分组密码的攻击方法与实例分析

李 超 孙 兵 李瑞林 著

科学出版社
www.sciencep.com

国防科技大学学术著作出版基金资助

分组密码的攻击方法与实例分析

李 超 孙 兵 李瑞林 著

科学出版社

北京

内 容 简 介

本书以美国AES计划和欧洲NESSIE计划等推出的著名分组密码算法为背景，系统地介绍分组密码的攻击方法和实例分析，包括差分密码攻击、线性密码攻击、高阶差分密码攻击、截断差分密码攻击、不可能差分密码攻击、积分攻击、插值攻击和相关密钥攻击等主要攻击方法的基本原理及其应用实例.

本书可以作为密码学专业和信息安全专业高年级本科生和研究生的选修课教材，也可以作为从事密码理论和方法研究的科技人员的参考书.

图书在版编目(CIP)数据

分组密码的攻击方法与实例分析/李超，孙兵，李瑞林著. —北京：科学出版社，2010

ISBN 978-7-03-026609-5

Ⅰ. ①分… Ⅱ. ①李… ②孙… ③李… Ⅲ. ①密码学 Ⅳ. ①TN918.1

中国版本图书馆 CIP 数据核字 (2010)第 017449 号

责任编辑：赵彦超 / 责任校对：郑金红

责任印制：钱玉芬 / 封面设计：王 浩

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

骏走印刷厂印刷

科学出版社发行 各地新华书店经销

*

2010 年 5 月 第 一 版 开本：B5(720×1000)

2010 年 5 月 第一次印刷 印张：15 1/2

印数：1—2 500 字数：295 000

定价：48.00 元

(如有印装质量问题，我社负责调换)

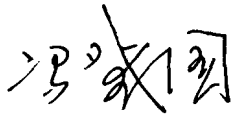
序

随着计算机网络和通信技术的飞速发展, 人们对信息的安全存储、处理和传输的需要越来越迫切, 信息的安全保护问题已经显得十分突出, 人们正面临着信息安全的巨大挑战. 作为信息安全理论和技术的基础, 密码学扮演着十分重要的角色. 分组密码作为对称密码学的重要分支, 已经在信息安全领域中得到了广泛应用.

分组密码的研究内容主要包括分组密码的设计与分析, 两者相互作用, 共同推动着分组密码理论的发展. 对于从事密码学相关领域的研究人员来说, 深刻理解分组密码的分析方法是从事分组密码理论与应用研究的前提. 自 20 世纪 90 年代差分密码分析和线性密码分析提出以来, 分组密码的分析理论有了长足的发展, 但有关分组密码分析的理论成果大多散落在国内外与密码学相关的学术会议论文集上, 专门讲述分组密码攻击方法的著作并不多见.

国防科技大学李超教授及其课题组多年来从事分组密码相关理论的研究, 取得了一系列学术成果, 这些成果相继发表在国内外重要期刊和学术会议论文集上, 引起了国内外密码学者的高度关注. 最近, 他们结合自身在密码分析方面的工作体会, 编写了《分组密码的攻击方法与实例分析》, 通过对一些具体密码算法的实例分析, 系统论述了分组密码攻击方法的基本原理与应用.

我相信, 本书的出版对于从事分组密码的理论与应用研究将具有十分重要的参考价值.



2009 年 12 月于北京

前　言

分组密码算法是许多密码系统的核心要素, 是保障信息机密性和完整性的重要技术. 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洞察, 才有可能设计出安全高效的分组密码.

近年来, 随着美国 AES 计划和欧洲 NESSIE 计划的实施, 针对美国高级加密标准 AES 和欧洲分组密码标准 Camellia, MSITY1 以及 SHACAL-2 等算法的安全性分析已经成为分组密码研究中的热点问题. 在 SHA-3 计划中, 可以发现很多基于分组密码组件设计的 Hash 函数, 有些算法甚至直接使用 AES 算法的组件. 之所以基于分组密码设计 Hash 函数, 是因为密码设计者对现有分组密码算法特别是 AES 算法的安全性有足够的信心. 对这些基于分组密码的 Hash 函数的安全性分析必将促进分组密码的设计与分析理论的发展. 因此, 掌握分组密码的分析技术对研究分组密码的安全性和 Hash 函数的安全性至关重要.

