消息认证码

Message Authentication Codes

裴定一 著

中国科学技术大学出版社

当代科学技术基础理论与前沿问题研究丛书

消息认证码

Message Authentication Codes

裴定一 著

中国科学技术大学出版社

内 容 简 介

保密和认证是信息安全的两个重要方面. 信息的认证用于鉴别信息的真伪,认证方法有无条件安全和计算安全两种类型. 本书主要研究无条件安全的认证理论,介绍了作者在这个领域的研究成果. 首先分别引入了三方(发方、收方和敌方)及四方(发方、收方、敌方和仲裁方)认证系统的完善认证概念,然后用组合设计的语言刻画了这两类完善认证码的结构,在此基础上找到了完善认证码的构造方法. 书中介绍了作者利用有理正规曲线构造的一类三方完善认证码,同时也介绍了其他构造完善认证码的方法,例如基于 t 设计、基于单位指标正交阵列和基于有限几何的构造方法. 本书最后两章研究具有保密功能的认证码的性质和构造方法,附录中简要介绍了基于 Hash 函数的消息认证码.

本书可供学习密码学的大学生、研究生作为教学参考书,也可供数学类专业学生和密码学研究人员参考.

图书在版编目(CIP)数据

消息认证码/裴定一著. 中国科学技术大学出版社,2009.5
(当代科学技术基础理论与前沿问题研究丛书:中国科学技术大学校友文库)
"十一五"国家重点图书
ISBN 978-7-312-02227-2

Ⅰ.消… Ⅱ.裴… Ⅲ.信息系统—安全技术—认证 Ⅳ.TP309

中国版本图书馆 CIP 数据核字(2009)第 060685 号

出版发行 中国科学技术大学出版社
地址 安徽省合肥市金寨路 96 号,230026
网址 http://press.ustc.edu.cn
印 刷 合肥晓星印刷有限责任公司
经 销 全国新华书店
开 本 710mm×1000mm 1/16
印 张 17.5
字 数 290 千
版 次 2009 年 5 月第 1 版
印 次 2009 年 5 月第 1 次印刷
印 数 1—2000 册
定 价 58.00 元

总　　序

侯建国

（中国科学技术大学校长、中国科学院院士、第三世界科学院院士）

大学最重要的功能是向社会输送人才.大学对于一个国家、民族乃至世界的重要性和贡献度,很大程度上是通过毕业生在社会各领域所取得的成就来体现的.

中国科学技术大学建校只有短短的50年,之所以迅速成为享有较高国际声誉的著名大学之一,主要就是因为她培养出了一大批德才兼备的优秀毕业生.他们志向高远、基础扎实、综合素质高、创新能力强,在国内外科技、经济、教育等领域做出了杰出的贡献,为中国科大赢得了“科技英才的摇篮”的美誉.

2008年9月,胡锦涛总书记为中国科大建校五十周年发来贺信,信中称赞说:半个世纪以来,中国科学技术大学依托中国科学院,按照全院办校、所系结合的方针,弘扬红专并进、理实交融的校风,努力推进教学和科研工作的改革创新,为党和国家培养了一大批科技人才,取得了一系列具有世界先进水平的原创性科技成果,为推动我国科教事业发展和社会主义现代化建设做出了重要贡献.

据统计,中国科大迄今已毕业的5万人中,已有42人当选中国科学院和中国工程院院士,是同期(自1963年以来)毕业生中当选院士数最多的高校之一.其中,本科毕业生中平均每1000人就产生1名院士和七百多名硕士、博士,比例位居全国高校之首.还有众多的中青年才俊成为我国科技、企业、教育等领域的领军人物和骨干.在历年评选的“中国青年五四奖章”获得者中,作为科技界、科技创新型企业界青年才俊代表,科大毕业生已连续多年榜上有名,获奖总人数位居全国高校前列.鲜为

人知的是，有数千名优秀毕业生踏上国防战线，为科技强军做出了重要贡献，涌现出二十多名科技将军和一大批国防科技中坚.

为反映中国科大五十年来人才培养成果，展示毕业生在科学研究中的最新进展，学校决定在建校五十周年之际，编辑出版《中国科学技术大学校友文库》，于2008年9月起陆续出书，校庆年内集中出版50种.该《文库》选题经过多轮严格的评审和论证，入选书稿学术水平高，已列为“十一五”国家重点图书出版规划.

