通信网络精品图书

信息安全实验教程

周亚建　郑康锋　武斌　杨义先　编著

電子工業出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

通信网络精品图书

信息安全实验教程

周亚建　郑康锋　武　斌　杨义先　编著

電子工業出版社
Publishing House of Electronics Industry
北京 • BEIJING

内 容 简 介

本书选择信息安全实验教学中最基本、最重要的内容作为切入点，深入到源代码层面讲解几种典型密码算法的原理，剖析各种信息安全技术实现的真实方式，旨在帮助读者深入探究算法与技术的原理，掌握运用所学知识解决各种具体问题的基本方法，提高其编程能力。本书分为密码学实验和网络安全实验两大部分，密码学实验部分包括实验 1 到实验 6，内容覆盖古典密码的加密与密码分析、DES 算法的加 / 解密原理及其差分分析和公钥密码算法的加 / 解密及信息隐藏等内容；网络安全实验部分由实验 7 到实验 14 组成，主要涉及常见的网络攻击技术（扫描、口令破解和嗅探 DoS/DDoS 攻击等）内容。

本书既可作为高等院校信息安全专业研究生、本科生的实验教材，也可作为电子信息与通信工程相关专业的教辅用书，还可供 IT 业工程技术人员学习参考。

图书在版编目（CIP）数据

信息安全实验教程 / 周亚建等编著. —北京：电子工业出版社，2013.1
（通信网络精品图书）
ISBN 978-7-121-18972-2

Ⅰ. ①信… Ⅱ. ①周… Ⅲ. ①信息系统—安全技术—高等学校—教材 Ⅳ. ①TP309

中国版本图书馆 CIP 数据核字（2012）第 274715 号

策划编辑：宋 梅
责任编辑：宋 梅
印 刷：北京中新伟业印刷有限公司
装 订：北京中新伟业印刷有限公司
出版发行：电子工业出版社
北京市海淀区万寿路 173 信箱 邮编 100036
开 本：787×1 092 1/16 印张：15 字数：384 千字
印 次：2013 年 1 月第 1 次印刷
印 数：3 000 册 定价：39.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888。

质量投诉请发邮件至 zlts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：（010）88258888。

前　言

信息安全的重要性不言而喻。随着电子政务、电子商务的进一步普及，政府、企业以及各种社会机构对信息安全专业人才的需求逐渐从数量向质量过渡，以适应日益复杂的网络应用环境。高等院校作为人才培养的基地，有义务向国家和社会输送一批又一批既掌握扎实理论基础、又具有较强动手能力的高素质专业人才。截至2009年，教育部共批准了70所高校设置信息安全类本科专业，其中的15所在2007年年底被教育部批准为“国家特色专业建设点”，从而能够把信息安全专业作为国家特色专业来建设。

但信息安全专业人才的培养绝非易事，原因在于信息安全涉及密码学、数学、计算机科学、通信工程和信息工程等多门学科的交叉，一方面知识体系庞杂、难于掌握，另一方面实践性很强。在长期的教学实践过程中，多数学校基本上形成一种共识：合理、完善的实验课程体系是帮助学生掌握理论体系、培养动手能力的有效途径，并纷纷加大了对实验室建设的投入，加强了实验课程体系的建设。于是，各种各样的实验教材如雨后春笋般地涌现出来。毋庸置疑，每本教材都是对其编著者教学理念和教学经验的总结，反映了不同学校对人才培养的侧重点和不同诉求。

本书几位编著者所在的北京邮电大学信息安全中心承担着本校信息安全本科专业建设和教学的任务，并有幸入选教育部的“国家特色专业建设点”。在多年信息安全实验教学过程中，几位编著者根据社会需求和学生特点的变化不断调整实验内容和实验教学方法，体会到实验教学的难点在于如何实现理论基础和动手操作的平衡：要么是学生的动手能力达不到用人单位的要求；要么是动手能力虽强，但只知其然、不知其所以然，缺乏创新能力。学生在走上工作岗位之后遇到的问题不可能都在上学期间由老师讲授过，需要自己发挥主观能动性，开展创新性思维寻求解决方案。要想具备这种能力，学生必须深刻理解各种信息安全技术和算法的基本原理，光靠“啃书本”不可能实现这一点，必须动手编程、调试程序和跟踪程序运行过程，才能了解每一种算法背后的真正机制。

