

# MAS 中信任和信誉系统的研究进展

贺利坚<sup>1,2</sup> 黄厚宽<sup>1</sup>

(北京交通大学计算机与信息技术学院 北京 100044)<sup>1</sup> (烟台大学计算机学院 烟台 264005)<sup>2</sup>

**摘要** 信任在人类社会的合作中起着非常重要的作用,在诸多领域也受到了广泛的关注。在开放多 Agent 系统(MAS)的研究中,引入了信任的方法,用于解决交互伙伴的选择问题。信誉与信任密切相关,可以视信誉为信任的信息来源之一,信誉系统是用于完成信任评价的机制。MAS 中信任研究应担负起发现计算实体之间信任的一般规律的重任。讨论了信任和信誉模型研究的内容、要求以及应用。在技术层面,信任表示有认知和数值两种观点,形成了集中式、分布式和混合式的体系结构,用于信任的汇总包括统计、概率、信念理论及模糊推理等方法。群体信誉、信息不准确、信息贫乏、异构模型互操作等问题有待进一步深入研究。

**关键词** 信任,信誉系统,信任表示,体系结构,信誉汇总

**中图法分类号** TP181 **文献标识码** A

## Trust and Reputation Mechanism in Multi-Agent System

HE Li-jian<sup>1,2</sup> HUANG Hou-kuan<sup>1</sup>

(School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China)<sup>1</sup>

(School of Computer, Yantai University, Yantai 264005, China)<sup>2</sup>

**Abstract** Trust plays an important role in the interaction of human society, and is focused on in many related research domains. Trust is also introduced to open multi-agent system to help agent choose interaction partner. Reputation can act as one of sources of trust information, which is closely related to trust. Reputation system is a mechanism to evaluate trust in MAS. Research on trust in MAS should take on responsibility of finding the general law of trust among computational entities. The content, requirements and applications of trust and reputation system were discussed in the paper. In terms of technology, there are cognitive and numerical opinions in trust representation, and centralized, distributed and mixed frameworks were adopted by the reputation system. The statistics, probability, belief theory and fuzzy reasoning methods were used to aggregate reputation information. At last, some problems were put forward to study further in the future, which cover group reputation, inaccurate report, lacking of information and interoperation between heterogeneous reputation systems.

**Keywords** Trust, Reputation system, Trust representation, Framework, Reputation aggregation

## 1 信任和信誉系统

许多计算机应用是开放的分布式系统,由分散在网络中的多个部件构成,没有集中的控制中心,环境不断地发生着变化。这些系统可以建模为开放的多 Agent 系统(Multi-Agent System, MAS)。在 MAS 中,能力受限、分布、自治的 Agent 需要通过合作,相互提供资源、服务或者信息,共同完成任务。然而,在这样的环境中,Agent 的合作面临着诸多困难。如何针对任务选择合适的合作伙伴,确定与合作伙伴交互的策略,成为一个非常基础性的问题<sup>[1]</sup>。

### 1.1 信任

人类社会是一个复杂的开放式系统,信任(Trust)是一种简化复杂系统中的决策的有效方法<sup>[2]</sup>,在人们的合作中起着

非常重要的作用。作为一种行之有效的社会机制,信任在心理学、经济学、管理学等领域也受到了广泛的关注<sup>[3-5]</sup>。在 MAS 中,对信任的研究和应用试图将以上领域的研究成果与计算科学结合起来,把在人类社会中运行良好的信任机制运用到计算科学领域中,最大化合作求解的收益,甚至由于信任机制的作用,培育出良好的合作环境<sup>[6]</sup>。Povey 认为信任是主体基于交互经验、信誉、角色,以及能力、意图等内部属性而形成的对客体特定行为的主观可能性的预期<sup>[7]</sup>。Gambetta 将信任定义为 Agent 要执行某一特定动作的主观概率的级别,并强调信任与动作的上下文环境密切相关<sup>[8]</sup>。在开放的 MAS 中的合作背景下:信任是在长期交互中形成的连续的度量;是一种主观性的度量,不同主体对同一客体有不同的信任值;是事前的估计,而不是准确的度量,主要是通过交互历史

到稿日期:2010-05-02 返修日期:2010-08-09 本文受国家“九七三”重点基础研究发展规划基金项目(2007CB307100,2007CB307106),北京市科技计划项目(Z09050600910902)资助。

贺利坚(1971—),博士生,副教授,CCF 会员,主要研究方向为多 Agent 系统, E-mail: sxhelijian@163.com; 黄厚宽(1940—),教授,博士生导师,CCF 会员,主要研究方向为人工智能、模式识别、数据库、数据挖掘以及多 Agent 系统。

中获得的经验得出的对未来交互结果的估计;是和交互目标相关联的。在具有信任机制的 MAS 系统中, Agent 在交互前需要通过信任评价,选择能满足其交互目的的合作者,决定何时、同谁以及如何进行交互<sup>[2]</sup>。

## 1.2 信誉和信誉系统

在 MAS 中,信誉(Reputation,也译作声誉)与信任密切相关。Agent 的信誉是基于其他 Agent 对该 Agent 过去行为的观察或收集的信息,做出的对其未来行为的期望<sup>[9]</sup>。信誉强调来自于公众,是第三方对信任客体的信任。在集中式的系统中,认证中心提供权威的、统一的公众观点,用来反映信任客体在公众中的映像;而在分布式的系统(如人类社会)中,基于参与者之间的交互,也会形成对客体局部的、多元的共同认识。

MAS 中的信任通常来源于两种途径<sup>[9]</sup>:一种称为直接信任,是评价 Agent 依靠自己的知识和以往交互中获得的直接经验判断得出的对目标 Agent 的信任,反映评价 Agent 的主观认识;另一种就是信誉,是通过收集评价 Agent 所处的社会中其他 Agent 对目标 Agent 的观点,经推理得出的信任。这里,信誉成为信任的一个组成部分;与此同时,信誉是基于别的 Agent 的直接信任产生的,因此信誉来自于信任。有的研究也在淡化信誉和信任之间的差别,例如,在著名的公共测试平台 ART 中<sup>[10]</sup>,不区分对评价目标的信任和信誉,而将 Agent 自己的观点作为证人观点的一部分,平等地对待所有的评定信息,这种处理尤其适用于 Agent 自身能力不足时的情形。

信誉系统(Reputation System)也称为信誉机制(Reputation mechanism),是用于支持信任评价的机制<sup>[11]</sup>。信誉系统通过在 Agent 之间收集、维护和传送评定报告,使评价 Agent 获得其他 Agent 的交互经验,从而增强自身的能力。信誉系统除了用于提高交互的收益之外,也常作为一种激励机制来有效地降低参与者诚实交易的成本,激励他们正面的合作行为。

信任系统和信誉系统的区别在于<sup>[12]</sup>:信任系统产生对信任客体的打分,反映的是对信任客体的主观性观点,而信誉系统产生的是整个群体对信任客体的信誉打分;信任传递路径和网络在信任系统中是显式的,而在信誉系统中经常不考虑传递性,或仅采用隐性的方式。这两种系统相互之间存在着交叉,往往也并不作特别的区分。

