

适应自组织管理模式的 P2P 资源共享网络激励模型的研究

刘 业, 杨 鹏, 刘林峰, 夏 勤

(东南大学计算机网络和信息集成教育部重点实验室, 江苏南京 210096)

摘 要: P2P 网络中的自组织管理模式使得节点的自私行为大量存在. 建立相应解决方案所面临的主要问题是难以继续维持 P2P 网络拥有的自组织及规模可扩放性等特性. 目前存在的信用管理模型或者激励模型多是采用泛洪方式获得网络其他节点的历史行为信息, 庞大的报文通信量和推导算法的高时间复杂度制约了所能应用的 P2P 网络的规模. 本文提出了一种利用随机相遇博弈理论指导建立的 P2P 资源共享网络激励模型 ResP2P. 该模型通过引入节点信誉及信誉恢复区分机制, 同时制定相关行为社会规范, 来激励理性节点为使其自身收益最大化而向整个网络贡献资源, 并且 ResP2P 所对应的分布式算法易于在自组织管理模式的网络环境中实施. 实验证实了 ResP2P 模型激励机制的有效性.

关键词: P2P 网络; 自组织; 可扩放性; 随机相遇博弈; 激励机制

中图分类号: TN393 **文献标识码:** A **文章编号:** 0372-2112 (2006) 11-2081-04

An Incentive Model Adapted to Self-Organizing Management Mode for Resource Sharing in P2P Networks

LIU Ye, YANG Peng, LIU Lir feng, XIA Qin

(The Key Laboratory of Computer Network and Information Integration, Ministry of Education, Southeast University, Nanjing, Jiangsu 210096, China)

Abstract: The self organizing management mode in P2P networks leads to a large amount of selfish behaviors among peers. The corresponding solutions to this problem could hardly keep the merits of P2P network simultaneously, such as self organization or dynamic scalability. Most of the proposed reputation management models or incentive ones use the flooding mechanism to learn historical behavior information of other peers, which causes excessive incurred packets, and thus limits the dynamic scalability. A novel incentive framework named as ResP2P based on random matching games theory is given in the paper. Peer reputation and its renewal mechanism, along with some essential social norms are introduced in ResP2P model, which stimulates rational peers to maximize their own utility and contribute their free resource. Experiments have verified the validity and efficiency of the incentive mechanism. The relative distributed algorithm can easily be deployed in a P2P networks and satisfied with self organization and scalability.

Key words: P2P networks; self organizing; scalability; random matching games; incentive mechanism

1 引言

P2P 网络中的自组织管理模式 (Chord, Pastry, SkipNet), 一方面在适应网络动态性的同时使得节点规模得以大幅提升, 另一方面也使得节点的自私行为大量存在. P2P 网络中由于没有权威第三方的控制, 节点的参与都是依靠自愿的原则, 节点总是希望能尽可能多的利用网络中的资源, 而对系统的贡献则很少, 比如提供很少的存储资源、计算资源等, 这样必然导致整个系统的可用资源存在很大的变数, 因此需要研究相应的机制来约束节点的自私行为, 从而保证整个网络性能的稳定. 约束节点行为的解决方案大致上有两类: (1) 信任管理模型, 通过获取所有节点的历史行为记录, 推导出节点的全

局信用度, 算法涉及所有节点历史行为记录的遍历及节点全局信用的推导过程. 具体实现中, 若是通过引入可信第三方进行操作的方式, 一方面破坏了 P2P 网络本身所拥有的自组织特性, 另一方面该超级第三方节点成为新的瓶颈, 限制了网络的规模; 若是通过泛洪方式, 则其所产生的庞大通信代价及低效率亦不适合规模较大的 P2P 网络环境; (2) 基于博弈的激励机制, 通过考察节点当前所有的行为策略对其收益函数的影响来决定策略的选择, 当前已经存在不少激励机制的研究.

