

# 联邦众包的激励机制

蒋晓倩

2023 年 5 月 22 日

## 摘要

激励推动式联邦众包 iFedCrowd 是在联邦学习的框架下实行众包项目，并引入激励机制，提高众包的质量和效率。本文重新设计了 iFedCrowd 的激励机制，在原有的基于两阶段 Stackelberg 博弈的激励机制的基础上，设置了时间阈值，更改了客户端的选择训练策略的选项，将对训练完成时间和数据新鲜度的评估交由任务发布者决定，提高了模型的泛化能力。本文的方法不仅简化了客户端模型，还提高了服务器端的计算效率，并且证明了客户端和服务端的结果为全局最优；最后，用实验结果证实，本文重新设计的激励机制不仅可以推动联邦众包项目高质量高效率地完成，还优于原有的激励机制。

## 1 Introduction

众包为解决机器学习所需要的大规模标注数据受单一的专业知识或者特定的地域和组织局限的问题提供了可行的途径 [3]。众包将任务分解或者经过设计交由大众群体来完成，并通过对完成的产出进行进一步筛选、分析等获得最终的结果 [1]。这些参与众包任务的工作者往往来自不同的地域、不同的行业，具有不同的背景。

然而，众包面临着一个开放性的挑战，一方面，如果没有一定的动机，工作者不会情愿贡献自己的热情参与到众包任务；另一方面，在传统的众包项目中，往往需要参与的工作者将标注的数据上传到云端，那么一些涉及到众包工作者敏感信息的标注数据一旦上传到云端，就意味着参与众包项目的工作者将面临着隐私泄露的风险（比如，位置感知众包技术 MCS，参与的工作者必须向服务提供商披露有关其当前真实位置的信息 [5]。），人们日益加强的隐私保护意识使得工作者更不愿意参与到众包项目中，即便参与

进来,遇到隐私敏感问题可能不会做出准确标注,从而降低众包的质量和效率。

解决以上问题有两种途径,一是解决隐私泄露问题,众包项目中最常用的隐私保护技术往往是以牺牲众包的质量为代价来实现的,例如,当参与者利用空间隐身来混淆他们的位置时,Pournajaf 等人使用隐蔽位置进行全局优化 [9],但是降低了众包的准确度,从而降低了众包的质量;Tang 等人提出了一种通过开发基于属性的高级加密,结合准备/在线加密和外包解密技术的方法,来实现工作者的隐私保护 [11],但是这无疑为服务平台增加了大量的解密开销,并且计算复杂;Wu 等人利用同态密码系统对用户的加密数据进行加权聚合,保证高精度同时保护隐私 [13],该方法不仅需要处理大规模数据,并且无法抵御推断攻击。二是足够的动机,使得工作者愿意承担风险,众包工作者参与众包活动的动机有经济、兴趣、荣誉等多方面驱动因素构成。而激励机制则是针对经济 [10][8]、兴趣 [6]、荣誉 [12] 等这些驱动因素制定相应的激励策略来吸引工作者并激发工作者的热情,但是一味的加强激励策略会增加任务发布者的成本。为了同时兼顾众包工作者的隐私保护问题和众包项目的效率和质量,Kang 等人提出了激励推动式的联邦众包 iFedCrowd(Incentive-boosted Federated Crowdsourcing)[4],将联邦学习框架应用于众包项目,并且设计了一个基于 Stackelberg 博弈的激励机制,在保护工作者隐私的同时,也激发了工作者的热情,有效提高了众包的效率。联邦学习 [2] 是一种新型的分布式学习框架,允许多个客户端通过迭代聚合模型更新来协作训练共享模型,而不暴露原始数据,同时,联邦学习还引入了密码学和可信硬件等技术,通过密码技术的理论安全性以及可信硬件的物理层面安全来保障用户隐私。Stackelberg 博弈 [7],是一个两阶段的完全信息动态博弈,主要思想是双方都是根据对方可能的策略来选择自己的策略以保证自己在对方策略下的利益最大化,从而达到纳什均衡。iFedCrowd 将众包平台和参与的工作者之间基于激励的互动建模为 Stackelberg 游戏,其中每一方都最大化自己的利润。推导出博弈的纳什均衡,以求得双方的最优解。然而,本文认为该激励机制存在以下缺陷:(1)该激励机制没有对时间做出合理的控制,第一,无法排除潜在的恶意工作者可能提交低质量的数据以获得快速奖励,第二,没有对因为工作者自身或者客户端设备问题以及网络质量等突发状况导致服务器端迟迟接收不到模型的情况做出相应的响应。(2)该激励机制给工作者分配奖励考虑了客户端收集到的数据新鲜度和完成模型训练更新所花费的时间这两个因素,目的在于在尽可能短的时间

