计算机研究与发展
Journal of Computer Research and Development
ISSN 1000-1239/CN 11-1777/TP
42(5):797~803, 2005

Grid 与 P2P 混合计算环境下基于推荐证据推理的信任模型

朱峻茂[1] 杨寿保[1] 樊建平[2] 陈明宇[2]
[1](中国科学技术大学计算机科学技术系 合肥 230026)
[2](中国科学院计算技术研究所 北京 100080)
(zhujm@mail.ustc.edu.cn)

A Grid & P2P Trust Model Based on Recommendation Evidence Reasoning

Zhu Junmao[1], Yang Shoubao[1], Fan Jianping[2], and Chen Mingyu[2]
[1](*Department of Computer Science and Technology, University of Science and Technology of China, Hefei* 230026)
[2](*Institute of Computing Technology, Chinese Academy of Sciences, Beijing* 100080)

Abstract Under mixed computing environment of grid and P2P (Grid & P2P), grid nodes provide the service with QoS guarantee. However, sharing computing resources of P2P nodes is the user's volunteer action without QoS guarantee. The user is not responsible for his actions. Therefore it's difficult to establish the trust relationship among users with traditional trust mechanism. Referring to social people trust relationship models, a grid & P2P trust model based on recommendation evidence reasoning is designed to solve the problem by building a recommendation mechanism in Grid & P2P and integrating the recommendation evidence with the D-S theory. Theoretical analysis and simulations prove that the model can tackle the trust problem under Grid & P2P in a simple and efficient way.

Key words grid computing; P2P computing; trust; recommendation; evidence reasoning

摘 要 在 Grid 与 P2P 混合计算环境(Grid & P2P)中,Grid 节点提供有 QoS 保证的服务,而 P2P 节点的计算资源属于自主贡献资源,不提供 QoS 保证,用户不为自己的行为承担任何责任,因此节点间的信任关系很难通过传统的信任机制来建立. 参考社会学的人际关系信任模型,通过在 Grid & P2P 中建立信任推荐机制,并利用 D-S 理论对推荐证据进行综合处理来解决该问题. 分析及仿真实验说明,基于推荐证据推理的信任模型可以简单有效地解决 Grid & P2P 中的信任问题.

关键词 网格计算;对等网络计算;信任;推荐;证据推理

中图法分类号 TP338.8

1 引 言

网格计算(grid computing)[1]和对等网络计算(P2P computing)[2]都强调提供普适的、廉价的计算服务,但是又有所区别,其中一个很重要的区别是 Grid 强调提供有 QoS 保证的服务,而 P2P 则不提供此 QoS 保证. 近来出现 Grid 和 P2P 计算相互融合的趋势,本文称之为 Grid 与 P2P 混合计算环境(Grid & P2P),其结构示意图如图 1 所示. 在图 1 中,上层的网格服务提供者 GP(grid provider)通常是可信的高性能计算节点,能够提供有 QoS 保障的计算服务. 下层的 P2P 服务提供者 PP(P2P provider)是地理上广泛分布的计算节点,运行在其上的计算任务不可再分. GP 是计算任务的调度者与管理者,是完成计算任务的主体部分,同时可以分

收稿日期:2003-12-04;修回日期:2004-03-29
基金项目:国家自然科学基金项目(60273041);国家"八六三"高技术研究发展计划基金项目(2002AA104560,2003AA1Z2070)

配计算任务到其他的 GP 节点和 PP 节点上,以构成一个协同的虚拟超级计算环境. 该混合环境既能提供有 QoS 保证的计算服务,又能整合散布在 Internet 上被自主贡献出来的不能提供服务质量保障的计算资源.

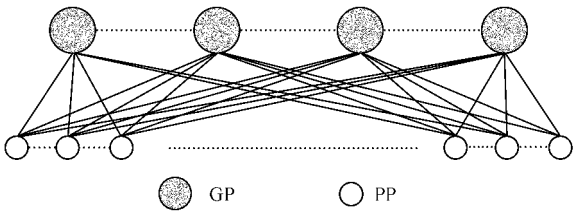

Fig. 1 Structure of mixed computing environment of Grid & P2P.

