

三方众包市场中的发包方-平台博弈机制设计

何雨橙 丁尧相 周志华
(计算机软件新技术国家重点实验室(南京大学) 南京 210023)
(heyc@lamda.nju.edu.cn)

Mechanism Design for Requester-Platform Strategies Under the Three-Party Crowdsourcing Market

He Yucheng, Ding Yaoxiang, and Zhou Zhihua
(National Key Laboratory for Novel Software Technology (Nanjing University), Nanjing 210023)

Abstract Crowdsourcing usually involves multiple parties of participants with their objectives. One of the fundamental challenges for crowdsourcing is to design effective mechanisms to make all the parties obtain their benefits besides competing. Even though fruitful previous studies have been conducted on this topic, they are usually based on the two-party crowdsourcing models under which one or more requesters and a crowd of workers are involved. However, in real-world applications, the requesters usually interact with the workers through the crowdsourcing platforms, making up the three-party crowdsourcing markets, under which the mechanism design for the requester-platform interaction doesn't receive previous study. In this work, we model the three-party crowdsourcing market as a game with incomplete information. We show that the Nash Equilibrium of this game can be found via regret minimization with proper online learning strategies. Under the single-requester setting, we show that the classical EXP3 algorithm is optimal for the requester, meanwhile, we propose a stronger strategy for the platform based on the counterfactual regret minimization technique. We also propose effective strategies for both platform and requesters in multiple requesters setting by generalizing the single-requester strategies. The performance of the proposed strategies is verified from experiments with both synthetic and real-world datasets.

Key words crowdsourcing; game theory; mechanism design; online learning; counterfactual regret minimization

摘 要 众包(crowdsourcing)通常涉及到目标各不相同的多个参与者,设计有效的众包机制,使得各个参与者在竞争中实现共赢,是众包理论研究中的基本问题之一.当前,众包机制设计通常基于发包方-标注者直接进行交互的两方博弈模型,而现实应用中,发包方与标注者之间往往通过平台进行交互,从而构成三方博弈下的众包市场.其中的发包方-平台博弈机制设计是过往众包研究中未曾涉及的全新问题.将三方众包市场建模为不完全信息博弈,并证明该博弈问题的 Nash 均衡可通过在线学习来最小化发包方和平台的累计遗憾而达到.在单发包方情形下,证明经典的 EXP3 算法对于发包方的最优性,并基于反事实遗憾最小化技术为平台设计了有效策略.同时,将单发包方情形下发包方和平台策略拓展到多发包方情形下并给出理论分析.合成及真实数据集上的实验验证了该方法的有效性.

关键词 众包; 博弈理论; 机制设计; 在线学习; 反事实遗憾最小化

中图法分类号 TP18

随着机器学习任务的规模逐渐增大,人们迫切需要对大规模数据进行收集.众包(crowdsourcing)作为低成本高效率的数据收集方式,受到了广泛欢迎.

众包研究的基本问题之一是设计有效的机制以使得参与者在竞争中实现共赢.当前,众包机制设计研究往往基于两方众包模型:发包方(requester)发布任务并支付标注者(workers)费用;标注者完成任务并收取报酬^[1].该模型的重要假设在于发包方和标注者可以直接进行交互.而现实应用中,如图1所示,发包方和标注者的交互往往以平台(platform)为中介,构成三方众包市场.其中,发包方将任务和报酬发布给平台,平台雇佣标注者进行标记,进而在将标记结果反馈给发包方的同时,赚取支付给标注者的费用和发包方支付给自己的费用之间的差价.显然,传统的两方众包模型无法对该过程进行建模,因此需要引入全新的三方众包模型进行研究.

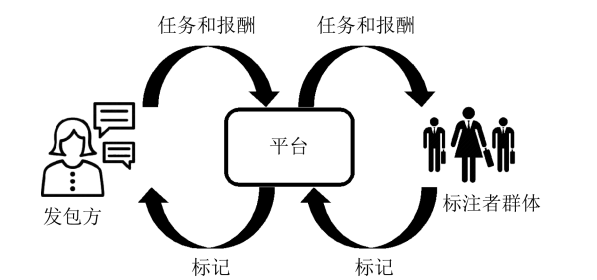


Fig. 1 The three-party crowdsourcing market
图1 三方众包市场示意图

相比于两方众包,三方众包的核心问题是发包方与平台之间的博弈:发包方希望支付较少的报酬同时获取准确率较高的标记;而平台则希望降低雇佣标注者的成本,同时从发包方处获取较多的报酬.这之中存在着复杂的博弈关系.一方面,发包方和平台既有合作也有竞争:双方都希望最大化标记的准确率,但在最小化或最大化发包方支付这一点上有冲突.另一方面,发包方和平台各自只能掌握自身信息,而无法直接观测到对方信息.在不完全信息下采取最优策略,对双方都是相当具有挑战性的问题.

本文开启三方众包市场中的发包方-平台博弈机制设计研究,主要贡献有4点:

- 1) 提出不完全信息博弈^[2]模型 CrowdMarket 对三方众包市场进行建模,并证明通过设计合适的在线学习策略可以近似达到该博弈的 Nash 均衡;
- 2) 在单发包方设定下,证明了 EXP3 算法^[3]为

发包方的最优策略,进而本文设计了基于反事实遗憾最小化(counterfactual regret minimization, CFR)技术^[4-5]的平台策略,证明该策略能够充分利用平台方的有效信息,具有比直接应用传统在线学习方法更强的理论保证;

- 3) 将单发包方的策略拓展到多发包方的情形,并给出了多发包方情形下的理论分析;
- 4) 通过合成及真实数据集上的实验验证了方法的有效性.

1 相关工作

众包研究的核心问题之一在于如何平衡标记质量和支付费用^[6-7].现有研究中,实现这一目标的方式可以分成2类:设计更好的标记推断方法以及设计更好的众包机制.标记推断已经有了丰富的研究,例如文献^[8-11].本文则着力于探讨机制设计问题.

两方众包机制设计已经得到了充分研究,其中包括任务与报酬的分配机制设计^[12-15],以及利用提供线索^[16]或跳过选项^[17-19]等方式提升高难度样本的标记质量.由于这些机制都是针对标注者的,因此这些策略同样可以应用于三方众包市场中作为平台与标注者之间的博弈机制.并且由于这方面的研究相对成熟,本文不再对此进行深入探讨.同时,也有大量的研究关注如何设计合适的支付策略以激励标注者给出高质量的标记.这方面的代表性工作包括文献^[20-23].在我们的问题设定中,这些激励机制可以被平台用于激励标注者给出更准确的标记,但是不能直接用于处理平台和发包方之间的博弈.

近期以来,一些工作也开始从应用的角度关注三方众包市场.例如文献^[24]提出了一种声誉评价方法以防止发包方和标注者的欺骗行为,而文献^[25]则将众包市场建模为一个3层的优化问题以最大化平台的健康度.但这些工作均不涉及博弈机制设计研究,因而它们的研究动机与本文有着显著的差异.

当前,将众包形式化为博弈问题,特别是预算有限条件下的收益最大化问题,是一个重要的理论研究方向^[26-28].这个方向上的现有研究主要集中于传统两方设定,本文所提出的三方众包模型,对拓展这一方向的研究内容可能起到有益的作用.此外,近期也有工作研究如何在广告市场中利用机器学习方法

从数据中学习有效的三方博弈机制^[29].本文的研究也将将这种新方法引入众包博弈问题提供了有益启示.

2 CrowdMarket 模型

2.1 基本定义

本节提出三方众包模型 CrowdMarket,将三方众包市场形式化为发包方、平台以及标注者三方的不完全信息博弈^[2].本节讨论单发包方情形,第4节中将对多发包方情形进行讨论.

单发包方情形下的 CrowdMarket 模型为持续 T 轮的博弈,在第 t 轮下:

1) 发包方发送一个包含 b 个任务的任务包,以及支付报酬 $x' \in (0, 1]$ 给平台.本文假设发包方只能从有限数量的 K 个报酬选项 $X_1, X_2, \dots, X_K$ 中选择合适的报酬.

2) 在收到报酬和任务包之后,平台选择 m' 个标注者完成标注任务.

