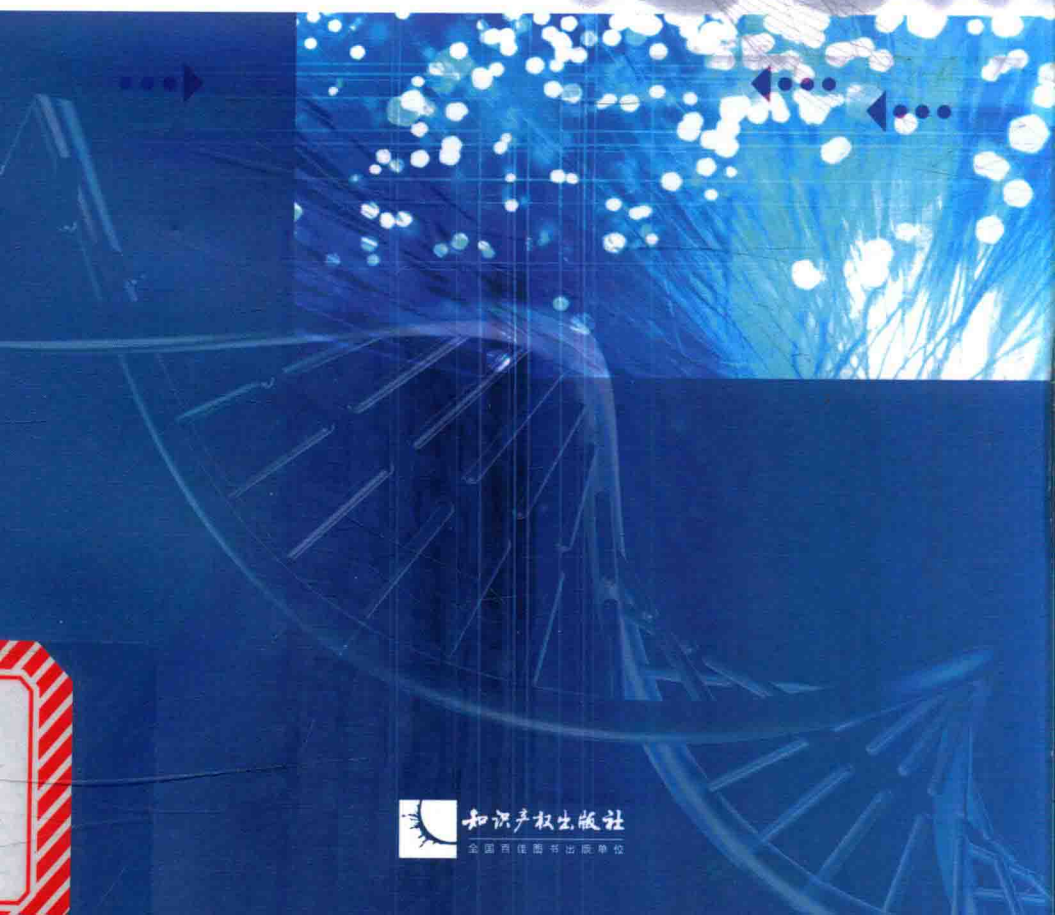
计算机基因学

——基于家族基因的网格信任模型

王铁方 著

COMPUTER GENICS

Family-gene Based Grid Trust Model



知识产权出版社
全国百佳图书出版单位

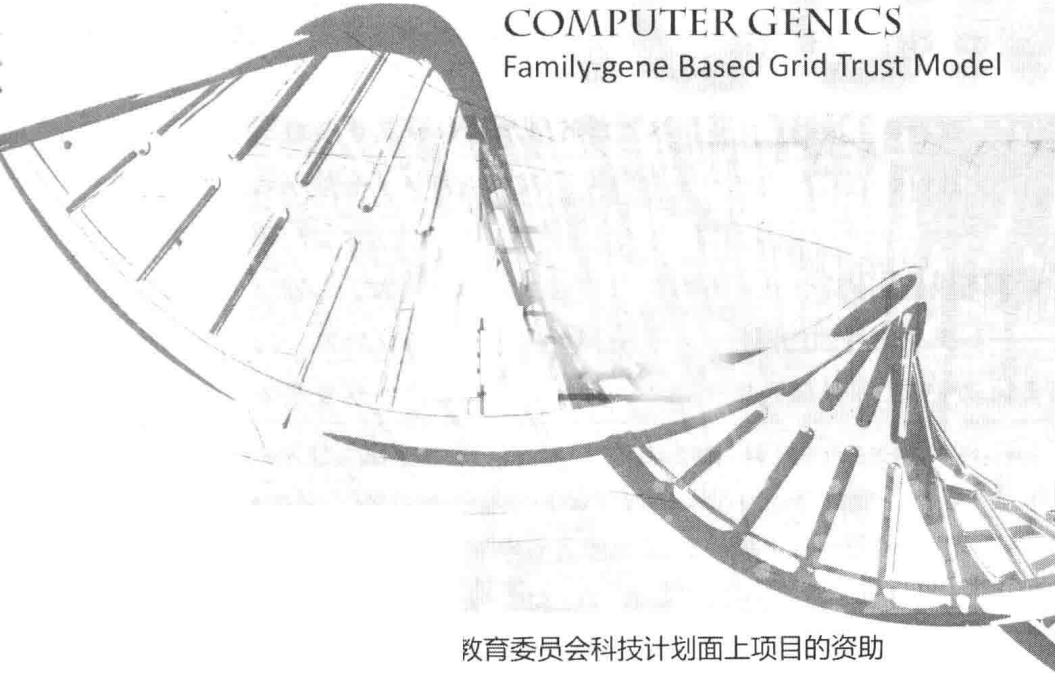
计算机基因学

——基于家族基因的网格信任模型

王铁方 著

COMPUTER GENICS

Family-gene Based Grid Trust Model



教育委员会科技计划面上项目的资助

知识产权出版社
全国百佳图书出版单位

图书在版编目（CIP）数据

计算机基因学——基于家族基因的网格信任模型 / 王铁方著. —北京：知识产权出版社，2016.3

ISBN 978-7-5130-3934-5

Ⅰ.①计… Ⅱ.①王… Ⅲ.①网格—安全技术—研究 Ⅳ.①TP393.08

中国版本图书馆 CIP 数据核字（2015）第 283891 号

责任编辑： 齐梓伊　　**责任出版：** 卢运霞

封面设计： 张　悦

计算机基因学

——基于家族基因的网格信任模型

王铁方　著

出版发行：知识产权出版社有限责任公司　　网　　址：http://www.ipph.cn

社　　址：北京市海淀区马甸南村 1 号（邮编：100088）　　天猫旗舰店：http://zscqcbs.tmall.com

责编电话：010-82000860 转 8176　　责 编 邮 箱：qiziyi2004@qq.com

