

分层组织管理下的众包质量演化机理

林 莉^{1 2}

(1. 北京工业大学计算机学院, 北京 100124; 2. 可信计算北京市重点实验室, 北京 100124)

摘 要: 针对众包系统质量控制方法设计需要提供理论依据的问题, 通过分析当前典型的众包分层组织管理模式, 建立了众包虚拟组织间任务协作的演化博弈模型. 基于该模型分析众包质量的演化稳定性, 讨论众包任务参与者协作行为的动态特征, 并给出制约众包质量演化稳定性的 3 个关键因素, 即完成众包任务获取的经济利益、任务请求者的消费效用以及任务完成者参与众包带来的安全风险对实现众包质量最大化的作用机理.

关键词: 众包质量; 分层组织; 演化稳定性

中图分类号: TP 393

文献标志码: A

文章编号: 0254-0037(2017)02-0237-07

doi: 10.11936/bjtxb2016070027

Crowdsourcing Quality Evolution Principles in Hierarchically-organized Networks

LIN Li^{1 2}

(1. College of Computer Science, Beijing University of Technology, Beijing 100124, China;

2. Beijing Key Laboratory of Trusted Computing, Beijing 100124, China)

Abstract: It is very important for crowdsourcing system to analyze crowdsourcing workers' collaborative behaviors. In this paper, the evolutionary process of crowdsourcing quality was studied based on classic hierarchically-organized mode. By establishing an evolutionary game model for crowdsourcing task collaboration among different virtual organizations, the evolutionary stability of crowdsourcing systems was analyzed and the dynamics of crowdsourcing workers' behaviors were discussed macroscopically. The key factors affecting the evolution of crowdsourcing system, including the economic benefits, consumer utilities from completing crowdsourcing tasks and risks from insecure participating, were suggested and how the factors work was presented. All these results together provide the theory basis for designing quality control methods of crowdsourcing system.

Key words: crowdsourcing quality; hierarchical organization; evolutionary stability

近年来, 众包(crowdsourcing)^[1]作为一种灵活有效的分布式解决问题方法, 在人机交互、自然语言处理、人工智能以及信息检索领域都得到广泛应用. Jeff 在 2006 年首次提出众包的概念^[2]“一个公司或机构把过去由员工执行的工作任务以自由自愿的形式外包给非特定的(而且通常是大型的)大众网络

的做法”. 与软件服务业传统的外包不同, 众包任务外派给互联网上不确定的群体, 而外包则是外派给确定的个体. 然而, 正是这种不确定性使得众包质量会动态变化; 同时, 众包任务参与者具有行为自主性, 存在大量理性参与者^[3]和恶意欺诈者^[4-5], 他们会以不同方式促进或破坏众包^[6], 导致任务完成结

收稿日期: 2016-07-06

基金项目: 国家自然科学基金资助项目(61502017); 国家“863”计划资助项目(2015AA016002)

作者简介: 林 莉(1979—), 女, 讲师, 主要从事网络计算、信息安全方面的研究, E-mail: linli_2009@bjut.edu.cn

果参差不齐,更引发众包质量^[7]演化的不确定性,给众包系统带来挑战.因此,如何控制众包质量朝合理方向演化,建立一个既有利于理性竞争、又减少恶意干扰的良性环境成为众包技术应用的关键问题之一.其中,对众包任务参与者自主行为的动态分析是实现众包质量可控演化的重要前提.

与人类社会活动类似,众包参与者基于“自愿参与,自主协同”的原则构造系统行为,包含丰富的节点间博弈.因而博弈论成为当前国内外学者理论分析众包节点协作行为的主要手段,可分为静态分析和动态分析2类.何云华等^[8]针对众包交通监测应用提出一种基于静态博弈分析的众包交通监测隐私保护机制. Yang等^[9]提出一种基于预算约束的经济激励机制去解决离线移动众包应用中的工作者自私问题. Sun等^[10]针对传感众包应用提出一种基于工作者行为分析的激励方法. Li等^[11]提出一种基于随机拍卖的方法去激励工作者参与智能手机感知的众包任务. Ranganathan等^[12]基于多人重复博弈模型分析了网络节点的文件共享行为. Dasilva等^[13]基于多人重复动态博弈方法分析自组织系统的节点协作行为.然而,这些工作要么侧重静态分析,忽略参与者具有的独立、匿名等动态特征;要么在动态分析中假设参与者完全理性,但在实际众包应用中经常出现由于软硬件错误导致参与者无意间犯错.此外,众包平台上有很多会泄露参与者诸如位置信息等隐私的任务^[14],同时,恶意参与者给众包任务带来的安全隐患也会影响理性参与者的行为^[15-17].因此,需要提供新的动态分析方法去研究众包任务参与者自主行为的演化原理.针对上述问题,本文采用有限理性博弈分析框架——演化博弈论,从演化的视角统一研究安全、经济、消费效用等因素对众包质量演化的影响.