入选作者中，有北京初创时期的毕业生，也有意气风发的少年班毕业生；有“两院”院士，也有IEEE Fellow；有海内外科研院所、大专院校的教授，也有金融、IT行业的英才；有默默奉献、矢志报国的科技将军，也有在国际前沿奋力拼搏的科研将才；有“文革”后留美学者中第一位担任美国大学系主任的青年教授，也有首批获得新中国博士学位的中年学者……在母校五十周年华诞之际，他们通过著书立说的独特方式，向母校献礼，其深情厚意，令人感佩!

近年来，学校组织了一系列关于中国科大办学成就、经验、理念和优良传统的总结与讨论.通过总结与讨论，我们更清醒地认识到，中国科大这所新中国亲手创办的新型理工科大学所肩负的历史使命和责任.我想，中国科大的创办与发展，首要的目标就是围绕国家战略需求，培养造就世界一流科学家和科技领军人才.五十年来，我们一直遵循这一目标定位，有效地探索了科教紧密结合、培养创新人才的成功之路，取得了令人瞩目的成就，也受到社会各界的广泛赞誉.

成绩属于过去，辉煌须待开创.在未来的发展中，我们依然要牢牢把握“育人是大学第一要务”的宗旨，在坚守优良传统的基础上，不断改革创新，提高教育教学质量，早日实现胡锦涛总书记对中国科大的期待：瞄准世界科技前沿，服务国家发展战略，创造性地做好教学和科研工作，努力办成世界一流的研究型大学，培养造就更多更好的创新人才，为夺取全面建设小康社会新胜利、开创中国特色社会主义事业新局面贡献更大力量.

是为序.

2008年9月

序

信息的保密和认证是信息安全的两个主要内容.C.E. Shannon 在20世纪40年代首先利用信息论的方法研究保密问题,提出了完善保密系统的概念.G. J. Simmons 在20世纪80年代将信息论的方法应用于研究认证问题,认证码成为构造无条件安全认证系统的密码学基础构件.Gilbert、MacWilliams 和 Sloane 构造了第一个认证码.在三方认证模型中,发方通过一个公共信道给收方发送消息,敌方企图假冒发方发送虚假消息欺骗收方.假定发方和收方互相信任,收方利用与发方约定使用的密钥可以判断所收到的消息是合法的还是虚假的.后来,人们放弃了发方和收方互相信任的假定,发方在发送一个消息后可能抵赖,收方可能谎称收到一个他捏造的消息,为了防止上述可能的欺骗行为,在方案中增添了一个可信的仲裁方,他可以仲裁发方和收方可能发生的争执,这就形成了四方认证系统.

在以上两种模型中,都利用欺骗成功概率的下界来评价一个认证系统.系统中所用的密钥个数也有一个下界,它依赖于欺骗成功概率.可以证明,当密钥个数达到这个下界时,欺骗成功概率也达到下界.密钥个数达到下界的认证系统称为是完善的.完善认证系统具有很规则的结构,可以用组合设计的语言刻画.

在 Simmons 研究的认证系统中,假定一个密钥(编码规则)只使用一次.本书作者研究了一个密钥可以多次使用的认证系统,发现这时的完善认证系统的结构与一类特殊的组合设计——强部分平衡设计密切相关,从而将完善认证系统的构造归结为强部分平衡设计的构造,建立

了认证理论与组合设计理论之间的一座桥梁.在此基础上,作者利用有限域上射影空间的有理正规曲线构造了一类新的强部分平衡设计,从而构造了一类新的完善认证码.

通过简洁明了、易于理解的数学语言,本书提供了一个将实际问题转化为数学问题的极好的例子.这个结果推进了认证理论的研究,开辟了组合设计在密码学中的一个新的应用领域,同时也提出了一个组合设计的新的研究课题——如何构造强部分平衡设计.

本书全面介绍了作者在这个领域所做的工作,同时也介绍了其他相邻的一些工作.全书的十一章安排如下:

第1章通过一些例子引入了认证系统和组合设计的概念,并概要介绍了本书研究认证系统的思路.第2章给出了欺骗成功概率的数学描述.

