中国科学: 信息科学 2022 年 第 52 卷 第 6 期: 971–991
SCIENTIA SINICA Informationis
《中国科学》杂志社
SCIENCE CHINA PRESS
理论计算机科学专题 · 论文

基于区块链和贝叶斯博弈的联邦学习激励机制

张沁楠[1], 朱建明[1], 高胜[1*], 熊泽辉[2], 丁庆洋[3], 朴桂荣[1]

1. 中央财经大学信息学院, 北京 100081, 中国
2. Pillar of Information Systems Technology and Design, Singapore University of Technology and Design, Singapore 487372, Singapore
3. 北京联合大学管理学院, 北京 100020, 中国
* 通信作者. E-mail: sgao@cufe.edu.cn

收稿日期: 2022–01–13; 修回日期: 2022–03–23; 接受日期: 2022–04–18; 网络出版日期: 2022–06–13

国家重点研发计划 (批准号: 2017YFB1400700)、国家自然科学基金 (批准号: 62072487)、北京市自然科学基金 (批准号: M21036)、北京联合大学教育科学研究课题 (批准号: JK202114) 和北京联合大学科研专项 (批准号: ZK30202101) 资助项目

摘要 联邦学习通过聚合多方本地模型成为数据共享的新模式. 现有的联邦学习激励机制有效缓解了完全信息下的数据供给不足问题, 但仍面临搭便车、不公平、不可信等挑战. 为此, 本文提出了一种基于区块链和贝叶斯博弈 (Bayesian game) 的不完全信息联邦学习激励机制, 通过量化数据供给方的成本效用与数据需求方的支付报酬对数据交易过程建模, 采用沙普利值 (Shapley value) 实现了数据供给方报酬分配的公平性. 在交易模型中考虑到参与个体的异质性与隐私保护, 将数据供给方的资源配置策略构建为不完全信息的贝叶斯博弈模型, 通过优化本地模型训练策略实现对数据供给方的激励作用. 本文进一步分析了激励机制的有效性与行动策略的可信性, 提出一种隐私保护的贝叶斯博弈行动策略共识算法 (privacy-preserving Bayesian game action strategy consensus algorithm, PPBG-AC), 该算法使数据供给方在基于区块链的数据交易平台下实现了贝叶斯纳什均衡. 方案对比与理论分析表明本文提出的不完全信息联邦学习激励机制保障了数据供给方利益分配的公平性与资源配置的可信性, 基于实际公开数据集的仿真实验与性能评估验证了激励机制的有效性.

关键词 联邦学习, 激励机制, 区块链, 贝叶斯博弈, 沙普利值, 不完全信息, 隐私保护

1 引言

在数字经济背景下, 数据已成为重要的商业资源和生产要素. 随着 6G 网络通信技术与智能边缘计算的发展, 海量数据通常以小规模、分散化、碎片化形式分布在不同行业与移动设备中. 目前虽然数据交易市场活跃, 但受制于数据的法律属性、产权规则、价值分配、供需信任等难题, 有效的数据要素流通市场尚未成规模, 其经济价值未能得到充分体现 [1]. 传统的以原始数据集为交易标的物的数据

引用格式: 张沁楠, 朱建明, 高胜, 等. 基于区块链和贝叶斯博弈的联邦学习激励机制. 中国科学: 信息科学, 2022, 52: 971–991, doi: 10.1360/SSI-2022-0020
Zhang Q N, Zhu J M, Gao S, et al. Incentive mechanism for federated learning based on blockchain and Bayesian game (in Chinese). Sci Sin Inform, 2022, 52: 971–991, doi: 10.1360/SSI-2022-0020

交易模式存在供给不足与多交易市场的套利问题 [2,3], 并且基于中心化的数据交易平台因信息不对称导致了交易过程的不公平与不可信. 联邦学习 (federated learning, FL) [4] 汇聚多方本地模型, 通过整合碎片化数据成为数据交易发展的新模式. 目前联邦学习已成为隐私计算领域的重要技术, 在新兴数据交易市场中具有广泛的应用前景.

然而, 参与联邦学习的数据供给方存在异质性, 即使某一方不执行任何本地训练, 也不能避免其使用各方协作生产的数据模型 [5]. 此外, 缺乏合理的激励机制与数据合规、隐私泄露等问题进一步导致了数据供给不足的现状. 在联邦学习数据模型协作训练过程中, 数据供给方直接进行中间参数交换可能导致模型提取攻击 [6]、成员推理攻击 [7] 与本地模型投毒攻击 [8] 等. 为了避免恶意攻击引发的隐私泄露与梯度爆炸, 可将训练好的梯度参数采用差分隐私加噪 [9,10] 并压缩 [11,12] 后上传至参数服务器进行模型聚合 [13]. 虽然差分隐私加噪保障了中间参数的隐私性, 但是联邦学习中多数据供给方间协作的公平性与可信性仍难以保证.

区块链 [14] 通过点对点网络构建了多方参与的分布式账本, 通过智能合约 [15] 可以实现可信、可配置、可编程的数据交易, 为基于联邦学习的数据交易市场提供了内生信任的交易环境 [16]. 已有研究基于区块链设计了联邦学习参与节点激励机制 [17], 通过自定义 Token 激励节点进行数据共享, 有效缓解了因供需双方信任鸿沟导致的供给不足问题. 但现有方案没有考虑到异质性数据供给方的不完全信息场景与数据模型的公共物品属性 [18], 忽视了数据交易过程的社会福利最大化问题与协作的公平可信性, 因此难以维系数据交易市场的稳定运行.

此外, 现有的激励机制大多只考虑了完全信息场景 [5,19~21], 即假设数据交易市场已知每个数据供给方真实的原始数据特征, 并且数据供给方彼此知道对方的模型训练资源配置策略. 然而, 在实际的数据交易市场中, 难免存在一些不愿意披露真实本地模型训练信息的数据供给方, 导致了激励机制设计的不公平与不可信风险. 因此, 如何在不完全信息场景下实现数据供给方公平的奖励分配机制与最优的资源配置策略是联邦学习激励机制设计的关键.

为此, 本文设计了不完全信息下的联邦学习激励机制, 构建了数据供给方本地模型训练中资源配置行动策略的贝叶斯博弈模型, 通过沙普利值 (Shapley value) 与联盟区块链实现了报酬分配的公平性与可信性. 考虑到贝叶斯博弈行动策略的一致性, 提出了一种隐私保护的贝叶斯博弈行动策略共识算法, 进一步提升了激励机制的有效性. 理论分析与实验结果表明, 数据供给方可以通过激励机制提升个人效用与社会福利, 并且随着数据质量的提升, 效用与福利的提升效果更加明显. 本文主要贡献总结如下.

- 构建了一个基于区块链的可信联邦学习数据交易模型, 通过量化数据供给方的成本效用确定数据需求方的支付价格, 实现联邦学习数据模型交易的匹配与撮合. 奖励分配阶段采用沙普利值实现了数据供给方报酬分配的公平性.
- 在不完全信息场景下提出一种基于贝叶斯博弈的联邦学习激励机制. 将数据供给方本地模型训练的资源配置策略构建为贝叶斯博弈模型, 通过求解预算约束下的社会福利最大化问题得到贝叶斯纳什均衡 (Bayesian Nash equilibrium, BNE), 实现数据供给方的资源优化配置.
- 为保障激励机制的有效性与行动策略的可信性, 提出一种基于区块链的隐私保护贝叶斯博弈行动策略共识算法 (privacy-preserving Bayesian game action strategy consensus algorithm, PPBG-AC), 通过贝叶斯博弈行动策略的投票共识保障了激励机制中资源配置策略的一致性与可信性. 理论分析与仿真实验验证了算法的收敛性与激励机制的有效性.

本文在第 2 节介绍了数据交易与联邦学习激励机制的相关研究工作. 第 3 节针对联邦学习数据交易过程建模, 量化了数据供给方的成本效用与数据需求方的支付报酬. 第 4 节介绍了基于贝叶斯博

弈的不完全信息联邦学习激励机制的设计细节. 第 5 节进行了方案对比与仿真实验, 评估了算法的收敛性与激励机制的有效性. 第 6 节对本文激励机制的设计工作进行了总结和展望.

2 相关工作

联邦学习自 2016 年提出之后成为分布式机器学习的新范式 [4], 已广泛应用于物联网、电子健康、智慧金融等领域 [22]. 联邦学习驱动的数据交易实现了数据供给方原始数据的本地化, 成为组织多方协作生产数据模型的有效方法. 然而, 现有研究大多乐观地假设数据供给方愿意贡献计算与存储资源参与联邦学习数据模型的加工协作 [5,17], 但是如果没有激励机制进行合理的经济补偿, 自利的数据供给方不愿意参与联邦学习 [23,24], 甚至故意作恶发起投毒攻击 [8] 扰乱协作秩序. 此外, 现有的完全信息场景下的联邦学习激励机制假设交易市场已知数据供给方的真实模型训练信息, 没有考虑到数据供给方的异质性与隐私保护需求, 不适用于实际的大规模数据交易市场. 同时, 数据供给方的设备资源受限与数据非独立同分布 (non-IID) [25~27] 特性导致的搭便车与不公平问题也进一步阻碍了数据供给方的参与. 为此, 本文针对联邦学习数据交易中的供给不足问题展开研究, 主要调研了数据交易与联邦学习激励机制的相关研究, 发现在满足数据供给方隐私保护需求的前提下, 实现报酬公平分配与资源优化配置有利于促进联邦学习数据模型的共享与协作.

在数据交易的相关工作中, 已有研究关注了数据交易定价 [2,28]、个人信息隐私保护 [29~31] 以及利益公平分配 [32,33] 等问题. 其中, Chen 等 [2] 提出了一种避免多交易市场套利的数据交易方法, 通过数据实例定价实现了数据供需方在多交易市场下的无套利交易. Liu 等 [28] 设计了一个基于区块链的数据交易市场, 通过两阶段斯塔克伯格博弈 (Stackelberg game) 设计了数据需求方与市场代理的定价与购买机制, 并验证了博弈均衡的存在性. Jiao 等 [34] 构建了一个基于联邦学习的服务交易市场, 采用近似策略证明保证了交易的可信性、个人理性以及计算效率, 通过反向多维拍卖机制和深度强化学习实现了社会福利最大化. 虽然已有方案构建了无线边缘计算场景中的联邦学习服务交易市场, 并且考虑了参与节点的资源配置以及交易套利问题, 但仍未基于数据的公共物品属性深入考虑数据供给方的异质性与多参与方协作的公平可信性.

已有联邦学习激励机制根据评估指标不同可以分为贡献度驱动、信誉值驱动与资源分配驱动三大类 [35], 主要采用的理论方法包括斯塔克伯格博弈 [19]、契约理论 [20] 以及拍卖机制 [21] 等. 在贡献度驱动方面, Feng 等 [36] 考虑到联邦学习过度依赖中心化参数服务器, 采用协作中继通信网络构建了服务器与移动设备之间的非同质博弈模型, 并将数据量与模型质量作为贡献度评估依据. 其中, 移动设备决定每单位数据的价格, 而服务器通过选择训练数据的大小实现社会福利最大化, 利用外点法 [37] 得到了博弈的纳什均衡. 在信誉值驱动方面, Kang 等 [20] 融合信誉评估机制与契约理论对参与联邦学习的节点进行约束, 引入信誉指标衡量数据供给方的可靠性, 采用多权重主观逻辑模型与区块链进行参与者的选择与信誉值的分布式管理, 满足了个人理性与激励相容, 仿真结果表明该方案相较于斯塔克伯格博弈模型能获得更高的任务发布者福利. Zeng 等 [21] 提出了一种基于多维采购拍卖理论的移动边缘设备激励机制, 并利用期望效用理论对数据供给方行为进行指导. 在资源分配驱动方面, 已有大量研究探讨了联邦学习中的资源和任务最优配置问题 [38~40]. Wang 等 [41] 采用沙普利值计算参与方贡献度, 构建了两阶段斯塔克伯格博弈模型, 分析了模型拥有者与参与方在完全信息与不完全信息场景下的资源分配机制. 虽然该工作分析了联邦学习中的不完全信息场景, 但并没有考虑到数据供给方因竞争关系存在的不完全信息情况. 总体而言, 已有工作提出了联邦学习驱动的数据服务交易市场, 关注到了数据供给方的异质性与隐私保护需求, 但是如何在不完全信息下设计公平有效的激励机制仍

需要进一步深入研究.