图 1 Grid 与 P2P 混合计算环境(Grid & P2P)示意图

在 Grid & P2P 中,PP 节点的资源共享是用户自愿行为,因而服务的质量无法得到保证. 为了保证服务质量,一种可行的方法是对每个 PP 节点评定信誉度,信誉度高的节点成为首选. 解决信任问题一般依赖于可信的第 3 方,只要用户持有该授权认证中心(certificate authority, CA)所颁发的证书即被认为是可信的,同时恶意用户必须承担(法律)责任. 然而,这种方案存在如下问题:①在 Grid & P2P 中,是否存在能为每个用户都接受的可信的第 3 方是值得怀疑的;②在 Grid & P2P 中,有一种排斥 CA 的倾向,用户自愿参与网络、自由交易并且不准备为自己的行为负(法律)责任;③即使能够鉴别对方用户的身份,也不能确保对方的行为符合我方的期望. 因此,在 Grid & P2P 中建立一种信任机制是十分必要的,有利于保证计算环境的高可用性和良性发展.

在社会网络中,信任关系是通过个体间的相互推荐而形成的,某一个体的信誉度一般取决于其他个体的推荐,与此同时推荐者的信誉度也决定了其推荐个体的信誉度. 这种互相依赖的信任关系构成了信任网络(Web of trust)[3,4]. 在该网络中,任何个体的信誉度都不是绝对可靠的,带有不确定性,但可以作为其他个体决定其交互行为的依据. 基于信任网络[5,6]的 Grid & P2P 系统与人际关系信任网络有很强的相似性[7,8],这表现在:①网络中的个体在与其他个体的交互中会留下历史信息;②个体对于交互对象具有充分的选择权;③个体一般不看重绝对的可靠性或服务质量,即个体可以忍受少量错误的选择带来的损失;④个体有义务为网络中的其他个体提供推荐信息. 在这种基于推荐的信任网络中,如果把一个 *GP* 节点对另一个 *PP* 节点的信誉度推荐视为一个证据,那么就不可避免地存在不确定性. 例如,一个 *GP* 节点从没有和一个 *PP* 节点交互过,那么该 *GP* 节点对该 *PP* 节点该如何推荐,显然既不是信任,也不是不信任,而是不确定. 而 D-S 证据理论能恰当地表示信息中的"不确定性",同时能够综合具有不同推荐信誉度的证据.

本文旨在借助社会学的人际关系信任模型,在 Grid & P2P 中,构造一个基于信任推荐和利用 D-S 理论进行推荐证据推理的信任模型,并给出该模型数学表述和实现方法.

本文的第 2 节介绍了相关工作;第 3 节介绍了推荐证据推理理论;第 4 节给出了基于推荐证据推理的信任模型;第 5 节分析模型中的若干关键问题;第 6 节给出仿真实验的结果与分析;最后总结全文,并指出了下一步的工作方向.

2 相关工作

目前的信任模型研究主要基于 P2P 环境,可以分为以下几类:

(1) 基于 PKI 的信任模型. 在这类系统中,中心节点的合法性通过 CA 颁发的证书加以保证.

这类系统往往是中心依赖的,具有可扩展性差、单点失效等问题. 这类系统的实例有 Onsale Exchange,eBay[9],eDonkey 等.

(2) 基于局部推荐. 在这类系统中,节点通过询问有限的其他节点以获取某个节点的信誉度,一般采取简单的局部广播的手段,其获取的节点信誉度往往是局部的和片面的. 如 Cornelli 对 Gnutella 的改进建议就是采用这种方法[10].

(3) 全局可信度模型. 为获取全局的节点可信度,该类模型通过相邻节点间相互满意度的迭代,从而获取节点全局的信誉度. Stanford 大学的 EigenRep[11]是目前已知一种典型的全局信任模型. EigenRep 的核心思想是,当节点 i 需要了解任意节点 k 的全局信誉度时,首先从 k 的交互节点(曾经与 k 发生过交互的节点 j)获知节点 k 的信誉度信息,然后根据这些交互节点自身的局部可信度(从 i 的角度看来)综合出 k 的全局信誉度. EigenRep 存在以下几个问题:

(1) 该协议没有考虑到信誉度本身所具有的不确定性. 在该模型中一个节点对另一个节点只有信任与不信任之分(用信誉度值的大小来表示),而没

有考虑到信任的不确定性.

(2) 该模型没有考虑惩罚因素. 模型没有对造成服务失败的节点在信誉度上做出惩罚.

(3) 该模型的协议实现没有考虑网络的性能开销,每次交易都会导致在全网络范围内的迭代,这在大规模网络环境中缺乏工程上的可行性.

本文借鉴了以 EigenRep 为代表的全局推荐思想,基于 D-S 证据推理理论,提出了一种新的 Grid & P2P 下的全局信任模型,并对模型进行了理论分析和仿真实验加以验证.

