

面向异构数据的半分布式联邦学习安全聚合

黄梅, 王玲玲, 张政胤, 刘宇飞, 孙艺桐

(青岛科技大学信息科学技术学院, 山东 青岛 266042)

摘要: 半分布式联邦学习结合集中式和分布式的优势, 提高了系统的可扩展性和灵活性。然而, 开放异构的联邦学习架构下, 半诚实参与者和异构本地数据的存在加剧了模型聚合过程中节点隐私泄露和模型难收敛的问题。一方面, 节点间复杂的邻接关系增加了本地数据隐私泄露的风险, 节点数据的隐私保护难以得到有效保障。另一方面, 本地模型之间的差异随着本地数据异构程度的增大而增大, 进而减缓了全局模型收敛速度, 降低了全局模型聚合的准确性。为解决上述问题, 提出了一种面向异构数据的半分布式联邦学习安全聚合方案, 在实现节点隐私保护的同时提升全局模型收敛效率。基于交替方向乘子法设计随机掩蔽机制, 在保证模型正确聚合的前提下避免任何实体访问单个节点的本地模型, 增强了本地数据的隐私保护。此外, 提出了一种双重加权聚合策略, 基于全局模型在不同节点样本上的损失设计节点簇内聚合权重, 根据不同簇对全局模型的贡献为其设计全局聚合权重。在3个公共标准数据集上进行的综合实验表明, 与先进的方案相比, 提升了数据非独立同分布时的模型收敛速度和准确率。

关键词: 半分布式联邦学习; 隐私保护; 加权聚合; 数据异构

中图分类号: TP309.2

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2025029

Secure aggregation for semi-decentralized federated learning under heterogeneous data

HUANG Mei, WANG Lingling, ZHANG Zhengyin, LIU Yufei, SUN Yitong

School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266042, China

Abstract: Semi-decentralized federated learning has gained attention for combining the advantages of centralized and decentralized approaches, thereby enhancing system scalability and flexibility. However, in the open heterogeneous federated learning architecture, the existence of semi-honest participants and heterogeneous data has aggra-

收稿日期: 2024-10-31; 修回日期: 2025-03-30

通信作者: 王玲玲, wanglingling@qust.edu.cn

基金项目: 国家自然科学基金(61802217); 山东省自然科学基金(ZR2023MF082); 青岛科技计划重点研发项目(22-3-4-xxgg-10-gx)

Foundation Items: The National Natural Science Foundation of China (61802217); The Natural Science Foundation of Shandong Province (ZR2023MF082); Qingdao Science and Technology Program Key Research and Development Project (22-3-4-xxgg-10-gx)

引用格式: 黄梅, 王玲玲, 张政胤, 等. 面向异构数据的半分布式联邦学习安全聚合[J]. 网络与信息安全学报, 2025, 11(3): 175-189.

Citation Format: HUANG M, WANG L L, ZHANG Z Y, et al. Secure aggregation for semi-decentralized federated learning under heterogeneous data[J]. Chinese Journal of Network and Information Security, 2025, 11(3): 175-189.

vated the problems of node privacy leakage and difficult model convergence during model aggregation. The differences between local models have increased with the degree of local data heterogeneity, which has further slowed down the global model's convergence and reduced its accuracy. Additionally, the complex network topology among nodes has heightened the risk of local data privacy leakage, making it difficult to ensure local data privacy protection. To tackle these issues, a secure aggregation scheme for heterogeneous data under a semi-distributed federated learning architecture was designed. This scheme aimed to enhance the global model's convergence performance while safeguarding node privacy. A random masking mechanism was developed based on the alternating direction multiplier method. This mechanism strengthened local data privacy protection and prevented access to individual node models. Moreover, a double-weighted aggregation strategy was proposed. In this strategy, intra-cluster aggregation weights for nodes were determined based on the loss of the global model on different node samples, and global aggregation weights for different clusters were established according to their contribution to the global model. Extensive experiments were carried out on three public standard datasets. The results demonstrate that, compared with advanced schemes, the proposed method improves model convergence speed and accuracy in the context of heterogeneous data.