文献 [1] 给出了一个 P2P 文件共享网络中的“囚徒困境”模型, 提出了一些建立激励机制的可行性思路, 不过并没有给出适应自组织管理模式的分布式算法. 文献 [2] 给出了一个 P2P 资源共享网络的非合作进化博弈模型, 节点通过计算收

益函数来决定是否加入 P2P 网络, 以及加入系统后贡献多少资源才能最大化其收益. 模型中通过泛洪机制来获取其它节点的资源贡献向量. 由于算法涉及所有节点的遍历工作, 因此该模型不适用于规模较大的 P2P 网络, 且均衡点的收敛过程是非常缓慢的. 当网络动态性比较明显的时候, 该模型容易造成网络可用资源整体性能的振荡. Feldman 和 J. Chuang 在 2004 年提出了利用非合作博弈道德风险理论中隐藏行动模型的相关方法, 通过同报文转发中继节点来制定契约^[3], 以此来约束多跳路由报文转发中 P2P 网络节点的自私行为. 该方法存在的缺点是报文发送端必须明确了解报文转发中继节点序列, 在该前提条件下才能同每一个报文转发节点制定契约, 这使得该类激励模型的应用范围受到了限制. Blanc 等人将多参与人非合作重复博弈理论中的随机相遇博弈模型应用到报文路由转发的节点激励机制的研究^[4]中, 该研究最大的不足在于没有考虑节点转发能力的差异性, 同时也没有提出具体的适应自组织管理模式的分布式算法.

当前国内外大部分研究只是考虑了网络节点的自组织及其动态性, 并没有过多考虑激励模型对 P2P 网络规模可扩充性的影响. 本文针对当前相关模型中存在的不足, 在 P2P 网络参与节点是理性且自私的假设条件下, 通过引入节点信誉及信誉恢复区分机制, 提出一种基于非合作重复博弈理论的资源共享激励模型 ResP2P, 模型中通过制定相关社会规范来约束节点的自私行为, 节点行为若偏离规范则其信誉将受损害, 从而不能使用其他节点所贡献出来的资源, 以此来激励理性节点为使其自身收益最大化而不发生偏离社会规范的行为, 即: 节点某一历史阶段的投机行为将导致其信誉受损, 并将在后续各个阶段受到与其互享资源的节点的持续的惩罚, 直至该节点度过名誉恢复期, 最终重新达到一个均衡状态. 模型中同时实现了基于节点资源贡献区分的信誉恢复机制, 即所贡献的所属资源愈多, 则信誉恢复愈快. 节点信誉的存取通过分布式哈希表 DHTP2P 的方式实现, 即节点信誉的管理是由网络中所有节点分担的, 维护了 P2P 网络的自组织特性及规模的可扩充特性.

2 P2P 资源共享网络中的“囚徒困境”分析

本文的分析场景为 P2P 网络资源共享应用, 设 $S = \{Sh, Un\}$ 表示节点可供选择的策略集合(共享, 不共享), 任意两节点在某一时间段的资源互享中收益函数的定义如表 1 所示. 解释如下: P2P 网络中任一节点 i 向对方 j 提供所属资源, 则存在资源消耗及资源使用状态的信息维护消耗等, 表示为 $-\beta_i \cdot R_i$, 若节点 i 使用节点 j 所贡献出来的资源, 则存在资源使用收益 R_j .

表 1 某一时间段资源互享收益函数

节 点	节点 j		
	策略	Sh	Un
i	Sh	$(R_j - \beta_i \cdot R_i, R_i - \beta_j \cdot R_j)$	$(-\beta_i \cdot R_i, R_i)$
	Un	$(R_j, -\beta_j \cdot R_j)$	$(0, 0)$

为使得收益函数更具有现实意义, 特作如下假设:

假设 1 \forall 节点 i, j , 不等式 $R_j - \beta_i \cdot R_i \geq 0$ 成立. 即: 在 P2P

资源共享网络中, 节点 i 在不同阶段可能是与不同的节点进行资源互享, 从节点 i 的角度出发, 若是所有节点都采用策略 Sh , 此时节点 i 的单阶段收益期望为 $R_j - \beta_i \cdot R_i$, 该值至少不应小于零, 否则, 节点加入 P2P 网络是没有任何意义的.

假设 2 若 $R_i > R_j$, 有 $\beta_i \cdot R_i > \beta_j \cdot R_j > 0$. 即: 节点所贡献的资源越多, 则该节点消耗越大.