3 推荐证据推理理论[12,13]

设 Q 是一个问题,其全部可能的答案用集合 Θ 来表示,在本文中仅考虑 $\Theta=\{T,\neg T\}$,其中 T 表示一个 GP 节点对 PP 节点信任,而 $\neg T$ 表示一个 GP 节点对 PP 节点不信任,其中 $t\in\{T,\neg T\}$ 是问题的正确答案,我们希望借助推荐网络找到正确的解答 t. 集合 Θ 称为问题 Q 的识别框架. 由于证据的不完备、不精确、不完全可靠等原因,无法确定 t,但基于推荐网络获得的推荐证据在一定的程度上确定 t 所处的各种范围是可能的. 基于这种思想,Shafer 定义了基本可信度函数的概念.

设变量 θ 论域 Θ 是一有限集,而 Θ 的所有子集构成的幂集记为 2^θ,则 θ 中的每一元素都对应一个关于 θ 的命题,其一般形式为"θ 的值在 A 中".

定义 1. 设 Θ 是识别框架,如果集函数 $m:2^\theta\to[0,1]$(2^θ 为 Θ 的幂集),满足下列条件:

(1) $m(\Phi)=0$;

(2) $\sum_{A\subseteq 2^\theta} m(A)=1$,

则把 m 称为识别框架 Θ 上的基本可信度函数;$\forall A\subseteq\Theta$,$m(A)$ 称为 A 的基本可信数.

基本可信度函数 m 可解释为证据的主观表示,即对于正确答案 t 的判断,一个推荐证据使 GP 节点产生一个识别框架 Θ 及 Θ 上的一个基本可信度函数 m,命题 $A\subseteq\Theta$ 为真,表示 $t\in A$,$m(A)$ 表示在 GP 看来该推荐证据支持命题 A、且由于证据不足不支持命题 A 的任何真子集的程度. 因此可得 $m(\{T\})+m(\{\neg T\})+m(\{T,\neg T\})=1$.

定义 2. 子集 $A\subseteq\Theta$,若 $m(A)>0$,则称 A 为函数 m 的焦元. 所有焦元的集合称为核,证据是由证据体$(A,m(A))$组成的.

节点 GPi 对节点 PPj 的局部信任度来自于节点 i 与 j 的交互历史. 识别框架为 $\Theta=\{T,\neg T\}$,其焦元为 $\{T\}$,$\{\neg T\}$,$\{T,\neg T\}$,相应地定义其基本可信数:

$$\alpha_{i1}=\frac{S_{ij}}{I_{ij}},$$

$$\alpha_{i2}=\frac{F_{ij}}{I_{ij}},$$

$$\alpha_{i3}=1-\alpha_{i1}-\alpha_{i2},$$

其中,I_{ij} 为节点 GPi 与 PPj 在最近某个固定时间 τ 内(τ 是一个可设定参数,视具体应用而定)实际交互的次数;S_{ij} 为在节点 GPi 看来交易成功的次数;F_{ij} 为在节点 GPi 看来交易失败的次数. 如果 $I_{ij}=0$,则表示节点 GPi 与 PPj 在最近某个固定时间 τ 内没有交互行为,此时 GPi 对 PPj 的局部信任是不确定的,此时设 $\alpha_{i1}=\alpha_{i2}=\sigma$,则 $\alpha_{i3}=1-2\sigma$(σ 也是一个可设定参数). 引入 τ 表示本模型更注重节点的行为时限性,同时也使模型更加灵活. I_{ij},S_{ij},F_{ij} 的信息需要节点 GPi 来维护,因为只有其本身才能判定 PPj 的某一次交互的行为是否符合节点 GPi 的期望.

定义 3. 设 Θ 是识别框架,$m:2^\theta\to[0,1]$(2^θ 为 Θ 的幂集)是 Θ 上的一个基本可信度函数,则称由

$$Bel(A)=\sum_{B\subseteq A} m(B),\ \forall A\subseteq\Theta$$

所定义的函数 $Bel:2^\theta\to[0,1]$ 为 Θ 上对应于 m 的信度函数.

由定义 3 可知,信度函数与基本可信度函数是相互惟一确定的,即有 $Bel(\{T,\neg T\})=m(\{T\})+m(\{\neg T\})+m(\{T,\neg T\})=1$,因此它们是同一证据的不同表示. 同时从定义 3 不难得出下列不等式:

$$\forall A\subseteq\Theta,\ Bel(A)+Bel(\bar{A})\leqslant 1,$$

其中,$\bar{A}$ 表示 A 关于 Θ 的补集,这是信度函数与概率函数的基本区别之一. 结合基本可信度函数的解释,信度函数值 $Bel(A)$ 可理解为推荐证据对命题 A 的总支持度或在该证据下 GP 有理由相信命题 A 的程度.