考虑到分组密码的攻击方法是密码学领域的难点问题, 且大量关于分组密码设计与分析的学术论文散见于密码学与信息安全的国际会议论文集, 特别是与密码学密切相关的 EUROCRYPT, CRYPTO, ASIACRYPT, FSE, CHES 和 SAC 等国际会议. 为便于国内从事分组密码理论与方法研究的研究生和年轻学者对分组密码的攻击方法有一个比较系统和深入的了解, 我们试图对国际上最近 20 年已有的分组密码攻击方法进行梳理, 通过对一些具体的分组密码算法的实例分析, 介绍分组密码攻击方法的原理与实践. 在写作过程中, 我们特别注重自身对分组密码攻击方法的理解, 书中部分内容包含了作者及其课题组成员近年来在分组密码攻击方面所取得的研究成果.

全书共分 10 章. 第 1 章给出分组密码的基本概念; 第 2 章介绍 8 个典型的分组密码算法; 第 3 章介绍差分密码分析的原理与实例分析; 第 4 章介绍线性密码分析的原理与实例分析; 第 5 章介绍高阶差分密码分析的原理与实例分析; 第 6 章介绍截断差分密码分析的原理与实例分析; 第 7 章介绍不可能差分密码分析的原理与实例分析; 第 8 章介绍积分攻击的原理与实例分析; 第 9 章介绍插值攻击的原理与实例分析; 第 10 章介绍相关密钥攻击的原理与实例分析.

冯克勤教授、裴定一教授和冯登国研究员等对本书的写作给予了极大的鼓励和支持, 科学出版社的责任编辑为本书的出版付出了辛勤的劳动, 在此表示深深的感谢! 全书的编写工作得到了国防科技大学理学院密码与信息安全实验室全体师生的积极配合, 特别是密码算法分析小组的魏悦川博士、张鹏博士、唐学海博士、王美一硕士等给予了全力协作和密切配合, 在此一并对他们表示衷心的感谢!

本书的出版得到国防科技大学学术著作出版基金和数学学科建设基金的资助. 此外, 本书部分成果来自课题组受资助的项目: 国家自然科学基金 (No: 60573028; 60803156)、信息安全国家重点实验室开放基金 (No: 01-07)、东南大学移动通信国家重点实验室开放基金 (No: w200805; w200807) 以及国防科技大学基础研究基金 (No: JC090201) 的资助, 在此一并表示感谢!

限于作者的水平, 书中难免存在不妥之处, 恳请读者批评指正.

作　者

2009 年 9 月 16 日

目　录

第 1 章　分组密码的基本概念

1.1　分组密码概述

分组密码是对称密码学的一个重要分支, 在信息安全领域发挥着极其重要的作用, 其研究的主要内容包括分组密码的设计和分析这两个既相互对立又相互统一的方面. 一方面, 针对已有的密码分析手段, 密码设计者总希望设计出可以抵抗所有已知攻击的密码算法; 另一方面, 对已有的密码算法, 密码分析者总希望可以找到算法的某些安全缺陷. 这两方面的研究共同推动了分组密码理论的发展.

分组密码的设计理念源于 Shannon1949 年发表的经典论文*Communication Theory of Secret System*[37], 其公开研究始于 20 世纪 70 年代末 DES 算法[14] 的公布, 分组密码理论及应用的飞速发展则得益于 20 世纪 90 年代末美国的 AES 计划[43] 和本世纪初欧洲的 NESSIE 计划[44].

Shannon 在文献 [37] 中从抵抗统计攻击的角度出发, 提出了设计加密算法的"混淆" 与 "扩散" 准则, 这一准则至今仍是设计分组密码所要遵循的重要原则之一. 此外, 他还创造性地从信息论的角度特别是信息熵出发构建数学模型以研究密码的安全性, 提出了 "完善保密性"、"唯一解距离" 和 "随机密码" 等诸多概念, 从而将密码学提升到了科学的范畴. 尽管如此, 在 20 世纪 70 年代以前, 对分组密码研究的公开文献微乎其微, 其理论研究相对滞后.

1977 年, 美国国家标准局 (National Bureau of Standards, NBS) 公布了著名的数据加密标准 DES(Data Encryption Standard) 算法. 尽管 DES 算法正逐步退出历史舞台, 但它对分组密码理论的发展起到了举足轻重的作用. 首先, 算法的公布促使民间开展了对分组密码的研究, 使得分组密码的设计与分析逐渐褪去神秘的面纱; 其次, 通过对 DES 算法安全性的研究, 分组密码的分析理论日渐成熟, 主要结果包括差分密码分析和线性密码分析两个方面.

在 CRYPTO 1990 上, Biham 等发表了对 DES 算法差分分析的论文[8]. 这篇文章发表后, 密码学界用差分密码分析的方法对几乎所有已知的密码算法进行了安全性分析. 1993 年, Matsui 在 EUROCRYPTO 上公布了对 DES 算法线性密码分析的结果[33]. 随后, 人们利用各种技巧改进了对 DES 算法的差分和线性密码分析, 结果表明, 完整 16 轮 DES 算法对差分和线性密码分析都是不免疫的.

