

基于多维历史向量的 P2P 分布式信任评价模型

谭振华¹⁾ 王兴伟²⁾ 程 维¹⁾ 常桂然³⁾ 朱志良¹⁾

¹⁾(东北大学软件学院 沈阳 110819)

²⁾(东北大学信息科学与工程学院 沈阳 110819)

³⁾(东北大学计算中心 沈阳 110819)

摘 要 P2P 网络的开放性、匿名性和对等性使 P2P 网络得到了广泛的应用,同时也滋生了大量的恶意节点.传统的中心化认证的信任模式无法满足 P2P 网络的分布式要求,分布式信任机制是 P2P 网络得以进一步发展的关键技术.文中基于社会网络的一些基本理论,提出了一种新的 P2P 分布式信任模型 MDHTrust.设计了多维交易历史向量模型及其分布式存储结构,并由此定义了局部和全局的多维因子计算方法,包括非线性时间因子、交易额、交易频率以及成功率;构建了局部信任度、全局信任度的计算方法,最终通过计算节点间交易历史的相关性,形成了具有相关性因子的全局信任度,即相关性信任度.仿真结果表明,MDHTrust 具有较快的信任收敛速度,是一种有效的信任模型.

关键词 对等网络;信任模型;P2P 安全;社群网络;网络安全;多维历史向量

中图法分类号 TP393 **DOI 号:** 10.3724/SP.J.1016.2010.01725

A Distributed Trust Model for Peer-to-Peer Networks Based on Multi-Dimension-History Vector

TAN Zhen-Hua¹⁾ WANG Xing-Wei²⁾ CHENG Wei¹⁾ CHANG Gui-Ran³⁾ ZHU Zhi-Liang¹⁾

¹⁾(Software College, Northeastern University, Shenyang 110819)

²⁾(College of Information Science and Engineering, Northeastern University, Shenyang 110819)

³⁾(Computing Center, Northeastern University, Shenyang 110819)

Abstract The peer-to-peer network has been popular in many distributed applications because of its openness, anonymous and reciprocity. Meanwhile, lots of malicious nodes appear in kinds of P2P systems, disturbing the trust of the P2P network. Since the traditional centralized trust mode can not satisfy the decentralized P2P network, distributed trust mechanism comes to be the key technology for the future P2P network. Based on some social network principles, this paper presents a new distributed trust model named MDHTrust (Multi-Dimension-History Trust Model). The authors designed a communication multi-dimension history vector and its distributed storage structure, and defined the computing methods for multi-factors, including nonlinear time factor, transaction currency amount factor, and frequency factor and success rate, both in local and global environments. Then, the authors constructed local trust and global trust with mathematical representations and computing methods based on user appraisals. On the basis of these jobs, a correlative trust model is presented by calculating nodes' correlative similarity multiplied with global trust degree. Simulations and analysis proved the MDHTrust has faster convergence speed comparing with some other trust models, and it is a feasible and effective model.

Keywords peer-to-peer network; trust model; P2P security; social network; network security; multi-dimension-history vector

收稿日期:2010-04-26;最终修改稿收到日期:2010-08-05.本课题得到国家自然科学基金(61070162,71071028,60802023,70931001)、高等学校博士学科点专项科研基金(20070145017)、中央高校基本科研业务费专项资金(N090504003,N090504006)资助.谭振华,男,1980年生,博士,讲师,主要研究方向为分布式网络安全、计算机系统结构等. E-mail: tanzh@mail.neu.edu.cn. 王兴伟,男,1968年生,博士,教授,博士生导师,主要研究领域为下一代互联网、移动 Internet 和 IP/DWDM 光 Internet 等.程 维,男,1970年生,博士研究生,讲师,主要研究方向为可信计算.常桂然,男,1946年生,博士,教授,博士生导师,主要研究领域为计算机系统结构、网络安全等.朱志良,男,1962年生,博士,教授,博士生导师,主要研究领域为复杂网络、混沌分形等.

1 引 言

P2P 覆盖网络^[1]是在现有 Internet 之上构建的一个完全位于应用层的对等网络系统,节点间具有很强的对等性、匿名性、开放性以及松耦合性等特点,近年来得到了快速的发展.然而,正是因为这些特点使得恶意节点能轻松进入 P2P 网络并发起攻击,提供欺骗服务,滥用网络资源等^[2-4],破坏 P2P 网络的信任系统.构建符合 P2P 网络特性的信任机制已经成为研究的热点.文献[5]明确指出,P2P 网络是下一代互联网体系结构的重要组成部分,在 P2P 网络中构建合理科学的信任模型是 P2P 网络继续发展的重要保障.

文献[6-7]最早提出个体 A 对个体 B 的信任是指个体 A 期望个体 B 为 A 服务(即执行 A 的利益所依赖的动作)的主观可能性.信任关系被认为是比授权关系更加本质的安全关系,人类社会之所以能够平稳健康地运行,很大程度上得益于个人、团体和组织之间的信任关系^[8].而在传统的网络环境或电子商务系统中,信任关系的建立依赖于可以信赖的第三方,但这种方式并不适合于 P2P 网络^[9].

因此,国内外的学者提出了很多种机制来解决 P2P 对等网络的可信安全问题^[9-20],在 P2P 信任机制的发展上做了很大的贡献.文献[10]最早提出了一种基于全局声誉的信任机制,节点可以在任意时刻抱怨其它节点,通过收集对某节点的所有抱怨信息来确定该节点的全局声誉,但该信任机制过于片面,并没有考虑影响信任度量的其它相关因素.文献[11]提出了一种信任模型 XRep,采用了节点 IP 检测的方法以识别共谋团体,认为属于同一 IP 聚类的节点集合为同一个共谋团体,尽管具有很好的收敛性,但是这种具有类似 IP 的节点为同一团体的节点的假定是不合理的.文献[12]引入了撒谎度的概念来提高信任估计的准确性,但没有提供具体的计算方法.文献[13]从经济学理论中获得灵感,对历史评价进行衰减计算,使得声誉稳定收敛.文献[14]提出了 EigenRep,任意节点的全局信任决定于与之发生过交易行为的其它节点对其的局部看法以及所有这些节点的局部信誉全局聚合,通过邻居节点间相互满意度的迭代来获取节点的可信度,但通信开销代价大,不适合全分布式 P2P 网络.文献[9]在迭代收敛性和模型安全性方面对 EigenRep 进行了改进,但改进后的模型仍然存在效率问题,且其安全性是通

过引入额外的认证机制和惩罚措施实现的.文献[15]基于模糊逻辑推理规则来计算节点的全局信任度,这种方法具有较高的恶意节点检测率,但其对抗的恶意行为较为简单,且不能对抗各种针对信任机制的攻击.文献[16]提出了一种基于局部声誉的信任机制,考虑了交易时间对信任度的影响,但没有给出交易时间影响因子的确定方法.文献[17]给出的是一种比较全面的信任机制,文中提出:一个节点同时具有局部声誉和全局声誉.局部声誉中引入了交易时刻和交易金额等因素,而全局声誉的计算引入了参与评价的节点个数、评价节点的交易总量、评价节点所给评价的可信度等信任因素,但在信任计算中信任值偏向高额交易,且其时间因子按时间进行线性的等比划分,虽然显示了当前时间的重要性,但时间因子的增长幅度是线性的,且没有明确的机制来防御团队恶意行为.文献[18]提出了一种基于节点反馈可信度的分布式 P2P 全局信任模型 FC-Trust,节点在提供反馈时,除了考虑反馈节点本身的全局信任度以外,还考虑该节点的反馈可信度,但并未考虑数据安全存储的分布式方式.文献[19]提出了一种 P2P 环境下的基于节点行为相似度的共谋团体识别模型,通过分析节点之间的行为相似度来排除共谋团体的干扰,但在通信模型上完全采用 Hash 算法来选择节点进行评分的存储,每个节点的评分数据都记录在另一个节点上,使得在计算信任的过程中通信开销较大.文献[20]提出了相关性信任模型 NBRTTrust,通过对节点通信历史数据进行相关性计算得到节点间的相关性因子,用以避免团队恶意节点的攻击.但该模型中的通信历史模型只考虑了通信的成败,并没有考虑其它 QoS 参数.

