密码学讲义

李 超 屈龙江 编著

科学出版社
www.sciencep.com

密码学讲义

李 超 屈龙江 编著

科学出版社
北京

内 容 简 介

本书从数学的角度较为系统地介绍了序列密码、分组密码和公钥密码的基本理论与方法. 利用周期序列的幂级数表示、根表示和迹表示研究了线性反馈移位寄存器序列及其变种的密码学性质; 利用图论和组合数学等工具研究了非线性反馈移位寄存器序列的状态图性质, 重点介绍 M 序列的存在性、构造与计数; 介绍了五类典型分组密码算法的加解密流程、分组密码的设计原理以及一些常见的分析方法; 讨论了 RSA 体制和椭圆曲线密码体制的基本原理及其相关的数学问题.

本书可以作为密码学与信息安全专业的本科生和研究生的教学用书, 也可以作为从事密码学和信息安全研究的科技人员的参考书.

图书在版编目(CIP)数据

密码学讲义/李超, 屈龙江编著. —北京: 科学出版社, 2010

ISBN 978-7-03-026385-8

Ⅰ. ① 密… Ⅱ. ① 李… ② 屈… Ⅲ. ① 密码-理论 Ⅳ. TN918.1

中国版本图书馆 CIP 数据核字(2010) 第 007025 号

责任编辑: 赵彦超 / 责任校对: 刘小梅

责任印制: 钱玉芬 / 封面设计: 黄华斌

科学出版社 出版

北京东黄城根北街 16 号

邮政编码: 100717

http://www.sciencep.com

新蕾印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2010 年 2 月第 一 版 开本: B5 (720 × 1000)

2010 年 2 月第一次印刷 印张: 12 3/4

印数: 1—30 00 字数: 247 000

定价: 38.00 元

(如有印装质量问题, 我社负责调换)

前　言

随着美国 AES 计划、欧洲 NESSIE 计划和 eSTREAM 计划的实施, 密码学的理论与方法研究遇到了前所未有的发展机遇和挑战. 如何设计安全、高效的对称密码体制和非对称密码体制, 已经成为信息安全领域中的重要问题.

序列密码是一类重要的对称密码体制, 它在加解密速度和硬件实现规模两方面具有明显优势, 非常适合在大量数据传输和资源受限的场合使用. 早期的序列密码研究主要围绕线性反馈移位寄存器序列及其扩展序列展开, 这些扩展序列包括前馈序列、非线性组合序列和钟控序列等. 由于线性问题具备一些良好的数学理论和方法作为支撑, 线性反馈移位寄存器序列及其扩展序列的密码学性质的研究取得了非常丰富的成果. 相比之下, 非线性反馈移位寄存器序列的密码学性质的研究所取得的成果就很少. 特别是 2004 年欧洲 eSTREAM 计划推出的序列密码算法, 大都采用非线性驱动和非线性迭代, 使得对这些序列密码算法的理论研究遇到了重大的挑战. 如何从理论上刻画 eSTREAM 计划获胜算法的安全性, 是国际密码学者需要解决的重要问题. 本书第一部分内容就是利用 Galois 域的基本理论与方法, 较为系统地研究 Galois 域上线性反馈移位寄存器序列及其扩展序列的密码学性质, 即本书的第 2 章和第 3 章. 然后利用图论和组合数学等数学工具研究 Galois 域上非线性反馈移位寄存器序列的一些密码学性质, 即第 4 章. 系统地掌握这些经典的序列密码设计与分析理论和方法, 对于进一步研究 eSTREAM 计划中的现代序列密码具有重要的意义.

分组密码是另一类重要的对称密码体制, 是保障信息机密性和完整性的重要技术. 由于与序列密码的实现机理不同, 分组密码的研究主要围绕分组密码的设计、分析、工作模式、快速实现和检测等方面展开. 分组密码的设计与分析是一对既相互对立又相互统一的矛盾体, 二者的互动决定了分组密码的发展. 分组密码的安全性分析为分组密码的设计提供了源源不断的新鲜思想, 而各种深思熟虑的设计又给分组密码的分析提出了严峻的挑战. 只有对分组密码分析具有深刻的理解和敏锐的感觉, 才有可能设计出安全有效的分组密码. 分组密码的工作模式是利用分组密码解决实际问题的密码方案, 直接影响分组密码在实际应用中的安全性和有效性. 分组密码的快速实现和检测是评估分组密码的重要组成部分, 为分组密码的设计、分析和应用提供重要的量化指标和技术参数. 美国 AES 计划和欧洲 NESSIE 计划的实施极大地推动了分组密码理论与方法的发展, 使得分组密码的研究从经验设计走向了理论设计的道路. 本书的第 5~7 章主要介绍分组密码的设计原理和分析方法,

以及基于这些设计原理所构造的一些典型的分组密码算法, 其目的是使读者掌握分组密码设计与分析的基本理论和方法, 以便对分组密码的设计与分析展开进一步的研究.

