ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.21, No.12, December 2010, pp.3042−3055
doi: 10.3724/SP.J.1001.2010.03717
© by *Institute of Software*, *the Chinese Academy of Sciences*. All rights reserved.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

虚拟计算环境中基于重复博弈的惩罚激励机制*

桂春梅+, 蹇 强, 王怀民, 吴泉源

(国防科学技术大学 计算机学院,湖南 长沙 410073)

Repeated Game Theory Based Penalty-Incentive Mechanism in Internet-Based Virtual Computing Environment

GUI Chun-Mei+, JIAN Qiang, WANG Huai-Min, WU Quan-Yuan

(School of Computer Science, National University of Defense Technology, Changsha 410073, China)

+ Corresponding author: E-mail: plantsperfum@yahoo.com.cn

Gui CM, Jian Q, Wang HM, Wu QY. Repeated game theory based penalty-incentive mechanism in Internet-based virtual computing environment. *Journal of Software*, 2010,21(12):3042−3055. http://www.jos.org.cn/1000-9825/3717.htm

Abstract: In order to construct a trustworthy computing platform for the next Internet, there needs to be normalizing and promoting of autonomic elements in order to have them collaborate actively with on another. A novel penalty-incentive mechanism, named PETrust, based on a repeated game theory, is given in this paper. This paper aims at providing a set of mechanisms, which the behavior of autonomic elements is normalized and is promoted to take the expected strategy. PETrust adjusts the degree of penalty by changing the reputation status. Theoretical analysis and simulation results show that PETrust can distinguish the different features of behavior, effectively, can punish and stifle malicious behavior, can improve the system's entire efficiency, can stimulate autonomic elements' honest trade enthusiasm, and can provide a better capacity of resisting collusive deception. Furthermore, PETrust presents both low time complexity and few incurred packets, which is favourable for engineering deployment and implementation.

Key words: autonomic element; reputation; collaboration; behavior criterion; penalty and incentive

摘 要: 如何促进网络中自主元素自觉规范行为、积极有序协作从而形成"可信的计算平台"是下一代互联网亟需解决的重点问题.提出一种基于重复博弈的惩罚激励机制 PETrust,旨在有效促进自主元素采取系统期望的诚实协作策略进行规范行为.PETrust 根据自主元素信誉特征的变化动态调整惩罚力度.理论分析和实验结果表明,PETrust 能够有效区分自主元素的不同行为特征,遏制和惩罚恶意行为,提高自主元素诚实交易的积极性和系统的整体效率,并具有更好的抵御共谋欺骗的能力.PETrust 还同时具备计算复杂度低、报文通信量小的特点,利于部署实施.

关键词: 自主元素;信誉;协作;行为规范;惩罚激励

中图法分类号: TP393 **文献标识码**: A

* Supported by the National Basic Research Program of China under Grant Nos.2005CB321800, 2005CB321804 (国家重点基础研究发展计划(973))

Received 2008-09-25; Revised 2008-12-18, 2009-03-30; Accepted 2009-07-07

© 中国科学院软件研究所 http://www.jos.org.cn

麦克尼利的预言“网络就是计算机”现在已基本成为现实.互联网业已成为人们生产生活中信息交流协作的重要基础设施.网上银行、网上购物、网上证券等网络服务和交易越来越流行并逐步呈现出取代传统柜台交易的势头.随着网格计算、普适计算、P2P 计算、Ad hoc 网络等分布式计算技术的深入研究和飞速发展,互联网资源共享与访问的一体化动态协作服务平台正逐步形成.

在发展的进程中,一些有趣而必然的现象伴随其中.由于虚拟计算环境中自主元素[2]的自组织特性[3]、理性和自私性,自主元素按照“自愿参与、自主协同”原则参加和构造系统行为,同时本能地恪守尽可能多地享用其他元素的资源和服务,尽可能少地共享自身资源和服务的意愿进行交互,从而追求自身利益最大化.因此在现有网络环境和监管机制下,随意中止服务、取消合作、搭便车、合伙欺骗的现象充斥其中;另一方面,以Gnutella网络应用为例,出于自私心理和规避风险的考虑,网络中66%的节点对整个系统没有任何贡献.值得指出的是,即便是最极端的“信息孤岛”的形式,自主元素仍然要承担自身基本运营维护成本、资源老化等自然损耗.因此,制定和发展一套行之有效的机制,在承认、包容和利用自主元素追求最大利益的自然特点的基础上有效规范和引导其行为,从而构建和谐、安全、透明的网络计算环境,促进互联网由“可能的计算平台”向“可信的计算平台”转变[1],是顺应发展趋势、符合自主元素个体利益追求和网络整体效能最优的有效途径.