第3章和第4章分别研究三方和四方认证系统,得到了欺骗成功概率的信息论下界,并由此得到了密钥个数的下界.描述了使密钥个数达到下界的完善认证系统的组合结构的特征,特别讨论了完善 Cartesian 码.此外,也讨论了欺骗成功概率的组合论下界.

第5章给出了基于有限域上有理正规曲线构造三方完善认证码的方法.对于一些特殊情况,给出了这类方案的编码规则的集合.

第6至第8章介绍可以用于构造完善 Cartesian 认证码的几个已知的组合设计,它们是 t 设计、单位指标正交阵列和基于有限几何的设计.

第9章给出了几个带仲裁(即四方)完善 Cartesian 码.

本书的最后两章研究同时具有认证和保密功能的系统,仍考虑一个密钥可以被多次使用.就 U(无序)保密和 O(有序)保密两种情况,分别引入完善保密的概念.第10章讨论 U 保密码及具有 U 保密功能的认证码的构造.第11章讨论 O 保密码及具有 O 保密功能的认证码的构造.

本书是在作者所著的 *Authentication Codes and Combinatorial Designs*(Chapman & Hall / CRC, 2006年)一书基础上改写而成.增补了一些新内容(如1.3节,5.5节),为了便于读者理解,多处论述作了简化.全书由原来的九章增为十一章.附录的内容更新为基于 Hash 函数的消息认证码,以便读者对消息认证码的研究有较全面的了解.

作 者

2007年12月于广州

目　次

第1章 引 言

1.1 什么是消息的认证

人们很早以前就知道保护敏感信息秘密的重要性.密写形成了一个称为密码学的领域,但长期以来它主要被应用在军事和政府部门.由于强大的信息技术的应用,现代信息的存储和传输已变得非常简便,大量信息在传输中很容易被他人介入,这对密码学提出了许多新的问题.例如,敌方不仅可以读到传输中的信息,也可以改变这些信息;或者敌方可以生成和发送一个虚假信息给收方,以期达到某种目的.

作为一个例子,我们想象一个窃贼可以非法接入银行的计算机通信线路.窃贼去银行在他的账户中存入100元后,银行要给中心计算机发一个消息,通知计算机在该窃贼的账户中增加100元.窃贼进入银行通信线路后,就可以改变银行所发消息中的存款数字,例如改为增加1 000元.另一种可能是窃贼将银行发出的消息收录下来,然后将同样的消息重复发给中心计算机,每发一次就在他的账户中增加100元.这个例子表明,必须有某种机制来检查,使得只有银行发出的消息才能被中心计算机接受.这就是我们所说的消息的认证.

当收方收到一个消息,例如一个电子邮件时,他关心谁是真正的发方,这消息的内容在传输中是否已被他人非法篡改.这就是消息认证所关心的两大问题.

保密和认证成为当代信息安全的两个重要方面.保密系统的作用是保护秘密,认证系统的作用则是使消息得到认证.保密和认证是信息安全中两件不同的事情,在有些情况下保密是主要的,但有些情况下就不是这样.一个认证系统可以带保密功能,也可能不带保密功能.在两个军事单位之间的通信,通常既需要保密又需要认证.下面是一个只需认证不需保密的例子.一个公司在马路上安装了很多停车计时表,由公司雇员经常从表中取出钱交给公司.公司老板担心雇员将取出的一部分钱装入自己口袋,所以计时表中所收集的钱数必须有一个认证,例如,可以让计时表打印一张纸条,上面有所收集的钱数,雇员必须把钱和纸条同时交给老板.雇员可以数清从计时表中取出多少钱,因而这个钱数对于雇员不是秘密.在有些类似情况下,敌方已经知道所认证的消息的内容,因而可以利用不带保密功能的认证系统.

在很多认证中,通常有发方、收方和敌方,发方想给收方发送信息,敌方想利用篡改发方所发的信息或制造一个虚假信息的手段欺骗收方.在有些情况下发方和收方可能是同一个人,例如对于计算机中所存储的数据文件的认证,所存储的数据可能包含一些敏感信息,例如人员的工资表、学生的成绩单或类似的数据.通过对数据文件的认证,发现其中任何可能有的非法改动.

