胡予濮　张玉清　肖国镇　编著

对称密码学

SYMMETRIC KEY CRYPTOGRAPHY

机械工业出版社
China Machine Press

对称密码学

胡予濮　张玉清　肖国镇　编著

国家自然科学基金资助项目（项目编号：60102004）
ISN 国家重点实验室开放课题基金资助项目
国防重点实验室预研基金资助项目

机械工业出版社

本书是信息与网络安全基础——对称密码学的专著，全面论述了对称密码学中的各种基本问题和最新研究进展。本书部分成果来源于国家自然科学基金资助项目，ISN 国家重点实验室开放课题基金资助项目和国防重点实验室预研基金资助项目。

著作论述了密码函数与信息泄露、序列密码和分组密码等内容，详细介绍了密码体制的安全性概念，密码函数的安全性指标——相关免疫函数、非线性性、差分与高阶差分，相关免疫函数的自然延伸——弹性函数；介绍序列的伪随机性，线性复杂度的计算，几何序列、对数序列、缩减序列及有特殊用途的稀疏序列；讨论简捷快速的分组密码体制，这是目前软硬件加密标准的主流。同时介绍几个著名的分组密码设计方案，包括 IDEA、RC5、RC6、Twofish 和著名的软件加密算法 SAFER+等。书中还给出几个重要的分组密码的 C -源程序代码。

本书面向信息安全领域的科研工作者，网络安全工程技术人员，信息安全专业的师生和从事通信、电子、计算机科学的科技人员等。

图书在版编目（CIP）数据

对称密码学/胡予濮等编著. —北京：机械工业出版社，2002.7
ISBN 7-111-10674-1

I. 对… II. 胡… III. 对称—密码—理论 IV.TN918.1

中国版本图书馆 CIP 数据核字（2002）第 054201 号

机械工业出版社（北京市百万庄大街 22 号 邮政编码 100037）
策　　划：胡毓坚
责任编辑：张　克
责任印制：路　琳

北京大地印刷厂 印刷 · 新华书店北京发行所发行

2002 年 8 月第 1 版 · 第 1 次印刷
1000mm×1400mm　B5 · 8.5 印张 · 388 千字
0 001—5 000 册
定价：25.00 元

前　言

信息安全的观念早已深入人心。现在已有许多密码学的书籍问世，许多“圈外人”在谈及 DES、RSA 时也能够略知一二。这就是说，在这里既不需要对读者进行密码学启蒙，也不必赘述密码学那充满传奇色彩的历史和五花八门的应用领域。

密码分为单钥密码（又称为对称密码）和公钥密码，单钥密码又分为序列密码和分组密码。他们各有特点，其设计原理、安全性机制也很不同。可以给出以下粗略的认识：

公钥密码是一个陷门单向函数，其主要原理是计算的单向可行性。比如离散指数与离散对数，数字背包的装载与卸载，大整数的组成与素分解等等。序列密码是一个伪随机序列发生器，分组密码是一个伪随机置换。

另外，单钥密码与公钥密码的应用领域也有很大的不同。序列密码体制以简捷、快速的生成算法，使其成为新一代移动通信的主流加密算法；分组密码体制也具有简捷、快速的特点，并且容易标准化，使其成为软硬件加密标准的主流。而公钥密码体制的运行速度则慢得多，占用的空间也大得多，一般用于数字签名、电子商务等计算资源相对宽松的领域。从以上可知，单钥密码与公钥密码的差别很大，而序列密码与分组密码则有许多相近之处，这是作者写本书的动机之一。

序列密码方面有三本影响最大的著作：Siegenthaler 的博士论文《流密码系统的设计方案》；Rueppel 的研究专著《流密码的分析与设计》；丁存生等人的《流密码学及其应用》。分组密码方面有四个著名的算法标准：美国 IBM 公司研制的“DES”；来学嘉博士设计的“IDEA”；密码学专家 James. L. Massey 设计的“SAFER”；Joan Daemen 和 Vincent Rijmen 共同提交的“RIJNDAEL”。分组密码的密码分析方面有两个著名的方法：E. Biham 和 A. Shamir 的“差分密码分析”；Matsui 等人的“线性密码分析”。此外单钥密码的研究还有许多新进展。以上这些都是本书的材料来源之一。本书的另一个材料来源是：本书作者以及西安电子科技大学信息安全研究所的同志们在这一领域新的研究成果。本书主要作者分别为研究密码学与网络安全方面的博士生导师、博士后及研究所所长等研究密码学与网络安全的专家和学者。

与其它密码学方面的著作和书籍相比，本书的写作力求新颖、深入、全面。很多高质量的学术论文被收入其中（甚至包括被录用但还未发表的论文），以便对相关领域的科学研究有更具体、更直接的帮助。因此本书不是一本密码学科普类读物。本书的另一个宗旨是工程实用，为此有七个著名分组密码算法的 C-源程序代码被收编。