以上模型有的是局部信任模型,如文献[11-12, 16],模型建立在两个节点之间的历史数据基础上;有的是全局信任模型,如文献[10, 14-15],模型建立在节点的所有历史数据基础上;还有一些是兼有两者的模型,如文献[17, 20],这些模型既进行局部计算又进行全局计算;还有一些模型考虑了信任中的一些个性化因素.文献[21]对当前的 P2P 信任模型进行了很好的总结和归类.可以看出,这些模型都具有自身的特点,但都有需要完善的地方,比方说都没有既考虑信任因素的多维化,又考虑信任数据的相关性,还考虑信任数据的分布式存储方式等.本文在这些模型工作的启发下,基于社群网络的一些基本理论,提出基于多维历史向量的 P2P 分布式信任模型 MDHTTrust(Multi-Dimension-History Trust Model).

人们发现,在社群网络中,评价个体的信任往往不是从单一的角度出发的;每个个体的信任数据除了存储在个体本身之外,往往还分布式地存储在与该个体相关的其它个体上;个体交往最频繁的对象往往与该个体处于同一团体中,因此,在 MDHTrust 模型中:(1) 为避免评价因子的单一化,设计了多维历史向量以及该向量的分布式存储结构;(2) 设计了综合多维因子的节点局部信任模型及其存储方式;(3) 针对单个恶意节点的欺骗行为,在局部信任基础上设计了全局信任模型的计算方法;(4) 针对团队恶意节点的欺骗行为,设计了节点之间的相关性因子,形成了新的相关性信任模型.根据模型算法,该模型可以有效地屏蔽单个恶意节点和团队恶意节点.仿真和分析表明了该模型的可行性和有效性.

本文第 2 节介绍多维历史向量的定义;第 3 节~第 5 节对 MDHTrust 中的局部信任模型、全局信任模型、相关性信任模型的构建过程进行论述;第 6 节进行仿真和实例说明;最后第 7 节对全文进行总结和展望.

2 多维历史向量

P2P 网络可以抽象为一个由点集 V 和边集 E 组成的图 $G = (V, E)$, MDHTrust (Multi-Dimension-History Trust model) 信任模型基于节点集 V 的通信历史向量建模.本文以节点的交易通信历史中与信任相关的有交易频率、成功计数、失败计数、交易额度、交易时间、用户满意度反馈等 6 个维度为出发点对通信历史向量进行建模,形式化定义该向量为 $MDHVector = \langle \varphi, s, f, \omega, t, \phi \rangle$,其中 φ 是交易频率, s 为交易成功计数, f 为交易失败计数, ω 为交易额度, t 为交易时间戳, ϕ 为用户评价反馈.本节将对该向量的基本数据结构进行定义,并阐述其分布式存储方式.

2.1 基本数据结构定义

定义 1. (1) t_m^{ij} 表示节点 i 与 j 进行第 m 次交易的时间戳.特别地, t_{nov}^{ij} 表示 i 与 j 的最近一次信任计算的时间戳;(2) t_m^i 表示 i 进行总的第 m 次交易的时间戳.特别地, t_0^i 表示节点 i 开始信任计算的初始时刻, t_{nov}^i 表示 i 最近一次的信任计算时间戳.

定义 2. $\omega_m^{ij} = \omega(i, j, t_m^{ij})$ 表示节点 i 与 j 在 t_m^{ij} 时刻进行交易的额度.

定义 3. (1) I_{all} 表示 P2P 网络的所有节点集合;(2) I_i 表示与节点 i 进行过交易的节点集合(I_i 可

以存储在节点 i 上,也可以从交易历史节点获取);(3) $I_{ij} = I_i \cap I_j$.

定义 4. s_{ij} 表示节点 i 与 j 交易成功的次数, f_{ij} 表示节点 i 与 j 交易失败的次数.具体的交易成功与失败的定义可以根据具体的网络进行定义.

定义 5. (1) $\lambda_{ij}^m = (s_{ij} / (s_{ij} + f_{ij}))_{[t_0^i \sim t_m^i]}$,取值范围 $[0, 1]$,表示到时刻 m 为止,节点 i 与 j 交易成功的比率.(2) $\lambda_i^m = \left(\sum_{c \in I_i} s_{ci} / \sum_{c \in I_i} (s_{ci} + f_{ci}) \right)_{[t_0^i \sim t_m^i]}$,取值范围 $[0, 1]$,表示到时刻 m 为止,节点 i 的总的交易成功率.

定义 6. $\phi_m^{ij} = \phi(i, j, t_m^{ij}) \in [-1, 1]$,表示节点 i 对节点 j 所提供的第 m 次交易的评价反馈,其中 t_m^{ij} 表示评价的时间戳.

定义 7. (1) $\varphi_m^{ij} = \varphi(i, t_m^{ij}) = \frac{(s_{ij} + f_{ij})}{(t_m^{ij} - t_0^i)}$ 表示节点 i 与 j 在 t_m^{ij} 时刻的交易频率,若 $t_m^{ij} = t_{nov}^{ij}$ 则表示到目前为止 i 与 j 的交易频率.(2) $\varphi_m^i = \varphi(i, t_m^i) = \frac{1}{(t_m^i - t_0^i)} \sum_{j \in I_i} (s_{ij} + f_{ij})$ 表示从 t_0^i 时刻到当前的 t_m^i 为止,节点 i 与其它节点进行的总的交易频率.

2.2 分布式存储结构

本小节定义了节点之间交易历史数据的分布式存储结构,每条历史数据被分布式地存储在了各个通信节点中.

定义 8. $TimeDBList(i, j) = \{ \langle m, \omega_m^{ij}, \phi_m^{ij}, t_m^{ij} \rangle \mid m \in \text{自然数} \wedge m \in [0, max] \}$ 用来存储节点 i 与 j 进行通信过程中与时间相关的向量列表,定义中 m 是 i 与 j 的交易顺序号, max 表示当前 i 与 j 的最大的交易顺序号,称为时间数据列表(注意:在本文的部分章节 $TimeDBList$ 被简称为 TDL).

不难看出, $TimeDBList(i, j)$ 是四元向量 $\langle m, \omega_m^{ij}, \phi_m^{ij}, t_m^{ij} \rangle$ 的集合,记录了节点 i 与 j 的交易序列、额度、评价和时间戳,其存储在节点 i 中.之所以选择存储在节点 i 中而不是节点 j 中,是为了避免节点 j 对评价等信息的恶意修改.图 1 是一个简单的 $TimeDBList(i, j)$ 的例子.