发行电话：010-82000860 转 8101/8102　　发 行 传 真：010-82000893/82005070/82000270

印　　刷：北京中献拓方科技发展有限公司　　经　　销：各大网上书店、新华书店及相关专业书店

开　　本：880mm×1230mm　1/32　　印　　张：7

版　　次：2016 年 3 月第 1 版　　印　　次：2016 年 3 月第 1 次印刷

字　　数：165 千字　　定　　价：28.00 元

ISBN 978-7-5130-3934-5

序　言

随着时代的进步，互联网的发展，网络新技术的涌现，通过网络共享各种软硬件资源，提供统一的开放的计算机信息服务已经成为一种趋势。但是由最初的终端计算机网络进化的 Internet 网络因为不能够高效地整合各种网络资源，导致其在向外提供服务的时候总是有些不顺畅。

网格可以解决上述的问题，网格把整个因特网整合成一台巨大的超级计算机，实现计算资源、存储资源、数据资源、信息资源、知识资源、专家资源的全面共享和协同工作，向用户提供方便快捷的网络服务。

在 Internet 上，网格技术在向人们描述美好前景的同时，也带来了巨大的挑战。与普通的网络计算环境不同，网格计算具有许多特殊性：网格环境复杂，各种资源都动态地连接到 Internet 上，而且不同网格节点之间的通信也是通过 Internet 来完成。同时，网格计算环境中的所有主体都可以动态地加入或撤离网格中的虚拟组织等。网格的这些特点使网络出现了新的问题，如假冒合法用户和服务器；虚假、恶意的节点提供虚假服务；存在自私节点，他们只是消耗资源，而不提供资源。在这样的环境下向用户提供可靠、安全的应用执行环境和信息共享服务，面临着更加严峻的安全技术