尽管在本书的写作之前做了充分的准备，但在书稿写作中仍感到时间仓促、水平有限，在此衷心地希望读者批评指正。愿作者的初衷能够被充分体现，愿密码学

学科能够在学术领域和应用领域比翼双飞。有关本书三位作者的详细情况请参见“中国密码”网站（www.ChinaCrypt.net）上的介绍，对本书的编写有什么具体的意见和建议，敬请用 E-mail 的方式和 WebMaster@ChinaCrypt.net 联系，我们将在以后的修订中进行修改。

编　者

目　录

前言

第一篇　密码函数与信息泄露 …… 1

第 1 章　密码体制的安全性 …… 2

1.1　密码体制与密码函数 …… 2

1.2　密码体制的安全性之一
完善保密性 …… 3

1.3　密码体制的安全性之二
计算安全性 …… 5

1.4　完善保密性与计算安全性的
比较 …… 5

1.5　完善保密性的获得
密钥协商 …… 6

1.5.1　优先提取 …… 6

1.5.2　信息协调 …… 9

1.5.3　保密增强 …… 10

1.6　优先提取与优先退化 …… 11

1.7　存在主动攻击时的保密
增强 …… 14

参考文献 …… 14

第 2 章　相关免疫函数 …… 16

2.1　相关免疫布尔函数及其
基本性质 …… 16

2.2　相关免疫阶与代数次数的
相互制约 …… 17

2.3　相关免疫布尔函数的构造
和计数 …… 19

2.3.1　结构定理 …… 19

2.3.2　一阶相关免疫布尔函数的
构造和计数 …… 20

2.3.3　高阶相关免疫布尔函数的
构造和计数 …… 22

2.4　布尔函数的广义相关
免疫性 …… 23

2.5　GF(q)上的相关免疫
函数 …… 24

参考文献 …… 29

第 3 章　弹性函数 …… 30

3.1　弹性函数及其
基本性质 …… 30

3.2　弹性函数的存在性与
惟一性 …… 32

3.3　弹性函数的构造 …… 33

3.4　仿射函数的弹性阶 …… 36

3.5　弹性阶上确界与代数次数的
关系 …… 37

参考文献 …… 40

第 4 章　非线性性 …… 41

4.1　布尔函数的非线性度和
线性度 …… 41

4.2　布尔函数非线性度与相关
免疫阶的关系 …… 43

4.3　高度非线性布尔函数
Bent 函数 …… 43

4.4　高度非线性均衡布尔函数的
构造 …… 46

4.4.1　$n=2^m$ …… 47

4.4.2　$n=2^m(2s+1)$，s 为奇数 …… 48

4.4.3　n 为奇数 …… 48

4.5　特征为 2 的域上的多输出
函数的非线性性 …… 48

4.6　一般有限域上的函数的
非线性性 …… 52

参考文献 ……55
第 5 章 密码函数的其它安全性设计 ……56
5.1 差分分布……56
5.2 特殊的差分分布：雪崩与扩散……57
5.3 高阶差分分布……58
5.4 高阶自相关性……59
5.5 正形置换……60
5.6 全距置换……61
参考文献 ……62
第二篇 序 列 密 码 ……63
第 6 章 序列密码的基础理论 ……64
6.1 序列的线性复杂度和最小周期……64
6.2 序列的根表示和迹表示……68
6.3 和序列与乘积序列……70
6.4 *m*-序列及其密码学特性……73
6.5 密钥序列的稳定性……76
6.6 线性递归序列的综合……77
6.6.1 B-M 算法……78
6.6.2 Games-Chan 算法 ……79
6.7 序列密码的研究现状简述……79
6.7.1 伪随机序列的生成现状 ……79
6.7.2 对序列密码的攻击现状 ……80
6.7.3 序列密码的某些非主流问题……80
参考文献 ……81
第 7 章 前馈序列 ……82
7.1 Bent 序列 ……82
7.2 几何序列……86
7.3 几何序列之例一 GMW 序列 ……88
7.4 几何序列之例二 瀑布型 GMW 序列 …… 89
7.5 No 序列 …… 92
参考文献…… 94
第 8 章 对数序列…… 96
8.1 对数序列的定义、定理 … 96
8.2 对数序列之例一 Legendre 序列 …… 99
8.3 对数序列之例二 *R* 为奇素数 …… 101
8.4 对数序列之例三 *R* 为 2 的幂 …… 102
8.5 对数序列的推广 广义 Jacobi 序列 …… 105
参考文献…… 107
第 9 章 钟控序列…… 108
9.1 Jennings 复合序列…… 108
9.2 停-走生成器 …… 114
9.3 Gunther 生成器 …… 116
9.4 缩减序列 互缩序列 …… 117
9.5 缩减序列 自缩序列 …… 119
9.6 缩减序列 广义自缩序列 …… 119
9.7 广义自缩序列的特例 … 126
参考文献…… 133
第 10 章 稀疏序列 …… 134
10.1 稀疏序列与信息隐藏 … 134
10.2 自缩乘积序列与自扩序列…… 135
10.3 基于 GF(q)上 *m*-序列的稀疏序列…… 136
10.4 基于乘方剩余符号的稀疏序列 …… 140
参考文献…… 142
第三篇 分组密码…… 143

