

# 基于无限重复博弈的 P2P 网络信任模型研究\*

王保玉<sup>1</sup>, 高承实<sup>1</sup>, 戴青<sup>1</sup>, 陈景林<sup>2</sup>, 刘洋<sup>3</sup>

(1. 解放军信息工程大学, 郑州 450004; 2. 舟山警备区, 浙江 舟山 316000; 3. 济南军区司令部, 济南 250000)

**摘要:** 为解决 P2P 电子商务环境中存在的安全问题, 针对现有一些信任模型的局限性, 提出一种基于无限重复博弈理论的信任模型。通过建立奖惩策略, 对节点的不合作行为, 依据博弈策略对节点进行惩罚, 同时对选择合作策略的节点进行奖励, 并根据欺诈行为次数设定不同的惩罚周期。理论分析和仿真实验表明, 合作策略会成为节点博弈的帕累托最优策略, 该策略模型能有效遏制和惩罚恶意行为, 提高节点诚实交易的积极性, 从而有效提高网络环境的安全性和稳定性。

**关键词:** P2P 网络; 信任模型; 重复博弈; 奖惩策略

**中图分类号:** TP393      **文献标志码:** A      **文章编号:** 1001-3695(2013)09-2802-03

**doi:**10.3969/j.issn.1001-3695.2013.09.061

## Research on trust model based on infinitely repeated games theory in P2P networks

WANG Bao-yu<sup>1</sup>, GAO Cheng-shi<sup>1</sup>, DAI Qing<sup>1</sup>, CHEN Jing-lin<sup>2</sup>, LIU Yang<sup>3</sup>

(1. PLA Information Engineering University, Zhengzhou 450004, China; 2. Zhoushan Garrison, Zhoushan Zhejiang 316000, China; 3. Jinan Military Area Command, Jinan 250000, China)

**Abstract:** To resolve the security problems in P2P networks, this paper proposed a trust model based on infinitely repeated game theory. Reward and punishment strategy in this paper not only punished the malicious behavior, but also incentivized the cooperation nodes, and the duration of penalty was determined in the times of fraudulent conduct. Mathematic analysis and simulation experiments show that the strategy of cooperation will be Pareto optimal strategy, which can punish and stifle available malicious behavior, and improve incentives of the nodes trading honestly, then the security and stability of the P2P network environment will be enhanced efficiently.

**Key words:** P2P networks; trust model; repeated game; incentive strategy

近年来,对等网络(P2P)因其开放性及其节点的匿名和高度自治的特性得到了广泛的应用和发展。与传统的网络模式相比,P2P的交互模式更直接、有效,它覆盖的面积也更广,提供的交易资源也更充足。然而随着P2P网络应用的日益普及,这些特点也导致了交易过程中存在大量的恶意行为和安全隐患,严重影响了P2P网络的服务质量。所以,如何建立有效的信任模型来提高系统整体可靠性,正逐渐成为P2P网络安全性能研究的焦点。P2P网络信任机制的研究中,传统的信任模型大多基于节点的历史行为,对节点的信誉值进行量化分析,从而计算节点的全局信任度,如EigenTrust<sup>[1]</sup>、PeerTrust<sup>[2]</sup>等信任模型,这在一定程度上反映了节点的诚信程度。但在现实生活中,交易双方面对的往往是复杂的网络环境,这种只分析单个节点的信任值,不仅开销大,而且实施起来较为困难,因此存在着很大的局限性。

基于博弈论<sup>[3]</sup>的信任模型正是基于这些原因提出的。博弈论是研究个体和群体选择最优行为策略的有效数学工具,而对等网络是社会群体在网络中的反映,每个节点都是P2P网络群体下的个体,它们的行为也具有社会个体的很多特性,因此对等网络中的节点行为亦可以采用博弈论来进行分析。利

用博弈理论分析节点行为策略的选择,即是从整个P2P网络的角度来分析节点行为策略选择之间的相互关系。

文献[4]给出了P2P文件共享网络中的“囚徒困境”博弈模型,提出了建立激励机制的可行性思路,但并没有给出具体的算法。文献[5]提出基于贝叶斯博弈的节点概率选择策略,对信息未知节点的策略选择提供了有效方法,但其计算复杂度较大。因此,本文提出了一种基于无限次重复博弈的P2P网络节点行为信任策略模型,利用重复博弈的贴现率和对节点行为的奖惩机制来分析节点的行为策略选择。