公钥密码是不同于序列密码、分组密码的密码体制, 其主要特点是将密码体制的加密密钥与解密密钥分开, 加密密钥可以公开, 而解密密钥需要保密. 公钥密码的设计依赖于计算困难的数学问题, 这些问题包括大整数因子分解问题、离散对数问题、格中向量问题、纠错码的译码问题、多变量多项式方程组的求解等. 自从 20 世纪 70 年代末公钥密码体制提出以来, 公钥密码的研究主要集中在公钥密码体制的高效实现及其所依赖的数学问题求解两个方面. 本书最后一章将较为详细地讨论 RSA 体制和椭圆曲线密码体制的加解密原理以及与这两种密码体制安全性有关的数学问题求解, 其目的是使读者掌握 RSA 体制和椭圆曲线密码体制的基本理论与方法, 因为这两类体制目前已经成为许多国家和地区的公钥密码标准算法.

本书根据第一作者 2002 年编写的密码学讲义改编而成, 原讲义已经在国防科技大学编码密码理论及其应用方向的研究生课程教学中使用多年. 本书得以出版, 首先感谢国防科技大学各部门的支持, 特别是国防科技大学研究生院教材建设基金和理学院数学学科建设基金的资助; 其次, 科学出版社的责任编辑为本书的出版做了大量的工作, 在此表示衷心感谢.

由于作者水平有限, 时间仓促, 书中不妥之处在所难免, 恳请读者指正.

编　者

2009 年 12 月于长沙

目　录

第1章 绪 论

随着信息时代的到来, 信息的传输、变换、压缩和存储等信息处理的有效性、可靠性和安全性已经成为当今信息处理中亟待解决的重要问题. 信息安全技术是维护信息保密性、完整性和可靠性的重要手段. 密码技术是信息安全的核心技术, 涉及密码算法的设计和分析、身份认证、数字签名、密钥管理等多种安全范围. 密码学是以密码技术作为研究对象的一门学科, 密码技术的研究与应用已有四千多年的历史, 但它作为一门科学却是 20 世纪 40 年代的事. 1948 年, Shannon 发表了一篇划时代的文章 "通信中的数学理论"[1], 将信息技术建立在严格的数学理论基础之上, 创立了 "信息论". 第二年, Shannon 又发表了另一篇重要文章 "保密系统中的通信理论"[2], 为密码学的发展奠定了理论基础, 使得密码学的研究与应用步入了科学的发展轨道, 产生了 "密码学".

1.1 密码学的基本概念

密码学是研究密码技术的一门学科, 密码技术主要涉及密码体制的设计、分析和应用. 密码体制通常用来隐蔽和保护需要保密的消息, 使未授权者不能提取信息. 被隐蔽的消息称作明文, 所有明文的集合一般用 P 表示. 隐蔽后的消息称作密文, 所有密文的集合记为 C. 将明文变换成密文的过程称为加密变换, 加密变换通常是在加密密钥的控制下进行的, 全体加密密钥的集合记为 K_1, 全体加密变换的集合记为 E. 对任意的加密密钥 $k_1 \in K_1$, 都唯一确定一个加密变换 $E_{k_1} \in E$. 类似地, 由密文恢复出明文的过程称为解密变换, 解密变换是在解密密钥的控制下进行的, 全体解密密钥的集合记为 K_2, 全体解密变换的集合记为 D. 对任意的解密密钥 $k_2 \in K_2$, 都唯一确定一个解密变换 $D_{k_2} \in D$. 密码体制的一般模型如图 1.1 所示.

从数学的角度来说, 一个密码体制就是指满足以下条件的六元组 (P, C, K_1, K_2, E, D):

(1) P 表示所有可能的明文组成的有限集;

(2) C 表示所有可能的密文组成的有限集;

(3) K_1 表示所有可能的加密密钥组成的有限集, K_2 表示所有可能的解密密钥组成的有限集;

(4) 对任意的 $k_1 \in K_1$, 都存在一个加密变换 $E_{k_1} \in E$; 同样, 对任意 $k_2 \in K_2$, 都存在一个解密变换 $D_{k_2} \in D$;

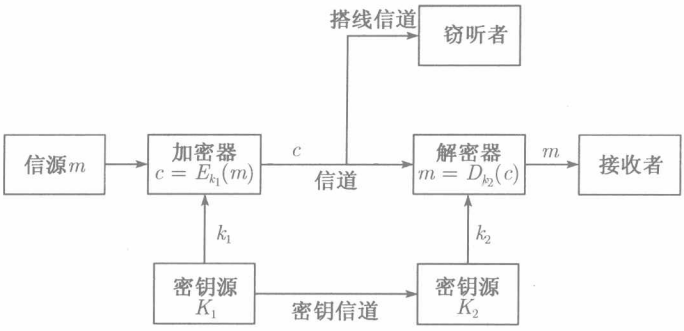

图 1.1　密码体制的一般模型

(5) 对任意的 $k_1 \in K_1(k_1$ 确定一个加密变换 $E_{k_1} \in E)$, 均存在唯一的 $k_2 \in K_2(k_2$ 确定一个解密变换 $D_{k_2} \in D)$, 使得对任意的 $m \in P$, 均有 $D_{k_2}(E_{k_1}(m)) = m$.