若发方和收方不是同一个人,他们可能是互相信任的,即他们不会相互欺骗.但在很多情况下可能不是这样,他们可能会相互欺骗.在证券市场中,客户向证券经纪人发出一个电子信息,要求他买入或卖出股票,这时在客户和经纪人之间可能会出现争执.若在客户发出买入某股票的请求后,该股票价格下跌,他可能会否认他所发出的请求.另一方面,经纪人也可能声称收到买股票的请求,但实际上客户并没有发出那样的请求.为了解决这类争执,必须有一个仲裁来判断哪一方是不诚实的.

如果敌方欺骗成功的可能很小,则认证系统就是安全的.一个系统的安全性若不信赖于敌方所拥有的计算能力,则该系统称为**无条件安全**.当假定敌方仅拥有有限的计算能力时系统是安全的,该系统称为**计算安全**.譬如数字签名是一个认证系统,它是计算安全,因它的安全性信赖于在假定敌方仅拥有有限的计算能力时某个困难问题不能有效求解(如大整数因子分解、离散对数的计算).本书限于讨论无条件安全的认证系统(带或不带保密

功能).

1.2 认证系统

为了简单起见,假定发方仅可能发送两个消息:进攻或撤退,这两个消息分别用字母 A 和 W 表示.发方可能用发数字"0"代表 A,数字"1"代表 W;或者用"0"代表 W,用"1"代表 A.所以发方有两个密钥:密钥 1 和密钥 2,生成表 1.1.

表 1.1

	0	1
密钥 1	A	W
密钥 2	W	A

发方每次通信时利用某一个密钥,只有收方知道他用的是哪一个密钥.当敌方观察到发方所发出的"0"或"1"时,他不知道它们的含意,因为不知道发方用的是哪一个密钥.所以,该系统具有保密功能.假定每个密钥被使用的概率是 1/2,敌方猜测所发报文的含意,他成功的概率是 1/2.

为了进一步的讨论,这里必须提及 Kerkhoff 假设,这是密码学所有领域都采用的一个著名假设.该假设说,对敌方保密的系统参数仅仅是正在使用的密钥,其他的所有参数,如系统的结构、系统的概率分布等,都是公开的.该假设对分析一个密码系统的安全性是合理的、必需的.

在发送消息之前,发方和收方约定所使用的密钥,按 Kerkhoff 假设,敌方知道表 1.1 的内容,但他不知道在某时刻使用哪个密钥.这时敌方仍有可能进行欺骗,在观察到发方发送的信号之前,敌方向收方发出"0"或"1",收方如果接受该信号,认为它是发方发来的,并据此采取行动.因而敌方使收方有了行动,尽管敌方事先并不知道收方有什么行动.这类欺骗称为假冒攻击.

敌方也可以采取另一类攻击,敌方可以将观察到的信号由"0"改为"1",或由"1"改为"0",篡改后的信号仍能被收方接受,并使收方采取与发方指示相违背的行动.这类欺骗称为替代攻击.

表 1.1 的系统没有认证功能.称发方要发送的消息(上例中的 A 和 W)为**信源**,发方真正发送的信号(上例中的"0"和"1")为**报文**.通过增加报文的个数可以得到具有认证功能的系统.如上取{A, W}为信源集合,{00, 01, 10, 11}为报文集合,考虑表 1.2 中所给出的系统.

