前　言

现代通信网络技术与信息技术的迅猛发展和日臻完善，已带来全球各行业、各领域和社会各角落的革命性变革。由此而衍生出的 Web 2.0 技术和科学 2.0 技术也成为“地球村”中的每一个参与主体（人、机、群，甚至类脑）所能交换，分享，投入，贡献数据、信息和知识以及享有这些产出和成果的基础性、关键性网络基础设施和载体。与此同时，以社交组织、内容媒介和智能计算为核心的社交媒体（social media）生态系统也在此基础上显示出综合化、多样化和智慧化特征，并进一步催生了这一崭新的网络和科学组织形态。

然而，用户在享有社交媒体所带来的便利的同时，也正遭受着非良序的社交组织形式和非友善的参与实体所带来的数据窃取、信息欺诈、隐私窥探和版权侵犯等严重困扰。社交媒体网络安全已成为近年来在信息技术、信息管理、社会科学等多学科领域的研究热点，也是交叉科学所共同面临的挑战性开放课题。

本书不仅是对国内外最新社交媒体网络安全相关安全理论与技术的全面分析和阐述，更重要的是，它是多媒体内容安全河南省高校科技创新团队和河南科技大学多媒体内容安全实验室近年来研究成果的汇集和精华提炼，并涵盖了对团队自主研发的“梦之声·飓风”多媒体社交网络原型系统 CyVOD 的实例剖析，展望了社交媒体安全研究新方向。全书分为五个部分：①社交媒体网络基础；②社交媒体网络用户信任评估；③社交媒体网络安全传播与风险管理；④多媒体社交网络的访问控制、委托授权和安全策略博弈；⑤多媒体社交网络原型系统和未来研究展望。本书所呈现出的最新理论研究成果将为解决社交媒体网络安全问题，实现现代信息社会媒体安全与数字内容产业的健康、良性和快速发展，提供系统的关键技术与理论支撑。

本书的出版受到河南科技大学学术专著出版基金资助。相关研究工作受到国家自然科学基金（No. 61370220）、河南省高校科技创新团队支持计划项目（No. 15IRTSTHN010）、河南省科技创新人才计划杰出青年基金（No. 134100510006）等的支持。

本著作的主要分工如下：河南省特聘教授张志勇博士负责本书的申请列选、论证和最终审定全稿，以及承担第 3 章、第 4 章、第 6 章、第 7 章、第 11 章的

撰写和全部参考文献的整理。此外，第 1 章、第 2 章、第 5 章由河南科技大学计算机系赵长伟博士撰写，第 8 章、第 9 章、第 10 章由河南科技大学计算机系王剑博士撰写。

这里，感谢河南科技大学“多媒体内容安全”研究团队的主要成员——研究生陈书全、杨丽君、陈庆丽、丰伟宁等所做出的科研工作及努力，以及研究生孙冉冉、韩林茜等对本书的编辑和校对工作给予的大力协助。此外，本书的最后完稿正值我在美国爱荷华州立大学计算机科学系（ISU CS）做访问教授（2015 ~ 2016 年）。其间，我加入了美国电气和电子工程协会（IEEE）前主席张可昭（Carl K. Chang）教授的软件工程实验室（SEL）从事社交媒体研究与开发，Carl 给了我很多宝贵意见和建议，特别是对我所做的几次 SEL Seminar 学术报告 Security, Trust and Risk：From Digital Rights Management to Multimedia Social Network；A Survey on Social Media Security, Privacy and Trustworthiness，给予了积极点评，这些都对本书的最终完稿具有非常重要的指导意义。在此，我对张可昭教授表示深深的感谢！

我们在本书的撰写过程中，力求精益求精，期间翻阅和参考了大量的国外相关文献资料，但因能力有限，难免仍有疏漏之处，敬请读者批评指正。

张志勇
于 ISU Atanasoff Hall
2016 年 4 月

目　录

第1章 绪 论

目前，媒体信息和社交网络呈现出深度“联姻”和加速发展的趋势，社交媒体网络（social media networks）已成为重要的互联网服务和关键应用之一。在社交媒体网络中，媒体信息通过社交网络用户之间建立的广泛的、动态的连接，以分享、推送、转发等方式传播，表现出速度快、便利程度高的显著优势。然而，当海量的媒体内容和爆炸式增长的社交用户共存于这个独特的虚拟社交网络世界时，又造成了安全策略配置难度增大、访问控制粒度要求提高等诸多问题，因此社交媒体网络安全更加凸显，亟待解决。社交媒体网络提供无处不在的数字媒体和信息服务的同时，如何保证媒体信息在合法的授权用户之间分享传播，防止非法的恶意分发和扩散，保护媒体信息拥有者的合法权益，并对媒体信息传播过程中可能面临的风险进行评估等，已成为社交媒体网络安全理论和技术研究的主要内容。