3 系统模型

在联邦学习数据交易场景中数据供给方既是全局模型的生产者也是消费者, 供需双方基于区块链数据交易平台进行交易撮合与利益分配. 本节构建了基于区块链的可信联邦学习数据交易模型, 引入交叉熵与新鲜度评估模型质量, 量化了数据供给方的成本效用与数据需求方的支付报酬, 采用沙普利值进行利益公平分配, 提出了不完全信息联邦学习激励机制的设计目标.

3.1 基于区块链的可信联邦学习数据交易模型

本模型考虑一个拥有 N 个数据供给方和 M 个数据需求方的联邦学习数据交易市场, 包括数据供给方集合 $P=\{P_1,P_2,\ldots,P_N\}$ 与数据需求方集合 $R=\{R_1,R_2,\ldots,R_M\}$. 每个数据供给方 P_i 拥有本地数据集 $D_i=\{d_1,d_2,\ldots,d_n\}$, 数据供给方集合 P 的总体数据集为 $\mathcal{D}=\cup_{i=1}^{N}D_i$. 考虑隐私保护与数据合规, 数据供给方不愿直接提供原始数据集 D_i, 而是将 t 轮迭代后的本地模型参数 $w_i^t=\{w_1,w_2,\ldots,w_m\}$ 提供给区块链上节点, 节点收到多方参数集合 $w^t=\cup_{i=1}^{N}w_i^t$ 之后通过模型聚合算法 [13] 计算全局模型 w_G, 经多轮迭代之后产生可交易的全局模型 w_G^* 满足在数据集 $\mathcal{D}$ 上的期望损失 $L(w_G)$ 最小化, 即

$$w_G^*=\arg\min_{w_G^r}\left\{\frac{1}{N}\sum_{i=1}^{N}L(f(w_G^r),D_i)\right\},\quad r\leqslant\frac{\sum_{j=1}^{M}B_j}{C_{\text{unit}}},\tag{1}$$

其中 $L(f(w_G^r),D_i)$ 是在给定参数 w_G^r 下模型 $f(\cdot)$ 在数据集 D_i 的损失值 [42], r 是满足成本预算约束的聚合次数, B_j 是对模型 w_G^* 有购买需求的数据需求方支付预算, C_{unit} 是模型聚合的单位计算成本. 数据异质性以及供给方在计算、存储、通信、资源等方面的不均衡特性, 导致了数据协作过程中的搭便车. 因此, 需要考虑数据供给方的成本与效用, 通过资源优化配置与公平奖励分配对数据供给方进行激励, 从而维系数据交易市场的稳定运行. 表 1 给出了本模型主要的符号及描述.

不同于传统中心化的数据交易平台, 本文构建了一个基于区块链的可信联邦学习数据交易模型, 如图 1 所示. 交易过程中数据需求方 R_j 向区块链节点发布数据需求, 链上节点接收到数据需求之后初始化模型参数 w_G^0 后广播给数据供给方集合 P, 确认参与的数据供给方 P_i 返回每次本地模型迭代的单位计算成本消息 $C_{\text{com}}(P_i)$ 后开始执行本地模型训练. P_i 经过 t 轮迭代之后将模型参数 w_i^t 上传至区块链数据交易平台通过智能合约进行参数验证与模型聚合 [13]. 在成本预算约束下链上节点不断更新模型直至收敛, 并将可交易的全局模型 $f(w_G^*)$ 访问接口 [43] 开放给数据需求方 R_j, R_j 支付报酬 B_j 后可获得全局模型 $f(w_G^*)$ 的访问权限. 区块链记录本地模型参数 w_i^t 及每轮聚合模型参数 w_G^r, 用于追溯恶意数据供给方并作为报酬分配依据. 系统模型的主要交易参与方介绍如下.

(1) 数据需求方 R_j. 负责发布数据需求 E_j 和支付预算 B_j, 基于区块链的数据交易市场收集需求并组织数据供给方生产数据模型 w_G^*, 链上节点验证模型质量后 R_j 获得模型访问权限并支付预算 B_j. 为了防止交易纠纷, 假设 R_j 已提前向平台缴纳押金以保障正常支付.

(2) 数据供给方 P_i. 根据初始全局模型 w_G^0 执行本地模型训练, t 轮本地迭代后得到本地模型参数 w_i^t, 本地迭代通常采用小批量随机梯度下降算法 [42] 更新模型参数, 即

$$\bigtriangledown g(w_i^{t-1},b_i)=\frac{1}{|b_i|}\sum_{j=1}^{|b_i|}\frac{\partial L(w_i^{t-1},d_j)}{\partial w_i^{t-1}},\tag{2}$$

表 1 符号及描述

Table 1 Notations and descriptions

Notation	Description
$P=\{P_1,P_2,\ldots,P_N\}$	The set of data providers
w_G^*	The tradable global model parameter
$\bigtriangledown\tilde{g}(w_i^t,b_i)$	The perturbed gradient of the local model w_i^t
ρ_i^r	The reward of P_i in r-round model aggregation
$Q(w_i^t)$	The quality evaluation function of model w_i^t
$U_i(r)$	The utility function of P_i in r-round model aggregation
θ_i	The device memory of P_i
β_i^t	The average memory consumption ratio of P_i in t iteration
ζ	The capacitance parameter
$C_{\rm com}(P_i)$	The computational cost of P_i in per iteration
$C_{\rm pr}(\bigtriangledown\tilde{g}(\cdot))$	The privacy cost of $\bigtriangledown\tilde{g}(\cdot)$
$C_{\rm unit}^r(P_i)$	The unit cost function of P_i in r-round model aggregation
$r(w_G^*)$	The aggregation rounds of w_G^*

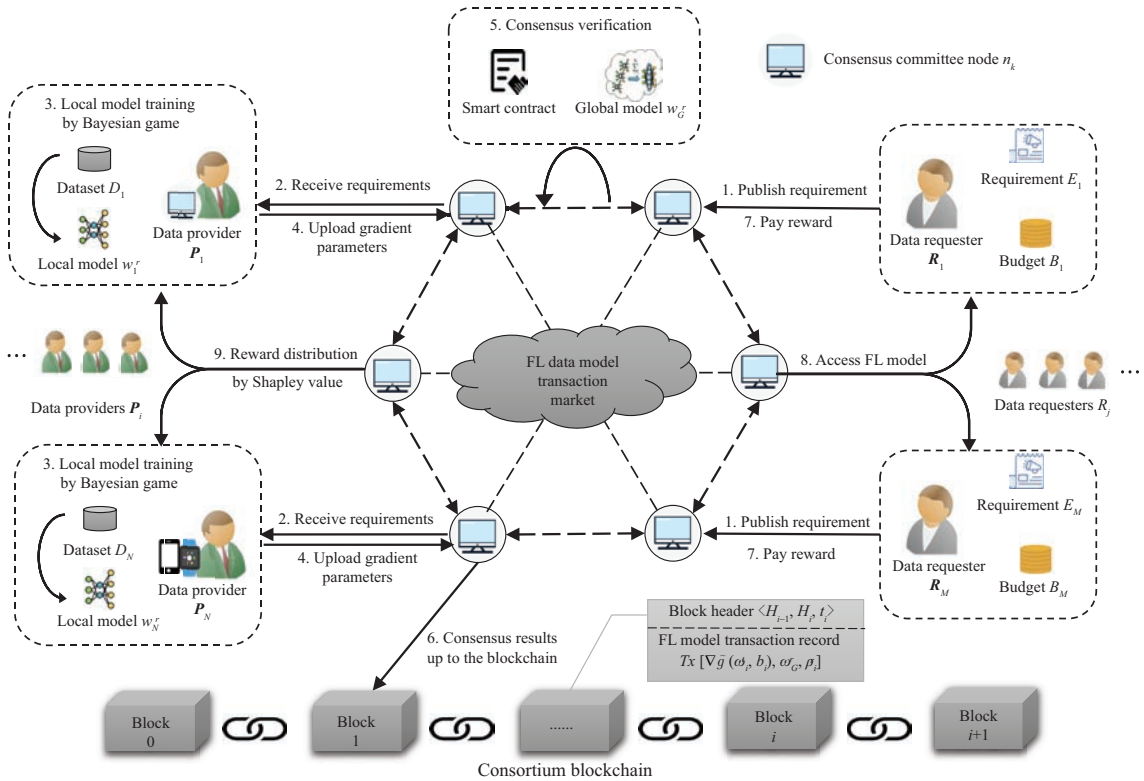

图 1 (网络版彩图) 基于区块链的可信联邦学习数据交易模型

Figure 1 (Color online) Trustworthy federated learning data transaction model based on blockchain

其中, b_i 是小批量数据样本, P_i 根据梯度值更新本地模型参数 $w_i^t = w_i^{t-1} + \alpha \bigtriangledown g(w_i^{t-1}, b_i)$. 为了避免模型提取攻击 [6] 与成员推理攻击 [7], 在 t 轮迭代之后, P_i 将梯度参数经差分隐私扰动 [9] 为 $\bigtriangledown \tilde{g}(w_i^t, b_i)$ 后上传至区块链节点执行模型聚合. 在第 r 轮模型聚合开始之后, 每个数据供给方访问区块链上数据获得上一轮的全局模型参数 w_G^{r-1}, 并基于 w_G^{r-1} 继续更新迭代本地模型.

(3) 区块链数据交易市场. 负责广播数据需求并进行交易记账. 当链上节点接收到数据需求方的请求后, 随机初始化一个全局模型 w_G^0 发布给数据供给方. 链上节点收集每轮数据供给方提供的本地模型参数 w_i^t, 执行模型聚合智能合约更新全局模型参数直至模型 w_G^* 收敛.

(4) 链上数据. 负责记载数据供给方提供的扰动后的本地模型梯度参数 $\bigtriangledown \tilde{g}(w_i^t, b_i)$、聚合后的全局模型参数 w_G^r 以及支付给 P_i 的报酬奖励 ρ_i^r, 将交易记录 $Tx = \{\bigtriangledown \tilde{g}(w_i^t, b_i), w_G^r, \rho_i^r\}$ 打包成区块 $\text{Block}_i = \{\langle H_{i-1}, H_i, t_i\rangle, \langle Tx = \{\bigtriangledown \tilde{g}(w_i^t, b_i), w_G^r, \rho_i^r\}\rangle\}$ 并通过共识机制完成记账, 链上数据保证模型参数的可信可追溯.

3.2 数据供给方的成本效用函数

数据供给方在本地模型协作加工过程中需要消耗计算存储资源, 在此过程中每个数据供给方隐私保护需求可能不同. 因此, 本小节综合考虑以上因素, 给出了数据供给方 P_i 的成本效用函数定义.

单位计算成本 $C_{\text{com}}(P_i)$. 数据供给方 P_i 在本地训练过程中需要占用内存及 CPU 计算资源. 假设数据供给方 P_i 用于本地数据模型训练时的 CPU 时钟频率为 f_i, 使用 $F = (f_1, f_2, \ldots, f_N)$ 定义生产数据模型 w_G^* 的数据供给方集合 P 的处理性能. 数据供给方 P_i 用于本地模型训练的设备基准内存为 θ_i, 第 t 轮本地迭代内存平均消耗比例 $\beta_i^t \in [0, 1]$, 内存实际消耗为 $\beta_i^t \theta_i$. 假设本地模型训练过程中每次迭代输入数据样本量大小相同, 基于处理器二次能耗模型 [24], 数据供给方 P_i 本地模型迭代的单位计算成本为

$$C_{\text{com}}(P_i) = \alpha \zeta c_i s_i f_i^2 + (1 - \alpha) \beta_i^t \theta_i, \tag{3}$$

其中, α 是计算成本调节因子, ζ 是 P_i 计算芯片组的有效电容参数 [44], c_i 是处理一个批大小数据所需的 CPU 周期, s_i 是每轮本地迭代所需的数据批样本大小.

