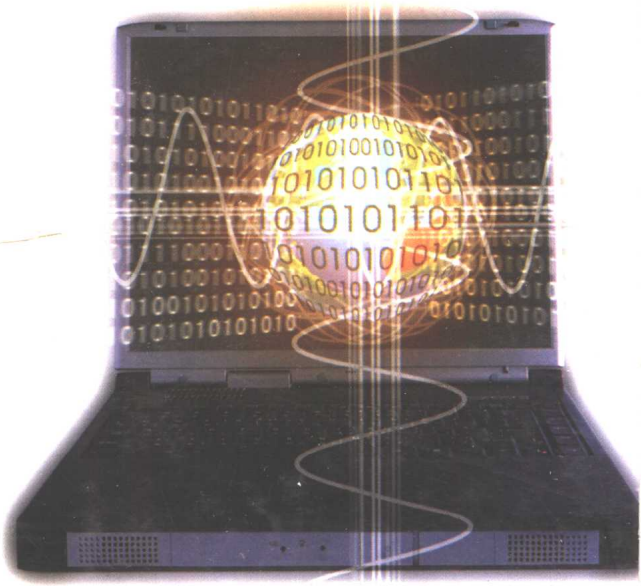
行为生物统计特征身份认证的理论与应用

朱树人　编著



国防科技大学出版社

本书获得长沙电力学院学术专著出版基金资助

行为生物统计特征身份认证的理论与应用

朱树人　编著

国防科技大学出版社

·长　沙·

内 容 简 介

本书重点介绍了基于网络应用的身份认证技术和行为生物统计特征身份认证技术，包括：计算机系统安全控制分布、身份认证理论与技术、动态手写签名识别的理论基础、生物测定加密技术、手写识别技术、认证信息的安全性、完整性、不可否认性、基于 INTERNET 的认证协议、基于 INTERNET 应用的客户认证协议、动态手写签名认证系统的设计等技术。本书特别强调生物测定加密技术在计算机网络信息系统中的应用，提供了常用的算法。可供计算机、通信与电子系统、信息工程、金融电子化等专业的工程技术人员使用和参考，亦可作大学高年级本科生和研究生的教材。

图书在版编目(CIP)数据

行为生物统计特征身份认证的理论与应用，朱树人编著 .—长沙：国防科技大学出版社，2002.6

ISBN 7-81024-861-8

Ⅰ.行… Ⅱ.朱… Ⅲ.行为生物统计特征—身份认证—计算机网络安全 Ⅳ.TP309

国防科技大学出版社出版发行

电话：(0731)4572640 邮政编码：410073

E-mail：gfkdcbs@public.cs.hn.cn

责任编辑：黄八一 责任校对：潘生

新华书店总店北京发行所经销

国防科技大学印刷厂印装

*

850×1168 1/32 印张：7 字数：175 千

2002 年 6 月第 1 版第 1 次印刷 印数：1-600 册

*

定价：14.00 元

前　言

计算机安全的研究，一般可分成物理安全和逻辑安全两个方面。物理安全研究包括：安全地放置设备，使之远离火灾、水灾、电磁辐射等；物理访问控制，如口令、指纹鉴别（鉴定用户身份及合法性）；安全管理；防止火灾和水灾等自然灾害；防止电磁泄露等。逻辑安全研究包括三个方面：信息的保密性（保护信息不被非授权泄露）；信息的完整性（保护信息不被非授权修改或破坏）；信息的可用性（避免拒绝授权访问，也称避免拒绝服务）。网络安全问题涉及许多方面，但对于大多数网络应用，尤其是电子商务这样的商业应用来说，身份认证是其服务过程中的关键环节，如果没有强有力的身份认证手段，许多商业性应用根本无法开展。

身份认证包括用户向系统出示自己的身份证明和系统查核用户的身份证明的过程，它们是判明和确定通信双方真实身份的两个重要环节。

单机状态下的身份认证有根据人的生理特征进行身份认证、根据约定的口令等进行身份认证以及采用硬件设备进行身份认证三种。

网络环境下，由于任何认证信息都是在网上传输的，所以其身份认证较为复杂，一方面，由于验证身份的双方一般都是通过网络而非直接交互，所以许多单机上的认证手段都无法实现。另一方面，大量的黑客随时随地都可能尝试向网络渗透，对认证信息进行攻击，所以网络身份认证必须防止认证信息在传输或存储过程中被截获、篡改和冒名顶替，同时也必须防止用户对身份的抵赖。