表 1.2

	00	01	10	11		00	01	10	11
密钥 1	A	—	W	—	密钥 3	—	A	W	—
密钥 2	A	—	—	W	密钥 4	—	A	—	W

该系统中有 4 个密钥.当选用密钥 1 时,报文"00"代表 A,"10"代表 W,而"01"和"11"不被使用.称"00"和"10"为密钥 1 的有效报文,而"01"和"11"为密钥 1 的无效报文.类似地考虑其他密钥,每个密钥有两个有效报文和两个无效报文.收方仅接受该时刻所采用的密钥的有效报文,而拒绝接受无效报文.每个报文对于四个密钥中的两个密钥是有效的.假定采用每个密钥的概率都相同(1/4).当敌方发起假冒攻击时,他任意挑选一个报文,例如"00",发给收方,"00"是该时刻所使用的密钥的有效报文的概率为 1/2.当敌方发起替代攻击时,他用一个虚假报文替代发方所发送的报文.假设敌方观察到发方所发送的报文是"00",敌方便知道该时刻所使用的密钥为密钥 1 或密钥 2,报文"10"和"11"是其他两个可能的有效报文,他任取其中一个报文发给收方,他的成功概率为 1/2.因此,表 1.2 中的系统具有认证功能.从表 1.2 可以看到,报文"00"和"01"永远传递信源 A,而报文"10"和"11"永远传递信源 W.所以,任何人观察到发方发的报文,就能知道该报文的含意.这是一个没有保密功能的系统.

如果把表 1.2 改为表 1.3,就可以得到具有保密功能的系统.

表 1.3

	00	01	10	11		00	01	10	11
密钥 1	A	—	W	—	密钥 3	—	W	A	—
密钥 2	W	—	—	A	密钥 4	—	A	—	W

应用类似于分析表 1.2 的系统的方法分析表 1.3 的系统，可以知道对于表 1.3 的系统，假冒攻击和替代攻击的成功概率仍然都是 1/2，但该系统带保密功能，因每个报文都可用于代表 A 和 W，敌方猜对发方所传输的信源的概率是 1/2.

1.3 欺骗成功概率和编码规则个数

设计认证系统的一个目标，是要使敌方欺骗成功概率尽可能小. 此外，由于编码规则个数越多，收、发双方传递编码规则所需的传输信息量和存储信息量也越大，所以在设计认证系统时，在达到一定的安全性的前提下，追求尽可能小的编码规则个数. 本节分别讨论不带保密的认证系统和一般(可以带保密)的认证系统两种情形.

1. 不带保密的认证系统

考虑一类特殊形式的不带保密的认证系统. 设 $\mathcal{S}=\{s_1, s_2, \cdots, s_k\}$ 为所有可能的 k 个信源. 若 $s \in \mathcal{S}$ 为发方所要传递的信源，发方实际发送的报文形如 (s, a)，a 称为 s 的认证字. 见到报文 (s, a) 也就知道它所代表的信源 s，所以这时没有保密. 设 $\mathcal{A}=\{a_1, a_2, \cdots, a_l\}$ 为所有可能的 l 个认证字的集合. 一个编码规则 e 是 $\mathcal{S}$ 到 $\mathcal{A}$ 的一个映射，每个信源对应一个认证字，代表信源 s 的报文就是 $(s, e(s))$. 所有编码规则的集合记为 $\mathcal{E}$. 这类认证系统称为系统 Cartesian 认证码(图 1.1)，它由三个集合 $\{\mathcal{S}, \mathcal{A}, \mathcal{E}\}$ 决定.

在应用中常见的一类消息认证码(MAC)中，选用一个带密钥的 Hash 函数 $H_k(x)$，信源 s 的认证字取为 $a = H_k(s)$. 攻击者见到报文 $(s, H_k(s))$ 后，若能找到另外一个信源 s'，使 $H_k(s) = H_k(s')$，则报文 $(s', H_k(s'))$ 也能被收方接受，这时该认证系统就不安全了. Hash 函数 $H_k(x)$ 的输入 x 的比特长度为任意，其输出 $H_k(x)$ 有固定的比特长度(例如 160 比特)，从理论上讲，一定存在另一个信源 s'，使 $H_k(s') = H_k(s)$.

但在实际上，要找到 s' 在计算上往往是困难的，所以该认证系统的安全称为计算安全.

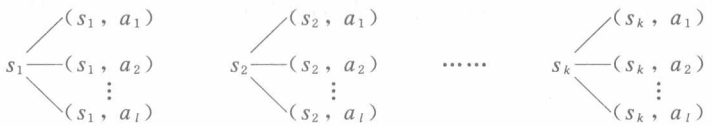

图 1.1 系统 Cartesian 认证码

在通信开始之前，收方和发方秘密约定选用一个编码规则 e. 分别以 P_0 和 P_1 代表敌方最佳假冒攻击和替代攻击的成功概率. 敌方进行假冒攻击时，他选定一个信源 s，但他不知道认证字 $e(s)$，只能在 $\mathscr{A}$ 中任取一个认证字 a，将 (s, a) 发给收方. 仅当 $a = e(s)$ 时，收方才能接受报文 (s, a). 假定敌方的攻击是最佳的，敌方选取的 a 等于 $e(s)$ 的概率至少为 $1/l$，所以敌方最佳假冒攻击成功概率