为了有效地帮助 Agent 选择未来的交互对象,大多数信任/信誉系统允许 Agent:(1)获得直接经验后可以评估对交互目标的信任;(2)表达对这一评估的确定程度;(3)以信誉报告的形式交换评估结果;(4)筛选和融合直接经验与信誉报告,以获得更准确、更有用的评估<sup>[13]</sup>。

## 1.3 MAS 中信任研究的意义

多 Agent 系统是采取社会学途径研究人工智能的产物。按照 M. Minsky 的思想, MAS 是一种设计大规模、分布且复杂系统的方法论,是一个适合描述大规模开放分布系统的软件方案。在由相互作用的自主 Agent 组成的大规模人工复杂自适应系统中,信任和信誉机制是一种重要的社会系统机制<sup>[14]</sup>。这样, MAS 中 Agent 信任的研究应该充当起计算世界中信任机制构建的代表。这里的 Agent 是一切交互实体的抽象,可以是个体,也可以是群体。从这个意义上讲, Agent

信任的研究应当承担着发现计算实体之间信任的一般规律的重任,注重建立分布的 Agent 来自主进行信任评价并完成合作的一般模型,进一步深入理解 MAS 中信誉系统本身的性质和特点,为所有采用信任机制的计算问题提供理论支持和实践指导。同时,相关研究要为研究具有信任机制的多 Agent 应用系统开发的方法学提供支持。

## 2 信任和信誉系统的研究现状

### 2.1 信任问题的研究内容

Ramchurn 对 MAS 中的信任进行了分类<sup>[1]</sup>,将信任分为个体层(Individual-level)和系统层(System-level)两个层次,这种划分能够体现信任问题研究的内容。信任的个体层是 Agent 推理能力的一部分,用于完成对交互伙伴的信任的评价,研究内容包括:(1)演化和学习模型: Agent 在持续的交互过程中,基于自己有限的经验对信任进行度量、选择交互的合作者、学习合作者的策略,从而追求更多的长期回报;(2)信誉模型:建立收集其他 Agent 的经验的手段,并计算对目标 Agent 的信任,信誉模型的核心是收集信息与对证人撒谎的识别与处理;(3)社会认知模型:建立对手的思维状态模型,从动机、能力等方面确定是否相信对方会诚实地进行交互。信任的系统层通过设计交互协议和工作机制来保证和引导 Agent 在行动的层面上诚实地进行合作。研究内容包括:(1)可信赖的交互机制:通过设计交互协议来约束 Agent 的交互行为,避免 Agent 在交互中撒谎和投机;(2)信誉机制:通过对有不良行为的 Agent 进行惩罚,对有良好行为的 Agent 给予激励,从而鼓励合作;(3)分布式安全机制:使用身份验证、授权等安全机制以及对 Agent 的标准化、信息加密、隐私保护等手段来保证信誉系统的安全性。信誉系统的研究重点是在个体层为 Agent 提供更好的识别、收集信任信息及基于信任推理和决策的能力。目前 MAS 中信任和信誉系统的研究要重视按照信任的来源和信誉的语义,构建多维的信任评价体系。

Artz 将信任研究分为 4 个主要领域<sup>[15]</sup>:(1)基于策略的信任:使用策略去建立信任,焦点在于管理和交换证书,强迫执行存取策略;(2)基于信誉的信任:利用信誉建立信任,组合与某实体过去交互的性能报告,从而评价未来的行为;(3)信任的一般模型:对信任进行建模和定义,包括其前提、条件、部件和输出,建模研究的范围包括从简单的存取控制策略到对能力、信念、风险、重要性、收益等的分析;(4)信任信息的来源。

### 2.2 信誉系统的性能要求

在建立信任问题的通用实验平台中,Fullam 等人提出在竞争环境下的信誉系统要符合以下要求<sup>[16]</sup>:(1)准确性: Agent 必须能够准确预知其他 Agent 的行为;(2)适应性:如果其他 Agent 突然改变合作的策略, Agent 必须能及时适应动态变化;(3)快速收敛:当新的 Agent 进入系统时,应能快速地为其建立模型,能及时阻止有恶意的 Agent 改变身份后重新进入系统;(4)多维:能用多种来源和类型的信息评价信誉;(5)效率:能用最小的成本和时间完成模型的创建。

Huynh 针对开放 MAS 中 Agent 数量多、可以自由进出等特点,提出了 4 方面的要求<sup>[9]</sup>:(1)能够提供多种有效的信任获取手段,特别是新加入的 Agent 也能够自主地获得有关信息;(2)每个 Agent 能够独立完成信誉的评价;(3)具有通用

性,能适合于不同的应用领域;(4)具有鲁棒性,以应对交互中可能的撒谎和同一证据被多次使用的问题。

### 2.3 信誉系统的应用

Artz 指出信任已经深入到了计算机科学的多个方面<sup>[15]</sup>: (1)人类用户、软件 Agent 以及日渐增多的提供服务的设备,均需要在各种应用和场合下得到信任;(2)计算机和用户之间的通信信道以及要交互的内容也需要信任度量;(3)信任可以被用来保护数据、发现不准确的信息、获得高质量的服务等;(4)信任在许多问题中充当激励手段,引导自治的实体采取对群体和个体均有利的行为。

电子商务是信誉系统应用的一个直接领域。除了在 eBay 等购物网站中使用的在线信誉系统<sup>[2]</sup>外, Kim 开发了一个完整的基于信任的消费者决策模型,用以支持消费者从站点购物时的决策过程<sup>[17]</sup>。Avegliano 等以消费者/企业市场为背景,提出并讨论了基于信誉的伙伴形成的方法<sup>[18]</sup>。方健雯建立了网上拍卖的信任机制<sup>[19]</sup>。童向荣等利用动态交互信任来有效检测对手的异常行为,防止欺骗<sup>[20]</sup>。

信任逐渐成为互联网中判断 Web 资源、站点和服务是否可靠的一般性方法<sup>[21]</sup>。利用 MAS 的信任机制,可以采用推荐系统的形式进行合作搜索,由 Agent 相互提供搜索服务<sup>[22-23]</sup>。Birukov 等开发的一个基于 Agent 的推荐系统,支持同一社区中具有相同兴趣的 Agent 合作搜索 Web 页面<sup>[24]</sup>。Krupa 等利用信任对维基百科的撰稿者进行评价<sup>[25]</sup>。信任在 Web Service 中选择服务提供者方面起到基础性的作用,相关的议题包括信任方的信心、与被信任方之间的竞争,以及被信任方的可靠性、集成性等<sup>[21]</sup>,Caballero 等建立的信任模型辅助完成了对 Web 服务的选择<sup>[26]</sup>。Zhu 等构造的信任模型能帮助服务 Agent 做出理性的选择<sup>[27]</sup>。

信任和信誉系统也应用于以网络为基础的各种新型计算模式中。在网格计算中,可以借助于信任判断获得资源的可靠性<sup>[28,29]</sup>。在 P2P 网络环境下, Xiong 建立了基于信誉的信任模型<sup>[30]</sup>。近两年受到关注的云计算针对服务的交付和使用模式,通过网络以按需、易扩展的方式获得所需的服务,这使得服务成为整个云计算的核心,服务实施的过程必须依靠相关实体之间的交互与协作完成。从这个角度看,信任机制在云计算乃至未来的基于互联网的云计算<sup>[31]</sup>中,将起到更加重要的作用。