第 11 章　分组密码的设计与安全性 …… *144*
11.1　分组密码的设计准则 …… *144*
11.1.1　安全性 …… *144*
11.1.2　简捷性 …… *145*
11.1.3　有效性 …… *146*
11.1.4　透明性和灵活性 …… *146*
11.1.5　加解密相似性 …… *146*
11.2　分组密码的设计技巧 …… *146*
11.2.1　计算部件 …… *146*
11.2.2　计算部件的组合 …… *148*
11.2.3　关于密钥长度 …… *150*
11.3　分组密码的工作模式 … *150*
11.4　典型攻击方法 …… *152*
11.4.1　朴素的攻击：穷举搜索 …… *152*
11.4.2　差分密码分析 …… *153*
11.4.3　线性密码分析 …… *155*
11.4.4　计时攻击和能量攻击 …… *155*
11.5　分组密码的随机算法 …… *155*
参考文献 …… *156*
第 12 章　分组密码 DES …… *157*
12.1　DES 概述 …… *157*
12.2　DES 的计算部件 …… *158*
12.2.1　初始置换 *IP* 与其逆置换 IP^{-1} …… *158*
12.2.2　扩充变换 *E* …… *158*
12.2.3　8 个 S 盒 S_1、S_2、…、S_8 …… *159*
12.2.4　置换 *P* …… *160*
12.3　DES 的加密算法和解密算法 …… *160*
12.4　DES 的 C-源程序代码… *162*
12.5　DES 的安全性 …… *175*
12.6　对 DES 的差分密码分析 …… *176*
参考文献 …… *179*
第 13 章　各种分组密码设计方案 …… *180*
13.1　分组密码 IDEA …… *180*
13.2　IDEA 的 C-源程序代码 …… *182*
13.3　RC5 与 RC6 …… *191*
13.4　RC5 与 RC6 的 C-源程序 …… *194*
13.5　Feistel 网络的变形 …… *201*
13.6　分组密码 Twofish 简介 …… *202*
13.7　Twofish 的 C-源程序 … *205*
参考文献 …… *216*
第 14 章　AES 算法（RIJNDAEL） …… *217*
14.1　AES 竞争过程及 RIJNDAEL 概述 …… *217*
14.2　RIJNDAEL 的数学基础和设计思想 …… *218*
14.2.1　有限域 $GF(2^8)$ …… *218*
14.2.2　系数在 $GF(2^8)$ 上的多项式 …… *218*
14.2.3　设计思想 …… *220*
14.3　算法说明 …… *220*
14.3.1　状态、密钥种子和轮数 *220*
14.3.2　轮函数 …… *221*
14.3.3　密钥扩展 …… *223*
14.3.4　RIJNDAEL 密码的加密算法 …… *225*
14.3.5　加解密的相近程度/解密算法 …… *225*

14.4 实现方面 …… 227
14.4.1 8位处理器 …… 227
14.4.2 32位处理器 …… 227
14.4.3 并行性 …… 228
14.4.4 解密算法的实现 …… 229
14.4.5 硬件适应性 …… 229
14.5 设计选择的诱因以及安全性分析 …… 229
14.6 RIJNDAEL的C-源程序代码 …… 231
参考文献 …… 240
第15章 分组密码SAFER+及其变形 …… 242
15.1 SAFER+概述 …… 242
15.2 SAFER+的C-源程序代码 …… 245
15.3 SAFER+M算法描述 … 255
15.4 SAFER+M与SAFER+计算量和数据量的比较 …… 257
15.5 SAFER+M的加解密相似性 …… 258
15.6 SAFER+M的安全性 … 260
15.6.1 线性层的扩散性能：与SAFER+比较 …… 260
15.6.2 SAFER+M的差分密码分析 …… 262
参考文献 …… 263

第一篇　密码函数与信息泄露

单钥密码体制的安全性基础是密码函数的伪随机性。它包括三个方面：（1）静态信息泄露的缓慢性和均匀性。具体有以下的安全性度量指标：非线性性、弹性、混淆性和扩散性。（2）动态信息泄露的缓慢性和均匀性。具体有以下的安全性度量指标：差分分布与高阶差分分布均匀性、积分分布的均匀性（包括 Square 分布的均匀性）、自相关性和互相关性。（3）前述两个方面各安全性指标的稳定性。以上这些安全性概念有的相互包含或相互交叉，有的相互制约，弄清楚它们的关系是安全性设计的重要前提。