基于上述考虑，本教程重新梳理了信息安全实验教学内容，结合源代码和详细的注释讲解算法（或技术）原理，引导学生深入到算法的实现过程中去理解其原理。由于篇幅的限制，不可能对所有的技术、理论和算法均进行深入的剖析，只能选择一些典型的、影响较大的算法来分析，关键是对方法论的讲解。这本质上是一种以点带面的方法：学生只要深入理解了书中所讲授的方法，同样能够举一反三，自己去研究、理解其他可能遇到的问题。

作为实现编著者教学理念的载体，本教程有以下几个鲜明的特色：

① 把攻、防统一起来考虑问题。

信息安全问题往往涉及攻击和防御这样两个矛盾的对立面：攻击方考虑的是如何把自己的攻击之矛打磨得无坚不摧，防御方则竭力去铸造坚不可摧的盾。本教程在讲解攻击的时

候，引导学生思考防御之策；而在讲解防御的时候，也不得不思考攻击之法。例如，对于古典密码算法和 DES 密码算法，不但要求学生掌握密码编码的算法，还要求他们能够编程实现密码分析的算法。

② 结合源代码讲解原理。

以最经典的密码算法 DES 为例，按照算法的原理把从明文输入到密文输出之间的加密流程分解为一系列关键步骤，每一步的原理结合相应的 C 语言源代码（含详细的注释）予以讲解。这样做的好处是，学生既理解了原理，又掌握了实现的方法。一旦遇到不同的应用需求，只需对现有的代码做或多或少的修改即可。

③ 采用软件或者开源软件构件实验环境。

学生们朝气蓬勃、思维活跃，随时随地可能产生新的想法，实验为他们提供了验证自己想法的测试环境。本教程涉及的实验仅依赖开源软件，甚至多数情况下要求学生自己编写程序，并不需要昂贵的设备作为支撑。这实际上也是在训练学生掌握自己创造实验条件的方法和技巧。

本教材在编写过程中引用了来自互联网的一些原理描述、源代码及注释，目的是服务于教学，为学生提供更优秀、更便于理解的教学素材和资源，作为正式出版物的参考文献（书籍、学术论文及学位论文）在每一实验的最后都做了标注。

本教材的编写得到了所在灵创团队的老师和研究生们的大力支持和协助，在此一并致谢！同时，由于编著者的水平有限，书中肯定存在这样或那样的问题，欢迎读者在使用过程中予以批评指正。

编著者

2012 年 10 月 21 日于北京

反侵权盗版声明

举报电话：（010）88254396；（010）88258888
传　　真：（010）88254397
E-mail：　dbqq@phei.com.cn
通信地址：北京市万寿路 173 信箱
　　电子工业出版社总编办公室
邮　　编：100036

目　录

第一篇　密码学实验

第二篇　网络安全实验

第一篇　密码学实验

密码学是信息安全的核心技术之一。人类的文明史、战争史无不伴随着密码领域两股力量的生死博弈：一股力量关心的是怎样使密码变成牢不可破的坚盾，而另一股力量则处心积虑地打造无坚不摧之利矛来摧毁任何坚固的密码。正是由于前者的智慧，古今中外、形形色色的密码算法浩如烟海，而后者的思想不仅留下了玛丽女王的叹息，也使得“一战”、“二战”的战局更加扑朔迷离。这场“道高一尺，魔高一丈”的对弈虽历经千年也未分出胜败，今后的斗争恐将更加激烈。信息安全专业的学生作为未来博弈的两方之一，只有熟悉双方的发展历史和趋势，才能在斗争中占据更加主动的位置。

本篇由古典密码入手，通过使用简单的换位与替换进行加密以及基于频度分析的密码攻击，消除初学者对密码的神秘感，进而引导学生采用现代技术构建更复杂的密码体制，同时掌握基于现代数学理论的密码分析技术。本篇内容覆盖了古典密码学实验、对称密码实验、非对称密码实验、信息隐藏和数字水印等内容。

本篇的实验内容需要密码学等先修课程作为理论基础。

实验 1　古典密码学实验

1.1　实 验 目 的

通过对古典密码算法的编程实现和频度分析，理解古典密码编码和分析的基本原理和方法。

1.2　实 验 原 理

古典密码的加密方法一般是文字置换，使用手工或机械变换的方式实现。最具代表性的古典密码体制主要包括单表代替密码、多表代替密码及转轮密码。凯撒密码就是一种典型的单表加密体制，多表代替密码包括 Vigenere 密码和 Hill 密码等，著名的 Enigma 密码就是第二次世界大战中使用的转轮密码。