$$P_0 \geqslant 1/l. \tag{1.1}$$

敌方进行替代攻击时，他首先截获发方发送的一个报文 (s, a)，然后选取另外一个信源 s' 替代 s，他同样不知道 $e(s')$，只能在 $\mathscr{A}$ 中任取一个认证字 a'，将 (s', a') 发给收方. 仅当 $a' = e(s')$ 时，收方才能接受报文 (s', a'). 敌方选取的 a' 等于 $e(s')$ 的概率至少为 $1/l$，所以敌方最佳替代攻击成功概率

$$P_1 \geqslant 1/l. \tag{1.2}$$

更确切地说，仅当发方和收方选用的编码规则 e 适合 $e(s) = a$ 时，敌方的假冒攻击才能成功. 假定每个编码规则被选用的概率相同，则敌方发报文 (s, a) 作假冒攻击的成功概率为

$$\frac{|\{e \in \mathscr{E} \mid e(s) = a\}|}{|\mathscr{E}|},$$

它的分子是使 s 的认证字为 a 的编码规则的个数，分母是编码规则的总数，敌方最佳假冒攻击成功概率为

$$P_0 = \frac{\max\limits_{s \in \mathscr{S}} \max\limits_{1 \leqslant i \leqslant l} |\{e \in \mathscr{E} \mid e(s) = a_i\}|}{|\mathscr{E}|}. \tag{1.3}$$

由于任一编码规则都会将信源 s 映射为 $\mathscr{A}$ 中某个认证字，故

$$\sum_{i=1}^{l} |\{e \in \mathscr{E} \mid e(s) = a_i\}| = |\mathscr{E}|. \tag{1.4}$$

因极大值不小于平均值，得到

$$\max_{1\leqslant i\leqslant l} |\{e \in \mathscr{E} \mid e(s) = a_i\}| \geqslant \frac{1}{l}\sum_{i=1}^{l} |\{e \in \mathscr{E} \mid e(s) = a_i\}|$$

$$= \frac{|\mathscr{E}|}{l},$$

利用式(1.3)同样得到式(1.1)，且可见其中等号成立(极大值等于平均值)的充分必要条件是对任一 $s \in \mathscr{S}$ 及 $1 \leqslant i \leqslant l$，

$$|\{e \in \mathscr{E} \mid e(s) = a_i\}| \tag{1.5}$$

是常数 $(|\mathscr{E}|/l)$.

类似地分析最佳替代攻击成功概率 P_1. 敌方观察到发方发送的报文 (s, a)，然后用报文 (s', a') $(s' \neq s)$ 进行替换，将它发给收方，收方能接受 (s', a') 的概率为

$$\frac{|\{e \in \mathscr{E} \mid e(s) = a, e(s') = a'\}|}{|\{e \in \mathscr{E} \mid e(s) = a\}|}.$$

当敌方观察到报文 (s, a) 时，他就能确定这时发方和收方选用的编码规则属于集合 $\{e \in \mathscr{E} \mid e(s) = a\}$. 进一步，如果发方和收方选用的编码规则属于集合 $\{e \in \mathscr{E} \mid e(s) = a, e(s') = a'\}$，则敌方的替代攻击就能成功. 所以，当敌方观察到报文 (s, a) 后，他的最佳替代攻击成功概率

$$P_1(s, a) = \frac{\max\limits_{s'\in\mathscr{S}, s'\neq s} \max\limits_{1\leqslant i\leqslant l} |\{e \in \mathscr{E} \mid e(s) = a, e(s') = a_i\}|}{|\{e \in \mathscr{E} \mid e(s) = a\}|}.$$

同样由于

$$\sum_{i=1}^{l} |\{e \in \mathscr{E} \mid e(s) = a, e(s') = a_i\}| = |\{e \in \mathscr{E} \mid e(s) = a\}|, \tag{1.6}$$

故