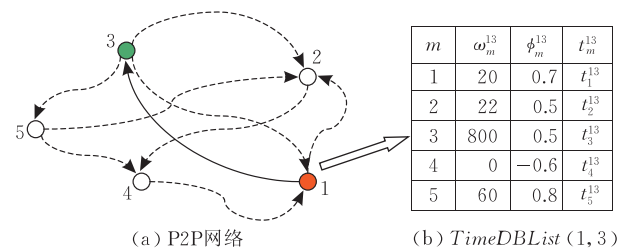


图 1 时间数据列表例: $TimeDBList(1, 3)$

图 1 中,节点 1 与多个节点进行了交易,示例给出了节点 1 与节点 3 进行交易的 $TimeDBList(1,3)$. 可以看出,到目前为止,节点 1 共计 5 次选择节点 3 进行交易,每次交易额度都不一样,其中最高的交易额度为 800;对每次交易,节点 1 给节点 3 的评价也不一样,最高为 0.8,最低为 -0.6;具体交易过程中的时间戳也同样记录了下来.事实上,每个节点中都存储着多个 $TimeDBList$,如上图中的节点 1 同时还选择与节点 2 进行了交易,那么节点 1 中除了有 $TimeDBList(1,3)$ 外,还存储了 $TimeDBList(1,2)$.

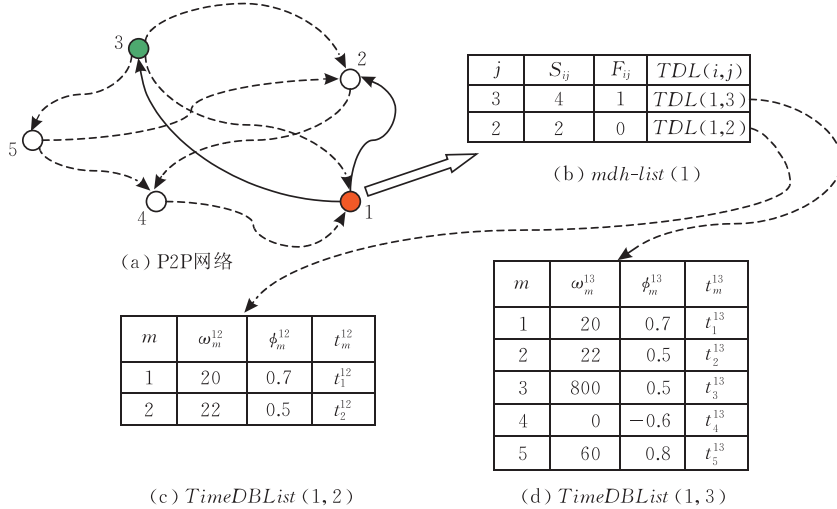


图 2 多维历史数据列表示例: $mdh-list(1)$

从图 2 中可以看出,节点 1 曾经与节点 {2,3} 进行过交易,其中,与节点 3 交易成功 4 次失败 1 次;与节点 2 交易 2 次,全部成功.对应的 $TimeDBList$ 被链式扩展.

3 局部信任度

局部信任度(local trust)是两个节点之间根据交易历史数据得出的一个节点对另一个节点的信任期望.本文所设计的局部信任度与时间因素、评价因素、交易额度以及交易频率相关.从社群网络的角度分析,而交易频率越高,交易额度越大,交易的评价越好,则节点间的信任关系越稳定.同时,交易时间距离计算声誉值的时间越远,交易评价对声誉值的影响应该越小.

下面,本文首先对局部信任度的时间因子、交易额度因子及交易频率因子的计算方法进行讨论和定义.

3.1 时间因子 tf

从时间对信任的影响程度来说,距离当前时刻

定义 9. 多维历史数据列表 $mdh-list(i)$ 用来存储当前节点 i 的通信历史, $mdh-list(i) = \{ \langle j, S_{ij}, F_{ij}, TimeDBList(i,j) \rangle | j \in I_i \}$.

$mdh-list(i)$ 是四元向量 $\langle j, S_{ij}, F_{ij}, TimeDBList(i,j) \rangle$ 的集合.每个四元向量记录了某个节点 j 与当前节点 i 的通信历史中交易成功和失败的次数以及节点 i 与对应的节点 j 之间的 $TimeDBList(i,j)$.每个节点维护一个 $mdh-list$ 列表,所有节点的 $mdh-list$ 列表综合起来则为一个完整的全网通信历史情况列表.图 2 是一个 $mdh-list(1)$ 的简单示例.

越近的交易的评价所占的权重越大,距离当前时刻越远的交易的评价所占的权重则越小^[16-17],即 $0 \leq tf_1 < tf_2 < \dots < tf_{now} \leq 1$. 在本文中,为体现当前时刻所进行的交易的重要性,令 $tf_{now} = 1$;一般时刻的 tf_m 采用直角三角形面积的方式进行计算,这种方式避免了简单的线性递增,进一步体现了“近”的重要性,具体计算方法如图 3 所示.令直角三角形 Δ_{now} 的面积为 1,其底是总的时间差“ $\delta(t) = t_{now}^{ij} - t_0^i$ ”,则其高为“ $2/\delta(t)$ ”;第 m 次交易的时间因子为 Δ_{now} 的一部分 Δ_m (阴影部分),显然,在面积上 $S_{\Delta_m} < S_{\Delta_{now}}$.

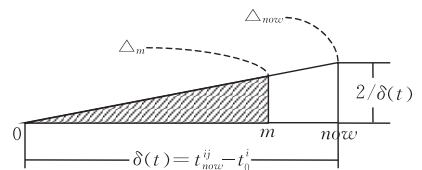


图 3 时间因子计算方法

定义 10. 时间因子 tf_m^{ij} 是节点 i 与 j 进行第 m 交易后所产生的信任数据在整个 i 与 j 交易的局部历史中的权重,

$$tf_m^{ij} = s(\Delta_m) = \frac{1}{2} \cdot (t_m^{ij} - t_0^i) \cdot \left(\frac{2 \cdot (t_m^{ij} - t_0^i)}{\delta(t)^2} \right) \\ = (t_m^{ij} - t_0^i / t_{now}^{ij} - t_0^i)^2 \quad (1)$$

显然, $0 \leq tf_1^{ij} < tf_2^{ij} < \dots < tf_{now}^{ij} = 1$.

3.2 交易额度因子 ωf

在现实的社会网络中,交易额度越大,一般越能引发人们的信任.同时,对于总额相同的交易,一般有三种情况,如图 4 所示.

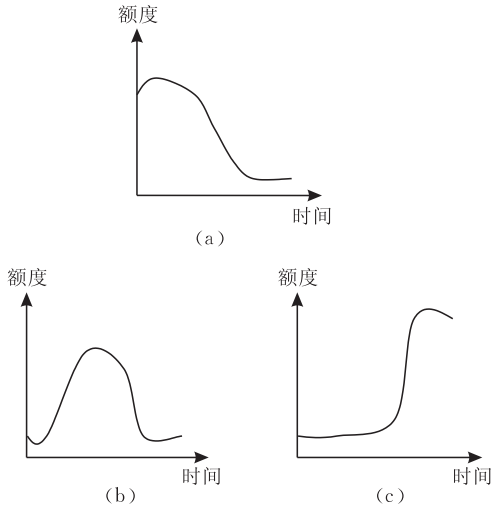


图 4 3 种大额交易情况

图 4(a)中,大额交易峰值出现在早期,(b)是出现在中间时段,而(c)则出现在近期.单从额度来讲,图中的 3 种情况下到目前时刻为止的交易总额一定,但额度的影响是不一样的,(c)的峰值被认为对当前状况最为有效,而离现在越久远的额度越不重要.因此,交易额度受时间因子的影响.