1.2.1　凯撒密码

凯撒密码（Caesar Cipher）是罗马扩张时期由朱利斯·凯撒（Julius Caesar）创造的，用于加密通过信使传递的作战命令。

（1）凯撒密码的加／解密原理

凯撒密码的加密算法极其简单，它将字母表中的字母移动一定位置而实现加密，如图 1.1 所示。

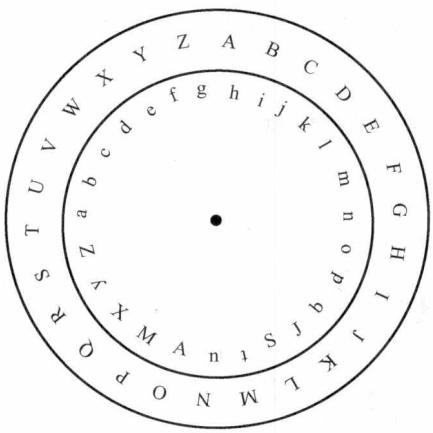

图 1.1　凯撒密码原理示意图

凯撒密码的加 / 解密过程等同于一个变换。如果把明文记为 m，密文记为 c，加密变换记为 $E(k, m)$（其中 k 为密钥），解密变换记为 $D(k, m)$，则加密过程可记为如下的变换：

$$c \equiv m+k \bmod n$$（其中 n 为基本字符个数）

同样，解密过程可表示为

$$m \equiv c+k \bmod n$$（其中 n 为基本字符个数）

对于计算机而言，n 可取 256 或 128，m、k、c 均为一个 8 bit 的二进制数。

这种加密算法的安全性不高，采用穷举法或者频率分析法都可轻易地破解。如果采用穷举法，最多只需要 255 次即可破译。本实验将重点讨论如何采用频率分析法进行破解。

（2）凯撒密码的频度分析原理

对于任何一种书面语言而言，不同的字母或字母组合出现的频率各不相同。如果以这种语言书写足够长的文本，都呈现出大致相同的特征字母分布。从图 1.2 所示的英语字母出现的频率不难发现，字母E出现的频率很高，而X则出现得较少。类似地，ST、NG、TH 以及 QU 等双字母组合出现的频率非常高，NZ 和 QJ 组合则极少。英语中出现频率最高的 12 个字母可以简记为“ETAOIN SHRDLU”。

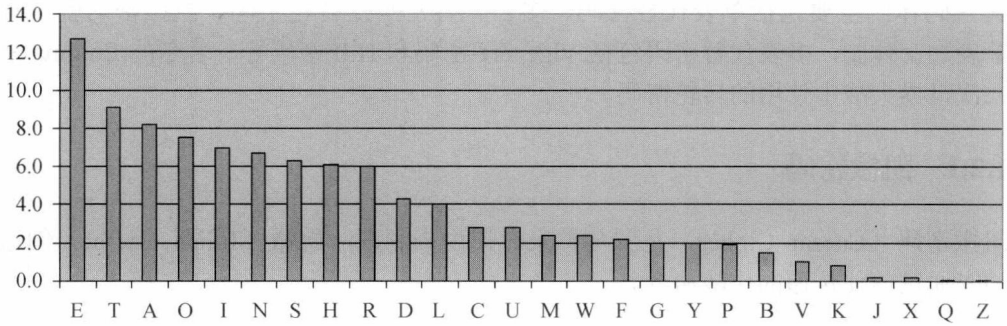

图 1.2　英语字母的频率

柯南·道尔（Arthur Conan Doyle）在其侦探小说《福尔摩斯归来记》中的“跳舞的人”篇中，讲述福尔摩斯在墙上看到 5 个跳舞的小人，如图 1.3 所示。他根据小人出现的频率，并和英语字母的频率相比照，猜出了其中一次跳舞人画所代表的字为“Never”，从而破解了字谜。