身份是指从行政、法律或经济、社会方面，确定一个人的地位或权利。身份证则是记载持证人必要情况的证明材料、官方文件或卡片，它可以表明持证人的个人身份。识别身份首先要解决证件本身的真伪。其次要确认持证人与证件的同一性，前者属于证件防伪技术，后者是本书讨论的用生物测定(统计)学方法，通过识别持证人的某种身体或行为特征，与证件所载有的资料比对，确认同一性，证明持证人的身份。生物测定(统计)学方法是根据个人唯一的生理特征或行为特征，来查证或辨认、识别某个活人身份的自动化方法。主要用于实体认证的生物特征有：指纹、视网膜、虹膜、面部特征以及其他行为生物特征(语音特征 、手写签名特征)等。

生物测量学源自基于生理学和行为特征的自动鉴别。这种鉴别方法在各方面都优于传统的口令和 PIN 号。个人识别必须是身体上存在的某一特征。基于生物测量技术的识别消除了记忆口令和携带标记的困难。由于作为信息媒介物的计算机使用的增加，限制访问敏感和个人数据是必要的。用生物测量技术代替 PIN 号能预防未授权的访问和欺诈使用 ATM、移动电话(手机)、智能卡、桌面计算机、工作站和计算机网络。PIN 号、口令可能被遗忘，基于标记的识别方法像护照、驾驶证可能被伪造和丢失。然而生物测量识别系统可以享受重新申请的乐趣。各种类型的生物测量系统可用在实时识别。最受欢迎的是脸型和指纹匹配，还有其他生物测量系统如 DNA、虹膜、视网膜、语音、面部温度记录、手型、手写签名等。

目前，备受瞩目的研究领域之一就是用于身份验证的统计技术领域。发表于 GartnerGroup 的 Web 杂志《 Business Technology Journal 》中预测的 10 大技术包括：生物测量学、桌面会议、数据挖掘、文档成像、电子现金、网络计算机、个人数字助理、push 技术、智能卡、语音识别。签名验证即属于这类统计技术。在大多数情

况下，签名验证需要借助于数字化仪来在线获取，并用光学扫描仪进行离线转换。这些接口要获取信号需要一定空间而且很笨重，并且使用起来也复杂，这就增大了整个确认系统的复杂性。与此相反，电子书写板个头较小且操作较简单，因而逐渐盛行在当前计算机技术界。手写输入接口获取签名进行身份验证是可行的，这种接口可让用户在比较自然和舒适的环境下与计算机进行交互。这种基于手写签名的身份验证系统可归结为完全基于笔迹可视化的计算机环境的一部分。

生物测定方法的重要性能是它的识别能力，其指标可以用一对术语来表示，即错误拒绝率(FRR)和错误接受率(FAR)。许多机器都有一个阈值，以期望 FRR 与 FAR 之间的平衡。如果把容许量设定点收紧，使冒名顶替者难以通过，则被允许进入的人，其通过也将变得很困难。相反，如果被允许的人非常容易通过，那么冒名顶替者也容易混入。

签字识别的最大优点是全世界的人们都已习惯这是辨认自身的方式。由于签字的自动化检验在金融界长期的潜在势头，签字动态是生物测定学方法开发中的热点。该领域已发布了上百个专利，一些大公司如 IBM、NCR 和 VISA 等，每家都拥有几项专利。现在有几家公司提供商业产品，不同厂家使用各自的技术，根据微小差异的原理，从不同方面观察签字的动态过程。签字动态的关键是区分签字的习惯部分，和几乎每次签字都有所变化的部分。静态图像也作为因素计算在内，还有一些装置能够捕获静态图像作为记录和复制之用。

笔迹识别虽然被广泛研究了许多年，但是到目前为止仍是一个尚待解决的问题。签名验证问题对于将识别错误率提高到与人类识别错误率可比的程度的研究人员来说，仍是一个挑战。针对签名验证的文献非常广泛，它们显示了研究的两个主要领域：离线系统和在线系统。离线系统处理签名的静态图像，比如签名动作

的结果，而在线系统处理签名的动态过程，比如签名本身的动作。

本书写作框架和大纲由我拟定。邹阿金副教授撰写了第二章的第一节至第三节，其余章节由我负责执笔。全书最后由我统一修改定稿。

在本书的写作出版中，得到了各方面的热情帮助：北京航空航天大学博士研究生导师李伟琴教授对本书的写作提出了很好的指导和建议；北京航空航天大学博士研究生黄冬泉副教授、北京航空航天大学博士研究生黄河同志提供了很多有价值的资料；长沙电力学院出版基金给予了经费上的支持；长沙电力学院副院长柳见成教授、科研处处长刘和云教授给予了热情的鼓励，关心着书稿的正式出版；长沙电力学院工程师贺株莉同志为本书收集和整理了大量文献资料；国防科技大学出版社黄八一编辑对本书给予了热情的帮助和扶持，为本书的出版费了许多心血；书中参阅了大量的国内外论著，吸收了不少同行学者的研究成果，引用了大量的数据和资料。特此一并致谢。