简捷性和透明性属于密码函数的实用设计。简捷性指的是密码函数的生成算法要简单快速；透明性指的是安全性要清晰可见，要能够被证明。

伴随着现代密码技术的发展，密码函数的设计与分析从来都是研究的热点。各种安全性的概念在过去 30 年内陆续提出，其中有些已经被深入地研究。在 80 年代，Ruepple 和 Siegenthaler 的研究已经表明非线性性和相关免疫性是相互制约的。1988 年肖国镇教授揭示了布尔函数代数次数与相关免疫阶的紧制约关系，同时开始了密码函数研究的频谱分析方法。此后密码函数的设计与分析如雨后春笋，其中包括冯登国博士在频谱分析方面的出色工作以及应用于非线性性的研究。但随着研究的深入，又相继出现许多新的问题，其中大部分问题将在本篇中给予讨论。

本篇的第 1 章介绍密码体制的安全性，给出两种不同的安全性概念以及它们的研究现状，特别是在完善保密方面的最新进展。第 2 章介绍相关免疫性，它是密码函数的安全性指标之一。该章论述了布尔函数相关免疫的局限性，构造和计数，以及广义相关免疫性和非布尔函数的相关免疫性。第 3 章介绍弹性函数。弹性函数原本是相关免疫函数概念的自然延伸，由于其应用领域和所包含的问题广泛得多，因此专门讨论。第 4 章介绍非线性性，它是密码函数最重要的安全性指标之一。本章介绍了非线性性与相关免疫性的相互制约，高度非线性函数的构造等。第 5 章介绍密码函数的其它一些安全性指标，包括差分分布、高阶差分分布、高阶自相关性，以及正形置换和全距置换。

第 1 章　密码体制的安全性

1.1　密码体制与密码函数

Shannon 在 1949 年发表的经典著作《保密系统的信息理论》中，对保密系统的运行做了如下的描述：

通信双方 Alice 和 Bob 通过一个安全信道进行相互协商，确定了一个共享的密钥 Z；

Alice 欲通过一个不安全的信道向 Bob 发送明文消息 X；Alice 使用钥控加密算法 $E_Z(\cdot)$将明文 X 变换为密文 Y，$Y=E_Z(X)$；Alice 通过不安全的信道将密文 Y 发送给 Bob；

Bob 使用钥控解密算法 $D_Z(\cdot)$将密文 Y 变换为明文 X，$X=D_Z(Y)$；

截听者 Eve 在不安全的信道上截获了密文 Y，他试图进行攻击（攻击的方式有：被动攻击，即破译密文 Y 得到明文 X 或密钥 Z；主动攻击，即毁坏或篡改密文以达到破坏明文的目的）。见图 1-1。

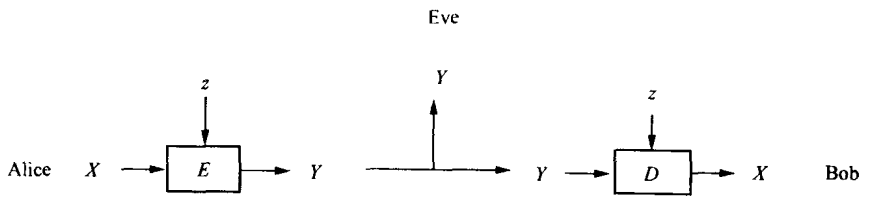

图 1-1　保密系统的运行

以上所描述的五位一体（X，Y，Z，$E_Z(\cdot)$，$D_Z(\cdot)$）就是一个密码体制，其中的 $E_Z(\cdot)$和 $D_Z(\cdot)$都是密码函数。一般地说，信息安全工程中的消息加密方案和认证方案统称为密码体制，所使用的函数统称为密码函数。

概括以上描述的例子，我们需要指出以下 4 点：

(1) Alice 和 Bob 之间的安全信道通常是低速率的，不能直接用于消息传递，否则就没有必要在不安全的信道上发送加密消息了。在安全信道上只能进行密钥协商，而且不能太频繁，这就是说安全信道的低速率使我们不能太频繁地更换密钥。

(2) 为消息加密方案中使用的密码函数，加密算法 $E_Z(\cdot)$和解密算法 $D_Z(\cdot)$的设计必须满足各种各样的置乱功能。比如，当两个密钥 Z_1 和 Z_2 不相同时，$E_{Z_1}(\cdot)$ 和 $E_{Z_2}(\cdot)$“差别很大”。又比如，当两个明文 X_1 和 X_2 “差别很小”时，它们对应的

密文 Y_1 和 Y_2“差别很大”；反之亦然。我们统称这些置乱功能为“伪随机性”。（注：作为认证方案中使用的某些密码函数，需要一些与置乱功能恰恰相反的认证及纠错功能，本书将不涉及）