1 众包系统的分层组织管理模式

众包任务参与者既包含长期签约工作者,也包含自由志愿者.签约工作者一般从较容易的任务开始做起,根据其对任务的理解和完成任务的表现被提升安排到不同岗位.众包应用系统可有效标识出最可靠的签约工作者作为管理员,先将任务派发给他们,再由他们将任务分发给各个工作者.这种分层组织管理^[18]是当前众包系统采用的主流模式,例如 Google 和 MetaWeb 的知识图谱众包应用.

如图1所示,在众包系统中,性能较高的节点作为上层管理节点称为汇点构成系统中的骨干层.每

一个汇点作为服务器,管理一系列的客户端(称为子点).这些子点本身自治,根据地域的划分或基于信任原则,选择最可信的汇点,从汇点上接受众包任务.一个汇点及选择它的多个子点构成的群体称为一个 Club.在实际众包系统中,Club 可以是一个真实的物理组织,也可以是一个由若干自治节点根据众包任务需求而临时成立的虚拟组织.汇点作为 Club 中子点的管理服务器,一方面,周期性地向其他汇点发布 Club 中工作者的信息;另一方面处理 Club 子点的众包任务请求,包括匹配众包工作者能力和传递众包需求者的投标报价等.不同汇点间可采用对等交互模式进行消息通信,由此,构建了由不同汇点管理的大规模众包系统(Club 联盟).

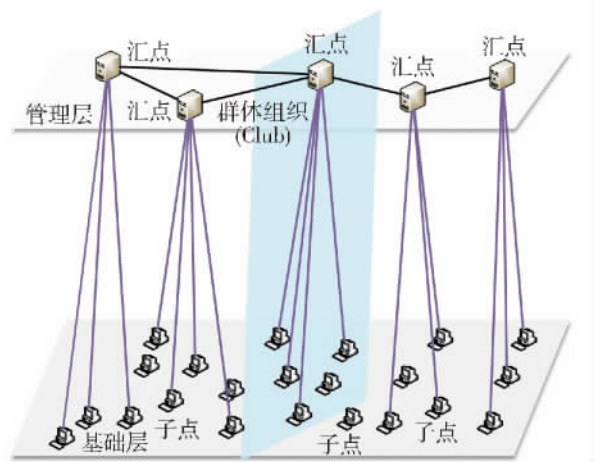


图1 典型的众包系统分层组织管理结构

Fig. 1 Classic layered organization management structure in crowdsourcing system

在上述分层组织模式下,众包工作者基于自愿参与、自主协同原则构造系统行为,包含丰富的节点间博弈.因而为确保 Club 联盟式众包应用中众包任务完成质量的演化可控性,必须先分析影响众包任务质量演化的因素及其作用原理,进而基于这些演化原理去设计质量控制方法.本文主要研究 Club 联盟式众包系统的质量演化原理.

2 众包质量的演化博弈分析

演化博弈论中最重要的2个基本概念是演化稳定策略(evolutionarily stable strategy, ESS)和复制子动态. Smith在文献[19]中给出了演化稳定策略的形式化定义,即对少数干扰(突变)仍具备稳健性的策略.复制子动态是由一个动态微分方程刻画的,用于描述某一特定策略在一个群体中被采用的比例随时间的变化率.类似适者生存原理,当一种策略

带来的期望得益高于群体所有个体的平均得益时, 这种策略就更容易在群体中发展; 随着群体中采用这种策略的个体比例增加, 其他个体更容易了解到其好处, 会改变原有策略, 这会提高群体中这种策略的增长速度.

在众包应用中, 众包任务完成质量的变化本质上是由网络中采用合作策略的节点比例变化而决定的. 由此, 本节先采用演化博弈分析方法建立 Club 联盟式众包系统任务协作的博弈模型, 然后基于该模型分析 Club 联盟式众包系统任务协作的演化稳定性, 进而得到众包系统任务质量的若干演化规律.

2.1 Club 间众包任务协作的博弈模型

在 Club 间众包任务协作的过程中, 由于子点具备行为的自主性, 因而一些子点可能为“合作”类型, 会按时提交正确的任务结果; 一些子点可能为“不合作”类型, 会为了尽可能付出最少的劳动而提交随意的答案甚至是虚拟结果. Club 子点类型无法事先确定, 会随着子点间策略的相互影响和战略调整而变化, 这使得 Club 联盟式众包质量的演化过程具备不确定性.

下面, 随机选择一对 Club, 比如 Club_i 和 Club_j, 建立它们之间的协作博弈模型. 在这个博弈中, 博弈方为 Club_i 和 Club_j.