1 自主元素可信机制分析

目前,围绕自主元素行为可信方面的研究已取得了很多成果,但对于如何有效促进自主元素采取诚实协作策略进行规范行为的机制性研究仍有不足.基于凭证的信任模型如PolicyMaker[4]在应用中不可避免地出现了一些问题,如:存在忽略信任的主观因素和不确定性因素、可扩放性差、单点失效等;集中式信任模型如Marsh的信任模型[5]、Manchala的信任模型[6]虽然比较简单,易于实现,但存在负载和健壮性问题,在依赖用户反馈的系统中,信任评估的准确性和对欺骗行为的识别能力都非常有限;分布式信任模型如基于Bayesian方法的模型[7],利用Bayesian网络推导节点的可信度,但对经验的可靠性依赖高;基于向量计算机制的Hassan模型[8]引入向量运算机制来描述信任关系,综合考虑自信(confidence)、历史、时间等因子来反映信任关系的动态性,但是向量机制的信任度计算比较复杂,不利于模型的实现,对于实体之间为了相互之间的利益在推荐时进行欺骗的行为不能抑制;EigenRep[9],PeerTrust[10]信誉模型需要遍历通过迭代方式计算节点的全局可信度,计算和通信开销大,模型收敛性值得商榷,同样对于欺骗和不诚实反馈抑制能力不足.

博弈论(game theory)[11]的相关思想和应用场景吻合网络自主元素在其生命周期内具有行为特征的生存场景,自主元素理性选择行为策略的关键在于考察系统对其行为的反馈,在一定条件下,元素基于追求利益最大化的考虑,倾向于采取收益较高的激励型策略,因此在无限次重复博弈中可以实现Pareto最优的合作均衡.文献[12]提出建立激励机制的有益思路未给出实现.文献[13]提出一种资源共享的非合作博弈模型,算法需遍历节点,对于网络规模的动态性和规模性适用性有限.文献[14]提出基于重复博弈的信用管理机制RGTrust,以“囚徒困境”模拟节点交易,算法具有良好的稳定性,信用计算时间复杂度及报文通信量与其他模型相比有明显优势,但是模型对于不同类型和不同信用历史节点的惩罚区分不明晰,没有探讨对于共谋等欺骗行为的抑制,缺乏对诚实交易的激励机制.文献[15]基于文献[14]的思想提出的DPTrust模型以节点最近时间窗口的信用量级作为对节点行为偏离的惩罚指标,能够在一定程度上区别施加对于不同类别节点的惩罚,但是指标粒度较粗,不考虑节点的历史信誉,使得投机节点有可乘之机;另外,对于网络节点行为多样性、欺骗性(例如:共谋团伙帮助同伙刷信用,从而度过惩罚期,并且通过多次此类投机行为牟取团伙整体收益)也没有进行进一步的探讨.

本文基于上述研究工作的建模思想,针对相关模型中存在的不足,构造一种虚拟计算环境下,基于非合作重复博弈的自主元素惩罚激励机制 PETrust,旨在提高模型对不良行为的动态适应能力和对信誉信息的有效聚合能力,根据自主元素行为特征和信誉状态区分交易行为的合作与偏离,并对偏离行为进行惩罚,鼓励自主元素积极参与诚实合作.给出模型的定义、机制、证明以及分布式实现.与相关模型进行实验比较,PETrust 能够准确甄别和区别对待自主元素诚实交易、偶尔偏离、摇摆投机和恶意行为,通过对博弈过程中个体收益、信誉变化在不同策略情形下的调整来引导自主元素采取系统所期望的行为,从而减少恶意行为的危害,对于共谋性欺骗具

© 中国科学院软件研究所 http://www.jos.org.cn

有较强抑制,同时激励计算环境中自主元素基于自身的利益判断,采取系统所期望的行为积极诚实协作,最终达成系统整体收益最优.

2 自主元素协作的囚徒困境分析

2.1 协作收益矩阵

在典型的囚徒困境特征的交易模型中,由于每个交易方的自私性和机会主义倾向,对于一次性博弈的策略组合,各方只关心一次性的收益,{不合作,不合作}是各方的最后策略选择,即选择尽可能大的可能性去享用他人的服务和资源,同时尽可能自保,以达到自身利益最大化.但是如果博弈重复多次,参与人则可能为了长远利益而牺牲眼前利益从而选择不同的策略均衡.无限次重复博弈理论[11]证明了当博弈重复无穷多次而不是一次时,存在着完全不同于一次博弈的子博弈精炼均衡.考虑囚徒困境博弈,此时{合作,合作}是一个子博弈精炼纳什均衡结果.

在以网格、P2P 等计算环境为代表的虚拟计算环境中,自主元素作为具备自主行为能力的基本资源管理单位,在其生命运转周期内,本着"自主参与、自愿协同"的原则与其他元素相互之间不断地发生交互,或以服务提供方的身份提供服务,或以服务消费方的身份享受服务,或者是以提供服务消费服务相混合的方式进行交互.

定义 1. 设 G 为阶段博弈,虚拟计算环境中自主元素之间的长期行为是阶段博弈的无限次重复博弈,记作 $G(\infty,\delta)$,δ为局中共同的贴现因子且 $0<\delta<1$.自主元素的行为选择策略集合 $Behavior=\{Co,Un\}$,表示实体可供选择的行为策略集合为{合作,不合作},则任意一对自主元素在某一时段的交互中收益矩阵定义见表 1.