由表 1 知, P2P 资源共享网络中任一节点在策略选择时最终会陷入集体理性和个体理性相矛盾的“囚徒困境”. 在没有有效约束的自组织管理模式下, 节点总是具有机会主义倾向, 因此, 不协作策略 Un 是其最佳选择, 即只想占用其它节点资源, 而自身不愿作出任何贡献, 以期其在某一时间段的收益最大化. 但是, 考虑到节点加入 P2P 资源共享网络的目的是为了更为广泛的利用整个 Internet 中的各类资源, 资源来源于各个参与节点, 而节点间的不协作策略选择使得 P2P 网络的整体收益达到最差, 这与 P2P 网络最初的研究目的是相背离的. 上述“囚徒困境”的形成是由于仅考虑了一个时间段的节点收益情况, 而在实际应用中, 参与节点的生命周期应该是长期的, 所以在制定节点收益函数的过程中应考虑长期的收益情况. 在诸多关于多参与人的非合作重复博弈理论的相关研究中, Kandori M. 在 1992 年提出的随机相遇博弈模型^[5]占有极其重要的地位. 本文即是通过随机相遇博弈理论进行扩展, 并在此基础上建立了资源共享激励模型 ResP2P.

3 ResP2P 资源共享激励模型

假设 P2P 资源共享网络中节点的加入、退出是随机的, 节点并不清楚自身会在哪一阶段离开 P2P 网络, 则该过程属于随机结束重复博弈模型, 它等同于无限次重复博弈; 节点进行行为策略选择时是理性且自私的; 节点所拥有的资源满足幂律分布, 亦称作 Zipf 分^[6], 即节点所拥有的资源在数量上是存在差异的; 节点设置对方节点信誉的过程中不存在欺骗行为. 为方便讨论, 我们将 P2P 网络的生命周期切分成多个连续的时间段, 假定节点在每一轮时间段中只与网络中的一个节点进行资源互享, 即同时以客户端(使用其它节点资源)、服务器端(提供自身资源)的类型出现; 在每一轮交易过程中, 节点所遇到的交易节点在 P2P 网络中是随机的.

设 $Z = \{0, 1, \dots, T\}$ 表示节点在共享资源过程中的信誉状态集合, 0 代表节点信誉守信状态, 非零状态代表节点失信, 数值标明了节点所处的信誉恢复程度; R 为节点所贡献的资源变量; $\lambda = \log\left(\frac{R}{R_j}\right)$ 表示节点 i, j 资源量的等级区分度, 其中参数 a 为节点间资源等级区分的粒度. 网络中任一节点在某一历史阶段的状态可表示为三元组 $\langle z, s, R \rangle$, 其中 $z \in Z, s \in S$.

定义 1 (行为社会规范). 若节点在某一阶段遵从公式 (1) 选取资源共享策略, 则称该节点在该阶段资源互享过程中行为友好, 反之则称该节点行为偏离.

$$\sigma(\langle z_i, z_j \rangle) = \begin{cases} (Sh_i, Sh_j) & \text{if } z_i = 0, z_j = 0 \\ (Un_i, Sh_j) & \text{if } z_i = 0, z_j \neq 0 \\ (Sh_i, Un_j) & \text{if } z_i \neq 0, z_j = 0 \\ (Un_i, Un_j) & \text{if } z_i \neq 0, z_j \neq 0 \end{cases} \quad (1)$$

定义 2 (信誉转换规则). 设 $t-1, t$ 表示两个连续的时间段, 若 $z_i(t-1) = 0$, 且此节点在 t 阶段行为友好, 则 $z_i(t) = 0$; 反之 $z_i(t) = 1$. 若 $z_i(t-1) \neq 0$, 且此节点在 t 阶段继续行为偏离, 则 $z_i(t) = 1$; 若 t 阶段行为友好, 则依据公式(2)设置节点信誉状态.