作为对自治和智能的实体间交互的抽象, MAS 中的信任机制在计算科学中也受到了关注。Ramchurn 研究了基于信任机制的 Agent 的任务分配<sup>[32]</sup>。Zhao 等将信任机制引入合同网中,从而改善合作求解的效果<sup>[33]</sup>,还基于 Agent 信任机制,解决各个 Agent 对待协商过程的态度问题,以达到更好的协商效果<sup>[34]</sup>。Griffiths<sup>[35]</sup>将信任用于联盟的形成,信任值可以根据 Agent 的个性动态更新, Vassileva 等<sup>[36]</sup>将信任用于联盟的形成,保证了系统的稳定性和可扩展性。童向荣等研究了长期联盟中的信誉问题<sup>[37]</sup>。

## 3 信任和信誉系统的技术

### 3.1 信任的表示和度量

对信任的表示和度量存在两种观点:认知观点和数值观点。

持认知观点的学者认为只有具备认知能力的 Agent 才会

有信任,信任由一系列的信念(belief)组成,信任是这些信念的函数<sup>[38]</sup>。在 Agent *a* 决定是否与 Agent *b* 交互前,需要基于多方面的信念完成对 *b* 的信任评价。以此为基础,构建信念的逻辑体系,完成信任的推理。Pinyol 等人<sup>[39]</sup>将认知信誉模型 Repage 整合到认知 BDI Agent 中,用信念逻辑获得 Repage 信息的语义,建立起一个完整的分级 BDI 模型,以执行 BDI 推理。Hubner 等人<sup>[40]</sup>将 Castelfranchi 提出的认知信任模型<sup>[41]</sup>整合到 BDI 框架中,并给出了实现方法。

数值观点也称为概率观点, Agent 可以利用过去自己的直接经验和获得的信誉信息做出主观性评价,预测可能的合作伙伴在将来执行某一特定行为的概率。对过去交互的评定值和对未来交互的期望都是用实数值表示的。Jurca 在博弈论的背景下,简单地用 0 和 1 表示评定,分别对应对抗和合作<sup>[42]</sup>。在线信誉系统中,普遍采用主观性度量方法,例如 eBay 中分别用 -1, 0, 1 表示否定的、中立的和肯定的评价<sup>[43]</sup>,有些系统将评定划分为更多的等级,如用 1-5 分别表示很差、差、一般、好、很好<sup>[2]</sup>。在与应用无关的信誉系统研究中,对交互的评定及信任值一般都规范化到域 $[-1, 1]$ 中,在此基础上,利用统计学、概率论等方法,汇总 Agent 的经验和搜集的信息,来计算对客体的信任值。

常采用多维的方法对交互进行度量和评价,这称为多上下文的度量。FIRE 模型<sup>[9]</sup>以五元组 $(a, b, c, i, v)$ 的形式支持用多个项目(*c*)评定同一交互,但没有给出多维评定的融合方法。ReGreT 模型<sup>[44]</sup>定义了一个本体结构,将多维度的评价组合为总的一般评价,但这种组合方法忽略了各维的联系且所有维度都必须做出评价。Sensoy 等<sup>[45]</sup>引入了一个更通用的本体共享交互经验,这种方法也可以引入到共享交互评定的信誉模型中。

基于认知的方法从社会心理学角度是一种很自然的观点,但是对其他 Agent 思维状态建模尚存在很多问题有待解决,采用认知的观点处理在开放式、大规模的 MAS 中的信任问题尚不合适<sup>[9]</sup>。数值观点的困难在于定量的主观评价并不总是恰当的。目前,也有些工作着眼于将这两种观点下的模型进行组合<sup>[46]</sup>。

### 3.2 信誉系统的体系结构

信誉系统的体系结构规定了收集、传播、计算信任的方法。根据交互后产生的评定记录的存储和汇总的方式不同,信誉系统的体系结构分为集中式、分布式和混合式 3 种。

集中式信誉系统设有一个信誉中心,负责收集与汇总所有的交互评定信息。每次交互完成后,交互双方均向信誉中心报告对交互的评定,信誉中心负责存储并汇总这些评定。在未来的交互之前,评价 Agent 可以从信誉中心查询目标 Agent 的信誉。目前电子商务领域中,多使用集中式系统,称为在线信誉模型。集中式方法的优点是实现简单,评价失败的风险小。不足之处包括:不能区分直接信任和信誉,不利于评价 Agent 个性化地利用信誉信息,不能快速适应环境的动态变化。另外,在线信誉模型在实际应用中因涉及太多的人为因素而无法发挥其应有的作用<sup>[2]</sup>。

越来越多的多 Agent 系统朝着大型化、开放、动态的方向发展,这些系统中不存在一个中心的权威机构,需要将信誉系统建立在分布式框架上。在分布式的信誉系统中, Agent 各自保存自己的评定,在交互之前,评价 Agent 通过询问一定范

围内的其他 Agent(称为证人),得到目标 Agent 的信誉,依靠存储的直接信任和获得的信誉信息,利用自己有限的的能力,独立地评价未来可能的合作伙伴。多数分布式的信誉系统采用直接经验和证人信息分别处理的模式进行处理<sup>[47]</sup>。一般认为,直接经验比证人信息更可靠,直接经验的权重大于证人信息的权重<sup>[9]</sup>。分布式信誉系统充分发挥了智能 Agent 自治、灵活的优势,适应了一些问题本身的环境约束,但也面临着一些突出的问题,例如:找不到证人会造成评价失败,证人视角不同会产生有偏差的评定,证人会为私利而撒谎或不提供信息等,这些问题也正是信誉系统研究的热点。

混合式信誉系统引入一些分布在 MAS 中专门服务于信誉传播和汇总的 Agent,如在 Jucar 等人提出的信誉系统<sup>[42]</sup>中,经纪人 R-Agent 负责购买 Agent 交互后产生的评定。在交互前,评价 Agent 要向 R-Agent 购买目标 Agent 的信誉信息,将信誉的传播视为信息的交易,实现评定信息的传播。Kravari 等人<sup>[48]</sup>提出了一个混合的信任模型 T-REX,以一种动态和灵活的方式将 Agent 个体的交互信任与来自证人的信誉组合在一起,通过认证中心(certified authority),保证证人诚实地提供报告,同时也降低了对存储成本和通信带宽的要求。混合式信誉系统发挥了集中式和分布式各自的优势,也有利于构造更有效的信誉机制。

### 3.3 信任的汇总

#### 3.3.1 统计方法

这种方法考虑构成信任评估的多方面因素,利用统计的手段,将多种因素综合起来,形成对 Agent 信任的评价。

Xiong 用交互后的反馈值、参与交互的总次数、证人的可信度 3 个参数和反映交互上下文与环境上下文重要性的两个因子计算信任,既基于历史信息计算将来交互的可能性,又考虑了交互所处的环境<sup>[30]</sup>。