(3) 对密码函数的另一个基本要求是稳定的安全性，当敌手已经知道了一部分密钥或一部分明 文时，整个密码体制不至于立刻崩溃。比如，密钥是一个 n 长的比特串，且敌手已经知道其中的 k 个比特。如果密码体制的安全性过分地依赖于这 k 个比特，那就很危险了。反之，如果密钥中的 n 个比特对密码体制安全性的贡献比较平均，则无论敌手知道的是哪 k 个比特，他所面对的破译难度是相同的。特别，如果密文独立于密钥中的任意 k 个比特（尽管 n 个比特的密钥已经完全确定了密文），则无论敌手知道的是哪 k 个比特，我们仍然能获得良好的安全性。我们将把这种性质称为相关免疫性或弹性。

(4) 为了达到密码算法商业化的目的，我们通常假设钥控加密算法 $E_Z(\cdot)$和钥控解密算法 $D_Z(\cdot)$是可以公开的；因此要求密码函数 $E_Z(\cdot)$和 $D_Z(\cdot)$的设计必须满足“唯密钥安全性”；这就是说，密文的保密性仅仅依赖于密钥的保密性，而与算法 $E_Z(\cdot)$和 $D_Z(\cdot)$的公开与否无关。

截听者 Eve 对保密系统的攻击强度依赖于他所掌握的信息。商用密码的迅速发展使我们必须假设 Eve 对密码体制是已知的，且 Eve 可以在不安全信道上任意地截取密文。另一个合理的假设是 Eve 始终不知道密钥（在某些特定的场合，可以假设 Eve 知道一部分密钥，但决不能知道全部密钥）。不能保证 Eve 对于其它秘密完全无知。在当今这样一个信息爆炸的时代，经常可能有一些“过时了的明文”被当作信息垃圾随意抛弃，而截听者 Eve 就可能从这些信息垃圾中找到攻击的突破口。对一个保密系统的攻击通常分为以下 3 种：

(1) 唯密文攻击：Eve 仅仅能够截取密文。

(2) 已知明文攻击：Eve 能够截取密文，而且能够获取一些“过时了的密文”所对应的明文。设 Eve 已知 k 组明文/密文对：(X_1, Y_1)、(X_2, Y_2)、…、(X_k, Y_k)。则 Eve 拥有了含 k 个方程的方程组$\{Y_j=E_Z(X_j),\ j=1\sim k\}$。Eve 希望从方程组中求解密钥 Z，或求解一个新截取的密文 Y_{k+1} 所对应的明文 X_{k+1}。从理论上来说，k 越大，Eve 成功的希望就越大。

(3) 选择明文攻击：Eve 能够截取密文，还能够获取一些“过时了的密文”所对应的明文，而且这些明文是 Eve 所希望的。这些“经过选择的”明文/密文对 (X_1, Y_1)、(X_2, Y_2)、…、(X_k, Y_k)，可能恰好暴露了密码函数的最弱之处，比如暴露密钥的某些位，等等；这使得方程组$\{Y_j=E_Z(X_j),\ j=1\sim k\}$变得更容易求解。

1.2 密码体制的安全性之一 完善保密性

本书基本上采用通用记号，比如：GF(q)表示特征为素数 p 的有限域，其中 $q=p^n$；

$P(\cdot)$表示概率；$P(\cdot \mid \cdot)$表示条件概率；$H(\cdot)$表示 Shannon 熵；$H(\cdot \mid \cdot)$表示 Shannon 条件熵；$I(\cdot;\cdot)$表示互信息；$W_H(\cdot)$表示 Hamming 重量，等等。

以下总设明文 X、密文 Y、密钥 Z 都是随机变量。由于

$$Y=E_Z(X),\quad X=D_Z(Y)$$

因此（X，Z）惟一确定了 Y，而（Y，Z）也惟一确定了 X。用信息论的语言就是

$$H(X \mid YZ)=0;\quad H(Y \mid XZ)=0 \tag{1-1}$$

定义 1-1 我们称密码体制是完善保密的，如果

$$I(X;Y)=0 \tag{1-2}$$

定理 1-1 完善保密的密码体制必然有

$$H(Z)\geqslant H(X) \tag{1-3}$$

证明 熵不等式如下所示。

$$\begin{aligned}H(X)&=H(X \mid Y)+I(X;Y)\\&=H(X \mid Y)\quad\text{（因为式(1-2)）}\\&\leqslant H(XZ \mid Y)\\&=H(Z \mid Y)+H(X \mid YZ)\\&=H(Z \mid Y)\quad\text{（因为式(1-1)）}\\&\leqslant H(Z)\end{aligned}$$

定理 1-1 得证。

关于完善保密，有以下的注解：

(1) 完善保密意味着明文随机变量 X 和密文随机变量 Y 相互独立。它的直观含义是：当攻击者不知道密钥时，知道对应的密文对于估计明文没有任何帮助。这是最强的安全性概念。

