计算机学术研究进展丛书

对等网和云计算中的行为信任

金 瑜 著

科学出版社

对等网和云计算中的行为信任

金 瑜 著

科 学 出 版 社

北 京

内 容 简 介

本书以著者在开放网络环境下的行为信任的研究成果为主线，详细介绍对等网和云计算环境下的行为信任的关键技术。第一部分，详细介绍对等网环境下行为信任管理的相关技术，包括各种算法设计、翔实的模拟实验配置、原型系统的设计与实现等。第二部分介绍云计算环境下的行为信任方面的研究成果，包括利用异常检测和多云技术，进行 SLA 信任管理，帮助云客户在服务运行时动态选择云服务提供者；提出能够满足云客户的 QoS 需求的信任数据存储和查找机制，帮助用户在服务启动前选择合适的云服务提供者；构建基于行为及记录的动态信任模型，解决云服务的动态变化问题等。

本书适合从事对等计算和云计算、信任管理系统、网络安全等相关研究和开发工作的人员阅读，也可以作为高校和研究所相关研究人员，以及计算机及其相关专业研究生和高年级本科生的参考教材。

图书在版编目(CIP)数据

对等网和云计算中的行为信任/金瑜著. —北京：科学出版社，2016.6

（计算机学术研究进展丛书）

ISBN 978-7-03-048409-3

Ⅰ.①对… Ⅱ.①金… Ⅲ.①计算机网络 Ⅳ.①TP393

中国版本图书馆 CIP 数据核字(2016)第 113605 号

责任编辑：闫 陶 杜 权 / 责任校对：董艳辉

责任印制：彭 超 / 封面设计：苏 波

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

武汉中科兴业印务有限公司印刷

科学出版社发行 各地新华书店经销

*

开本：B5(720×1 000)

2016 年 5 月第 一 版 印张：10 1/4

2016 年 5 月第一次印刷 字数：210 000

定价：50.00 元

（如有印装质量问题，我社负责调换）

前　言

近年来，以对等网和云计算为代表的开放网络计算环境成为学术界、产业界以及政府等各界关注的焦点。对等网和云计算通过 Internet 让用户共享各种资源，如信息、软件、平台、计算周期和存储等，具有极大的应用价值，对中小型企业和大型 IT 企业，甚至普通个人用户都有极大的经济价值，对促进国民经济发展具有重要作用。

然而，对等网和云计算是一把双刃剑，为企业和个人带来机遇的同时，也带来了新的挑战。要使对等网和云计算能够正常提供服务，必须满足一个前提：所有服务和请求都是真实、可信的。然而，由于对等网和云计算环境天生具有自治、开放和动态等特性，实体间存在信任缺失问题。例如，恶意服务提供者可以发布虚假服务广告，声称其服务是高质量的，却提供伪劣的服务，甚至是恶意服务，从而造成消费者难以选择到真正高质量的服务。另外，恶意服务消费者也可以对服务提供者发动恶意攻击，滥用服务等，造成诚实、可信的服务提供者不能为其他可信实体提供正常服务。为了促进对等网和云计算等应用，防止上述恶意行为，必须在服务提供者和服务消费者之间建立信任关系。

目前，开放网络计算环境下的信任研究大多数关注身份信任，主要采用认证和授权技术，即所谓的“硬安全”。然而，身份信任是静态的，是一种基本信任关系，不能充分反映开放网络计算环境下实体高度动态和自治特性。因此，要在对等网和云计算环境下建立实体间的信任关系，仅有身份信任是不够的，必须研究动态的行为信任模型，与身份信任相结合。也就是采用“软安全”方法，分析实体历史行为，预测其将来行为，动态反映实体行为变化。

著者对该领域的国内外同类书籍进行分析后发现，它们的研究内容一般都很宽泛，没有对某个具体研究方向进行深入研究；另外，它们基本都是从事信任方面的理论研究，没有涉及具体的分布式应用。针对这些问题，本

书从以下两方面来切入：第一，仅对开放环境下的行为信任进行深入研究；第二，在研究行为信任时，考虑的是目前广泛应用和最为热门的对等网和云计算环境。撰写本书的目的是为读者提供一部深入研究开放网络环境下行为信任的书籍，希望对他们今后的行为信任学习和研究提供一定的帮助。同时，本书也是著者多年科研工作的总结。

本书共 8 章，主要内容如下。

第 1 章从理论和技术层面，介绍和分析对等网中行为信任的相关知识，包括信任定义、信任分类、声誉系统结构以及 P2P 信任模型关键问题等。为后续章节提供理论和技术上的支撑。

