

信息与通信工程研究生规划教材



网 络信息安全理论与技术

Theory and Technology for Network Information Security

蒋 睿 胡爱群 陆哲明 等编著

华中科技大学出版社

<http://www.hustp.com>

TP393.08/241

2007

信息与通信工程研究生规划教材

网络信息安全理论与技术

Theory and Technology for
Network Information Security

蒋 睿 胡爱群 陆哲明 编著
宋宇波 秦中元 黄 杰 裴文江

华中科技大学出版社
(中国·武汉)

图书在版编目(CIP)数据

网络信息安全理论与技术/蒋 睿,胡爱群 陆哲明等编著. —武汉:华中科技大学出版社,2007年11月

ISBN 978-7-5609-4287-2

I. 网… II. ①蒋… ②胡… ③陆… III. 计算机网络-安全技术-研究生-教材
IV. TP393.08

中国版本图书馆CIP数据核字(2007)第163007号

网络信息安全理论与技术

蒋 睿 胡爱群 陆哲明等编著

责任编辑:张志华

封面设计:潘 群

责任校对:陈 骏

责任监印:周治超

出版发行:华中科技大学出版社(中国·武汉)

武昌喻家山 邮编:430074 电话:(027)87557437

录 排:华中科技大学惠友文印中心

印 刷:湖北新华印务有限公司

开本:850mm×1065mm 1/16

印张:27.75

字数:585 000

版次:2007年11月第1版

印次:2007年11月第1次印刷

定价:38.80元

ISBN 978-7-5609-4287-2/TP·639

(本书若有印装质量问题,请向出版社发行部调换)

内容提要

什么是信息安全的基础理论一直是困扰学术界的一个问题。除了密码学理论这个方向外,目前关于其他用来解决信息安全问题的技术的基础理论还没有形成共识。本书力求在当前层出不穷的信息安全技术间,阐述并归纳出这些技术赖以支撑的理论基础,并把这些理论分成密码学理论、安全协议理论、模式识别理论、信息隐藏理论以及可信计算理论。

本书分为五个部分。第一部分是关于密码学理论,主要阐述密码学数学基础和常用密码算法理论;第二部分是关于安全协议理论,主要阐述安全协议的形式化分析方法与安全协议的分析、设计理论,同时介绍相应的安全协议;第三部分是关于模式识别理论,主要论述模式分类基础理论,并在此基础上介绍常用的模式识别方法;第四部分是关于信息隐藏理论,主要论述信息的隐藏理论,同时介绍相应的技术应用;第五部分是关于可信计算导论,主要论述可信计算的概念与架构。本书安排的这五个部分具有强烈的代表性,既考虑到信息安全知识的覆盖面,又兼顾到先进性和实用性,以信息安全的各个方面最基础、最典型的理论为主要内容,便于教学。各部分按照先介绍基础理论,后讲解应用的方式安排。这是本书不同于目前已有教材的地方。

本书在每一部分后面附有参考文献和适量的有代表性的习题。参考文献方便读者查阅原始文献,习题可以让读者更好地理解书上讲解的内容。

Abstract

It is the problem puzzled the academe all the time for what is the fundamental theory of the information security. Beside the cipher theory, there are many other technologies dealing with the information security. However, the fundamental theories of these technologies are not commonly acknowledged. In this book, we will try our best to expatiate and conclude the fundamental theories, which can support the secure information technologies emerged in endlessly. We classify the theories into five parts, which include the cipher theory, the security protocol theory, the pattern recognition theory, the information hiding theory and the trusted calculation theory.

This book is divided into five parts. The first part is the cipher theory, which mainly expatiates the mathematic theory of the cipher and introduce some common cipher algorithm. The second part is the security protocol theory, which mainly expatiates the formal method for security protocol and the theory of protocol analysis and design. Also, the corresponding security protocols are introduced in this part. The third part is the pattern recognition theory, which mainly discuss the fundamental theory of pattern classifying and introduce the common method for pattern recognition. The fourth part is the information hiding theory, which mainly discuss the fundamental theory of information hiding and introduce the corresponding applications. The last part is the trusted calculation theory, which mainly introduce the concepts and framework of the trusted calculation. The five parts arranged in this book have strong characters of delegation, they consider not only the coverage of the information security knowledge, but also have the characters of advancement and practicability. It is also convenient for teaching. The difference of this book between the other teaching material is the arrangement method, for which is discussing the theory first, and introducing the application second.