本章主要对社交媒体网络的概念、发展历程及其安全问题与安全技术等进行介绍。

1.1 社交媒体网络背景

1.1.1 社交媒体网络概念

社交媒体网络是一种将媒体信息和社交网络相结合，借助于社交网络特性分享、传播和推送（多）媒体信息的网络组织新形态，也是涵盖相关软件、工具、站点、平台和服务的统称。目前，社交媒体网络已成为综合化程度高、应用广泛、深受用户喜欢的互联网应用和服务之一，常见的社交媒体网络如 Facebook、Twitter、YouTube、腾讯微信（空间）、优酷土豆网等，都提供了相关（多）媒体信息服务。在社交媒体网络中，媒体信息、社交用户和用户之间的连接关系是社交媒体网络的重要组成部分。其中，社交用户可以通过社交网站构建自己公开的或半公开的用户简档信息，并依据用户自身的喜好建立好友关系列表，设置自己

的访问偏好或被访问要求的权限。用户之间的连接可以有多种类型，如好友关系、点击、转发、评论等。媒体信息主要通过用户之间的好友关系推送、分享和转发。

作为提供媒体信息服务的社交网络，社交媒体网络（G）可看作由节点、边和媒体信息构成的一种复杂的图，记为 $G(V, E, I)$。其中，节点 V 代表个人或组织；边 E 代表社会关系连接；I 代表通过社交网络传播的媒体信息。与一般的社交网络类似，社交媒体网络的连接结构也具有一定的特征，如无标度网络、小世界网络的特征。社交媒体网络的连接特点为媒体信息的快速传播提供了可能，通过社交媒体网络，可以便捷、安全地分享和传播媒体信息内容（Pang and Zhang，2015）。社交媒体网络媒体内容的传播具有传播速度快、点击次数高、影响力大等特点，同时，社交媒体网络也增加了数字内容传播方式，使信息传播的形式更加多样化。

1.1.2 社交媒体网络发展进程

社交媒体网络属于社交网络的一种类型，它是在互联网高度发展，具有社交和媒体传播功能的情况下出现的一种新的服务模式。社交媒体网络的发展可以追溯到互联网发展初期的社交网络。在互联网出现之前，社交网络这一概念已经被提出。但早期的社交网络是现实的社交用户之间的联系，与基于互联网平台的虚拟的社交网络相比，在社交方式和信息传递的方面存在明显的差异。目前的社交网络一般指基于互联网的社交网络。

社交网络源自网络社交。电子邮件是网络社交的起点，也是互联网早期的重要应用之一。应该说，互联网从已出现就具有了社交功能，但早期的社交网络由于提供的服务单一，常被作为互联网的一种应用或功能，其社交特性并没有引起人们足够的重视。

当提及社交网络时，人们的第一印象往往是社交站点和社交软件。以社交站点和社交软件为代表，社交网络的发展大致经历了五个阶段。

（1）早期概念化阶段

这一阶段以 SixDegrees 为代表。SixDegrees. com 是第一个被用户广泛认可的社交网络站点，始于 1997 年，其允许用户提供建立自己的简档和好友关系列表，并为用户建立好友连接和传递信息服务。SixDegrees 的服务模式吸引了上百万的用户，但由于没有可持续的商业计划，SixDegrees. com 站点于 2000 年被关闭。对于站点被关闭的原因，其创建者认为是因为 SixDegrees 是超越时代的。当时的社

交网络为用户提供的服务很少，用户也没有扩展社交网络和添加陌生用户的需求（Boya and Ellison，2010）。与 SixDegrees 同期的社交站点还包括：AsianAvenue、LiveJournal. AIM ，ICQ、BlackPlanet 和 MiGente 等。