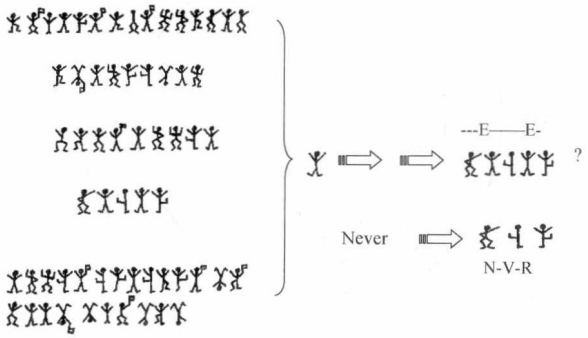

图 1.3　跳舞的人

对凯撒密码进行频度分析的源代码请参见附录 A.1。

1.2.2　维吉尼亚密码（Vigenère Cipher）

由于频率分析法可以有效地破解单表替换密码，法国外交官、密码学家维吉尼亚（Blaise De Vigenère）于 1586 年提出了一种多表替换密码，即维吉尼亚密码。维吉尼亚密码引入了“密钥”的概念，即根据密钥来决定用哪一行的密表来进行替换，以此来对抗字频统计。如果使用 m 个字母组成密钥，维吉尼亚密码的密钥空间大小为 26^m，即使 m 值很小，使用穷尽密钥搜索方法也需要很长的时间。例如，当 m=6 时，密钥空间的大小超过 3×10^8，这样的密钥量已经超出了使用手工计算进行穷尽搜索的能力范围。由于破译的难度很高，维吉尼亚密码也因此获得了很高的声望。数学家查尔斯·路特维奇·道奇森（笔名路易斯·卡罗）在 1868 年所编的《字母表密码》（The Alphabet Cipher）中称其是不可破译的密码。1917 年，《科学美国人》将维吉尼亚密码称为“无法被转化的”密码。然而，查尔斯·巴贝奇完成了破译的工作，但他没有将此发表。之后，弗里德里希·卡西斯基（Friedrich Kasiski）于 19 世纪完全破解并发表了他的方法。

（1）维吉尼亚密码加／解密原理

为了抵抗频度分析强度，维吉尼亚密码由一些偏移量不同的恺撒密码组成，而且密码使用表格法生成。这一表格包括了 26 行字母表，每一行都由前一行向左偏移一位得到，如图 1.4 所示。在加密过程中由密钥决定具体使用哪一行字母表进行编码。

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

图 1.4　维吉尼亚表格

例如，假设明文为 SUBMARINE，选择某一关键词并重复而得到密钥（当关键词为 OPENER 时，密钥为 OPENEROPENER）。对于明文的第一个字母 A，对应密钥的第一个字母 O，于是使用表格中第 *O* 行字母表进行加密，得到密文第一个字母 G。类似地，明文第二个字母为 U，对应密钥的第二个字母 P，于是使用表格中第 *P* 行进行加密，得到密文第二个字母 J。以此类推，可以得到：

明文	S	U	B	M	A	R	I	N	E
密钥	O	P	E	N	E	R	O	P	E
密文	G	J	F	Z	E	I	W	C	I

解密的过程正好与加密相反。例如，根据密钥第一个字母 O 所对应的第 *O* 行字母表，发现密文第一个字母 G 位于第 *S* 列，因而明文第一个字母为 S。密钥第二个字母 P 对应第 *P* 行字母表，而密文第二个字母 J 位于此行第 *U* 列，因而明文第二个字母为 U。以此类推便可解出其余明文。

用数字 0～25 代替字母 A～Z，维吉尼亚密码的加密文法可以写成同余的形式：

$$C_i \equiv P_i + K_i \pmod{26}$$

解密方法则能写成：

$$P_i \equiv C_i - K_i \pmod{26}$$

（2）维吉尼亚密码分析

破译包括维吉尼亚密码在内的所有多表密码仍然基于字母频率，但类似于破解单表替换密码的直接频率分析却并不适用，因为同一个明文在密钥的作用下可能被加密成不同的密文。

由于使用的密钥是循环重复的，破译维吉尼亚密码的关键在于确定密钥的长度。如果知道了密钥的长度，密文就可以看作交织在一起的凯撒密码，而其中每一个都可以单独破解。通常，使用卡西斯基试验和弗里德曼试验都可得到密钥的长度。

（3）卡西斯基试验

弗里德里希·卡西斯基于 1863 年首先发表了完整的维吉尼亚密码的破译方法，称为卡西斯基试验（Kasiski Examination），其基本思想是：类似 the 这样的常用单词有可能被同样的密钥加密，从而在密文中重复出现。例如，明文中不同的 CRYPTO 可能被密钥 ABCDEF 加密成不同的密文：