Keywords: semi-decentralized federated learning, privacy protection, weighted aggregation, heterogeneous data

0 引言

联邦学习^[1]作为一种分布式机器学习范式,通过数据不出本地的聚合方式实现了用户隐私保护。然而,开放异构的联邦学习场景中,用户数据通常是非独立同分布(non-independent and identically distributed, Non-IID)的,这将导致本地模型差异显著,传统均匀加权聚合无法有效捕捉关键节点贡献,会降低模型准确度。半分布式联邦学习^[2](semi-decentralized federated learning, SDFL)结合设备到设备(device to device, D2D)和设备到服务器(device to server, D2S)的交互方式,提高了模型聚合的可扩展性和灵活性。D2D交互可促进邻居节点间的本地模型达成共识^[2],进而削弱异构数据对全局模型造成的负面影响。然而,参与节点之间的D2D交互受限于节点之间的邻接关系,这不仅增加了本地模型达成共识的难度,还降低了模型收敛速度和泛化性能。此外,这种复杂的交互方式也带来了隐私安全问题,半分布式架构中,节点的D2D通信依赖邻接关系,导致隐私泄露风险随拓扑复杂度增加(如多跳路径增多、攻击面扩大)而增大。例如,稀疏拓扑下节点需要通过多个中间节点传递模型,可能被半诚实节点合谋截获,导致模型训练过程中隐私泄露。

在SDFL架构中,半诚实参与方可能会对其他用户的隐私信息发起推理攻击。为保护节点隐私,同态加密(homomorphic encryption, HE)^[3]

和差分隐私(differential privacy, DP)^[4]等隐私保护技术被广泛应用于联邦学习系统,虽然可在一定程度上保护数据隐私,却与SDFL的复杂拓扑结构存在设计上的冲突。在节点D2D的交互方式下,基于HE对本地模型进行加密的方法大幅增加节点的计算开销以及集群内部节点间的通信负载。基于DP对本地模型添加不可消除的扰动,会导致聚合模型的噪声随着节点间交互轮次的增加而不断累积,加大了全局模型的精度损失。轻量级的随机掩蔽技术虽然可通过隐藏数据实现隐私保护,但其需要与其他客户端交互生成成对的可抵消掩蔽值,与SDFL架构的复杂交互模式难契合。因此如何在保证数据隐私的前提下,设计一种安全聚合方法,是SDFL面临的首要挑战。

簇内的D2D交互使得节点只能与其邻居节点通信,而没有统一的全局模型,这意味着需要经过多轮的交互才能在簇内达成共识,使各节点拥有相同的模型。然而,现有的加权矩阵设计方法^[4-6]仅利用节点之间的邻接关系作为权重度量因子,实际聚合效果等同于为每个节点的本地模型赋予相同的权重。因此,现有方案并未充分考虑本地模型对全局模型的实际贡献,在处理非独立同分布数据集时,难以保障模型聚合的公平性。所以,亟须一种新的高效聚合策略在SDFL架构下实现数据异构的分布式共识,提高模型的收敛性能。

为解决上述问题,本文首先给出了一种基于交替方向乘子法(alternating direction method of

multipliers, ADMM) 的零和随机掩蔽机制。节点根据邻接关系使用 ADMM 为外邻节点更新掩蔽值, 确保每个节点的内邻掩蔽值之和为零。因此, 系统中的所有节点可通过交互得到聚合后的模型, 而不能访问任何单个节点的本地模型, 保护了数据的隐私性。其次, 提出了一种双重加权策略, 根据节点之间的邻接关系和全局模型在不同节点的局部样本上的损失值来设计簇内模型聚合的加权邻接矩阵。在 Non-IID 数据分布下损失大的客户端样本不一定对全局模型贡献小, 若对损失大的客户端本地模型赋予较小的聚合权重, 会使损失较大的本地样本不容易被全局模型学习, 进而阻碍全局模型收敛。因此本文以内邻损失均值作为损失阈值参数, 在阈值范围内, 损失越大, 权重越大, 适当增加损失稍大的本地模型的簇内聚合权重, 使全局模型充分学习此类节点的本地样本。并在全局模型聚合阶段遵循损失越小, 贡献越大的原则, 对不同簇的模型进行加权聚合, 加速全局模型的收敛并提高全局模型的泛化性能。