定义 4. 设 $\bar{A}$ 是 A 关于 Θ 的补集,则称由 $Pl(A)=1-Bel(\bar{A})$ 所定义的函数 $Pl:2^\theta\to[0,1]$ 为似然函数. 似然函数 $Pl(A)$ 表示不反对命题 A 的程度.

定义 5. 设两个推荐证据是完全独立的,它们在识别框 Θ 上对应的基本可信度函数分别为 m_1 和 m_2,定义函数 m_{12} 是识别框上的一个基本可信度

函数:

$$m_{12}(\Phi)=0,$$
$$m_{12}(A)=k\sum_{X\cap Y=A}m_1(X)m_2(Y),\ A\neq\Phi, \tag{1}$$

其中,

$$k=\Big(\sum_{X\cap Y\neq\Phi}m_1(X)m_2(Y)\Big)^{-1}. \tag{2}$$

式(1)称为合成 m_1 和 m_2 的 Dempster 规则,记为 $m_{12}=m_1\oplus m_2$,它反映了 m_1 和 m_2 对应的两个推荐证据对命题 A 的联合支持程度. 其中 k 为归一化因子,保证 m_{12} 的值域规格化到标准空间[0,1]上.

4 基于推荐证据推理的信任模型

令 $\Theta=\{T,\neg T\}$ 为识别框,假定有 n 个 GP 节点 $GP_1,\cdots,GP_n$,m 个 PP 节点 $PP_1,\cdots,PP_m$,节点 $GP_i(1\leqslant i\leqslant n)$ 推荐节点 $PP_j(1\leqslant j\leqslant m)$ 的信誉度信息 $\{T\}$,$\{\neg T\}$ 以置信度 α_{i1},α_{i2} 为真,置信度 α_{i1},α_{i2} 满足:

$$\alpha_{i1}+\alpha_{i2}\leqslant 1. \tag{3}$$

同时,n 个 GP 节点的推荐地位可以不同,可以使用权值来加以区分. 设 GP_i 的权值为 $\omega_i(i=1,\cdots,n)$,满足:

$$\sum_{1\leqslant i\leqslant n}\omega_i=1,\ \omega_1,\cdots,\omega_n\geqslant 0. \tag{4}$$

计算节点 $PP_j(1\leqslant j\leqslant m)$ 信誉度的输入参数如表 1 所示,下面就是利用这些推荐信息来获得识别框 Θ 上各个命题的可信度. 为了解决这一问题,首先必须对每个 GP 节点提供的推荐证据建立各自的基本可信度函数模型;然后再把这些基本可信度函数按照 Dempster 规则综合为一个统一的基本可信度函数.

Table 1 Input Parameter of Model Based on Recommendation Evidence Reasoning

表 1 推荐证据推理模型的输入参数

GP Node	Weight	$\{T\}$	$\{\neg T\}$
GP_1	ω_1	α_{11}	α_{12}
GP_2	ω_2	α_{21}	α_{22}
□	□	□	□
GP_n	ω_n	α_{n1}	α_{n2}

4.1 *GP* 节点的基本可信函数建模

(1) 自身节点的基本可信函数

一个 GP 节点对一个 PP 节点的信誉度评价会根据两个方面的信息来综合,一是其自身与该 PP 节点的直接历史交互记录,二是其他 GP 节点对该 PP 节点的推荐信誉度. 一般来说自身评价会更看重一些,因而相应地其权值更大.

首先建立自身节点信息源 GP_k 的基本可信度函数,记该基本可信函数为 $m(A\mid GP_k)$,$A\subseteq\Theta$. $m(A\mid GP_k)$ 表示 GP 节点对命题 A 提供的支持强度,由表 1 已知,GP_k 提供的信息 $\{T\}$,$\{\neg T\}$ 为真的置信度分别是 α_{k1},α_{k2}. 本文就是借助这个置信度 α_{kj} 建立支持强度 $m(\{\theta_1\}\mid GP_k)$,其关系如下:

$$m(\{\theta_1\}\mid GP_k)=\lambda\alpha_{kl},\ l=1,2, \tag{5}$$

其中,系数 λ 为区间(0,1]上的一个常系数. 由于 α_{kl} 满足式(3),因此可得

$$m(\{T\}\mid GP_k)+m(\{\neg T\}\mid GP_k)\leqslant 1.$$

为使 m 成为基本可信度函数,补充定义:

$$m(\{T,\neg T\}\mid GP_k)=1-(m(\{T\}\mid GP_k)+m(\{\neg T\}\mid GP_k)). \tag{6}$$

此时,根据定义 1 可知,$m(\cdot\mid GP_k)$ 是一个基本可信度函数,它的焦元至多包括 $\{T\}$,$\{\neg T\}$ 和整个识别框 $\Theta=\{T,\neg T\}$,这个函数就是对自身节点的基本可信度函数.