3) 标注者各自选择采取低等级的努力或者高等级的努力以完成任务.高努力下准确率为 $p^* > 1/2$;而低努力下标注者随机猜测以返回标记,准确率为 $1/2$.之后每位标注者各自独立地为每个任务进行标注,并将结果反馈给平台.

4) 平台将每位标注者给出的标记通过多数投票法进行集成,并将集成后标记返回给发包方.集成后标记的实际平均准确率记为 a' .发包方从返回的标记中得到的收益取决于该准确率以及支付给平台的报酬,因此可以表示为 $a' - x'$.

5) 平台从发包方的支付中取出一部分作为参与了本轮标注的标注者的报酬,其余作为自己的收益.

为了激励标注者给出高质量的标记,平台需要得知标注者给出的标记的准确程度,这就需要发包方向平台提供反馈.因此,我们引入了这样一个步骤:发包方在每一轮结束之后会向平台反馈该轮标记的准确率.参考文献^[17],可以设计激励机制来确保发包方会诚实地反馈标记准确率,在3.1节中我们将会给出该激励机制并进行分析.另外,本文假设平台只能有部分轮次得到真实标记.因此,平台不能在任意轮次中直接推断出标记准确率.

每一轮中参与者可能选择的动作组成的集合称为动作集.具体而言,发包方的动作集为 $A_{\text{req}} = \{(x, a') : x \in \{X_1, X_2, \dots, X_K\}, a' \in [0, 1]\}$,其中 x 代表支付给平台的报酬, a' 代表汇报的准确率;平台的

动作集为 $A_{\text{pla}} = (m, c) : m \in \{1, 2, \dots, N\}, c \in \mathbb{R}^N$,其中 m 代表平台选择标注的人数, c 代表为每一位标注者支付的报酬组成的向量($m < N$ 时,多出的向量元素值为0), N 代表平台能选择的标注者最大人数;标注者的动作集为 $A_{\text{wor}} = \{\text{高等级努力, 低等级努力}\}$.在博弈过程的每一轮中,每位参与者可能会以一定的概率分布从各自的动作集中随机选择1个动作.这样随机选择时,每个动作被选中的概率分布称为(混合)策略.每位参与者的所有策略组成的集合称为策略空间.分别记发包方、平台和标注者群体的策略以及策略空间为 $\sigma_{\text{req}}^t \in \Sigma_{\text{req}}, \sigma_{\text{pla}}^t \in \Sigma_{\text{pla}}, \sigma_{\text{wor}}^t \in \Sigma_{\text{wor}}$.

2.2 CrowdMarket 博弈的分解

CrowdMarket 模型是一个多方非零和博弈^[30],这是博弈论中最难分析的博弈类型.但是,该模型有特殊的结构:发包方和标注者之间并没有直接的交互.利用这一点,我们可以将 CrowdMarket 模型分解成2个部分:发包方和平台之间的博弈以及平台和标注者之间的博弈.其中,平台和标注者之间的博弈与传统的两方众包中发包方和标注者之间的博弈是类似的,因此平台需要设计机制以激励标注者采取高等级的努力.并且,需要同时引入激励机制使发包方诚实反馈准确率.

3 平台激励机制设计

本节介绍平台需要采用2个激励机制:1)激励发包方诚实反馈准确率信息的机制;2)激励标注者采取高等级努力的机制.

3.1 激励发包方诚实反馈的机制

在 CrowdMarket 模型中,平台可以获得发包方反馈的准确率信息 a'^t .但是,发包方可能会试图通过反馈一个虚假的准确率 a'^t 进行欺诈.为了防止这一点,受到文献^[17]中“没有免费的午餐”(no-free-lunch)原则的启发,我们为平台设计了一个惩罚机制以防止发包方欺诈.

本文假设在 CrowdMarket 博弈过程中,平台可以随机选取某些轮次通过第三方得到真实标记,而在这些轮次中对发包方反馈的标记准确率进行验证.如果验证发现发包方反馈的准确率不正确,即发包方在该轮进行了欺诈,那么平台将会对发包方进行一定的惩罚.

我们首先定义指示变量序列 $y^t, t = 1, 2, \dots, T$ 如下:

$$y^t = \begin{cases} 1, & \text{如果发包方在第 } t \text{ 轮反馈的收益真实,} \\ 0, & \text{如果平台在第 } t \text{ 轮没有进行检查,} \\ -1, & \text{如果发包方在第 } t \text{ 轮反馈的收益虚假.} \end{cases}$$

惩罚机制可以表示为指示变量序列的一个函数 $f(y^1, y^2, \dots, y^T)$: 在众包博弈结束后, 发包方需要支付 $f(y^1, y^2, \dots, y^T)$ 给平台作为欺诈的惩罚. 为了确保公平性, 惩罚机制需要满足 2 个基本条件:

第 1 个条件是平台不能在发现发包方欺诈的情形下进行惩罚; 相反, 如果发包方被发现在每轮都作弊了, 那么需要支付最大数额的惩罚金额 $F_{\max}$. 该条件可以形式化为定义 1.

定义 1. 如果惩罚机制 f 满足:

1) 若 $y^t \neq 1$ 对所有 $t=1, 2, \dots, T$ 成立且存在 t 使得 $y^t = -1$, 则 $f(y^1, y^2, \dots, y^T) = F_{\max}$.

2) 若 $y^t = 1$ 对所有 $t=1, 2, \dots, T$ 成立, 则 $f(y^1, y^2, \dots, y^T) = 0$.

则称 f 满足“没有免费的午餐”条件.

第 2 个条件是发包方应该支付惩罚当且仅当其被发现有欺诈行为. 该条件可形式化为定义 2.

定义 2. 如果惩罚机制 f 满足: 对于任意轮次 $t \in \{1, 2, \dots, T\}$ 以及任意

$$(y^1, \dots, y^{t-1}, y^{t+1}, \dots, y^T) \in \{-1, 0, 1\}^{T-1}$$

都有

$$\begin{aligned} f(y^1, \dots, y^{t-1}, 1, y^{t+1}, \dots, y^T) &= \\ f(y^1, \dots, y^{t-1}, 0, y^{t+1}, \dots, y^T) &\leq \\ f(y^1, \dots, y^{t-1}, -1, y^{t+1}, \dots, y^T), \end{aligned}$$

那么称 f 满足“激励相容”条件.

基于这 2 个条件, 我们提出如下机制:

$$f(y^1, y^2, \dots, y^T) = F_{\max} \left(1 - \prod_{t=1}^T I(y^t \neq -1) \right), \quad (1)$$

其中, $I(y^t \neq -1)$ 为指示函数.

式(1)表明: 如果被发现在众包过程中有欺诈行为, 发包方会支付最大数额的惩罚. 显然, 式(1)所述的机制满足我们提出的 2 个条件. 我们可以进一步证明定理 1 成立, 定理 1 表明了由式(1)定义的惩罚机制是满足我们所要求的 2 个条件的唯一机制.

定理 1. 当 $F_{\max} \geq T^2$ 时, 式(1)机制是唯一满足“没有免费的午餐”条件和“激励相容”条件的机制, 且发包方最优策略是每轮诚实反馈收益信息.

证明. 假设 f 是同时满足定理要求的 2 个条件的机制. 我们只需要证明: 当所有的指示变量 $y^i \neq -1$ 时 $f(y^1, y^2, \dots, y^T) = 0$, 否则 $f(y^1, y^2, \dots, y^T) = T$ 即可.

1) 当 $\forall i \in [t]: y^i \neq -1$ 时, 我们假设 $y^1, y^2, \dots, y^T$ 中有 k 个 1 和 $T-k$ 个 0. 不失一般性地, 可以假设前 k 个变量取值为 1, 其余变量取值为 0. 由“没有免费的午餐”可知:

$$\begin{aligned} f(\underbrace{1, \dots, 1}_{k \text{ 个}}, 0, \dots, 0) &= f(\underbrace{1, \dots, 1}_{k+1 \text{ 个}}, 0, \dots, 0) = \dots = \\ f(1, 1, \dots, 1) &= 0. \end{aligned}$$

2) 平台仅在第 t 轮进行了 1 次检查, 且 $y^t = -1$. 根据“没有免费的午餐”条件, 该情况下惩罚为 $F_{\max}$. 又由“激励相容”条件易知, 任何存在 $y^i \neq -1$, $i=1, 2, \dots, T$ 的情况下, 惩罚均不应小于该情况, 也为 $F_{\max}$.