由于我们的水平有限，编写的疏漏和错误在所难免。诚恳地欢迎同行专家和使用本书的同志提出宝贵的批评和意见。

朱树人

2002 年 1 月

目　　录

第一章 绪 论

随着微电子、光电子、计算机、通信和信息服务业的发展，Internet 已经得到了广泛的应用，从过去主要用于科研和简单信息发布向商业化大踏步地迈进，尤其是以 Internet 为支撑平台的电子商务已经引起了社会各界及商家的普遍重视，并正在成为网络经济时代的主要标志之一。

在创建 Internet 的初期，安全性并不是 Internet 初创人所考虑的一个首要问题，他们假定网络或者是专用的、孤立的和物理上安全的，或者就是完全开放的，但是随着 Internet 应用越来越商业化，特别是电子商务的应用，在这个开放网络上传输的保密信息也越来越多，网络的安全性已引起学术界的普遍重视。尽管已有许多密码技术保护通信信息的安全，但是其中许多实现都是专用的，要求通信双方使用相同的软件。目前已有一些建议标准，容许不同系统在 Internet 上安全地交换数据，IETF(Internet Engineering Task Force)工作组正在考虑采纳这些标准。

网络安全问题涉及许多方面，但对于大多数网络应用，尤其是电子商务这样的商业应用来说，身份认证是其服务过程中的关键环节，如果没有强有力的身份认证手段，许多商业性应用根本无法开展。

在 Internet 进入迅速发展之初，也就是基于 Internet 的商业活动开始大规模启动的时候，数字认证技术便已引起了人们的关注。数字认证技术的目标是要在 Internet 上解决“我是谁”的问题，就如同我们每一个人都会拥有一张证明个人身份的身份证或

是驾驶执照，以“表明”我们的身份或某种资格。当 Internet 逐渐由科研、教育工具转变为一种商业工具之后，这种“表明”变的更加重要。

身份认证包括用户向系统出示自己的身份证明和系统查核用户的身份证明的过程，它们是判明和确定通信双方真实身份的两个重要环节。

单机状态下的身份认证有根据人的生理特征进行身份认证、根据约定的口令等进行身份认证以及采用硬件设备进行身份认证三种。

网络环境下，由于任何认证信息都是在网上传输的，所以其身份认证较为复杂，一方面，由于验证身份的双方一般都是通过网络而非直接交互，所以许多单机上的认证手段都无法实现。另一方面，大量的黑客随时随地都可能尝试向网络渗透，对认证信息进行攻击，所以网络身份认证必须防止认证信息在传输或存储过程中被截获、篡改和冒名顶替，同时也必须防止用户对身份的抵赖。

第一节　计算机系统安全与安全控制分布

网络化和分布式是目前计算机系统的主要特点之一。信息共享与资源共享带来了巨大的经济效益，但同时也伴随着更严重和更复杂的安全问题。研究在网络和多用户系统环境下的计算机安全问题是目前最主要的研究课题之一。

计算机安全的研究，一般可分成物理安全和逻辑安全两个方面。

物理安全研究包括：安全地放置设备，使之远离火灾、水灾、电磁辐射等；物理访问控制，如口令、指纹鉴别（鉴定用户身份及合法性）；安全管理；防止火灾和水灾等自然灾害；防止电磁泄漏等。

逻辑安全研究包括三个方面:信息的保密性(保护信息不被非授权泄漏);信息的完整性(保护信息不被非授权修改或破坏);信息的可用性(避免拒绝授权访问,也称避免拒绝服务)。

本书主要涉及逻辑安全的研究。以后所提计算机安全实际是指逻辑安全。

更通俗地说,计算机安全的研究就是研究控制计算机使用的方式。对实际计算机系统,就是研究应当有什么样的安全保护需求(或目标),以及如何采取硬件、软件、人工一体化的手段达到这些安全保护目标。前者是安全策略的研究问题,后者是安全策略的实施问题或安全控制与安全管理问题。