（2）交友网站兴起阶段

这一阶段以 Friendster 为代表。Friendster 始于 2002 年，其免费服务方式吸引了大量的用户，随着注册用户数的增加，该网站的注册人数已经超过了软件可以负载的规模，结果导致网站运行速度变得很慢。Friendster 通过限制用户的功能来缓解这一问题，结果使竞争对手乘机取而代之。但 Friendster 并没有销声匿迹，只是它的用户基础转移到了海外，如今，Friendster 超过 70% 的访问流量都来自东南亚。据网站研究机构 Alexa 分析，Friendster 是菲律宾最受欢迎的网站，在马来西亚、新加坡和印度尼西亚则排名第二。该网站目前的注册用户人数是 2003 年的 10 倍，日网页浏览量是 2003 年的 20 倍，而且，网站还在快速发展。

（3）社交网络流行阶段

这一阶段以 MySpace 为代表。MySpace. com 成立于 2003 年 9 月，是目前全球第二大的社交网站。MySpace 的创意最初来自 Intermix 的 Chris DeWolfe 和 Tom Anderson。由于 Intermix 收购的 ResponseBase 很多成员来自 X-drive，因此，他们有很强的针对网络个人用户的开发和推广经验。MySpace 为全球用户提供了一个集交友、个人信息分享、即时通信等多种功能于一体的互动平台。经过 4 年的高速发展，现已拥有超过 2 亿的注册用户，并且正在以每天新增 23 万注册用户的速度继续增长。

（4）社交图阶段

社交图阶段以 Facebook 为代表。Facebook 是一个社交网络服务网站 ，于 2004 年 2 月上线。主要创始人为马克·扎克伯格。Facebook 是世界排名领先的照片分享站点，截至 2013 年 11 月每天上传约 3. 5 亿张照片。目前，随着 WhatsApp 和 Facebook 账号的互通，其用户数目将超过 10 亿人。Facebook 作为典型的 SNS（社交网络服务）网站发展案例，其成功与 SNS 顺应了当今网络潮流有关。

（5）云社交阶段

云社交（cloud social）的概念由利科国际咨询集团“著云台”分析师团队提出。云社交是一种物联网、云计算和移动互联网交互应用的虚拟社交应用模式，以建立著名的“资源分享关系图谱”为目的，进而开展网络社交。云社交的主要特征就是把大量的社会资源统一整合和评测，构成一个资源有效池向用户按需提供服务。参与分享的用户越多，能够创造的利用价值就越大。著云网也于 2012 年 1 月 5 日在美国上线。

综观社交网络的发展过程，可以看出，社交网络的发展一直遵循“低成本替代”原则。网络社交一直在降低人们社交的时间和物质成本，或者说是降低管理和传递信息的成本。与此同时，网络社交一直在努力通过不断丰富的手段和工具，替代传统社交来满足人类这种社会性动物的交流需求，并且正在按照从“增量性的娱乐”到“常量性的生活”这条轨迹不断接近基本需求。图 1-1 给出了社交站点创见或重建的时间进程表。