计算机技术的发展是促使密码理论不断进步的又一重要因素. 计算机技术, 特别是并行计算和分布式计算的发展使得穷尽搜索 DES 算法的 56 比特密钥成为可

能, 加上差分密码分析和线性密码分析技术的出现, 56 比特密钥的 DES 算法逐渐不能满足人们的安全需求. 1997 年, 美国国家标准技术研究所 (National Institute of Standard Technology, NIST) 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, 即 AES(Advanced Encryption Standard) 计划. 1998 年, NIST 宣布接受 15 个候选分组密码算法并提请全世界密码研究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 5 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 Rijndael 作为高级加密标准 (AES).

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 ——NESSIE(New European Schemes for Signatures, Integrity, and Encryption) 计划, 以适应 21 世纪信息安全发展的全面需求. 该计划为期 3 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 Camellia, MISTY1, SHACAL-2 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.

在 AES 计划和 NESSIE 计划中, 密码学界对分组密码的设计与分析理论都进行了广泛而深入的研究, 分组密码理论日趋完善, 人们对设计出安全高效的分组密码算法较有信心. 也正是因为人们对分组密码算法安全性具有足够的信心, 在 SHA-3 计划中[45], 超过半数的 Hash 函数都采用了分组密码的设计理念, 甚至直接采用分组密码的组件. 随着 SHA-3 计划的实施, 分组密码的设计与分析理论必将得到更进一步的发展.

1.2　分组密码的设计原理

分组密码的数学模型如下:

记 $\mathbb{F}_2$ 为二元域, $\mathbb{F}_2^n$ 和 $\mathbb{F}_2^m$ 分别为 $\mathbb{F}_2$ 上的 n 和 m 维向量空间, $S_K \subseteq \mathbb{F}_2^m$, 那么一个以 $\mathbb{F}_2^n$ 为明文和密文空间、S_K 为密钥空间的分组密码就可以表示为如下两个映射:

$$E : \mathbb{F}_2^n \times S_K \to \mathbb{F}_2^n, \quad D : \mathbb{F}_2^n \times S_K \to \mathbb{F}_2^n.$$

上述两个映射满足对任意 $k \in S_K$, $E(\cdot, k)$ 和 $D(\cdot, k)$ 都是 $\mathbb{F}_2^n$ 上的置换, 并且互为逆置换. 通常称 $E(\cdot, k)$ 为固定密钥 k 时的加密函数, $D(\cdot, k)$ 为固定密钥 k 时的解密函数. 上述模型中明文和密文的长度均为 n, 而密钥的长度为 $l = \log_2 |S_K|$.

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密变换. 一个好的分组密码应该是既难破译又容易实现, 也就是说, 加密函数 $E(\cdot,k)$ 和解密函数 $D(\cdot,k)$ 是很容易计算的, 但要从方程 $y=E(x,k)$ 或 $x=D(y,k)$ 中解出 k 应该是一个困难问题.

1.2.1 分组密码的设计原则

分组密码的设计通常遵循如下两个原则: 安全性原则和实现原则.

安全性原则包含混淆原则、扩散原则和抗现有攻击原则. 混淆原则是指所设计的密码应该是明文、密文和密钥三者之间的依赖关系非常复杂以至于攻击者无法理出相互之间的关系, 从而这种依赖性对密码分析者来说是无法利用的; 扩散原则是指所设计的密码应该使得明文和密钥的每一比特影响密文的许多比特, 从而便于隐蔽明文的统计特性, 该准则强调输入的微小改变将导致输出的多位变化; 抗现有攻击的原则是指所设计的密码应该抵抗已有的各种攻击方法.

实现原则包含软件实现原则和硬件实现原则. 软件实现原则是指密码算法应该尽可能使用子块和简单的运算, 比如采用 8, 16, 32 位的字进行模加运算、移位运算或者异或运算等; 硬件实现原则是指密码算法应该保证加密和解密的相似性, 即加密和解密过程应该仅仅是密钥的使用方式不同, 以便同样的器件既可以用来加密也可以用来解密.

通常采用迭代手段使得设计出的算法符合上述原则: 一种方法是构造密码学性质强的迭代函数, 从而可以减少迭代次数; 另一种方法是构造密码学性质相对弱的迭代函数, 但迭代次数相对较多. 在实际构造中通常采用后者, 即把密码学性质较弱的函数迭代多次以满足安全性原则和实现原则.