挑战。

信任模型主要用于提供建立和管理网络信任关系的整体框架，研究实现用户与信任机构之间、信任机构与信任机构之间相互信任的基本准则和方法，并解决信任凭证的产生、信任关系的建立和传递等问题。

信任模型能够动态地描述相互信任的基本准则和方法，从而使网格系统变得更加健壮、安全，是解决网格安全问题的一个重要手段，是实现可信网络和信任计算的重要基石，对于实现高可信的网络服务，支持网络服务所要即所得具有重要的理论价值和现实意义。对于构建高可信的网格计算系统，对等计算网络（Peer to Peer，以下简称 P2P 网络）等新型的网络系统具有较好的参考价值和应用前景。

然而目前存在的几种网格环境下信任模型如网状的、层状的、对等的和单 Certificate Authority（以下简称 CA）等模型，都是基于公共密钥基础设施（Public Key Infrastructure，以下简称 PKI）的理论，虽然这些模型初期在电子商务、电子政务等领域得到了一定的应用，但随着网格技术的发展，信息化建设的不断深入和扩大，模型缺陷日渐暴露，进一步的应用会受到限制，其主要原因在于以下几点：

（1）在这类模型中，中心节点的合法性通过 CA 颁发的证书加以保证，但是证书主体信息不明确，难以区别现实中的同名实体；

（2）PKI 采用的是 CA 证书，CA 证书是一种识别证书，只能进行身份认证，不能进行授权，也不能进行访问控制，而且还需要在线运行；

（3）每次认证过程需要第三方中介机构的介入，使得这类信任模型不仅认证过程复杂，而且可扩展性差、计算量大、效率低、成

本高昂。

针对上述网格信任模型之不足，受人类自然信任、基因遗传、人体免疫系统能识别并排斥外源性抗原异物等原理启发，在借鉴PKI及特权管理基础设施（Privilege Management Infrastructure，以下简称PMI）机理的基础上，吸取传统信任模型之精华，以网格信任体系中的身份认证、授权管理与访问控制等关键技术为主要研究对象，以提高网格信任机制的安全性、可靠性、灵活性和效率为目的，本书提出了一种基于家族基因的网格信任模型（a Family - gene Based model for Grid Trust，以下简称FBGT），并描述了它的应用。这种模型按照家族基因确定家族成员的血源关系，并把这种血缘关系的远近作为家族中某个家族成员（族长）对各种事务（身份鉴别、访问控制）进行管理时和与其他家族成员交往（访问资源）时的一种依据。

本书主要创新如下：

（1）提出了一种可自证实的集身份认证和访问控制于一体的基因证书机制。在这种证书里，克服了原有的“扁平的”识别证书不能执行相应的任务、不能进行授权、不能进行访问控制，只能进行身份认证的缺陷，实现了将身份认证、访问控制和权限管理集为一体，这种证书既可以携带数据也可以携带对数据的操作权限和操作方法，即证书＝数据＋权限＋方法，是一种“立体的”证书。

（2）提出了基于家族基因的网格信任方法，它不同于传统PKI方法，具有以下特点：模仿人类血亲生物的特点，用人类家族中各成员血亲关系的远近作为对各个不同用户的信任计算程度。这种的方法能够反映当前主体信任程度，解决了传统信任计算模型存在的问题。

（3）建立了一种基于家族基因的网格信任模型，给出了网格环

境下家族基因、变异基因、族长、孩子、兄弟信任、遗传信任、基因鉴别和基因指派等形式化描述。根据信任的定义，以家族基因为原理的信任模型满足网格环境对信任的需求，符合信任的不对称、传递等属性，解决了信任凭证（基因证书）的产生、信任关系的建立和传递等问题，为建立基于家族基因的网格信任新体系提供了理论基础。

本书虽然是一本有关网格信任和计算机基因学的专著，但是语言通俗易懂，深入浅出，公式也不是很多，还有一些机器学习、数理基础、密码技术等入门级的知识，所以既适合机器学习和信息安全的专业人士研读，又适合想了解这方面知识的大学生和研究生看。