综上即是该机制为唯一满足条件的机制.

接下来考虑该机制下发包方作弊能得到的额外收益. 考虑极端情况: 发包方只需要在其中 1 轮作弊, 未被发现即可得到最大收益 R , 被发现则要另外支付惩罚 T . 再设发包方不作弊得到的收益为 r , 平台检查的轮数所占比例为 p . 由于平台至少检查 1 次, 因此 $p \geq 1/T$. 对发包方而言, 不作弊是最优策略当且仅当

$$p(R - F_{\max}) + (1-p)R \leq r,$$

解不等式得 $p \geq (R-r)/F_{\max}$. 由于在 CrowdMarket 模型中, $R-r \leq T$, 因而上述不等式恒成立. 这表明诚实反馈是发包方的最优策略. 证毕.

3.2 激励标注者采取高等级努力的机制

作为博弈的参与方之一, 标注者在博弈中的目标也是最大化自身的期望收益. 因此平台只需要设计一个支付机制, 使得标注者采取高等级努力时的期望收益最大即可. 与平台激励发包方诚实反馈的机制类似, 平台可以利用第三方反馈的真实标记检查标注者给出标记的准确率, 并据此判断标注者是否采用低等级努力. 下文中为简化记号, 假定标注者 i 在 $t \leq T$ 轮中参与了标注, 构造指示变量序列 y_i^s , $s=1, 2, \dots, t$ 如下:

$$y_i^s = \begin{cases} 1, & \text{如果标注者 } i \text{ 在第 } s \text{ 次采取高等级努力,} \\ 0, & \text{在第 } s \text{ 次平台未检查,} \\ -1, & \text{如果标注者 } i \text{ 在第 } s \text{ 次采取低等级努力.} \end{cases}$$

标注者 i 获得的总收益为

$$u(y_i^1, y_i^2, \dots, y_i^t) = \sum_{s=1}^t c_i I(y_i^s \neq -1). \quad (2)$$

式(2)采用了一个简单的机制: 如果被发现在标注过程中有采取低等级努力行为, 标注者将无法得到该轮报酬, 否则等价于每一次标注都得到报酬 c_i . 由于标注者的目标是最大化其收益, 因而有定理 2:

定理 2. 在式(2)机制下,标注者的最佳策略为在所有轮次均采用高等级努力进行标注.

证明. 由于标注者的目标为最大化其收益,在任意 $t \leq T$ 轮中,显然只有当标注者采取高等级努力时,所收获的单轮回报最大,进而知定理结论成立. 证毕.

注意到,由于标注者所标记的样本是有限的,因而通过标注是否正确判断其努力程度存在一定错误的可能性.在实际问题中,不难通过允许用户对自身努力程度被错判的情况进行申诉,来消除该错误.

在本节给出的激励机制的保证之下,平台可以保证在每一轮中:1)发包方诚实汇报准确率;2)标注者选择高等级努力,因而给出的标记的准确率为 p^* ;3)平台向每位被分配任务的标注者支付报酬 c_i ,由于标注者之间相互等价(能力与标注策略均相同),每位标注者的报酬均相同.在上述标注者策略已经固定的情况下,后文将集中研究如何设计发包方和平台之间的博弈策略.另一方面,不妨将发包方与平台已经确定的动作从动作集中移除,以重点研究还未确定的动作:发包方-平台博弈的每一轮中,发包方的动作是选择支付给平台的报酬,平台的动作是选择分配的标注者人数.为此,在下文中我们重新定义发包方和平台的动作集分别为 $A_{\text{req}} = \{X_1, X_2, \dots, X_K\}$, $A_{\text{pla}} = \{1, 2, \dots, N\}$.发包方和平台的策略空间也相应地重新定义.

4 发包方-平台博弈策略设计

本节介绍发包方和平台在 CrowdMarket 博弈中采取的博弈策略.在 4.1 节我们证明发包方和平台可以使用在线学习算法最小化自身遗憾;4.2 节和 4.3 节分别给出发包方和平台基于在线学习算法的策略.

4.1 基于在线学习的博弈策略

每个参与者的目标均为最大化自身的累计收益,这等价于最小化自身的“遗憾”:

$$R_i^T = \max_{\sigma_i^* \in \Sigma_i} \sum_{t=1}^T (u_i(\sigma_i^*, \sigma_{-i}^t) - u_i(\sigma_i^t, \sigma_{-i}^t)),$$

其中, $i \in \{\text{req}, \text{pla}\}$, σ_i^t 表示局中人 i 在第 t 轮采用的策略, σ_{-i}^t 表示除对应局中人 i 之外其他局中人在第 t 轮采取的策略, $u_i(\sigma_i^t, \sigma_{-i}^t)$ 表示局中人 i 的收益.称策略组合 $\{(\sigma_{\text{req}}^t, \sigma_{\text{pla}}^t)\}_{t=1}^T$ 是 ϵ -最优的,当且仅当 $\frac{R_{\text{req}}^T}{T} \leq \epsilon_1$, $\frac{R_{\text{pla}}^T}{T} \leq \epsilon_2$, 且 $\epsilon = \max\{\epsilon_1, \epsilon_2\}$.容易看出,当 ϵ -最优性满足时,博弈各方都满足了遗憾最小化

的性质.而下述定理给出了策略组合的最优性与 Nash 均衡之间的联系.

定理 3. 如果 $\{(\sigma_{\text{req}}^t, \sigma_{\text{pla}}^t)\}_{t=1}^T$ 是 ϵ -最优的,那么对应的平均策略组合 $(\bar{\sigma}_{\text{req}}, \bar{\sigma}_{\text{pla}})$ 是一个 ϵ -Nash 均衡^[2], 其中: $\bar{\sigma}_i = \frac{1}{T} \sum_{t=1}^T \sigma_i^t, i \in \{\text{req}, \text{pla}\}$.

证明. 首先以平台为例进行证明.由 ϵ -最优性的定义有

$$\frac{R_{\text{pla}}^T}{T} = \frac{1}{T} \max_{\sigma_{\text{pla}}^* \in \Sigma_{\text{pla}}} \sum_{t=1}^T (u_{\text{pla}}(\sigma_{\text{req}}^t, \sigma_{\text{pla}}^*) - u_{\text{pla}}(\sigma_{\text{req}}^t, \sigma_{\text{pla}}^t)) \leq \epsilon,$$

进而,由于平台收益是线性函数,我们可以得出:

$$\begin{aligned} u_{\text{pla}}(\sigma_{\text{req}}, \sigma_{\text{pla}}) + \epsilon &= \frac{1}{T} \sum_{t=1}^T u_{\text{pla}}(\sigma_{\text{req}}^t, \sigma_{\text{pla}}^t) + \epsilon \geq \\ \frac{1}{T} \max_{\sigma_{\text{pla}}^* \in \Sigma_{\text{pla}}} \sum_{t=1}^T u_{\text{pla}}(\sigma_{\text{req}}^t, \sigma_{\text{pla}}^*) &= \max_{\sigma_{\text{pla}}^* \in \Sigma_{\text{pla}}} u_{\text{pla}}(\sigma_{\text{req}}, \sigma_{\text{pla}}^*). \end{aligned}$$

发包方的效用函数可类似地表示为策略的线性函数,从而同理可证

$$u_{\text{req}}(\sigma_{\text{req}}, \sigma_{\text{pla}}) + \epsilon \geq \max_{\sigma_{\text{req}}^* \in \Sigma_{\text{req}}} u_{\text{req}}(\sigma_{\text{req}}^*, \sigma_{\text{pla}}). \quad \text{证毕.}$$

由于 Nash 均衡状态代表了各方策略同时达到稳定状态,因而定理 1 表明,博弈各方只需采用能够对遗憾进行最小化的学习策略,就能在竞争中合作共赢.

4.2 发包方策略

发包方策略设计面临的主要挑战在于缺少信息:发包方无法得知之前轮次中被分配任务的标注者的具体数目.因而,发包方很难对标注者能力 p^* 及平台策略 σ_{pla}^t 进行估计.因此,发包方需要采用能保证最坏情形下收益的策略.

