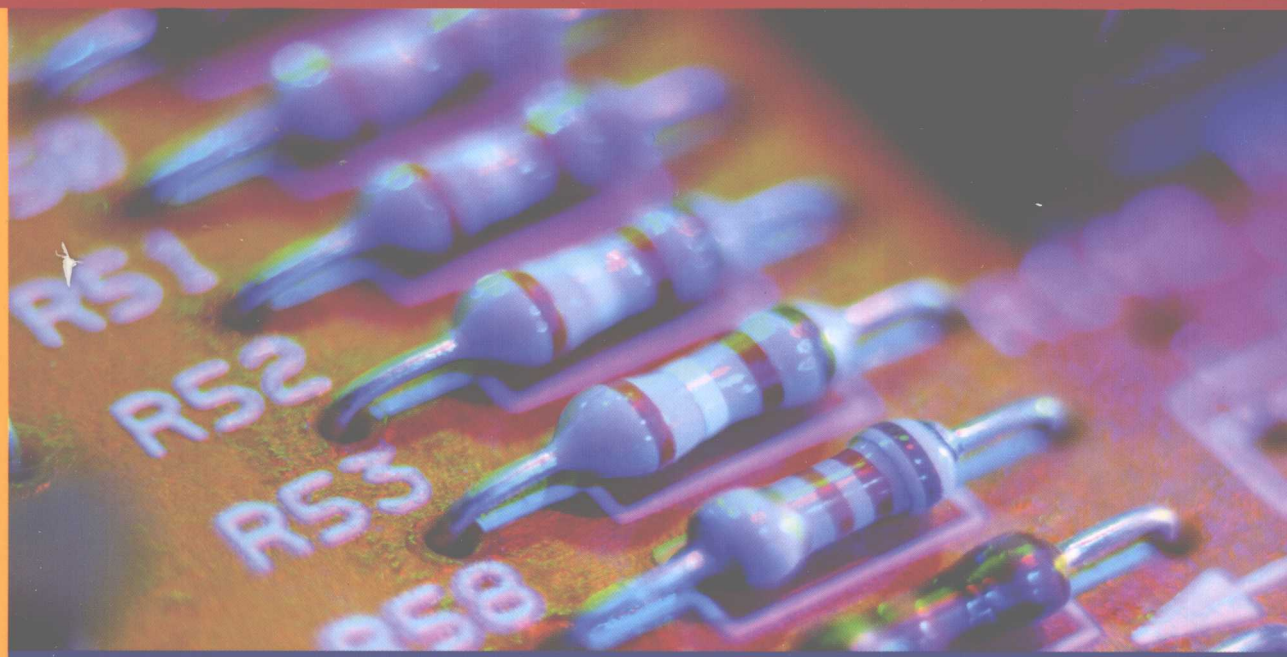
高等院校信息与通信工程系列教材

密码学理论与技术



范明钰 王光卫 编著

清华大学出版社

高等院校信息与通信工程系列教材

密码学理论与技术

范明钰 王光卫 编著

清华大学出版社
北京

内 容 简 介

本书从基本概念入手，通过手工密码算法、机械密码算法，建立密码算法的概念；从算法的设计和分析两条线索，指出密码学的对抗和发展状况。全书主要分为以下4个部分：第1部分介绍与密码学相关的基本概念(第1章)。第2部分介绍古典密码学，重点介绍古典密码设计和分析留下的经验和教训(第2、3章)。第3部分介绍现代密码，包括对称密码中的分组密码(第4章)和序列密码(第5章)、Hash算法(第6章)和公钥密码(第7章)。第4部分介绍密码算法的使用与发展，包括密钥管理过程(第8章)和网络时代的密码(第9章)。