同时约定交易的额度的价值与到目前时刻为止的成功率相关,如交易额为 100,但 λ_{ij}^m 只有 0.1,则其价值额度为 $100 \times 0.1 = 10$.这种约定可以刺激节点努力提高自身的交易成功率.

定义 11. 交易价值额度 v_m^{ij} 是 i 与 j 的第 m 次交易额度与 m 时刻的时间因子、交易成功率的乘积,是调整后的交易额度,

$$v_m^{ij} = \omega_m^{ij} \cdot t_m^{ij} \cdot \lambda_{ij}^m \quad (2)$$

定义 12. 交易额度因子 ωf_m^{ij} 是第 m 次交易的价值额度在总价值额度中所占的比重,取值范围为 $[0, 1]$,令 now 为当前交易序号(以下同),则

$$\omega f_m^{ij} = v_m^{ij} / \sum_{m \in [1, now]} v_m^{ij} \quad (3)$$

3.3 交易频率因子 ϕf

在社会网络中,人们普遍认为,通信频率越高,对其信任的正面影响越大.所以,本文在局部信任上

引入交易频率因子,本文在定义 7 中已经给出了频率的计算方法,但频率的值不是 $[0, 1]$ 的范围,因此对定义 7 的频率进行归 1 化处理,形成频率因子.

定义 13. 交易频率因子 ϕf_m^{ij} 表示 i 与 j 第 m 次交易后的频率在总的交易频率中所占的比重.

$$\phi f_m^{ij} = \phi_m^{ij} / \sum_{m \in [1, now]} \phi_m^{ij} \quad (4)$$

3.4 局部信任度及其存储

根据以上分析与因子的计算,我们定义局部信任度如下.

定义 14. 节点间局部信任度 TL_{ij} ,表示单独从节点 i 的角度看节点 j 的信任度,综合了评价、时间因子、额度因子、频率因子,取值范围 $[-1, 1]$.规定在初始时刻(t_0^i)令 max 为当前 i 与 j 通信的总次数,则

$$TL_{ij} = \begin{cases} 0, & \text{当 } t_0^i \text{ 初始时刻时} \\ \sum_{m \in [1, max]} (t f_m^{ij} \cdot \omega f_m^{ij} \cdot \phi f_m^{ij} \cdot \phi_m^{ij}) / max, & \text{其它情况} \end{cases} \quad (5)$$

不难发现,如果包含交易额度因子中的成功率因子,则上面关于 TL_{ij} 的计算公式中是 5 个小数的乘积,使得最终的计算结果很小;为了方便计算和结果显示,在不丢失数据本质(保留计算逻辑与顺序)且不影响数据结果的排序的情况下,采用开五次方的形式对 TL_{ij} 的最终结果进行数据放大,对公式进行修订.修订后的 TL_{ij} 计算公式如下:

$$TL_{ij} = \begin{cases} 0, & \text{当在 } t_0^i \text{ 初始时刻时} \\ \sqrt[5]{\sum_{m \in [1, max]} (t f_m^{ij} \cdot \omega f_m^{ij} \cdot \phi f_m^{ij} \cdot \phi_m^{ij}) / max}, & \text{其它情况} \end{cases} \quad (6)$$

定义 15. 列表 $TL-list$ 用来存储节点之间的局部信任度, $TL-list(i) = \{ \langle j, TL_{ij}, t_{last}^{ij} \rangle | j \in I_i \}$.

由定义 15 可知, $TL-list(i)$ 是一个三元组的集合,这个三元组包括与 i 交易的节点、 i 节点对这些节点进行的局部信任度评价以及进行该评价的时间戳 t_{last}^{ij} .在 MDHTrust 模型中,使用 $TL-list$ 列表存储所有节点间的局部信任度信息,并且按照当前的 P2P 网络规则分布式地存储在各个节点之上,每个节点负责与该节点相关的 $TL-list$ 的信息的存储,而所有节点的 $TL-list$ 信息在当前 P2P 中形成一个大的分布式的 $TL-list$ 网.如图 5 是一个简单的 $TL-list$ 示例.

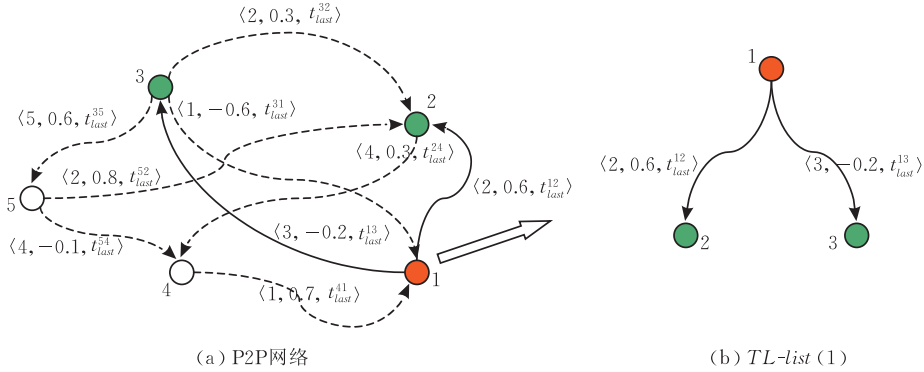


图 5 节点间局部信任度的存储 TL-list 示例

在该 TL-list 网中, $TL-list(1) = \{ \langle 2, 0.6, t_{last}^{12} \rangle, \langle 3, -0.2, t_{last}^{13} \rangle \}$, 表示节点 1 给节点 2 的局部信任度评价为 0.6, 节点 1 给节点 3 的局部信任度为 -0.2 . 每当节点 i 选择与 j 进行一次交易时, 都会对信任度进行计算, 这个时候则对 $TL-list(i)$ 进行更新.

4 全局信任度

很显然, 节点的局部信任度具有一定的局限性, 因为局部信任度是根据两个节点之间的历史进行计算的, 这样无法避免恶意修改或恶意评价. 因此本文定义了节点的全局信任度(Global Trust), 且根据局部信任度列表 TL-list 可以求得某个节点的全局信任度. 节点 i 的全局信任度 TG_i 是所有与 i 进行过交易的节点 I_i 对 i 的局部评价的综合值, 如图 6 所示.

4.1 全局信任度因子

为了计算全局信任度, 本文对全局信任度的影响因素进行考虑. 由于计算某节点 i 的全局信任度的时候, 将获取集合 I_i 对节点 i 的局部信任度, 因此, 涉及到节点 i 的面向 I_i 的总的成功率、总的频率以及全局时间因子, 其中前两项因子已经在定义 5(2))、

定义 7(2)中分别进行了定义, 这里我们对全局时间因子进行讨论.