$$z_i(t) = \begin{cases} z_i(t-1) + \lambda + 1 & \text{if } \{z_i(t-1) + \lambda + 1\} < T + 1 \\ 0 & \text{if } \{z_i(t-1) + \lambda + 1\} > T + 1 \end{cases}$$

$$\text{其中 } \lambda = \begin{cases} 0 & \text{if } \lambda \leq 0 \\ \lambda & \text{if } \lambda > 0 \end{cases} \quad (2)$$

在信誉转换规则中, 节点某一历史阶段的偏离行为, 需要经历一个信誉恢复阶段, 才能恢复至信誉守信状态. 在信誉恢复期中, 节点只有行为良好, 才能不断使得其信誉向清白靠拢. 由公式(1)可以看出, 在信誉恢复期, 该节点与其他信誉清白的节点互享资源的时候, 必须不断向对方提供资源, 而“没资格”占用对方资源, 即必须经历了一个 T 时间跨度的惩罚期. 在信誉恢复期, 若节点再次出现行为偏离, 则重置 T 时间跨度的惩罚过程, 即本信誉转换规则中对节点的偏离行为的惩罚不累积. 信誉转换规则中参数的变化使得信誉恢复规则具有按照节点贡献资源量进行区分的能力.

假设 3 设 \bar{R} 表示整个 P2P 网络中节点所拥有资源量的平均值, \forall 节点 i , 不等式 $\bar{R} - \beta_i \cdot R_i \geq 0$ 成立.

在推论 1 的证明中, 我们将假设 1 的约束条件放宽为假设 3. 节点 i 在不同阶段可能是与不同的节点进行资源互享, 若是所有节点都采用策略 Sh , 节点 i 单阶段收益期望为 $\bar{R} - \beta_i \cdot R_i$, 该值至少不应该小于零, 否则, 节点加入 P2P 网络是没有任何意义的.

推论 ResP2P 激励模型中, 理性节点考虑其长期收益, 最大化其长期平均收益函数的最优行为策略是依据行为社会规范中行为友好的规定进行策略选择. (证明略)

在信誉转换规则中, 后续惩罚时间跨度由参数 T 决定, 当 T 较大的时候, 节点出于自身收益的考虑, 是不会轻易选择使得行为偏离的策略的. 对于在某一阶段资源互享过程中的任意节点 i 、节点 j , 假设 $R_i > R_j$, 则有 $R_j - \beta_i \cdot R_i < R_i - \beta_j \cdot R_j$, 节点 j 的收益明显是大于节点 i 的收益的, 因此在公式(2)中, 通过 $\lambda = \log_a(R_i/R_j)$ 来区别对待提供不同量级资源的节点, 以此来弥补不同量级的节点在互享过程中的收益差, 资源量级高的节点在惩罚恢复期内, 若行为友好, 则存在较大的恢复步长, 此信誉恢复区分机制的目的也在于鼓励系统中的节点多贡献自身的闲置资源. 节点信誉恢复区分机制在一定程度上弥补了节点互享过程中的收益差的同时, 也引入了新的问题, 即节点在行为偏离后, 临时增加其共享资源, 以缩短信誉恢复时间, 待信誉恢复后, 则继续提供较少的共享资源. 为解决这一问题, 我们设定如下补充规则:

规则 1 节点所提供的共享资源在连续的时间段内, 考察其变化率是否超过一定的阈值. 若节点处于信誉恢复期且其资源变化率超过阈值, 则将公式(2)中参数 λ 的计算过程替换为:

$$\lambda = \left\{ \lambda, \frac{1}{\lambda} \right\}_{\min}$$

规则 2 准入条件, 对于欲加入 P2P 资源共享网络的节

点, 其所共享出来的资源量必须要超过一个底限值.

规则 1 考虑到在 P2P 资源共享网络中, 节点所提供的共享资源在连续的时间段内, 一般不会出现很大的变化. 对于一个理性的节点来讲, 是不会在信誉恢复期, 减少其自身的共享资源, 毕竟在此种情况下会延长其信誉恢复速度; 若节点处于信誉清白状态, 则不影响整个模型的稳定性. 规则 2 的制定, 则是为了避免在式子 $\lambda = \log_a(R_i/R_j)$ 中出现分母无限小的情况.