隐私成本 C_{pr}. 为了最大限度减少本地模型上传时额外的隐私泄露, 本模型考虑了不同梯度参数的隐私成本, 主要通过差分隐私预算 [45] 来衡量. 假设数据供给方 P_i 将差分隐私扰动后的梯度参数 $\bigtriangledown \widetilde{g}(w_i^t, b_i)$ 上传至区块链节点, $\bigtriangledown \widetilde{g}(w_i^t, b_i)$ 通过拉普拉斯噪声扰动得到

$$\bigtriangledown \widetilde{g}(w_i^t, b_i) = \bigtriangledown g(w_i^t, b_i) + \left\langle \text{Lap}\left(\mu, b = \frac{\Delta f}{\epsilon}\right)\right\rangle, \tag{4}$$

其中, μ, b 分别是 Laplace 分布的位置参数和尺度参数, Δf 是噪声局部敏感度, ϵ 是差分隐私预算.

考虑不同数据供给方提供的梯度参数对全局模型贡献不同 [46], 可通过 1- 范数计算不同梯度对全局模型的贡献程度. 由此, 加噪梯度参数 $\bigtriangledown \widetilde{g}(w_i^t, b_i)$ 的隐私成本 $C_{\text{pr}}(\bigtriangledown \widetilde{g}(w_i^t, b_i))$ 可通过下式计算:

$$C_{\text{pr}}(\bigtriangledown \widetilde{g}(w_i^t, b_i)) = \epsilon_i \ln(1 + \| \bigtriangledown \widetilde{g}(w_i^t, b_i) \|), \tag{5}$$

其中, 将隐私预算 i 作为隐私成本计算主要依据, 采用对数函数实现隐私成本与梯度贡献影响的正相关, 并满足隐私成本大于 0.

成本函数 $C_{\text{unit}}^r(P_i)$. 基于以上成本要素的量化, 数据供给方 $P_i \in P$ 参与第 r 轮联邦学习数据模型交易的成本函数 $C_{\text{unit}}^r(P_i)$ 可以定义为

$$C_{\text{unit}}^r(P_i) = (\ln C_{\text{com}}(P_i) + C_{\text{pr}}(\bigtriangledown \widetilde{g}(w_i^t, b_i))). \tag{6}$$

效用函数 $U_i(r)$. 本文定义数据供给方效用主要评估依据是模型质量 [47] 与新鲜度. 由于在分类预测问题中交叉熵损失函数优于均方误差 [48, 49], 因此对于标签数据 (x_i, y_i), 采用交叉熵损失函数 $H(f(x_i), y_i) = -\sum_{i=1}^{n} y_i \log f(x_i)$ 判断模型质量变化情况. 在强凸损失函数条件下, 损失越小则模型与数据集的拟合度更好且模型质量更高.

考虑到数据模型价值的特殊性, 随着时间的流逝数据模型的价值随之降低, 因此本文引入可信时间戳量化模型新鲜度. 为了保障时间戳的可验证性, 数据供给方完成本地模型 w_i^k 迭代之后, 需要向可信硬件 SGX 中的内容容器 enclave [50] 请求一个计时器 $T_i^k(w_i^k)$, 模型 w_i^k 的模型新鲜度 $F(w_i^k) = \lg T_i^k(w_i^k)$. 由此, 模型 w_i^t 的质量评估函数 $Q(w_i^t)$ 可以被建模为

$$Q(w_i^t) = \frac{\mu_0 F(w_i^k)}{\mu_1 + H(f(x_i), y_i)}, \tag{7}$$

其中, 模型效用参数 $\mu_0, \mu_1 \geqslant 0$, 可以根据损失函数、神经网络结构以及数据分布的不同进行设置 [27].

本文选取 u_0 为模型隐藏层个数, u_1 为模型输出层个数, 输出层主要取决于标签数据集. 采用 P_i 参与第 r 轮自适应聚合之后的模型质量 $Q(w_G^r(P_i))$ 与 P_i 未参与聚合的模型质量 $Q(w_G^r(P\backslash\{P_i\}))$ 的差值作为效用函数 [51]. 基于 $Q(w_i^t)$ 的强凸函数属性和初始化全局模型边界 $Q(w_i^0)$, 数据供给方 $P_i \in P$ 在参加第 r 轮模型聚合后提升的效用函数可建模为

$$U_i(r) = \rho_i^r [Q(w_G^r(P_i)) - Q(w_G^r(P\backslash\{P_i\}))], \tag{8}$$

其中, $Q(w_G^r(P_i))$ 是数据供给方 P_i 参与聚合的模型质量评估结果, $Q(w_G^r(P\backslash\{P_i\}))$ 是 P_i 未参与模型聚合的评估结果, ρ_i^r 是数据供给方 P_i 在第 r 轮模型聚合中的报酬奖励, 如果 $U_i(r) < 0$, 则代表 P_i 需要支付价格补偿以换取更高质量的模型参数, 避免数据供给方的搭便车行为.

3.3 数据需求方的支付报酬函数

在数据模型交易过程中, 数据需求方 R_j 基于有限的预算 $B_j \in \mathbb{R}_+$ 发布数据需求 E_j, 而数据供给方 P_i 通过联邦学习参与数据模型 w_G^* 的加工. 为了激励更多的数据供给方 P_i 参与联邦学习模型共享, 数据交易平台需要将数据需求方预算 B_j 公平分配给协作贡献者以激励数据供给方参与. 为了防止低质量数据模型的引入或者恶意数据供给方的投毒攻击 [8], 需要保证协作中数据供给方贡献度评估的可靠性与报酬分配的公平性.

本文引入区块链记录每轮聚合的本地模型参数 w_i^t, 采用沙普利值 [52] 进行供给方报酬的公平分配. 在报酬分配阶段, 假设第 r 轮数据模型 w_G^r 的数据供给方集合 S_r 是联盟博弈中的局中人集合. 其价值函数 $v(P_i)$ 由数据供给方 $P_i \in P$ 在参加第 r 轮模型聚合后提升的效用函数 $U_i(r)$ 决定, 即

$$v(P_i) = \begin{cases} U_i(r), & P_i \in S_r, \\ 0, & P_i \notin S_r, \end{cases} \tag{9}$$

其中, S_r 是在数据交易平台注册过的数据供给方集合 P 的任意子集, 即 $S_r \subseteq P$. 基于异步训练假设, 数据供给方 P_i 可以随时加入或者退出 S_r, 将每次参与聚合的报酬求和之后, 即可得到总计报酬奖励. 当价值函数为负时, 数据供给方需要支付价格补偿以换取更高质量的模型参数. 在数据供给方的联盟博弈 $(S_r, v(P_i))$ 中, 任意数据供给方 P_i 的沙普利值 $\xi_i(S_r, v(P_i))$ 可通过下式计算:

$$\xi_i(S_r, v(P_i)) = \sum_{S_r \subseteq P\backslash\{P_i\}} \frac{|S_r|!(N - |S_r| - 1)!}{N!} \times [v(S_r \cup \{P_i\}) - v(S_r)], \tag{10}$$

其中, N 为数据供给方集合 P 的元素个数, $P\backslash\{P_i\}$ 为不包含 P_i 的数据供给方集合. 如果数据供给方 P_i 的沙普利值 ξ_i 为零, 则 P_i 对数据模型 w_G^* 的贡献度为零, 即为哑元 (dummy player) [53].

本模型将数据需求方的预算支付定义在两个预算约束条件之下, 以满足预算均衡和与个人理性.

约束 1 (模型聚合预算约束). 满足 w_G^* 的模型聚合成本小于初始数据需求方预算手续费, 即

$$r(w_G^*) \leqslant \frac{\sum_{j=1}^{M} B_j \times \sigma}{\sum_{k=1}^{T} C_{\mathrm{com}}(n_k)}, \tag{11}$$

其中, $r(w_G^*)$ 是数据模型 w_G^* 的聚合轮数, B_j 是第 j 个期望购买 w_G^* 的数据需求方的购买预算, $C_{\mathrm{com}}(n_j)$ 是交易平台的共识委员会节点 n_j 执行共识验证的单位计算成本, T 是参与共识的节点个数, M 是支付购买预算的数据需求方个数, $\sigma \in [0,1]$ 是数据需求方支付预算中的手续费比例.

约束 2 (本地训练预算约束). 满足数据供给方 P_i 本地模型训练成本小于数据需求方预算中的训练成本, 即

$$k(w_i^t) \leqslant \frac{\sum_{j=1}^{M} B_j \times (1-\sigma)}{r(w_G^*)|P_\pi| C_{\mathrm{unit}}^r(P_i)}, \tag{12}$$

其中, $k(w_i^t)$ 是数据供给方 P_i 训练本地模型 w_i^t 时执行的本地迭代次数, $C_{\mathrm{unit}}^r(P_i)$ 是 P_i 参与第 r 轮模型聚合所消耗的单位成本, $|P_\pi|$ 是模型聚合中选择数据供给方的个数, $1-\sigma$ 是预算中用于支付数据供给方的报酬比例. 数据供给方 $P_i \in P$ 参与第 r 轮数据模型聚合的报酬函数可建模为

$$\rho_i^r = \frac{\xi_i(S_i, v(P_i)) \sum_{j=1}^{M} B_j \times (1-\sigma)}{r(w_G^*)}, \tag{13}$$

其中, $\xi_i(S_i, v(P_i))$ 是 P_i 在第 r 轮模型聚合中的由数据交易平台区块链节点共识验证的沙普利值.

3.4 联邦学习激励机制的设计目标

假设数据供给方 P_i 在上传本地模型参数时需要提供一个声明消息摘要 $M(\gamma_i, C_{\mathrm{unit}}^r(P_i))$ 至数据交易平台. 其中声明 γ_i 是 P_i 训练本地模型 w_i^t 所迭代的次数, 声明 $C_{\mathrm{unit}}^r(P_i)$ 是 P_i 在第 r 轮模型聚合中训练本地模型 w_i^t 所消耗的单位成本. 激励机制的目标是使得数据供给方的报酬分配以及资源配置相关的消息摘要 $M(\gamma_i, C_{\mathrm{unit}}^r(P_i))$ 满足如下性质.

P1 个人理性. 每个数据供给方 $P_i \in P$ 在参与数据交易后的效用和利润为正, 即

$$U_i(r) \geqslant 0, \rho_i^r - k(w_i^t) \times C_{\mathrm{unit}}^r(P_i) \geqslant 0. \tag{14}$$

P2 预算均衡. 所有数据供给方成本之和小于所获得的报酬之和, 所有数据交易平台区块链节点的聚合成本之和小于数据需求方预算手续费之和, 整个数据交易市场无需第三方投资, 即

$$r(w_G^*) \sum_{k=1}^{T} C_{\mathrm{com}}(n_j) \leqslant \sum_{j=1}^{M} B_j \sigma, \tag{15a}$$

$$r(w_G^*) \sum_{i=1}^{|P_\pi|} k(w_i^t) \times C_{\mathrm{unit}}^r(P_i) \leqslant \sum_{j=1}^{M} B_j (1-\sigma). \tag{15b}$$

P3 社会福利最大化. 数据交易市场需要满足福利最大化从而维系交易市场的稳定运营. 数据供给方需要选择使社会福利最大化的供给方案, 来优化本地模型迭代次数 $k(w_i^t)$. 预算约束下的社会福

利最大化问题可通过下式定义:

$$\arg\ \max_{k(w_i^t)} \sum_{i=1}^{N} (\rho_i^r - k(w_i^t) \times C_{\text{unit}}^r(P_i)), \tag{16a}$$

$$\text{s.t. } r(w_G^*) \leqslant \frac{\sum_{j=1}^{M} B_j \times \sigma}{\sum_{k=1}^{T} C_{\text{com}}(n_k)},\ k(w_i^t) \leqslant \frac{\sum_{j=1}^{M} B_j \times (1-\sigma)}{r(w_G^*)|P_\pi| C_{\text{unit}}^r(P_i)}. \tag{16b}$$

P4 激励相容. 在实现社会福利最大化的同时满足个人理性 [23].

P5 隐私保护. 数据模型交易过程中通过分配隐私预算 ϵ 保障梯度参数隐私, 采用差分隐私噪声注入 [9] 方式, 避免中间参数隐私泄露而遭受成员推理攻击 [7]. 将隐私预算作为成本要素纳入成本函数满足不同数据供给方的隐私需求.