明文	C	R	Y	P	T	O	I	S	S	H	O	R	T	F	O	R	C	R	Y	P	T	O	G	R	A	P	H	Y
密钥	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D
密文	C	S	A	S	X	T	I	T	U	K	S	W	T	G	Q	U	G	W	Y	Q	V	R	K	W	A	Q	J	B

此时，明文中重复的元素在密文中并不重复；如果把密钥换为 ABCD：

明文	C	R	Y	P	T	O	I	S	S	H	O	R	T	F	O	R	C	R	Y	P	T	O	G	R	A	P	H	Y
密钥	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
密文	C	S	A	S	T	P	K	V	S	I	Q	U	T	G	Q	U	C	S	A	S	T	P	I	U	A	Q	J	B

此时，卡西斯基试验就能产生效果，而且明文的段落越长，试验就越有效，因为密文

中重复的片段可能会更多，通过下面的密文就能破译出密钥的长度：

DYDUXRMHTVDVNQDQNWDYDUXRMHARTJGWNQD

其中，两个 DYDUXRMH 的出现相隔了 18 个字母。因此，可以假定密钥的长度是 18 的约数，即长度为 18、9、6、3 或 2。而两个 NQD 则相距 20 个字母，意味着密钥长度应为 20、10、5、4 或 2。取两者的交集，则可以基本确定密钥长度为 2。

再如下面的密文：

ISWZPNQCKMYYYJKAYYEZFFSWEESSPGZXQAHF
ISWZPNQCKMTVYJOACVEHAESAZRLTPQIZMXOT
QSWMCVUDSIJGGDEUWAZRSFXWILKUEJQLDACB
GDLYJXMYLMDQKZMPLDILQEMWFSWDPAZEZQNW
DYWDZXFSAEEAZJDUELVPTMCEKWSEEFURZFSW
DPXACQAFKMXWAWVEZFSDBGDLAYUQXGDPEKWS
EEFURZFSWDPOUEZKZMYLQNPQQDEMJTQYGUVA
ZOGRWAWPVUEQAFJQJGGCOMJZAHQAFKTJDKAD
MNWPJGGCWKPKAYEQZZPTVKZMQGWDVFAHLTLL
USSPXAZPGZJGGOSDWAZRKAEZQCWKZMMCWITL
TEZMEDAZCAYQAFJRLUQLKUQQAFJQYWHPJTFJ
FLKUQQAFJQYWHPJPZOZDZMWDUMWFSWAYWRZJ
KZMISGBTFOSEEJGGDGREDKMMFDMDPARQJAHF
UDKTZOZEZQYAITDXVFAHLTLLKZMMCWZZVDPS
YPJ

在里面，重复序列有 ISWZPNQCKM、BGDL、SEEFURZFSWDP、JGGC、LKUQQAFJQYWHPJ 和 VFAHLTLL 等，而且每个重复间隔都能被 3 整除，密钥应该有 3 个字母。

一旦确定了密钥的长度，密文就能被划分为多个组，组数与密钥长度对应。由于每一组内所有密文的密钥（偏移量）对应于维吉尼亚密码密钥的相应字母，每一组其实就是一个凯撒密码，可采用频率分析的方法将密文破译。

柯克霍夫方法作为卡西斯基试验的改进，由奥古斯特·柯克霍夫（Auguste Kerckhoffs）提出。它将每一列的字母频率与转换后的明文频率相对应而得出每一列的密钥字母。一旦密钥中每一个字母都能确定，就能很简单地破译密文，从而得到明文。如果维吉尼亚字母表表格本身是杂乱的而不是按通常字母表排序的话，柯克霍夫方法就会无效，但卡西斯基试验和重复指数对于决定密钥长度仍旧是有效的。

在上例中，因为密钥有 3 个字母，可把密文分为 3 组进行破解，每一组有 169 个字母：

① 把第 1、4、7、10、13……个字母分为一组，称之为 L1；

② 把第 2、5、8、11、14……个字母又分为一组，称之为 L2；

③ 余下的归另一组，称之为 L3。

接下来，用 169 乘以各个字母的标准百分比即可得各字母的标准个数，如表 1.1 所示。