4 ResP2P 激励模型分布式协议

分布式协议的主要任务是在自组织管理的网络环境中完成相应的节点信誉信息的存取操作. Reputation Get(int node_{id}) 分布式哈希表中基本读操作; Set(int node_{id}, Reputation node_{id}. rep) 分布式哈希表中基本写操作; 伪代码中用到的另外几个主要变量包括, rep 为节点的信誉变量, str 为节点所采取的资源共享策略变量.

// 设节点 i 在某一阶段与节点 j 互享资源, 节点 i 中 ResP2P 激励模型分布式协议代码如下:

(1) i . repGet(i); j . repGet(j);

(2) select_strategy_bySocialNorm(i . rep, j . rep); // 节点遵从定义 1(行为社会规范)依据双方节点当前的信誉值, 选择相应的行为友好的资源共享策略.

(3) share_resource_byStrategy(i , j , i . str, j . str); // 节点按照所选择的策略进行资源互享, 并同时得到相关收益.

(4) repCalculation(j . rep); // 根据收益结果检查节点是否行为友好, 并根据定义 2(信誉转换规则)重新计算出对方节点的信誉.

(5) set_reputation(j , j . rep); // 利用 DHT 写操作, 在 P2P 网络中设置节点 j 新一阶段的信誉值.

相应地, 节点 i 在新一阶段的信誉值由节点 j 设置. 在整个信誉读写过程中, 信誉信息由不相关的第三方节点管理, 避免了节点本身设置其信誉的安全隐患. Chord 路由算法本身所具有的良好性能特性保证了系统可以在有限的开销内完成信誉存取操作.

5 试验及其结论

试验场景设计如下: P2P 网络中节点所拥有的资源呈现 Zipf 分布^[7], 节点在加入 P2P 网络的整个生命周期中进行资源互享, 以从中获得收益, 节点并不知道什么时候离开 P2P 网络. 本试验过程主要目的是为了考察在 P2P 资源贡献激励模型中引入信誉机制的有效性、信誉区分机制对节点信誉恢复速度的影响. ResP2P 激励模型参数设置: 资源区分参数; 惩罚期跨度.

图 1 直观地反映出了节点收益变化(试验持续时间为 100 个连续的时间段), 较之于遵从社会规范行为友好的收益, 节点除了在某个特定的行为偏离阶段存在短期“获利”外, 在后续阶段, 其偏离行为受到的惩罚是持续的. 图 1(a) 中节点的贡献资源量为 1, 节点每次出现行为偏离后, 需经历一个 ResumPeriod_a = 20 的信誉恢复期; 图 1(b) 中节点的贡献资源量为 20, 节点信誉失信后, 需经历一个 ResumPeriod_b ∈ [4, 6] 不

等的信誉恢复期,由于整个网络节点资源呈现 Zipf 分布,该节点信誉恢复期的数学期望为 $E(\text{ResumePeriod}_i) = 4$.