P6 资源浪费最小化. 现有的数据模型交易市场大多通过数据收集、清洗、集成、整合等流程处理数据导致数据交易成本虚高、资源浪费. 因此, 本文在联邦学习分布式机器学习框架下, 量化了数据供给方成本效用与数据需求方支付报酬, 在预算约束下协作生产数据模型, 根据数据需求方预算动态调整训练过程中的本地模型资源开销, 最大程度避免资源浪费, 促进 "双碳" 目标达成.

4 不完全信息联邦学习激励机制

考虑到数据供给方的异质性与隐私保护需求, 本文设计了不完全信息下的联邦学习激励机制. 下文首先介绍激励机制流程, 然后将数据供给方本地模型训练的资源配置策略构建为贝叶斯博弈模型分析纳什均衡与社会福利, 提出一种隐私保护的贝叶斯博弈行动策略共识算法.

4.1 激励机制的流程描述

在联邦学习数据交易过程中, 基于区块链的数据交易平台 T_B 收集数据需求方 R_j 的需求消息摘要 $\mathcal{M}_j(E_i, B_i)$ 后通过数据交易平台 T_B 确定手续费比例 σ. 每轮模型聚合中, 数据供给方之间的本地训练策略可以构建为贝叶斯博弈 $P^{\text{BG}} = [P, \Omega, \mathcal{A}, \mathcal{T}, \mu, \succeq]$, 其中, P 是参与数据模型 w_G^* 协作加工的数据供给方集合, Ω 是 P 的状态空间集合, $\mathcal{A}$ 是 P 的行动空间集合, $\mathcal{T}$ 是 P 所能观察到的信号类型集合, μ 是状态空间为 Ω 时的先验概率分布, $\succeq$ 是 P 在笛卡尔积 $\Omega \times \mathcal{A}$ 上的行动偏好.

不完全信息联邦学习激励机制. 基于区块链的数据交易平台接收到数据需求方 R_i 的需求消息摘要 $\mathcal{M}_j(E_i, B_i)$ 后, 初始化模型参数 w_G^0 与本地模型初始迭代次数 δ_k, 并将初始化行动策略消息摘要 $\mathcal{M}_i^0(w_G^0, \delta_k)$ 分发给数据供给方 P_i. 数据供给方 P_i 收到消息 $\mathcal{M}_i^0(w_G^0, \delta_k)$ 后先返回一个确认参与数据交易的单位计算成本消息 $C_{\text{com}}(P_i)$ 后开始执行本地模型迭代, P_i 执行 δ_k 次本地模型迭代后将行动策略 $\mathcal{A}_i(\gamma_i^r, c_{\text{unit}}^i)$ 及经差分隐私加噪 [9] 的本地模型梯度参数 $\bigtriangledown\widetilde{g}(w_i^{t-1}, b_i)$ 发送至区块链平台共识节点. 共识节点接收到行动策略 $\mathcal{A}_i(\gamma_i^r, c_{\text{unit}}^i)$ 后执行贝叶斯博弈行动策略共识算法 (详见 4.4 小节), 并将贝叶斯纳什均衡行动策略 $\mathcal{A}_i(\gamma_i^{\text{NE}}, c_{\text{unit}}^i)$ 和相应奖励返回至各数据供给方 P_i, P_i 根据行动策略 $\mathcal{A}_i(\gamma_i^{\text{NE}}, c_{\text{unit}}^i)$ 执行本地模型训练, 直至模型收敛或者满足预算约束条件.

- $\gamma^{\text{NE}} = \{\gamma_1^r, \gamma_2^r, \ldots, \gamma_N^r\}$ 代表数据交易平台期望数据供给方集合 $P = \{P_1, P_2, \ldots, P_N\}$ 在第 r 轮模型聚合过程中的本地模型迭代行动策略集合, 通过供给方不完全信息下的贝叶斯博弈确定;
- $\rho_i^r \in \mathbb{R}^+$ 代表数据供给方 P_i 在第 r 轮模型聚合中上传梯度参数 $\bigtriangledown\widetilde{g}(w_i^{t-1}, b_i)$ 后获得的报酬奖励, 主要取决于模型聚合后效用 $U_i(r)$, 并采用沙普利值 $\xi_i(S_i, v(P_i))$ 保障利益的公平分配;

• 本地模型迭代行动策略 γ^{NE} 由数据交易平台链上节点 $T_B=\{n_1,n_2,\ldots,n_T\}$ 经贝叶斯行动策略共识验证后广播给数据供给方集合 P;

• 在报酬分配阶段, 数据交易平台共识委员会节点 $T_B=\{n_1,n_2,\ldots,n_T\}$ 根据区块链上记载的模型参数以及沙普利值计算报酬奖励 ρ_i^r, 并进行共识验证保证结果的可信性.

激励机制中数据供给方 P_i 需要按照数据交易平台经过共识验证后的行动策略 γ_i^r 执行本地模型训练, 数据供给方动态调整本地模型迭代次数, 以达到贝叶斯纳什均衡. 在此过程中, 数据供给方 P_i 作为不完全信息的博弈参与方不清楚竞争者 P_j 的本地模型训练时的资源配置策略 γ_j^r 与其成本函数 $C_{\mathrm{unit}}^r(P_j)$. 而数据交易平台链上节点作为行动决策者, 根据每轮参与聚合的数据供给方模型质量计算贝叶斯纳什均衡的资源配置行动策略并经过共识后返回给数据供给方. 在贝叶斯博弈决策与共识阶段, 数据供给方可以选择继续之前的本地模型迭代直至接收到新的行动策略后进行变更, 从而提高联邦学习执行效率. 联邦学习每轮模型聚合中数据供给方的贡献度由记载在链上的模型质量与新鲜度判断, 并通过沙普利值确定每个参与交易的数据供给方的奖励分配比例, 从而实现数据交易中利益的公平分配. 基于区块链的不完全信息联邦学习激励机制描述详见算法 1.

Algorithm 1 Incomplete information federated learning incentive mechanism based on blockchain

Input: Data requesters R demand message abstract sample space $\{\mathcal{M}_j(E_j,B_j)\}_{j=1}^M$;

Input: Data providers P local model action strategy sample space $\{\mathcal{A}_i(\gamma_i^{r-1},C_{\mathrm{unit}}^i)\}_{i=1}^N$;

Output: Local model training action strategy $\gamma^{\mathrm{NE}}=\{\gamma_1^r,\gamma_2^r,\ldots,\gamma_N^r\}$ and reward vector $\rho^r=\{\rho_1^r,\rho_2^r,\ldots,\rho_N^r\}$ of data providers P in the r-round model aggregation;

Data market platform T_B receives demand message abstract $\{\mathcal{M}_j(E_j,B_j)\}_{j=1}^M$;
Data market platform T_B broadcasts initial action strategy $\mathcal{A}_i^0(w_G^0\leftarrow f(\cdot),\gamma_i^0\leftarrow\delta_k)$;
//Data provider P_i returns confirmation message and executes local model iterations;
for $P_i\in P$ **do**
  if $\delta_k\leqslant t\leqslant\gamma_i^r$ **then**
    Compute $L(f(w_i^{t-1},x_i),y_i)$ with label data $d_i=(x_i,y_i)$;
    Compute gradient $\bigtriangledown g(w_i^{t-1},b_i)=\frac{1}{|b_i|}\sum_{j=1}^{|b_i|}\frac{\partial L(w_i^{t-1};d_j)}{\partial w_i^{t-1}}$;
    Update local model $w_i^t=w_i^{t-1}+\alpha\bigtriangledown g(w_i^{t-1},b_i)$;
  else
    Perturb local model gradient $\bigtriangledown\widetilde{g}(w_i^{t-1},b_i)=\bigtriangledown g(w_i^{t-1},b_i)+\langle\mathrm{Lap}(\mu,b=\frac{\Delta f}{\epsilon})\rangle$;
    Broadcast perturbed gradient $\bigtriangledown\widetilde{g}(w_i^{t-1},b_i)$ to data market platform T_B;
  end if
end for
//Data market platform T_B executes model aggregation and computes reward for data providers P;
while $T_r\leqslant T_{\mathrm{limit}}$ **do**
  for $n_i\in T_B$ **do**
    Compute global model $w_G^r=w_G^{r-1}+\alpha\frac{\sum_{i=1}^N\bigtriangledown\widetilde{g}_i(w_i^{t-1},b_i)}{N}$;
    Compute local model training action strategy $\gamma^{\mathrm{NE}}=\{\gamma_1^r,\gamma_2^r,\ldots,\gamma_N^r\}$ by Bayesian game (Subsection 4.2);
    Compute unit reward vector $\rho^r=\{\rho_1^r,\rho_2^r,\ldots,\rho_N^r\}$ by Shapley value;
  end for
end while
return Local model training action strategy $\gamma^{\mathrm{NE}}=\{\gamma_1^r,\gamma_2^r,\ldots,\gamma_N^r\}$ and unit reward vector $\rho^r=\{\rho_1^r,\rho_2^r,\ldots,\rho_N^r\}$;

4.2 数据供给方行动策略的贝叶斯博弈

本文将数据供给方与区块链数据交易平台的行动策略与消息互动构建为一个不完全信息的贝叶

斯博弈 (Bayesian game) [54], 通过求解贝叶斯纳什均衡得到每个数据供给方本地模型训练的资源配置行动策略 $\gamma^{\mathrm{NE}}=\{\gamma_1^r,\gamma_2^r,\ldots,\gamma_N^r\}$, 数据供给方行动策略的贝叶斯博弈定义如下.

(1) 数据交易平台 T_B 初始化行动: 数据交易平台 T_B 接收到需求消息摘要 $\mathcal{M}_j(E_i,B_i)$ 后, 将初始化行动策略 $\mathcal{A}_i^0(w_G^0\leftarrow f(\cdot),\gamma_i^0\leftarrow\delta_k)$ 广播给已注册的数据供给方 P_i.

(2) 数据供给方贝叶斯博弈: 数据供给方 P_i 将本地模型迭代行动策略 $\mathcal{A}_i(\gamma_i^r,c_{\mathrm{unit}}^r(P_i))$ 发送至联盟链共识节点, 其中本地模型迭代次数 γ_i^r 为 P_i 的行动策略选择, $c_{\mathrm{unit}}^r(P_i)$ 是单位成本函数. 假设参与数据模型交易的数据供给方集合 P 的行动策略组合为 $\mathcal{A}$, 其期望效用可通过下式定义:

$$\bar{U}^r=E[U_1^r(\mathcal{A}_1(\gamma_1^r,c_{\mathrm{unit}}^r(P_1))),U_2^r(\mathcal{A}_2(\gamma_2^r,c_{\mathrm{unit}}^r(P_2))),\ldots,U_N^r(\mathcal{A}_N(\gamma_N^r,c_{\mathrm{unit}}^r(P_N)))]. \tag{17}$$

数据供给方贝叶斯博弈的目标就是通过优化行动策略 $\mathcal{A}$, 求解贝叶斯纳什均衡 $P^{\mathrm{BNE}}(\bar{U}^r)$ 得到数据供给方行动策略的最优组合. 数据供给方集合 P 的贝叶斯博弈定义如下.

定义1 (贝叶斯博弈 (Bayesian game, BG)) 数据供给方集合 P 本地模型训练资源配置行动策略的贝叶斯博弈可以定义为 $P^{\mathrm{BG}}=[P,\Omega,\mathcal{A},\mathcal{T},\mu,\succeq]$, 其中,

- 参与者 Players: 参与数据模型 w_G^* 交易的所有数据供给方 $P_i,P_i\in P$;
- 状态空间 Ω: 数据供给方 P_i 的成本函数 $c_{\mathrm{unit}}^r(P_i)\in\mathbb{R}_+$;
- 行动策略 $\mathcal{A}$: 数据供给方集合 P 的本地模型训练资源配置行动策略 $\mathcal{A}=\{\mathcal{A}_i(\gamma_i^r,c_{\mathrm{unit}}^r(P_i))\}_{i=1}^N$;
- 信号类型空间 $\mathcal{T}$: 数据供给方集合 P 接收到的信号类型 $\{\tau_1,\tau_2,\ldots,\tau_N\}$;
- 先验概率分布 μ: 数据供给方集合 P 选择行动策略 $\mathcal{A}$ 的先验概率分布 $\{\mu_1,\mu_2,\ldots,\mu_N\}$;
- 偏好效用函数 $\succeq$: P_i 在 $\Omega\times\mathcal{A}$ 上的行动偏好, 由报酬函数 ρ_i^r 与成本函数 $C_{\mathrm{unit}}^r(P_i)$ 决定.