本书可供工科类计算机、电子信息、通信等学科的本科学生和研究生使用。

图书在版编目(CIP)数据

密码学理论与技术/范明钰，王光卫编著. —北京：清华大学出版社，2008.10
(高等院校信息与通信工程系列教材)
ISBN 978-7-302-18195-8

Ⅰ. 密…　Ⅱ. ①范…　②王…　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108262号

责任编辑：陈国新
责任校对：梁　毅
责任印制：李红英
出版发行：清华大学出版社　　**地　　址**：北京清华大学学研大厦A座
http://www.tup.com.cn　　**邮　　编**：100084
社　总　机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质 量 反 馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：北京国马印刷厂
经　　销：全国新华书店
开　　本：185×260　**印　张**：11.5　**字　数**：278千字
版　　次：2008年10月第1版　　**印　　次**：2008年10月第1次印刷
印　　数：1～3000
定　　价：23.00元

出版说明

信息与通信工程学科是信息科学与技术的重要组成部分。改革开放以来，我国在发展通信系统与信息系统方面取得了长足的进步，形成了巨大的产业与市场，如我国的电话网络规模已位居世界首位，同时该领域的一些分支学科出现了为国际认可的技术创新，得到了迅猛的发展。为满足国家对高层次人才的迫切需求，当前国内大量高等学校设有信息与通信工程学科的院系或专业，培养大量的本科生与研究生。为适应学科知识不断更新的发展态势，他们迫切需要内容新颖又符合教改要求的教材和教学参考书。此外，大量的科研人员与工程技术人员也迫切需要学习、了解、掌握信息与通信工程学科领域的基础理论与较为系统的前沿专业知识。为了满足这些读者对高质量图书的渴求，清华大学出版社组织国内信息与通信工程国家级重点学科的教学与科研骨干以及本领域的一些知名学者、学术带头人编写了这套高等院校信息与通信工程系列教材。

该套教材以本科电子信息工程、通信工程专业的专业必修课程教材为主，同时包含一些反映学科发展前沿的本科选修课程教材和研究生教学用书。为了保证教材的出版质量，清华大学出版社不仅约请国内一流专家参与了丛书的选题规划，而且每本书在出版前都组织全国重点高校的骨干教师对作者的编写大纲和书稿进行了认真审核。

祝愿《高等院校信息与通信工程系列教材》为我国培养与造就信息与通信工程领域的高素质科技人才，推动信息科学的发展与进步做出贡献。

北京邮电大学

陈俊亮

2004 年 9 月

前　言

本书主要作为密码学课程或信息安全类专业的本科生、研究生的教学参考书。

全书从基本概念入手，通过手工密码算法、机械密码算法建立了密码算法的概念。依照算法的设计和分析两条线索，指出了密码学的对抗和发展状况。

在现代密码学部分介绍了分组密码和序列密码的基本原理、设计和分析方法；在公钥密码算法部分，介绍了公钥密码体制的数学原理和公钥算法的使用方法。

在最后两章中分析了密钥管理技术和网络给密码学带来的挑战和发展机遇。

本书每章都附有习题和思考题。

本书共分为 4 个部分：

第 1 部分介绍与密码学相关的基本概念(第 1 章)。

第 2 部分介绍古典密码学，包括手工密码和机械密码，重点介绍古典密码设计和分析留下的经验和教训(第 2、3 章)。

第 3 部分介绍现代密码学，首先介绍对称密码中的分组密码(第 4 章)和序列密码(第 5 章)，重点介绍算法概念、设计原理和分析；其次介绍 Hash 算法(第 6 章)，重点分析其原理；最后介绍公钥密码(第 7 章)。

第 4 部分介绍密码算法的使用和发展，首先介绍密钥管理过程(第 8 章)，重点介绍分组密码算法的工作模式、密钥管理体制；其次介绍网络时代的密码(第 9 章)，重点介绍网络的发展给密码学带来的挑战、在网络中密码技术的使用以及网络时代给信息保护带来的新课题。

