机代码中得到广泛应用，其计算公式是：

$$c_i = (m_i + k_i) \bmod 2 \qquad i = 1, 2, 3, \cdots \tag{1-6}$$

式中，m_i、k_i、c_i 都是 0 或 1。

密钥一般是取一个周期很长的伪随机二元码序列。加密后，密钥的随机性掩盖了明文的可读性，使密文变的不可理解。为了得到较长的密钥，可以找两个不太长的密钥，各自循环着使其中一个对另一个加密，当两密钥长度互素时，可得到长度为二者之积的密钥。

1.2.7 普莱费厄(Playfair)加密算法

普莱费厄加密算法把引入密钥词组的单表密码与棋盘密码结合。例如密钥是 five stars，按单表方式排序后写成 5×5 方阵：

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

若明文为

M ＝Play fair cipher was actually invented by wheat stone

则先将明文两两分组：

pl, ay, fa, Ir, ci, ph, wa, sa, ct, ua, lq, ly, in, ve, nt, ed, by, wh, ea, ts, to, ne

分组时注意：

(1) 若分组出现两字母相同时，则在两字母之间插入一个 q。

(2) 若总字符个数为单数，则在最后那个不配对的字母后添一个 q。

加密方法如下：

(1) 若分组 $m_1 m_2$ 在方阵同行时，密文 $c_1 c_2$ 分别是 m_1 和 m_2 右面的字母，其中第一列可视为第五列的右面。

(2) 若分组 $m_1 m_2$ 处在方阵同一列时，密文 c_1 和 c_2 分别是 m_1 和 m_2 下面的字母，第一行可视为第五行的下面。

(3) 若 $m_1 m_2$ 不同行也不同列时，c_1 和 c_2 是 $m_1 m_2$ 为对角的矩阵块的另两个角，并且 c_1 与 m_1 同行，c_2 与 m_2 同行。

按此加密方法，上例中分组被加密为

QK, BW, IT, VA, AS, OK, IG, IC, TA, WT, QZ, KZ,
AW, ES, MA, FK, KE, XG, IB, CF, RM, PI

1.2.8 希尔(Hill)加密算法

将明文中长为 L 的字符分组 $\boldsymbol{m}$ 通过线性变换 $\boldsymbol{k}$，变为密文中长为 L 的字符分组 $\boldsymbol{c}$。字符分组以列矢量表示：

$$\boldsymbol{c} = \boldsymbol{k} \cdot \boldsymbol{m} \bmod 26 \tag{1-7}$$

有时，字符分组以行矢量表示，此时应采用的变换式为

$$\boldsymbol{c} = \boldsymbol{m} \cdot \boldsymbol{k} \bmod 26 \tag{1-8}$$

例如，$\boldsymbol{m}$=Hill，四个字母序号为

$$7, 8, 11, 11$$

若

$$\boldsymbol{k}=\begin{pmatrix} 8 & 6 & 9 & 5 \\ 6 & 9 & 5 & 10 \\ 5 & 8 & 4 & 9 \\ 10 & 6 & 11 & 4 \end{pmatrix}$$

则

$$\boldsymbol{c}=\boldsymbol{k}\cdot\boldsymbol{m}=\begin{pmatrix} 8 & 6 & 9 & 5 \\ 6 & 9 & 5 & 10 \\ 5 & 8 & 4 & 9 \\ 10 & 6 & 11 & 4 \end{pmatrix}\cdot\begin{pmatrix} 7 \\ 8 \\ 11 \\ 11 \end{pmatrix}=\begin{pmatrix} 24 \\ 19 \\ 8 \\ 23 \end{pmatrix} \bmod 26$$

即

$$\boldsymbol{c}=\mathrm{YTIX}$$

解密可由反变换：

$$\boldsymbol{m} = \boldsymbol{k}^{-1}\boldsymbol{c} \bmod 26 \tag{1-9}$$

来计算，$\boldsymbol{k}^{-1}$是 $\boldsymbol{k}$ 的逆矩阵，且作模 26 处理。掌握密钥的用户不难求出 $\boldsymbol{k}^{-1}$。本例中：

$$\boldsymbol{k}^{-1}=\begin{pmatrix} 23 & 20 & 5 & 1 \\ 2 & 11 & 18 & 1 \\ 2 & 20 & 6 & 25 \\ 25 & 2 & 22 & 25 \end{pmatrix}$$