在贝叶斯博弈中, 每个数据供给方 P_i 的利润函数为 $\psi_i(\mathcal{A}_i,\mathcal{A}_{-i},\gamma_i^r,C_{\mathrm{unit}}^r(P_i))=\rho_i^r-\gamma_i^rC_{\mathrm{unit}}^r(P_i)$, 在不完全信息博弈情况下, 竞争者 P_j 不知道 P_i 的成本函数 $C_{\mathrm{unit}}^r(P_i)$. 在一个纯策略贝叶斯纳什均衡求解中, 每个数据供给方 P_i 会针对其他数据供给方策略的先验概率分布 μ 作出最优行动决策.

命题1 所有参与联邦学习数据模型 w_G^* 交易的数据供给方集合 P 的行动策略组合 $\mathcal{A}$ 是贝叶斯博弈的一个贝叶斯纳什均衡解当且仅当对于所有的 $P_i\in P$ 和以正概率发生的 $\gamma_i^r\in\mathbb{R}_+$ 满足

$$\begin{aligned}&E_{\gamma_i^r,C_{\mathrm{unit}}^r(P_i)}[U^r(\mathcal{A}_i(\tilde{\gamma}_i^r,C_{\mathrm{unit}}^r(P_i)),\mathcal{A}_{-i}(\tilde{\gamma}_{-i}^r,C_{\mathrm{unit}}^r(P\backslash\{P_i\})))]\\&\geqslant E_{\gamma_i^r,C_{\mathrm{unit}}^r(P_i)}[U^r(\mathcal{A}_i'(\tilde{\gamma}_i^r,C_{\mathrm{unit}}^r(P_i)),\mathcal{A}_{-i}'(\tilde{\gamma}_{-i}^r,C_{\mathrm{unit}}^r(P\backslash\{P_i\})))],\end{aligned} \tag{18}$$

对于所有 $\mathcal{A}_i\in\mathcal{A}$ 成立, 其中 $E_{\gamma_i^r,C_{\mathrm{unit}}^r(P_i)}$ 是数据供给方 P_i 在与竞争对手 $P\backslash\{P_i\}$ 贝叶斯博弈中的期望效用, 可通过区块链上的交易信息 $Tx=\{\triangledown\tilde{g}(w_i^t,D_i),w_G^r,\rho_i^r\}$ 计算.

证明 必要性: 对于数据供给方 P_i 以正概率发生的某个 $\mathcal{A}_i(\tilde{\gamma}_i^r,C_{\mathrm{unit}}^r(P_i))$ 不成立, 那么 P_i 在收到本地模型迭代行动策略 $\mathcal{A}_i(\tilde{\gamma}_i^r,C_{\mathrm{unit}}^r(P_i))$ 后, 可以改变策略, 从而使其效用更高, 这与贝叶斯纳什均衡的事实矛盾.

充分性: 如果对于所有以正概率发生的 $\mathcal{A}_i(\tilde{\gamma}_i^r,C_{\mathrm{unit}}^r(P_i))$ 都成立, 那么 P_i 选择其他策略得到的效用不可能高于选择策略 $\mathcal{A}_i(\tilde{\gamma}_i^r,C_{\mathrm{unit}}^r(P_i))$ 的效用.

定义2 (贝叶斯纳什均衡) 对于贝叶斯博弈 P^{BG} 的纳什均衡解 $P^{\mathrm{BNE}}(\bar{U}^r)$, 如果数据供给方 P_i 选择采用本地模型迭代轮数为 γ_i^r 参与模型聚合, 则将其记为 $\mathcal{A}_i(\gamma_i^r,c_{\mathrm{unit}}^r(P_i))=1$. 假设第 r 轮模型聚合中数据供给方选择比率 $\pi=1$, 无论其他数据供给方 P_j 是否参与聚合, P_i 参与第 r 轮模型聚合的期望利润为 $(\rho_i^r(\gamma_i))^2-\gamma_i^r\times C_{\mathrm{unit}}^r(P_i)$, 期望效用为 $\bar{U}^r(\mathcal{A}_i(\gamma_i^r,c_{\mathrm{unit}}^r(P_i)))$. 因此, 贝叶斯博弈 P^{BG} 达

到贝叶斯纳什均衡状态 $P^{\mathrm{BNE}}(\bar{U}^r)$ 是当 P_i 参与第 r 轮模型聚合时当且仅当 $\rho_i^r(\gamma_i)$ 满足

$$\rho_i^r(\gamma_i) \geqslant \sqrt{\frac{\gamma_i^r \times C_{\mathrm{unit}}^r(P_i)}{1-\mathrm{Prob}(\mathcal{A}_i(\gamma_i^r, c_{\mathrm{unit}}^i)=1)}}, \quad 0<\mathrm{Prob}(\mathcal{A}_i(\gamma_i^r, c_{\mathrm{unit}}^i)=1)<1, \tag{19}$$

其中 Prob(·) 函数代表 $\mathcal{A}_i(\gamma_i^r, C_{\mathrm{unit}}^r(P_i))=1$ 的概率, $\gamma_i^r \times C_{\mathrm{unit}}^r(P_i)$ 代表采用 $\mathcal{A}_i(\gamma_i^r, C_{\mathrm{unit}}^r(P_i))=1$ 行动策略的成本开销.

根据分割线原则 (cutoff rule) [54], 当数据供给方 P_i 满足贝叶斯纳什均衡条件时, P_i 的最优选择是 $\mathcal{A}_i(\gamma_i^r, C_{\mathrm{unit}}^r(P_i))=1$. 当被选择的数据供给方数量 $|P_\pi|=1$ 时, 则 $\mathcal{A}_i(\gamma_i^r, C_{\mathrm{unit}}^r(P_i))=1$ 与 $\mathcal{A}_i(\gamma_i^r, C_{\mathrm{unit}}^r(P_i))=0$ 之间是无差异的. 当 $|P_\pi|>1$ 时, P_i 的分割线总是 (弱) 大于 $\sqrt{\gamma_i^r \times C_{\mathrm{unit}}^r(P_i)}$. 因为每个 P_i 都希望不改变策略, 从而可以实现在降低成本开销的情况下, 得到更高质量的聚合模型.

4.3 激励机制所满足的特性分析

本小节从个人理性、预算均衡、社会福利最大化等方面对不完全信息联邦学习激励机制进行分析, 证明本文激励机制的优越性.

命题2 (个人理性) 在贝叶斯博弈 P_i^{BNE} 的任意纳什均衡 $\mathcal{A}_i^{\mathrm{NE}}(\gamma_i^r, C_{\mathrm{unit}}^r(P_i))$ 行动策略下, 每个数据供给方 P_i 拥有非负报酬与效用, 即 $\rho_i^r(\gamma_i), \bar{U}^r \geqslant 0$.

证明 对于 $P_i \in P$, $\rho_i^r(\gamma_i) \geqslant \rho_i^0(\gamma_i)=0, \bar{U}^r(\mathcal{A}_i(\tilde{\gamma}_i^r, C_{\mathrm{unit}}^r(P_i))) \geqslant \bar{U}^0(\mathcal{A}_i(\tilde{\gamma}_i^0, C_{\mathrm{unit}}^r(P_i)))=0$.

命题3 (预算均衡) 在贝叶斯博弈 P_i^{BG} 的任意纳什均衡 $\mathcal{A}_i^{\mathrm{NE}}(\gamma_i^r, C_{\mathrm{unit}}^r(P_i))$ 行动策略下, 所有数据供给方的成本开销之和小于数据需求方总预算 $\sum_{j=1}^M B_j$, 即

$$r(w_G^*)\left[\sum_{k=1}^T C_{\mathrm{com}}(n_k) + |P_\pi|\gamma_i^r C_{\mathrm{unit}}^r(P_i)\right] \leqslant \sum_{j=1}^M B_j. \tag{20}$$

证明 将 $\mathcal{A}_i^{\mathrm{NE}}(\gamma_i^r, C_{\mathrm{unit}}^r(P_i))$ 与预算 $\sum_{j=1}^M B_j$ 带入式 (16b) 求和可得 $r(w_G^*)\sum_{k=1}^T C_{\mathrm{com}}(n_k) + \gamma_i^r r(w_G^*)|P_\pi|C_{\mathrm{unit}}^r(P_i) \leqslant \sum_{j=1}^M B_j$, 提取公因子 $r(w_G^*)$ 后式 (20) 得证.

命题4 (社会福利最大化) 如果数据供给方 P_i 按照行动策略 $\mathcal{A}_i(\gamma_i^r, C_{\mathrm{unit}}^r(P_i))$ 进行本地模型迭代, 则必定可以达到纳什均衡状态, 即存在可行解 $(\gamma_i^{\mathrm{NE}}, \lambda^{\mathrm{NE}})$ 使得在预算约束下实现社会福利最大化.

证明 数据供给方 P_i 在第 r 轮模型聚合过程中的收益为 ρ_i^r, 成本包括本地模型训练成本 $\gamma_i^r \times C_{\mathrm{unit}}^r(P_i)$ 以及模型聚合成本 $\sum_{k=1}^T C_{\mathrm{com}}(n_k)$, 总预算为 $\sum_{j=1}^M B_j$, 则社会福利最大化可定义为

$$\arg\max_{\gamma_i^r} \sum_{i=1}^N(\rho_i^r - \gamma_i^r \times C_{\mathrm{unit}}^r(P_i)) - \sum_{k=1}^T C_{\mathrm{com}}(n_k), \tag{21a}$$

$$\text{s.t. } r(w_G^*)\left[\sum_{i=1}^N \gamma_i^r \times C_{\mathrm{unit}}^r(P_i) + \sum_{k=1}^T C_{\mathrm{com}}(n_k)\right] \leqslant \sum_{j=1}^M B_j. \tag{21b}$$

为了求解约束条件下的凸函数最大化问题, 采用拉格朗日 (Lagrange) 乘数法求解. 将预算约束下社会福利最大化问题的拉格朗日函数构建为 $L(\gamma_i^r, \lambda) \to \mathbb{R}$, 即

$$\begin{aligned} L(\gamma_i^r, \lambda) = &\sum_{i=1}^N(\rho_i^r - \gamma_i^r \times C_{\mathrm{unit}}^r(P_i)) - \sum_{k=1}^T C_{\mathrm{com}}(n_k) \\ &+ \lambda\left(\sum_{j=1}^M B_j - r(w_G^*)\left[\sum_{i=1}^N \gamma_i^r \times C_{\mathrm{unit}}^r(P_i) + \sum_{k=1}^T C_{\mathrm{com}}(n_k)\right]\right), \end{aligned} \tag{22}$$

其中, λ 是拉格朗日乘数向量, 对 γ_i^r 和 λ 的一阶偏导数方程组进行求解, 即

$$\begin{cases} \sum_{i=1}^{N} C_{\text{unit}}^r(P_i) + \lambda r(w_G^*) \sum_{i=1}^{N} C_{\text{unit}}^r(P_i) = 0, \\ \sum_{j=1}^{M} B_j - r(w_G^*)[\sum_{i=1}^{N} \gamma_i^r \times C_{\text{unit}}^r(P_i) + \sum_{k=1}^{T} C_{\text{com}}(n_k)] = 0, \end{cases} \tag{23}$$

可得 $L(\gamma_i^r, \lambda)$ 的可行解 $(\gamma_i^{\text{NE}}, \lambda^{\text{NE}})$, 即

$$\gamma_i^{\text{NE}} = \frac{\sum_{j=1}^{M} B_j}{r(w_G^*) \sum_{i=1}^{N} C_{\text{unit}}^r(P_i)} - \frac{\sum_{k=1}^{T} C_{\text{com}}(n_k)}{\sum_{i=1}^{N} C_{\text{unit}}^r(P_i)}, \ \ \lambda^{\text{NE}} = -\frac{1}{r(w_G^*)}. \tag{24}$$

基于式 (21a) 的上凸函数性质 (二阶导数小于 0), 可证明当数据供给方 P_i 的本地模型迭代行动策略为 $\mathcal{A}_i(\gamma_i^{\text{NE}}, C_{\text{unit}}^r(P_i))$ 时, 数据供给方满足预算约束下的社会福利最大化且不倾向于改变策略, 达到了贝叶斯纳什均衡状态.