In the end of five parts of the book, there are corresponding references and some typical exercises. The readers can find the references conveniently, and they can understand the contents of the book better through doing exercises.

总 序

随着信息时代的到来,人类已经生活在信息的“海洋”之中,信息和通信已渗入我们生活的各个方面。近年来,我国的电信产业以10%以上的年增长率迅猛发展,“中国制造”的通信产品广泛进入了全球市场。另一方面,信息和通信领域的理论与技术获得了迅速发展,不少技术难题已取得实质性突破,技术进步和产业发展相互推动、相互促进。

产业的发展带来了对人才,特别是高层次专业人才的巨大需求。信息与通信工程是我国工科门类中应用前景广阔、招生量比较大的学科,对我国的现代化建设起着非常重要的作用。其中的通信与信息系统更是近几年硕士研究生报考的热门专业之一。随着硕士研究生的不断扩招,研究生教育成为一个突出的问题。鉴于通信学科的迅猛发展,广大科技工作者和硕士、博士研究生迫切需要学习与掌握信息和通信的现代理论与技术。目前本专业的研究生教材已有一些,其中亦不乏典范之作,但专门针对研究生读者成系列出版的尚为少见。其中的一个原因是各校研究生课程设置自成体系,各校之间不尽相同,这为研究生教材的建设和推广造成了困难。

有鉴于此,来自清华大学等十多所高校、科研单位的教授和专家相约聚首,对通信专业研究生课程体系设置进行探讨,尝试从各校现有的课程体系中提取共同性的知识结构框架,并结合他们多年的教学实践积累,编写一套针对通信专业研究生,兼顾高年级本科生的系列教材,为研究生教育做一点工作。

本系列研究生教材针对性强,知识覆盖较为全面,相信该系列教材的出版将会为读者系统掌握通信科学、信息科学的基础理论与技巧,以及本领域的先进技术方法和现代技术手段提供相对便捷的途径,对培养具有从事通信科学、信息科学以及相关领域的科研与开发和教学工作能力的人才提供有力的手段,对本专业研究生教学起到积极的推动作用。

本系列教材的作者均来自信息和通信学科实力较强的院校,不仅有较为丰富的教学经验,而且在研究方向和地域分布上具有一定的代表性。我有感于他们对教育事业的热忱、对教书育人的执著,遂为之序。

中国工程院院士 李乐民

2007年8月

前 言

近 10 年来,随着网络技术的普及应用和信息化进程的不断推进,网络信息安全问题日显突出。在信息安全技术的发展过程中,密码学得到了长足的发展,保密通信理论已经成为信息安全的主要技术支撑。然而,网络信息安全的需求远不只是保密通信这个方面,还涉及信息的隐蔽性、完整性、真实性、可控性等多个方面。

在信息的隐蔽性方面,信息隐藏技术已经发展起来,如数字水印技术可以起到保护数字版权的作用。在信息的完整性和真实性方面,信息的摘要技术和数字签名技术已经能够支撑电子政务等应用。

信息的可控性是近几年来发展起来的技术,主要包括信息内容的可检查能力和从信息系统中获取信息的能力,是网络安全监管的重要内容。可信计算技术是近几年的研究热点之一,其目的是为了防止信息系统中的资源被未经授权修改。该技术也可以归类到信息的可控性这一分支。它既涉及密码技术,又涉及操作系统技术,是目前被认为最有希望解决病毒攻击等问题的技术方向,因此该技术得到当前国内外信息安全界的高度重视。

当前国内信息安全类教材在上述各个方面的介绍各有侧重。但还没有一本教材全面地讲解信息安全涉及的各种主流技术。另外,由于目前信息安全是一种应用牵引的学科,它主要以解决信息安全应用问题为目标,涉及的技术非常广泛,包括保密通信、入侵检测、协议分析、内容安全、病毒防护等。作者在多年研究以及与同行交流的基础上认为,可以将密码学理论和模式分类理论作为信息安全这门学科的主要基础理论。这一观点供大家讨论。究其原因,信息安全的主要任务是保护信息的安全,其主要工作就是信息的保密性和信息的防窃取。前者的基础是密码学,后者是要进行入侵检测、行为分析、病毒防护等,它的基础理论主要是模式分类。正是基于这点考虑,本书把密码学和模式分类作为网络信息安全的理论基础进行讲解。