1 相关工作

1.1 异构数据下的加权聚合

分布式联邦学习通过节点间均匀的分布式聚合, 能够更有效避免单点故障问题, 同时减少网络中出现的拥塞或时延^[7]。然而, 在没有中心节点协调的稀疏 D2D 拓扑中, 模型收敛性能会受到数据异构的显著影响^[8], 簇内聚合的模型不足以适用于所有的本地推理任务。Lin 等^[2]提出 SDFL 架构的两阶段聚合方案, 通过调整双向 D2D 通信轮次和全局聚合周期来实现次线性收敛率。在此基础上, Parasnis 等^[5]基于邻接关系的列随机加权矩阵为 D2D 聚合分配权重。Wang 等^[6]根据双向 D2D 交互拓扑结构生成对称邻接矩阵进行模型加权聚合。Yemini 等^[9]通过邻接矩阵设计聚合权重, 由与服务器有通信链路的客户端进行其内邻模型的加权聚合并上传至服务器。然而这些方案仅根据节点的邻接关系设置聚合权重, 并未考虑 Non-IID 数据下本地模型对全局模型的实际贡献。在极端 Non-IID 场景下, 均匀权重会放大噪声模型的负面影响。Sun 等^[10]根据样本量设计聚合权重, 并通过集群内多次交互本地模型减轻异构数据的影响。仅根据样本量设计静态权重是无法根据本

地模型对全局模型的贡献进行动态调整的。

文献[11]表明在 Non-IID 数据分布下使用类似联邦平均的聚合算法时会降低收敛速度。文献[12-15]基于训练样本量、局部模型和全局模型之间的相似性、节点的贡献或节点设计加权聚合算法加速模型收敛。上述聚合策略侧重减少 Non-IID 数据分布对测试精度和模型收敛性的负面影响^[16], 且权重均由中心服务器基于明文梯度计算, 不支持用户数据的隐私保护。

1.2 隐私保护的模型聚合

现有分布式联邦学习隐私保护方案大多基于 HE^[17-18]和 DP^[19-22]技术。具体而言, Liu 等^[18]利用同态加密和智能合约, 实现加密的梯度聚合和模型更新, 增强了梯度的隐私保护。Chen 等^[19]使用局部差分隐私, 通过 Pair-Wise Gossip 算法实现异步分层聚合, 但需节点间两两协商, 难以在单向 D2D 链路中实现。Gao 等^[22]使用差分隐私技术对梯度参数进行隐私保护, 按照邻接关系设置行随机加权邻接矩阵进行加权聚合。以上工作的局限性在于, 同态加密和差分隐私本身带来的性能开销和计算误差在端到端传播的过程中会被不断加大。

基于一次一密和秘密分享的随机掩蔽机制^[23-25]需要节点间协商出可相互抵消的成对掩码, 不适用于 SDFL 场景多邻居的复杂 D2D 交互模式, 节点难以协商出可以和多个邻居节点相抵消的掩码。Shrestha 等^[26]将每轮的模型聚合结果放到区块链系统中保护隐私, 但忽略了集群内部模型更新的隐私保护。Jeon 等^[27]指出交替方向乘子法的对偶变量在首轮之后具有零和的性质, 利用其作为掩蔽值设计了遵循差距约束的基于组的通信方法, 实现了轻量级的隐私保护, 但需分组之间交互共识变量, 并且任意节点不能连续两轮在同一个分组。与已有工作不同, 本文基于 ADMM 对偶变量零和的性质设计随机掩蔽机制, 消除了节点不能连续在同一分组的限制, 且无须分组之间交互共识变量, 通过设置所有节点内邻对偶变量和为零, 达到保护节点数据隐私的目的。