引理1 (纳什均衡鞍点) 对于 $L(\gamma_i^r, \lambda)$ 的鞍点 $(\gamma_i^{r*}, \lambda^*)$, 当数据供给方 P_i 的本地模型迭代行动策略为 $\mathcal{A}_i^r(\gamma_i^{r*}, C_{\text{unit}}^r(P_i))$ 时, 可以达到贝叶斯纳什均衡状态.

证明 根据斯莱特条件 (Slater condition) [5], 福利最大化问题强对偶性条件成立. 假设 $L(\gamma_i^r, \lambda)$ 的拉格朗日鞍点用 $(\gamma_i^{r*}, \lambda^*)$ 表示, 满足 $L(\gamma_i^r, \lambda^*) \leqslant L(\gamma_i^{r*}, \lambda^*) \leqslant L(\gamma_i^{r*}, \lambda)$, 对于任意的 $\gamma_i^r \in [0, \bar{\gamma_i^r}]^N$, 且 $\lambda \in \mathbb{R}_+$ 成立, 则鞍点 $L(\gamma_i^{r*}, \lambda^*)$ 满足预算约束下福利最大化的卡鲁什 – 库恩 – 塔克 (Karush-Kuhn-Tucker, KKT) 条件为

$$\frac{\partial L(\gamma_i^{r*}, \lambda^*)}{\partial \gamma_i^{r*}} = \sum_{i=1}^{N} C_{\text{unit}}^r(P_i) + \lambda^* \left(r(w_G^*) \sum_{i=1}^{N} C_{\text{unit}}^r(P_i) \right) = 0, \tag{25a}$$

$$\lambda^* \left\{ \sum_{j=1}^{M} B_j - r(w_G^*) \left[\sum_{i=1}^{N} \gamma_i^{r*} \times C_{\text{unit}}^r(P_i) + \sum_{k=1}^{T} C_{\text{com}}(n_k) \right] \right\} = 0, \tag{25b}$$

$$r(w_G^*) \left[\sum_{i=1}^{N} \gamma_i^{r*} \times C_{\text{unit}}^r(P_i) + \sum_{k=1}^{T} C_{\text{com}}(n_k) \right] \leqslant \sum_{j=1}^{M} B_j. \tag{25c}$$

当 $\lambda^* \neq 0$ 时 $\gamma_i^r = \gamma_i^{r*}$ 同时满足式 (21) 和 (23). 因此, 可以证明如果 $(\gamma_i^{r*}, \lambda^*)$ 是 $L(\gamma_i^r, \lambda)$ 的鞍点, 也是贝叶斯博弈 P_i^{BG} 的纳什均衡解.

4.4 贝叶斯博弈行动策略共识算法

为了保障激励机制的有效性与博弈行动策略的可信性, 本文进一步提出一种隐私保护的贝叶斯博弈行动策略共识算法 PPBG-AC, 该算法使数据供给方能在基于区块链的数据交易平台协调下实现贝叶斯纳什均衡 P_i^{BNE}. 算法主要解决贝叶斯博弈中行动策略的不可信问题以及中间梯度参数广播造成的隐私泄露. PPBG-AC 的主要思想是将数据供给方的消息摘要上传至联盟链共识委员会, 通过链上节点共识保障行动策略结果的不可篡改和可追溯特性. 通过贝叶斯博弈求解社会福利最大化问题直至收敛于拉格朗日鞍点, 最后广播数据供给方的行动策略建议, 保障策略的一致性与可信性.

PPBG-AC 共识算法基于消逝时间量的共识算法 (proof-of-elapsed-time, PoET) [55] 进行交易时间戳验证, 共识委员会节点调用 “CheckTimer” 智能合约检验时间戳. 联盟链共识委员会通过委托权益证明 (delegated proof of stake, DPoS) 算法 [56] 遴选高可靠共识节点, 并由共识委员会达成贝叶斯博弈行动策略的一致意见后在区块链网络中进行全网广播. 共识委员会内部采用实用拜占庭容错共识算法

(practical Byzantine fault tolerance, PBFT) [57] 保障贝叶斯博弈中行动策略共识消息传播的可靠性与共识节点的容错性.

在贝叶斯博弈行动策略共识过程中, 区块链共识委员会首先获取数据供给方噪声注入 [9] 后的消息摘要进行验证, 对验证通过的交易进行排序并求解社会福利最大化问题鞍点, 即贝叶斯纳什均衡. 共识节点收到待验证交易之后由共识委员会进行投票共识, 获得半数以上节点投票通过的贝叶斯纳什均衡解, 即可广播至数据供给方, 在此之前数据供给方继续之前的本地模型迭代直至接收到新的行动策略再进行变更. 消息验证过程中需要同时验证数据供给方签名与模型新鲜度时间戳, 并随机抽取模型参数通过公共数据集进行模型质量验证, 避免数据供给方提供虚假的模型参数. 在不完全信息联邦学习激励机制约束下, 数据供给方倾向于将更高质量的模型尽早提交至数据交易平台.

PPBG-AC 算法共识过程如下: 数据交易平台 T_B 接收模型迭代行动策略 $\{\mathcal{A}_i(\gamma_i^{r-1}, C_{\text{unit}}^{r-1}(P_i))\}_{i=1}^N$, 链上节点接收到行动策略后采用 CheckTimer 智能合约进行时间戳校验, 根据可信时间戳对交易排序后求解社会福利最大化鞍点, 即贝叶斯纳什均衡 P_i^{BNE}. 将计算结果广播至共识委员会进行投票共识, 共识结束后区块链给出本地模型行动策略建议 γ_i^r. 数据供给方 P_i 根据 γ_i^r 执行本地模型训练, 不断迭代直至聚合模型 w_G^* 收敛或满足预期计算时间上限. PPBG-AC 共识算法描述如算法 2 所示.

Algorithm 2 Privacy-preserving Bayesian game action strategy consensus algorithm (PPBG-AC)

Input: Data providers P model training action strategy sample space $\{\mathcal{A}_i(\gamma_i^{r-1}, C_{\text{unit}}^{r-1}(P_i))\}_{i=1}^N$;
Output: Local model training action strategy consensus results $\gamma^{\text{NE}} = \{\gamma_1^r, \gamma_2^r, \ldots, \gamma_N^r\}$;
Data market platform T_B receives local model training action strategy $\{\mathcal{A}_i(\gamma_i^{r-1}, C_{\text{unit}}^{r-1}(P_i))\}_{i=1}^N$;
//Consensus committee nodes verify signature and execute CheckTimer smart contract to verify timestamp;
if Verify(Sign(P_i)) and CheckTimer($T_i^k(w_i^k)$) **then**
    $\text{sort}_{T_i^k(w_i^k)}$;
    Solving the saddle point of social welfare maximization $L(\gamma_i^{r*}, \lambda^*)$;
    γ^{NE} append γ_i^{r*};
    for $n_k \in C_{\text{committee}}$ **do**
        $\text{vote}(\gamma_i^r) \to C_{\text{committee}}$;
    end for
    if $\sum_{i=1}^T \text{vote}(\gamma^{\text{NE}}) \geqslant \frac{1}{2}|C_{\text{committee}}| + 1$ **then**
        return $\gamma^{\text{NE}} = \{\gamma_i^r\}_{i=1}^N$;
    else
        return 0;
    end if
else
    return 0;
end if

5 方案对比与性能评估

本节从方案对比与收敛性能评估两个方面对激励机制进行分析评估. 方案对比主要基于已有的联邦学习激励机制与本文在不完全信息下的基于贝叶斯博弈的激励机制进行对比. 在性能评估方面, 采用在联邦学习常用的公开图片数据集 MNIST 和 CIFAR-10 进行算法的收敛性能评估.

5.1 方案对比

对比已有的跨孤岛联邦学习激励机制 (cross-silo federated learning incentive mechanism,

表 2 方案对比

Table 2 Comparison of the schemes

	CSFL [5]	SGFL [19]	CTFL [20]	FMore [21]	Our BGFL
Theory method	No-cooperative game	Stackelberg game	Contract theory	Auction theory	Bayesian game
Privacy-preserving	✓	×	✓	✓	✓
Public goods properties	✓	×	×	×	✓
Blockchain-based	×	×	✓	×	✓

CSFL) [5]、基于斯塔克伯格博弈的联邦学习激励机制 (Stackelberg game-based federated learning incentive mechanism, SGFL) [19]、基于契约理论的联邦学习激励机制 (contract theory-based federated learning incentive mechanism, CTFL) [20]、基于多维拍卖理论的联邦学习激励机制 (multi-dimensional auction theory-based incentive scheme, FMore) [21] 与本文基于贝叶斯博弈的联邦学习激励机制 (Bayesian game-based federated learning incentive mechanism, BGFL) 在理论方法、隐私保护等方面的特性, 如表 2 所示. 对比已有相关工作, 本文基于贝叶斯博弈理论实现了不完全信息的联邦学习激励机制, 更适用于有异质性数据供给方参与的联邦学习数据交易场景. 不完全信息博弈的优势主要在于避免实际数据交易场景中异质性数据供给方的完全信息约束, 可以激励更广泛的数据供给方参与. 本文在基于区块链的去中心化平台架构下, 考虑了多数据需求方访问数据模型时的公共物品属性, 在不完全信息的贝叶斯博弈理论基础上实现了数据供给方的资源优化配置与奖励公平分配, 进而有利于提高数据供给方参与的积极性以及全局模型的收敛性能.

5.2 收敛性能评估

性能评估采用 Python v3.6.10 和 PyTorch v0.4.1 实现多数据供给方的联邦学习数据模型协作训练. 实验测试数据集使用 MNIST [58] 和 CIFAR-10 [59] 数据集, MNIST 包含 60000 个训练样本和 10000 个测试样本, CIFAR-10 包含 50000 个训练样本和 10000 个测试样本. 采用小批量随机梯度下降算法 [42] 迭代优化本地模型, 神经网络模型选用多层感知机 (multi-layer perceptron, MLP) [60] 和卷积神经网络 (convolutional neural networks, CNN) [61] 进行预测模型的迭代训练. 模型聚合采用自适应模型聚合算法 (FedAdp) [13] 实现全局模型的参数更新. 实验代码基于轻量级的联邦学习框架 [42] 模拟数据供给方的协作训练过程, 实验参数设置详见表 3. 实验环境为 Intel (R) Core (TM) i7-9700 CPU, 3.00 GHz, 16 GB RAM, 操作系统为 Windows10. 在本文设计的基于贝叶斯博弈的联邦学习激励机制约束下, 对比两种设置下的模型收敛情况, 其中每组实验结果为执行 10 次的平均值.

设置 1. 不同数据供给方协作. 考虑到数据集的有效性, 将 MNIST 数据集等分给 10 个数据供给方. 从数据供给方集合 $P=\{P_0,\ldots,P_9\}$ 中随机选取 3 个数据供给方 $P_\pi=\{P_1,P_8,P_0\}$ 参与数据交易. 在设置 1 中, 对比 MLP 模型在 3 个数据供给方的所有协作组合的准确率与损失收敛情况 (图 2), 发现异质性数据供给方参与数据交易时准确率之间存在差异, 通过数据协作可以提升低质量数据供给方的模型质量, 避免高质量数据供给方的过拟合. 3 个数据供给方 $P_\pi=\{P_1,P_8,P_0\}$ 在模型聚合 100 次之后准确率达到 95.34%, 损失值收敛于 0.1616, 相较于单节点的模型上收敛效果有显著提升, 这表明了在本文激励机制作用下多数据供给方参与的联邦学习全局模型具有较好的收敛性.