对一个密码体制来说, 如果加密密钥与其相应的解密密钥实质相同 (简称相同), 即加密密钥与解密密钥可以容易地相互推出, 就称该密码体制为对称密码体制或者私钥密码体制, 这里 "对称" 的意思是指加密密钥和解密密钥的地位是对称的, "私钥" 的意思是指在使用私钥密码体制的过程中加密密钥和解密密钥一定得保密, 不能让未授权者获知. 相反地, 如果加密密钥与其相应的解密密钥实质不同 (简称不同), 即由加密密钥求解密密钥或者由解密密钥求加密密钥都是十分困难的, 就称该密码体制为非对称密码体制或者公钥密码体制. 顾名思义, 在公钥密码体制中, 加密密钥和解密密钥的地位是不对称的, 加密密钥可以公开, 而解密密钥需要保密. 私钥密码体制根据加密方式的不同又可以分为序列密码体制和分组密码体制. 将明文消息按字符逐位加密者称为序列密码体制, 简称序列密码. 将明文消息分组并逐组进行加密者称为分组密码体制, 简称分组密码.

从密码学的发展历史来看, 密码学经历了四个重要的发展时期.

1949 年以前是科学密码学的前夜时期. 这一阶段主要是密码技术的简单应用, 密码的设计与分析大都是密码专家凭直觉进行的, 缺乏必要的理论支持. 以 Vigenere 密码和 Hill 密码为代表的古典密码体制现在看来都是不安全的.

1949~1975 年是密码学发展的神秘时期. 这一时期的密码学充满神秘色彩, 密码学研究大都集中在军队和国防安全部门, 密码学的成果也被禁锢在这些国家安全部门, 公开发表的密码学成果并不多见, 以移位寄存器序列为代表的序列密码在这一时期得到了较快发展.

1976~1996 年是密码学发展的关键时期. 这一时期密码学发生了几件大事, 1976 年, Diffie 和 Hellman 发表了 "密码学的新方向"[3], 提出了公钥密码的思想, 产生了公钥密码学; 1977 年, 美国国家标准局公布了数据加密标准 DES(Data Encryption Standard)[4], 揭开了密码学的神秘面纱, 推动了分组密码的发展; 1978 年,

Rivest, Shamir 和 Adleman 提出了第一个实用的公钥密码体制 ——RSA 体制 [5], 极大地促进了公钥密码的发展.

1997 年至今是密码学发展的辉煌时期. 经历了近 50 年的积累, 密码学在世纪之交步入了一个高速发展的时期. 1997 年, 美国国家标准技术研究所 (National Institute of Standard and Technology, NIST) 推出 AES(Advanced Encryption Standard) 计划 [6], 在世界范围内征集美国的高级加密标准, 以代替 DES 算法. 2000 年, NIST 推选比利时人 Daemen 和 Rijmen 设计的 Rijndael 算法 [7] 为高级加密标准. 继美国 AES 计划以后, 欧洲相继启动了 NESSIE 计划和 ECRYPT 计划 [8,9], 在世界范围内征集欧洲新世纪的各类密码标准, 包括序列密码标准、分组密码标准、公钥密码标准和数字签名标准等. 这些计划的兴起, 使得密码设计与分析从经验设计走上理论设计的道路, 密码学理论与方法得到重大发展.

从密码学的研究内容来看, 密码学包括密码编码学和密码分析学两个分支. 密码编码学主要研究密码体制的设计理论与方法, 重点解决信息的保密性问题, 即研究如何设计一个密码体制来对明文信息进行加解密处理. 密码分析学主要研究密码体制的分析理论与方法, 重点解决加密信息的破译问题, 即研究在不知道解密密钥的情况下, 如何从密文中恢复明文消息. 密码体制的安全性涉及多方面的因素, 比如所采用的密码算法的安全强度、密钥管理的安全措施、密码体制的工作模式等. 密码算法的安全强度是保证密码体制安全性的根本要素, 因此密码算法的设计一直是密码编码学研究中的主要问题. 序列密码的设计就是在密钥的控制下寻找一种生成方式, 产生具有类似随机序列特性的密钥流序列, 用来对明文信息序列进行加解密变换; 分组密码的设计就是找到一种算法, 能在密钥的控制下从一个足够大且足够好的置换子集合中简单而迅速地选出一个置换, 用来对当前的明文进行加解密变换. 目前人们在设计一个密码时, 大都基于如下的 Kerckhoff 假设: 密码体制的安全性应该不依赖于对密码算法的保密, 而只依赖于对密钥的保密. 尽管在不知道密码算法的情况下, 恢复密码的密钥是一件更加困难的事情, 但作为密码设计者来说, 不能低估密码分析者的能力, 在设计一个密码算法并对它的安全性进行评估时, 应该假定密码算法是公开的.