Table 1 Autonomic elements collaboration income matrix

表 1 自主元素协作收益矩阵

Strategy	*Co*	*Un*
Co	$(v-c_1-c_2,v-c_1-c_2)$	$(-c_1-c_2,\eta-c_1)$
Un	$(\eta-c_1,-c_1-c_2)$	$(-c_1,-c_1)$

表 1 中:

- 自主元素在网络中存在计算、通信、存储以及运行和维护等成本,记作$-c_1$;
- 自主元素交易中采取合作策略时,存在资源占用及消耗等成本,记作$-c_2$;
- 当采取投机策略而对方合作时,则没有$-c_2$这个损耗,并且能够获得一个η的收益;
- 如果各方友好合作,双方在付出正常成本$-c_1-c_2$的同时,均获得收益v.

例如,在文件共享系统的一次交易过程中,任意节点向对方提供资源,则存在网络成本及合作成本(如共享文件带来的资源占用等),记为$-c_1-c_2$;若对方节点使用了该共享资源而自身并未友好合作(如合理评价、及时付费或资源交换等),则只付出网络成本并可获得投机收益,记为$\eta-c_1$;若双方在一次交易过程中均采取友好合作策略,则都获得合作收益;双方均不采取合作时,则只存在一个网络成本而没有任何收益.

2.2 相关假设

假设 1. 自主元素是理性且自私的,均以获取自身总体最大收益为目的.

假设 2. 在上述收益中,$v-c_1-c_2\geqslant 0$ 成立.

虚拟计算环境中自主元素之间的友好行为对合作双方均为有利,通过合作各方获得的收益至少不应该是一个负值.

假设 3. 在上述收益中,$\eta>v-c_2>(-c_2+\eta)/2>0$ 成立.

在单阶段博弈中,自主元素出于自身考虑,总是具有投机倾向,*Un* 策略成为其最佳选择.

假设 4. 自主元素的交易策略选择可以描述为无限期重复博弈.

自主元素在其网络生存过程中,并不确知自身生命周期的终点以及会在哪一个阶段离开网络.该过程属于随机结束重复博弈模型,其认知是自身会永远处于网络行为之中.从交易策略选择的角度来看,等同于无限次重

© 中国科学院软件研究所 http://www.jos.org.cn

复博弈.借鉴文献[12−15]思想,引入非合作重复博弈理论对合作行为的形成机理和实施条件进行剖析.重复博弈激励自主元素友好合作,合作是集体理性的结果,合作整体利益大于每个成员单独经营的收益之和.通过大量的重复博弈以及高效的合作保障机制的配合,形成高昂的单边背叛行为成本和有效的行为认同模式.协作策略{*Co*,*Co*}成为 Pareto 最优结果.

3 PETurst 模型

自组织系统具有动态开放的本质特点,自主元素行为自治、多变,总处于追求自身利益最大化的状态之中.在交易中,它们期望尽可能少地提供服务和资源,同时又想尽可能多地使用他人提供的服务和资源.本文机制尊重并抓住自主元素交易行为本质特点,根据自主元素在参与系统中的行为表现,通过一定的策略来引导节点按照系统所期望的方式参与到信誉系统中,制定策略的原则是积极诚实协作的节点可以获取更大收益,如信誉值的提高、更多的协作机会、更高的收益率等.

3.1 自主元素信誉特征描述

定义 2. 自主元素i第l次交易的可信辨识度$D_{trust}(i,l)\in\{0,1\}$,0 代表自主元素i本次交易是可信的,1 代表不可信.通常情况下,自主元素在设置交易对方可信辨识时不存在自欺行为.

定义 3. 自主元素i第l次交易的信誉状态$R_{status}(i,l)\in[0,1]$,1 代表实体信誉最好,数值越低表示实体信誉状况越差.

定义 4. 自主元素 i 在第 l 次交易的行为策略如果符合公式(1)的规范,则称自主元素行为友好,记为$departure(i,l)=0$;如果不符合上式规范,则称其行为偏离,记为 $departure(i,l)=1$.

$$B_{policy}(i,l)=\begin{cases}Co, & D_{trust}(j,l-1)=0, R_{status}(j,l-1)\geqslant ReV\\ Un, & D_{trust}(j,l-1)=1 \text{ or } R_{status}(j,l-1)<ReV\end{cases} \tag{1}$$

其中,$D_{trust}(j,l-1)$为自主元素j在第$l-1$ 次交易的可信辨识,$R_{status}(j,l-1)$为自主元素j在第$l-1$ 次交易的信誉状态,ReV是一个信誉状况阈值.

自主元素通常希望交易对方的信誉状态良好且交易可信,因此,自主元素 i 的策略选择不仅与元素 j 在第$l-1$ 次交易的可信辨识有关,同时与 j 当前信誉状况有关.目前,大多数研究工作中,i 的行为策略或者只考虑 j 的可信辨识度,或者只考虑 j 在最近一次交易中的信誉状况,带有一定片面性.通过稍后的介绍可以看到,本模型中 j 的信誉状态是通过综合考察 j 在历次交易中的行为得到的,因此能更加客观地表示节点的信誉状况.

定义 5. 虚拟计算环境中任一自主元素在其生命周期的某一阶段的信誉特征可以用三元组$\langle D_{trust},R_{status},B_{policy}\rangle$来表示.其中,可信辨识初值$D_{trust}(i,0)=0$,信誉初值$R_{status}(i,0)=1$.