在单个 Club 内部, “合作”类型的子点采用合作策略, “不合作”类型的子点采用不合作策略. 记 Coop 为合作策略, NonCoop 为不合作策略, 即 Club_i 和 Club_j 的子点策略集合都为 {Coop, NonCoop}.

令 Child_i 和 Child_j 为 Club_i 和 Club_j 的任一对子点. 对于某个时间段, 记 U_k, k = i, j 为该时间段内 Child_i 和 Child_j 都采取 NonCoop 时各方的正常效用, ΔU_k, k = i, j 为该时间段内 Child_i 和 Child_j 都采用 Coop 时所产生的额外效用. 其中, ΔU_k 可通过双方采用 Coop 时所获得的正效用 PU_k 和负效用 NU_k 之差来计算, 即

$$\Delta U_k = PU_k - NU_k, k = i, j$$

其中

$$PU_i = B_{-i} + B_i, PU_j = B_{-j} + B_j$$

$$NU_i = C_{-i} + C_i, NU_j = C_{-j} + C_j$$

即 PU_k 分成 2 部分: 子点 Child_k 从另一子点提交正确任务结果获取的好处 B_{-k} 和自身提交正确任务结果获取的经济利益 B_k; NU_k 也分成 2 部分: 子点 Child_k 获得另一子点提交正确任务结果的代价 C_{-k} 和完成众包任务承担的风险 C_k, 如可体现为完成众包任务会泄露的个人隐私信息^[13] 以及完成众包任务所产

生的系统开销, 这里 k = i, j.

显然, B_{-k} - C_{-k} 表示 Child_k 担当众包任务需求者时, 从 Child_{-k} 完成的任务结果中获取的效用, 简称为消费者效用; B_k - C_k 表示 Child_k 担当众包任务完成者时, 从自己完成任务上获取的利益, 简称为提供者效用. 由此, 给出 Club_i 和 Club_j 间众包任务协作博弈的效用矩阵, 如表 1 所示.

表 1 Club_i 和 Club_j 间众包任务协作博弈的效用矩阵
Table 1 Utility matrix between Club_i and Club_j in crowdsourcing collaboration game

子点	Child _j	
	Coop	NonCoop
Child _i	Coop (U _i + ΔU _i , U _j + ΔU _j) (U _i + B _i - C _i , U _j)	(U _i , U _j)
	NonCoop (U _i , U _j + B _j - C _j) (U _i , U _j)	(U _i , U _j)

分析上述效用矩阵. 容易看出, 若期望 Child_i 和 Child_j 众包任务协作, 即 (Coop, Coop) 为 Child_i 和 Child_j 间博弈的一个纳什均衡, 那么必须有 ΔU_k > 0, k = i, j, 即它们采用 Coop 策略都产生非负的额外效用.

2.2 Club 间众包任务协作的演化稳定性分析

为了发现 Club_i 和 Club_j 间众包任务协作的演化规律, 下面考虑 Club_i 和 Club_j 策略类型比例的变化. 假设在 Club_i 和 Club_j 中, 采用 Coop 的子点比例分别为 x 和 y, 即 Club_i 和 Club_j 中分别有 1 - x 和 1 - y 比例的子点采用 NonCoop, 那么 Club_i 中采用 Coop 子点的期望效用、采用 NonCoop 子点的期望效用和 Club_i 的平均期望效用依次为

$$U_i^{Coop} = y(U_i + \Delta U_i) + (1 - y)(U_i + B_i - C_i)$$

$$U_i^{NonCoop} = yU_i + (1 - y)U_i$$

$$U_{iAverage} = xU_i^{Coop} + (1 - x)U_i^{NonCoop}$$

由此得到 Club_i 的复制子动态方程为

$$\begin{aligned} dx/dt &= x(U_i^{Coop} - U_{iAverage}) = \\ &= x[U_i^{Coop} - xU_i^{Coop} - (1 - x)U_i^{NonCoop}] = \\ &= x(1 - x)[U_i^{Coop} - U_i^{NonCoop}] = x(1 - x)[y(U_i + \Delta U_i) + \\ &= (1 - y)(U_i + B_i - C_i) + yU_i + (1 - y)U_i] = \\ &= x(1 - x)(y\Delta U_i + B_i - C_i - yB_i + yC_i) = x(1 - \\ &= x)[y(B_{-i} + B_i - C_{-i} - C_i) + B_i - C_i - yB_i + yC_i] = \\ &= x(1 - x)[y(B_{-i} - C_{-i}) + B_i - C_i] \quad (1) \end{aligned}$$

方程 (1) 刻画了 Club_i 中所有子点众包任务协作行为的动态性, 对其做渐近稳定性分析有如下结论.