本书的 4 个部分基本上是前后相互关联，但也相互独立。内容由浅入深没有重叠，因而既可以按顺序从概念入手学习，也可以先从实例开始而后理论。

本书的编写历时近 4 年，在此过程中，多次得到魏正耀院士的关心和指导，书中的主要思路也得到了魏院士的指导，在此表示诚挚的谢意。

参与本书编写的主要人员有朱大勇博士、王庆先博士以及实验室学生李欣、周文锦、沈丹。其中第 2 章和第 5 章的大部分内容是李欣同学编写的，第 7 章主要内容是周文锦同学编写的，书中大部分习题是沈丹同学编写的。谨在这里向他们表示诚挚的谢意。

编　者

于电子科技大学

目　录

第 1 章　密码学中的基本概念

本章主要内容是密码学的基本概念，包括密码学的应用、密码算法的基本概念、密码编码学和分析学中的基本概念、密码学的信息论基础、密码学的起源和发展以及密码算法的安全性和复杂性的概念。

密码学在公元前 400 多年就已经产生了，正如《破译者》一书中所言，人类使用密码的历史几乎与使用文字的时间一样长。

密码学的起源可以追溯到人类刚刚出现，并且尝试去学习如何通信的时候，为了确保相互间通信的机密，开始是有意识地使用一些简单的方法来加密信息，如通过一些（密码）象形文字相互传达信息。接着由于表音和表意文字的出现和使用，确保通信的机密性就成为一种艺术。而随着国家、政权、军事力量的建立，密码学在重要信息的交流传递方面起到了越来越重要的作用。

随着数字化和网络技术不断深入到社会各个方面，人们对信息安全的重要性认识不断提高，而在信息安全中起着举足轻重作用的密码学也就成为信息安全中不可或缺的重要部分。今天密码学已经逐步揭开了神秘的面纱，进入了寻常百姓的日常生活之中。密码学的研究应用前景十分广阔，这个总是秘而不宣的重要角色，在人类的发展中将起到不可估量的作用。

当今世界各主要国家的政府都十分重视密码工作，其中一些国家设立庞大机构，拨出巨额经费，集中数以万计的专家和科技人员，投入大量的高速电子计算机和其他先进设备进行工作。与此同时，各民间企业和学术界也对密码学日益重视，不少数学家、计算机学家和其他有关学科的专家也投身于密码学的研究行列，更加快了密码学的发展。

1.1　术语

密码学的英文名称为 cryptography，此英文单词为两个古希腊词根的组合：crypto（秘密）和 graphein（书写），意为密写，故密码学是一门研究秘密书写的科学，是以认识密码变换的本质、研究密码保密与破译的基本规律为对象的学科。

密码学的另一种定义是一门与信息安全密切相关的数学科学，是信息安全的核心。通俗地讲，密码学将信息表述为不可读的方式，但通过秘密的方法可将信息恢复出来。密码学提供的最基础的服务是使合法通信者进行信息的交互，而其他人员难以获得通信内容。

密码学一般包括两个对立统一的分支学科：密码编码学和密码分析学。研究密码变

化的规律并用之于编制密码以保护秘密信息的学科，称为密码编码学。研究密码变化的规律并用之于分析密码以获取信息情报的学科，称为密码分析学，也叫密码破译学。前者是实现对信息保密的，后者是实现对信息反保密的，密码编码学与密码分析学相反相成，共处于密码学的统一体中。

现代密码学除了包括密码编码学和密码分析学两个主要学科外，还包括一个新产生的分支——密码密钥学。它是以密码体系最核心部分的密钥作为研究对象的学科。密钥管理是一种规程，它包括密钥的产生、分配、存储、保护、销毁等环节，因而在密码体系中密钥管理至关重要。上述三个分支学科构成了现代密码学的主要学科体系。

1. 发送者和接收者

试图发送消息的一方称为发送者(sender)。

消息的接收对象称为接收者(receiver)。

发送者和接收者都称为用户。经过认可的用户为合法用户，否则为非法用户。

2. 明文和密文

尚未隐藏或未被加密的信息称为明文(plaintext)或消息(message)，用 P 或 M 表示。明文是未被隐藏的信息，即是发送者准备发送的原文信息。明文的集合称为明文信息空间，用 S_P 表示。