3.2 惩罚-激励机制

对于自组织虚拟计算环境而言,由于无法直接控制自主元素的参与行为,使得系统常常面临严重的可用性问题,这使得提出有效的机制、激励节点积极贡献和可靠服务、抑制不良行为的影响成为必须.本文提出的惩罚-激励机制 PETrust 的核心思想在于,发现自主元素的危害行为并根据其行为特征决定惩罚力度,惩罚力度并非持续不变,而是根据自主元素的历史行为和累积信誉信息不断动态调整.从无限期重复博弈的角度来看,任何短期的机会主义行为的所得都是微不足道的,交互的自主元素有积极性为自身建立一个良好的信誉;同时,不同行为特征(理性或恶意)的自主元素随着生存周期增加,其信誉特征将逐渐出现分化,从而有利于系统进行甄别和防范,达到鼓励诚实合作、遏制和惩罚投机行为的目的.

自主元素的某次投机行为将导致其信誉受损,并且进入惩罚期接受一定时间和频次的持续惩罚.在惩罚期中,自主元素只有进行一定次数的诚实交易才能使受损信誉逐步恢复,并在惩罚期结束后重新达到平衡状态.为了进一步区分惩罚力度,惩罚期的时长、惩罚期内必须满足的最小诚实交易次数应与节点的信誉状态、对系统的贡献值(如开放的资源、提供的计算能力等)以及投机行为的危害程度(如消耗的资源和时间、造成的网络拥塞、对交易对方造成的损害等)有关,并随着节点在惩罚期内的表现(信誉状态、是否可信交易等)动态调整.这

© 中国科学院软件研究所 http://www.jos.org.cn

样的模型称为“惩罚-激励”机制.

图 1 表示了惩罚-激励机制的基本工作原理.自主元素的生存周期可分为若干至少包含一次交易行为的交易阶段,每个阶段的时长不必相等但不可小于某一固定阈值τ.在网络运行初期,假设所有节点工作于正常期(normal period),其行为可信.若自主元素i的某次交易不可信($D_{trust}(i,l)=1$),则进入惩罚期(punishment period).此时,节点必须保证至少连续完成θ_i次诚实交易才可结束惩罚期,恢复到正常期,显然惩罚期时长不小于$\theta_i\tau$.惩罚期包含的阶段数量θ_i称为惩罚计数因子.惩罚-激励机制通过调整θ_i区分对不同偏离行为的惩罚力度.θ_i的取值与自主元素i在本次交易之前固定$w(w>1)$个交易阶段时间窗口内的历史交易行为有关,其计算方法为

$$\theta_i(w)=\left\lceil \log_\alpha\left(\frac{\sum_{k=1}^{w}N_k}{\sum_{k=w}^{1}\sum_{l=1}^{N_k}\left[R_{status}(i,k,l)\times\lambda_1(i,k,l)\times\lambda_2^{-1}(i,k,l)\times\sigma^{\sum_{j=1}^{w-k}N_{w-j+1}+N_k-l}\right]}\right)\right\rceil \tag{2}$$

其中,

- N_k表示第k个阶段产生的交易次数;
- $\lambda_1(i,k,l)$表示自主元素i在第k个阶段的第l次交易的贡献,且$0<\lambda_1(i,k,l)<1$;
- $\lambda_2(i,k,l)$表示自主元素i在第k个阶段的第l次交易对系统造成的危害,且$1<\lambda_2(i,k,l)$;
- $R_{status}(i,k,l)$为自主元素i在第k个阶段的第l次交易的信誉状态;
- $\sigma(0<\sigma<1)$称为比重区分因子,对于自主元素 i 的所有考察的历史交易,新近发生的交易在惩罚计数因子中所占比重更大;
- α称为惩罚力度因子,$\alpha(\alpha>1)$越小,惩罚计数因子越大.

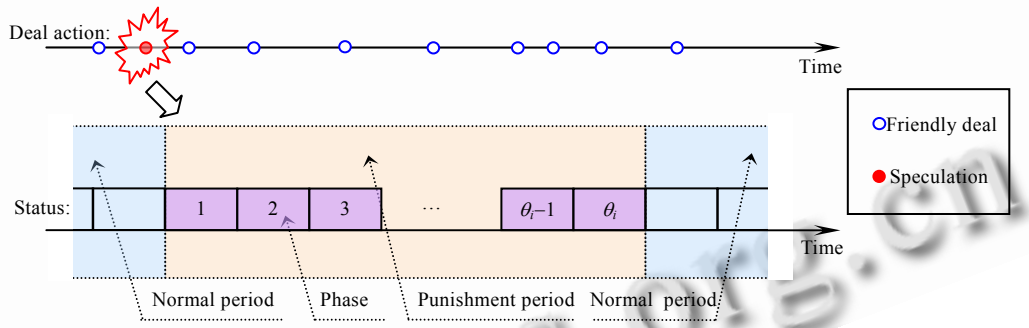

Fig.1 Sketch map of the penalty-incentive mechanism

图 1 惩罚-激励机制示意图

惩罚期内每个阶段结束后,需要根据自主元素在该阶段内完成交易的情况按照相应规则重新计算惩罚期内剩余阶段数量n.为描述方便,假设$n_i(t)$为自主元素i在惩罚期内当前剩余阶段数量,$n_i(t+1)$为新的阶段数量,N_t为当前阶段自主元素i的交易次数,$D_{trust}(i,N_t,l)$为自主元素i当前阶段第l次交易的可信辨识度,则