2 系统设计

2.1 系统模型

本节描述了半分布式联邦学习系统模型, 如图1所示, 符号说明见表1。

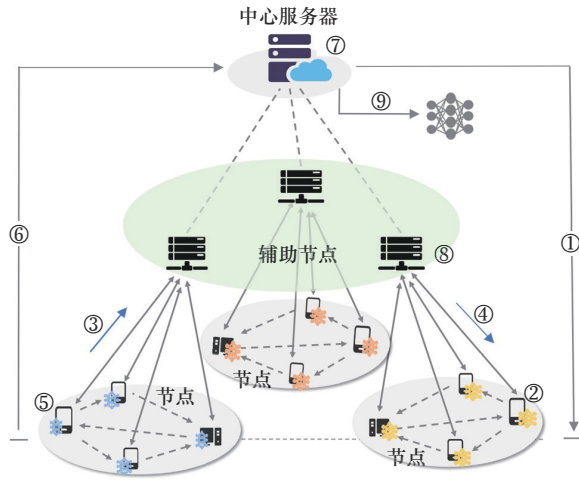


图1 系统模型

Figure 1 System model

表1 符号说明

Table 1 Key notations

符号	含义
C_l	第 l 个簇
AP_l	第 l 个簇的辅助节点
CS	中心服务器
P_i	第 i 个节点
D_i	P_i 所拥有的数据集
N_i^-	P_i 的内邻
N_i^+	P_i 的外邻
θ_i^t	P_i 在第 t 轮的本地模型
θ^t	系统全局模型
x_i	P_i 的本地共识变量
λ_{ij}	P_i 对于 P_j 关于共识问题的对偶变量
ρ	用于控制共识速度的惩罚参数
z_i	P_i 的内邻共识变量均值
$f_i(\theta^t)$	P_i 基于模型 θ^t 的损失
$f_i^{\text{avg}}(\theta^t)$	P_i 内邻损失均值
$[\cdot]$	掩蔽符号
w_{ij}	P_i 对于外邻 P_j 聚合时的权重参数
M_{ik}	P_i 聚合时为内邻 P_k 分配的权重
S	全局聚合采样率
S_l	全局聚合时 C_l 的权重
η	联邦学习的学习率
τ	全局聚合频率

系统中共有 n 个节点，将它们分为 m 个簇 $\{C_1, C_2, \dots, C_m\}$ ，并为每个簇 C_l 分配一个辅助节点

AP_l 。中心服务器 (central server, CS) 负责聚合所有簇的模型以生成全局模型。 C_l 中共有 $|C_l|$ 个节点，节点 P_i 拥有本地数据集 D_i 。

系统中存在设备到服务器和设备到设备两类交互。在设备到服务器交互中，进行全局聚合的节点参与与 CS 的上行链路通信。设备到设备交互网络表示为有向图 $G = (\{n\}, E)$ ，其中 $\{n\}$ 为点集， E 为边集。存在 m 个由参与节点组成的分簇 $\{(C_1, E_1), (C_2, E_2), \dots, (C_m, E_m)\}$ ，任意两个簇之间无通信链路，即 $E = \bigcup_{l=1}^m E_l$ 。在有向图 G 中，若存在从节点 $P_i \in \{n\}$ 到 $P_j \in \{n\}$ 的有向边，则存在从 P_i 到 P_j 的通信链路。因此， P_i 为 P_j 的内邻， P_j 为 P_i 的外邻。 P_i 的内邻集合可表示为 N_i^- ，其外邻集合可表示为 N_i^+ 。

2.2 工作流程

本文方案的工作流程如下：①中心服务器发布学习任务并进行系统初始化；②节点执行本地训练获得本地模型，并据此更新本地共识变量，计算损失并用对偶变量掩蔽损失值和本地共识变量；③节点向辅助节点发送掩蔽的损失值、掩蔽的本地共识变量以及内邻列表；④辅助节点为各节点计算内邻损失均值及内邻共识变量均值；⑤节点更新对偶变量，计算权重参数，向外邻发送模型聚合所需参数，聚合内邻模型并更新本地模型；⑥被采样节点向服务器发送内邻聚合模型；⑦服务器计算簇模型均值；⑧辅助节点计算簇内损失均值；⑨服务器更新全局模型并广播。

2.3 安全模型

系统中的节点 P_i 、辅助节点 AP_l 以及中心服务器均为半诚实实体，即它们诚实地遵循协议，但对参与节点的敏感数据感到好奇，进而被动地发起隐私推理攻击，以获取客户的隐私数据。

2.4 设计目标

本文旨在实现异构数据鲁棒的半分布式联邦学习安全聚合方案，可在保护参与节点数据隐私的基础上，确保全局模型收敛。首先，可防止半诚实参与方通过掩蔽的本地梯度和损失值推断参与节点的隐私数据；其次，保证在数据异构场景下全局模型仍能实现高效聚合。

3 前置知识

3.1 交替方向乘法

交替方向乘法结合了对偶上升法的可分解性和乘子法的稳健性，解决了 $\min f(\mathbf{x}) + g(\mathbf{z})$ 形式的凸优化问题，约束条件为 $\mathbf{Ax} + \mathbf{Bz} = \mathbf{c}$ 。其中 $\mathbf{x} \in \mathbf{R}^n$, $\mathbf{z} \in \mathbf{R}^m$, $\mathbf{A} \in \mathbf{R}^{p \times n}$, $\mathbf{B} \in \mathbf{R}^{p \times m}$, $\mathbf{c} \in \mathbf{R}^p$, $f(\mathbf{x})$ 和 $g(\mathbf{z})$ 是凸函数。上述凸优化问题的增广拉格朗日矩阵可表示为：

$$L_\rho(\mathbf{x}, \mathbf{z}, \boldsymbol{\lambda}) = f(\mathbf{x}) + g(\mathbf{z}) + \boldsymbol{\lambda}^\top (\mathbf{Ax} + \mathbf{Bz} - \mathbf{c}) + \frac{\rho}{2} \|\mathbf{Ax} + \mathbf{Bz} - \mathbf{c}\|_2^2$$