另一方面, 作为密码分析者来说, 其任务就是在不知道密钥的情况下对密码体制进行攻击, 试图获取密钥. 在 Kerckhoff 假设下, 可以将密码攻击分为如下四种类型:

唯密文攻击: 密码分析者拥有一个或更多用同一个密钥加密的密文, 通过对这些截获的密文进行分析得出明文或密钥.

已知明文攻击: 除待解密的密文外, 密码分析者拥有一些明文和用同一个密钥加密这些明文的密文, 通过对这些已知明文和相应密文的分析来恢复密钥.

选择明文攻击: 密码分析者可以随意选择自己想要的明文并加密, 根据选择的

明文和相应的密文来恢复密钥.

选择密文攻击：密码分析者可以随意选择自己想要的密文并解密, 根据选择的密文和相应的明文来恢复密钥.

在对密码进行分析的过程中, 存在两种分析或破译的概念, 一种是 "现实破译", 另一种是 "理论破译". 如果一个分析算法是实际可行的, 即用它在现有的计算资源下, 可以完全恢复一个密码的明文或密钥, 就说该分析方法现实破译了该密码. 对于一个 n 比特密钥的密码来说, 如果一个分析算法可以在低于 2^n 次加密后找到正确的密钥, 则称该算法理论破译了这个密码. 从历史上看, 有很多密码都是先被理论破译, 然后在理论破译的指导下, 最终被现实破译. 另外, 从硬件实现的角度看, 如果一个 128 比特密钥的密码可以在 2^{100} 次加密运算内被理论破译, 那么, 完全可以设计出一个 100 比特密钥的密码, 其安全强度和 128 比特密钥密码相当, 也就是说, 原算法从资源上造成了一定的浪费. 由于序列密码、分组密码和公钥密码的加解密机制不同, 对这些不同密码进行安全性分析的手段也不全相同. 从理论上讲, 如果序列密码中密钥流序列是随机序列, 那么就认为它的安全性是最高的. 而在分组密码的设计中, 如果它使用的置换是一个随机置换, 那么该置换的安全强度是最高的. 然而, 在实际设计序列密码或者分组密码时, 很难构造出一条随机序列或者一个随机置换, 现实生活中的序列密码算法都不是随机序列生成器, 分组密码算法也不是随机置换, 所以分析这些私钥密码算法的实质可以归纳为以下两个问题：

(1) 如何寻找密码算法有效的区分器, 使得该区分器可以将密钥流序列与随机序列区分开, 或者将分组密码所用的置换与随机置换区分开.

(2) 密码算法的有效区分器通常是一些与密钥有关的方程, 如何列出这些方程并求解.

一般而言, 对于给定的私钥密码算法, 如果能够找到该密码算法的一个有效的区分器, 则通过代数或者统计的方法求解这些区分器, 可以得到该密码的部分或全部密钥. 因此, 寻找有效的区分器是私钥密码算法分析中的一个主要问题. 如差分密码分析和线性密码分析等都是寻找某些密码算法好的区分器, 可以实现对该密码算法的破译. 公钥密码算法的安全性主要依赖于某个困难的数学问题, 对公钥密码算法的分析相对来说比较单一, 通常就是寻找求解困难问题的有效算法.

当比较不同攻击算法的优劣时, 最主要的指标是数据复杂度、时间复杂度和空间复杂度. 数据复杂度是指为了实现一个特定的攻击所需要的数据总和. 时间复杂度是指密码分析者为了恢复密钥, 对采集到的数据进行分析和处理所消耗的时间. 在具体计算时, 通常将时间复杂度折合成加密的次数. 空间复杂度是指为得到密钥所需要的存储空间. 攻击一个密码算法, 有两个最基本的方法：一是穷尽搜索, 二是字典攻击. 所谓穷尽搜索, 就是在已知一个明密文对 (m, c) 的情况下, 用所有可能的密钥对该明文加密, 若所得密文与 c 相同, 则认为该密钥是正确密钥, 否则淘汰;

所谓字典攻击, 就是攻击者将所有可能的明密文对 (m,c) 存储下来, 当攻击者看到密文 c 时, 直接通过查表的方式得到 m. 假设一个密码算法的密钥长度为 k, 明文长度为 n, 那么穷尽搜索的数据复杂度、时间复杂度和空间复杂度分别为 1, 2^k 和 1; 对字典攻击而言, 数据复杂度、时间复杂度和空间复杂度分别为 2^n, 1 和 2^n.