(2) 推荐节点的基本可信函数

设 GP_k 是自身节点,那么其他的 GP 节点 $GP_i(i\neq k)$ 都称为推荐节点. 下面就来建立推荐节点的基本可信函数 $m(\cdot\mid GP_i)(i\neq k)$. 设 GP_i 对 $\{T\}$,$\{\neg T\}$ 的置信度为 α_{i1},α_{i2},那么可以相应地把推荐节点的基本可信度函数定义为

$$m(\{T\}\mid GP_i)=\frac{\omega_i}{\omega_k}\lambda\alpha_{i1},$$
$$m(\{\neg T\}\mid GP_i)=\frac{\omega_i}{\omega_k}\lambda\alpha_{i2}, \tag{7}$$
$$m(\{T,\neg T\}\mid GP_i)=1-(m(\{T\}\mid GP_i)+m(\{\neg T\}\mid GP_i)),\ i\neq k. \tag{8}$$

(3) 基本可信度函数矩阵

进一步简化记号,对任意的 GP 节点 $GP_i(i=1,\cdots,n)$,令

$$m_i^1=m(\{T\}\mid GP_i),$$
$$m_i^2=m(\{\neg T\}\mid GP_i),$$
$$m_i^0=m(\{T,\neg T\}\mid GP_i).$$

由上面建立的 n 个 GP 节点的 n 个基本可信度函数可简明地用如下矩阵 $\boldsymbol{M}$ 来表示:

$$\boldsymbol{M}=\begin{bmatrix} m_1^0 & m_1^1 & m_1^2\\ m_2^0 & m_2^1 & m_2^2\\ \square & \square & \square\\ m_n^0 & m_n^1 & m_n^2\end{bmatrix},$$

其中，第 i 行表示 GP_i 的基本可信度函数（$i=1,\cdots,n$）.

4.2 证据推理算法

现在采用 Dempster 规则把矩阵 $\boldsymbol{M}$ 中的 n 个基本可信度函数综合为一个统一的基本可信度函数. 为此假定 n 个 GP 节点是独立的. 由于这 n 个基本可信度函数具有一个共同的特殊性质，即它们的焦元都至多为单元素集 $\{T\}$，$\{\neg T\}$和识别框 Θ 本身，因此下面给出的 Dempster 算法是一个线性时间的算法，而一般的 Dempster 规则是指数时间的算法.

递归算法如下：

把前 i 个 GP 节点定义为下列集合：

$s(i)=\{GP_1,GP_2,\cdots,GP_i\}$，$i=1,\cdots,n$.

此时矩阵 $\boldsymbol{M}$ 中前 i 列的 i 个基本可信度按 Dempster 规则综合得到新的基本可信度函数记为

$$m^1_{S(i)}=m(\{T\}\mid S(i)),$$
$$m^2_{S(i)}=m(\{\neg T\}\mid S(i)), \tag{9}$$

$$m^0_{S(i)}=m(\{T,\neg T\}\mid S(i)). \tag{10}$$

下面的公式不难用‘交集列表法’直接得出.

对 $S(2)=\{GP_1,GP_2\}$的综合，有以下公式：

$$m^1_{S(2)}=k_{S(2)}[m^1_1m^1_2+m^1_1m^0_2+m^0_1m^1_2], \tag{11}$$

$$m^2_{S(2)}=k_{S(2)}[m^2_1m^2_2+m^2_1m^0_2+m^0_1m^2_2], \tag{12}$$

$$m^0_{S(2)}=k_{S(2)}m^0_1m^0_2, \tag{13}$$

$$k_{S(2)}=1-(m^1_1m^2_2+m^2_1m^1_2). \tag{14}$$

显然，$m^0_{S(1)}=m^0_1$，$m^1_{S(1)}=m^1_1$，$m^2_{S(1)}=m^2_1$，所以可用类似的方法综合 $S(i+1)=\{GP_1,\cdots,GP_i,GP_{i+1}\}$，并得到如下的递归公式：对任意的 $i=1,2,\cdots,n-1$，有：