本书的读者对象主要是信息安全学科的研究生,通过对本书的学习使他们能够充分了解信息安全的理论、技术范畴及今后的研究方向。同时,本书可作为信息学科高年级本科生关于信息安全专业的基础教材,为进一步从事信息安全技术的研究和工作奠定基础。书中的信息隐藏、安全协议、模式识别三个章节可以作为博士研究生的学习内容。

本书编写分工如下。第一部分是关于密码学理论,主要由宋宇波博士撰写;第二部分是关于安全协议理论,主要由蒋睿博士撰写;第三部分是关于模式分类理论,主要由秦中元博士撰写;第四部分是关于信息隐藏理论,主要由陆哲明教授撰写;第五

部分是关于可信计算导论,主要由黄杰副教授撰写。全书由胡爱群教授统稿和修订。裴文江教授参与了本书部分章节的起草和审稿工作。

由于信息安全学科起步较晚,内容覆盖面很宽,本书难以覆盖所有的知识点。虽然每一个部分的撰写人都是在该方向上具有专业特长的科研人员,有一定的教学经验,但也难免存在疏漏甚至错误之处,恳请读者批评指正,并欢迎与我们联系,以便今后修订时提高本书的质量。对读者提出的宝贵建议,我们深表感谢!

编 者
2007 年 8 月

目 录

第一部分 密码学	(1)
引言	(1)
第 1 章 密码学基础	(3)
1.1 概率论	(3)
1.1.1 基本概念	(3)
1.1.2 随机变量	(4)
1.1.3 概率分布	(5)
1.1.4 Markov 不等式	(6)
1.1.5 Chebyshev 不等式	(6)
1.2 信息论	(7)
1.2.1 熵	(8)
1.2.2 联合熵	(9)
1.2.3 条件熵	(10)
1.2.4 互信息	(10)
1.2.5 冗余度	(11)
1.2.6 密钥含糊度和唯一解距离	(12)
1.3 计算复杂度	(13)
1.3.1 算法的复杂性	(14)
1.3.2 确定性图灵机	(16)
1.3.3 判决问题和语言	(20)
1.4 代数学和数论	(22)
1.4.1 群	(22)
1.4.2 环	(26)
1.4.3 域	(26)
1.5 数论	(27)
1.5.1 整数的唯一分解定理	(27)
1.5.2 同余与剩余类	(31)
第 2 章 对称密码算法	(36)
2.1 引言	(36)
2.2 古典密码	(37)
2.2.1 代换密码	(37)

2.2.2 置换密码	(44)
2.2.3 乘积密码	(46)
2.3 流密码算法	(46)
2.4 DES 算法	(48)
2.4.1 DES 加解密算法	(48)
2.4.2 DES 算法的安全性分析	(52)
2.4.3 3 重 DES 算法	(53)
2.5 AES 算法	(54)
2.5.1 AES 的数学基础和设计思想	(54)
2.5.2 AES 算法描述	(56)
2.6 分组密码操作模式	(63)
第 3 章 公钥密码和数字签名	(66)
3.1 Merkle-Hellman 背包密码系统	(67)
3.2 RSA 密码系统	(69)
3.3 Rabin 密码系统	(70)
3.4 ElGamal 密码系统	(71)
3.5 McEliece 公钥密码系统	(72)
3.6 椭圆曲线密码系统	(73)
3.6.1 椭圆曲线定义	(73)
3.6.2 椭圆曲线运算规则	(74)
3.6.3 椭圆曲线密码算法	(79)
3.7 数字签名	(80)
3.7.1 数字签名算法	(80)
3.7.2 GOST 数字签名算法	(81)
3.7.3 基于离散对数问题的数字签名算法	(82)
3.7.4 Ong-Schnorr-Shamir 签名算法	(83)
3.7.5 ESIGN 签名算法	(83)
3.7.6 特殊数字签名	(84)
第 4 章 量子密码学	(87)
4.1 引言	(87)
4.2 量子力学原理	(88)
4.2.1 量子和量子态	(88)
4.2.2 微观粒子的波粒二象性	(89)
4.2.3 波函数的统计解释	(89)
4.2.4 量子态叠加原理	(89)
4.2.5 海森堡测不准原理	(90)
4.2.6 薛定谔方程	(91)
4.3 量子信息论	(91)