对于节点 i , 集合 I_i 中的节点对 i 的局部评价都有最终的时间戳记录 (t_{last}^{ji}), 与 3.1 节讨论的依据类似, 本文认为距离当前时刻越近的局部信任评价所占的权重应该越大. 因此, 仍然采用 3.1 节所讨论的直角三角形面积法对全局时间因子 gf 进行定义.

为方便计算, 将 I_i 节点集对节点 i 的局部信任评价按时间戳从小到大顺序排序, 形成排序后的局部信任度评价列表: $\rho(i) = \{ \langle j, TL_{ji}, t_j^i \rangle \mid j \in I_i \}$, 其中 t_j^i 是 I_i 中对应的某个 j 节点对节点 i 最后进行局部信任计算时间戳 t_{last}^{ji} . 令 $\rho(i)$ 排序中最近一次 (最大索引项) 的局部评价的全局时间 (记为 t_{max}^i) 因子为 1, 即 $gf_{max}^i = 1$ ($num = max$ 为最大编号), 具体的计算方法与 3.1 节所述类同, 得到如下定义.

定义 16. 全局时间因子 gf_j^i 表示 I_i 中的节点 j 对节点 i 的局部信任度评价在 i 的全局评价中所占的时间比重,

$$gf_j^i = (t_j^i - t_0^i / t_{max}^i - t_0^i)^2 \tag{7}$$

显然, gf_j^i 中离现在越近的所占的时间比重越大, 取值范围在 $[0, 1]$.

4.2 全局信任度

为了能够表示节点 i 的信任度综合值, 必须采

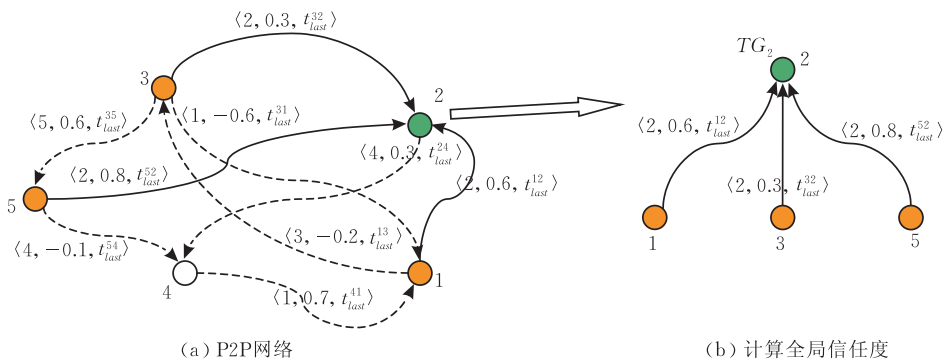


图 6 全局信任度的分布式计算

用一些策略来进行全局信任度的计算,这些策略需要符合两个条件:(1)计算的节点顺序不影响计算结果,即计算具有对等公平性;(2)由于节点间局部信任度的范围在 $[-1, 1]$,因此,节点的全局信任度值也应该在 $[-1, 1]$.在这两个条件下,本文结合全局成功率因子 λ_i^{now} 、全局频率因子 φ_{now}^i 、全局时间戳因子 gf_j^i 采用了平均值法来对局部信任度进行均值以计算全局信任度.因为涉及到3个0~1的小数因子乘积运算,为了方便结果的读取与分析,与局部信任处理方法类似,本文也对全局信任值进行开三次方放大,最终公式如下:

$$TG_i = \sqrt[3]{\left(\sum_{c \in I_i} (gf_c^i \cdot TL_{ci}) / \text{Count}(I_i)\right) \cdot \lambda_i^{now} \cdot \varphi_{now}^i} \quad (8)$$

其中, $\text{Count}(I_i)$ 为 I_i 的节点数量.可以看出,全局信任度的取值范围在 $[-1, 1]$.值得说明的是,还可以选取其它的方法来求节点的全局信任度,如迭代加权和法等,只要满足计算要求即可.

5 相关性信任度

从理论上来看,全局信任模型虽然具有全局性,能够避免单个节点的恶意行为;然而,全局信任模型无法防御团队恶意节点的协同作案.倘若恶意节点之间互相抬高评价,那么,对于一个可信节点来说,无法直接从全局信任度的值来区分恶意与否了.因此,本文构建了相关性信任模型.

5.1 相关性因子

节点之间的评价的相关度用来描述节点之间相关联的程度,这种关联程度主要是通过公共的第三方节点群体对这两个节点之间评价的相关程度来体现.这和社会网络的情况是类似的,我们在判断某个节点是否和自己是同一类人的时候,往往是通过双方共同的朋友来进行判断的.比方说, A 与 B 并不认识,但都拥有共同的朋友 $\{C, D\}$,则 A 可以通过朋友 $\{C, D\}$ 来判断 B 与自己的相关程度,如果 $\{C, D\}$ 与 A, B 的交往历史都很接近,则 A, B 之间具有较大的相关性.

邓爱林在文献[22]中总结了度量相关性的多种方法,主要包括3种方法:余弦相似性、相关相似性以及修正的余弦相似性.这些方法都属于统计学的范畴,主要是通过用户对某些项目的评价来度量相关性.根据其他用户的观点产生对目标节点的推荐列表,是基于这样一个假设:如果用户对一些项目的

评分比较相似,则他们对其它项目的评分也比较相似.

由于根据节点之间的通信历史可以得出节点之间的局部信任度 $TL-list$ 网,则可以根据 $TL-list$ 网来计算节点之间的相关程度.

度量相关性一般用矩阵来表述基本评价参数,为了和节点之间的相关性联系起来,由 $TL-list$ 网可以转换形成局部信任度矩阵 $\mathbf{R}(n \times n)$.

定义 17. 信任度矩阵 \mathbf{R} 是由节点间局部信任度 $TL-list$ 网转换而来的 $n \times n$ 的矩阵, n 是节点总数.矩阵中的每一个元素 r_{ij} 的定义如下:

$$r_{ij} = \begin{cases} TL_{ij}, & \text{若 } i, j \text{ 存在交易历史} \\ \epsilon, & \text{若 } i, j \text{ 不存在交易历史} \\ 1, & \text{条件是 } i=j \end{cases} \quad (9)$$

如果节点 i 对 j 没有交易历史则对应的 r_{ij} 值则为 ϵ (设定的非零小值),而节点对自己的评价则为1.信任度矩阵 \mathbf{R} 为

$$\mathbf{R} = (r_{ij})_{n \times n} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix}.$$

在此基础上,则可以使用相关性计算方法计算节点 i 与节点 j 之间的相关性.本文采用相关相似性来计算节点 i 和节点 j 之间的相关程度,典型的计算相关相似性的方法是PCC(Pearson Correlation Coefficient)方法^[23].已知 I_{ij} 是节点 i 和节点 j 共同访问过的节点,或者从 $TL-list$ 网的角度说是节点 i 和节点 j 共同作出过信任评价的节点.根据PCC公式, \bar{R}_i 表示节点 i 对 I_{ij} 中所有节点的平均评价, $\bar{R}_i = \sum_{c \in I_{ij}} TL_{ic} / n$,则得到节点间相关性的计算公式:

$$\text{sim}(i, j) = \frac{\sum_{c \in I_{ij}} (TL_{ic} - \bar{R}_i)(TL_{jc} - \bar{R}_j)}{\sqrt{\sum_{c \in I_{ij}} (TL_{ic} - \bar{R}_i)^2} \sqrt{\sum_{c \in I_{ij}} (TL_{jc} - \bar{R}_j)^2}} \quad (10)$$

$\text{sim}(i, j)$ 取值范围在 $[0, 1]$ 之间. $\text{sim}(i, j)$ 的值越大,表示节点 i 和 j 之间的相关性越大.但是,上述公式中可能出现分母为零的问题,这个时候认定为节点之间无法计算相关性;同时,节点的相关度 $\text{sim}(i, j)$ 取决于节点 i 和节点 j 的公共节点集合 I_{ij} .如果 I_{ij} 的节点数目很少甚至可能没有的时候,节点间的相关性则很难计算出来.因此,本文规定,当 I_{ij} 的节点数量 $\text{Count}(I_{ij})$ 少于所要求的最少节点数量 τ (可以根据实际情况按比例或数量设计)时,

则认为 $sim(i, j) = 0$.