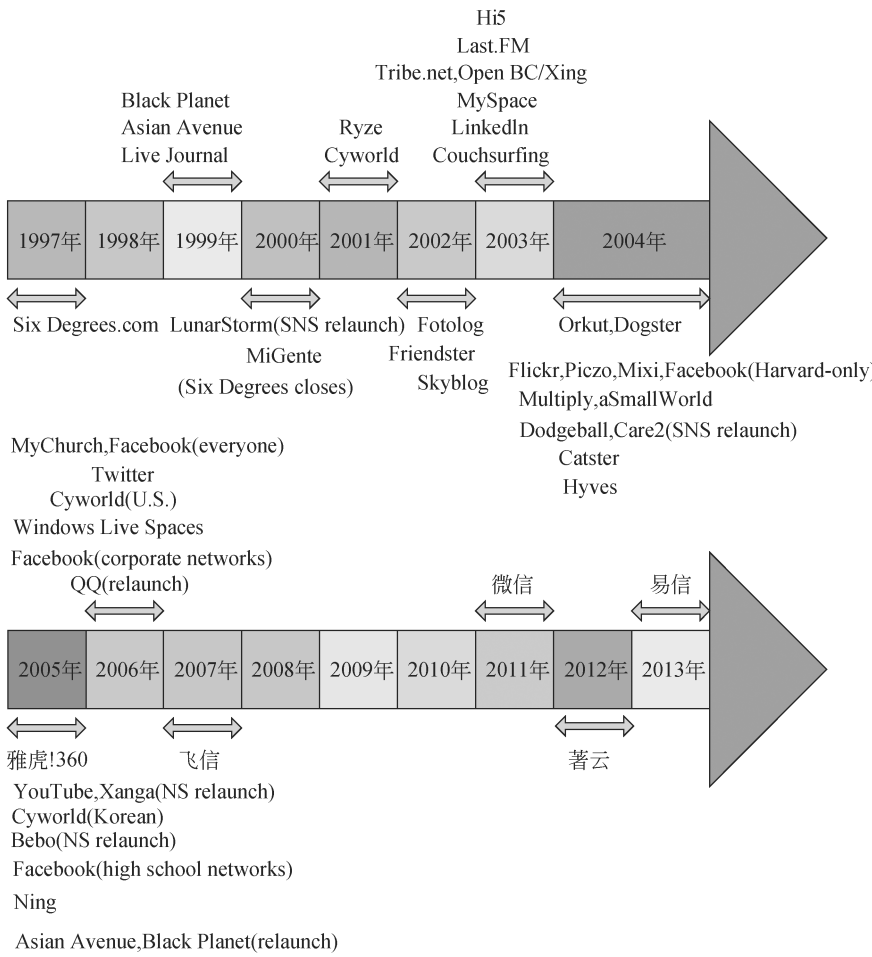

图 1-1　社交站点的初始发展历程

目前常用的社交软件工具、服务和网络平台有 Facebook、Twitter、YouTube、

Vine、Instagram、Google +、Linkedin、Yelp、腾讯 QQ（Tencent QQ）、微信（Wechat）等。根据全球统计平台 Statista. com 的统计结果显示，2016 年全世界范围内的社交媒体用户数量将达到 21. 3 亿人，是 2010 年 9. 7 亿人的 2. 2 倍，并且在 2018 年这一数字将达到 24. 4 亿人。其中，全球社交媒体市场排名第一的 Facebook 截至 2016 年 1 月，其注册用户数已达 10 亿人，每月的活跃用户数已达 15. 5 亿人。排名第三位的腾讯 QQ 用户达到 8. 6 亿人，仅排名第 20 位的 Linkedin（领英）也拥有 1 亿人稳定的用户群体。

1. 1. 3　社交媒体网络安全问题

社交媒体网络安全主要涉及三个方面的问题，社交用户安全、社交媒体安全和社交平台安全。其中，社交用户安全问题主要包括个人数据和隐私信息泄露，即用户账号被窃取、隐私信息被窥探，以及通过它们所造成的其他间接安全问题和损失；社交媒体安全主要涉及保护媒体内容不被非法篡改、使用和扩散，充分保障媒体信息版权拥有者的合法权益，防止非授权使用或恶意分享；社交平台安全则是确保社交用户和媒体内容赖以生存的软硬件基础设施、功能及服务的基础性安全。

据 Webroot[①] 调查显示，社交网站用户更容易遭遇财务信息丢失、身份信息被盗和恶意软件感染等安全威胁，而且其严重性可能超乎用户自己的想象。与用户个人隐私泄露不同，社交网络中信息的传播方式，尤其是社交媒体网络中媒体信息的传播方式可能使传统的网络安全策略失效，社交媒体网络中的安全问题需要进一步进行研究。

社交媒体网络安全的目标是在为用户提供便捷、低成本社交和媒体信息共享的前提下，保护用户个人数据、隐私信息不被泄露，保证拥有版权的媒体信息仅在合法的、许可的用户之间传播和使用，保障媒体社交基础平台安全、可信、可控。

社交媒体网络主要涉及的安全理论和关键技术包括社交用户信任评估、媒体传播安全机制、媒体访问控制模型、社交网络风险评估、社交网络安全策略博弈，以及多媒体数字版权管理和隐私保护等多个研究领域和方面。1. 2 节将对社交媒体网络安全技术进行介绍。