1.3 密码分析举例

以破译密码为目的的学问叫密码分析学。越来越多的事实告诉我们，密码分析是密码学不可分割的组成部分。每研制出一种新密码的同时，就应当讨论它的抗攻击性能；每当某种密码被破译的同时，也就启示密码学家发现原先密码系统的漏洞和改进方法。

根据掌握数据的多少，密码分析可以分为以下几种：

(1) 唯密文攻击：仅掌握若干被同一密钥和同一加密算法得到的密文，想导出明文。

(2) 已知明文攻击：不仅掌握若干密文，还知道相应的明文，想从中找到密钥，从而对任何其他同类加密的密文都能破解。

(3) 选择明文攻击：不仅掌握密文，还有多个可供选择的明文。显然攻击力度更强了，或破译更容易了。

下面举几个破译传统密码的例子。

1.3.1　对单表置换密码的分析

单表密码系统的漏洞在于明文中相同的字符一定对应密文中同样的字符，明文中出现概率最高的字符，必然对应于密文中出现最多的字符，因此利用自然语言文字的概率统计性质容易找到明、密文字符的对应关系。下面以两个实例来展示分析过程。

对大量英语文章的统计发现，各个字母出现的概率是不相同的。对单个字母的统计结果列于表 1.4 中(概率值为 $x\%$)。

表 1.4　英文文章中各个字母出现的概率　　(%)

字母	a	b	c	d	e	f	g	h	i	j	k	l	m
概率	8.56	1.39	2.79	3.78	13.04	2.89	1.99	5.28	6.27	0.13	0.42	3.39	2.49
字母	n	o	p	q	r	s	t	u	v	w	x	y	z
概率	7.07	7.97	1.99	0.12	6.77	6.07	10.5	2.49	0.92	1.49	0.17	1.99	0.08

【例 1】 已知仿射密码的密文为[11]

FMXVEDRAPHFERBNDKRXRSREFMORUDSDKDVSHVUFEDKAPRKDLYEVLRHHRH

共 57 字。对应的明文是什么？

解：先统计各字符(特别是高频字符)出现的频度：

R(8)，D(7)，E、H、K(各 5)，F、S、V(各 4)

(1) 根据出现频率大小，不妨设 R→e，D→t，R 序号为 17，D 序号为 3，e 为 4，t 为 19。将之代入加密的仿射方程：

$$y = e_k(x) = ax + b \bmod 26 \tag{1-10}$$

得
$$\begin{cases} 4a + b \bmod 26 = 17 \\ 19a + b \bmod 26 = 3 \end{cases}$$

解得
$$\begin{cases} a = 6 \\ b = 19 \end{cases}$$

由于 $\gcd(a, 26)=2$，表明此密码不正确(正确的 a 应当与 26 互素)。

(2) 重新假设 R→e，E→t，E 的序号为 4，从而得

$$\begin{cases} 4a + b \bmod 26 = 17 \\ 19a + b \bmod 26 = 4 \end{cases}$$

解得
$$\begin{cases} a = 13 \\ b = 9 \end{cases}$$

由于 $\gcd(a, 26)=13$，因此还是不正确。

(3) 再次假设 R→e，K→t，K 序号为 10，从而得

$$\begin{cases} 4a + b \bmod 26 = 17 \\ 19a + b \bmod 26 = 10 \end{cases}$$

解得
$$\begin{cases} a = 3 \\ b = 5 \end{cases}$$

由于 $\gcd(a, 26)=1$，因此有可能是正确的结果。

(4) 由加密算法：

$$y = 3x + 5 \bmod 26$$

得到解密算法

$$x = \frac{y-5}{3} \bmod 26 = 9y - 19 \bmod 26$$

式中：$3^{-1} = 9 \bmod 26$(3 的模逆元是 9)。

(5) 用这个结果解译密文，看能否得到有意义的明文。结果是：

algorithms are quite general definitions of arithmetic processes

得到了有意义的明文，表明破译正确。

【例 2】 已知单表加密的 280 字密文为

GJXXN GGOTZ NUCOT WMOHY JTKTA MTXOB YNFGO GINUG JFNZV
QHYNG NEAJF HYOTW GOTHY NAFZN FTUIN ZBNEG NLNFU TXNXU
FNEJC INHYA ZGAEU TUCQG OGOTH JOHOA TCJXK HYNUV OCOHO
UHCNU GHHAF NUZHY NCUTW JUWNA EHYNA FOWOT UCHNP
HOGLN FQZNG FOUVC NVJHT AHNGG NTHOU CGJXY OGHYN ABNTO
TWGNT HNTXN AEBUF KNFYO HHGIU TJUCE AFHYN GACJH OATAE
IOCOH UFQXO BYNFG