FIRE 模型<sup>[9]</sup>是独立于应用的分布式信誉系统模型,采用 4 种部件共同完成信任的评价:(1)基于直接经验的交互信任(Interaction Trust, IT);(2)根据评价 Agent 与目标 Agent 角色定义的基于规则的信任(Role-based Trust, RT);(3)由证人提供的信息计算得到的证人信誉(Witness Reputation, WR);(4)由目标 Agent 主动向评价 Agent 提供的第三方参考(Certified Reputation, CR)。前两种部件是基于评价者直接经验的直接信任,而后两种是基于其他 Agent 提供的信息的信誉评价。

FIRE 模型中各种部件均采用同样的形式描述信息,也采用同一公式进行求解。设  $K$  为部件名称,  $K \in \{I, R, W, C\}$ , 分别对应上述 4 个部件,  $R_k$  为各部件使用的评定信息数据库,  $v_i$  是评定值, 权值  $\omega_K(r_i)$  与交互时间等因素有关。则利用  $K$  为部件求得的 Agent  $a$  对 Agent  $b$  针对评价项目  $c$  的信任为:

$$T_K(a, b, c) = \frac{\sum_{r_i \in R_K(a, b, c)} \omega_K(r_i) \cdot v_i}{\sum_{r_i \in R_K(a, b, c)} \omega_K(r_i)}$$

最后,对各部件采用加权平均的方法合并,求出 Agent  $a$  对 Agent  $b$  针对评价项目  $c$  的信任  $T(a, b, c)$ 。初始时,直接经验的权重大于信誉信息,在随后的运行中,部件的权重可以自适应调整。

ReGreT 模型<sup>[44]</sup>依靠直接信任以及来自社会结构等多方面的信息综合完成信任评价。ReGreT 模型有效地解决了直

接信任与信誉的组合问题,充分利用可获得的各种与信誉有关的信息,既评价信任又给出可靠性度量。

ReGreT 模型首先针对每一个特征  $\varphi$  评价直接信任和信誉。评价者 Agent  $a$  对目标 Agent  $b$  的直接信任  $DT_{a \rightarrow b}(\varphi)$  及其可靠性  $DTRL_{a \rightarrow b}(\varphi)$  由历史记录得到的评定  $Imp(o_i, \varphi)$  按与时间有关的权重函数  $\rho(t, t_i)$  加权累加得到。

ReGreT 模型中的信誉评价与 Agent 所处的社会结构密切相关。信誉分为 4 类:(1)证人信誉(witness reputation)根据证人提供的信息计算得出;(2)邻居信誉(neighbourhood reputation)利用评价 Agent 对邻居的直接信任与邻居同目标 Agent 的社会关系(竞争或合作),借助事先规定的模糊规则计算得到;(3)系统信誉(system reputation)基于 Agent 在系统中的角色和一般性质,根据应用领域直接定义;(4)为防止以上信息全部缺失而使信誉评价失败,引入了默认信誉值(default reputation)。在此基础上,计算出 Agent 对目标 Agent 的信誉评价  $R_{a \rightarrow b}(\varphi)$ 。

在 ReGreT 模型中,如果对目标有绝对可靠的直接经验,将完全使用直接经验,由各种信誉加权求和得来的信誉只是作为补充。由直接信任和信誉共同构成的信任为:

$$Trust_{a \rightarrow b}(\varphi) = DTRL_{a \rightarrow b}(\varphi) \cdot DT_{a \rightarrow b}(\varphi) + (1 - DTRL_{a \rightarrow b}(\varphi)) \cdot R_{a \rightarrow b}(\varphi)$$

此外, Gómez 评价信任时使用了 3 种部件:直接交互信任(direct trust, DT)、基于广告的信任(advertisements-based trust, AT)和基于推荐的信任(recommendations-based trust, RT)<sup>[49]</sup>。多数模型假设不信任与自己的经验有差异的信息,而 Gómez 用这种差异预测其他 Agent 行为的变化。将 DT, AT 和 RT 3 个部件的评价结果按可靠性加权平均组合到一起,得到对目标的信任。Lam<sup>[50]</sup>着眼于半竞争环境,认为 Agent 仅凭信任(或信誉)和期望效用进行决策,均可能受骗而损害自己的收益。对于交互过程中信息的接受者,构造了信任模型(trust model),用于决定是否相信发送者的消息;对于发送者,构造了诚实模型(honest model),根据对方的特征,决定在交互过程中是否撒谎,以增加自己的收益。Lam 还提供了在动态环境中对各种参数(文中称态度)自适应调整的方法。Rettinger 等<sup>[51]</sup>将机器学习的方法用于信任的推理,提出了一种有限隐藏理性的信任模型 IHRTM,它使用基于统计的理性学习方法,实现和学习上下文敏感的信任。Liang<sup>[52]</sup>在总结各种信誉模型的基础上,利用抽象、通用的分布式信任推理模型对各种算法进行了仿真,得出的结论之一是:复杂的评定汇总的算法并不总比简单的出色。统计方法具有简单性,因此在进一步的研究中应该受到重视,这也符合 Agent 计算资源受限的前提。

#### 3.3.2 概率方法

信任是 Agent 对交互合作者将来执行某一特定动作的估计,是一种主观的概率。面对环境的不确定性和信息的不完全性,概率论方法成为一种可行的选择:在选择交互伙伴之前,可以通过概率的期望值进行估计。

Teacy 的 TRAVOS 模型<sup>[53]</sup>使用概率的方法,对交互伙伴的信任进行建模。模型采用 0-1 分布,考虑了 Agent 之间过去的交互,当个体的经验不足时,利用了从第三方获得的信誉信息。在 TRAVOS 模型中,在不完全信息的前提下将  $t$  时该信任者  $a_r$  对被信任者  $a_w$  的信任  $\tau_{a_r, a_w}$  定义为在给定  $a_r$  的

交互知识  $O_{a_r, a_{tr}}^{t, t}$  的前提下,对  $B_{a_r, a_{tr}}$  的期望值。在这里,交互知识  $O_{a_r, a_{tr}}^{t, t}$  是  $a_r$  观察得到的截止  $t$  时的交互产生的结果的集合; $B_{a_r, a_{tr}}$  代表被信任者  $a_{tr}$  在接下来与信任者  $a_r$  可能的交互中完成承诺的概率。

Mui 使用 Bayesian 方法,从社会学和进化论的角度给出了一个信任和信誉的计算模型<sup>[54]</sup>。Mui 的模型将互惠的行动作为重要的因素。互惠是两个 Agent( $a_i$  和  $a_j$ )之间的一种对称的变量,互惠程度越高, $a_i$  和  $a_j$  对彼此间互惠动作的期望越大。用  $\theta_{ji}(c)$  度量  $a_i$  采取惠及  $a_j$  的动作的可能性,即  $a_i$  在与  $a_j$  相关的社会网络中的信誉, $c$  是感兴趣的上下文;用  $D_{ji}(c)$  表示  $a_j$  与  $a_i$  之前的交互历史。 $a_j$  可以在  $D_{ji}(c)$  的前提下,计算对  $a_i$  的信任,即  $a_j$  对  $a_i$  互惠动作的期望  $\tau(c) = E[\theta(c) | D(c)]$ 。