内完成联邦众包任务，这个机制可以有效地提高即时性的众包项目的效率，然而尽可能新鲜的数据和尽可能短的完成时间对于提高非即时性众包项目效率的作用不大，反而使得模型更加复杂。（3）该模型在计算服务器端效益时，一方面需要对隐函数进行求解，增加了服务器模型的复杂度和计算成本，另一方面对服务器模型的求解不能确定是否为全局最优。为了解决以上问题，本文重新设计了激励机制，本文的激励机制依然是基于 stackelberg 博弈，将对完成时间和数据新鲜度的把控交给任务发布者裁定，对参与众包任务的工作者发放奖励只考虑客户端模型精度这一个因素。

本文设计的激励机制除了修复 iFedCrowd 的缺陷外，还有以下贡献：（1）本文的激励模型按贡献分配奖励，只按客户端训练模型达到的精度分配奖励，维护了众包市场的公平性。（2）本文的模型对众包工作者的质量进行了初步筛选，排除了低质量工作者以牺牲精度换取更短训练时间从而换取丰厚的奖励的可能，增强了 iFedCrowd 算法的健壮性。（3）本文的激励机制规则简单，具有良好的可解释性，有利于长期留住高质量的众包工作者，体现了激励推动联邦众包的责任性。

## 参考文献

- [1] Daren C Brabham. Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence*, 14(1):75–90, 2008.
- [2] Dashan Gao, Xin Yao, and Qiang Yang. A survey on heterogeneous federated learning. *arXiv preprint arXiv:2210.04505*, 2022.
- [3] Jeff Howe et al. The rise of crowdsourcing. *Wired magazine*, 14(6):1–4, 2006.
- [4] Xiangping Kang, Guoxian Yu, Jun Wang, Wei Guo, Carlotta Domeniconi, and Jinglin Zhang. Incentive-boosted federated crowdsourcing. *arXiv preprint arXiv:2211.14439*, 2022.
- [5] Jong Wook Kim, Kennedy Edemacu, and Beakcheol Jang. Privacy-preserving mechanisms for location privacy in mobile crowdsensing: A survey. *Journal of Network and Computer Applications*, page 103315, 2022.

- [6] Deng Li, Luan Yang, Jiaqi Liu, and Hui Liu. Considering decoy effect and fairness preference: An incentive mechanism for crowdsensing. *IEEE Internet of Things Journal*, 6(5):8835–8852, 2019.
- [7] Tao Li and Suresh P Sethi. A review of dynamic stackelberg game models. *Discrete & Continuous Dynamical Systems-B*, 22(1):125, 2017.
- [8] Wei Li, Cheng Zhang, and Yoshiaki Tanaka. Privacy-aware sensing-quality-based budget feasible incentive mechanism for crowdsourcing fingerprint collection. *IEEE Access*, 8:49775–49784, 2020.
- [9] Layla Pournajaf, Li Xiong, Vaidy Sunderam, and Slawomir Goryczka. Spatial task assignment for crowd sensing with cloaked locations. In *2014 IEEE 15th International Conference on Mobile Data Management*, volume 1, pages 73–82. IEEE, 2014.
- [10] Yaron Singer and Manas Mittal. Pricing mechanisms for crowdsourcing markets. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1157–1166, 2013.
- [11] Wenjuan Tang, Kuan Zhang, Ju Ren, Yaoxue Zhang, and Xuemin Sherman Shen. Privacy-preserving task recommendation with win-win incentives for mobile crowdsourcing. *Information Sciences*, 527:477–492, 2020.
- [12] Runhua Wang, Feng Zeng, Lan Yao, and Jinsong Wu. Game-theoretic algorithm designs and analysis for interactions among contributors in mobile crowdsourcing with word of mouth. *IEEE Internet of Things Journal*, 7(9):8271–8286, 2020.
- [13] Haiqin Wu, Liangmin Wang, and Guoliang Xue. Privacy-aware task allocation and data aggregation in fog-assisted spatial crowdsourcing. *IEEE Transactions on Network Science and Engineering*, 7(1):589–602, 2019.