4.3.1 量子位	(92)
4.3.2 多量子位系统	(92)
4.3.3 测量原理	(93)
4.3.4 量子态不可克隆定理	(93)
4.4 量子密码学	(93)
4.4.1 无噪声的 BB84 方案	(94)
4.4.2 有噪声的 BB84 方案	(96)
4.4.3 B92 方案	(97)
4.4.4 E91 方案	(98)
4.4.5 六态方案	(100)
4.5 量子密钥分发协议安全性分析	(100)
习题	(102)
参考文献	(103)
第二部分 安全协议	(107)
引言	(107)
第 5 章 安全协议分析与设计模型	(110)
5.1 引言	(110)
5.2 Dolev-Yao 模型	(111)
5.2.1 引言	(111)
5.2.2 层叠协议模型	(113)
5.2.3 名字标签协议模型	(117)
5.3 BAN 逻辑模型	(122)
5.3.1 引言	(122)
5.3.2 形式化模型	(122)
5.3.3 形式化认证的目标	(126)
5.4 串空间模型	(127)
5.4.1 引言	(127)
5.4.2 串空间	(128)
5.4.3 入侵者	(131)
5.4.4 理想和诚实	(132)
5.4.5 正确性概念	(135)
5.4.6 认证测试方法	(135)
5.5 预言机模型	(137)
5.5.1 引言	(137)
5.5.2 预备知识	(138)
5.5.3 分布式安全通信模型	(139)
5.5.4 实体认证	(141)

第 6 章 典型安全协议	(145)
6.1 引言	(145)
6.2 IPSec 协议	(146)
6.2.1 体系结构	(146)
6.2.2 认证头 AH	(148)
6.2.3 封装安全载荷 ESP	(151)
6.2.4 Internet 密钥交换 IKE	(156)
6.3 SSL 和 TLS	(162)
6.3.1 SSL 的分层结构	(162)
6.3.2 SSL 记录协议层	(163)
6.3.3 SSL 握手协议层	(164)
6.3.4 传输层安全协议 TLS	(168)
6.4 Kerberos 协议	(170)
6.4.1 Kerberos 协议的结构	(170)
6.4.2 Kerberos 交换	(171)
6.4.3 Kerberos 票据标志	(173)
6.5 X.509 证书及协议	(174)
6.5.1 X.509 证书	(174)
6.5.2 X.509 认证协议	(177)
6.6 RADIUS 协议	(178)
6.6.1 RADIUS 协议的特点	(179)
6.6.2 RADIUS 协议的运行过程	(179)
6.6.3 RADIUS 数据包	(181)
6.6.4 RADIUS 计费协议	(185)
第 7 章 安全协议形式化分析与设计	(187)
7.1 引言	(187)
7.2 基于 BAN 逻辑的安全协议形式化分析	(187)
7.2.1 Kerberos 协议分析	(187)
7.2.2 CCITT X.509 认证协议分析	(191)
7.3 基于串空间模型的安全协议形式化分析	(192)
7.4 基于预言机模型的安全协议形式化分析	(198)
7.5 安全协议的形式化设计	(203)
7.5.1 基于类 BAN 逻辑的认证协议设计	(203)
7.5.2 基于串空间模型的安全协议设计	(211)
习题	(219)
参考文献	(219)

第三部分 模式识别	(223)
引言	(223)
第 8 章 模式识别数学基础	(226)
8.1 随机向量及其分布	(226)
8.1.1 随机向量	(226)
8.1.2 分布函数	(226)
8.1.3 随机向量的数字特征	(227)
8.2 多维正态分布	(228)
8.2.1 一维正态密度函数	(228)
8.2.2 多维正态密度函数	(228)
8.2.3 多维正态密度函数的变换	(229)
8.3 向量和矩阵运算	(230)
8.3.1 向量的内积	(230)
8.3.2 向量的外积	(230)
8.3.3 微商	(230)
8.3.4 矩阵的迹和特征向量	(231)
8.4 最优化算法	(232)
8.4.1 拉格朗日乘数法	(232)
8.4.2 梯度下降法	(233)
8.5 小结	(233)
第 9 章 统计模式识别	(234)
9.1 聚类分析	(234)
9.1.1 特征的选取	(235)
9.1.2 距离的测度	(235)
9.1.3 聚类准则	(236)
9.1.4 聚类算法	(237)
9.2 特征选择与提取	(239)
9.2.1 特征提取的必要性	(240)
9.2.2 类别可分性测度	(240)
9.2.3 离散 K-L 变换	(241)
9.3 贝叶斯判别理论	(244)
9.3.1 最小错误率贝叶斯判别	(244)
9.3.2 最小风险贝叶斯判别	(245)
9.3.3 正态分布下的贝叶斯判别	(246)
9.4 线性分类器	(247)
9.4.1 线性可分时的分类器	(247)
9.4.2 广义线性分类器	(248)