(2) 定理 1-1 表明，完善保密的密码体制其密钥的不确定性要不小于消息的不确定性。比如，当明文 X 是 n 比特长的均匀分布随机变量，为了达到完善保密，密钥 Z 的长度必须至少是 n 比特长；而且为了用 n 比特长的密钥达到完善保密，密钥也必须是均匀分布的随机变量。这意味着完善保密的密码体制需要消耗大量的密钥。

(3) 完善保密的密码体制是存在的。比如，当明文 $X=(x_1, x_2, \cdots, x_n)$是 n 比特长的均匀分布随机变量，密钥 $Z=(z_1, z_2, \cdots, z_n)$也是 n 比特长的均匀分布随机变量，加密算法为 $Y=X\oplus Z$，其中 $\oplus$ 为逐比特异或运算。由于 $\oplus$ 是群运算，故容易看出 Y 是 n 比特长的均匀分布随机变量，且 X 和 Y 相互独立。

(4) 由概率统计和信息论的知识知道，为了实现完善保密，通信双方必须在每一次传递秘密消息时，所用的密钥对于敌手来说都是完全未知的。这就是说，要传递一个新的消息，必须首先更新密钥。我们称这种体制为一次一密制。完善保密的密码体制的密钥一般不能用于多次加密。比如在注解（3）的例子中，设两个明文

$X^{(1)}$ 和 $X^{(2)}$ 都是 n 比特长的均匀分布随机变量，它们用同一个密钥 Z 进行加密，分别得到密文随机变量 $Y^{(1)}=X^{(1)}\oplus Z$ 和 $Y^{(2)}=X^{(2)}\oplus Z$。设 $X^{(1)}$，$X^{(2)}$，Z 相互独立。则容易看到：$X^{(1)}$，$X^{(2)}$，$Y^{(1)}$ 相互独立；$X^{(1)}$，$X^{(2)}$，$Y^{(2)}$ 相互独立。但由于 $X^{(1)}\oplus X^{(2)}\equiv Y^{(1)}\oplus Y^{(2)}$，故 $(X^{(1)}, X^{(2)})$ 与 $(Y^{(1)}, Y^{(2)})$ 不相互独立。这说明在注解（3）的例子中，重复使用密钥是不能得到完善保密的。

(5) 定理 1-1 使我们想到一个问题：对于明文随机变量 X，是否存在“最节省密钥的”完善保密的密码体制？即是否存在密码函数 $E_Z(\cdot)$和 $D_Z(\cdot)$，使得 $H(Z)=H(X)$？这是一个编码的问题，我们不详细讨论，基本结论是：在 $H(X)\to+\infty$的过程中，总有完善保密的密码体制使得 $H(Z)$任意接近 $H(X)$。

综上所述，我们对完善保密的密码体制有了这样的认识，要想实现完善保密，必须有时时更新的密钥；只要有时时更新的密钥，就一定能实现完善保密。

1.3 密码体制的安全性之二 计算安全性

一个密码体制（X，Y，Z，$E_Z(\cdot)$，$D_Z(\cdot)$），如果破译所需的代价太大而难以实现，这个密码体制就称为计算安全的。这里的“代价”通常指的是计算复杂度，有时也可包括经济代价。计算复杂度原本分为时间复杂度和空间复杂度，但由于并行计算技术的发展，在许多情况下可以进行时空转换，故一般不分时空，统称为计算复杂度。计算安全性已经有了多种定义，在这些定义中分别使用了概率图灵机、多项式时间确定性等概念，我们这里不详细叙述。下面只以 RSA 公钥密码体制作为例子来演示计算安全性。

设 p、q 是两个大素数，$n=pq$。设 e、d 是两个正整数，$ed\equiv 1 \mod (p-1)(q-1)$。设 Alice、Bob、Eve 三人都知道（$n$，$e$），只有 Bob 知道 d，无人知道（p，q）。因此（n，e）为 Bob 的公钥，d 为 Bob 的私钥。当 Alice 欲向 Bob 发送明文消息 X 时，她计算密文

$$Y\equiv X^e \bmod n$$

并将密文 Y 发送给 Bob；Bob 计算明文

$$X\equiv Y^d \bmod n$$

此时 Eve 能够截获密文 Y，但由于不知道素分解 $n=pq$，因此没有有效的算法由（n，e）得到 d，虽然（n，e）惟一确定了 d。这样，Eve 对明文 X 的估计近乎于盲目地随机猜测。当然，如果知道素分解 $n=pq$，由（p，q，e）得到 d 是很容易的，只需要使用欧几里德算法和孙子定理。