### 1 信任博弈模型

#### 1.1 模型的基本假设

P2P网络系统具有动态、连续、不确定的系统状态,每个节点之间的关系是对等的,节点可以自由地进行交互,而节点进行交互协作的基础是信任。所以,信任策略的选取既是为了保证节点自身利益,同时也是为了保证网络环境的安全性和稳定性。为便于对节点策略分析研究,本文作以下假设:

- 假设1 交易节点行为是理性的,且追求自己利益最大化。
- 假设2 节点在交易之前认为对方节点在博弈中不存在

**收稿日期:** 2012-11-15; **修回日期:** 2013-01-04      **基金项目:** 国家部委基金资助项目(9140c130103120c13062)

**作者简介:** 王保玉(1986-),男,河南周口人,硕士研究生,主要研究方向为网络安全(wby8697@126.com);高承实(1973-),男,硕士,博士,主要研究方向为网络安全、网络计算;戴青(1963-),男,副教授,硕士,主要研究方向为网络安全、多媒体应用技术;陈景林(1985-),男,助理工程师,主要研究方向为信息安全;刘洋(1985-),男,助理工程师,主要研究方向为信息安全。

欺诈行为。

假设 3 节点交易遵循相同的规则和过程。

假设 4 节点对自己何时离开 P2P 网络环境无法预知,即不清楚博弈过程何时终止,所以从博弈策略的角度来看,此博弈构成无限次重复博弈。

### 1.2 模型的相关定义

**定义 1** 节点策略。P2P 网络节点交易过程中,节点策略集  $S = \{ \text{合作}, \text{不合作} \}$ , 简记为  $\{ \text{Co}, \text{Do} \}$ 。节点  $i$  在第  $t$  阶段的策略记为  $S_i$ 。

**定义 2** 博弈  $G$  重复进行  $T$  次,并且每次重复进行  $G$  之前,每个参与人都观察到以前博弈的结果,这种博弈过程称为  $G$  的  $T$  次重复博弈,记做  $G(T)$ 。其中, $G$  称做  $G(T)$  的原博弈,从而无限次重复博弈就是  $G(\infty)$ 。

**定义 3** 重复博弈  $G(T)$  中,尽管每个参与人都是理性的、前瞻的,但是他们并不是充分有耐心的。因此引入贴现因子  $\delta$ , 如果节点  $i$  在第  $t$  阶段的收益为  $g_i(t)$ , 那么  $T$  次重复博弈中它的总收益就是  $g_i = \sum_{t=1}^T \delta^{t-1} g_i(t)$ , 无限重复次博弈中,总收益为

$$g_i = \sum_{t=1}^{\infty} \delta^{t-1} g_i(t) \quad (1)$$

**定义 4** 重复博弈中的触发策略是指参与人之间首先合作,如果在某一阶段对方不合作,则自己也用永远不合作来报复或者惩罚对方。触发策略是重复博弈能够实现合作和提高均衡效率的关键机制。

### 1.3 重复博弈信任模型策略的描述

节点博弈策略选择的优劣体现在节点的收益结果,即节点通过此种策略选择获得的好处。而 P2P 系统中节点的长期交易行为可描述为其中任一对节点之间阶段博弈行为的无限次重复博弈。设节点阶段博弈  $G$  的收益矩阵如表 1 所示。

表 1 博弈收益矩阵

		节点 2	
		Co	Do
节点 1	Co	$\pi, \pi$	$-\omega, \bar{\omega}$
	Do	$\bar{\omega}, -\omega$	$0, 0$

其中: $\bar{\omega} > \pi > \omega > 0, \pi$  为一次博弈中双方都选择合作策略时的收益;当一方策略为合作,而另一方的策略为不合作时,  $-\omega$  为选择合作策略节点的收益,  $\bar{\omega}$  为选择不合作策略节点的收益;当双方都选择不合作策略时,交易无法完成,故双方收益均为 0。

## 2 模型的证明和分析

### 2.1 模型策略的证明

在经典的“囚徒困境”中,存在唯一的 Nash 均衡为博弈双方都选择不合作策略。但如果博弈无限地进行下去,博弈双方都会采用相互合作的策略进行博弈,合作策略成为了唯一的 Nash 均衡,这在经济学中的非合作重复博弈理论中被称为子博弈精练 Nash 均衡,是得到证明存在的。