注意到发包方可以将博弈过程建模为赌博机 (bandit) 在线学习问题:发包方可能选择的支付动作 A_{req} 可视作赌博机摇臂,博弈产生的收益可以作为选择特定摇臂之后得到的收益.进而,发包方可以采用文献[3]提出的 EXP3 算法作为其策略.下面的定理给出了任何发包方策略的遗憾下界,进而验证了 EXP3 策略的最优性.

定理 4. 在最坏情况下,任何发包方策略的期望遗憾至少为 $\frac{1}{20} \min\{\sqrt{KT}, T\}$.

证明. 考虑这样一种情况,标注者能力 $p=1$,即标注者总是返回真实标记;平台通过以一定概率随机反转集成后标记的方式控制返回给发包方的标记准确率.假设当发包方选择支付给平台的报酬为 x 时,平台翻转标记的概率为 $1-q(x)$,即平台提供的标记的期望准确率为 $a=q(x)$.

根据文献[3]中的定理 5,如果发包方得到的效用 $u=q(x)-x$ 满足以下条件,则发包方的遗憾至少为 $\frac{1}{20}\min\{\sqrt{KT}, T\};1)$ 存在一个最好的支付选择 x_0 使得发包方的期望收益为 $\alpha+\epsilon$;2)当选择除了 x_0 之外的其他支付时,发包方获得的期望支付为 α .

考虑函数 $q(x)=x+\alpha+\epsilon I(x=x_0), I(\cdot)$ 表示指示函数,并且 α 足够小从而使得 $q(x)\in[0,1]$.显然此时发包方的期望收益 $u=q(x)-x=\alpha+\epsilon I(x=x_0)$ 满足以上条件. 证毕.

这与 EXP3 的遗憾上界同阶,因而 EXP3 是发包方在缺乏信息的条件下所能采用的最优策略.值得注意的是,平台也可以由定理 4 得知发包方会使用 EXP3 策略.在 4.3 节我们会展示这一优势是如何帮助平台制定策略的.

4.3 平台策略

尽管平台也可以使用 EXP3 策略并得到 $O(\sqrt{T})$ 阶的遗憾上界,但是这样的策略并不是最优的.其原因在于 EXP3 并不能利用到平台掌握的全部信息:由于平台知道每一轮分配的标注者的具体数目,再结合发包方反馈的收益信息,就可以推断标注者能力及后续的标注准确率;同时,再和发包方使用 EXP3 算法这一信息结合,平台就可以进一步预测出后续发包方会选择的支付策略;进而,平台可以直接模拟整个博弈过程求解出最优策略.但是,基于历史信息估计标注者能力是有误差的,特别是在博弈的早期阶段,

此时信息很不充足,对标注者能力的估计误差会很大.解决该挑战的思路是,在每轮博弈中基于文献[4]提出的 CFR 技术模拟之后的博弈过程.基于 CFR 的平台博弈策略(以下简称 CFR 策略)如算法 1 所示:

算法 1. CFR 策略.

- ① 初始化置信区间 $I^1=[0,1], P^1=\{p:p\in I^1\};$
- ② 初始化历史记录 $H=\emptyset;$
- ③ for $t=1,2,\cdots,T$
- ④ 运行 CFR 模拟,根据式(3)更新策略 $\sigma_{\text{pla}}^t;$
- ⑤ 根据策略 σ_{pla}^t 选择分配的标注者人数 $m^t;$
- ⑥ 完成众包并接收发包方反馈的收益 $a^t;$
- ⑦ 将 (m^t,a^t) 添加到 H 之中;
- ⑧ 由历史 H 计算选 m^t 个标注者时发包方的平均收益 $\bar{f};$

- ⑨ 由历史 H 计算标注者的平均标注者能力 $\bar{p};$
- ⑩ 更新 I^{t+1} 为
- $$I^{t+1}=\left[\bar{p}-\sqrt{\frac{1}{bt}\ln\frac{2}{\delta}},\bar{p}+\sqrt{\frac{1}{bt}\ln\frac{2}{\delta}}\right];$$
- ⑪ 更新 $P^{t+1}=\{p:p\in I^{t+1}\};$
- ⑫ end for

该策略的关键思想是通过模拟后续博弈过程求解出最小化反事实遗憾的策略(算法 1 行④).假设在第 t 轮平台运行了一个总轮数为 S 的模拟博弈;平台的反事实收益 $\tilde{u}_{\text{pla}}(\sigma_{\text{req}}^t,\sigma_{\text{pla}}^t)$ 定义为平台在模拟博弈中得到的期望累计收益;在模拟博弈中,标注者能力是从可能取值的集合 P^t 中随机挑选的;进而,平台的反事实遗憾可表示为

$$R_{\text{pla}}^S(P^t,m)=\sum_{s=1}^{S(t)}\sum_m[\sigma_{\text{pla}}^s]_m(\tilde{u}_{\text{pla}}(\sigma_{\text{req}}^s,m)-\tilde{u}_{\text{pla}}(\sigma_{\text{req}}^s,\sigma_{\text{pla}}^s)),$$

其中, $(\sigma_{\text{req}}^s,m)$ 代表这样的策略组合:发包方策略为 σ_{req}^s ,而平台则始终选择 m 个标注者, σ_{req}^s 代表模拟博弈中的发包方策略; $[\sigma_{\text{pla}}^s]_m$ 代表平台策略中选择 m 个标注者的概率.在模拟博弈的第 s 轮,平台采取遗憾匹配(regret matching)算法^[31] 计算其策略.令

$$R_{\text{pla}}^{s-1,+}(P^t,m)=\max\{R_{\text{pla}}^{s-1}(P^t,m),0\},$$

则有

$$[\sigma_{\text{pla}}^s]_m=\begin{cases} \frac{R_{\text{pla}}^{s-1,+}(P^t,m)}{\sum_m R_{\text{pla}}^{s-1,+}(P^t,m)}, & R_{\text{pla}}^{s-1,+}(P^t,m)>0, \\ 1/N, & \text{其他情况,} \end{cases} \tag{3}$$

其中, N 为平台最多可选择的标注者人数.

另一方面,直接应用原始的 CFR 算法无法达到理想效果.这是由于对于平台而言,标注者能力在博弈开始是未知的,必须通过对标注者能力进行有效估计,才能加快 CFR 算法的收敛速度.为了能更精确地得到对标注者能力的估计,算法 1 中引入了标注者能力估计步骤,利用 Hoeffding 不等式逐渐缩小标注者能力的可能取值范围(算法 1 行⑤~⑪,行⑩中 b 为每轮任务数, δ 为置信系数).随着博弈轮数的增加,参数区间会越来越紧,从而起到逐渐缩小标注者能力的可能取值集合的效果(算法 1 行⑪).下述定理表明应用该策略可以达到更紧的遗憾界.

定理 5. 当博弈总轮数 T 充分大时,以至少 $1-\delta$ 的概率,算法 1 中的策略的期望遗憾上界为 $O(\log T)$.

证明. 如果平台可以使用真实能力 p^* 进行模拟, 由文献[4]可知, 模拟得到的平台策略 $\sigma_{\text{pla}}^t(p^*)$ 和平台的最优策略 σ_{pla}^* 之间的差距将完全由模拟深度决定: 在第 t 轮, 在模拟博弈中由式(2)得到的策略满足

$$u_{\text{pla}}(\sigma_{\text{req}}^t, \sigma_{\text{pla}}^*) - u_{\text{pla}}(\sigma_{\text{req}}^t, \sigma_{\text{pla}}^t(p^*)) \leq O\left(\sqrt{\frac{1}{S(t)}}\right). \quad (4)$$

但是平台仅仅可以知道一个可能的标注者能力的取值集合 P^t . 因此, 模拟得到的结果是不准确的, 其中的误差会随着模拟深度的增加逐渐累积. 当模拟深度较大时, 累积的误差也会很大. 文献[32]中的定理 2.1 表明, 由 P^t 模拟得到的策略 $\sigma_{\text{pla}}^t(P^t)$ 和由 p^* 模拟得到的策略 $\sigma_{\text{pla}}^t(p^*)$ 之间的差异有关系:

$$u_{\text{pla}}(\sigma_{\text{req}}^t, \sigma_{\text{pla}}^t(p^*)) - u_{\text{pla}}(\sigma_{\text{req}}^t, \sigma_{\text{pla}}^t(P^t)) \leq S^2(t)\epsilon^t, \quad (5)$$

其中, ϵ^t 表示模拟过程和真实过程之间有差异的平均概率. 结合式(4)和式(5)可知, 总的遗憾上界为

$$R_{\text{pla}}^T \leq \sum_{t=1}^T \left[O\left(\sqrt{\frac{1}{S(t)}}\right) + S^2(t)\epsilon^t \right].$$

令 $g = \min\{p_i - p_j \mid p_i, p_j \in P^t\}$, $T_0 = \frac{4}{bg^2} \ln \frac{2}{\delta}$,

考虑 $t < T_0$ 和 $t \geq T_0$ 这 2 种情况.