概率论的方法由先验的经验预测未来行为,要求的数据量一般较多,需要 Agent 之间有较频繁的交互。另外,在开放 MAS 中,历史交互数据未必符合特定的分布规律,尤其在数据量不充足的情况下。这些问题制约了概率论方法的效力。

### 3.3.3 信念理论

Yu 和 Singh<sup>[55]</sup> 利用 Dempster-Shafer 证据理论建立了信任的计算框架。信念理论在信任计算中主要用于解决信任的传递性问题,涉及到的基本问题包括信任在传播链中如何传递,以及多个来源的信任如何进行合并的问题,Dempster 规则提供了组合不同来源证据的经典算子。

为表示在信任判断中必须满足的语义约束,Jøsang<sup>[56]</sup> 借鉴 D-S 证据理论,将不确定性引入到他的观点模型中,提出了主观逻辑(subjective logic)。一个观点可以是信任,也可以是不信任和不确定,以此对计算机系统信任进行表示和推理。进一步 Jøsang 还给出了信任传递和合并的算法。其中,每个传递的信任路径用折扣算子(discounting operator)来计算,而两条信任路径的组合用共识算子(consensus operator)计算。Wang 等<sup>[11]</sup> 在对信任网络形式化表示的基础上,定义了信任的合并算子(concatenation operator)和聚集算子(aggregation operator),通过赋予这些算子的一些重要性质,将信任网络中的信任评定正确地合并起来,从而限制了恶意 Agent 提供错误信息。

### 3.3.4 模糊推理方法

信任是用于解决不完全和不确定环境中决策问题的,模糊方法是解决此类问题的途径之一。Rubiera 等<sup>[57]</sup> 定义了一种信誉的模糊模型,其能够结合 Agent 个体的经验及其他 Agent 提供的参考信息。模型中用到了其他 Agent 的模糊信誉表示个性,使得 Agent 之间的交互接近于人类思维的方式。模糊集聚使得具有相同偏好的 Agent 能够聚合为一个群体,由此引出了一种非常有效的分布式通信方式。另外,在 ReGreT 模型<sup>[44]</sup> 中,证人信誉、系统信誉等也使用到了对模糊信息的表示。

## 4 几个相关问题

### 4.1 群体信誉

在 MAS 的信任模型中,信任对象可以是单个的 Agent,也可以是一个由多 Agent 构成的群体<sup>[2]</sup>。群体信誉(group reputation)可以由其中每一个成员的个体信誉的统计值获得,也可以由外部的其他成员将群体当作一个整体进行评价。Hales<sup>[58]</sup> 揭示了人工社会中群体信誉的作用,通过将 Agent

分组来评价每个群体的信誉。这种群体信誉在特定条件下成为一种促进善行规范的强有力的机制。在经济学和社会学中,团队信誉(team reputation)<sup>[59]</sup> 和群体信誉<sup>[3]</sup> 也得到了充分的重视。然而,在各种 MAS 的信任模型中,一般并不区分个体信誉和群体信誉。

在 MAS 中,Agent 之间往往能够形成一定的结构,以利于它们之间的相互合作,这些结构包括群体(group)、团队(team)、联盟(coalition)、组织(organization)等。董向荣等认为,Agent 组织信誉的研究可以完善 Agent 组织研究,能够为群体 Agent 交互与合作提供理论基础<sup>[60]</sup>,他还给出一种 Agent 组织信任计算模型 CMAOT<sup>[61]</sup>。Kastidou 等<sup>[62]</sup> 设计了一种机制,即在允许 Agent 迁移的环境中,社区间交换 Agent 的信誉信息,其中使用支付函数(payment function)激励社区诚实地交换信息。Hermoso 等人<sup>[63]</sup> 分析了组织结构如何支持 Agent 的选择过程,当使用有利于组织结构的信任机制时,可以改善 Agent 的决策。

在计算科学领域,Agent 群体信誉还没有得到充分的重视。而计算模型和软件工程的进展已经对这一问题提出了要求。例如互联网计算和云计算将是互联网资源利用的新模式,也将引发面向大众的软件工程变革<sup>[31]</sup>。在这样的环境中,个体 Agent 和各种不同形式的群体 Agent 并行存在,软件构件之间合作更加深入,由此产生新的需求。将人类社会群体信誉机制引入计算领域,或者建立起适合计算领域的新的群体信誉模型都具有一定的现实意义。

### 4.2 信息不准确问题

在信任评价过程中需要共享评价信息,而在交互前由证人提供的报告与交互后得到结果之间往往存在差异,这就是信息不准确问题。信息不准确问题具体表现为信息错误(error)、有偏(bias)和不完全(incomplete)<sup>[44]</sup>。在集中式的信誉系统中,由于专门的信誉中心的存在,合作双方的历史能够快速集中起来,提供不准确信息易于被发现,理性证人不会通过撒谎来破坏自己良好的声誉。在分布式的信誉系统中,可以依据信息的精确程度将信誉系统分为 3 个层次<sup>[44]</sup>:(1) Level 0:不考虑欺骗行为;(2) Level 1:Agent 可能隐藏信息或提供有偏见的信息,但从撒谎;(3) Level 2:模型中有处理撒谎的机制。已有的研究大多建立在 Level 0 层,即假设 Agent 诚实地交换信息。

解决信息不准确问题的思路之一是精心设计交互协议与机制,迫使参与交互的 Agent 诚实地交换信息。这样做尽管增加了交互的代价,并失去了一些灵活性,但总体效果还是明显的。但是,在实际系统中,这些手段并不总是有用的<sup>[12]</sup>,需要通过统计或机器学习的手段检测和去除不公平的评定。

解决信息不准确问题的另一种思路是将证人提供的信息同自己积累的观察值进行纵向(考察证人不同时期的报告)和横向(不同证人在同一时期的报告)比较后做出判断,如将同历史记录差异大的新评定视为不公平的评定<sup>[53,64]</sup>。FIRE 模型<sup>[9]</sup> 中评价 Agent 将每次得到的评定值与交互后得到的实际观察值的差值记录下来作为直接经验,从而调整 WR 和 CR 部件的权重,使模型具有自适应功能。这一思路的实现需要积累足够的信息,当可供对比的数据很少时,对不准确信息的识别效果不能令人满意。

“和历史及多数观点不一致就是不准确的”是一个非常强的假设。而产生不准确问题的原因包括目标 Agent 的性能变

化、证人完成准确评价的能力、证人的故意操纵等多个方面<sup>[9]</sup>,在实际应用中还会涉及人为的因素。将信息不准确问题全归因于证人是公平的,也不利于问题的解决。贺利坚等人<sup>[65]</sup>通过用 Agent 间的距离代表 Agent 之间的关系,用目标 Agent 的操作半径抽象与证人无关的信息不准确因素,提出了一种评价者 Agent 可以从自己的角度修改证人 Agent 的报告的方法,并在交互后利用自适应算法,学习到更加准确的目标 Agent 的操作半径,从而保证评价者 Agent 获得更高的收益。

#### 4.3 信息贫乏问题

在信任和信誉系统中还存在信息贫乏的问题,即:评价者 Agent 无法获得足够的信息来评价未来的合作者。即使是在 Agent 数目大、交互频繁的场所,由于交互的种类、主题众多的原因,获得足量符合评价目的的信息也不是一件容易的事情。