设置 2. 不同隐私预算. 本文在数据供给方的成本函数中考虑了隐私成本, 对于不同数据供给方的隐私需求进行仿真模拟. 在 MLP 模型下不同隐私预算对数据模型收敛性与分布式异步联邦学习 AsyFL [62] 对比, 如图 3 所示. 当数据供给方的隐私预算分别设置为 $\epsilon_1=1,\epsilon_2=3,\epsilon_3=4,\epsilon_4=5$ 时, 全

表 3　实验参数设置

Table 3　Experiment parameter settings

Parameter	Value	Parameter	Value
Number of data providers	$N = 10$	Number of data requesters	$M = 5$
Learning rate	$\alpha_t = 0.01$	Total budget of data requesters	$B = 5000$
Activation function	ReLU	Model utility parameters	$\mu_1 = 10, \mu_2 = 2$
Model training batch size	bs = 64	Initial local model epochs	$\delta_k = 5$
CPU cycle and clock frequency	$c_i = 5, f_i = 3$	The capacitance parameter	$\zeta = 2$
SGD momentum	0.5	Consensus committee	$T = 7$
Regulatory factor	$\alpha = 0.5, \mu = 0.8, \sigma = 0.2$	Other parameters	$\beta = 0.3, \theta = 16$

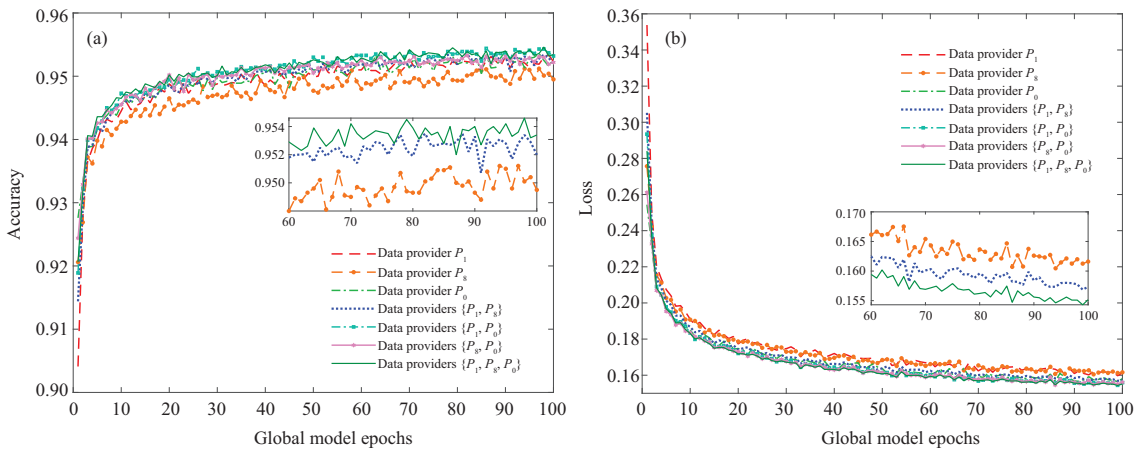

图 2　(网络版彩图) 在 3 个数据供给方参与的场景下不同数据供给方协作时模型准确率与损失

Figure 2　(Color online) The convergence of FL global model when different data providers cooperate under three data providers participating scenario. (a) Test accuracy; (b) test loss

局模型聚合 100 次之后的模型准确率分别为 94.12%, 94.31%, 94.68%, 95.36%, 模型损失分别为 0.1973, 0.1829, 0.1616, 0.1541. 在相同隐私预算设置下全局模型聚合 100 次之后的模型准确率相较于 AsyFL 模型 [62] 提升了 4.73%. 本文激励机制在考虑隐私成本的前提下, 通过贝叶斯博弈调整数据供给方本地模型迭代行动策略, 在优化数据供给方资源配置的同时提升了聚合模型的准确率.

进一步, 在设置 2 下对比了 4 个数据供给方在不同数据集下 MLP 与 CNN 模型聚合 100 次的总计成本效用和社会福利情况, 如图 4 所示. 随着隐私预算的增加, 在不同的模型与数据集下数据供给方的成本都有明显的增加. 在模型为 MLP 数据集为 MNIST 的设置中, 当隐私预算从 $\epsilon_1 = 1$ 增加至 $\epsilon_4 = 5$ 时数据供给方平均成本增加 16.53%, 而平均效用仅增加 0.11%. 这是因为随着隐私预算的增加, 模型噪声扰动减小, 模型质量提升, 弥合了部分隐私成本的增加. 通过纵向对比数据供给方参与每轮迭代的单位效用, 发现随着隐私预算的增加数据供给方单位效用明显增加, 当 $\epsilon_2 = 3$ 时, 数据供给方参与第 1 次与第 100 次模型聚合的单位效用从 42.74 增加至 45.28, 增加了 5.89%. 因此, 随着数据供给方的参与, 每轮参与模型聚合的单位效用不断提升也反映了对数据供给方的激励作用.

在仿真实验中, 数据供给方的成本效用采用了系统模型中定义的成本效用量化方法, 即成本采用二次能耗模型结合隐私成本计算, 效用通过模型交叉熵与新鲜度量化计算, 并对 100 轮模型聚合之后的成本效用进行求和. 联邦学习模型聚合训练采用 python asyncio 实现多线程数据供给方的协作训练,

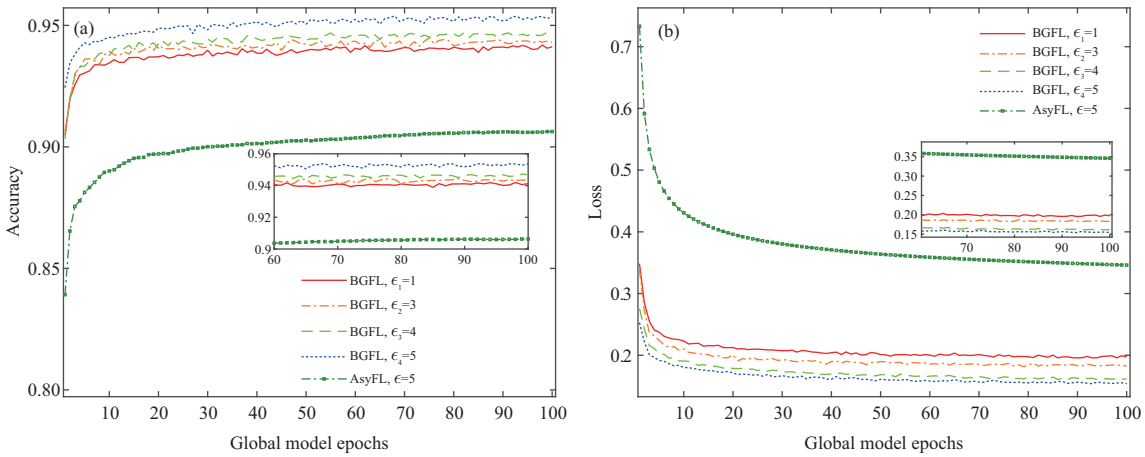

图 3 (网络版彩图) 不同隐私预算对联邦学习全局模型收敛性的影响对比

Figure 3 (Color online) The comparison of different privacy budgets on the convergence of federated learning global model. (a) Test accuracy; (b) test loss

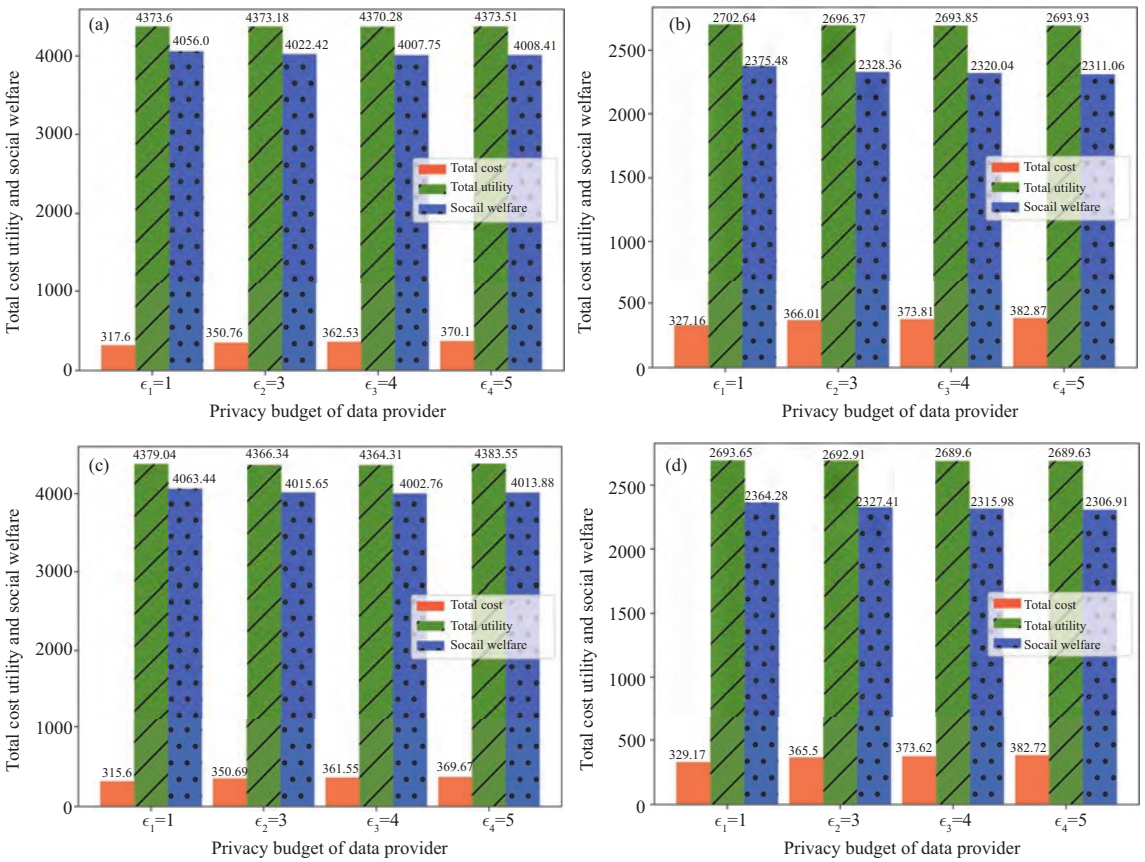

图 4 (网络版彩图) 不同模型与数据集下数据供给方的总计成本效用和社会福利对比

Figure 4 (Color online) Total cost utility and social welfare comparison of data providers under different models and datasets. (a) MLP MNIST; (b) MLP CIFAR-10; (c) CNN MNIST; (d) CNN CIFAR-10

基于联盟区块链的数据交易平台形成了数据供给方本地模型与最新聚合模型参数的链上数据池, 数据供给方将训练好的本地模型参数加入数据池, 并换取最新的全局模型参数. 此过程时间开销主要取决于数据供给方消耗的本地训练时间, 表明了本方案的实用性.

6 总结

数字经济背景下, 如何在保障数据资产隐私安全的前提下建立公平、可信的数据交易市场已经成为备受关注的问题. 基于此, 本文提出了一种基于区块链和贝叶斯博弈的联邦学习激励机制来解决实际数据交易市场中的供给不足问题, 消除了对于数据供给方的完全信息假设约束. 本文采用贝叶斯博弈分析了不完全信息下的数据供给方的本地模型训练资源配置策略, 通过求解社会福利最大化鞍点得到贝叶斯博弈的纳什均衡. 理论分析表明在本文激励机制的作用下, 联邦学习数据交易市场可以实现社会福利最大化、个人理性以及预算均衡. 通过轻量级联邦学习框架模拟多数据供给方协作的仿真实验验证了本文激励机制的有效性与实用性. 未来工作中将进一步探索数据供给方的多维度异质性, 并探索不完全信息下的其他博弈模型的求解与应用.