用某种方法伪装消息以隐藏它的内容的过程称为加密(encryption)。

被加密后的消息称为密文(ciphertext)，用 C 表示。所有的密文构成密文信息空间，用 S_C 表示。

3. 密钥

明文到密文的转换往往由一些特殊的函数完成，控制这些函数的参数称为密钥(key)，用 K 表示。所谓密钥，是指由用户事先选定的较短的字符或数字序列，其作用近似于打开保险箱的钥匙。所有密钥的集合构成密钥空间，用 S_K 表示。密钥空间中不相同密钥的个数称为密钥体制的密钥量，它是衡量密码体制安全性的一个重要指标。

密钥是一个数值，它和加密算法一起生成特别的密文。密钥本质上是非常非常大的数。密钥的长度尺寸用比特(bit)来衡量，1024bit 密钥代表的数是非常巨大的。在公开密钥加密方法中，密钥的长度越大，密文就越安全。

然而，公开密钥的尺寸和传统加密方法中密钥的尺寸是不相关的。传统 80bit 密钥的强度等同于 1024bit 的公钥，传统 128bit 密钥的强度等同于 3000bit 的公钥。在同种加密算法中，密钥越大越安全。但是传统方法和公开密钥方法所用的加密算法不一样，因此它们的密钥尺寸不能直接比较。

公钥和私钥是算术相关的，仅凭公钥推算出私钥是非常困难的。然而如果有足够的时间和计算能力，总是可能导出私钥的。这使得选择合适尺寸的密钥变得非常重要。为了安全需要足够大的密钥，而为了速度则要用小的密钥。

4. 加密与解密

加密是在密钥 K 的作用下，把明文 P 从明文信息空间 S_P 对应到密文信息空间 S_C 的一种变换，记该变换过程为

$$E_K : S_P \rightarrow S_C$$

则明文和密文的关系可表示为

$$C = E_K(P)$$

密文传送到接收者，合法用户利用密钥对密文 C 进行与加密变换相反的逆变换，称为解密变换，用 D_K 表示。解密变换是把密文 C 从密文信息空间 S_C 对应到明文信息空间 S_P 的变换为

$$D_K : S_C \rightarrow S_P$$

逆变换的过程称为解密或译密。解密变换的目的是恢复出明文 P：

$$P = D_K(C) = D_K[E_K(P)] \tag{1-1}$$

上式中的 E_K 和 D_K 为可逆变换对。K 不同，E_K 和 D_K 也不同。可见，信息的保密性完全依赖于密钥 K 的保密性。

5. 密码体制

一个完整的密码体制(cryptosystem)由 5 部分组成：

- 明文信息空间 S_P
- 密文信息空间 S_C
- 密钥空间 S_K
- 加密变换族 E_K
- 解密变换族 D_K

密码体制应满足以下 3 个一般性要求：

- 加、解密变换对所有密钥都一致有效。
- 体制必须是简单易行的，应易于找到密钥用于逆变换。
- 体制的安全性仅依赖于密钥的保密性而不能依赖于加、解密算法的强度。

一个好的密码体制则应至少满足以下两个条件：

- 在已知明文 P 和密钥 K 时，计算 $C=E_K(P)$容易；在已知密文 C 和密钥 K 时，计算 $P=D_K(C)$容易。
- 在不知密钥 K 时，不可能由密文 C 推知明文 P。

对于一个密码体制，如果能够根据密文确定明文或密钥，或者能够根据明文及其生成的密文确定密钥，这个密码体制就是可以破译的，反之则为不可破译的。

6. 鉴别、完整性和抗抵赖

除了提供机密性外，密码学在网络时代还有其他作用：

(1) 鉴别

解决身份冒充问题。消息的接收者应该能够确认消息的来源，发送者不可能伪装成

他人。同样，发送者也要能够确认接收者是否是自称的，接收者不可能伪装成他人。

(2) 完整性

解决篡改问题。消息的接收者应该能够验证在传送过程中消息没有被修改。

(3) 抗抵赖

发送者或接收者事后不可能虚假地否认他发送或接收到的消息。

这些功能使通过计算机进行的信息交流，就像面对面交流一样安全可靠。某人是否就是他说的人，某人的身份证明文件(驾驶执照、医学学历或者护照)是否有效，声称从某人那里来的文件是否确实是从那个人那里来的，这些事情都是通过鉴别、完整性检验和抗抵赖来实现的。