可以充分利用 MAS 中 Agent 之间形成的社会网络去解决信息贫乏的问题。由于信任双方存在着商业或社会的交易,或者都是某共同体中的参与者,它们之间存在着一些直接或间接的关系,从而形成一定的社会网络。通过社会网络分析可以增加 Agent 对所处社会环境的知识。这种方法尤其适用于评价者 Agent 无直接交互经验的场合,但一般需要以信任关系具备传递性为前提,即使这种传递性带有一定的折扣。Mui 的模型<sup>[54]</sup>中信任信息直接来自社会网络。Hang 等人定义了聚集(aggregation)、串联(concatenation)和选择(selection)3个算子,以使信任在社会网络中有效、正确地传输<sup>[66]</sup>。在 Battiston 的模型<sup>[67]</sup>中,Agent 利用彼此间的动态信任关系,在社会网络上搜索推荐信息,提高了其搜索效率。ReGreT 模型<sup>[44]</sup>在社会网络的基础上,构造了多种信誉部件,还有些模型<sup>[68]</sup>研究社会网络中信任关系的动态进化。

在现实中,当特定的信息缺失时,人们总是利用类似的信息进行替代。这种方法也被广泛用于解决信任/信誉系统中信息贫乏的问题。Caballero 等利用任务的相似度建立信任模型,完成了对相似 Web 服务的选择<sup>[26]</sup>,解决了在特定任务上从未直接交互或参考信息严重缺乏的问题。李景涛等<sup>[69]</sup>改进了判断评分行为相似程度的算法,建立的模型较好地解决了评分信息稀疏性的问题。Tavakolifard 等人<sup>[70]</sup>针对 Web 中由于信息稀少,无法准确预测两个不熟悉的主体之间的信任值的问题,在将信任推理与用户相似性结合的基础上,提出了 TILLIT 模型,以此来解决信息贫乏的问题。

#### 4.4 异构模型互操作问题

目前提出了不少的信任/信誉系统,分别在不同的领域开展应用。但是,各种系统中采用的术语不一致,并缺乏统一的性能测试标准<sup>[9,10]</sup>,这给模型测试以及需要具有异构信誉模型的应用系统的互操作带来非常大的困难。

信任概念不清晰是导致研究混乱的原因之一。众多研究者致力于信任概念的研究,其中包括 Mayer 等利用语义学的方法对不同领域的研究成果进行分析,总结、提炼出信任的概念模型<sup>[71,72]</sup>。由于面对的背景不同,是否需要构造这种“普适的”信任逻辑模型尚待商榷。要在 Mayer 概念模型基础上建立可计算的信任模型,需要解决对模型中各种因素的度量。但是,能力、善意、正直等属性本身主观性很强,目前的建模手段难以给出实用的方案。

有些学者希望建立与应用无关的信誉系统模型,从而使

之适用于不同的领域,例如 FIRE 模型<sup>[9]</sup>和 Regret 模型<sup>[44]</sup>。但这些模型只是提供了一套统一的信任表示、传播与汇总机制,并未考虑与已有其他模型的结合。ART<sup>[16]</sup>的提出使各种信誉系统的测试和竞赛有了一个统一的平台,利用 ART 规定的方法和通信协议,研究人员自行实现 Agent 内部的信誉系统,实现同其他异质 Agent 之间的交互。然而,ART 在信息的表示形式、交互方式等诸多方面设置的限制,使很多信誉机制和测试内容不方便甚至不能够进行下去<sup>[6,26]</sup>。

Casare 等提出了信誉的功能性本体 FORe<sup>[73]</sup>,描述了信誉系统中的关键概念,明确了信誉的属性、Agent 在信誉系统中的角色以及信誉的处理。FORe 是在归纳了心理学与人工智能等领域文献的基础上提出的,目的在于为不同信任模型的互操作提供基础。Vercouter 和 Brandão 等研究了目前主要信誉系统与 FORe 之间的映射,并将之用于处理异质 Agent 信誉系统的互操作<sup>[74,75]</sup>。Nardin 提出一个方便不同信誉模型互操作的面向服务的体系结构 SOAR<sup>[76]</sup>,即使用外部的本体映射服务 OMS 与 Agent 内部的翻译器模块,支持信誉模型中的概念与公共本体之间的相互映射和转换。He<sup>[77]</sup>在 FORe 和已有的信誉系统基础上,提出了一个用于支持异构信誉模型互操作的平台,提出 4 种形式的推荐信息,为开发有多种信誉传播机制的多 Agent 的应用系统提供支持。

**结束语** 交互是开放 MAS 中的 Agent 合作的基本要求,引入信任和信誉系统的目的就在于通过这种在社会科学研究中行之有效的方法,选择出能保证收益的交互伙伴,乃至营造良好的合作环境。本文就信任和信誉的概念、研究意义、研究现状、具体技术,以及目前的几个热点问题进行了综述。由于开放 MAS 环境的开放、大规模、复杂等性质,以及 Agent 本身在计算、存储、通信等能力方面的限制,尚有不少问题待深入研究,尤其是在结合人工智能、社会学、管理学等学科的方法和成果方面,尚有不少工作可做,这也正是作者所在团队的努力方向。

#### 参 考 文 献

- [1] Ramchurn S D, Huynh D, Jennings N R. Trust in Multi-Agent Systems[J]. The Knowledge Engineering Review, 2004, 19(1): 1-25
- [2] Jøsang A, Ismail R, Boyd C. A Survey of Trust and Reputation Systems for Online Service Provision[J]. Decision Support Systems, 2006, 43(2): 618-644
- [3] 阿克塞尔罗德. 合作的进化(修订版)[M]. 吴坚忠,译. 上海:上海人民出版社,2007
- [4] 罗德里克. M. 克雷默,汤姆. R. 泰勒. 组织中的信任[M]. 管兵,等译. 北京:中国城市出版社,2003
- [5] 郑也夫. 信任论[M]. 北京:中国广播电视出版社,2001
- [6] 贺利坚,黄厚宽,张伟. 多 Agent 系统中信任和信誉系统研究综述[J]. 计算机研究与发展, 2008, 45(07): 1151-1160
- [7] Povey D. Developing electronic trust policies using a risk management model[C]//Proceedings of the 1999 CQRE Congress. LNCS 1740. Berlin: Springer, 1999: 1-16
- [8] Gambetta D. Can We Trust Trust? [M]//Gambetta D. Trust: Making and Breaking Cooperative Relations, electronic edition, Department of Sociology, University of Oxford, 2000: 213-237
- [9] Huynh T D. Trust and Reputation in Open Multi-Agent Systems[D]. Southampton: Electronics and Computer Science, University of Southampton, 2006
- [10] Fullam K, Klos T B, Muller G, et al. The Agent Reputation and