9.4.3	Fisher 准则函数	(249)
9.4.4	感知器算法	(250)
9.4.5	梯度下降法	(252)
9.4.6	支持向量机的概念	(252)
9.5	神经网络	(255)
9.5.1	神经网络概念的引入	(255)
9.5.2	神经网络的基本概念	(258)
9.5.3	反向传播算法	(260)
第 10 章	句法模式识别	(261)
10.1	形式语言理论	(261)
10.2	句法结构的自动机识别	(263)
10.3	基元的提取	(264)
10.4	形式语言在模式识别中的应用	(265)
第 11 章	模式识别的安全应用	(268)
11.1	防火墙和入侵检测	(268)
11.1.1	防火墙概述	(268)
11.1.2	防火墙体系结构	(268)
11.1.3	数据包过滤技术分析	(270)
11.1.4	过滤规则的制定	(270)
11.2	人脸识别	(271)
11.2.1	基于主成分分析的人脸识别	(271)
11.2.2	人脸识别过程和示例	(273)
	习题	(274)
	参考文献	(275)
第四部分	信息隐藏技术	(279)
	引言	(279)
第 12 章	隐写术概论	(282)
12.1	基本概念	(282)
12.2	语义隐写术	(285)
12.3	技术隐写术	(288)
12.3.1	纯隐写术、密钥隐写术和公钥隐写术	(288)
12.3.2	时空域隐写术	(289)
12.3.3	变换域隐写术	(289)
12.3.4	信道隐写术	(289)
第 13 章	数字水印技术概述	(291)
13.1	数字水印技术的需求背景	(291)
13.2	数字水印概念	(293)

13.3	数字水印系统的基本框架	(294)
13.4	数字水印及处理技术的分类	(295)
13.5	数字水印技术的应用和特性	(296)
13.5.1	数字水印技术的应用	(296)
13.5.2	数字水印技术的特性	(298)
第 14 章	常用信息隐藏技术	(301)
14.1	时空域隐藏方法	(301)
14.1.1	加性和乘性嵌入规则	(301)
14.1.2	基于最不重要位替换和位平面工具	(302)
14.1.3	利用统计特征	(303)
14.1.4	替换	(305)
14.1.5	量化	(306)
14.1.6	关系	(307)
14.1.7	回声隐藏	(309)
14.1.8	时间标度修改	(310)
14.1.9	幅度调制	(311)
14.1.10	利用字移和行移的文本水印算法	(312)
14.1.11	基于字符间距的水印算法	(312)
14.2	DCT 变换域隐藏方法	(312)
14.2.1	非自适应加性和乘性嵌入方式	(314)
14.2.2	基于量化的嵌入方式	(315)
14.2.3	相位调制	(316)
14.2.4	替换或交换嵌入方式	(316)
14.2.5	基于关系的嵌入方式	(317)
14.2.6	基于零树结构的嵌入方法	(320)
14.2.7	修改直流分量	(323)
14.3	DWT 变换域隐藏方法	(323)
14.3.1	加性和乘性嵌入方式	(325)
14.3.2	基于量化的嵌入方式	(327)
14.3.3	基于替换的嵌入方式	(330)
14.3.4	基于关系的嵌入方式	(331)
14.3.5	基于树结构的嵌入方法	(336)
14.3.6	多分辨率嵌入方式	(338)
14.4	DFT 变换域隐藏方法	(341)
14.4.1	基于幅度调制的嵌入方法	(343)
14.4.2	基于相位调制的嵌入方法	(343)
14.4.3	基于量化的嵌入方法	(344)
14.4.4	直接修改复系数的嵌入方法	(346)