命题 1 若 y = (C_i - B_i) / (B_{-i} - C_{-i}), 那么使用 Coop 的 Club_i 子点的任何比例都是稳定的. 若

$y \neq (C_i - B_i) / (B_{-i} - C_{-i})$ 那么:

1) 对于 $B_{-i} - C_{-i} > 0$ 的情形, 当 $y > (C_i - B_i) / (B_{-i} - C_{-i})$ 时 $x^* = 1$ 是稳定的, 即 Coop 是 ESS; 当 $y < (C_i - B_i) / (B_{-i} - C_{-i})$ 时 $x^* = 0$ 是稳定的, 即 NonCoop 是 ESS.

2) 对于 $B_{-i} - C_{-i} < 0$ 的情形, 当 $y > (C_i - B_i) / (B_{-i} - C_{-i})$ 时 $x^* = 0$ 是稳定的, 即 NonCoop 是 ESS; 当 $y < (C_i - B_i) / (B_{-i} - C_{-i})$ 时 $x^* = 1$ 是稳定的, 即 Coop 是 ESS.

证明: 因为 $y \in [0, 1]$, 故须对 $(C_i - B_i) / (B_{-i} - C_{-i})$ 分类讨论:

1) $0 < (C_i - B_i) / (B_{-i} - C_{-i}) < 1$

如果 $y = (C_i - B_i) / (B_{-i} - C_{-i})$, 那么 $dx/dt = 0$ 始终成立, 即 $Club_i$ 中采用 Coop 子点的所有比例都不随时间变化, 它们都是稳定状态.

如果 $y \neq (C_i - B_i) / (B_{-i} - C_{-i})$ 那么只有当 $x = 0$ 和 $x = 1$ 时才有 $dx/dt = 0$, 即 $x^* = 0$ 和 $x^* = 1$ 是 2 个均衡状态. 也就是说, 当 $Club_i$ 中的子点干扰使 x 低于 x^* 时 $dx/dt > 0$ 必须成立; 当 $Club_i$ 中的子点干扰使 x 高于 x^* 时 $dx/dt < 0$ 必须成立. 对于 $B_{-i} - C_{-i} > 0$ 的情形, 当 $y > (C_i - B_i) / (B_{-i} - C_{-i})$ 时, $x^* = 1$ 处导数值为负, 即 Coop 为 ESS; 当 $y < (C_i - B_i) / (B_{-i} - C_{-i})$ 时, $x^* = 0$ 处导数值为负, 即 NonCoop 是 ESS. 对于 $B_{-i} - C_{-i} < 0$ 的情形, 结论正好相反.

2) $(C_i - B_i) / (B_{-i} - C_{-i}) > 1$

因为 $y \neq (C_i - B_i) / (B_{-i} - C_{-i})$ 且 $y < (C_i - B_i) / (B_{-i} - C_{-i})$, 对于 $B_{-i} - C_{-i} > 0$ 的情形 $x^* = 0$ 处导数值为负, 即 NonCoop 为 ESS; 对于 $B_{-i} - C_{-i} < 0$ 的情形 $x^* = 1$ 处导数值为负, 即 Coop 为 ESS.

3) $(C_i - B_i) / (B_{-i} - C_{-i}) < 0$

因为 $y \neq (C_i - B_i) / (B_{-i} - C_{-i})$ 且 $y > (C_i - B_i) / (B_{-i} - C_{-i})$, 对于 $B_{-i} - C_{-i} > 0$ 的情形 $x^* = 1$ 处导数值为负, 即 Coop 为 ESS. 对于 $B_{-i} - C_{-i} < 0$ 的情形 $x^* = 0$ 处导数值为负, 即 NonCoop 为 ESS.

由命题 1 可知, 当 $Club_j$ 中采用 Coop 的子点比例发生变化时, $Club_i$ 的进化稳定策略随之改变.

类似地, 对于 $Club_j$ 有命题 2, 证明过程不再赘述.

命题 2 如果 $x = (C_j - B_j) / (B_{-j} - C_{-j})$, 那么 $Club_j$ 中使用 Coop 的子点的任何比例都是稳定的. 如果 $x \neq (C_j - B_j) / (B_{-j} - C_{-j})$, 对于 $B_{-j} - C_{-j} > 0$ 的情形, 当 $x > (C_j - B_j) / (B_{-j} - C_{-j})$ 时 $y^* = 1$ 是稳

定的, 即 Coop 为 $Club_j$ 的 ESS; 当 $x < (C_j - B_j) / (B_{-j} - C_{-j})$ 时, $y^* = 0$ 是稳定的, 即 NonCoop 为 $Club_j$ 的 ESS. 对于 $B_{-j} - C_{-j} < 0$ 的情形, 结论正好相反.

根据命题 1、2, 可把 $Club_i$ 和 $Club_j$ 策略类型比例变化关系用一个坐标平面图表示 (见图 2).