1.4 完善保密性与计算安全性的比较

以下是完善保密性与计算安全性的比较结果：

(1) 计算安全性的安全强度弱于完善保密性。

(2) 从当前来看，计算安全性比完善保密性要现实得多。具有计算安全性的密码体制允许相同密钥的重复使用，从而大大减少了密钥协调所需的通信量。

(3) 完善保密的密码体制，由于拥有源源不断的更新密钥，因此不需要精心地设计密码函数，只需要简单的群运算来加密和解密即可。具有计算安全性的密码体制则不然。如果加密算法是简单的群运算 $Y=X\otimes Z$，则 $Z=X^{-1}\otimes Y$。这就是说，当 Eve 进行已知明文攻击时，只需要一组明文/密文对（X，Y）就解出了密钥。具有计算安全性的密码体制要能抵抗已知明文攻击，即当 Eve 已知 k 组明文/密文对：(X_1, Y_1)、(X_2, Y_2)、…、(X_k, Y_k)时，虽然方程组

$$Y_j=E_Z(X_j),\ j=1\sim k$$

可能已经完全确定了密钥 Z，但无法将 Z 解出。这意味着密码函数 $E_Z(\cdot)$和 $D_Z(\cdot)$需要精心地设计。

(4) 长期以来由于完善保密难以实现，故人们在信息保密方面的努力主要集中在计算保密上。近年来形势渐渐有所改变，显示出以下两个特征：①人类计算能力越来越强，其中包括芯片技术的飞速发展和量子计算机的问世，这一切预示着“计算安全”似乎越来越不可靠。②大量有扰信道（比如卫星信道和广播信道）的开通，使得通信伙伴之间能够共享源源不断的互信息；只要使用信息处理技术，将这些互信息中敌手已知的部分去掉，保留并协调敌手未知的部分，通信伙伴之间就获得了源源不断的密钥流，因而实现了一次一密，达到“完善保密”。从这个观点来看，完善保密性要比计算安全性优越。

1.5 完善保密性的获得　密钥协商

获得完善保密性，指的是要获得源源不断的共享密钥。我们总是假设通信伙伴之间共享着源源不断的信息，比如新闻、音乐等；通信伙伴各自掌握的数据可能有差别，且这些信息可能也在与敌手共享。要将这些共享的信息剪辑成源源不断的共享密钥，需要相互协调、纠错，并去除敌手已知的部分。这项工作分为优先提取、信息协调、保密增强等几个部分。

1.5.1 优先提取

优先提取协议的功能是使通信伙伴之间共享的信息被优先纠错，该协议的作用原理是：通信伙伴之间是能够相互协商的，而敌手只能被动地截听。M. J. Gander 和 U. M. Maurer[1, 2] 提出了一个比特对检验协议，这是迄今为止最有效的优先提取协议。协议描述如下。

设 Alice 拥有一个对称分布无记忆比特串 $X=X_1X_2\cdots$（即随机变量 X_1、X_2、…相互独立，各自等概地取值 0 和 1）。

比特串 X 通过一个 BSC 信道（二元对称无记忆信道）发送给 Bob，Bob 收到比特串为 $Y=Y_1Y_2\cdots$，向 Bob 发送的误比特率为 p，即有条件概率 $P(Y_k\neq x \mid X_k=x)=p$，$0<p<\frac{1}{2}$。

Eve 通过另一个独立的 BSC 信道截获比特串 X，Eve 收到比特串为 $Z=Z_1Z_2\cdots$，截获的误比特率为 q，即有条件概率 $P(Z_k\neq x \mid X_k=x)=q$，$0<q<\frac{1}{2}$。

然后 Alice 和 Bob 通过一条无扰的可认证信道交换信息（Eve 在这条信道上只能截听）。

对于 $k=1,2,\cdots$：

Alice 计算 $U_k=X_{2k-1}+X_{2k}$，并将 U_k 发送给 Bob；（注意此时的加法运算是比特异或）

Bob 计算 $V_k=Y_{2k-1}+Y_{2k}$，并将 V_k 发送给 Alice；

如果 $V_k=U_k$，则 Alice 删去自己比特串中的 X_{2k}，Bob 删去自己比特串中的 Y_{2k}；此时 Eve 也不得不删去自己比特串中的 Z_{2k}；

如果 $V_k\neq U_k$，则 Alice 删去自己比特串中的 $X_{2k-1}X_{2k}$，Bob 删去自己比特串中的 $Y_{2k-1}Y_{2k}$；此时 Eve 也不得不删去自己比特串中的 $Z_{2k-1}Z_{2k}$。

这样 Alice 获得了删减后的比特串 $X^{(1)}=X_1^{(1)}X_2^{(1)}\cdots$；Bob 获得了删减后的比特串 $Y^{(1)}=Y_1^{(1)}Y_2^{(1)}\cdots$；Eve 也通过截听对应地获得了删减后的比特串 $Z^{(1)}=Z_1^{(1)}Z_2^{(1)}\cdots$。这一过程称为比特对检验协议。