$$m^j_{S(i+1)}=k_{S(i+1)}[m^j_{S(i+1)}m^j_{i+1}+m^j_{S(i+1)}m^0_{i+1}+m^0_{S(i+1)}m^j_{i+1}],\ j=1,2, \tag{15}$$

$$m^0_{S(i+1)}=k_{S(i+1)}m^0_{S(i+1)}m^0_{i+1}, \tag{16}$$

$$k_{S(i+1)}=1-(m^1_{S(i+1)}m^2_{i+1}+m^2_{S(i+1)}m^1_{i+1}). \tag{17}$$

在式(15)，(16)中，当 $i=n-1$ 时，就得到全部 n 个基本可信度函数的综合基本可信度函数，记此基本可信度函数为 m，则有：

$$m(\{T\})=m^1_{S(n)},\ m(\{\neg T\})=m^2_{S(n)}, \tag{18}$$

$$m(\{T,\neg T\})=m^0_{S(n)}, \tag{19}$$

$$m(A)=0,\ \forall A\neq\{T\};$$
$$A\neq\{\neg T\};A\neq\{T,\neg T\}.$$

这个函数 m 即为推荐证据推理模型的输出.

5 模型中的几个关键问题

(1) PP 节点信誉度初值的设定，即参数 σ 的选择问题，其参数的设定应满足如下原则：①σ 不能过小，以防初次欲提供计算服务的 PP 节点因信誉度太低而永远得不到计算任务，最终导致该计算节点被饿死；②同时 σ 也不能过大，因为如果 σ 过大，那么一个节点就可能以欺骗的方式骗得一个计算任务后，再重新以一个新的身份再次骗得一个计算任务.

(2) 对不诚实的 PP 节点的惩罚问题. 假设一个“有预谋”的 PP 节点在开始阶段每次都提供诚实的计算任务，一段时间以后该节点的信誉度会相当得高，此时该节点开始以下列方式提供计算服务：一次提供诚实服务，一次提供不诚实服务，这样的行为对该 PP 节点的信誉度并不会造成多大的影响，因此必须引入对欺骗行为的惩罚机制. 修改基本可信数定义如下：

$$\alpha_{i1}=\frac{S_{ij}}{S_{ij}+\rho F_{ij}},\ \alpha_{i2}=\frac{\rho F_{ij}}{S_{ij}+\rho F_{ij}},$$

其中，$\rho\geqslant 1$，称为惩罚系数. 当 $\rho=1$ 时即为原来的定义.

(3) GP 节点的权值赋值问题. 在模型中，对不同的推荐 GP 节点赋予了不同的权值，自身节点 GP_k（调度节点）权值定义为 ω_k，$0<\omega_k<1$，其他 GP 节点的权值为$(1-\omega_k)/(n-1)$. 为了防止 GP 节点的不诚实行为，故意提高一个 PP 节点的信誉度，在每次的计算任务结束时，自身 GP 节点重新计算在自身节点看来其他各个 GP 节点的相对推荐权值，对有不诚实行为的 GP_i 节点可以降低其权值，给出惩罚数 ε，$\varepsilon<\omega_i$，其中的 ω_i 是 GP_i 节点的当前权值. 有两种策略处理 ε：①如果自身节点认为在目前的混合计算环境中，不存在其他 GP 节点协同欺骗行为，那么把该惩罚数 ε 均分给其他的节点；②反之，自身节点将把该惩罚数 ε 留给自己.

(4) PP 节点的选择问题. 假设自身节点 GP_i 在若干个满足计算任务需求的 PP 节点中，设为 $\{PP_1,PP_2,\cdots,PP_q\}$，每次都把计算任务分配给信誉度最高的节点，会出现服务热点现象，某一 PP 节点一直很忙，而其他也符合条件的节点则可能分配不到计算任务而饿死，致使负载不平衡，这与 Grid & P2P 计算的初衷相背离. 在本信任模型中，采用如下方式加以解决：在 $\{PP_1,PP_2,\cdots,PP_q\}$信誉度高的 $\min(N,\lfloor q/2\rfloor)$（$N$ 是一个可设定参数）个

PP 节点中随机选取一个 *PP* 节点,把计算任务分配到该节点上,从而可以有效地避免计算服务的热点现象.