第 2 章针对目前 P2P 单层信任模型信任收敛慢、不能反映信任时间相关特性和可扩展性差等缺点，提出一个时间相关的、基于信任簇的两层信任模型 TLT。给出一个基本的两层信任管理框架，介绍簇内服务信任、簇内代理信任以及簇间信任管理等功能，没有考虑虚假回馈、行为信任数据安全性问题，将这两个问题分别放在第 3 章和第 4 章描述。

第 3 章针对目前绝大多数 P2P 行为信任模型不能处理策略回馈和不递交回馈的问题，提出一个新的基于双 rating 的回馈处理机制。描述回馈处理机制细节，接下来给出回馈处理机制在 TLT 中的实现算法，然后进行详细实验分析。

第 4 章针对现有 P2P 系统中的信任数据管理协议存在信任轮询效率低、信任数据不安全等问题，提出一个新的基于簇投票的信任数据管理协议 SP2PRep。给出协议工作步骤和设计，并对协议安全性和性能进行分析。描述 SP2PRep 中信任数据安全性管理方法在 TLT 中的应用。

第 5 章结合第 2～4 章的内容，设计和实现了基于信任簇的两层信任模型原型系统 TLTP，并从安全性和性能两方面对 TLTP 进行评价，为今后 P2P 信任模型深入研究和应用奠定实验基础。

第 6 章介绍了云计算环境下的行为信任相关知识。云环境下的行为信任机制大致可分为如下两类：基于声誉的信任机制、基于 SLA 的信任机制，并对这两种信任进行详细分析。

第 7 章针对目前单云的云计算环境存在的缺点（在服务运行时均只选择一个服务提供者提供服务，即使发现该服务违背了 SLA 约定，也不能无缝切换到另一云服务，因此存在服务不可用空隙，同样会给 CSC 带来损失），提出 Stadam——一种新的适用于云环境的基于异常检测和多云的 SLA 信任模型。给出异常检测算法、SLA 管理等，并进行了详细实验分析。

第 8 章针对大规模云计算环境下云用户很难找到满足自己个性化意愿

的云服务，设计和实现一个 SLA 个性化云服务推荐系统。提出以“块”为单位的聚类方法，能够基于滑动窗口对持续演化的数据流进行聚类，并且可以应用在多记录环境中。该方法应用在基于 SLA 聚类的个性化云服务推荐系统中，可以帮助云用户在纷繁复杂的云市场找到满意的云服务。

本书的出版得到了国家自然科学基金（基金编号：61303117）的经费资助，在此表示感谢。在撰写和出版本书过程中得到科学出版社的支持和帮助，在此一并表示感谢。

由于著者水平有限，书中难免存在不足之处，敬请广大读者批评指正。

金　瑜

2015 年 12 月 24 日

目　录

第1章 对等网中行为信任技术基础

本章对对等网中行为信任相关理论和技术基础进行阐述。首先介绍信任的几种不同定义，并从中导出它们的工作定义；接着介绍信任分类；接下来描述信任分值在P2P系统中的用途；然后介绍目前行为信任系统两种不同的体系结构，分析它们各自的优缺点；最后着重介绍目前P2P行为信任模型如何解决关键问题。

1.1 信任定义

信任在人类社会中起着很重要的作用，可以用来解决由于开放性和自由性带来的不确定性和不可控性。在计算机辅助的合作和电子商务中，由于合作和交易双方在时间和空间上分离，同样存在不确定性和不可控性[1]，可计算信任模型能够缓解这个问题[2]。

信任有多种表现形式，在不同上下文有不同含义，因此给信任下定义是一项有挑战性的工作。到目前为止，信任还没有统一的定义。下面介绍几种有代表性的定义。

定义1.1 Gambetta[3]认为信任是一种主观可能性(probability)，通过它，个体(A)期待另一个体(B)从事一个给定的、A能从中获取利益的活动。

这个定义非常简单，仅包含信任方(A)对可信方(B)的依赖性和可信方的可靠性(reliability)。因此这种定义又称为可靠性信任。

定义1.2 McKnight和Chervany[4]给出了外延更广的定义，他们认为，信任是一方在即使有可能出现负面后果，但还是感觉相对安全的情形下，对某人和某物的依赖程度。