当 $t < T_0$ 时, 由于平台每一轮的收益在 $[0, 1]$ 之间, 故平台每一轮的遗憾必定在 $[0, 1]$ 之间.

当 $t \geq T_0$ 时, 区间长度会缩短至仅有一个 $\hat{p} \in I^t$, 并且 $\Pr[\hat{p} = p^*] \geq 1 - \delta$, 从而令 $S(t) = O(t^2)$, 即可使得以至少 $1 - \delta$ 的概率, 平台的遗憾有如下的一个上界:

$$R_{\text{pla}}^T \leq \sum_{t=1}^{T_0} r(t) + \sum_{t=T_0+1}^T r(t) \leq T_0 + \sum_{t=T_0+1}^T \sqrt{\frac{1}{S(t)}} \leq T_0 + O(\log T),$$

其中, $r(t) = O\left(\sqrt{\frac{1}{S(t)}}\right) + S^2(t)\epsilon^t$. 证毕.

定理 5 说明算法 1 的遗憾界显著优于 EXP3 策略, 表明该策略充分利用了前文所述的额外信息.

5 多发包方情形下的策略

本节讨论多发包方 CrowdMarket 模型的博弈机制. 如果标注者群体可以同时为所有的发包方提供服务, 那么平台只需要和每一个发包方单独进行 CrowdMarket 博弈, 此时多发包方和单发包方的情形是完全一致的. 但是如果标注者群体在同一时间

只能为部分发包方提供服务, 那么发包方之间需要竞争服务的使用权. 因此, 本节针对后一种情况进行研究.

不失一般性, 假设一共有 n 个发包方参与博弈, 平台在每一轮中为出价最高的发包方提供服务. 同时, 假设单个发包方只能知道自己是否成功获得服务, 而无法得知其他发包方的出价以及任务完成准确率信息. 易知, 在任何一轮当中, 出价最高而得到服务的发包方的收益和单发包方的情形相同, 而未得到服务的发包方的收益则为 0.

与单发包方条件下类似, 在多发包方条件下, 发包方仍然面临着缺乏决策信息的问题: 不仅无法得知平台如何雇佣工人, 而且无法得知其他发包方的情况. 因而, 可以类似地证明发包方的最优策略为使用 EXP3 算法. 接下来我们证明: 平台也仍可以使用 CFR 策略模拟多发包方的情形, 以优化自身的遗憾.

定理 6. 在有 n 个发包方情形下, 当博弈总轮数 T 充分大时, 以至少 $1 - \delta$ 的概率, CFR 策略的期望遗憾上界为 $O(n(\log n + \log T))$.

证明. 通过和定理 5 相似的证明方法可知, 当平台为单个发包方服务轮数超过阈值 $T_1 = \frac{4}{bg^2} \ln \frac{2n}{\delta}$ 时, 可以至少 $1 - \delta/n$ 的概率准确估计与该发包方对应的标注者能力. 并且, 可知平台的遗憾表示为

$$R_{\text{pla}}^T = \sum_{i=1}^n \sum_{t=1}^T [u(\sigma_{\text{req}, i}^t, \sigma_{\text{pla}}^*) - u(\sigma_{\text{req}, i}^t, \sigma_{\text{pla}}^t)].$$

当 $T \geq nT_1$ 时, 至少有一个发包方在至少 T_1 轮赢得服务, 而对于赢得服务小于 T_1 轮的发包方, 与其博弈的遗憾界是常数阶. 进而知以至少 $1 - \delta$ 的概率有

$$R_{\text{pla}}^T \leq n(T_1 + O(\log T)) = O(n(\log n + \log T)).$$

证毕.

6 实验验证

本节对发包方-平台策略性能进行验证. 具体而言, 本节实验对 3 点进行验证: 1) 对于发包方, 当平台使用强对抗性的策略时, EXP3 策略是否有好的表现; 2) 对于平台, CFR 策略是否能利用额外信息给出更好的结果; 3) 对于多发包方的情形, 发包方 EXP3 策略及平台 CFR 策略是否依然适用. 实验中发包方和平台的动作集设定为: 1) 发包方可能的支付选项为 $\{0.1, 0.2, 0.3\}$; 2) 平台可能选择的标注者人数为 $\{1, 3, 5\}$; 3) 平台雇佣每个标注者的成本 $C = 0.01$.

本节实验使用 8 个二分类真实数据集:1)BM 数据集^[33],该数据集中标注者对语料给出正面或负面情绪标记;2)TEMP 数据集^[34],该数据集中标注者对 2 件事是否是先后发生的进行标记;3)WVSCM 数据集^[8],该数据集中标注者对图片中人脸是否微笑进行标记;4)WB 数据集^[35],该数据集中标注者对图片中的水鸟是否是鸭子进行标记;5)SpamCF 数据集^[36],该数据集中标注者对一个 AMT 平台上的任务是否是垃圾任务进行标注;6)MediaEval 数据集^[36],该数据集中标注者对给定图片是否和时尚有关进行标注;7)MEHCB 数据集^[37-38],该数据集中标注者对搜索请求和网页是否有关进行标记;8)RTE 数据集^[34],该数据集中标注者对文本之间是否有蕴含关系进行标注.实验所用 8 个数据集的相关信息如表 1 所示.

6.1 单发包方策略验证

本节在单发包方情形下,将发包方 EXP3 策略与 ϵ -贪心策略($\epsilon=0.05$)及固定策略(始终固定在最高支付)进行性能对比.同时,假设平台使用高对抗性甚至是作弊性质的策略,因为我们需要验证即便在最坏情形下 EXP3 策略仍然有效.

Table 1 Information About Datasets

表 1 实验所用数据集的相关信息

数据集	样本数	标注者数	标注者能力
BM	1 000	83	0.696 0
MediaEval	3 532	208	0.917 6
MEHCB	3 532	492	0.773 8
RTE	800	164	0.856 2
SpamCF	101	153	0.653 5
TEMP	462	76	0.917 7
WB	108	39	0.759 3
WVSCM	160	64	0.700 0

在我们的实验中,平台和发包方用各自的策略进行持续 T 轮的 CrowdMarket 博弈.轮次上限 T 的取值分别设定为 10,15,⋯,40.本实验中假设平台可获取真实的标注者能力 p^* ,从而可以通过以一定概率翻转标记的方式准确控制返回给发包方的标记准确率.并且,平台会采用如下强对抗性的策略以诱导发包方提高支付:平台以一定的概率 q 翻转标记,之后如果平台收到了更高的报酬则会逐渐降低 q 的取值.平台采用的这一策略要求发包方逐渐提