① http：//www. webroot. com。

1.2 社交媒体网络安全技术

1.2.1 社交用户信任评估

按照 Gambetta（1988）的定义，信任是一个实体对另一个实体在给定的时间范围内、特定的上下文条件下执行某个特定行为能力的主观概率水平度量，即对目标实体在待定行为发生之前的一个可信程度的度量。在社交媒体网络中，社交用户信任评估是对用户的可信任性进行评价，以保障媒体内容在信任用户之间进行传播，阻止媒体内容分享给信任等级不达标的用户。信任评估的核心是依据用户属性信息或行为特征评价用户的信任等级。通过建立信任模型，可以对各种网络环境下实体之间潜在的信任信息进行度量和评估，为网络系统提供一种相对柔性的安全度量机制（Rasmusson and Jansson，1996）。

在信任评估研究方面，1994 年，Marsh（1994）首次对信任程度和内容的概念进行区分，针对信任关系的主观特性，建立了信任度量的数学模型。针对开放网络系统，Blaze 等（1999）提出通过可信的第三方提供额外的安全信息进行安全决策，并首次提出信任管理的思想。在用户信任评估方面，针对不同的网络环境，研究人员建立了不同的信任评估模型。

1. 开放网络下的用户信任评估模型

Beth（1994）信任模型是一个有代表性的开放网络下的用户信任评估模型。该模型基于贝叶斯概率理论，将信任划分为直接信任和间接信任，利用信任相关经验评估信任关系，并根据推荐路径定义不同的信任合并规则，推导和合成不同推荐信任路径上的信任关系。在 Beth 信任模型中，推荐信任路径上信任关系的派生如图 1-2 所示，模型中仅利用肯定经验进行直接信任评估，采用算术平均的方法合成推荐信任，模型缺乏可靠性和完整性。

在开放系统用户信任评估研究方面，近年来大量研究致力于信任和信誉机制，以简化开放网络环境中的用户间复杂交易。其中，Agudo 等（2010）考虑不同上下文场景中的信任计算及语义，基于信任尺度提出了信任度量模型，并将模型应用到在线的信任和信誉管理系统中，该模型通过合并在不同上下文场景下收集到的信任的语义信息，为用户提供一个基于用户偏好的个性化信任推荐；唐文和陈钟（2003）基于模糊集合理论定义连接、合并算子和信任关系的推导规则来

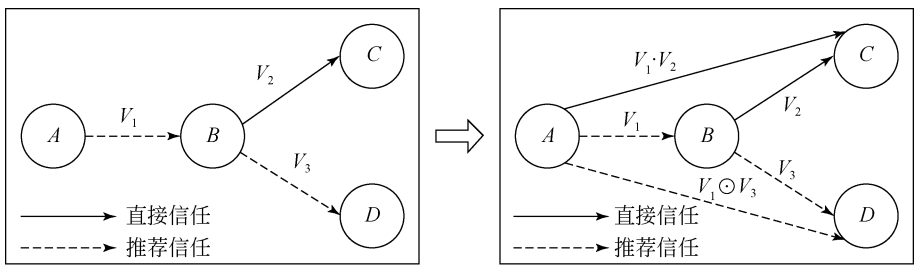

图 1-2　信任关系的派生

计算信任向量，为开放网络场景下的信任管理建立了可参考的框架模型；Bharadwaj 和 Shamri（2009）提出了包括信任和信誉两方面的模糊计算通用模型，将互惠和历史经验应用于信任建模，该模型将信任和信誉的概念组合到推荐系统中，与现有的影视推荐系统相比，提高了推荐信息的准确性，具有更好的评估效果，但该模型未考虑数据的稀疏性和模型的可扩展性问题。

2. 典型网络环境下的用户信任评估

信任机制的典型的网络场景有：多智能网络系统 MAS（muti-agent system）、P2P 对等网络、普适与网格计算场景、移动自组网和电子商务社区与虚拟企业等。针对不同的环境，可采用不同的信任评估模型。

（1）多智能网络系统

在多智能体系统中，利用合理的信任评估模型来计算代理之间的信任信息，并对目标代理的可相信程度进行评估，以保障系统安全高效的运行。FIRE 模型是针对多智能体系统的信任评估模型 Dong-Huynha 等（2006）。如图 1-3 所示，在 FIRE 模型中，评估实体首先询问与自己相邻的熟悉实体，这些实体再询问自己的邻居，反复下去直到找到和目标实体有直接交互经验的推荐实体，最终形成信任关系。

FIRE 模型主要贡献在于，将由直接交互历史经验信息得到的直接信任和实体在系统中所具有信誉值区别开来进行最终信任值的合成。但是，它是一个静态参数的系统，不能够适应网络环境的动态改变特性。