这个定义不仅包含信任方对可信方(B)的依赖性、可信方的可靠性，还指出信任方愿意承担一定风险(risk)的态度。因此这种定义又称为决策

信任。

定义 1.3 信任是一个 peer 基于自己的直接经验，对另一个 peer 能力、诚实和可靠性的信念。

这种定义虽然包含设计和实现信任系统细节，但不能反映信任的时间相关性和上下文依赖特性。

定义 1.4 信任是一个 peer 在给定上下文、给定时间点，基于自己的直接经验，对另一个 peer 能力、诚实和可靠性的信念。

1.2 信任分类

信任大致可以分为两类：身份信任(identity trust)和行为信任(behavior trust)[5]。身份信任又可以分为对象(object)身份信任和实体(entity)身份信任[6]。对象指静态和动态数据，静态对象是存储在本地的数据，动态对象是处于传输过程中的数据；实体则是指对象源端和目的端，前者是数据发送者，后者是数据接收者。行为信任则是指对象和实体行为信任。图 1.1 给出了这种分类方法。

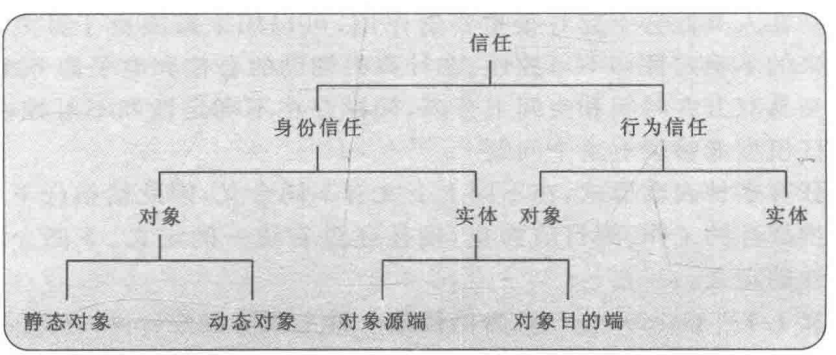

图 1.1 信任分类

建立身份信任的目的是验证实体和对象的真实性，可以采用加密、数据隐藏、数字化证书、认证和访问控制等技术，因此身份信任又称为“硬安全”。建立实体身份信任时，主要关注数据对象源端或目的端的真实性；前者关注数据发送者是否是所期望的源端，后者则关注数据目的端是否有访问本地资源的授权[7]。对象身份信任则关注对象的隐私性(privacy)和完整性(integrity)。隐私性指对象机密性，表现在两方面：①存储在实体本地的静态对象不能被未授

权用户访问；②传输的动态对象不能被数据接收者以外的其他实体访问。对象完整性也表现在两方面：①存储在实体本地的静态对象不能被蓄意或恶意更改；②目的端接收到的数据与源端发送的数据是相同的，在任何中间环节都没有被修改。可以使用数据加密、IPSec(IP security)[8] 和 SSL(secure sockets layer)[9] 等技术建立对象的身份信任。

然而，行为信任对实体和对象“可信度”关注的内容要比身份信任更为广泛。行为信任通过分析实体和对象历史行为，估计其将来行为，又称为“软安全”。通常从三方面观察一个实体的行为属性：可访问性，即实体最大上线时间；性能，即实体响应时间；诚实度。对象行为信任关注数据对象在运行机上的行为，包括是否遵守运行机系统规则、策略和隐私等。要注意的是，仅仅是包含动态或可执行数据的对象才需要建立行为信任，因为静态对象不会影响运行机上的规则和策略，也就没有必要建立行为信任。

表 1.1 给出了身份信任和行为信任的区别。身份信任是任何系统建立信任的基础，没有身份信任，就不能建立行为信任[10-12]；建立身份信任非常简单，信任层次要么是 0，要么是 1，例如，用户登录系统时，要提供自己的身份，系统对其建立身份信任。

表 1.1　身份信任和行为信任比较

信任属性	身份信任	行为信任
重要性	基础	上层
计算复杂性	简单	复杂
变化性	很少	频繁
建立方式	授予	赚取
被取代性	能	不能

通过识别和验证其身份实现，如果验证通过，则信任层次是 1，否则为 0。由此可以看出，身份信任建立过程只是验证其身份，其计算成本是固定的；而行为信任所需计算成本要大得多，需要随着时间而动态变化。身份信任很少变化，除非又拥有了一个新身份；而行为信任经常变化。身份信任是由系统管理员授予的，而行为信任不是依靠分配，必须通过自身服务赚取。身份信任可以很容易地被取代，如果实体丢失身份以后，它需向系统管理员申请一个新身份即可；而行为信任是不可被取代的，如果实体由于误行为，其“可信度”很快就降低了，很难再达到以前的高可信度。