- Trust(ART) Testbed Architecture[C]//Alguer, Italy. Proceedings of the 8th International Conference of the ACIA, CCIA 2005. IOS Press, 2005; 389-396
- [11] Wang Y, Singh M P. Trust Representation and Aggregation in a Distributed Agent System[C]// Boston, Massachusetts, USA. Proceedings of Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference. AAAI Press, 2006
- [12] Jøsang A. Trust and Reputation Systems [C]// Aldini A, Gorrieri R, eds. *Foundations of security analysis and design IV*, LNCS 4677. Berlin; Springer, 2007; 209-245
- [13] Reece S, Rogers A, Roberts S, et al. A Multi-Dimensional Trust Model for Heterogeneous Contract Observations [C]// Proceedings of 22nd AAAI Conference on Artificial Intelligence. Vancouver, Canada, 2007; 128-135
- [14] 金士尧, 黄红兵, 范高俊. 面向涌现的多 Agent 系统研究及其进展[J]. 计算机学报, 2008, 31(6): 881-894
- [15] Artz D, Gil Y. A survey of trust in computer science and the Semantic Web[J]. *Journal of Web Semantics*, 2007, 5(2): 58-71
- [16] Fullam K, Klos T B, Muller G, et al. A specification of the Agent Reputation and Trust(ART) testbed; experimentation and competition for trust in agent societies[C]// Utrecht, Netherlands. 4rd International Joint Conference on Autonomous Agents and Multiagent Systems(AAMAS 2005). ACM, 2005; 512-518
- [17] Kim D J, Ferrin D L, Rao H R. A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents[J]. *Decision Support Systems*, 2008, 44(2): 544-564
- [18] Avegliono P, Sichman J S. Using the RePart Simulator to Analyze Different Reputation-Based Partnership Formation Strategies within a Marketplace Scenario[M]// Falcone R, Barber S, Sabater-Mir J, et al. LNCS 5396; Trust in Agent societies. Berlin; Springer, 2008; 226-243
- [19] 方健雯. 网上拍卖信任机制的建立[D]. 上海: 复旦大学, 2006
- [20] 童向荣, 黄厚宽, 张伟. Agent 动态交互信任预测与行为异常检测模型[J]. 计算机研究与发展, 2009, 46(8): 1364-1370
- [21] Grandison T, Sloman M. A Survey of Trust in Internet Application[J]. *IEEE Communications Surveys and Tutorials*, 2000, 3(4): 2-16
- [22] Wei Y Z, Jennings N R, Moreau L, et al. User evaluation of a market-based recommender system[J]. *Autonomous Agents and Multi-Agent Systems*, 2008, 17(2): 251-269
- [23] He L, Huang H, Zhang W, et al. ALRS; agent based literature recommendation system[C]// Wuhan, China. Proceedings of International Workshop on Intelligent Systems and Applications (ISA2009). IEEE Computer Society, 2009; 1485-1488
- [24] Birukov A, Blanzieri E, Giorgini P. Implicit: An Agent Based Recommendation System for Web Search[C]// Utrecht, Netherlands. Proceeding of the 4th International Joint Conference on Autonomous Agents and MultiAgent Systems(AAMAS2005). ACM Press, 2005; 618-624
- [25] Krupa Y, Vercouter L, Fred J, et al. Trust Based Evaluation of Wikipedia's Contributors[C]// Proceedings of Engineering Societies in the Agents World X, 10th International Workshop (ESAW 2009). LNCS 5881. Berlin; Springer, 2009; 148-161
- [26] Caballero A, Botía J A, Gómez-karmeta A F. Trust and Reputation Model based on WSMO[C]// Hakodate, Japan. Proceeding of Trust Workshop in 5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006). ACM, 2006; 9-18
- [27] Zhu M, Jin Z. An Agent-Based Trust Model for Service-Oriented Systems[C]// Falcone R, Barber S, Sabater-Mir J, et al. LNCS 5396; Trust in Agent societies. Berlin; Springer, 2008; 162-181
- [28] Foster I, Jennings N R, Kesselman C. Brain Meets Brawn: Why Grid and Agents Need Each Other[C]// New York, USA. 3rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2004). IEEE Computer Society, 2004; 8-15
- [29] 张煜, 林莉, 怀进鹏, 等. 网络环境中信任-激励相容的资源分配机制[J]. 软件学报, 2006, 17(11): 2245-2254
- [30] Xiong L. A reputation-based trust model for peer-to-peer e-commerce communities [C] // Proceedings of IEEE International Conference on E-Commerce(CEC2003). 2003; 275-284
- [31] 李德毅, 张海粟. 超出图灵机的互联网计算[J]. 中国计算机学会通讯, 2009, 5(12): 8-16
- [32] Ramchurn S D, Dash R K, Giovannucci A, et al. Trust-based mechanisms for robust and efficient task allocation in the presence of execution uncertainty[J]. *Journal of AI Research*, 2009, 35: 119-159
- [33] Zhao X, Huang H, He L. The Contract Net Protocol with Trust Model in Open Multi-Agent System[C]// Wuhan, China. Proceedings of International Conference on Computational Intelligence and Software Engineering(CISE 2009). IEEE, 2009; 1-4
- [34] 赵翔, 黄厚宽, 邵翀. 基于 Agent 信任机制的一体化协商研究[J]. 北京交通大学学报, 2009, 33(5): 95-98
- [35] Griffiths N, Luck M. Coalition formation through motivation and trust[C]// Melbourne, Australia. Proceedings of Second International Joint Conference on Autonomous Agents & Multiagent Systems, AAMAS 2003. ACM, 2003; 17-24
- [36] Vassileva J, Breban S, Horsch M. Agent Reasoning Mechanism for Long-Term Coalitions Based on Decision Making and Trust [J]. *Computational Intelligence*, 2002, 18(4): 583-595
- [37] Tong X, Huang H, Zhang W. Agent long-term coalition credit [J]. *Expert Systems with Applications*, 2009, 36(5): 9457-9465
- [37] Falcone R, Pezzulo G, Castelfranchi C. A Fuzzy Approach to a Belief-Based Trust Computation [C] // Bologna, Italy. Trust, Reputation, and Security; Theories and Practice. AAMAS 2002 International Workshop, LNCS 2631. Berlin; Springer, 2002; 73-86
- [39] Pinyol I, Sabater-Mir J, Dellunde P, et al. Reputation-based decisions for logic-based cognitive agents[J]. *Autonomous Agents and Multi-Agent Systems*
- [40] Hubner J F, Lorini E, Vercouter L, et al. From cognitive trust theories to computational trust[C]// Proceedings of the 12th International Workshop on Trust in Agent Societies. Budapest, Hungary, 2009
- [41] Castelfranchi C, Falcone R. Principles of Trust for MAS; Cognitive Anatomy, Social Importance, and Quantification[C]// Paris, France. Proceedings of the Third International Conference on Multiagent Systems (ICMAS 1998). IEEE Computer Society, 1998; 72-79
- [42] Jurca R, Faltings B. Towards incentive-compatible reputation management[C]// Trust, reputation and security; theories and practice. LNCS 2631. Berlin; Springer-Verlag, 2003; 13-24
- [43] Jøsang A. Online Reputation Systems for the Health Sector[C]// Proceedings of ehPASS, The National e-Health Privacy and Security Symposium. Brisbane, Australia, 2006
- [44] Sabater J. Trust and reputation for agent societies[D]. Barcelona, Spain; Universitat Autònoma de Barcelona, 2003
- [45] Sensoy M, Yolum P. A context-aware approach for service selection using ontologies[C]// Hakodate, Japan. Proceeding of Trust Workshop in 5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2006). ACM, 2006;