此外，在多智能体系统中，Kolaczek（2008）基于多智能体网络度量标准的描述，利用代理间的信任声明和网络当前的状态，对信任等级的主观逻辑一致性进行修正。Rettinger 等（2011）提出了无限隐藏信任模型，该模型利用 Dirichlet 过程混合模型和统计关系学习机制处理历史观察信息与上下文环境信息，来实现

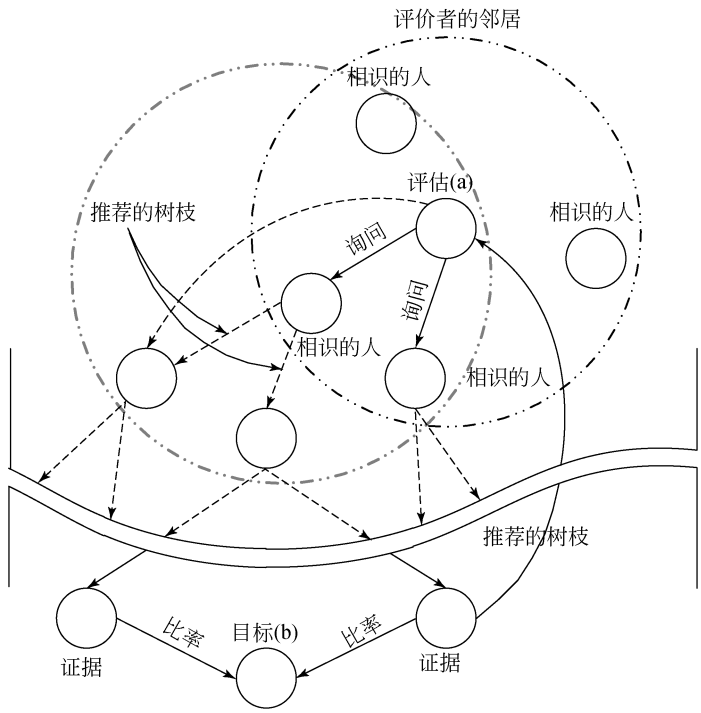

图 1-3　推荐信任形成过程

上下文敏感信任的识别与推导。

（2）P2P 对等网络

针对 P2P 对等网络的分布式特性，Kamvar 等（2003）提出了 EigenTrust 模型来防止 P2P 文件共享网络中恶意节点分发和传播虚假的文件，该模型运用了迭代求全局信任矩阵的思想，试图为每个节点提供一个全局信任值，但此算法的计算复杂度过于庞大。当 P2P 网络规模过大时，利用迭代方式求全局信任矩阵现实中难以实现，且模型中关于推荐信任的处理采用直接信任值相乘的形式获得，不能很好地体现信任的弱传递性。

为了识别并去除 P2P 网络环境中恶意节点和搭便车的节点，Mármol 等（2006）利用蚁群算法提出了 TACS 模型。节点经过蚂蚁算法选择出最优的服务提供商并取得相关的服务，如果对服务不满意，将会对路径上的节点进行惩罚处理，离恶意节点越近的节点受到的处罚会越重，即减少其费洛蒙含量。TACS 模型在信任评估算法的选择上比较新颖，算法能够动态地适应环境的变化，并且能够对恶意推荐路径上节点进行不同程度的惩罚，但模型中只考虑节点只提供一类

服务，不适合节点提供多种服务的网络环境。

Yu 等（2004）在开放分布式 P2P 网络中利用证据理论来表示和传递网络系统中代理之间的评价信息。通过组合本地证据（基于先前直接与被评估代理的历史合作经验）和其他相关的代理提供的证据来评估目标代理的可信程度。与 Yu 等（2004）的研究相关，朱峻茂等（2005）提出了 Grid 和 P2P 混合场景下基于证据理论的信任模型，但两者采用不同的数学形式构造证据。

（3）普适与网格计算场景

传统计算机网络中，安全通常依靠认证中心、证书目录或者一些预先安装的密钥和程序。然而，这些安全机制不再适用于普适与网格计算环境。Denko 等（2011）对普适计算环境中基于概率的信任模型和信任管理方案进行研究。将 β 分布作为贝叶斯推理的先验分布，利用定向信任管理方案对实体的可信程度进行评估。Denko 的信任模型对实体间的历史经验在数量上没有明确的限制，能够很好地处理和运用的推荐信任信息，评估实体能力的可信的程度，识别恶意实体，为实体选择可信的对象进行交互。当系统中具有一定比例的虚拟推荐者和虚假推荐信息时，信任评估算法能够随之进行动态的调整，仍然保持较好的鲁棒性。但是值得注意的是，模型对实体自身的计算或者储存能力要求较高，难以应用到实际环境当中。