14.4.5	CDMA 嵌入方式	(346)
14.4.6	基于能量关系的嵌入方法	(347)
14.4.7	修改窗函数	(348)
14.5	其他变换域隐藏方法	(350)
14.5.1	Fourier-Mellin 变换	(350)
14.5.2	离散分数傅里叶变换域嵌入技术	(353)
14.5.3	哈德码变换域嵌入技术	(358)
14.5.4	Gabor 变换域嵌入技术	(360)
14.6	压缩域隐藏方法	(361)
14.6.1	JPEG 压缩域嵌入技术	(362)
14.6.2	MPEG 压缩域嵌入技术	(364)
14.6.3	VQ 压缩域嵌入技术	(369)
第 15 章	信息隐藏技术应用	(375)
15.1	版权保护和版权跟踪	(375)
15.1.1	引言	(375)
15.1.2	沿时间轴嵌入水印的视频水印算法	(378)
15.1.3	仿真实验	(382)
15.2	内容认证	(385)
15.3	隐秘通信	(390)
15.4	多媒体信息检索	(393)
15.4.1	特征抽取	(393)
15.4.2	离线多功能水印嵌入	(395)
15.4.3	采用不同查询策略的在线图像检索	(396)
15.4.4	实验结果	(397)
习题		(399)
参考文献		(400)
第五部分	可信计算导论	(409)
引言		(409)
第 16 章	可信计算平台	(412)
16.1	可信平台模块	(412)
16.2	可信计算平台	(416)
第 17 章	可信计算平台的应用	(420)
习题		(426)
参考文献		(426)

第一部分 密码学

引言

什么是密码学? 简单地说,它是关于消息保密的研究的科学。它既是一门科学,也是一门非常古老的艺术,某些学者认为密码学出现在发明书写的时候。如今,密码学随着计算机的广泛应用得到快速发展。纵观其发展历史,可以分为三个阶段:从古代到 1949 年前称为第一阶段,这时的密码学只是一门艺术而非科学;从 1949 年到 1976 年称为第二阶段,1949 年香农的一篇名为《Communication Theory of Secrecy Systems》的论文开创了密码学的科学时代,在这篇论文中,香农创建了基于信息论的密码系统数学模型,从理论上证明密钥数目至少要和可能的消息数目一样多时,才能实现完全保密(perfect security);从 1976 年至今称为第三阶段,1976 年 W. Diffie 和 M. E. Hellman 在发表的《New Directions in Cryptography》论文中首次提出了公钥密码学,将密码学带入快速发展阶段。

各个研究领域都有自己的语言,使用特定的术语以方便理解待研究的对象。术语密码学(cryptology)由 James Howell 于 1645 年提出,该词来源于希腊词根“cryptos”和“logos”,意为有隐藏含义的单词。密码学分为密码编码学(cryptography)和密码分析学(cryptanalysis)两个分支。密码编码学由英国人 Thomas Browne 于 1658 年发明,graphy 来源于希腊词根“graphein”,意为书写。密码编码学即是关于秘密书写的研究的科学。从事密码编码学研究的人称为密码编译专家(cryptographer)。原始的消息称为明文(plaintext),被伪装的消息则称为密文(ciphertext)。最终封装并被发送的消息称为密码(cryptogram)。将明文转换为密文的过程称为加密(encrypt)或是译成密码(encipher)。而将密文转换为明文的相反过程称为解密(decrypt)或是解译(decipher)。密码分析学则与密码编码学相对立,它是关于用数学方法进行密码破解的研究的科学。analysis 来源希腊词根“analyein”,意为解开。从事密码分析的人称为密码分析专家(cryptanalyst)。

进行加密和解密的数学函数称为密码算法(algorithm)或密码(cipher)。它通常包含两个函数:一个用于加密,另一个用于解密。古代密码算法的保密性是基于算法细节的保密,显然这并不安全。现代密码算法函数中引入称为密钥(key)的参数,密钥值的范围即为密钥空间(keyspace),该参数只有通信双方秘密拥有。荷兰人 Kerchoffs 在 1883 年提出了密码算法设计的一个重要原则:“算法的秘密必须全部隐藏在密钥中”。即算法的安全性是基于密钥的安全性而非算法本身的安全性。假设