一、安全策略与安全模型

安全策略是安全体系结构中最重要的成份。根据美国国防部《可信计算机系统评价准则(TCSEC)》的定义,安全策略是规定一个组织如何管理、保护和分配敏感信息的法律、规则和实践经验的集合。由于自然语言的模糊性会导致策略理解的多义性和不准确性,因此需要构造形式化的安全策略模型(简称安全模型),以便对安全策略所表达的安全需求进行简单、抽象和无歧义的形式化描述。安全模型是设计和开发高可信度安全系统的前提和指南,是系统安全形式化说明与验证的关键一步,是判断系统安全性和检测系统安全缺陷的重要依据。表 1 - 1 是 TCSEC 给出的各个可信等级的计算机系统的特点。从表中可看出:为了达到 B2 级或 B2 以上的级,系统必须有形式化安全模型。虽然 TCSEC 是根据军事安全保密的需求而提出的,但它能说明研究安全模型的重要性和必要性。

表 1-1 TCSEC 的等级

类别	名称	主要特征
A1	验证设计	形式化的最高级描述和验证，形式化的隐通道分析，非形式化的代码一致性证明
B3	安全区域	存取监视器(安全内核)，高抗入侵能力
B2	结构化保护	形式化模型，隐通道约束，面向安全的体系结构，较好的抗入侵能力
B1	有标识的安全保护	强制访问控制，安全标识，删去安全相关的缺陷
C2	受控存取保护	单独的可说明性，广泛的审计，附加软件包
C1	任意安全保护	自主访问控制，在共同工作的用户中防止事故
D	低级保护	不分等级

在安全策略上，与安全模型构造技术息息相关的技术包括：

确立安全策略的理论与方法：对于不同的应用，如何自动或半自动地确定相适应的安全策略。

安全模型到安全策略实施的映射与验证。安全策略、安全模型及安全策略实施之间的映射关系如图 1.1 所示。安全模型架起了安全策略与安全策略实施之间的桥梁。

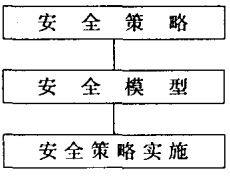


图 1.1 安全策略、安全模型及安全策略实施之间的映射关系

二、安全策略的实施与安全控制分布

安全策略的实施机制，简称安全机制，如，密码机制(加密，解密)、密钥管理机制、数字签名机制、访问控制机制等都是安全机

制。安全机制有很多种，大体可分为三类：预防机制（如访问控制机制）、检测机制（如病毒检测机制，入侵检测系统）和恢复机制（如系统备份）。一般来说，安全策略的实施需要结合一类或多类安全机制。

实施了某种安全策略的计算机系统称为满足该安全策略的计算机系统。可信计算机系统（Trusted Computer System）是采用了充分的软硬件完整性措施的系统，它可以同时处理一些安全的信息。

可信计算基（Trusted Computing Base 简称 TCB），是计算机系统内保护机制的总体，包括硬件、固件和软件，他们结合起来负责安全策略的实施。它建立了基本的保护环境，提供了可信计算机系统要求附加的用户服务。TCB 能正确实施安全策略的能力仅仅取决与 TCB 内的机制和系统管理人员正确地输入有关安全策略的参数（如用户密级）有关。

一个系统的安全控制分布在硬件、固件、操作系统、工具、应用程序、人工（系统安全员）等各个层次上（图 1.2）。下面就上面四个层次作简要说明。

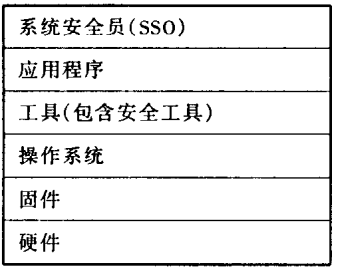


图 1.2　安全控制分布

逻辑安全的目标就是确保信息的保密性、完整性和可用性（即避免拒绝服务）。更通俗地说，计算机安全的研究，就是研究控制计算机使用的方式。对实际计算机系统，就是研究应当有什么样

的安全保护需求(或目标),以及如何采取硬件、软件、人工一体化的手段达到安全保护目标。前者是安全策略的研究问题,后者是安全策略的实现问题或安全控制与安全管理问题。

安全策略通常用自然语言表示。但是,自然语言的模糊性会导致策略理解的多义性和不准确性,因此需要构建形式化安全策略模型,以便对安全策略所表达的安全需求进行抽象和无歧义的形式化描述。形式化安全策略模型是设计和开发高可信度安全系统的前提和指南,是系统安全形式化说明与验证的依据。因此它的研究具有非常重要的意义。