从密码学的应用来看, 在公钥密码体制出现以前, 密码学的主要应用就是进行数据的加解密处理, 以达到保护秘密信息的目的. 密码技术作为保护秘密信息的一种主要手段, 最早运用于军事和政治斗争, 但是随着计算机出现, 特别是计算机和通信设备的广泛应用以及计算机网络的飞速发展, 工业、商业、金融、私营部门, 乃至个人都有很多机密或敏感数据需要存储和交换, 而在存储和传输的过程中, 这些机密或敏感数据却正在受到日益严重的安全威胁, 采用密码技术对这些数据进行加密保护和实现保密通信已经被公认为是迄今唯一最有效的手段. 因此, 密码技术和保密通信迅速从官方走向民间, 并得到前所未有的发展和应用. 随着公钥密码体制的提出, 密码学的应用范围正在不断扩大. 公钥密码技术除了提供信息的加解密外, 还具有对信息来源进行鉴别、保证信息的完整性和不可否认性等安全功能, 这些安全功能都是通过数字签名实现. 数字签名的作用有两点: 一是因为自己的签名难以否认, 从而确认了文件已签署这一事实; 二是因为签名不易仿冒, 从而确认了文件的完整性. 数字签名的原理是将要传送的明文通过一种函数运算 (通常指 Hash 函数) 转换成报文摘要, 报文摘要经签名后与明文一起传送给接收方, 接收方将接收的明文产生新的报文摘要与发送方发来的报文摘要比较, 比较结果一致表示明文未被改动, 如果不一致表示明文已被篡改. 公钥密码技术另一个重要应用就是数字证书, 数字证书是指网络通信中标志通信各方身份信息的一系列数据, 是网络正常运行所必须的“电子身份证”. 数字证书一般采用交互式询问回答, 在询问和回答过程中采用密码算法进行加解密处理. 目前采用密码技术的带 CPU 的智能卡在网络安全通信中应用非常广泛. 除此之外, 私钥密码算法和公钥密码算法相结合的混合密码体制在 SSL(安全套接字层)、SET(安全电子交易)、S/MIME(安全电子邮件)等安全通信标准和 IPsec 安全协议等诸多网络安全协议中得到重要应用. 所谓混合密码体制, 就是指利用公钥密码算法来传递私钥密码算法的会话密钥, 而采用私钥密码算法来加解密大数据量的明文信息. 采用混合密码体制进行保密通信, 是因为公钥密码算法加解密速度慢, 不适合大数据量的信息加解密, 而私钥密码体制虽然加解密速度快, 但需要事先通过安全信道传送会话密钥.

1.2 序列密码概述

序列密码是一类重要的私钥密码体制, 它的产生源于 1917 年 Gilbert Vernam 提出的 “一次一密” 密码体制. 该密码体制是指满足如下条件的密码模型: 设 $n \geqslant 1$

是正整数, $P = C = K = F_2^n$, 对于任意给定的明文 $m \in P$, 随机选取密钥 $k \in K$, 定义加密变换为: $c = E_k(m) = m \oplus k$, 相应的解密变换为: $m = D_k(c) = c \oplus k$. "一次一密" 的意思就是要求密钥空间 K 中的密钥都是按随机方式使用的, 每一个密钥被使用的概率为 $\frac{1}{2^n}$, 并且对任意 $m \in P$ 和 $c \in C$, 一定存在密钥 $k \in K$, 使得 $E_k(m) = c$, 也就是说, 一次一密密码体制具有完善保密性, 或者无条件安全性. 序列密码的基本思想是: 利用种子密钥 k 通过密钥流产生器 f 生成一个密钥流序列 $z = z_0 z_1 z_2 \cdots$, 并使用如下的规则对明文序列 $x = x_0 x_1 x_2 \cdots$ 加密:

$$y = y_0 y_1 y_2 \cdots = E_{z_0}(x_0) E_{z_1}(x_1) E_{z_2}(x_2) \cdots .$$

密钥流由密钥流产生器 f 按如下方式产生: $z_i = f(k, \sigma_i)$, 这里 σ_i 是密钥流生成器中的记忆元件在时刻 i 的状态, 即 f 是依赖于密钥 k 和状态 σ_i 的函数. 根据状态 σ_i 是否依赖于明文序列, 将序列密码分为同步和自同步两种形式. 自同步序列密码要求密钥流产生器依赖于明文序列, 因而较难从理论性分析其安全性. 目前实际使用的序列密码主要是同步序列密码, 即密钥流产生器与明文序列无关. 由于理论上已经证明了一次一密的密码体制是无条件安全的, 我们当然希望按照一定规则生成的密钥流序列能够具有随机序列类似的特性. 于是, 周期、线性复杂度、相关函数、元素分布和统计特性就成为衡量序列密码安全性的重要参数. 从 20 世纪 60 年代开始, 人们设计了多种形式的序列密码, 但主流依然是线性反馈移位寄存器序列 (Linear Feedback Shift Register, LFSR) 及其扩展序列, 包括前馈序列、非线性组合序列和钟控序列等. Galois 理论与方法在序列密码的研究中起着十分重要的作用. 比如, 利用 Galois 域上的母函数理论和方法可以完整地刻画线性移位寄存器序列的周期分布和线性复杂度分布; 利用 Galois 域上序列的根表示理论和方法可以有效地控制前馈序列的周期和线性复杂度; 利用 Galois 域上序列的迹表示理论和方法可以确定某些非线性组合序列的周期和线性复杂度. 20 世纪 80 年代后, 一方面, 人们利用 Galois 域上的迹函数理论和方法构造了多种序列密码: Bent 序列、GMW 序列、No 序列、Kasami 序列和 Kumar-Moreno 序列等, 这些序列的共同特点是: 周期大, 线性复杂度高, 相关性能好. 相对于移位寄存器序列及其变种序列而言, 这些迹函数生成序列的相关函数是可计算的. 相关函数的计算和元素统计特性的刻画一直是序列密码研究中的难点问题, 即使是在保密通信、扩频通信和 CDMA 通信系统中广泛使用的 m 序列, 它们的互相关特性到目前为止也没有得到完整的描述. 另一方面, 人们对于 Galois 环上的本原序列及其权位序列的密码特性研究给予了极大的关注, 在本原序列及其权位序列的周期特性、线性复杂度特征、元素分布和综合算法等方面得到了很多深刻的结果, 这些结果的取得离不开 Galois 环的理论与方法.