定理 1-2 设比特串 $X^{(1)}$与 $Y^{(1)}$ 的误比特率为 $p^{(1)}$，比特串 $X^{(1)}$与 $Z^{(1)}$ 的误比特率为 $q^{(1)}$。则

$$p^{(1)}=\frac{p^2}{(1-p)^2+p^2}<p;\quad q^{(1)}=q$$

证明 设 $\oplus$ 为比特异或运算。取 Alice 的比特串 $U=U_1U_2\cdots$，Bob 的比特串 $V=V_1V_2\cdots$，Eve 的比特串 $W=W_1W_2\cdots$，其中 $U_k=X_{2k-1}+X_{2k}$，$V_k=Y_{2k-1}+Y_{2k}$，$W_k=Z_{2k-1}+Z_{2k}$。首先注意到，以下九个比特串

$$X,\ Y,\ Z,\ U,\ V,\ W,\ X^{(1)},\ Y^{(1)},\ Z^{(1)}$$

都是对称分布无记忆比特串（即各位相互独立，各自等概地取值 0 和 1）。

比特串 U 与 V 的误比特率 r 为以下的条件概率

$$\begin{aligned}r&=P(Y_{2k-1}+Y_{2k}\neq x \mid X_{2k-1}+X_{2k}=x)\\&=P(Y_{2k-1}Y_{2k}\text{与}X_{2k-1}X_{2k}\text{之间误比特的个数为 1} \mid X_{2k-1}+X_{2k}=x)\\&=P(Y_{2k-1}Y_{2k}\text{与}X_{2k-1}X_{2k}\text{之间误比特的个数为 1})\quad\text{（因为是 BSC 信道）}\\&=2p(1-p)\end{aligned}$$

同理，比特串 U 与 W 的误比特率 s 为条件概率：$s=2q(1-q)$。

以下计算 $X^{(1)}$ 与 $Y^{(1)}$ 的误比特率 $p^{(1)}$，以及 $X^{(1)}$ 与 $Z^{(1)}$ 的误比特率 $q^{(1)}$。

$$p^{(1)}=P(Y_1^{(1)}\neq x \mid X_1^{(1)}=x)$$

$$=\sum_{k=1}^{+\infty}P(Y_1^{(1)}\neq x;U_i\neq V_i,i=1\sim k-1;U_k=V_k\Big|X_1^{(1)}=x)$$

$$=2\sum_{k=1}^{+\infty}P(Y_1^{(1)}\neq x;U_i\neq V_i,i=1\sim k-1;U_k=V_k;X_1^{(1)}=x)$$

$$=2\sum_{k=1}^{+\infty}P(Y_{2k-1}\neq x;U_i\neq V_i,i=1\sim k-1;X_{2k-1}+X_{2k}=Y_{2k-1}+Y_{2k};X_{2k-1}=x)$$

$$=2\sum_{k=1}^{+\infty}P(U_i\neq V_i,i=1\sim k-1;X_{2k-1}\neq Y_{2k-1};X_{2k}\neq Y_{2k};X_{2k-1}=x)$$

$$=2\sum_{k=1}^{+\infty}P(U_i\neq V_i,i=1\sim k-1)P(X_{2k-1}\neq Y_{2k-1};X_{2k}\neq Y_{2k};X_{2k-1}=x)$$

$$=\sum_{k=1}^{+\infty}P(U_i\neq V_i,i=1\sim k-1)P(X_{2k-1}\neq Y_{2k-1};X_{2k}\neq Y_{2k})\quad\text{（因为 BSC 信道）}$$

$$=p^2\sum_{k=1}^{+\infty}P(U_i\neq V_i,i=1\sim k-1)$$

$$=p^2\sum_{k=1}^{+\infty}(2p(1-p))^{k-1}=\frac{p^2}{(1-p)^2+p^2}<p$$

注意到比特串 Y 与 Z 关于比特串 X 条件独立（这是因为比特串 X 分别通过两个相互独立的 BSC 信道发送，分别收到比特串 Y 和比特串 Z），因此有

$$q^{(1)}=P(Z_1^{(1)}\neq x \mid X_1^{(1)}=x)$$

$$=\sum_{k=1}^{+\infty}P(Z_1^{(1)}\neq x;U_i\neq V_i,i=1\sim k-1;U_k=V_k\Big|X_1^{(1)}=x)$$

$$=2\sum_{k=1}^{+\infty}P(Z_1^{(1)}\neq x;U_i\neq V_i,i=1\sim k-1;U_k=V_k;X_1^{(1)}=x)$$

$$=2\sum_{k=1}^{+\infty}P(Z_{2k-1}\neq x;U_i\neq V_i,i=1\sim k-1;U_k=V_k;X_{2k-1}=x)$$

$$=2\sum_{k=1}^{+\infty}P(U_i\neq V_i,i=1\sim k-1)P(U_k=V_k;X_{2k-1}=x;Z_{2k-1}\neq x)$$