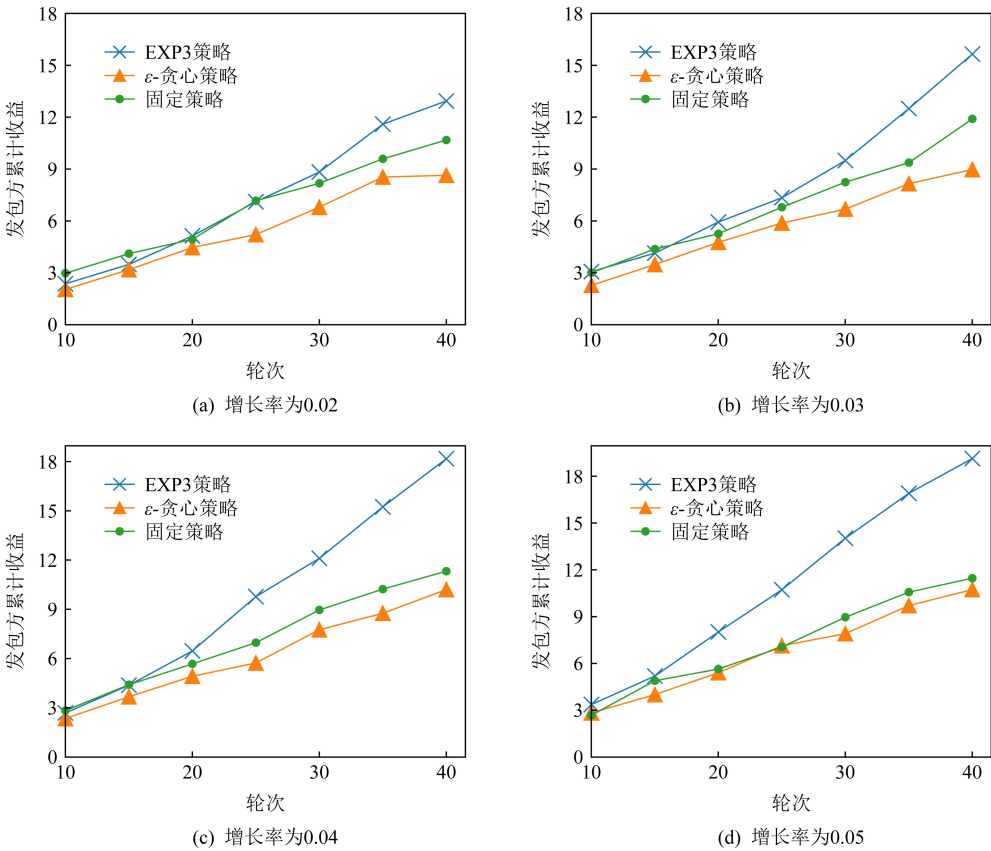


Fig. 2 The cumulative rewards of requester strategies under the single requester setting

图 2 单发包方情形下发包方策略的累计收益对比

高支付的报酬而不能只用贪心策略.本次实验中设定初始轮次中 $q = 0.50$,每次收到更高报酬后 q 的降低量分别设为 $0.02, 0.03, 0.04, 0.05$.在每组参数下我们重复实验 50 次并汇报平均累计收益.实验结果如图 2 所示,实验结果可见,当平台使用强对抗性的策略时,发包方使用 EXP3 策略获得的累计收益总是比 ϵ -贪心策略和固定策略要好,这表明了 EXP3 策略的有效性.

6.2 多发包方策略验证

本节验证了 3 个发包方情形下,发包方使用 EXP3 策略的有效性.博弈过程的参数设定同 6.1 节.

为了验证 EXP3 策略的有效性,在实验中我们令 3 个发包方分别使用 EXP3 策略、 ϵ -贪心($\epsilon = 0.05$)策略和固定策略,平台使用的策略为 CFR 策略.我们令 3 个发包方在 CrowdMarket 博弈中使用不同策略相互竞争,胜出的发包方所使用的策略就是这