1.2.2 分组密码的结构

目前通用的密码算法都采用了迭代结构, 根据算法采用结构的不同, 现行主要结构可分为 Feistel 结构、SPN 结构和 Lai-Massey 结构等.

(1) Feistel 结构

Feistel 结构是 20 世纪 60 年代末 IBM 公司的 Feistel 和 Tuchman 在设计 Lucifer 分组密码时提出的, 后因 DES 算法的广泛使用而流行.

对于分组长度为 $2n$ 的 r 轮 Feistel 结构的密码, 参考图 1.1, 加密流程如下:

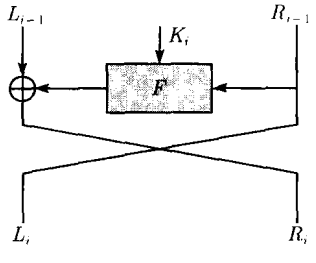


图 1.1 Feistel 结构示意图

给定 $2n$ 比特的明文 P, 首先将其分为左右两个 n 比特部分, 不妨记 L_0 是 P 的左边 n 比特, R_0 是 P 的右边 n 比特, 则 $P=L_0R_0$. 然后根据如下规则, 进行 r

轮完全相同的运算. 对于 $i = 1, 2, \cdots, r$, 令

$$\begin{cases} L_i = R_{i-1}, \\ R_i = L_{i-1} \oplus F(R_{i-1}, K_i), \end{cases}$$

这里 "$\oplus$" 表示异或运算, $F : \mathbb{F}_2^n \times \mathbb{F}_2^m \to \mathbb{F}_2^n$ 是轮函数, $K_1, K_2, \cdots, K_r$ 是由种子密钥 K 根据密钥扩展方案得到的轮密钥, m 为轮密钥的长度. 在加密的最后一轮, 不需要做 "左右交换", 即密文为 $C = R_r L_r$, 这主要是为了使算法加密和解密流程一致.

在密码设计中, 加解密一致的算法在实现时往往可以节省资源. 但注意到 Feistel 结构的密码扩散较慢, 因为算法至少需要两轮才有可能改变输入的每一比特.

(2) SPN 结构

SPN 结构每轮一般由一个轮密钥控制的可逆非线性函数 S 和一个可逆线性变换 P 组成. SPN 密码的结构非常清晰, S 变换层起混淆作用, P 变换层起扩散作用. 与 Feistel 结构相比, SPN 结构数据扩散更快, 而且, 当给出 S 变换层和 P 变换层的某些安全性指标后, 设计者可以给出算法抗差分密码分析和线性密码分析的可证明安全, 但 SPN 结构密码的加解密通常不具有一致性, 从而在实现时需要更多的资源.

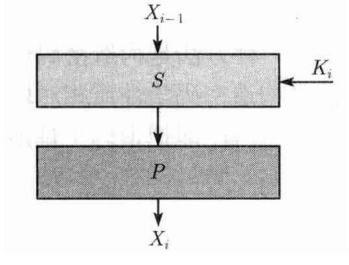


图 1.2　SPN 结构示意图

对于分组长度为 n 的 r 轮 SPN 结构密码, 参考图 1.2, 加密流程如下:

给定 n 比特的明文 P, 令 $P = X_0$; 然后根据如下规则, 进行 r 轮完全相同的运算. 对于 $i = 1, 2, \cdots, r$, 令

$$\begin{cases} Y = S(X_{i-1}, K_i), \\ X_i = P(Y). \end{cases}$$

在 SPN 结构密码中, 最后一轮中的 P 变换通常由密钥加代替.

(3) Lai-Massey 结构

Lai-Massey 结构是 Lai 和 Massey 设计 IDEA 算法时提出的一种结构, Junod 和 Vaudenay 根据 Lai-Massey 结构设计了 FOX 算法. 通常情况下, Lai-Massey 结构也具有加解密一致的优点.