$$n_i(t+1)=\begin{cases}\theta_i, & \text{if } \sum_{l=1}^{N_t}D_{trust}(i,N_t,l)>0, n_i(t)=0\\ \max\{\theta_i,n_i(t)+1\}, & \text{if } \sum_{l=1}^{N_t}D_{trust}(i,N_t,l)>0, n_i(t)>0\\ n_i(t)-1, & \text{if } \sum_{l=1}^{N_t}D_{trust}(i,N_t,l)=0, n_i(t)>0, r(s)\leqslant p\\ n_i(t), & \text{if } \sum_{l=1}^{N_t}D_{trust}(i,N_t,l)=0, n_i(t)>0, r(s)>p\\ 0, & \text{otherwise}\end{cases} \tag{3}$$

其中,$r(s)$为随机函数,p称作信用常模且 $0<p<1$.正常情况下,若自主元素i在惩罚期内当前阶段交易可信,则剩余阶段数量会以步长 1 逐阶段减少,直到$n=0$ 时惩罚期结束;反之,若自主元素i在惩罚期内某次交易行为不可信,

© 中国科学院软件研究所 http://www.jos.org.cn

则根据本次交易之前w个阶段的时间窗口重新计算惩罚计数因子θ_i,并取θ_i和$n_i(t)+1$ 中的最大者作为节点i在惩罚期内的剩余阶段数量.显然,自主元素i在惩罚期内的偏离行为必然导致惩罚期延长,自主元素i必须完成更多次数的可信交易才能通过惩罚期.可以看到,自主元素在惩罚期内至少需要进行θ_i次可信交易,且惩罚期时长不小于$\theta_i\tau$;其次,即使当前阶段的交易可信,剩余阶段数量减少的概率只有p,不仅增加了自主元素的代价,而且使实体无法预知博弈次数,从而防止自主元素间通过共谋欺骗快速结束惩罚.

当$k>0$时,$R_{status}(i,k,l)$表示自主元素i在惩罚期第k阶段第l次交易完成时的信誉状态;当$k=0$时,$R_{status}(i,k,l)$表示自主元素i在正常期第l次交易完成时的信誉状态;$n_i(0)$表示自主元素i惩罚期阶段数量初值;$n_i(t)$表示自主元素i当前剩余惩罚阶段数量,则$R_{status}(i,k,l)$设置规则可表示为

$$R_{status}(i,k,l)=\begin{cases}\min(R_{status}(i,k,l-1)+\varepsilon,1), & \text{if } departure(i,l)=0,k=0\\ R_{status}(i,k,l-1)-ReD, & \text{if } departure(i,l)=1,k=0\\ R_{status}(i,1,1)+ReA[n_i(0)-n_i(t)]/n_i(0), & \text{if } k>0\end{cases} \tag{4}$$

其中,

- 参数ε为非惩罚期内可信节点提供一个很小的信誉值增量;
- ReD 是对实体在其行为偏离时给予的一个信誉惩罚值;
- ReA 则是对实体经过努力度过惩罚期而给予的一个信誉奖励值.

从上述规则可以看到,若自主元素在惩罚期内行为偏离,则信誉状态会逐阶段减少,甚至可能低于惩罚期初期的信誉状态.若其行为友好,则信誉状态会逐渐恢复.但考虑到对实体行为曾经偏离的惩戒,通常有 $ReA<ReD$,因此不会超过之前正常期的信誉状态.

自主元素 i 第 l 次交易的可信辨识设置规则如下:

$$D_{trust}(i,l)=\begin{cases}0, & \text{if } D_{trust}(i,l-1)=0, departure(i,l)=0 \text{ or } D_{trust}(i,l-1)=1, departure(i,l)=0, n_i(t-1)=1\\ 1, & \text{if } departure(i,l)=1 \text{ or } departure(i,l)=0, n_i(t-1)>1\end{cases} \tag{5}$$

这表明自主元素可信辨识较之信誉状态是一个相对微观的标识,它能够更紧密地反映实体实时可信信息.

总之,自主元素在整个生命周期内信誉特征三元组$\langle D_{trust},R_{status},B_{policy}\rangle$随着自主元素交易行为和状态不断动态变化.对于因行为偏离而进入惩罚期的自主元素,PETrust可根据其行为特征决定惩罚力度,并且保留自主元素历史信誉信息,记录于信誉特征之中,作为未来策略选择和计算惩罚力度的依据.随着交易次数的增加,信誉特征三元组$\langle D_{trust},R_{status},B_{policy}\rangle$能够更好地反映节点的行为特征,有效甄别自主元素的诚实交易、偶尔偏离、摇摆投机和恶意行为,从而鼓励正常节点的规范行为,并减少恶意节点对系统的危害.

3.3 PETrust模型分布式实现

为实现PETrust模型在虚拟计算环境下的分布式计算,需要一套非集中式的数据管理方案实现自主元素信誉信息的分布式存储与查询.我们基于文献[16]提出的Terrace拓扑设计了面向PETrust的信誉信息分布存储机制.Terrace拓扑是一种基于分布哈希表(distributed hash table,简称DHT)的结构化拓扑,我们将其作为PETrust模型的底层拓扑支撑.利用Terrace所具有的地址冗余容错能力[16],可以确保信誉信息不会因为个别Terrace拓扑节点的失效或退出而受到损失.