故,公式修订为

$sim(i, j) =$

$$\left\{ \begin{array}{l} 0, \sqrt{\sum_{c \in I_{ij}} (TL_{ic} - \bar{R}_i)^2} \sqrt{\sum_{c \in I_{ij}} (TL_{jc} - \bar{R}_j)^2} = 0, \\ \text{或 } Count(I_{ij}) < \tau; \\ \left| \frac{\sum_{c \in I_{ij}} (TL_{ic} - \bar{R}_i)(TL_{jc} - \bar{R}_j)}{\sqrt{\sum_{c \in I_{ij}} (TL_{ic} - \bar{R}_i)^2} \sqrt{\sum_{c \in I_{ij}} (TL_{jc} - \bar{R}_j)^2}} \right|, \text{其它情况} \end{array} \right. \quad (11)$$

式(11)就是节点间相关性的计算方法,取值范围在 $[0, 1]$ 之间. $sim(i, j)$ 的值越大,表示节点 i 和 j 之间的相关性越大.

5.2 相关性信任度

本文提出基于多维历史的节点相关性信任度来表示节点的可信程度.相关性信任度不仅仅考虑到目标节点的全局信任度,还考虑源节点与目标节点之间的相关性.

定义 18. 基于历史相关性的信任度 $MDHT_{ij}$, 表示节点 i 给 j 的相关性信任度评价,是从节点 i 看来, j 的全局信任度 TG_j 与 (i, j) 之间的相关性因子 $sim(i, j)$ 的乘积 $sim(i, j) \cdot TG_j$, 且当 $sim(i, j) = 0$ 时, $MDHT_{ij}$ 取 TG_j 与 0 之间的最小值.

$$MDHT_{ij} = \begin{cases} \min(0, TG_j), & sim(i, j) = 0 \\ sim(i, j) \cdot TG_j, & sim(i, j) > 0 \end{cases} \quad (12)$$

根据定义,相关性信任度 $MDHT_{ij}$ 具有如下性质.

(1) 全局性. 因为相关性信任度使用了全局信任度,因此具有全局信任度全局性的特点.从本质上来讲,相关性信任度仍然是一种全局信任度.这样能有效避免单个恶意节点欺骗或攻击的行为.

(2) 客观性. 由于相关性信任度考虑到了节点之间的相关性因素,因此,即使某个恶意节点团体集体伪造了高的信任值,也将因为原节点与目标节点之间的相关性很低,而使得乘积运算后变得很小.因此,相比全局信任度而言,相关性信任度具有更高的客观性.

6 仿真性能分析与实例说明

6.1 仿真性能分析

本文通过仿真对 MDHTrust 模型的准确性进行检验,仿真平台用 C# 语言进行设计与实现,通

过多线程方式模拟分布式通信.定义了 3 类节点:(1) A 类.这类节点是 P2P 系统中的正常节点,提供正常服务以及做出正常的服务评价;(2) B 类.这类节点是 P2P 系统中的单个恶意节点,提供虚假服务并且做出虚假评价,但不存在协同行为;(3) C 类.这类节点是 P2P 系统中的恶意团队节点成员,提供虚假服务,但对正常节点进行虚假评价,而对其团队成员节点信任进行夸大,具有协同欺骗行为.

(1) 实验 1. 局部信任度有效性验证.

首先对 MDHTrust 模型中的局部信任度模型的有效性进行测试,如图 7 所示.实验中的所有节点全为 A 类节点,进行等额交易(为了方便观察局部信任度因子的有效性),从中观察到高频率通信的 A 类节点和低频率通信的 A 类节点都具有良好的信任收敛性,其中在高频交易环境下, A 类节点能更快的提升自身的局部信任,而低频交易环境下其信任度受到频率因子的影响;结果表明 MDHTrust 中的局部信任度模型是有效的.

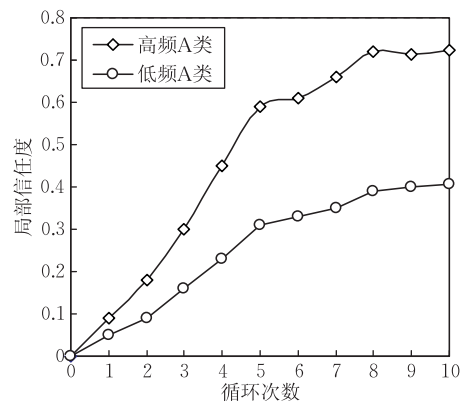


图 7 局部信任度测试

(2) 实验 2. 全局信任度有效性验证.

随机生成若干个 A 类、B 类、C 类节点,将 2000 份具有可信标记的文件摘要信息按 SHA-1 散列编码方式存储在 200 个 A 类节点中,将 1000 份具有虚假标记的文件摘要信息散列在 100 个 C 类节点中,将 500 份虚假标记文件摘要存放在 50 个 B 类节点中.之后启动随机通信,根据公式,每个节点的初始全局信任度被设置为 0,循环测试,观察指定的一个 A 类节点、B 类节点和 C 类节点的全局信任度的走势,实验结果如图 8 所示.

图 8 中, B 类节点的全局信任度迅速下降,而 A 类和 C 类并未出现明显差异.该结果表示 MDHTrust 模型中的全局信任模型可以有效并迅速地判断出单个恶意节点(B 类节点)的攻击和欺骗行为,但却无

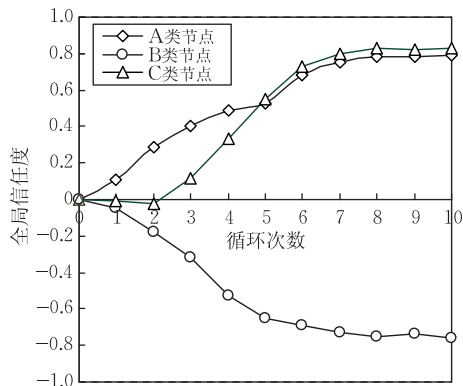


图 8 三类节点的全局信任度测试

法对团队恶意节点的欺骗行为进行判断。

(3) 实验 3. 相关性信任度有效性验证.