对于分组长度为 $2n$ 的 r 轮 Lai-Massey 结构密码, 参考图 1.3, 加密流程如下:

给定 $2n$ 比特的明文 P, 首先将其分为左右两个 n 比特部分, 不妨记 L_0 是 P 的左边 n 比特, R_0 是 P 的右边 n 比特, 则 $P = L_0 R_0$. 然后根据如下规则, 进行 r 轮完全相同的运算. 对于 $i = 1, 2, \cdots, r$, 令

$$\begin{cases} T = F(L_{i-1} \oplus R_{i-1}, K_i), \\ L_i = L_{i-1} \oplus T, \\ R_i = R_{i-1} \oplus T. \end{cases}$$

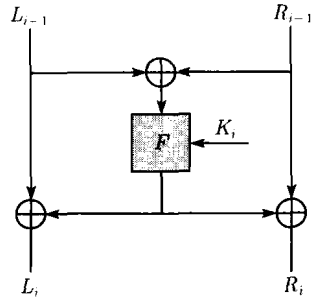


图 1.3 Lai-Massey 结构示意图

整体结构是分组密码算法的重要特征, 不同结构对轮函数的选取以及各种平台上的性能都有很大的影响. 除了上述三种主流结构外, 整体结构还包括广义 (非) 平衡 Feistel 结构、MISTY 结构以及各种结构的混合使用. 另外, 很多密码算法的轮函数采用了不同的结构. 如 Camellia 算法整体采用 Feistel 结构, 但轮函数采用了 SPN 结构; FOX 算法整体采用 Lai-Massey 结构, 轮函数采用 SPS 结构; SMS4 算法采用了广义非平衡 Feistel 结构, 轮函数采用了 SPN 结构. 设计一个算法采用何种结构主要依赖于算法的性能要求、子模块的构造以及整体结构的安全性等因素.

1.3 分组密码的分析方法

衡量一个密码算法的安全性有两种基本方法: 一是实际安全性, 二是无条件安全性, 又称理论安全性. 实际安全性是根据破译密码系统所需的计算量来评价其安全性的, 如 RSA 系统的安全性就是基于大整数分解的困难性, 但是必须注意到, 随着计算机技术发展, 现在固若金汤的 RSA 系统在不久的将来也可能不堪一击. 理论安全性则与对手的计算能力或时间无关, 破译一个密码算法所做的任何努力都不会优于随机选择即碰运气.

如果攻击者能够恢复算法的密钥, 这样的攻击称为 "密钥恢复攻击". 如果攻击者能够根据已有的信息, 在不知道密钥的情况下可以预测某些未知明文所对应的密文, 也可以称这个算法被部分攻破. 如无特殊说明, 本书所说的攻击都指 "密钥恢复攻击".

1.3.1 密码分析中常见的假设和原则

研究密码算法的理论安全性往往都基于 Kerckhoffs 假设:

Kerckhoffs 假设 在密码分析中, 除了密钥以外, 密码分析者知道密码算法的每一个设计细节.

根据 Kerckhoffs 假设可知, 密码算法的安全性应依赖密钥的保密性, 而不是算法本身的保密性.

在密码体制的基本模型中, 密码分析者的任务就是获取适量的明文及其相应的

密文, 通过分析这些明文和密文得到密钥信息. 根据攻击环境的不同可以将密码攻击分为如下四种类型:

(1) 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.

(2) 已知明文攻击. 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.

(3) 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.

(4) 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.

1.3.2　强力攻击

对任意一个分组密码, 都存在如下 4 种攻击方法:

(1) 穷尽密钥搜索. 在唯密文攻击下, 攻击者利用所有可能的密钥对一个或多个密文进行解密, 直至得到有意义的明文; 在已知明文攻击或选择明文攻击时, 攻击者利用所有可能的密钥对一个已知明文加密, 直到加密结果与正确的密文相符合. 穷尽密钥搜索理论上可以破译任何分组密码算法, 但它的效率是最低的, 在实际密码分析中, 通常将穷尽密钥搜索与其他分析方法结合使用.

(2) 字典攻击. 攻击者收集明密文对, 并将它们编排成一个 "字典", 当看到一个密文时, 攻击者检查这个密文是否在字典中, 如果在, 则攻击者获得该密文对应的明文.

(3) 查表攻击. 该方法是选择明文攻击, 攻击者利用所有可能的密钥对同一个明文加密, 将密钥和对应的密文存储起来. 当获得该明文及相应密文后, 攻击者只需从存储表中找到相对应的密钥即可.

(4) 时间–空间权衡攻击. 这是一种选择明文攻击方法, 由 Hellmma 提出[18], 通过结合使用穷尽密钥搜索攻击和查表攻击, 在选择明文攻击中用时间换取空间, 因此该方法比穷尽密钥搜索攻击的时间复杂度小, 比查表攻击的空间复杂度低.

除了上述 4 种通用的攻击方法, 在对一个具体的分组密码算法进行安全性分析时, 根据算法的特点, 往往会有其他不同的攻击方法. 而比较不同攻击算法的优劣, 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; 空间复杂度是指为了完成攻击所需要的存储空间.

一般而言, 假设分组长度为 n, 密钥长度为 k, 则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, 字典攻击的数据复杂度、时间复杂度和空