6 仿真及其结果分析

(1) 模型有效性仿真实验

该仿真实验用来评测前面构造的信任模型. 在仿真实验中,把计算服务简化为文件共享服务,即 *GP* 节点从 *PP* 节点上下载共享的文件,下载文件的真实性是其判断一次交互是否成功的惟一标准. 假设 *PP* 节点有两类:一类是诚实的节点,即它提供真实可信的文件下载服务;另一类是不诚实的节点,即它有可能提供一个假文件下载服务,也有可能根本就不提供文件下载服务,也可能在下载的文件中含有恶意代码(例如病毒). 在实验中,假定该网络是理想的,即任意一个节点可以找到任意文件及其声称为该文件拥有者的所有节点(文件可能并不真实). 用户的行为较为简单,即从所有声称拥有其所需文件的节点中选择信誉度最高的节点,从该节点上下载文件.

对于规模为 10000 个节点的仿真网络,其中 100 个 *GP* 节点、9900 个 *PP* 节点,我们设定的文件总数为 100000 个,$\sigma = 0.48$, $\omega_k = 0.5$, $\varepsilon = \omega_i/2$, $\lambda = 1$. 我们将 100000 个文件随机分配到所有 10000 个节点,并保证每个文件至少被一个诚实的节点拥有. 每个用户在整个仿真过程中必须完成 100 次下载,每次下载目标为从其不曾拥有的文件中随机选择一个并试图进行下载. 下载的成功使得该节点拥有该文件,失败的下载不会增加该节点拥有的文件.

仿真实验的最终评价标准是整个仿真系统成功下载的次数. 显然,在理想的状态下(所有的节点都是诚实守信的节点),成功下载的次数为 10^6. 实验结果如图 2 所示,本文提出的信任模型对不诚实的节点所产生的影响有相当的抑制作用,当不诚实节点所占的比例达到 50% 时,整个系统的成功下载的次数仍能达到 816087.

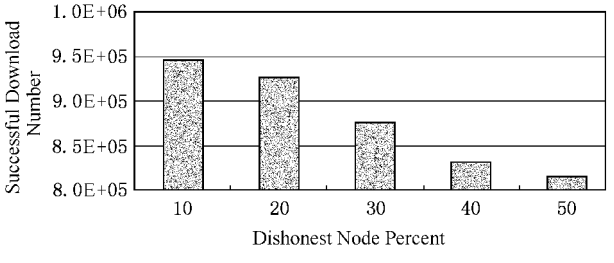

Fig. 2 Successful download number according to dishonest node percent.

图 2 在不同比例的不诚实节点情况下文件成功下载次数

(2) 服务热点现象实验

如果每次下载都选择信誉度值最大的节点,出现服务热点的现象,本仿真实验就是为了验证这一现象. 仿真环境同上,为了简化实验,假定整个下载网络是理想状态,节点有可能不提供服务,也有可能提供含有恶意代码的文件,但是只要提供服务,每次下载的文件都是真实的. 把某个文件 10 份随机地分布到 10 个节点上,然后这 10 个节点提供该文件的下载服务,每个节点对该文件的下载请求为 100 次(每次下载成功后,该节点并不提供此文件的下载服务). 实验结果如图 3 和图 4 所示,采用基于信誉度的概率下载($N = 5$)有效地消除了计算热点现象,解决了负载平衡问题.

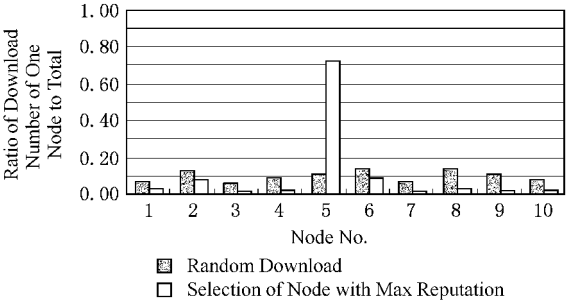

Fig. 3 Comparison between random download and choosing node with max reputation.

图 3 自由下载与选择信誉度最大值下载的对比实验

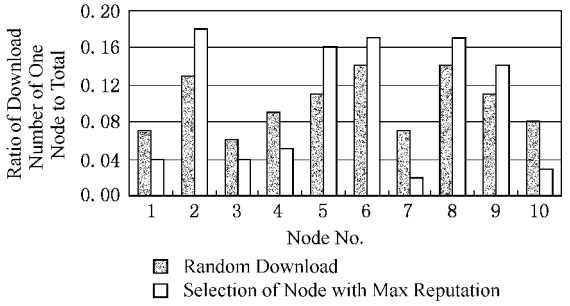

Fig. 4 Comparison between random download and probability mechanism based on trust model.