首先,在实验 2 的环境下,观察一个 A 类节点 a_1 和 A 类节点 a_2 、C 类节点 c_2 的相关性因子,结果如图 9 所示。

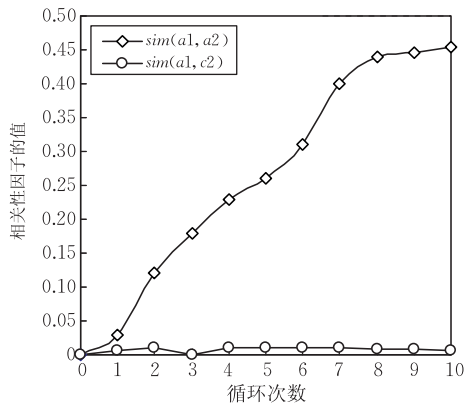


图 9 相关性因子测试

图 9 结果表明,随着不断的实验,节点 a_1 与 a_2 的相关性因子越来越高,而 a_1 与 c_2 的相关性因子一直处于低等水平,表明相关性因子的设计合理。

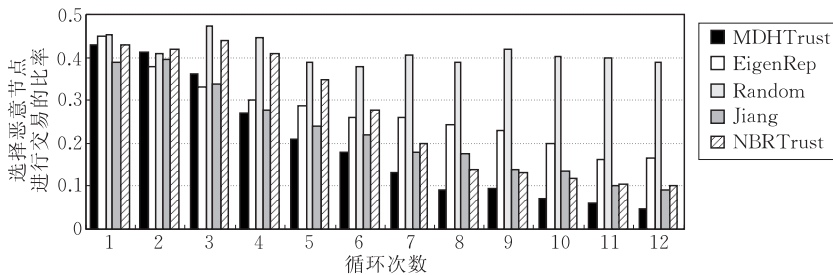


图 11 不同模型选择恶意节点进行交易的比例

从图 11 的实验结果看,随机模型具有最差的收敛性,而 MDHTrust 的收敛效果较好。虽然 EigenRep 模型和 Jiang 的模型都能有效收敛,但由于这两个模型都没有引入相关性参数,使得其对实验环境中的团队恶意节点(C 类节点)的协同行为反应较慢。尽管在 Jiang 模型中引入了对评价的信任判断来处

同时,实验观察 a_1 对 a_2 、 a_1 对 c_2 的相关性信任度评价,结果如图 10 所示。

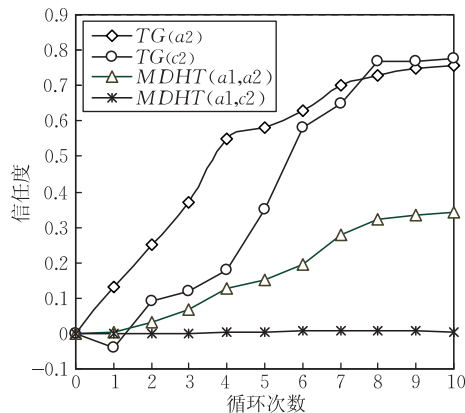


图 10 相关性信任度 MDHT(a_1, a_2) 与 MDHT(a_1, c_2) 测试

结果表明,由于有团队的影响, a_2 和 c_2 都具有高的全局信任度,但是,由于加入了相关性因子,使得最终的 MDHT(a_1, a_2) 与 MDHT(a_1, c_2) 能清楚地区分可信节点与团队恶意节点,相关性信任度具有更高的客观性。

(4) 实验 4. 与其它模型的性能比较.

本文将 MDHTrust 中的相关性信任模型与 EigenRep 模型^[14]、Jiang 的模型^[17]、NBRTrust 模型^[20]、随机模型进行比较实验.实验统计在不同信任模型下所有 A 类节点选择从恶意节点(B 类和 C 类)进行交易的次数在总交易次数中所占的比重(如所有 A 类节点总共交易了 100 次,其中有 20 次选择了 B 类和 C 类交易,则比例为 20%).实验结果如图 11 所示。

理协同恶意行为,同时通过线性时间因子对评价进行了因式处理,但相比 MDHTrust 的相关性因子和非线性时间因子,其收敛速度相对较慢。NBRTrust 模型因本身较为简单,只计算了成功率和相关性因子,这使得其收敛效果较差。该结果表明,相比 EigenRep 模型、Jiang 模型、NBRTrust 模型和随机模型,本文所

做的 MDHTrust 模型具有多维参数和相关性因子, 具有更好的收敛性能, 能更好地屏蔽恶意节点(单个恶意节点与团队恶意节点), 是一个有效的模型。

6.2 应用方向与实例说明

MDHTrust 信任模型在用户评价的基础上, 集成了时间因子、频率因子、额度因子、交易成功率因子以及相关性因子, 可以产生局部信任度、全局信任度和相关性信任度. 可以适合 P2P 文件共享系统、P2P 搜索引擎系统、多媒体共享系统以及 P2P 移动电子商务系统等应用类型, 通过 MDHTrust 模型能够获得一个较为全面的信任数据列表. 在具体应用实例上, 集成 MDHTrust 的系统需要为 MDHTrust 开辟单独的通信端口进行分布式通信, 用以实施节点通信历史数据的分布式数据存储与通信. 在通信历史的基础上, 对各个因子进行独立计算和显示. 用户可以得到每个节点当前的用户评价、时间因子、频率因子、额度因子、交易成功率因子以及相关性因子的情况, 用以帮助用户做出选择. 同时, 集成 MDHTrust 模型的 P2P 系统实时更新局部信任度、全局信任度和相关性信任度的数据曲线, 使用户对当前待选择的节点有较全面的了解. 针对用户每次交易之前的节点搜索请求, 系统根据 MDHTrust 模型所提供的数据为用户提供节点推荐列表, 用户可以直接接受系统推荐或结合本身的判断进行最终选择。

7 结论与展望

本文提出了一种新的 P2P 分布式信任模型 MDHTrust. 设计了节点多维通信历史向量及其分布式存储结构 *TimeDBList* 和 *mdh-list*, 综合两个节点间交易历史中的评价、时间因子、额度因子、频率因子, 提出了局部信任度计算方法; 综合全局节点交易历史中的时间因子、频率因子和成功率因子, 提出了全局信任度计算方法及通过分布式计算节点信任评价行为的相关性信任度的计算方法. 值得一提的是, 所设计的时间因子为非线性因子, 而在交易额因子引入了价值额度的概念. 仿真分析表明, 相比其它模型, MDHTrust 信任模型具有较快的信任收敛速度, 能够有效抵御单个恶意节点和团队恶意节点的欺骗行为, 是一种有效的信任模型。

当然, 信任模型的研究是一个长期的过程, 仍然有很多问题需要在未来的工作中进一步完善, 比如: (1) P2P 信任模型下的分布式通信策略; (2) 研究社会网络中关于社群信任的相关策略, 用以指导和完善信任模型; (3) 研究基于交易内容的信任评价策略, 比如基于交易价格的分布式评价策略等。