1.2 密码学的应用

由于在很长的时间内，密码仅限于军事、政治和外交的用途，密码学的知识和经验也仅掌握在与军事、政治和外交有关的密码机关手中，再加上通信手段比较落后，所以不论密码理论还是密码技术发展都很缓慢。

随着科学技术的进步，信息交换的手段越来越先进，信息交换的速度越来越快，信息交换的内容越来越广泛，信息交换的形式越来越多样化，信息交换的规模也越来越大。到了 20 世纪 70 年代，随着信息的激增，对信息保密的需求也从军事、政治和外交等领域，逐步扩展到民用和商用领域，从而导致了密码学知识的广泛传播。计算机技术和微电子技术的发展，为密码学理论的研究和实现提供了强有力的手段和工具。进入 20 世纪 80 年代以后，随着网络的兴起，对密码理论和技术的研究更是呈爆炸性增长的趋势，密码学在雷达、导航、遥控、遥测等领域占有重要地位。除此之外，密码学正渗透到通信、电力、金融、医疗、卫生、交通等各行业的管理信息系统，甚至到个人和家庭等领域，而且保密的作用也已不再仅仅是保密，还有认证、完整性检验和抗抵赖等新的功能。

对普通的家庭来说，生活中许多地方需要保密，如各种银行密码、信用卡密码、网络账号和密码等。

1.3 密码算法的概念及其分类

密码算法(algorithm)也叫密码(cipher)，是用于加密和解密的数学函数。通常情况下，有两个相关的函数，一个用作加密，另一个用作解密。

密码算法有多种分类方法。如果算法的保密性是依赖于保持算法的秘密，这种算法称为受限制的算法。受限制的算法具有历史意义，但按现在的标准，它们的保密性已远远不够。大的或人员经常变换的用户组织不能使用它们，因为每有一个用户离开这个组织，其他用户就必须改换另外不同的算法。如果有人无意暴露了这个秘密，所有人都必须改变他们的算法。

不利的是，受限制的密码算法不可能进行质量控制或标准化。每个用户组织必须有他们自己的唯一算法。这样的组织不可能采用流行的硬件或软件产品。由于窃听者可以

买到这些流行产品并学习这些算法，于是用户不得不自己编写算法并予以实现，如果这个组织中没有好的密码学家，那么他们就无法知道他们是否拥有安全的算法。

尽管有这些主要缺陷，受限制的算法对低密级的应用来说还是很流行的，用户或者没有认识到或者不在乎他们系统中内在的问题。

如果算法的保密性是依赖于保持密钥的秘密，这种算法称为非受限制的算法，也称基于密钥的算法。在非受限密码算法中，根据密钥的特点，加密算法分为两类：秘密密钥算法和公开密钥算法。如果没有特殊说明，在本书中，密码算法均指非受限制的算法。

1.3.1　对称密码算法

秘密密钥算法通常称之为对称密码算法或传统密码算法，也称单密钥算法。对称密码算法要求发送者和接收者在安全通信之前，商定一个密钥。其算法的安全性依赖于密钥，密钥一旦泄露，整个安全系统都要崩溃。换句话说，在使用对称密码加密算法时，只要通信需要保密，密钥就必须保密。

对称密码算法的加、解密可表示为

加密：$E_K(P)=C$

解密：$P=D_K(C)$

式中均使用同一密钥 K。

对称算法可分为两类。一类只对明文中的单个比特（有时对字节）运算的算法称为序列算法（stream algorithm）或序列密码（stream cipher）。另一类算法是对明文的一组比特运行运算，这些比特组称为分组（block），相应的算法称为分组算法（block algorithm）或分组密码（block cipher）。现代计算机密码算法的典型分组长度为 64 比特——这个长度大到足以防止分析破译，但又小到足以方便使用（在计算机出现前，算法普遍到每次只对明文的一个字符运算，可认为是序列密码对字符序列的运算）。