由于时间仓促，难免有一些错误，如有发现，请指出，笔者将会非常感谢，并将在下一版本中进行改正。对于本书这种机器算法感兴趣的读者，也可来信与我交流。笔者邮箱 wangtiefang@126.com。

目　录

第一章　信任相关技术概述

人无信而不立，信任是构建现实和谐社会的基石，也是构建网络和谐的中坚力量。网络世界如同人类社会，用户、客户机、服务器、网络等实体相互作用，关系错综复杂，这些实体间的信任关系构建是推动网络技术迅速发展的源动力。本章主要围绕计算机网络中信任的概念、信任技术、信任模型等方面的内容展开介绍。

一、信任

目前信任的理论研究涉及心理学、生物基因学、计算机科学、信息安全、人类学、社会学、政治学、经济学及历史学等多个领域。在社会科学中，信任被认为是一种依赖关系。值得信任的个人或团体意味着他们遵循实践政策、道德守则、法律和其先前的承诺。

（一）信任溯源

信任一词由来已久，从儒家的经典文献中，可以发现许多关于信任的论述。在儒家伦理文化强调的“三纲五常”中，“信任”被作为“五常之一”而加以强调。据学者统计，在儒家经典《论

语》中，“信”字出现了38次[①]。按照学者的分析，“信”在孔子看来，不仅是一种政治理念，同时也是处理社会关系的一个基本原则。

关于信任一词的解释，《史记·蒙恬列传》中描述，始皇甚尊宠蒙氏，信任贤之。《南史·荀伯玉传》也有，高帝重伯玉尽心，愈见信任，使掌军国密事。可见，信任就是相信并加以任用。

信任一词在心理学、生物基因学、计算机科学、信息安全、人类学、社会学、政治学、经济学及历史学等众多学科中被应用。各个学科的专家也曾试图从其研究的领域对信任予以界定。从各个学科的文献中可以看出信任与血缘关系、亲密程度、可预测、可靠性、能力、义务、爱、责任、等概念密切相关。

按照《现代汉语词典》的释义，信任是“相信而敢于托付”。在社会科学中，信任被认为是一种依赖关系。尼克拉斯·卢曼（Niklas Luhmann）认为，“信任是一个社会复杂性的简化机制”。罗素·哈丁（Russell Hardin）认为，“就某一事情而言，说我信任你，意味着关于该事情我有理由期望你为了我的利益行事，你的利益暗含我的利益”。伯纳德·巴伯（Bernard Barber）则认为“信任从来不是完全充分的”。国内学者白春阳认为：“信任是一种多层次、多维度的社会心理现象。信任作为一种交往态度，产生于交往过程中并作用于交往行为中，如果不相信交往对方，或是不相信交往行为能够产生合意的结果，那就不会有交往活动的发生。”同时“信任作为一种心理和态度，具有某种感染或扩散的特征。一次的受骗受害产生的不信任，可以抵消先前建立起的信任，还能够扩大

① 马得勇：“信任、信任的起源与信任的变迁”，载《开放时代》2008年第4期，第63－80页。

为较大范围的不信任。”

信任在英文中为 trust，这个词于 1200 年左右被引入英语，直接源自古北欧语的“traust”，意为帮助，依赖。按照韦氏词典的解释：“Firm reliance on the integrity, a ability, or character of a person or thing（对某人或者某事的正直、能力或者性格的坚定依靠）。”按照牛津词典的解释：“The firm belief in the reliability or truth or strength of an entity（对于一个主体的可靠性或者真实性或者实力的坚定信心）。”

综上所述，古文溯源、词典释意和专家的看法给出了信任的内涵，这些观点可以总结为：对实体的某方面行为的可靠性、依赖性、安全性等能力的坚定依靠。但是在计算机的学科中，这一描述还是显得不够确切。因为如果按照上述的定义对信任进行表述：当实体 S 假定实体 T 严格按照实体 S 所期望的那样行动，则实体 S 信任实体 T。从这个定义看出信任涉及假设、期望和行动，这意味着信任很难定量的测量，信任与风险是并存的。