eSTREAM 计划是 2004 年欧洲启动的一项密码计划, 目的是在世界范围内征集欧洲新世纪序列密码标准算法. 在提交的候选序列密码算法中, 产生了许多序列密码新的设计思想. 比如, 在 eSTREAM 计划进入第 3 轮竞赛的候选算法当中, 面向硬件的算法有 DECIM, Edon80, F-FCSR, Grain, Mickey, Moustique, Pomaranch 和 Trivium; 面向软件的算法有 CryptMT, Dragon, HC, LEX, NLS, Rabbit, Salsa20 和 SOSEMANUK. 通过对这些序列密码算法的分析, 我们注意到这样一个事实: 面向硬件的序列密码算法 DECIM, F-FCSR, Grain, Mickey 和 Trivium 主要采用线性反馈移位寄存器 (LFSR)、非线性反馈移位寄存器 (NFSR) 和带进位的反馈移位寄存器 (FCSR) 为主要密码组件. 比如, 瑞典 Hell 等提交的 Grain 算法采用一个 80 级 LFSR、一个 80 级 NFSR 和一个 5 个变量的非线性滤波函数构成; 英国 Babbage 等提交的 Mickey-128 算法利用一个 80 级的 LFSR 来控制另一个 80 级的 NFSR 的状态更新; 法国 Arnault 等提交的 F-FCSR 算法则采用一个简单的线性滤波函数对 FCSR 生成的序列进行过滤来生成密钥流; 比利时 Canniere 等提交的 Trivium 算法采用三个相互关联相互影响的 NFSR. 而面向软件的序列密码算法 Dragon, HC, LEX, Salsa20 和 SOSEMANUK 主要采用新的设计思想和新的设计理念. 有的序列密码算法融入了分组密码的设计思想, 比如, 比利时 Biryukov 提交的 LEX 算法的基本思想就是从分组密码中提出一部分内部状态作为流密码的密钥流, 该方案对每个分组密码都通用, 但需要具体分析哪些内部状态可以提取; 澳大利亚 Dawson 等提交的 Dragon 算法可以看作是分组密码的输出反馈模式 (OFB) 构成的流密码, 该方案包含了一个 1024 比特的 NFSR 和一个状态更新函数; 美国 Bernsterin 提交的 Salsa20 算法在总体结构上借鉴了 AES 的设计风格, 采用了行变换和列变换的方式来提高系统的安全性; 法国 Berbain 等提出的 SOSEMANUK 则巧妙地将流密码 SNOW 2.0 和分组密码 SERPENT 结合起来. 有的序列密码算法融入了著名序列密码算法 RC4 的设计理念, 比如, 新加坡 Hongjun Wu 提交的 HC-256 算法借鉴 RC4 的设计理念, 采用两个 1024 比特的内部状态, 类似于 RC4 中的 SWAP 函数. 有的序列密码算法与 Hash 函数有关, 比如, Salsa20 算法的主要组件就是基于 Hash 函数设计的, 尽管后来密码分析学者指出 Salsa20 采用的主要组件作为 Hash 函数来说是不安全的, 但 Salsa20 算法作为序列密码算法来说是最受大家欢迎的. 目前对这些序列密码算法的实现 (软件实现和硬件实现) 性能分析较为详尽, 但关于这些算法的安全性分析, 很少有较为详尽的令人信服的密码分析. 主要原因是对于那些面向硬件的序列密码算法, 大都采用非线性反馈移位寄存器 (NFSR) 作为主要的密码组件, 同时为了提高系统的实现速度, 面向字的非线性逻辑往往采用字加法、异或和循环移位等三种基本运算, 并且通常含有分组密码的某些基本组件 (S 盒、T 函数和 P 置换), 使得理论分析难度加大. 对于那些面向软件的序列密码算法, 由于采用了一些新的设计思想和理念, 使得那些分析经典序列密码的工具难以

分析这些新型的序列密码.