如何解密以恢复明文？

解：(1) 先统计密文中各字母出现的次数：(由大到小排列)

N(36)，H(26)，O(25)，G(23)，T(22)，U(20)，F(17)，A(16)，Y(14)，C(13)，J(12)，X(9)，E(7)，Z(7)，W(6)，B(5)，I(5)，Q(5)，V(4)，K(3)，L(2)，M(2)，P(1)

英文文章中的高概率字母 e、t、a、o、n、i、r、s、h 和中概率字母 d、l、u、c、m 之间的概率值有一明显的间断。由此我们大体可确定密文中出现频度较高的 9 个字符 N、H、O、G、T、U、F、A、Y 的对应范围，并且大体能肯定 N 就是 e。

(2) 再来统计密文中各字符的前缀与后缀的出现频度。下表列出了密文中出现频度较高的 9 个字符的前后连缀关系。表格中的每组数分别表示该行、该列两字符的两种不同排序方式所出现的次数。比如 N 行 G 列交叉处的(5，4)表示 NG 出现 5 次，GN 出现 4 次。

	N	H	O	G	T	U	F	A	Y	推测
N	0，0	1，3	0，0	5，4	4，0	5，0	7，3	5，0	0，9	e
H	3，1	2，2	4，5	1，2	1，4	1，1	0，2	1，1	10，0	t
O	0，0	5，5	0，0	10，6	7，1	1，0	1，1	2，0	0，3	i
G	4，5	2，1	6，4	2，2	0，0	0，2	0，2	2，0	0，0	?
T	0，4	4，1	1，7	0，0	0，0	3，4	0，1	3，2	0，0	n
U	0，5	1，1	0，1	2，0	4，3	3，2	0，0	0，0	0，0	a
F	3，7	2，0	1，1	3，0	1，0	2，3	0，0	0，3	1，0	?
A	0，5	1，1	0，2	0，2	2，3	0，0	4，0	0，0	0，1	o
Y	9，0	0，10	3，0	0，0	0，0	0，0	0，1	0，1	0，0	h

① 由这些数据推测，O、U、A 可能是元音 i、a、o，这是因为它们互连的可能很小；OA 出现两次，OU 出现一次，其他 UO、UA、AO、AU 均无出现，所以 O→i，A→O；又由 OU 出现一次，NU 出现 5 次，UN 不出现，可猜到 U→a(因 ea 常见 ae 极少，ia 也是存在的)。

② 根据 n 是高频辅音，且 n 的前面是元音的概率高达 80%，现在 T 频度为 22，前面是 O. U. A 的占 17 次，由此可判定 T→n。

③ YN 出现 9 次，NY 不出现，已经判知 N→e，于是可判 Y→h。根据 th 常见而 ht 罕见，及 HY 出现 10 次，YH 一次也没有，知 H→t。

④ 高概率集中只剩下 r 和 s 尚看不出与密文前 9 位中剩余的 F 和 G 有何对应关系。

(3) 接下来就把 280 个字母中已找到的 159 个先写出来，空出尚未找到对应关系的字母，待猜测填空(注意：大写表示密文字符，小写表示已经找到的明文字符)。

① 由 Mith 猜出 M→w，于是由 nKnown→nknown 知 K→k；由 intJition 知 J→u。

② 将已判定的字母填入置换对照表中就有：

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
U N Y O K T A H J M

之后的 HJ. M 已呈现出字母表的顺序，于是可猜到中间空的是 L→v；而前面的 rs 应当是 FG，后面的 xy 则是 PQ；最前面既然 U→a，可猜到 V→b。

③ 再次将猜到的字母代入密文，就得到：

G J X X N G GO T Z N U C O T W M O H Y J T K T A M T X O B Y N F G
s uXXe s si n Ze a Ci n Ww i t h unk n o w n ci Bh e r s

O G I N U G J F N Z V Q H Y N G N E A J FH Y O T W
i s Ie a s ure Zb y t h e s e Eo urt h i n W