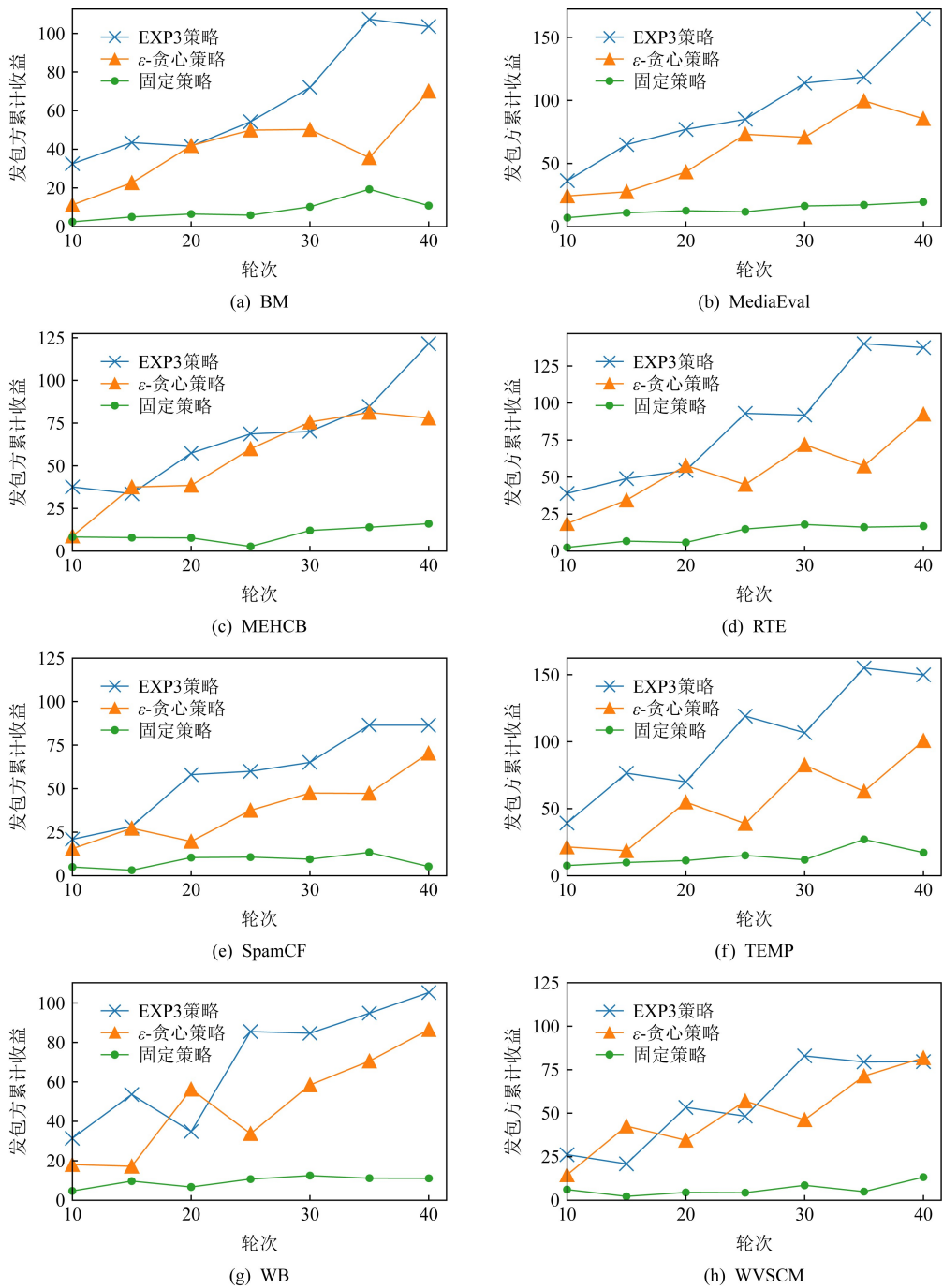


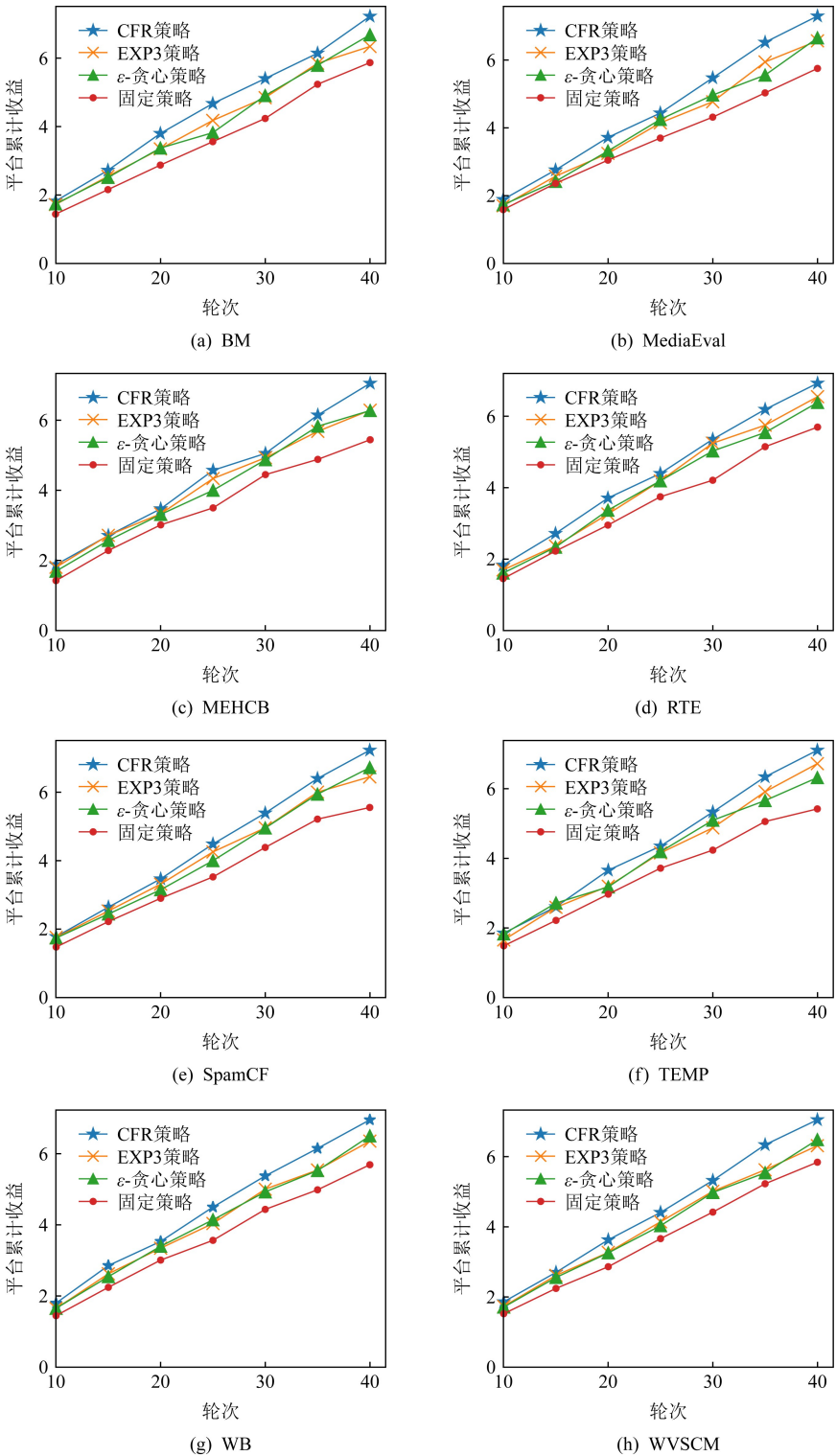
Fig. 3 The cumulative rewards of requester strategies under the multiple requester setting

图 3 多发包方情形下各发包方策略的累计收益对比

3 个策略中最优的策略.为了保证公平性,所有发包方使用的数据集都是一样的,以确保标注者能力对于所有发包方是一致的.我们在 8 个数据集上进行了实验,每个数据集上重复 10 次.发包方累计收益的平均值如图 3 所示.可以发现,使用 EXP3 策略在绝大多数时间内都能获得最多的收益,这表明 EXP3 策略在多发包方的情形下依然适用.

6.3 平台策略验证

为了验证平台在单发包方与多发包方情形下使用 CFR 策略的性能,我们将 CFR 策略和 EXP3 策略、 ϵ -贪心($\epsilon=0.05$)策略以及固定策略进行了对比,发包方的策略固定为 EXP3 策略.博弈过程的参数设置与 6.1 节相同.单发包方情形下我们在 8 个数据集上进行了实验,在多发包方情形下则测试了 2 组



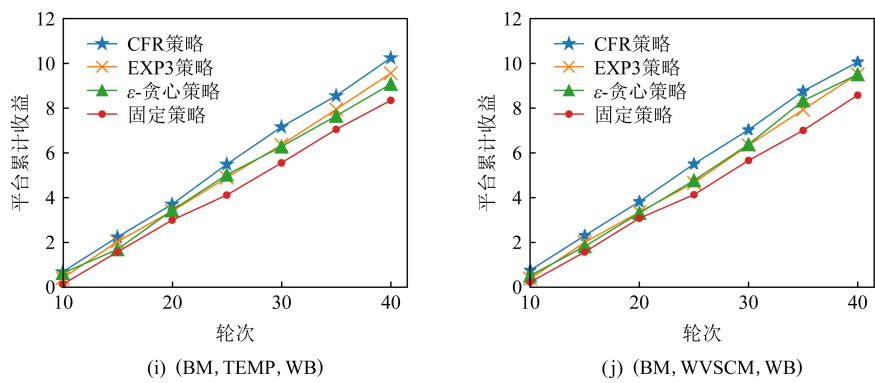


Fig. 4 The cumulative rewards of the platform strategies under the single requester and multiple requester settings

图 4 单发包方与多发包方情形下的平台策略累计收益对比

发包方数据集的组合,所有实验结果如图 4 所示.每张子图展示了平台的累计收益.实验结果显示:无论在哪个数据集上,性能最好的策略均为 CFR 策略,其次是 EXP3 策略,再次是 ϵ -贪心策略,排名最后的是固定策略.这表明利用到了额外信息的 CFR 策略确实能取得更好的效果.

表 2 展示了单发包方情形下,当平台使用不同策略,发包方使用 EXP3 策略时,40 轮之后平台和发包方的合计累计收益.结果表明平台使用 CFR 策略可以使得双方的累计收益达到最大.综合上述结果可知,CFR 策略是最适合于合作的策略.注意到平台使用 CFR 策略是在发包方反馈准确率信息时的最优策略,而平台使用 EXP3 策略是在发包方不反馈准确率信息时的最优策略.因此,表 2 中平台使用 CFR 策略时,累计收益超过 EXP3 策略,这表明 2.1 节中引入的反馈步骤可以提升双方的累计收益,对于双方的合作有促进作用.

Table 2 Total Rewards of Platform and Requester After 40 Round with Different Strategies of Platform

数据集	CFR 策略	EXP3 策略	ϵ -贪心策略	固定策略
BM	27.35	26.74	26.56	25.80
MediaEval	36.07	35.53	35.65	34.69
MEHCB	30.49	29.70	29.72	28.90
RTE	33.60	33.11	32.94	32.30
SpamCF	25.52	24.98	24.96	24.17
TEMP	36.07	35.55	35.61	34.72
WB	29.72	29.22	29.14	28.40
WVSCM	27.45	26.86	26.83	26.01

注:黑体数字表示最优结果.

6.4 实验结果分析

在本节中,我们利用仿真数据集和真实数据集进行了实验,对本文提出的基于在线学习方法的三方众包市场发包方-平台博弈策略进行了验证.实验结果表明,在符合 CrowdMarket 模型假设的条件下,本文提出的单发包方及多发包方策略不仅能优化自身的累计收益,而且能达到促进博弈双方合作共赢的目的.这验证了本文提出策略的有效性.另一方面,在实际应用中,也可能存在数据不符合 CrowdMarket 模型假设的情况.由于相关数据集的缺乏,难以验证在这一条件下本文方法的实际效果.我们会在未来研究中探索这一问题.

7 结束语

本文针对三方众包市场中的发包方-市场机制设计问题进行理论研究,提出三方众包市场模型 CrowdMarket.在该模型的基础上,针对单发包方和多发包方的设定,研究平台和发包方的策略设计和理论分析.真实数据集上进行的实验验证了本文所提出的策略的有效性.我们相信本文的研究结果可以激发更多针对三方众包市场的研究,有助于更好地理解现实应用中众包产业的市场行为.

作者贡献声明:何雨橙调研整理文献,实施方法研究,完成实验,撰写论文;丁尧相设计研究方案,实施方法研究,修订论文;周志华提出研究选题,指导方法研究与论文撰写支持.