1.2.1 身份信任模型

信任模型 PGP 和 X.509 及其应用 PolicyMaker 和 KeyNote 都是关注建立实体身份信任。PGP 和 X.509 是基于公钥加密技术，使用数字化证书进行认证的两种主要信任模型。PGP 可以用于邮件保密，防止非授权者阅读邮件，它还能对邮件加上数字签名，从而使收信人可以确认邮件发送者，并能确信邮件没有被篡改。PGP 假定证书权威之间没有集中式或层次关系，但 X.509 是一个具有严格层次的信任模型，它以可信第三方（证书权威）颁发的公钥证书为核心，以公钥密码体系和数字化签名为基础，为开放环境下的网络交互提供身份认证。

IPSec [8] 是一组给 IPv6 和 IPv4 数据包提供高质量、可互操作的、基于密码学的安全协议，它们确保在 IP 网络上进行安全通信。它使用了严格的密码认证协议和加密算法来保证 IP 通信完整性和保密性。即通过网络层加密和认证技术，IPSec 能建立端到端数据对象身份信任，确保数据对象传输时的私密性和完整性。

安全套接字层（secure sockets layer，SSL）信任模型 [9] 为 Web 浏览器提供应用加密技术，确保使用了 SSL 的应用发送数据对象具有机密性和完整性。SSL 协议是建立在 TCP/IP 之上，提供基于 C/S 模式的网络应用安全通信开放式协议。它能使用户与服务器之间的通信不被攻击者窃听，始终对服务器进行认证，还可选择用户认证。SSL 协议使用不对称加密技术实现会话双方之间信息的安全传递，可以保证信息传递的保密性、完整性，还可以鉴别会话双方的身份。SSL 协议要求建立在可靠传输层协议（TCP）之上。它的优点在于它与应用层协议无关。

以上这些身份信任模型都没有考虑实体和对象行为信任，是一种静态信任模型，因此也就不能根据它们历史行为改变其信任层次。

1.2.2 行为信任模型

行为信任是通过观察实体和对象历史行为，估计它们将来行为，其信任层次不是静态的，而是随着时间流逝动态改变的。行为信任模型又可以分为基于声誉的和基于激励的模型。

1. 基于声誉的行为信任模型

目前，基于声誉的行为信任模型在 P2P 信任模型中占绝大多数。当节点与另一节点交互前，它需要管理两方面信息：一是自身经历，又称为直接信任；二是来自其他节点的推荐，又称为间接信任（声誉）。该节点通过综合这两方面信息，选取某一节点作为服务提供者进行交易。根据服务提供者选取策略不同，基于声誉的模型又可以分为两种：最大值模型和域值模型。在最大值模型里，选取综合信任分值最高的节点进行交互。这种模型总是选择信任分值最高的节点提供服务，容易导致负载不均衡[13]，并且系统中新节点利用率不高，而在域值模型里，任何响应节点都可以作为服务提供者，只要其综合信任超过预定域值。这种模型就是负载均衡的，并且新节点利用率高。因此现有基于声誉的 P2P 信任模型绝大多数是域值模型。

由于基于声誉的模型是本文研究重点，故将其细节问题放在 1.5 节介绍。

注意：如果没有特殊说明，本书的对等网中的行为信任是指基于声誉的行为信任模型，又称为声誉模型。

2. 基于激励的行为信任模型

基于激励的模型主要是模拟现实社会中的经济体制，使用虚拟支付概念来激励买卖双方进行诚实交易或汇报真实回馈[14-19]。

文献[17]提出了一个安全虚拟支付系统 PeerPay。PeerPay 里存在一个可信的、集中的代理，为节点交易提供桥梁作用，如计算、发布和收取虚拟货币，还要负责探测节点恶意行为（如多次消费同一虚拟货币）。节点加入系统后，需从代理处购买其签名的虚拟货币，才能进行买卖活动；然后节点就可以使用该虚拟货币从其他节点处购买服务；服务完毕后，服务购买者在虚拟货币上使用私钥签名，并附上一个序列号，作为本次服务购买凭证；服务提供者可以在任何时候以购买凭证从代理处换取虚拟货币。

文献[19]使用了另一种形式的虚拟支付系统，不过支付对象不是服务，而是声誉报告。在文献[19]里有一批特殊代理从事声誉信息买卖。当一个节点（如 peerA）与另一节点（如 peerB）交易前，为了降低风险，peerA 需要了解 peerB 的服务声誉信息。因此，peerA 可以向某代理购买 peerB 的声誉信息，假设所花费的成本为 Y。当交易完成后，peerA 也可以向代理售卖 peerB 本次服务的报告，假设得到的费用为 X。如果 peerA 的报告与 peerB 的真实服务性能相符时，X 就大于或等于 Y，否则 X 小于 Y。因此节点汇报真实回馈时，其虚拟货币就会增加，反之会减少。