Marsh 于 1994 年在其博士论文中针对多代理中的信任与协作问题，系统地阐述了信任的形式化问题，为信任在计算机领域尤其是互联网领域的应用奠定了基础。

然而到目前为止，对于计算机领域内的信任，学术界并没有一个准确而统一的定义，不同的文献对于信任按照各自解决问题的角度不同而有着不同的解释。据 Welty 等人统计，截至 2001 年，对信任的各种不同定义就达到 65 种之多。Grandison 和 Sloman 对各种形式的信任进行了综合分析，给出了信任的定义为“针对某个实体能够在某种给定的环境中可靠、可依赖性、安全采取行动的一种能力，一种坚定的信念”。他们认为信任是由多种属性构成，包括可靠性、可依赖性、真实性、实力性、安全性、诚实性等，需要根据

信任所处的具体环境进行相应的考虑和定义。

而 Dimitrakos 从另一个方面对信任给出释义，针对 T 方能够在给定的时间和环境下与 X 相关活动中可信赖性的表现，S 方对于 T 方关于服务 X 的信任是指 S 方对于 T 方的一种可预测信念。在这里，S、T 指的是一个实体（计算机）、一个人或者进程，服务指的是推荐、发行证书、各种事务等。

综上所述，不同学科领域，对于信任的理解也不尽相同，但是基本共识有以下几点：

信任表明了一个实体的诚实、真实、能力以及可信赖程度；

信任会随着时间延续而导致对实体认知程度的下降而衰减；

信任是建立在对实体历史行为认知的基础之上；

信任是实体间相互作用的依据。

显然，信任程度会随着实体间的多次接触而动态变化。信任的获得除了通过直接的相互认知以外，还可以通过间接途径（第三方的推荐）来获取。这里的第三方可以是社会认可的权威部门，也可以是双方实体认可的机构。

根据上面的分析，本书对信任作出如下定义：信任表明对实体身份的确认和其行为的期望，既是对实体的历史行为的直接认知，又是其他实体对该实体的推荐。信任可以随着实体行为而动态变化且随着时间延续而增强或者衰减。

（二）可信计算

可信计算（Trusted Computing）是一种信息系统安全新技术。可信计算的思想源于人类社会，是把人类社会成功的管理经验用于计算机信息系统和网络空间，以确保计算机信息系统和网络空间的安全可信。可信计算的总体目标是提高计算机系统的安全性，现阶

段的主要目标是确保：系统数据的完整性、数据的安全存储和平台可信性的远程证明。

可信计算产品主要用于：安全风险控制，使发生安全事件时的损失降至最小；安全检测与应急响应，及时发现攻击并采取相应措施；电子商务，减少电子交易的风险；数字版权管理，停止数字媒体的非法复制与传播，等等。

可信计算的基本思想是：首先，在计算机系统中建立一个信任根，信任根的可信性由物理安全、技术安全、管理安全共同确保；其次，建立一条信任链，从信任根开始到硬件平台，到操作系统，再到应用，一级测量认证一级，一级信任一级，把这种信任扩展到整个计算机系统，从而确保整个计算机系统的可信。

可信计算主要关注以下五个方面：第一，签注密钥（Endorsement key），签注密钥是一个 2048 位的 RSA 公共和私有密钥对，它在芯片出厂时随机生成并且不能改变。这个私有密钥永远在芯片里，而公共密钥用来认证及加密发送到该芯片的敏感数据；第二，安全输入输出（Secure input and output），安全输入输出是指电脑用户和他们认为与之交互的软件间受保护的路径。当前，电脑系统上恶意软件有许多方式来拦截用户和软件进程间传送的数据，如键盘监听和截屏；第三，储存器屏蔽（Memory curtaining），储存器屏蔽拓展了一般的储存保护技术，提供了完全独立的储存区域。例如，包含密钥的位置。因为操作系统自身也没有被屏蔽储存的完全访问权限，所以入侵者即便控制了操作系统信息也是安全的；第四，密封储存（Sealed storage），密封存储通过把私有信息和使用的软硬件平台配置信息捆绑在一起来保护私有信息。意味着该数据只能在相同的软硬件组合环境下读取。例如，某个用户在他们的电脑上保存一首歌曲，而他们的电脑没有播放这首歌的许可证，他们就不能播放这首