本书第一部分内容就是利用 Galois 域的基本理论与方法, 较为系统地研究 Galois 域上线性移位寄存器序列及其扩展序列的密码学性质, 即本书的第 2 章和第 3 章. 然后利用图论和组合数学等代数工具研究 Galois 域上非线性移位寄存器序列的一些密码学性质, 即本书的第 4 章. 系统地掌握这些经典序列密码设计和分析的理论与方法, 对于进一步研究 eSTREAM 计划中的现代序列密码的特性具有十分重要的意义.

1.3　分组密码概述

分组密码是指将明文消息经过编码表示后的二进制序列 $x_1, x_2, \cdots, x_i, \cdots$ 划分成固定长度为 n 比特的组:

$$X_i = (x_{(i-1)n+1}, x_{(i-1)n+2}, \cdots, x_{(i-1)n+n}), \quad i = 1, 2, 3, \cdots,$$

各组在密钥 $k = (k_1, k_2, \cdots, k_r)$ 的控制下变换成长度为 n 的二进制序列:

$$Y_i = (y_{(i-1)n+1}, y_{(i-1)n+2}, \cdots, y_{(i-1)n+n}), \quad i = 1, 2, 3, \cdots,$$

所以分组密码具有以下严格的数学定义:

定义 1.1　记 F_2^n 和 F_2^r 分别为 F_2 上 n 维和 r 维向量空间, $S_K \subseteq F_2^r$, 则数据长度为 n 的分组密码定义为满足下列条件的映射 $E : F_2^n \times S_K \to F_2^n$: 对每个 $k \in S_K$, $E(\cdot, k)$ 是一个从 F_2^n 到 F_2^n 的置换.

通常情况下, 称 $E(\cdot, k)$ 为密钥 k 时的加密函数, $E(\cdot, k)$ 的逆置换 $D(\cdot, k)$ 为密钥 k 时的解密函数. $t = \log_2 |S_K|$ 称为分组密码的密钥长度或者密钥规模, 它是度量分组密码安全性的重要指标. 一般情况下, 通常取 $S_K = F_2^r$, 这时 r 就是分组密码的密钥长度. 从上述定义可以看出, 分组密码本质上是由密钥所控制的从明文空间到密文空间的置换, 它与序列密码的区别在于:

(1) 序列密码具有记忆功能, 分组密码却没有. 序列密码的密钥流 $z = z_0 z_1 z_2 \cdots$ 在 i 时刻的密钥 z_i 由密钥 k 和当前状态所确定, 而输入的明文可能影响寄存器中内部记忆元件的存储状态.

(2) 序列密码将明文序列按一定的长度分组 (通常是一个字符), 然后各组用相关但不同的密钥进行加密, 产生相应的密文, 相同的明文分组会因在明文序列中的位置不同对应不同的密文分组. 而分组密码将明文序列分组后, 每一组用相同的密钥进行加密, 产生相应的密文, 相同的明文分组不管处在明文序列的什么位置, 总是对应相同的密文分组.

分组密码是对称密码学中一个重要分支, 分组密码的研究始于 20 世纪 70 年代, 研究的主要内容包括两个方面: 分组密码的设计和分组密码的分析. 分组密码的设计与分析是既相互对立又相互依存的两个方面. 从分组密码的发展来看, 正是由于这种相互对立, 才促进了分组密码的飞速发展. 分组密码早期研究主要围绕美国数据加密标准 DES 进行, 推出了许多类似于 DES 的分组密码, 比如, LOKI 密码、FEAL 密码和 GOST 密码等. 1977~1990 年间, 由于对 DES 密码的攻击没有取得突破性进展, 分组密码的设计与分析发展较为缓慢. 20 世纪 90 年代以后, 由于针对 DES 算法的差分密码攻击和线性密码攻击的提出, 人们不得不研究新的密码结构. 由 Xuejia Lai 和 Massey 设计的国际数据加密算法 IDEA[10] (International Date Encryption Algorithm) 打破了 DES 类密码的垄断局面, 其设计思想主要是混合使用来自不同群中运算. 随后出现的 Square 密码、Shark 密码和 Safer-64 密码都采用了结构非常清晰的代替–置换网络, 这类密码的特点是算法每一轮由较为明确的混淆层 (S 盒) 和扩散层 (P 置换) 组成, SPN 结构的最大优点是能够从理论上证明一个密码算法能否抵抗差分密码攻击和线性密码攻击. 1997 年, NIST 为了履行其法定职责, 发起了一场推选用于保护敏感的联邦信息的对称密码算法的活动, 在世界范围内征集美国的高级加密标准 AES, 以取代 DES 算法. 1998 年, NIST 宣布接收 15 个候选分组密码算法并提请全世界密码学者协助分析这些候选算法, 包括对每个算法的安全性和效率特性进行初步检验. NIST 考察了这些初步的研究结果, 选定 MARS, RC6, Rijndael, Serpent 和 Twofish 等五个分组密码算法作为参加决赛的算法, 经公众对决赛算法进行进一步的分析和评论, 2000 年, NIST 决定推荐比利时人 Daemen 和 Rijmen 设计的 Rijndael 密码作为 AES 算法. 继美国推出 AES 计划以后, 欧洲于 2000 年启动了新欧洲签名、完整性和加密计划 (New European Schemes for Signatures, Integrity and Encryption, NESSIE), 以适应 21 世纪信息安全发展的全面需求. 该计划为期三年, 主要目的就是通过公开征集和进行公开透明的测试、评估, 提出一套高效的密码标准, 以保持欧洲工业界在密码学研究领域的领先地位. 2003 年, NESSIE 工作组公布了包括分组密码、公钥密码、认证码、杂凑函数和数字签名等在内的 17 个标准算法, 其中 MISTY1, Camellia, SHACAL 三个分组密码算法连同 AES 算法一起作为欧洲新世纪的分组密码标准算法. 这些计划的兴起, 使得分组密码研究从经验设计走向理论设计的道路, 分组密码理论得到飞速发展, 同时, 分组密码理论的发展也带动了密码学其他分支的发展. 2004 年, 在欧洲 eSTREAM 计划的序列密码标准算法的征集活动中, 涌现了一大批基于分组密码的工作模式构造的新的序列密码算法, 如 Salsa20, LEX 等都是基于这种模式设计的. 随着 MD 系列 Hash 函数碰撞攻击的成功, 人们越来越关注基于分组密码所构造的 Hash 函数类, 比如 Tiger 算法、Whirlpool 算法和 Grindahl 算法等. 尽管在当前密码学领域已有很多公认的安全强度高的分组密码算法, 人们仍力求设计新