网格计算场景中，信任信息具有模糊性并缺乏完备性。Song 等（2004）首先利用模糊推理计算节点的局部信任度然后利用逻辑理论合成全局信誉值，建立了模糊信任模型，解决模糊信任的问题。

（4）移动自组网

在移动自组网（MANETs）中，由于不存在基础设施，源节点与目标节点如果不在各自的传输信息范围内，那么他们之间的通信就必须依靠中间节点进行信息的转发，且节点的计算、存储、电力资源非常有限。网络中存在着自私节点，这些节点只利用其他节点有限的资源来为自己转发数据且拒绝为他们提供服务。为了提高网络的效率及去除网络中存在的自私节点，信任模型被引入用以寻找那些自私节点，鼓励节点之间的相互合作来提高网络的整体性能。

Luo 等（2009）在移动自组网络环境中利用模糊逻辑建立起 RFSTrust 模糊信任模型。RFStrust 中每个节点为其他节点维护一个数据包转发信息表，将信任分为直接信任和间接信任，并分别建立不同的模糊隶属函数来评估信任值。RFStrust 模型指出间接信任是由源节点和目标节点之间直接信任和推荐信任合成而来的，并根据推荐路径的长短将间接信任分为不同的等级，如图 1-4 所示，信任路径中有两次信任推荐的过程，形成了二级间接信任，然后将直接信任和不同

等级的间接信任合成为节点的全局信任值。

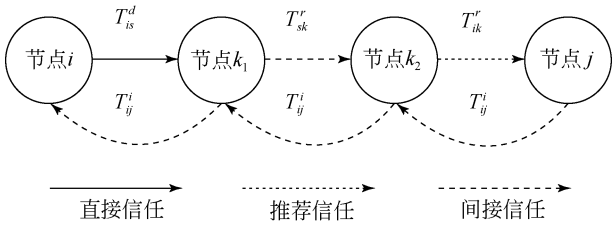

图 1-4　二级间接信任的合成

RFStrust 模型利用模糊逻辑能够较好地体现信任的主观性，提高移动自组网中节点之间通信的效率和降低节点的平均资源消耗量，虽然给出了多推荐路径上合成间接信任，但其并为对最长有效信任路径长度进行合理的限制，如果移动自组网的规模较大，那么考虑所有的推荐路径来求得全局信任值将会消耗网络自身大量的资源，实际应用中难以实现。

针对自组织网络的开放性、动态性，以及易遭受安全威胁等问题，Wang 等（2011）以信任的个性化为研究重点来捕获信任的主观性，提出了一个多信任路径的多维证据信任管理系统 Metrust（multi- dimensional evidence- based trust management system with multi-trusted paths），在任意复杂的信任关系图中利用多信任路径进行信任值的计算，并允许用户从大量信任参数中选择合适参数，对信任模型具有一定的启发作用。在该模型中同时提出两个开放性问题：①如何寻找一系列合适的信任路径形成信任关系图；②如何确保得到的信任关系图用于计算信任是合理的。

（5）电子商务社区与虚拟企业

在电子商务社区（e-commerce communities）平台下，信任信誉机制被应用到社区中来，用来评估参与实体的可信程度，引导实体与可信的实体进行交易，减少欺骗性交易的发生。然而在电子商务系统中关于信任的精确数字模型的建立和应用仍尚未令人满意。

该方面研究中，Manchala（1998）介绍了电子商务中可量化的信任概念，提出的基于模糊逻辑的信任模型，使用模糊逻辑来合并信任矩阵核实事务，还提出了可以在电子商务中保护信任信息所涉及的个人隐私的信任协议。为了较好地描述信任的主观模糊性，Jsang（1999）基于证据理论提出一个称为观点（opinion）的信任度量尺度，该度量尺度由信任、不信任和不确定三部分组成，且三部分之和为 1，利用观点来表达信任与推荐信任，并分别使用折扣操作符（discounting