**定理 1** 如果阶段博弈  $G$  有唯一的 Nash 均衡,那么对任意有限次  $T$ , 重复博弈  $G(T)$  有唯一的子博弈完美结局:在每一阶段取  $G$  的 Nash 均衡策略。

**定理 2** 完全信息静态博弈  $G$  中,  $(e_1, e_2, \dots, e_n)$  和  $(x_1, x_2, \dots, x_n)$  分别表示子博弈完美 Nash 均衡的得益和可实现的任意得益。如果  $x_i > e_i$  对  $\forall i \in N$  都成立,则必定存在  $\delta \in (0, 1)$ , 当  $\delta$  充分接近 1 时,无限次重复博弈  $G(\infty, \delta)$  一定存在子博弈完美 Nash 均衡。

该定理表明,在无限次重复博弈中,如果参与人有足够的耐心,那么对双方均有利的合作均衡则可以作为子博弈完美 Nash 均衡出现。也即只要参与人之间交易结束的概率足够小,交易具有长期性,且交易双方对未来收益的贴现足够高,则双方将从长期利益出发来维持相互合作,从而实现在静态博弈中无法实现的合作解。下面给出 P2P 网络中节点间无限重复博弈中合作策略为子博弈精练 Nash 均衡的具体证明。

**证明** a) 首先验证合作策略是否为子博弈完美均衡。只须验证如下几个方面:在合作状态,节点双方任何一方偏离将不会给它带来好处;一旦节点其中一方偏离合作,选择欺诈行为,它必定会为了极大化自己利益而在受到惩罚之后回到合作状态;如果节点一方偏离,节点另一点采取惩罚才对自己有利,防止共谋行为的出现。

(a) 假设节点 1 偏离诚信状态,那么它将承受一个周期的惩罚,倘若折扣因子充分地接近于 1,对于节点 1 来说,偏离诚信不是一个正确的策略。因为若不偏离,两个周期的盈利和应当为  $\pi + \pi\delta$ ;若节点 1 偏离,它先获得  $\bar{\omega}$ , 然后受到一个周期的惩罚,这两个周期的盈利和为  $\bar{\omega} - \omega\delta$ , 当  $\delta$  比较大时,即当  $\delta > \frac{\bar{\omega} - \pi}{\omega + \pi}$  时,有  $\pi + \pi\delta > \bar{\omega} - \omega\delta$ 。

(b) 若节点 1 在  $k$  阶段选择偏离,如果它甘愿受惩罚并在之后回到合作状态,那么它的得益总和为

$$g_1 = \sum_{t=1}^{\infty} \delta^{t-1} g_1(t) = \pi + \delta\pi + \dots + \delta^{k-1} \pi + \delta^k \bar{\omega} - \delta^{k+1} \omega + \delta^{k+2} \pi + \dots = \frac{\pi}{1-\delta} + \delta^k (\bar{\omega} - \pi) - \delta^{k+1} (\omega + \pi) \quad (2)$$

如果节点 1 偏离策略剖面的给定,如它坚持再取欺诈一次,那么它面临的结局当为  $\{ \text{Do}, \text{Do} \}, \{ \text{Co}, \text{Co} \}$  与永远的  $\{ \text{Co}, \text{Do} \}$ , 其相应的得益总和为

$$g'_1 = \sum_{t=1}^{\infty} \delta^{t-1} g_1(t) = \pi + \delta\pi + \dots + \delta^{k-1} \pi + \delta^k \bar{\omega} + 0 - \delta^{k+2} \omega + \delta^{k+3} \pi + \dots = \frac{\pi}{1-\delta} + \delta^k \bar{\omega} - \delta^{k+2} \omega - \delta^{k+1} (1 + \delta + \delta^2) \pi \quad (3)$$

$$g_1 - g'_1 = \delta^{k+1} (\delta\omega + \delta\pi - \omega) \quad (4)$$

当  $\delta$  较大时,即  $\delta > \frac{\omega}{\omega + \pi}$ , 显然有  $g_1 - g'_1 > 0$ 。由此可以看出,节点 1 在选择欺诈行为后,它必定会为了极大化自己利益而在受到惩罚后回到合作状态。

(c) 当节点 1 因欺诈而受惩罚时,节点 2 的得益为

$$g_2 = \sum_{t=1}^{\infty} \delta^{t-1} g_2(t) = \pi + \delta\pi + \dots + \delta^{k-1} \pi - \delta^k \omega + \delta^{k+1} \bar{\omega} + \delta^{k+2} \pi + \delta^{k+3} \pi + \dots = \frac{\pi}{1-\delta} + \delta^{k+1} (\bar{\omega} - \pi) - \delta^k (\omega + \pi) \quad (5)$$