图 4 自由下载与基于信誉度的概率选择下载对比实验

7 结 论

本文在 Grid & P2P 环境下构造了一个基于信任推荐和推荐证据推理的信任模型,给出了信任推荐模型以及推荐证据推理算法. 分析和仿真实验表

明,该模型克服了已有的信任模型的若干局限性,简单有效,具有良好的工程可行性. 模型的安全问题、推荐证据冲突处理等问题是下一步的研究方向.

参 考 文 献

1 Ian Foster, Carl Kesselman, Steven Tuecke. The anatomy of the grid. International J. Supercomputer Applications. http://www.globus.org/research/papers/anatomy.pdf, 2001

2 Andy Oram. Peer to Peer: Harnessing the Power of Disruptive Technologies. New York: O'Reilly & Associates, 2001

3 R. Khare, A. Rifkin. Weaving a Web of trust. World Wide Web, 1997, 2(3): 77~112

4 G. Caronni. Walking the Web of trust. In: Proc. the 9th Workshop on Enabling Technologies (WET ICE '2000). Los Alomitos, CA: IEEE Computer Society Press, 2000

5 T. Beth, M. Borcherding, B. Klein. Valuation of trust in open networks. In: Proc. the 3rd European Symposium on Research in Computer Security. Brighton: Springer-Verlag, 1994. 3~18

6 A. Abdul-Rahman, S. Hailes. A distributed trust model. In: Proc. the 1997 New Security Paradigms Workshop. Cumbnia: ACM Press, 1997. 48~60

7 J. Scott. Social Network Analysis: A Handbook. Oxford: SAGE Publications, 2000

8 S. Wasserman. Social network analysis: Methods and applications. Cambridge: Cambridge University Press, 1994

9 P. Resnick, R. Zeckhauser. Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. NBER Workshop on Empirical Studies of Electronic Commerce, California, 2000

10 F. Cornelli. Choosing reputable servents in a P2P network. The 11th Int'l World Wide Web Conf., Honolulu, Hawaii, 2002

11 S. Kamvar. EigenRep: Reputation management in P2P networks. Stanford University, Tech. Rep.: SCCM-02-16, 2002

12 G. Shafer. A Mathematical Theory of Evidence. Princeton, NJ: Princeton University Press, 1976

13 Kari Sentz. Combination of evidence in dempster-shafer theory. Sandia National Laboratories, Tech. Rep.: SAND2002-0835, 2002

Zhu Junmao, born in 1974. Ph.D. candidate. His main research interests include computer system, grid computing, and network.

朱峻茂,1974 年生,博士研究生,主要研究方向为计算机体系结构、网格计算、计算机网络.

Yang Shoubao, born in 1947. Professor and Ph. D. supervisor. His main research interests include computer system, distributed computing, information security, and cryptology(syang@ustc.edu.cn).

杨寿保,1947 年生,教授,博士生导师,主要研究方向为计算机体系结构、分布式计算、信息安全与密码学.

Fan Jianping, born in 1960. Professor, and Ph. D. supervisor. His main research interests include structure of computer system, high performance computing(fan@ict.ac.cn).

樊建平,1960 年生,研究员,博士生导师,主要研究方向为计算机体系结构、高性能计算.

Chen Mingyu, born in 1972. Associate professor, His main research interests include structure of computer system, and high performance computing(cmy@ict.ac.cn).

陈明宇,1972 年生,副研究员,主要研究方向为计算机体系结构、高性能计算.

Research Background

This Paper is supported by the National Science Foundation (No. 60273041) and the National "863" Hi-Tech Research & Development Project Foundation (No. 2002AA104560 and No. 2003AA1Z2070).

On the one hand, both grid computing and P2P computing focus on providing pervasive and inexpensive computing services. On the other hand, there are differences between them on application mode, quality of service, security management. Above all, grid computing provides service with QoS guarantee, while P2P doesn't provide such service. Recently the tendency to the mixture of grid computing and P2P is coming out which is defined as Grid & P2P in this paper.

Under mixed computing environment of grid and P2P (Grid & P2P), grid nodes provide the service with QoS guarantee. However, sharing computing resources of P2P nodes is the user's volunteer action without QoS guarantee. The user is not responsible for his actions. Therefore it's difficult to establish the trust relationship among users with traditional trust mechanism. Referring to social people trust relationship model, a Grid & P2P trust model based on recommendation evidence reasoning is designed to solve the problem by building a recommendation mechanism in Grid & P2P and integrating the recommendation evidence with D-S theory. Theoretical analysis and simulation tests prove that the model can tackle the trust problem under Grid & P2P in a simple and efficient way.