如图 2 (a) 所示, 当 $Club_i$ 和 $Club_j$ 均获取非负的消费者效用, 即 $B_{-i} - C_{-i} > 0$ 且 $B_{-j} - C_{-j} > 0$, 当 $Club_i$ 和 $Club_j$ 中采用 Coop 的初始比例落在 ② 区时, 随着时间的推移, $Club_i$ 和 $Club_j$ 间的博弈最终演化到 2 个 Club 的所有子点都合作完成众包任务的状态; 当初始比例落在 ③ 区, $Club_i$ 和 $Club_j$ 间的博弈最终将演化到所有子点都不合作完成众包任务的状态; 而当初始比例落在 ①④ 区时, 演化方向是不确定的. 简而言之, 在这种情况下, $Club_i$ 和 $Club_j$ 或者都协作完成众包任务以发挥其带来的最大效用, 或者都不参与众包任务, 自行优化配置以增加效用.

如图 2 (b) 所示, 当 $Club_i$ 和 $Club_j$ 子点均获取负的消费者效用, 即 $B_{-i} - C_{-i} < 0$ 且 $B_{-j} - C_{-j} < 0$, 当 $Club_i$ 和 $Club_j$ 中采用 Coop 的初始比例落在 ①④ 区时, 随着时间推进, $Club_i$ 和 $Club_j$ 间博弈最终将演化到一个 Club 的子点都协作完成众包任务而另一个 Club 的子点都不参与众包任务的状态. 因此, 每一个 Club 都可能存在利用对方采用 Coop 而相机地采用 NonCoop 以增加本 Club 效用的意图.

如图 2 (c) (d) 所示, 当 $B_{-i} - C_{-i} > 0$ 且 $B_{-j} - C_{-j} < 0$, 或者 $B_{-i} - C_{-i} < 0$ 且 $B_{-j} - C_{-j} > 0$ 时, $Club_i$ 和 $Club_j$ 间的众包任务协作没有明确的演化方向.

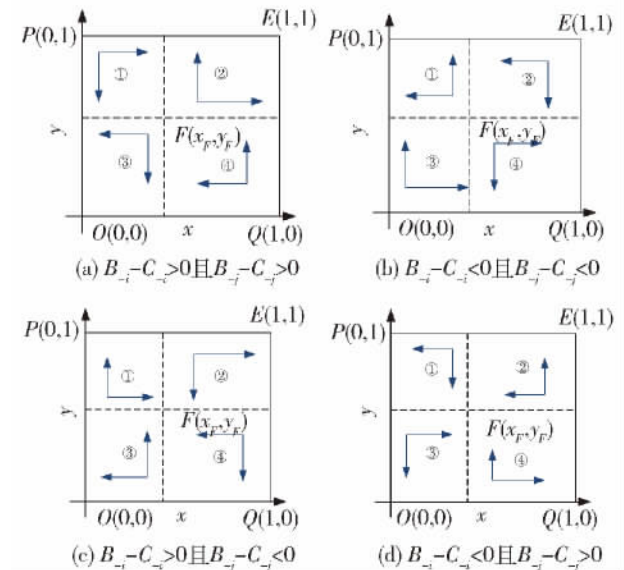


图 2 $Club_i$ 和 $Club_j$ 策略类型比例变化关系

Fig. 2 Policy type proportion change in $Club_i$ and $Club_j$

下面验证图 2 的规律,得到命题 3.

命题 3 当 $B_{-i} - C_{-i} > 0$ 且 $B_{-j} - C_{-j} > 0$ 时, $(Coop, Coop)$ 和 $(NonCoop, NonCoop)$ 是 $Club_i$ 和 $Club_j$ 众包任务协作博弈的 ESS; 当 $B_{-i} - C_{-i} < 0$ 且 $B_{-j} - C_{-j} < 0$, $(Coop, NonCoop)$ 和 $(NonCoop, Coop)$ 是 $Club_i$ 和 $Club_j$ 众包任务协作博弈的 ESS; 当 $B_{-i} - C_{-i} > 0$ 且 $B_{-j} - C_{-j} < 0$ 或 $B_{-i} - C_{-i} < 0$ 且 $B_{-j} - C_{-j} > 0$ 时, $Club_i$ 和 $Club_j$ 众包任务协作博弈无 ESS.