若节点 2 不去惩罚节点 1,而节点 1 在欺诈之后立即回到诚信状态,此时节点 2 的得益总和为

$$g'_2 = \sum_{t=1}^{\infty} \delta^{t-1} g_2(t) = \pi + \delta\pi + \dots + \delta^{k-1} \pi - \delta^k \omega + \delta^{k+1} \pi + \delta^{k+2} \pi + \delta^{k+3} \pi + \dots \quad (6)$$

$g_2 - g'_2 = \delta^{k+1} (\bar{\omega} - \pi)$ , 因为  $\bar{\omega} > \pi$ , 显然有  $g_2 > g'_2$ 。由此可见,为极大化自身得益,节点 2 必定会按照策略剖面的规定去惩罚节点 1。

综上可以证明,此策略是子博弈完美 Nash 均衡。

b) 由于此博弈是个无限重复博弈,任何一个阶段开始的子博弈与原博弈的结构相同,因此,此策略构成了无限囚徒困境重复博弈的子博弈精练 Nash 均衡。

### 2.2 相关参数的选取

通过分析表明,在任意时刻  $t$ , 历史无非有三种情况:要么

是双方都合作状态,要么节点 1 欺诈,要么节点 2 欺诈。对应于这三种子博弈的平均收益分别为  $(\pi, \pi), (\delta(\omega + \pi) - \omega, \bar{\omega} - \delta\omega), (\bar{\omega} - \delta\omega, \delta(\omega + \pi) - \omega)$ , 它们之中没有一个帕累托 (Pareto) 优于另外一个,因此,这个策略剖面是 WRP<sup>[3]</sup> (重复博弈的弱不可改进性)。在无限重复囚徒窘境博弈中,得到上述有效 WRP 得益的关键是使用策略组合 {Co, Do} 惩罚恶意节点,而惩罚者从中又得到了奖励。在确定的情况中,惩罚越严厉,越有利于节点选择诚信行为,因为均衡情况下,欺骗从来不会发生,所以惩罚实际上是没有成本的。但在实际生活中,节点可能因为一些客观的原因导致交易无法完成,如果总是触发惩罚,就会冤枉好人,过重的惩罚反而会导致交易无法进行。因此,惩罚程度的选择会对节点的行为产生较大的影响。惩罚既不能太大,也不能过小而达不到惩罚作用。

本文主要从贴现因子  $\delta$  和惩罚周期  $l$  的设定和选择来分析。对于交易过程中首次偏离合作策略或偶尔偏离的节点,在执行惩罚中,选择较短的惩罚周期  $l$ ; 对于经常选择欺诈行为或者周期性选择欺诈行为的节点,  $\delta$  会随之下降,使节点在与之交易过程中更加慎重,并选取较长的惩罚周期  $l$ , 一旦节点有欺诈行为,加大对此恶意节点的惩罚力度。本文设定  $l = 2n - 1$ , 其中  $n$  为节点在交易过程中选择欺诈行为的次数。

### 3 仿真实验

#### 3.1 仿真实验环境

本文在单机环境下实现了模型的仿真,仿真实验的系统环境为 Windows XP SP3 操作系统, CPU 为 AMD Athlon™ II × 2 250 Processor 3.01 GHz, 内存 2 GB。为了测试本模型的有效性,本文利用 MATLAB6.5 软件建立了一个仿真平台,该平台可以模拟大规模的动态 P2P 网络环境中节点之间的交易过程,所用编程语言为 VC++。系统中选择所有节点数为 1 000, 每个节点发起 500 次合作请求,实验以各个节点随机交易开始。

本文中博弈信任模型节点的策略选择是基于理性行为的,但在实际中,由于主观因素和客观原因,节点并不总是选择理性行为,因此,在仿真实验的系统中将节点类型大致分为三类:

- a) 理性节点,这类节点总是坚持博弈均衡策略。
- b) 摇摆型恶意节点,其周期性选择合作策略,在偏离策略剖面时,能友好接受惩罚,但过一段时期又会采取欺诈行为。
- c) 纯粹型恶意节点,这类节点在前期坚持博弈均衡策略,积累较高的信誉,然后在某一时段集中选择欺诈行为。