参考文献

1 Xu X, Li K O B, Tian X X. Research progress of data production factors. Economic Perspect, 2021, 4: 142–158 [徐翔, 厉克奥博, 田晓轩. 数据生产要素研究进展. 经济学动态, 2021, 4: 142–158]
2 Chen L, Koutris P, Kumar A. Towards model-based pricing for machine learning in a data marketplace. In: Proceedings of the International Conference on Management of Data, 2019. 1535–1552
3 Lin B R, Kifer D. On arbitrage-free pricing for general data queries. Proc VLDB Endow, 2014, 7: 757–768
4 Yang Q, Liu Y, Cheng Y, et al. Federated learning. Synthesis Lectures Artif Intell Machine Learn, 2019, 13: 1–207
5 Tang M, Wong V W S. An incentive mechanism for cross-silo federated learning: a public goods perspective. In: Proceedings of IEEE Conference on Computer Communications, 2021. 1–10
6 Liu B, Ding M, Shaham S, et al. When machine learning meets privacy: a survey and outlook. ACM Comput Surv, 2022, 54: 1–36
7 Shokri R, Stronati M, Song C, et al. Membership inference attacks against machine learning models. In: Proceedings of IEEE Symposium on Security and Privacy, 2017. 3–18
8 Fang M, Cao X, Jia J, et al. Local model poisoning attacks to byzantine-robust federated learning. In: Proceedings of the 29th USENIX Security Symposium, 2020. 1605–1622
9 Dwork C. Differential privacy. In: Proceedings of the 33rd International Colloquium on Automata, Languages and Programming, 2006. 1–12
10 Dwork C. Differential privacy: a survey of results. In: Proceedings of International Conference on Theory and Applications of Models of Computation, 2008. 1–19
11 Karimireddy S P, Rebjock Q, Stich S, et al. Error feedback fixes signSGD and other gradient compression schemes. In: Proceedings of International Conference on Machine Learning, 2019. 3252–3261
12 Lin Y, Han S, Mao H, et al. Deep gradient compression: reducing the communication bandwidth for distributed training. 2017. ArXiv:1712.01887
13 Zhu J M, Zhang Q N, Gao S, et al. Privacy preserving and trustworthy federated learning model based on blockchain. Chin J Comput, 2021, 44: 2466–2486 [朱建明, 张沁楠, 高胜, 等. 基于区块链的隐私保护可信联邦学习模型. 计算机学报, 2021, 44: 2466–2486]
14 Zheng Z, Xie S, Dai H N, et al. Blockchain challenges and opportunities: a survey. Int J Web Grid Serv, 2018, 14: 352–375
15 Zheng Z, Xie S, Dai H N, et al. An overview on smart contracts: challenges, advances and platforms. Future Generation Comput Syst, 2020, 105: 475–491
16 Kim H, Park J, Bennis M, et al. Blockchained on-device federated learning. IEEE Commun Lett, 2019, 24: 1279–1283

17 Weng J, Weng J, Zhang J, et al. DeepChain: auditable and privacy-preserving deep learning with blockchain-based incentive. IEEE Trans Dependable Secure Comput, 2019, 18: 2438–2455
18 Stiglitz J E. Knowledge as a global public good. In: Global Public Goods: International Cooperation in the 21st Century. New York: Oxford University Press, 1999. 308: 308–325
19 Sarikaya Y, Ercetin O. Motivating workers in federated learning: a Stackelberg game perspective. IEEE Netw Lett, 2019, 2: 23–27
20 Kang J, Xiong Z, Niyato D, et al. Incentive mechanism for reliable federated learning: a joint optimization approach to combining reputation and contract theory. IEEE Internet Things J, 2019, 6: 10700–10714
21 Zeng R, Zhang S, Wang J, et al. FMore: an incentive scheme of multi-dimensional auction for federated learning in MEC. In: Proceedings of IEEE 40th International Conference on Distributed Computing Systems, 2020. 278–288
22 Li T, Sahu A K, Talwalkar A, et al. Federated learning: challenges, methods, and future directions. IEEE Signal Process Mag, 2020, 37: 50–60
23 Li Y, Courcoubetis C, Duan L. Recommending paths: follow or not follow? In: Proceedings of IEEE Conference on Computer Communications, 2019. 928–936
24 Tran N H, Bao W, Zomaya A, et al. Federated learning over wireless networks: optimization model design and analysis. In: Proceedings of IEEE Conference on Computer Communications, 2019. 1387–1395
25 Huang Y, Chu L, Zhou Z, et al. Personalized cross-silo federated learning on non-iid data. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2021. 35: 7865–7873
26 Sattler F, Wiedemann S, Müller K R, et al. Robust and communication-efficient federated learning from non-i.i.d. data. IEEE Trans Neural Netw Learn Syst, 2019, 31: 3400–3413
27 Wang H, Kaplan Z, Niu D, et al. Optimizing federated learning on non-iid data with reinforcement learning. In: Proceedings of IEEE Conference on Computer Communications, 2020. 1698–1707
28 Liu K, Qiu X, Chen W, et al. Optimal pricing mechanism for data market in blockchain-enhanced Internet of Things. IEEE Internet Things J, 2019, 6: 9748–9761
29 Zheng X. Data trading with differential privacy in data market. In: Proceedings of the 6th International Conference on Computing and Data Engineering, 2020. 112–115
30 Jung K, Lee J, Park K, et al. PRIVATA: differentially private data market framework using negotiation-based pricing mechanism. In: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, 2019. 2897–2900
31 Khapre S P, Dhasarathan C, Puviyarasi T, et al. Blockchain-based data market (BCBDM) framework for security and privacy: an analysis. In: Applications of Big Data in Large-and Small-Scale Systems. Hershey: IGI Global, 2021. 186–205
32 Zhao Y, Yu Y, Li Y, et al. Machine learning based privacy-preserving fair data trading in big data market. Inf Sci, 2019, 478: 449–460
33 Liu Z, Hacigümüs H. Online optimization and fair costing for dynamic data sharing in a cloud data market. In: Proceedings of the ACM SIGMOD International Conference on Management of Data, 2014. 1359–1370
34 Jiao Y, Wang P, Niyato D, et al. Auction mechanisms in cloud/fog computing resource allocation for public blockchain networks. IEEE Trans Parallel Distrib Syst, 2019, 30: 1975–1989
35 Zhan Y, Zhang J, Hong Z, et al. A survey of incentive mechanism design for federated learning. IEEE Trans Emerg Top Comput, 2021. doi: 10.1109/TETC.2021.3063517
36 Feng S, Niyato D, Wang P, et al. Joint service pricing and cooperative relay communication for federated learning. In: Proceedings of International Conference on Internet of Things and IEEE Green Computing and Communications and IEEE Cyber, Physical and Social Computing and IEEE Smart Data, 2019. 815–820
37 Yang X Q. An exterior point method for computing points that satisfy second-order necessary conditions for a $C^{1,1}$ optimization problem. J Math Anal Appl, 1994, 187: 118–133
38 Wang S, Tuor T, Salonidis T, et al. Adaptive federated learning in resource constrained edge computing systems. IEEE J Sel Areas Commun, 2019, 37: 1205–1221
39 Nishio T, Yonetani R. Client selection for federated learning with heterogeneous resources in mobile edge. In: Proceedings of IEEE International Conference on Communications, 2019. 1–7
40 Khan L U, Pandey S R, Tran N H, et al. Federated learning for edge networks: resource optimization and incentive

mechanism. IEEE Commun Mag, 2020, 58: 88–93
41 Wang Z, Hu Q, Li R, et al. Incentive mechanism design for joint resource allocation in blockchain-based federated learning. ArXiv:2202.10938
42 McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data. In: Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 2017. 1273–1282
43 Peters S E, McClennen M. The paleobiology database application programming interface. Paleobiology, 2016, 42: 1–7
44 Kang J, Xiong Z, Niyato D, et al. Incentive design for efficient federated learning in mobile networks: a contract theory approach. In: Proceedings of IEEE VTS Asia Pacific Wireless Communications Symposium, Singapore, 2019. 1–5
45 Abadi M, Chu A, Goodfellow I, et al. Deep learning with differential privacy. In: Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, 2016. 308–318
46 Liu R X, Cao Y, Yoshikawa M, et al. FedSel: federated SGD under local differential privacy with top-k dimension selection. In: Proceedings of the 25th International Conference on Database Systems for Advanced Applications, 2020. 485–501
47 Pandey S R, Tran N H, Bennis M, et al. A crowdsourcing framework for on-device federated learning. IEEE Trans Wireless Commun, 2020, 19: 3241–3256
48 Deng J, Guo J, Xue N, et al. ArcFace: additive angular margin loss for deep face recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019. 4690–4699
49 de Boer P T, Kroese D P, Mannor S, et al. A tutorial on the cross-entropy method. Ann Oper Res, 2005, 134: 19–67
50 Jauernig P, Sadeghi A R, Stapf E. Trusted execution environments: properties, applications, and challenges. IEEE Secur Privacy, 2020, 18: 56–60
51 Karimireddy S P, Kale S, Mohri M, et al. SCAFFOLD: stochastic controlled averaging for federated learning. In: Proceedings of International Conference on Machine Learning, 2020. 5132–5143
52 Conitzer V, Sandholm T. Computing Shapley values, manipulating value division schemes, and checking core membership in multi-issue domains. In: Proceedings of the 19th National Conference on Artificial Intelligence, 16th Conference on Innovative Applications of Artificial Intelligence, 2004. 4: 219–225
53 Liu Y R, Ke J M, Jiang H, et al. Improvement of POS consensus mechanism in blockchain based on Shapley value calculation. Comput Res Dev, 2018, 55: 2208–2218 [刘怡然, 柯俊明, 蒋瀚, 等. 基于沙普利值计算的区块链中 PoS 共识机制的改进. 计算机研究与发展, 2018, 55: 2208–2218]
54 Harsanyi J C. Games with incomplete information played by “Bayesian” players part II. Bayesian equilibrium points. Manage Sci, 1968, 14: 320–334
55 Chen L, Xu L, Shah N, et al. On security analysis of proof-of-elapsed-time (PoET). In: Proceedings of International Symposium on Stabilization, Safety, and Security of Distributed Systems. Cham: Springer, 2017. 282–297
56 Fan X, Chai Q. Roll-DPoS: a randomized delegated proof of stake scheme for scalable blockchain-based internet of things systems. In: Proceedings of the 15th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, 2018. 482–484
57 Castro M, Liskov B. Practical Byzantine fault tolerance. In: Proceedings of the 3rd USENIX Symposium on Operating Systems Design and Implementation, 1999. 173–186
58 Deng L. The MNIST database of handwritten digit images for machine learning research. IEEE Signal Process Mag, 2012, 29: 141–142
59 Abouelnaga Y, Ali O S, Rady H, et al. CIFAR-10: KNN-based ensemble of classifiers. In: Prcoeedings of International Conference on Computational Science and Computational Intelligence (CSCI), 2016. 1192–1195
60 Pal S K, Mitra S. Multilayer perceptron, fuzzy sets, and classification. IEEE Trans Neural Netw, 1992, 3: 683–697
61 Vedaldi A, Lenc K. MatConvNet: convolutional neural networks for MATLAB. In: Proceedings of the 23rd ACM International Conference on Multimedia, 2015. 689–692
62 Gao S, Yuan L P, Zhu J M, et al. A blockchain-based privacy-preserving asynchronous federated learning. Sci Sin Inform, 2021, 51: 1755–1774 [高胜, 袁丽萍, 朱建明, 等. 一种基于区块链的隐私保护异步联邦学习. 中国科学: 信息科学, 2021, 51: 1755–1774]

Incentive mechanism for federated learning based on blockchain and Bayesian game

Qinnan ZHANG[1], Jianming ZHU[1], Sheng GAO[1*], Zehui XIONG[2], Qingyang DING[3] & Guirong PIAO[1]

1. *School of Information, Central University of Finance and Economics, Beijing* 100081*, China*;
2. *Pillar of Information Systems Technology and Design, Singapore University of Technology and Design, Singapore* 487372*, Singapore*;
3. *School of Management, Beijing Union University, Beijing* 100020*, China*
* Corresponding author. E-mail: sgao@cufe.edu.cn

Abstract Federated learning (FL) has become a new form of data sharing by aggregating multi-party local models. Although the existing FL incentive system has reduced insufficient data supply under comprehensive information, it still confronts issues including free-riding, unfairness, and unreliability. Therefore, this paper proposes an incomplete information FL incentive mechanism based on blockchain and Bayesian games. The data transaction process is modeled by quantifying the cost-utility of the data providers and the payment reward of the data requesters, in which Shapley value is used to realize the fairness of reward distribution of data providers. We consider the heterogeneity and privacy protection of participating individuals. The data providers' resource allocation strategies are built as a Bayesian game model, which optimizes the local training strategy to realize the incentive effect on the data providers. Furthermore, we consider the effectiveness of the incentive mechanism, a privacy-preserving Bayesian game action strategy consensus algorithm (PPBG-AC) is proposed, which enables the data providers to realize Bayesian Nash equilibrium under a data trading platform based on blockchain. The comparison and analysis of the schemes reveal that the incentive mechanism presented in our paper assures benefit distribution fairness and resource allocation credibility. Simulation experiments and performance evaluations based on real datasets demonstrate the effectiveness of our incentive mechanism.

Keywords federated learning, incentive mechanism, blockchain, Bayesian game, Shapley value, incomplete information, privacy protection