的安全高效的算法, 如最近设计的 FOX 算法和 CLEFIA 算法等, 这些算法继承了 AES 算法的优点, 同时也有自己独特的设计技巧, 对这些新算法的安全性进行分析大大充实了现有的分组密码的设计与分析理论.

本书的第 5~7 章主要介绍分组密码的设计原理和分析方法, 以及基于这些设计原理所构造的一些典型分组密码算法, 其目的是使读者掌握分组密码的基本理论和方法, 以便进一步对分组密码的设计与分析方法展开研究.

1.4　公钥密码概述

序列密码和分组密码由于具有加解密速度快、容易实现等诸多优点, 被广泛应用于现代通信系统中的数据加密. 由于序列密码和分组密码均为私钥密码体制, 其加密密钥与解密密钥是一致的. 在应用这些私钥密码体制之前, 通信双方必须通过安全信道协商密钥, 这一点在实际通信时是很难做到的. 比如假定通信双方相距很远, 他们希望通过电子邮件来进行保密通信, 在这种情况下, 通信双方可能没有一个合理的安全信道. 另外, 假设一个通信系统中有 n 个用户采用私钥密码体制进行保密通信, 他们中任意两人之间都必须拥有一个保密的密钥, 这样就需要 C_n^2 个密钥, 当 n 比较大时, 如何安全管理这 C_n^2 个密钥就是一个非常棘手的问题.

1976 年, Diffie 和 Hellman 为解决私钥密码体制中的密钥管理问题, 创造性地提出了一种密钥交换协议, 允许在不安全的信道上通过通信双方交换信息, 安全地达成一致的密钥, 这就是著名的 Diffie-Hellman 交换协议. Diffie-Hellman 交换协议作为公钥密码的起源利用了如下事实: 在一个高阶有限群中, 已知元素 a 和正整数 n, 计算 $a^n = b$ 是一件容易的事, 但反过来, 已知元素 a 和 $b = a^n$, 计算 n 是一件十分困难的事. 所谓一个问题是困难的或者说难解的, 直观上讲, 就是不存在一个计算该问题的有效算法. 换句话说, 按照目前的计算能力, 计算一个难解的问题所需要的时间是非常长的, 譬如 100 年甚至更长. 理论上来说, 对于某个问题而言, 如果存在一个算法, 使得求解该问题所需要的时间是输入数据长度的多项式函数, 就称该算法是一个有效算法, 相应的问题就称为易解问题, 否则就称该问题是难解的. 一般而言, 计算一个难解的问题所需要的时间通常是输入数据长度的一个指数函数. 因此, 对于一个难解的问题, 随着输入数据长度的增大, 进行计算所需要的时间将会急剧增加. 到目前为止, 还没有人能够严格证明哪一个问题是难解问题. 对于某些问题, 比如大整数分解问题、离散对数的求解问题等, 经过多年的努力仍然没有找到进行计算的有效算法, 使得人们认为这些问题是难解的.

公钥密码体制的提出与难解问题有关, 其理论基础就是所谓的陷门单向函数: 设 $y = f(x)$ 是一个函数, t 是一个与 f 有关的参数. 如果对于任意给定的 x, 计算 $y = f(x)$ 是容易的; 并且对任意给定的 y, 当不知道参数 t 时, 计算 x 使得 $f(x) = y$