- [46] Staab E, Engel T. Combining Cognitive with Computational Trust Reasoning[C] // Trust in Agent societies, LNCS 5396. Berlin; Springer, 2008; 99-111
- [47] Aberer K. P-grid: A self-organizing access structure for P2P information systems[C] // Trento, Italy. Cooperative Information Systems, Proceedings of 9th International Conference, CoopIS 2001, LNCS 2172. Berlin; Springer, 2001; 179-194
- [48] Kravari K, Malliarakis C, Bassiliades N. T-REX: A hybrid agent trust model based on witness reputation and personal experience [C] // Bilbao, Spain. Lecture Notes in Business Information Processing, vol. 61: E-Commerce and Web Technologies, 11th International Conference (EC-Web 2010). Berlin; Springer, 2010; 107-118
- [49] Gómez M, Carbó J, Earle C B. Honesty and trust revisited; the advantages of being neutral about other's cognitive models[J]. Autonomous Agents and Multi-Agent Systems, 2007, 15(3): 313-335
- [50] Lam K, Leung H. A Trust/Honesty Model with Adaptive Strategy for Multiagent Semi-Competitive Environments [J]. Autonomous Agents and Multi-Agent Systems, 2006, 12(3): 293-359
- [51] Rettinger A, Nickles M, Tresp V. Statistical relational learning of trust[J]. Machine Learning
- [52] Liang Z, Shi W. Analysis of ratings on trust inference in open environments[J]. Journal of Performance Evaluation, 2008, 65(2): 99-128
- [53] Teacy W T L, Patel J, Jennings N R, et al. TRAVOS: Trust and Reputation in the Context of Inaccurate Information Sources [J]. Autonomous Agents and Multi-Agent Systems, 2006, 12(2): 183-198
- [54] Mui L, Mohtashemi M, Halberstadt A. A Computational Model for Trust and Reputation[C] // Ralph H. Sprague. Proceedings of the 35th Hawaii International Conference on System Sciences (HICSS-35 2002). Los Alamitos, California: IEEE Computer Society, 2002; 188-196
- [55] Yu B, Singh M P. An evidential model of distributed reputation management[C] // Bologna, Italy. The First International Joint Conference on Autonomous Agents & Multiagent Systems (AAMAS 2002). New York, USA: ACM Press, 2002; 294-301
- [56] Jøsang A. A Logic for Uncertain Probabilities[J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2001, 9(3): 279-311
- [57] Rubiera J C, Lopez J M M, Muro J D. A fuzzy model of reputation in multi-agent systems[C] // Montreal, Quebec, Canada. Proceedings of the fifth international conference on Autonomous agents (AGENTS'01). New York, USA: ACM Press, 2001; 25-26
- [58] Hales D. Group Reputation Supports Beneficent Norms [J]. Journal of Artificial Societies and Social Simulation, 2002, 5(4)
- [59] Czarnitzki D, Stadtmann G. Uncertainty of outcome versus reputation: empirical evidence for the First German Football Division [J]. Empirical Economics, 2002, 27(1): 101-112
- [60] 童向荣, 张伟. 基于信任和声誉的 Agent 组织信誉[J]. 计算机科学与探索, 2007, 1(3): 325-330
- [61] 童向荣, 黄厚宽, 张伟. 一种 Agent 组织信任计算模型[J]. 北京交通大学学报, 2009, 33(5): 91-94
- [62] Kastidou G, Larson K, Cohen R. Exchanging reputation information between communities; a payment-function approach[C] // Pasadena, California, USA. Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI 2009). Morgan Kaufmann Publishers Inc., 2009; 195-200
- [63] Hermoso R, Centeno R, Billhardt H, et al. Extending virtual organizations to improve trust mechanisms[C] // Proceedings of 7th International Joint Conference on Autonomous Agents & Multiagent Systems (AAMAS 2008). Estoril, Portugal, IFAAMAS, 2008; 1489-1492
- [64] Jurca R, Faltings B. Minimum payments that reward honest reputation feedback[C] // Ann Arbor, Michigan, USA. Proceedings 7th ACM Conference on Electronic Commerce (EC-2006). ACM, 2006; 190-199
- [65] He L, Huang H, Zhang W. Revising and Learning: A New Approach to Inaccurate Ratings in Trust Model [J]. Journal of Computational Information Systems, 2010, 6(3): 679-687
- [66] Hang C, Wang Y, Singh M P. Operators for Propagating Trust and their Evaluation in Social Networks[C] // Budapest, Hungary. Proceedings of 8th International Joint Conference on Autonomous Agents & Multiagent Systems (AAMAS 2009). 2009; 1025-1032
- [67] Battiston S, Walter F E, Schweitzer F. Impact of Trust on the Performance of a Recommendation System in a Social Network [C] // European Conference on Complex Systems 2006 (ECCS '06). Oxford; Oxford Press, 2006
- [68] Walter F E, Battiston S, Schweitzer F. A model of a trust-based recommendation system on a social network[J]. Autonomous Agents and Multi-Agent Systems, 2008, 16(1): 57-74
- [69] 李景涛, 荆一楠, 肖晓春, 等. 基于相似度加权推荐的 P2P 环境下的信任模型[J]. 软件学报, 2007, 18(1): 157-167
- [70] Tavakolifard M, Herrmann P, Knapskog S J. Inferring Trust Based on Similarity with TILLIT[C] // Ferrari E, Li N, Bertino E, et al. Trust Management III; IFIP Advances in Information and Communication Technology. Vol. 300. Boston: Springer, 2009; 133-148
- [71] Mayer R C, Davis J H, Schoorman F D. An Integrative Model of Organizational Trust [J]. Academy of Management Review, 1995, 20(3): 709-734
- [72] Schoorman F D, Mayer R C, Davis J H. An Integrative Model of Organizational Trust: Past, Present, and Future[J]. Academy of Management Review, 2007, 32(2): 344-354
- [73] Casare S, Sichman J. Towards a Functional Ontology of Reputation[C] // Utrecht, Netherlands. 4rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2005). ACM, 2005; 505-511
- [74] Brandão A, Vercouter L, Casare S, et al. Extending the ART Testbed to Deal with Heterogeneous Agent Reputation Models [C] // Honolulu, Hawaii, USA. The Workshop on Trust in Agent Societies at The Sixth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007). IFAAMAS, 2007; 7-13
- [75] Vercouter L, Casare S J, Sichman J S, et al. An experience on reputation models interoperability based on a functional ontology[C] // Hyderabad, India. Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI'07). 2007; 617-622
- [76] Nardin L G, Brandão A, Sichman J S, et al. SOARI: A Service Oriented Architecture to Support Agent Reputation Models Interoperability[C] // Trust in Agent societies. LNCS 5396. Berlin; Springer, 2008; 292-307
- [77] He L, Huang H, Zhang W. A Framework to support Heterogeneous Reputation Models Interoperation in Multi-Agent System [C] // Shenyang, China. Proceedings of the second IEEE International Conference on Advanced Computer Control (ICACC 2010). IEEE Computer Society, 2010; 183-188