以上两个系统都是利用了一个可计费系统激励节点提供诚实服务和回馈，采用虚拟支付手段购买服务和声誉报告。尽管货币机制提供了一个清晰的经济模型，但在实际中并不可行，并且以上两个系统均使用了可信第三方，即代理，这与 P2P 系统设计初衷是相悖的。

1.3 信任分值用途

在 P2P 系统中，信任分值主要有两种用途[20,21]：①服务提供者选择，如果有多个服务响应者可以提供相同服务，在最大值模型里按照它们的综合信任值排序，选择综合信任值最高的 peer 来提供服务，如果交易失败，则选择综合信任值第二的 peer，如此下去，直到得到满意内容和服务；在域值模型里，从综合信任值高于给定域值的所有服务响应者中随机选择一个提供服务，如果交易失败，再依据相同规则选取另一个节点提供服务，直到交易成功，服务提供者选择可以隔离和孤立恶意 peer，避免由于交易失败造成的资源和网络带宽浪费；②调停需求冲突，当有多个 peer 竞争同一资源时，资源所有者可以使用信任调解资源访问冲突，高优先级赋予信任状况好的资源请求者，用信任实现了访问控制，具有激励作用。目前，基于声誉的 P2P 信任模型的研究主要集中在第一种情况，选择好的服务提供者，保证资源请求者接收到高质量内容和服务，提高成功交易率。

1.4 声誉系统分类

声誉系统体系结构决定了 rating、信任与声誉分值在节点间的传递方式。目前主要有两种体系结构：集中式结构和分布式结构。

1.4.1 集中式结构

目前，大多数拍卖网站的声誉系统都是采用集中式结构，如 eBay[22]、Yahoo![23] 和 Amazon[24] 等。集中式体系结构的声誉系统工作原理如图 1.2 所示。系统中一对节点（如 peer3 与 peer1）发生交易以后，peer3 根据本次交易结果以 rating 的形式对交易伙伴 peer1 的服务可靠性进行评价，并将此 rating 汇报给一个集中式声誉服务器。声誉服务器收集到该 rating 以后，结合以前其他

节点(如 peer7 和 peer5)给 peer1 的 rating,计算 peer1 服务可靠性的总体声誉分值,然后使该声誉分值对系统中所有成员公开。下一次,系统中其他节点(如 peer2)就可以使用这一信息,选择服务声誉较好的节点(如 peer1)提供服务,从而提高了服务成功率。

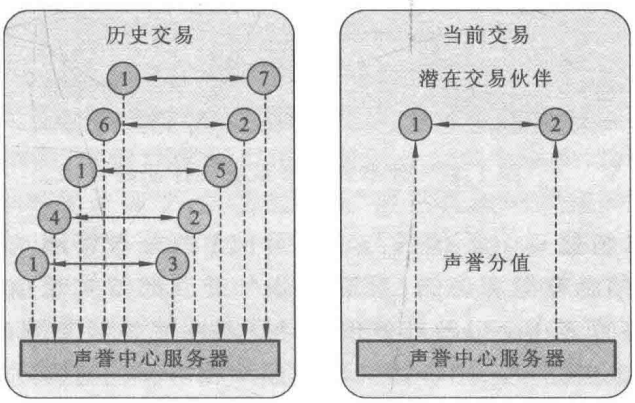

图 1.2　集中式声誉系统工作原理

集中式声誉系统包含两个重要组件:集中式通信协议和声誉计算引擎。集中式通信协议负责实现节点和声誉服务器之间的通信,包括节点向声誉服务器汇报交易伙伴的 rating 以及从声誉服务器处获得其他节点声誉信息;声誉计算引擎仅在声誉中心服务器上实现,根据收集的 ratings 以及其他信息计算系统中每个参与者的声誉分值。

1.4.2　分布式结构

在分布式声誉系统里,没有集中场所存储和访问信任与声誉值。rating 的提供、收集和存储以及信任与声誉分值计算都要依赖 peer 之间相互合作。分布式声誉系统工作原理如图 1.3 所示。当交易完成以后,节点将对交易伙伴的 rating 存储在本地,或者依据一定规则(如 DHT),存储在系统中其他节点上(不是声誉中心)。如果目前有一节点(如 peer2)想要与另一节点(peer1)进行交易,为了保证本次交易成功,peer2 需要向系统中发出声誉轮询,请求以前与 peer1 交易过的节点(peer3、peer7 和 peer5)汇报 rating。peer2 根据收集的 rating,计算 peer1 综合信任分值,然后决定是否与其进行交易。

分布式声誉系统也包含两个重要组件:分布式信任与声誉数据管理协议、信任与声誉计算引擎。分布式信任与声誉数据管理协议负责实现各种信