1.3.2　公开密钥算法

公开密钥加密法可以解决密钥发布的问题，公开密钥的概念由 Whitfield Diffie 和 Martin Hellman 在 1975 年提出。现在也有证据表明英国情报机关先于 Diffie 和 Hellman 几年发明了这种方法，但是却作为军事秘密不为人知，并且没有什么有价值的成果。

公开密钥算法也叫非对称密码算法或双密钥算法。

公开密钥算法不要求通信双方共享一个密钥，其用作加密的密钥不同于用作解密的密钥，而且解密密钥不能根据解密密钥计算出来。用于加密的密钥可向所有使用者公开，使用加密密钥加密后的信息只有用相对应的解密密钥才可解密。

公开加密算法可用下式表示：

加密：$E_{Ke}(P)=C$

解密：$P=D_{Kd}(C)$

其中 K_e 和 K_d 分别代表加密和解密密钥。

公钥加密法的主要优势在于可以让事先没有安全通道的人安全地交换信息。收、发

双方通过安全通道共享密钥的前提条件不存在了，所有的通信中只包含了公钥，私钥是不会传输或共享的。

由于传统的加密方法曾经是传送秘密信息的唯一手段，保持安全通道和发布密钥的高昂费用将其应用范围限制在能够负担得起这些费用的用户中，比如政府或者大银行。公钥加密法是加密技术的革命，它可以为普通人提供较强的加密手段，因而可以说公钥加密法是密码学发展史上的一个里程碑。

1.3.3 Hash 算法

Hash 算法也称为消息摘要或单向函数，是密码学中的一种重要的算法。它是许多安全认证协议的重要组成部分，是实现有效、安全可靠的数字签名和认证的重要工具。

Hash 算法是一种将任意长度的输入消息计算产生出一个固定长度的输出的数学变换，即消息 m 的 Hash 为 $h(m)$。这类算法具有以下特点：

- 对于任何消息，计算 $h(m)$相对来说较为容易，这意味着要使用该函数，并不需要占用太多的计算时间。
- 给定 $h(m)$，寻找一个消息使得其 Hash 值为 $h(m)$的难度与穷举所有可能的 m 并计算 $h(m)$的难度相比，不会有明显的差别。
- 虽然理论上存在很多不同数值，其 Hash 值都是 $h(m)$，但是要找到两个 Hash 结果相同的数值，从计算的角度来说是很困难的。

要从密码学的角度认为一个 Hash 函数是安全的，其必备条件如下：

- 找到一个消息，使其消息摘要为一预先给定的消息摘要值，在计算上是不可行的。
- 以现有的计算能力，不可能找到两个具有相同消息摘要的消息。
- 给定一个消息，不可能找到另一个消息与其具有相同的消息摘要。

1.4 密码编码学的基本概念

密码编码学的主要目的是确保明文、密钥等秘密信息不被窃听者或攻击者窃听或破译。这里有一个前提假设，即以上非法人员完全能够截获收发者之间的通信。

密码编码学希望能够解决在下述环境下即信息的存储（可能为非授权者接触）、信息的交换（可能被冒用或抵赖）以及信息的传输（可能被截获）过程中，信息的安全保护问题。

密码编码系统应具有以下独立的特征：

（1）转换明文为密文的运算类型

有的算法的保密性可能依赖于保护算法本身，称为受限制的密码算法；也有的算法仅依赖于使用算法时所采用的密钥，称为基于密钥的密码算法。本书仅讨论基于密钥的密码算法。

（2）所用的密钥类型

如果发送方和接收方使用相同的密钥，这种密码称为对称密码、单密钥密码或传统密码。如果发、收双方使用不同的密钥，这种密码就称为非对称密码、双钥密码或公钥密码。