证明: 联立 $Club_i$ 和 $Club_j$ 的复制子动态方程得到微分方程组

$$\begin{cases} dx/dt = x(U_i^{Coop} - U_{iAverage}) = x(1-x) [y(B_{-i} - C_{-i}) + B_i - C_i] = f(x, y) \\ dy/dt = y(U_j^{Coop} - U_{jAverage}) = y(1-y) [x(B_{-j} - C_{-j}) + B_j - C_j] = g(x, y) \end{cases}$$

令 $f(x, y) = g(x, y) = 0$, 容易得到, 在平面 $S = \{(x, y) | 0 \leq x, y \leq 1\}$ 内上述方程组有 5 个均衡点: $O(0, 0)$ 、 $P(0, 1)$ 、 $Q(1, 0)$ 、 $E(1, 1)$ 和 $F(x_F, y_F)$, 其中 $x_F = (C_j - B_j) / (B_{-j} - C_{-j})$, $y_F = (C_i - B_i) / (B_{-i} - C_{-i})$. 根据 Freedman^[20] 提出的方法, 非线性微分方程组均衡点的稳定性可通过其雅可比矩阵的局部稳定性分析得到. 对 O, P, Q, E, F 进行局部稳定性分析, 其结果如表 2~5 所示. 证毕!

表 2 $B_{-i} - C_{-i} > 0$ 且 $B_{-j} - C_{-j} > 0$ 时 均衡点 O, P, Q, E, F 的局部稳定性分析

Table 2 Local stability analysis of equilibrium point O, P, Q, E, F when $B_{-i} - C_{-i} > 0$ and $B_{-j} - C_{-j} > 0$

均衡点	det J 的符号	tr J 的符号	结果
$O(0, 0)$	+	-	汇结点
$P(0, 1)$	+	+	源结点
$Q(1, 0)$	+	+	源结点
$E(1, 1)$	+	-	汇结点
$F(x_F, y_F)$	-	0	鞍点

表 3 $B_{-i} - C_{-i} < 0$ 且 $B_{-j} - C_{-j} < 0$ 时 均衡点 O, P, Q, E, F 的局部稳定性分析

Table 3 Local stability analysis of equilibrium point O, P, Q, E, F when $B_{-i} - C_{-i} < 0$ and $B_{-j} - C_{-j} < 0$

均衡点	det J 的符号	tr J 的符号	结果
$O(0, 0)$	+	+	源结点
$P(0, 1)$	-	+	汇结点
$Q(1, 0)$	+	-	汇结点
$E(1, 1)$	+	+	源结点
$F(x_F, y_F)$	-	0	鞍点

表 4 $B_{-i} - C_{-i} > 0$ 且 $B_{-j} - C_{-j} < 0$ 时 均衡点 O, P, Q, E, F 的局部稳定性分析

Table 4 Local stability analysis of equilibrium point O, P, Q, E, F when $B_{-i} - C_{-i} > 0$ and $B_{-j} - C_{-j} < 0$

均衡点	det J 的符号	tr J 的符号	结果
$O(0, 0)$	-		鞍点
$P(0, 1)$	-		鞍点
$Q(1, 0)$	-		鞍点
$E(1, 1)$	-		鞍点
$F(x_F, y_F)$	+	0	中心点

表 5 $B_{-i} - C_{-i} < 0$ 且 $B_{-j} - C_{-j} > 0$ 时 均衡点 O, P, Q, E, F 的局部稳定性分析

Table 5 Local stability analysis of equilibrium point O, P, Q, E, F when $B_{-i} - C_{-i} < 0$ and $B_{-j} - C_{-j} > 0$

均衡点	det J 的符号	tr J 的符号	结果
$O(0, 0)$	-		鞍点
$P(0, 1)$	-		鞍点
$Q(1, 0)$	-		鞍点
$E(1, 1)$	-		鞍点
$F(x_F, y_F)$	+	0	中心点

3 众包质量最大化的影响因素及其作用原理

众包技术的主要目标是实现众包任务质量最大化, 即实现众包系统中的双赢, 争取更多 $(Coop, Coop)$ 结果的出现.

根据图 2 可知: 首先, 须使各 Club 中子点获取非负的消费者效用; 其次, 正如图 2(a) 所示, 即便各 Club 中子点已获取非负的消费者效用, 当 $Club_i$ 和 $Club_j$ 中采用 Coop 的初始比例落在①④区时, 该群体博弈的演化方向是不确定的. 根据命题 3 可进一步得到初始比例落在①④区的质量演化规律, 即图 3. 正如图 3 所示, 当 $Club_i$ 和 $Club_j$ 中采用 Coop 的初始比例落在区域 $PFQE$, 随着时间的推移, $Club_i$ 和 $Club_j$ 构成的众包任务协作系统最终演化到 $Club_i$ 和 $Club_j$ 中所有子点都合作完成众包任务的状态; 当这个初始比例落在区域 $PFQO$ 时, 众包系统将最终演化到 $Club_i$ 和 $Club_j$ 中所有子点都不参与众包任务的状态. 为叙述方便, 给出定义 1.