由 suXXess 可猜出 X→c，于是由 XiBhers 知 B→p，得 ciphers；还可由 Eour 猜 E→f 得 four；thinW 猜 W→g，得 thing。

④ 再次回到单表对照关系上，就知密钥词组是：NEW YORK CITY。去掉重复的 Y，后面按顺序接字母表中尚未出现的字符，就是：NEW YORK CIT ABDFGHJLMPQS。最后将 UVXZ 循环移位到前面去，得到与字母表对应的置换密钥是 UVXZNEWYORKCIT-ABDFGHJLMPQS。

(4) 按此对照表，可将密文全文译出为

Success in dealing with unknown ciphers is measured by these four things in the order named, perseverance, careful methods of analysis, intuition, luck. The ability at least to read the language of the original text is very desirable but not essential. Such is the opening sentence of Parker Hitt's Manual for the Solution of Military Ciphers.

意思是：破译一未知密码是否成功，可由以下四个因素来衡量，按顺序为：毅力、审慎的分析方法、直观和运气。阅读原文文字的起码能力是需要的，然而不是必不可少的。这是 Parker Hitt 的"军事密码破译指南"一书的开场白。

1.3.2 对维吉利亚密码的分析[2,5]

维吉利亚密码属多表密码，明文相同的字符变换到密文中未必是相同的字符，只有当

这两个相同字符距离等于密钥长度时，它们对应的密文字符才会相同。因此分析密文中两个相同字符出现的位置，就是寻找密钥长度的关键。

1. 密钥长度的初步判断

首先统计密文中不同距离处出现相同字符的次数。例如密文是如下的280个字符：

UFQUIUDWFHGLZARIHWLLWYYFSYYQATJPFKMUXSSWWCSVFAEVW
WGQCMVVSWFKUTBLLGZFVITYOEIPASJWGGSJEPNSUETPTMPOPHZS
FDCXEPLZQWKDWFXWTHASPWIUOVSSSFKWWLCCEZWEUEHGVGLRL
LGWOFKWLUWSHEVWSWTTUARCWHWBVTGNITJRWWKCOYFGMILR
QESKWGYHAQENDIULKDHZIQASFMPRGWRVPBUIQQDSVMPFZMVEGE
EPFODJQCHZIUZZMXKZBGJOTZAXCCMUMRSSJW

从头到尾考察密文中所有的符号。统计距离为1的两个相同符号出现的次数，就是数出相邻两字符相同的次数，结果是20次；统计距离为2的两个相同符号出现的次数，也就是中间只隔一个符号的两相同符号的次数，结果共出现52次。再统计距离为3的两个相同符号出现的次数，得到是32次。下表列出了两个相同字符出现在距离20以内的各种情况的次数：

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
20	52	32	28	50	18	13	15	11	21	5	8	7	8	16	3	9	5	4	12

可以看到，距离为2和5时，两相同符号出现的次数最多，由此可以推测，密钥长度可能为2或5。但是长度为2的密钥极少有人取，于是初步判定密钥长度为5。

2. 密钥长度的确认

根据初步判断，可以把密文依次按列排序，每列5个字符，构成$m=5$行的阵列。也就是说，把原来的序列像发扑克牌似地依次分配到5行中：

UUGIWYJUWAGVUGTFGPTOFPKWPVKCUGGWHTCVTKGQGNKQPVQM
VPQUKOCR

FDLHYYPXCEQSTZYAGNPPDLDTWSWRELWLETWTJCMEYDDARPQPE
FCZZTCS

QWZWYQFSSVCWBFOSSSTHCZWHISWZHROUVUHGROISHIHSGBDFG
OHZBZMS

UFALFAKSVWMFLVEIJUMZXQFAUSLWGLFWWAWNWTLKAUZFWUSZ
EDZMGAUJ

IHRLSTMWFWVKLIIWEEPSEWXSOFCEVLKSSRBIWPRWELIMRIVMEJIX
JXMW

如果密钥长度确实是5，则若设密钥为$k=k_1k_2k_3k_4k_5$，那么阵列的第1行应当是明文中序号为1，6，11，16……的相应字符用第1位密钥k_1移位加密的结果，第2行是明文中序号为2，7，12，17……的相应字符用第2位密钥k_2移位加密的结果…… 由于同一行的各个元素都是由明文字符通过相同距离的移位产生的，而这种单表加密方式的字符替换不改变原来的概率分布，因此每行仍然应当呈现出与明文英语相同的统计规律。然而，如果