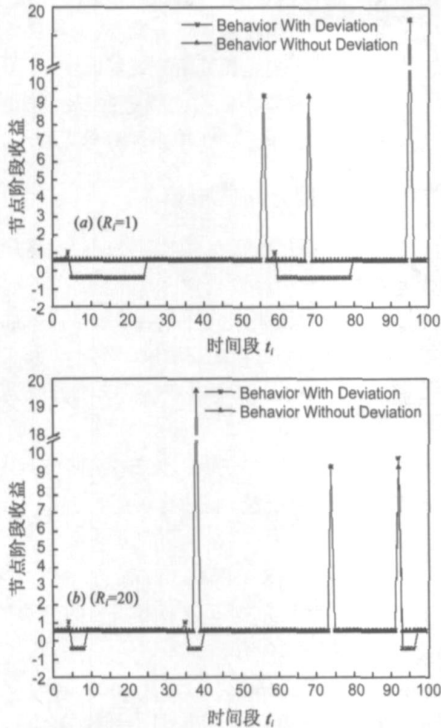


图1 节点行为偏离/行为友好~收益变化图

导致理性节点信誉失信的主要原因有两个方面:(1)节点自身策略选择,即选择了使得行为偏离的策略;(2)由于网络的不稳定因素,例如,由于网络拥塞导致网络的一段时间不可达,使得资源互享过程失败,从而导致资源提供节点的信誉受损.节点信誉恢复速度与参数 T 相关,即整个网络对偏离行为的惩罚力度随着 T 的增加而增大.随着 T 的增大,节点出于自身长期收益最大化的考虑,是不会轻易选择行为偏离策略的,另一方面, T 的增加也导致了由于网络不稳定因素而引起的节点信誉失信的修正速度,从而使得网络的整体收益再次达到稳定值的恢复时间延长.图2给出了节点资源贡献量对其信誉恢复周期的影响情况,由图1知节点在信誉恢复期

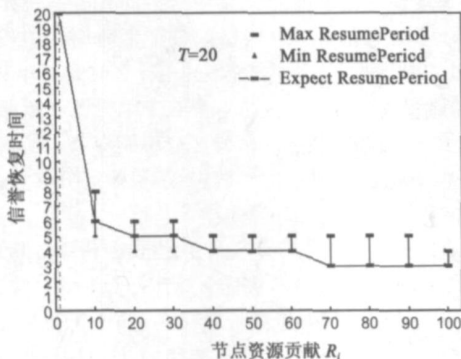


图2 节点信誉恢复周期/资源贡献量~变化关系图

间的收益是非常差的,尽快度过信誉恢复期是每个理性节点的最佳选择,因此信誉恢复区分机制的实施能够鼓励节点多贡献自身的可用资源.

6 结束语

理论和试验结果表明,在节点是理性且自私的假设条件下,ResP2P 资源共享激励模型是有效的,因为对于理性的参与节点来说,能够使长期收益最大化的策略是不偏离 ResP2P 模型中规定的社会行为规范,同时为了加快其信誉恢复速度应尽可能多地贡献其可用资源.本论文的资源共享激励模型适用于资源无差异的应用环境,诸如:计算资源(CPU)、存储资源、带宽资源等,而对于内容相关的文件资源等,则需要进一步研究节点间缘于内容兴趣度不同而导致的效用函数的差异.现实的 P2P 网络环境中,节点的生命周期是多样的,如何建立一套同时适应不同长短生命周期节点的统一激励机制,将是进一步的研究重点.

参考文献:

- [1] K Ranganathan, M Ripeanu, A Sarin, I Foster. To Share or not to Share An Analysis of Incentives to Contribute in File Sharing Environments [A]. International Workshop on Economics of Peer to Peer Systems [C]. Berkeley, CA, 2003. 151-155.
- [2] C Buragohain, D Agrawal, S Suri. A game theoretic framework for incentives in p2p systems [A]. In Proc. 3rd Intl. Conf. on Peer to Peer Computing [C]. Los Alamitos, CA: IEEE Computer Society Press, 2003. 48-56.
- [3] M Feldman, J Chuang. Hidden Action in Multi-Hop Routing [A]. 2nd Workshop on Economics of Peer to Peer Systems [C]. Cambridge, Mass: Harvard University, June 2004. 117-126.
- [4] A Blanc, Y Liu, A Vahdat. Designing Incentives for Peer to Peer Routing [A]. 2nd Workshop on Economics of Peer to Peer Systems [C]. Cambridge, Mass: Harvard University, June 2004. 374-385.
- [5] Kandori M. Social Norms and Community Enforcement [J]. The Review of economic studies. 1992, 59 (1): 61-80.
- [6] Jacky Chu, Kevin Labonte, Brian Neil Levine. Availability and Locality Measurements of Peer to Peer File Systems [A]. In Proc. ICom: Scalability and Traffic Control in IP Networks II Conferences [C]. Boston, MA, July 2002. 310-321.
- [7] L A Adamic, B A Huberman. Zipf's law and the Internet [J]. Glottometrics, 2002, 3: 143-150.

作者简介:

刘业男, 1977年生, 江苏建湖人, 博士研究生, 主要研究方向为高性能网络、对等网络、分布对象计算等. E-mail: yliu@seu.edu.cn

杨鹏男, 1975年生, 博士研究生, 主要研究方向为新一代网络体系结构、分布计算环境等.