通过 Terrace,网络中的所有节点投影到一个逻辑 *d-tree* 上,并赋予节点全局唯一的逻辑地址,逻辑地址的编码基数为 *d-tree* 的阶.例如,如果用八进制表示逻辑地址,则 d 为 8.从根节点起,每一个节点的子节点都是该节点逻辑地址在下一个基数位的列举,而节点自身包含当前基数位的所有列举(根节点例外,它没有第 0 位子树).通过均匀的 Hash,节点可以将对象(或对象索引)投影到同样的逻辑地址空间.

定义 6. 设DHT为任意均匀的Hash函数,网络中任意节点i的标识在Terrace中的投影称为i的档案点$FilePoint_i$.即$FilePoint_i=DHT(ID_i)$.ID_i为节点i的全局唯一标识符.

基于Terrace,网络中的任意节点同时也是其他节点的档案点.档案点$FilePoint_i$至少具有如下4项功能:1) 存储、计算和更新自主元素i信任值和交易信息等相关数据.2) 根据公式(1)判断自主元素i是否行为偏离,并进一步判断是否进入或结束惩罚期.其中,惩罚期内剩余阶段数量根据公式(2)和公式(3)计算得到.3) 根据公式(4)计

© 中国科学院软件研究所 http://www.jos.org.cn

算当前交易完成时自主元素i的信誉状态.为减少存储开销,$FilePoint_i$仅根据i最近w个交易阶段时间窗口内的交易行为计算其信誉状态,因为i早期交易行为对其最近的表现说服力较弱.这种方案也易于实现,只要档案点为该自主元素维护一个先进先出队列淘汰那些较早时间的交易记录即可.4) 向其他节点提交i的信誉特征$\langle D_{trust}, R_{status}, B_{policy}\rangle$以及每一次的交易信息等数据.

事实上,PETrust模型的工作机理是通过档案点之间的协同完成的.因此,每个档案点至少需要包含一个如表 2 所示的数据结构.其中:节点$FilePoint_i$是网络中节点i的档案点,ID_i为i的标识;$N_1,\ldots,N_w$代表节点i在各阶段执行的交易次数,$D_{trust}(i,0)$和$R_{status}(i,0)$分别表示节点i当前的可信辨识度和信誉状态;$ID_{j,*,*}$为与节点i交易的各节点的标识,$n_i(1),\ldots,n_i(w)$为节点i的当前剩余阶段数量,其值大于 0 表示节点i尚处于惩罚期;P_k^i为节点i的公钥.

Table 2 Data structure for FilePointi

表 2 FilePointi 的数据结构

ID_i			$D_{trust}(i,0)$	$R_{status}(i,0)$	$B_{Policy}(i,0)$	P_k^i
N_1	$\lambda_1(i,1,1)$ $\lambda_1(i,1,2)$ … $\lambda_1(i,1,N_1)$	$\lambda_2(i,1,1)$ $\lambda_2(i,1,2)$ … $\lambda_2(i,1,N_1)$	$D_{trust}(i,1,1)$ $D_{trust}(i,1,2)$ … $D_{trust}(i,1,N_1)$	$R_{status}(i,1,1)$ $R_{status}(i,1,2)$ … $R_{status}(i,1,N_1)$	$ID_{j,1,1}$ $ID_{j,1,2}$ … $ID_{j,1,N_1}$	$n_i(1)$
N_2	$\lambda_1(i,2,1)$ $\lambda_1(i,2,2)$ … $\lambda_1(i,2,N_2)$	$\lambda_2(i,2,1)$ $\lambda_2(i,2,2)$ … $\lambda_2(i,2,N_2)$	$D_{trust}(i,2,1)$ $D_{trust}(i,2,2)$ … $D_{trust}(i,2,N_2)$	$R_{status}(i,2,1)$ $R_{status}(i,2,2)$ … $R_{status}(i,2,N_2)$	$ID_{j,2,1}$ $ID_{j,2,2}$ … $ID_{j,2,N_2}$	$n_i(2)$
…			…			…
N_w	$\lambda_1(i,w,1)$ $\lambda_1(i,w,2)$ … $\lambda_1(i,w,N_w)$	$\lambda_2(i,w,1)$ $\lambda_2(i,w,2)$ … $\lambda_2(i,w,N_w)$	$D_{trust}(i,w,1)$ $D_{trust}(i,w,2)$ … $D_{trust}(i,w,N_w)$	$R_{status}(i,w,1)$ $R_{status}(i,w,2)$ … $R_{status}(i,w,N_w)$	$ID_{j,w,1}$ $ID_{j,w,2}$ … ID_{j,w,N_w}	$n_i(w)$

自主元素之间的匿名性和数据传输安全性可从以下 3 方面得到保证:

1) 由于Hash函数的单向性,因此$FilePoint_i$无法知道i的物理地址.
2) 由于自主元素加入拓扑时获取逻辑地址是随机的,无法根据某个特征(如 IP 地址等)预先决定节点的逻辑地址,因此,任意自主元素 i 不能选择自己的标识 ID_i,以便使 ID_i 正好是存储网络中某个自主元素 j 的信任值的逻辑地址.
3) 为抑制冒名欺骗,要求任意节点i具有各自的公钥P_k^i和私钥S_k^i对,同时要求任意节点i在提交对节点j的信任值和交易信息时同时提交自身的IP地址IP_i和节点j的公钥P_k^i.文献[16]称这种IP地址与节点标识匹配的认证方法为$Cent_{IP\text{-}ID}$机制.在$Cent_{IP\text{-}ID}$机制下,对于节点u冒充节点i企图欺骗节点j的档案点的情况,冒名者u将面临两难的局面:如果u提交自身的IP地址IP_u,j的档案点$FilePoint_j$可以由$Terrace$网络通过i的档案点$FilePoint_i$获取i的公钥P_k^i,然后用P_k^i和任意一个整数r构造一个对u的挑战(challenge),从而使u暴露;如果u提交i的IP地址IP_i,$FilePoint_j$可以通过与i交互发现i近期并没有提交对j的信任值和交易信息,从而使u暴露.$Cent_{IP\text{-}ID}$机制对冒名欺骗具有较好的防范能力.

在实现自主元素匿名管理的基础上,图 3 给出了PETrust模型的分布式算法伪码.算法分别从服务提供节点和档案点的角度出发,讨论了单次交易过程中信任信息和交易信息的计算和更新过程以及消息发送机制.可以看出,在NodeAction部分,伪码中交易节点仅根据本地信息进行策略选择,对节点状态的计算和判断则由档案点负责完成,档案点只需维护固定阶段窗口内节点交易信息.从FilePointAction部分伪码中可以看出,$Node_i$与其档案点并未发生交互,满足了匿名性要求.与EigenRep和PeerTrust等模型不同,PETrust中自主元素的全局信任度的计算无需全网范围内的迭代,而且算法中信任数据的定位和查询操作具有可以证明的消息复杂度上界[14].也可以证明,在规模为n的网络中,任意自主元素i可以在$O(\log n)$的消息复杂度内将其对自主元素j的信任评分写入

© 中国科学院软件研究所 http://www.jos.org.cn

$FilePoint_j$,同样,任意自主元素i可以$O(\log n)$的消息复杂度从$FilePoint_j$获取j的信任数据(如信誉状态R_{status}).每个节点需要维护的路由表(finger table)的规模仅为$O(\log n)$,这与消息复杂度为$O(n^2)$的EigenRep相比,在效率上有了很明显的提高,更易于部署实施.

```
Algorithm NodeAction.
Begin
  INFO←GetTransactionInfo(FilePoint_j,REQ)      //get j's last deal information from file node
  Node_i.Behavior←GetBehavior(INFO)             //select trade strategy according to formula (1)
  RESULT←Transaction(Node_i,Node_j)             //trading
  SetTransactionInfo(FilePoint_j,RESULT)        //send j's current trade information to its file node
  Receive ACK message from FilePoint_j          //receive ACK
End
Algorithm FilePointAction
Begin
  While receive REQ message from Node_i do
    SendTransactionInfo(Node_i,INFO)            //send last deal information to node i
  While receive RESULT message from Node_i do
    Sign←Departure(RESULT)                      //judge j is right or not according to formula (1)
    trust←Dtrust(RESULT,Sign)                   //judge if this trade is trustworthy according to formula (6)
    reputation←Rstatus(RESULT,Sign)             //compute reputation status according to formula (5)
    If Sign=1 then                              //if departure=true
        Update the count of residual punish phases N_j  //compute penalty phases according to formula (2), formula (3)
    Else
      If N_j>0 and the current phase is expire then     //if j is in penalty period and current phase falls due
        Update the count of residual punish phases N_j
    Update Node_j status
    Send ACK message to Node_i                  //send ACK
End
```

Fig.2 Pseudocode for the PETrust model

图 2 PETrust 模型伪码

3.4 PETrust模型分析

定理 1. 自主元素基于长期利益的考虑,采用 PETrust 前述策略(简称 PE 策略),构成虚拟计算环境下自主元素间无限次囚徒困境型重复博弈的一个子博弈精炼纳什均衡.

证明:

(1) 首先证明交互实体采取 PE 策略构成一个纳什均衡.

当实体j坚持PE策略时,考察实体i坚持PE策略的收益π_p以及实体i在某阶段偏离的最大可能收益π_d.这两个收益如果满足$\pi_p>\pi_d$,则实体i会自然地选择坚持PE策略.π_p,π_d按照无限次博弈中所有阶段收益的贴现值$\sum_{i=1}^{\infty}\delta^{i-1}\pi_i$ 计算,可得如下不等式:

$$\frac{v-c_1-c_2}{1-\delta}>\frac{(v-c_1-c_2)(1-\delta^k)+(\eta-c_1)\delta^k(1-\delta)-(c_1+c_2)\delta^{k+1}(1-\delta^n)+(v-c_1-c_2)\delta^{k+n+1}}{1-\delta} \tag{6}$$

求解,得到$\delta>\sqrt[n]{\frac{\eta+c_2}{v(n+1)}}$.即,如果满足该条件,那么当实体$j$坚持 PE 策略并且没有首先不合作,实体$i$不会选择首先不合作.