参 考 文 献

- [1] Touch J. Overlay networks. *Computer Networks*, 2001, 36(2): 115-116
- [2] Zhang P, Helvik B E. Modeling QoS in P2P file-sharing with benign and malicious peers by stochastic activity networks//*Proceedings of the 7th IEEE Consumer Communications and Networking Conference*. Las Vegas, NV, USA, 2010: 461-465
- [3] Mekouar L, Iraqi Y, Boutaba R. Peer-to-peer's most wanted: Malicious peers. *Computer Networks*, 2006, 50(4): 545-562
- [4] Zhang Q, Sun Y, Liu Z, Zhang X, Wen X Z. Design of a distributed P2P-based grid content management architecture//*Proceedings of the 3rd Annual Communication Networks and Services Research Conference*. Halifax, NS, Canada, 2005: 339-344
- [5] Esaki H. A consideration on R & D direction for future Internet architecture. *International Journal of Communication Systems*, 2010, 23(6-7): 694-707
- [6] Gambetta D. Can we trust trust? //Gambetta D ed. *Trust: Making and Breaking Cooperative Relations*. Oxford: Basil Blackwell, 1990: 213-238
- [7] Rahman A A, Hailes S. Supporting trust in virtual communities//*Proceedings of the 33rd Hawaii International Conference on System Sciences*. Maui, USA, 2000: 4-7
- [8] Wang Shou-Xin, Zhang Li, Li He-Song. Evaluation approach of subjective trust based on cloud model. *Journal of Software*, 2010, 21(6): 1341-1352(in Chinese)
(王守信, 张莉, 李鹤松. 一种基于云模型的主观信任评价方法. *软件学报*, 2010, 21(6): 1341-1352)
- [9] Dou Wen, Wang Huai-Min, Jia Yan et al. A recommendation-based Peer-2-Peer Trust model. *Journal of Software*, 2004, 15(4): 571-583(in Chinese)
(窦文, 王怀民, 贾焰等. 构造基于推荐的 Peer-to-Peer 环境下的 Trust 模型. *软件学报*, 2004, 15(4): 571-583)
- [10] Aberer K, Despotovic Z. Managing trust in a peer-to-peer information system//*Proceedings of the 10th International Conference on Information and Knowledge Management*. Atlanta, GA, USA, 2001: 1-7
- [11] Damiani E, Vimercati S D C, Paraboschi S. A reputation-based approach for choosing reliable resources in peer-to-peer networks//*Proceedings of the 9th ACM Conference on Computer and Communications Security*. Washington, DC, USA, 2002: 207-216
- [12] Despotovic Z, Aberer K. Maximum likelihood estimation of peers' performance in P2P networks//*Proceedings of the 2nd Workshop on the Economics of Peer-to-Peer Systems*. Harvard University, Cambridge, 2004: 1-9
- [13] Huberman B A, Wu F. The dynamics of reputations. *Journal of Statistical Mechanics: Theory and Experiment*, 2004, (4): 1-17
- [14] Kamvar S D, Schlosser M T. EigenRep: Reputation management in P2P networks//*Proceedings of the 12th International World Wide Web Conference*. Budapest, Hungary, 2003: 123-134

- [15] Song S S, Hwang K, Zhou R F, Kwok Y K. Trusted P2P transactions with fuzzy reputation aggregation. *IEEE Internet Computing*, 2005, 9(6): 24-34
- [16] Yuan Wei, Li Jin-Sheng, Hong Pei-Lin. Distributed Peer-to-Peer trust model and computer simulation. *Journal of System Simulation*, 2006, 18(4): 938-942(in Chinese)
(袁巍, 李津生, 洪佩琳. 一种 P2P 网络分布式信任模型及仿真. *系统仿真学报*, 2006, 18(4): 938-942)
- [17] Jiang Shou-Xu, Li Jian-Zhong. A reputation-based trust mechanism for P2P E-commerce systems. *Journal of Software*, 2007, 18(10): 2551-2563(in Chinese)
(姜守旭, 李建中. 一种 P2P 电子商务系统中基于声誉的信任机制. *软件学报*, 2007, 18(10): 2551-2563)
- [18] Hu Jian-Li, Wu Quan-Yuan, Zhou Bin et al. Robust feedback credibility-based distributed P2P trust model. *Journal of Software*, 2009, 20(10): 2885-2898(in Chinese)
(胡建理, 吴泉源, 周斌等. 一种基于反馈可信度的分布式 P2P 信任模型. *软件学报*, 2009, 20(10): 2885-2898)
- [19] Miao Guang-Sheng, Feng Deng-Guo, Su Pu-Rui. Colluding clique detector based on activity similarity in P2P trust model. *Journal on Communications*, 2009, 30(8): 9-20(in Chinese)
(苗光胜, 冯登国, 苏璞睿. P2P 信任模型中基于行为相似度的共谋团体识别模型. *通信学报*, 2009, 30(8): 9-20)
- [20] Tan Zhen-Hua, Cheng Wei, Chang Gui-Ran et al. A distributed trust model for P2P overlay networks based on correlativity of communication history. *Journal of Northeastern University(Natural Science)*, 2009, 30(9): 1245-1248(in Chinese)
(谭振华, 程维, 常桂然等. 基于通信历史相关性的 P2P 网络分布式信任模型. *东北大学学报(自然科学版)*, 2009, 30(9): 1245-1248)
- [21] Li Yong-Jun, Dai Ya-Fei. Research on trust mechanism for Peer-to-Peer network. *Chinese Journal of Computers*, 2010, 33(3): 390-405(in Chinese)
(李勇军, 代亚非. 对等网络信任机制研究. *计算机学报*, 2010, 33(3): 390-405)
- [22] Deng Ai-Lin, Zhu Yang-Yong, Shi Bo-Le. A collaborative filtering recommendation algorithm based on item rating prediction. *Journal of Software*, 2003, 14(9): 1621-1628(in Chinese)
(邓爱林, 朱扬勇, 施伯乐. 基于项目评分预测的协同过滤推荐算法. *软件学报*, 2003, 14(9): 1621-1628)
- [23] Benesty J, Chen J D, Huang Y T. On the importance of the Pearson correlation coefficient in noise reduction. *IEEE Transactions on Audio, Speech & Language Processing*, 2008, 16(4): 757-765



TAN Zhen-Hua, born in 1980, Ph. D., lecturer. His current research interests include distributed network security and computer architecture.

WANG Xing-Wei, born in 1968, Ph.D., professor, Ph. D. supervisor. His current research interests include NGI, mobile wireless Internet, and IP/DWDM optical Internet, etc.

Background

The P2P technological framework would be a key component of the future Internet system, and the peer-to-peer network has been popular in many large-scale distributed applications because of its openness, anonymous and reciprocity. However, various kinds of malicious appear in kinds of P2P systems, disturbing the trust mechanism of the P2P network. More and more researchers make contributions to the security technologies, and distributed trust model is one of the key technologies.

Based on some social network principles, this paper presented a new distributed trust model named MDHTrust (Multi-Dimension-History Trust Model).

This work is supported by the National Natural Science

CHENG Wei, born in 1970, Ph. D. candidate, lecturer. His current research interests include complex networks and SoC design.

CHANG Gui-Ran, born in 1946, Ph. D., professor, Ph. D. supervisor. His current research interests include computer architecture and network security, etc.

ZHU Zhi-Liang, born in 1962, professor, Ph. D. supervisor. His current research interests include chaos information processing and complex network.

Foundation of China under grant Nos. 61070162, 71071028, 60802023 and 70931001; the Specialized Research Fund for the Doctoral Program of Higher Education under grant No. 20070145017; the Fundamental Research Funds for the Central Universities under grant No. N090504003 and No. N090504006. These projects focus on trusted computing to solve security technologies of the next generation Internet, including P2P security. Authors of this paper work for distributed network security, next generation Internet and so on. They began the research on P2P network security since 2000, and have achieved many scientific research achievements. This paper contributes to the P2P trust model which can be applied on P2P E-commerce, P2P sharing system and P2P search engine.