#### 3.2 博弈均衡策略对不同类型节点的平均收益值影响

当交易过程中发生恶意欺诈行为时,分析比较三类节点的平均收益变化情况。本文假设选取若干个交易周期,将收益矩阵的参数设定为  $\pi = 10, \bar{\omega} = 15, \omega = 5, \delta = 0.8$ , 结果如图 1 所示。在博弈均衡策略中,三类节点在前期表现较好时,其平均收益值基本相同。当面临欺诈行为时,理性节点严格按照策略剖面给定的策略执行,虽然节点收益值会有一定损失,但由于按照奖励策略对其进行了补偿,其平均收益不会产生太大影响,并能很快回升到正常水平。而摇摆节点周期性选择恶意交易,在首次实行恶意交易时,节点的平均收益值会得到提升,但随之而面临的是一定周期的惩罚,节点通过友好配合惩罚使其收益值再次得到提升,当节点缓慢积累了较高信誉时,再次实施欺诈行为,按照惩罚策略,策略会加大对节点的惩罚周期,使其收益值快速下降。而纯恶意节点在实施欺诈行为后,虽然收益值得到提升,但由于其不按照策略剖面接受惩罚,收益值会

迅速下降,信誉也随之越来越低,最终节点无法在网络环境中生存下去。

#### 3.3 恶意节点对交易成功率的影响

在不同恶意节点比例下比较本文的博弈信任模型与 EigenTrust 模型和 PeerTrust 模型中的交易成功率,结果如图 2 所示。使用本文模型在恶意节点增加的情况下,交易成功率仍能维持在较高水平,而其他两个模型会随着恶意节点比例的增加对交易成功率产生较大的影响。这是因为,本文模型中引入了奖惩策略,恶意节点如若选择偏离合作策略,就会遭到网络中其他节点的惩罚,且惩罚周期会随着恶意交易次数的增加而增长,如果节点一意孤行,继续偏离策略剖面,它将会永远找不到合作节点。

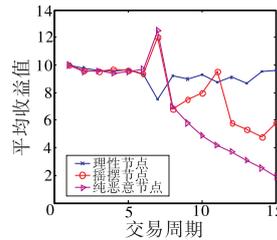


图1 三类节点随交易周期变化的平均收益值比较

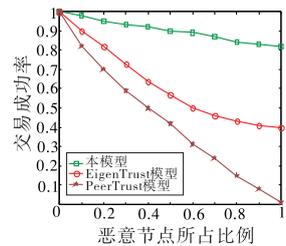


图2 不同模型下交易成功率变化对比

### 4 结束语

本文基于无限重复博弈理论的方法,提出了一种基于奖惩策略的 P2P 网络信任模型,对网络中节点行为策略进行了分析、比较和验证,有效避免了节点的恶意欺诈行为,同时在一定程度上避免了节点的共谋行为。仿真实验表明,该模型能有效提高系统的整体安全性和稳定性。当然,本文中对奖励程度的设定没有具体的分析和讨论,有待进一步深入研究。

#### 参考文献:

- [1] KAMVAR S D, SCHLOSSER M T, GARCIA-MOLINA H. The EigenTrust algorithm for reputation management in P2P networks[C]//Proc of the 20th International World Wide Web Conference. 2003: 640-651.
- [2] LI Xiong, LIU Ling. A reputation-based trust model for peer-to-peer commerce communities [C]//Proc of IEEE Conference on e-Commerce. New York: IEEE Computer Society, 2003: 275-284.
- [3] 施锡铨. 博弈论[M]. 上海: 上海财经大学出版社, 2005: 132-253.
- [4] RANGANATHAN K, RIQEMU M, SATIN A, et al. To share or not to share: an analysis of incentives to contribute in file sharing environments [C]//Proc of International Workshop on Economics of Peer-to-Peer Systems. 2003.
- [5] 李致远, 王汝传. 一种移动 P2P 网络环境下的动态安全信任模型 [J]. 电子学报, 2012, 40(1): 1-7.
- [6] BURAGOHAIM C, AGRAWAL D, SURI S. A game theoretic framework for incentives in P2P systems [C]//Proc of the 3rd International Conference on Peer-to-Peer Computing. Washington DC: IEEE Computer Society, 2003: 48-56.
- [7] MARINA N. A note on the nucleolus and the kernel of the assignment game [J]. International Journal of Game Theory, 2004, 33(1): 55-65.
- [8] 徐海滔, 郑相全, 齐守青, 等. 一种新的基于博弈论的 P2P 激励机制 [J]. 计算机应用研究, 2008, 25(9): 2787-2789.
- [9] 张妮妮, 陈绥阳, 余洋. 基于博弈论的 P2P 激励机制 [J]. 计算机工程, 2011, 37(15): 89-91, 102.
- [10] AHMED A F. Classifying and tracking free riders in multimedia-based systems [J]. Journal of Universal Computer Science, 2010, 16(10): 1368-1387.
- [11] ANGEL E, BANPES E, BLIN L, et al. Fair cost-sharing methods for the minimum spanning tree game [J]. Information Processing Letters, 2006, 100(1): 29-35.