当实体j首先选择了不合作,在随后的惩罚期,实体i如果坚持PE策略,则每阶段的收益是$\eta-c_1$;如果i选择其他策略,i在任何单阶段的收益不会大于$\eta-c_1$.因此,i有积极性坚持PE策略.同样,在惩罚期之后,i坚持PE策略所能获得的收益也使i有积极性坚持PE策略.因此,PE策略是博弈双方最佳应对策略,构成一个纳什均衡.

(2) 由于博弈重复无限次,从任何一个阶段开始的子博弈与原博弈的结构相同,原博弈的纳什均衡也构成每一个子博弈上的纳什均衡.

综合(1)、(2),我们证明了 PE 策略是自主元素间无限次囚徒困境型重复博弈的一个子博弈精炼纳什均衡,pareto 最优(合作,合作)是每一个阶段的均衡结果,实体走出了一次性博弈时的困境.

© 中国科学院软件研究所 http://www.jos.org.cn

上述结果表明,虚拟计算环境下,自主元素之间的交互在无限次的假设前提下,实体如果有足够的耐心,以长远利益为着眼点,任何短期的机会主义行为的所得都是微不足道的,交互的实体有积极性为自身建立一个好的信誉,同时也有积极性惩罚对方的机会主义行为.

4 仿真实验与分析

以P2P网络应用为例,考察PETrust的性能.为此,我们在Peersim[17]仿真实验平台上,采用Java语言实现了PETrust模型.假设节点加入P2P网络后进行文件共享并从中受益,整个生存周期中节点的行为理性而自私,并且并不知道何时会离开网络,节点间共享的资源在效用上是对等且互补的.我们通过实验考察PETrust的有效性.

令 P2P 网络的规模 N=2000,实验中每次仿真选取 200 个连续的时间段作为节点交易阶段,交易只发生在阶段内,每个阶段随机选取一定比例的节点进行交易.交易是否成功主要看是否提供文件下载以及下载文件是否真实.节点在每个阶段以概率 0.1 发出一个文件下载请求后进入等待状态,拥有该请求文件的节点接收到请求后发出响应,如果接收到多个请求,则发送节点随机选择一个请求节点作出响应;若接收到多个响应,则从随机选择的一个响应节点处下载文件.如果下载失败,则在下一阶段再次发出请求并重新下载文件,直到下载成功.实验在每个阶段结束后收集一次数据,记录交易信息并更新相关节点的信誉信息,然后仿真进入下一个阶段.实验对 DHT 机制进行了简化,即随机指定一个节点作为某节点的信用档案点,并假设交易信息的获取没有开销.

实验 1. 节点信誉状态变化比较.

选取文献[14]提出的RGTrust、文献[15]提出的DPTrust与本文PETrust进行比较.我们同样在Peersim平台上加入了代码,实现了RGTrust和DPTrust模型.上述模型均采用囚徒困境重复博弈来描述节点交易行为,通过DHT方式存取节点博弈后的信用信息.不同之处在于:RGTrust通过随机概率值p决定惩罚力度(惩罚期长度),目的是使节点出于自身收益的考虑,不轻易选择行为偏离策略;DPTrust通过考察节点最近时间窗口内的交易行为来区分调控惩罚力度,但不考虑节点的信誉值,仅仅根据节点局部可信辨识计算惩罚周期,不区分惩罚期内的多次行为偏离.需要指出的是,由于RGTrust和DPTrust模型以交易的次数计算惩罚周期,不考虑交易间隔时间,为公平起见,实验规定每个阶段只进行一次交易.实验 1 中,PETrust环境参数设置见表 3,除了RGTrust的惩罚概率p_2=0.2外,其余模型相关参数均与PETrust相同.

Table 3 Experiments parameters of PETrust

表 3 PETrust 实验参数

Parameter	Value	Parameter	Value
Number of nodes (N)	2 000	Number of phases (T)	200
The length of a phase (τ)	5	Incoming parameter (η)	0.7
Incoming parameter (c_1)	0.05	Incoming parameter (v)	0.8
Incoming parameter (c_2)	0.5	Number of phases in the window (w)	8
Reputation threshold value (ReV)	0.7	Reputation incentive value (ReA)	0.08
Reputation penalty value (ReD)	0.1	Proportion division factor (σ)	0.8
Penalty degree factor (α)	1.25	Reputation norm (p_1)	0.95

Table 4 Comparison of different model features

表 4 不同模型特点比较

Model	Description method for trading behavior	Embody style for punishment degree	Departure in punishment period	Correlation between punishment and historical reputation	Length of punishment period	Distinguish granularity of punishment
RGTrust	Repeated game	Random p	Without considering	Unrelated	Related to trade count	Coarse
DPTrust	Repeated game	Punishment count	Without considering	Unrelated	Related to trade count	Relatively coarse
PETrust	Repeated game	Dynamic adjusting of punishment phases	Prolonged punishment period	Related	Related to trade count and punishment phases	Relatively fine

图 3 比较了 3 种模型在不同阶段行为偏离概率($p_{departure}$=0.1,0.2,0.4)下节点信誉状态变化.可以看出,由于 3

© 中国科学院软件研究所 http://www.jos.org.cn