在计算机安全研究早期,保密性研究就得到了极大的重视,也取得了成功,提出了一些比较成功的保密性安全模型,如 Bell - LaPadula 模型。目前基于 Bell - LaPadula 模型和多级安全机制而实施的保密性技术已经比较完善。进入 20 世纪 80 年代后,人们逐步认识到完整性问题的重要性,进行了深入研究,提出了一些完整性模型,如 Biba 模型、Lipner 模型、Clark - Wilson 模型。已见到的非形式化的 Clark - Wilson 模型构造了一个较好的完整性安全策略框架。避免拒绝服务研究难度大,进展甚微,目前这方面还没有公认成功的模型。

在对计算机安全模型的研究中,我们发现现有的安全模型存在许多缺陷,主要体现在:①他们只能表示特殊的安全策略或某一类安全策略,特别是不能同时面向保密性、完整性和避免拒绝服务。②不适用于含有并发、不确定等特性的分布式系统。③模型的描述能力和分析能力在其所声称的范围内仍然有限。例如,Bell - LaPadula 模型未能体现最小权力原则。Clark - Wilson 模型未考虑多粒度动态层次的完整性,未考虑时序限制、并发限制、不可中断状态流限制,同时它把责任分离这个难题交给了安全管理人员,构成了该模型理论上虽无安全漏洞,实际上的漏洞却难以发现和预防的结果。④模型的柔性能力低。

第二节　生物统计身份认证系统综述

一、生物统计身份认证系统

目前已经有多种通过独一无二的生理特征鉴别实体的技术方法。一些方法用指纹、手形、手、脸或者视网膜的脉管结构、虹图、音频图、步态、甚至整脸视频识别。

这些生理特征中的每一种对每个个体而言都是独有的。即使遗传上相同的双胞胎在几乎所有的这些识别符中也有区别。通过利用这些生物度量可以产生一种可以唯一地识别任何个体的数学表现形式。

生物统计学被典型应用的第一种方法就是鉴别系统。一个使用者可以有密码、用户名或者其他标识符，包含生物统计学的数字表记录。用户可以用他们的指纹、声频图代替口令。系统处理这些输入并把它与模板比较。因为皮肤是有弹性的(如：指纹，肿胀的手指的指纹)，手形、声频(感冒)识别，系统必须有灵活度，而不仅仅是将一个图像叠加到一个上面来决定是否匹配，匹配的衡量关系即可信度，通常用 0 ~ 100，大多数系统都有一个可调整的匹配门限值，来决定承认(匹配 = 真实)和拒绝(匹配 = 虚假)。

1. 指纹

现今最古老最健全的技术是指纹鉴定。在 20 世纪 60 年代许多警察和政府办事处就开始了对他们收集的原始底图指纹卡片进行数字处理。随着计算效率和扫描技术进步，拍摄一个指纹并将它同模板相比较的能力变得更实际。在典型的系统中，手指是在

一块一平方英尺的区域中扫描的。扫描器以每英尺 500 点，每像素 80 × 80 环节强度的灰度工作。

图像拍摄后，在探测它的任何特征前，先进行强化处理，图像被调整为 1bit 黑白图像，其中每个线条都变为一个像素的宽度。

然后，指出一个经过强化处理图像的一些细节，这些细节是一些点，显示线条在哪里结束或分叉。然后，给这些参考点一个基于类型的数值，如终端分叉，探测的数据。然后这些参考被存储起来当作样板，通常占据 400 ~ 1024 字节数据。

2. 手形

20 世纪 60 年代后期至 20 世纪 70 年代早期，Robert P. Miller 发明了一种测量手特征的设备，这种设备还能够记录下一些独一无二的特征，以备于以后用来比较，他因此获的了美国专利。

系统用一个固定焦长的摄像机系统、镜子来取得侧面积和上部景象，一个定位别针来确保用户的手放于适当位置的工作台。

拍到的图像是放于 32000 像素区域的手的影子，然后这个图像要经过 96 项测量，变为 9 字节模板。

3. 虹图

所有虹识别系统都使用一种单算法来产生虹码，这种虹码由剑桥大学的 John Daugman 发明。不像许多其他生物统计学识别符，当瞳孔伸张，虹仅有一些固定的变化。Daugman 取到一个由红外摄像机拍到的虹的图像，然后用综合积分方程对它进行处理。这些方程自动补偿了当瞳孔膨胀和收缩所产生的变形。通过将四个条件积分方程运用到系统中，一个 256 字节虹码便产生了。

虹识别享有当前技术最安全的地位。由于缺乏决策标准而产生错误接受率或错误拒绝率都是 120 万分之一。