参 考 文 献

[1] Zhang Jing, Wu Xindong, Shen V S. Learning from crowdsourcing labeled data: A survey [J]. Artificial Intelligence Review, 2016, 46(4): 543-576

- [2] Nisan N, Roughgarden T, Yardos E, et al. Algorithm Game Theorem [M]. Cambridge, UK: Cambridge University Press, 2007
- [3] Auer P, Cesa-Bianchi N, Freund Y, et al. The Nonstochastic multiarmed bandit problem [J]. SIAM Journal on Computing, 2002, 32(1): 48-77
- [4] Zinke V M, Johanson M, Bowling M, et al. Regret minimization in games with incomplete information [C] // Proc of the 20th Int Conf on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc, 2007: 1729-1736
- [5] Hu Yujing, Gao Yang, An Bo. Online counterfactual regret minimization in repeated imperfect information extensive games [J]. Journal of Computer Research and Development, 2014, 51(10): 2160-2170 (in Chinese)
(胡裕靖, 高阳, 安波. 不完美信息扩展式博弈中在线虚拟遗憾最小化[J]. 计算机研究与发展, 2014, 51(10): 2160-2170)
- [6] Wang Lu, Zhou Zhihua. Cost-saving effect of crowdsourcing learning [C] //Proc of the 25th Int Joint Conf on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2016: 2111-2117
- [7] Wang Wei, Zhou Zhihua. Crowdsourcing label quality: A theoretical study [J]. Science China Information Sciences, 2015, 58(11): 1-12
- [8] Whitehill J, Wu Tingfan, Bergsma J, et al. Whose vote should count more: Optimal integration of labels from labels of unknown expertise [C] //Proc of the 22nd Int Conf on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc, 2009: 2035-2043
- [9] Dekel O, Shamir O. Vox Populi: Collecting high-quality labels from a crowd [C/OL] //Proc of the 22nd Annual Conf on Learning Theory. New York: ACM, 2009 [2021-10-11]. <http://www.learningtheory.org/colt2009/papers/037.pdf>
- [10] Raykar V C, Yu Shipeng, Zhao L H, et al. Learning from crowds [J]. Journal of Machine Learning Research, 2010, 11 (43): 1297-1322
- [11] Yan Yan, Rosales R, Fung G, et al. Modeling annotator expertise: Learning when everybody knows a bit of something [C] //Proc of the 13th Int Conf on Artificial Intelligence and Statistics. Brookline, MA: Microtome Publishing, 2010: 932-939
- [12] Von Ahn L, Dabbish L. Labeling images with a computer game [C] //Proc of the 22nd SIGCHI Conf on Human Factors in Computing Systems. New York: ACM, 2004: 319-326
- [13] Ho C J, Vaughan J W. Online task assignment in crowdsourcing markets [C] //Proc of the 26th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2012: 45-51
- [14] Karger D R, Oh S, Shah D. Budget-optimal task allocation for reliable crowdsourcing systems [J]. Operation Research, 2014, 62(1): 1-24
- [15] Zheng Yudian, Wang Jiannan, Li Guoliang, et al. QASCA: A quality-aware task assignment system for crowdsourcing applications [C] //Proc of the 41st ACM SIGMOD Int Conf on Management of Data. New York: ACM, 2015: 1031-1046
- [16] Han Bo, Yao Quanming, Pan Yuangang, et al. Millionaire: A hint-guided approach for crowdsourcing [J]. Machine Learning, 2019, 108(5): 1-28
- [17] Shah N B, Zhou Dengyong. Double or nothing: Multiplicative incentive mechanisms for crowdsourcing [J]. Journal of Machine Learning Research, 2016, 17(1): 5725-5776
- [18] Zhong Jinhong, Tang Ke, Zhou Zhihua. Active learning from crowds with unsure option [C] //Proc of the 24th Int Conf on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2015: 1061-1067
- [19] Ding Yaoxiang, Zhou Zhihua. Crowdsourcing with unsure option [J]. Machine Learning, 2018, 107(4): 749-766
- [20] Jurca R, Faltings B. Mechanisms for making crowds truthfully [J]. Journal of Artificial Intelligence Research, 2009, 34(1): 209-253
- [21] Hu Zehong, Liang Yitao, Zhang Jie, et al. Inference aided reinforcement learning for incentive mechanism design in crowdsourcing [C] //Proc of the 32nd Int Conf on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc, 2018: 5512-5522
- [22] Witkowski J, Parkes D C. Peer prediction without a common prior [C] //Proc of the 13th ACM Conf on Electronic Commerce. New York: ACM, 2012: 964-981
- [23] Jain S, Gujar S, Bhat S, et al. An incentive compatible multi-armed-bandit crowdsourcing mechanism with quality assurance [J]. arXiv preprint, arXiv:1406.7157, 2014
- [24] Gamage D, Ballav A, Goyal S, et al. Boomerang: Rebounding the consequences of reputation feedback on crowdsourcing platforms [C] //Proc of the 29th Annual Symp on User Interface Software and Technology. New York: ACM, 2016: 625-637
- [25] Qiu Chenxi, Squicciarini A, Rajtmajer S. Rating mechanisms for sustainability of crowdsourcing platforms [C] //Proc of the 28th ACM Int Conf on Information and Knowledge Management. New York: ACM, 2019: 2003-2012
- [26] Singer Y, Mittal M. Pricing mechanisms for crowdsourcing markets [C] //Proc of the 22nd Int World Wide Web Conf. New York: ACM, 2013: 1157-1166
- [27] Goel G, Nikzad A, Singla A. Mechanism design for crowdsourcing markets with heterogeneous tasks [C] //Proc of the 2nd AAAI Conf on Human Computation and Crowdsourcing. Palo Alto, CA: AAAI Press, 2012: 77-86

- [28] Gravin N, Jin Yaonan, Lu Pinyan, et al. Optimal budget-feasible mechanisms for additive valuations [J]. ACM Transactions on Economics and Computation, 2020, 8(4): 1-15
- [29] Liu Xiangyu, Yu Chuan, Zhang Zhilin, et al. Neural Auction: End-to-end learning of auction mechanisms for e-commerce advertising [C] //Proc of the 27th ACM SIGKDD Conf on Knowledge Discovery and Data Mining. New York: ACM, 2021: 3354-3364
- [30] Vamvoudakis K G, Lewis F L. Multi-player non-zero-sum games: Online adaptive learning solution of coupled Hamilton-Jacobi equations [J]. Automatica, 2011, 47(8): 1556-1569
- [31] Hart S, Mas-Colell A. A simple adaptive procedure leading to correlated equilibrium [J]. Econometrica, 2000, 68(5): 1127-1150
- [32] Ross S, Bagnell D. Efficient reductions for imitation learning [C] //Proc of the 13th Int Conf on Artificial Intelligence and Statistics. Brookline, MA: Microtome Publishing, 2010: 661-668
- [33] Mozafari B, Sarkar P, Franklin M J, et al. Active learning for crowd-sourced databases [J]. arXiv preprint, arXiv: 1209.3686, 2014
- [34] Snow R, O' Connor B, Jurafsky D, et al. Cheap and fast—but is it good? Evaluating non-expert annotations for natural languages tasks [C] //Proc of the 13th Conf on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2008: 254-263
- [35] Welinder P, Branson S, Belongie S, et al. The multidimensional wisdom of crowds [C] //Proc of the 23rd Int Conf on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc, 2010: 2424-2432
- [36] Sheshadri A, Lease M. SQUARE: A benchmark for research on computing crowd consensus [C] //Proc of the 1st AAAI Conf on Human Computation and Crowdsourcing. Palo Alto, CA: AAAI Press, 2013: 156-164
- [37] Tang Wei, Lease M. Semi-supervised consensus labeling for crowdsourcing [C] //Proc of the 34th SIGIR Workshop on Crowdsourcing for Information Retrieval. New York: ACM, 2011: 36-41
- [38] Dekel O, Shamir O. Good learners for evil teachers [C] //Proc of the 26th Annual Int Conf on Machine Learning. New York: ACM, 2009: 233-240



He Yucheng, born in 1996. PhD candidate. Student member of CCF. His main research interests include human-machine synergistic learning, machine learning theory and algorithms.

何雨橙, 1996 年生. 博士研究生. CCF 学生会员. 主要研究方向为人机协同机器学习、机器学习理论与算法.



Ding Yaoxiang, born in 1988. PhD. His main research interests include machine learning theory and algorithms.

丁尧相, 1988 年生. 博士. 主要研究方向为机器学习理论与算法.



Zhou Zhihua, born in 1973. PhD, professor and PhD supervisor. Member of Academy of Europe. Fellow of ACM, AAAI, IEEE, CCF, CAAI, etc. His main research interests include artificial intelligence, machine learning, and data mining.

周志华, 1973 年生. 博士, 教授, 博士生导师. 欧洲科学院院士. ACM, AAAI, IEEE, CCF, CAAI 等的会士. 主要研究方向为人工智能、机器学习和数据挖掘.