定义 1 令 x 和 y 分别为 $Club_i$ 和 $Club_j$ 中采用 Coop 的节点比例, 如果从 (x, y) 出发, $Club_i$ 和 $Club_j$ 间博弈最终演化到所有子点都正确完成众包任务, 即都采取 Coop, 称 (x, y) 为最大质量可达态, 同时由

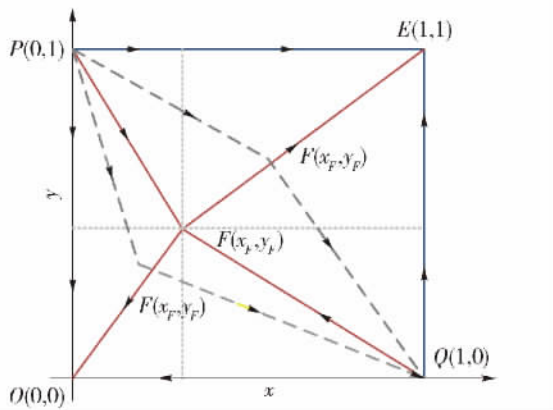


图3 Club间众包系统的动态演化图

Fig. 3 Dynamic evolution of club-based crowdsourcing system

所有最大质量可达态构成的区域(如图3的 $PFQE$)称为最大质量可达区。

下面结合图3,对制约众包系统演化的3个关键因素进行分析,并讨论它们对众包系统质量控制的影响。

1) 众包节点的消费者效用 $B_{-k} - C_{-k}$, $k = i, j$. 由图3可以看到,当节点充当消费者角色时,随着各点消费效用的增加, F 点向左下移动,区域 $PFQE$ 的面积增加,众包系统演化到稳定点 E 的可能性随之增加。另外,当节点获得另一子点提交正确任务结果的代价 C_{-k} 越小,它的消费效用就越高,而在众包任务协作系统中,节点获得另一子点提交正确任务结果的代价主要用于对其他节点参与众包任务协作的激励。因此,当Club的汇点在为该Club中节点设计众包任务激励机制时,首先需要满足其子点消费效用的最大化。尤其,当一个Club同时与多个Club存在众包任务协作博弈时,诸如生物领域的蛋白质折叠等众包任务,此时的激励机制设计必须在促使多方正确完成众包任务的同时,确保用户消费效用的提高。

2) 完成众包任务承担的风险 C_k 。风险 C_k 与众包环境紧密相关,表现为节点选择合作策略后众包环境给其带来的风险水平,如可体现为完成众包任务会泄露的个人信息^[13]以及完成众包任务所产生的系统开销。这主要是基于对众包系统的数据安全和隐私保护的考虑。由图3可知,当 C_k 增加时, F 点向右上移动,区域 $PFQO$ 的面积增加,即系统演化到所有节点都合作参与众包任务的可能性减少。因此,需要从参与众包任务后系统资源的安全性和众包环境的可信性等方面来考虑如何降低风险。由

于相同的众包环境可能带来不同的风险水平,因而,汇点在决策Club间众包任务协作关系时,必须考虑发布任务者的可信度,降低众包任务协作给所有子点带来的风险。

3) 完成众包任务产生的利益 B_k 。由图3可知,当参与众包任务带来的风险 C_k 固定时,随着 B_k 的增加, F 向左下移动,区域 $PFQE$ 的面积增加,即众包系统演化到稳定点 E 的概率增加,越来越多的参与者会选择合作策略。因此,在承担相同风险的前提下,汇点应该尽力保证提供参与者的利益极大化。此外,由图3也容易看出,当风险 C_k 变化时,极大化参与者的利益 B_k 并不一定总使得区域 $PFQE$ 的面积增加。因此,汇点需要考虑众包环境的动态性,既关心本Club子点获得的利益,又关心其承担的风险,这样才能最终促进Club联盟建立稳定的众包任务协作关系,使众包质量朝最大化方向发展。

4 结论

1) 本文基于演化博弈论的方法,研究分层组织管理模式下众包系统质量的演化原理。研究结果表明:众包系统质量的演化方向受系统的初始状态影响,任务完成者参与众包任务获取的经济利益、参与众包任务带来的安全风险以及任务发布者的消费效用都是制约实现众包质量最大化的关键要素。

2) 下一步可以以本文分析为理论依据,设计众包质量控制策略或方法,并针对现有的众包应用实例实验检验提出策略或方法的效果。力图通过追求任务请求者消费效用和经济利益的极大化,以及尽量减少任务完成者参与众包带来的安全风险,使得更多的众包系统初始状态成为最大质量可达态,增加众包系统演化到所有节点都参与众包任务协作的可能性,实现众包质量最大化的目标。

参考文献:

- [1] 冯剑红,李国良,冯建华. 众包技术研究综述[J]. 计算机学报, 2015, 38(9): 1713-1725.
FENG J H, LI G L, FENG J H. A survey on crowdsourcing [J]. Chinese Journal of Computer, 2015, 38(9): 1713-1725. (in Chinese)
- [2] HOWE J. The rise of crowdsourcing [J]. Wired Magazine, 2006, 14(6): 1-4.
- [3] MASON W A, WATTS D J. Financial incentives and the "performance of crowds" [J]. ACM SIGKDD Explorations Newsletter, 2009, 11(2): 100-108.
- [4] CARSTEN E, ARJEN P V. Increasing cheat robustness of

- crowdsourcing tasks [J]. *Information Retrieval*, 2013 (2): 49-57.
- [5] FAN Y, SUN H L, ZHU Y M, et al. A truthful online auction for tempo-spatial crowdsourcing tasks [C] // *The IEEE Symposium on Service-Oriented System Engineering*. Washington, DC: IEEE Computer Society, 2015: 332-338.
- [6] XU A F, FENG X N, TIAN Y. Revealing, characterizing, and detecting crowdsourcing spammers: a case study in community Q&A [C] // *The 34th IEEE Conference on Computer Communications*. Amsterdam: Computer Communications, 2015: 2533-2541.
- [7] 张志强, 逢居升, 谢晓芹, 等. 众包质量控制策略及评估算法研究 [J]. *计算机学报*, 2013, 36(8): 1636-1649.
ZHANG Z Q, PANG J S, XIE X Q, et al. Research on crowdsourcing quality control strategies and evaluation algorithm [J]. *Chinese Journal of Computer*, 2013, 36(8): 1636-1649. (in Chinese)
- [8] 何云华, 孙利民, 杨卫东, 等. 基于博弈分析的众包交通监测隐私保护机制 [J]. *电子与信息学报*, 2016, 38(2): 340-346.
HE Y H, SUN L M, YANG W D, et al. Enhancing privacy preserving for crowdsourced monitoring—a game theoretic analysis based approach [J]. *Journal of Electronics & Information Technology*, 2016, 38(2): 340-346. (in Chinese)
- [9] YANG D J, XUE G L, FANG X, et al. Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing [C] // *The 18th Annual International Conference on Mobile Computing and Networking*. New York: ACM, 2012: 173-184.
- [10] SUN J, MA H. A behavior-based incentive mechanism for crowd sensing with budget constraints [C] // *The IEEE International Conference on Communications*. Washington, DC: IEEE Computer Society, 2014: 1314-1319.
- [11] LI J, ZHU Y M, HUA Y Q, et al. Crowdsourcing sensing to smartphones: a randomized auction approach [C] // *The IEEE International Symposium on Quality of Service*. Washington, DC: IEEE Computer Society, 2015: 219-224.
- [12] RANGANATHAN K, RIPEANU M, SARIN A, et al. Incentive mechanisms for large collaborative resource sharing [C] // *The 4th IEEE/ACM International Symposium on Cluster Computing and the Grid*. Washington, DC: IEEE Computer Society, 2004: 1-8.
- [13] DASILVA L A, SRIVASTAVA V. Node participation in adhoc and peer-to-peer networks: a game-theoretic formulation [C] // *The First Workshop on Games and Emergent Behaviors in Distributed Computing Environments With PPSN*. New York: Springer, 2004: 1-8.
- [14] CHEN F, ZHAN S B, SHI G J, et al. Crowdsourcing leakage of personally identifiable information via Sina microblog [C] // *The First International Conference on Internet of Vehicles*. New York: Springer International Publishing, 2014: 262-271.
- [15] SHEN Y, HUANG L S, LI L, et al. Towards preserving worker location privacy in spatial crowdsourcing [C] // *The IEEE Global Communications Conference*. Washington, DC: IEEE Computer Society, 2015: 1-6.
- [16] GONG Y M, WEI L B, GUO Y X, et al. Optimal task recommendation for mobile crowdsourcing with privacy control [J]. *IEEE Internet of Things Journal*, 2016, 5(3): 745-756.
- [17] ZHU B, ZHU S, LIU X J, et al. A novel location privacy preserving scheme for spatial crowdsourcing [C] // *The 6th IEEE International Conference on Electronics Information and Emergency Communication*. Washington, DC: IEEE Computer Society, 2016: 34-37.
- [18] KOCHHAR S, MAZZOCCHI S, PARITOSH P. The anatomy of a large-scale human computation engine [C] // *The ACM SIGKDD Workshop on Human Computation*. New York: ACM, 2010: 10-17.
- [19] SMITH J M. *Evolution and the theory of games: III* [M]. Cambridge: Cambridge University Press, 1982.
- [20] FRIEDMAN D. Evolutionary games in economics [J]. *Econometrica*, 1991, 59(3): 637-666.

(责任编辑 梁 洁)