其中， $\boldsymbol{\lambda}$ 为对偶变量， $\rho > 0$ 为惩罚参数。ADMM 的迭代求解过程如下：

$$\begin{aligned} \mathbf{x}^{k+1} &= \arg \min_{\mathbf{x}} L_\rho(\mathbf{x}, \mathbf{z}^k, \boldsymbol{\lambda}^k) \\ \mathbf{z}^{k+1} &= \arg \min_{\mathbf{z}} L_\rho(\mathbf{x}^{k+1}, \mathbf{z}, \boldsymbol{\lambda}^k) \\ \boldsymbol{\lambda}^{k+1} &= \boldsymbol{\lambda}^k + \rho(\mathbf{Ax}^{k+1} + \mathbf{Bz}^{k+1} - \mathbf{c}) \end{aligned}$$

本文利用 ADMM 迭代更新本地共识变量和对偶变量，在保护隐私的同时促进簇内模型共识。

3.2 基于交替方向乘子法的模型共识

簇内节点共同执行 SDFL 任务，节点 P_i 经本地训练得到本地模型 θ_i^t ，簇 C_l 内平均模型可表示为 $\theta_{C_l}^t$ 。簇内分布式训练目标为使簇内所有节点的本地模型达成共识，令 $\mathbf{x}$ 为与 θ^t 同维度的变量，该问题表示为 $\min \sum_{i=1}^{|N_l|} \|\mathbf{x} - \theta_i^t\|_2^2$ ，其最优解

为平均值 $\mathbf{x}^* = \frac{1}{|N_l|} \sum_{k=1}^{|N_l|} \theta_k^t$ 。为了解决这个问题，为 P_i 引入本地变量 $\mathbf{x}_i$ 及共识变量 $\mathbf{z}_i$ ，上式的等效

问题写为 $\min \sum_{i=1}^{|N_l|} \|\mathbf{x}_i - \theta_i^t\|_2^2$ ，目标为 $\mathbf{x}_i = \mathbf{z}_i$ ，其增

广拉格朗日函数可表示为：

$$L_\rho(\mathbf{x}_i, \boldsymbol{\lambda}_{ij}, \mathbf{z}_i) = \sum_{i=1}^{|N_l|} \left(\|\mathbf{x}_i - \theta_i^t\|_2^2 + \boldsymbol{\lambda}_{ij}^\top (\mathbf{x}_i - \mathbf{z}_i) + \frac{\rho}{2} (\|\mathbf{x}_i - \mathbf{z}_i\|_2^2) \right)$$

其中， $\boldsymbol{\lambda}_{ij}$ 为 P_i 对该问题的对偶变量， $\rho \in \mathbf{R}$ 为惩罚参数。遵循标准的 ADMM 算法， $\mathbf{x}_i$ 、 $\mathbf{z}_i$ 、 $\boldsymbol{\lambda}_{ij}$ 更新可形式化为如下形式：

$$\mathbf{x}_i^t = \frac{1}{2 + \rho} \left(2\theta_i^t - \sum_{j \in N_l^+} \boldsymbol{\lambda}_{ij}^{t-1} + \rho \mathbf{z}_i^{t-1} \right) \quad (1)$$

$$\mathbf{z}_i^t = \frac{1}{|C_l|} \sum_{i \in C_l} \mathbf{z}_i^t = \frac{1}{|N_l|} \sum_{k \in N_l^+} \left(\mathbf{x}_k^t + \frac{1}{\rho} \boldsymbol{\lambda}_{ki}^{t-1} \right) \quad (2)$$

$$\boldsymbol{\lambda}_{ij}^t = \boldsymbol{\lambda}_{ij}^{t-1} + \rho(\mathbf{x}_i^t - \mathbf{z}_i^t) \quad (3)$$

受节点之间交互限制，本文使每个节点与其内邻节点达成共识，本地模型趋于内邻模型均值，即 $\mathbf{x}_i^t = \mathbf{z}_i^t$ 。若簇内所有节点均与内邻节点达成共识，则整个簇内部达成共识，即 $\mathbf{z}^t = \mathbf{z}_i^t$ 。

3.3 基于随机掩蔽的隐私保护机制

基于随机掩蔽的隐私保护机制包含 3 步：①节点 P_i 和 P_j 根据密钥协商协议协商出一对可相互抵消的随机掩蔽；②各节点利用协商的随机掩蔽值保护梯度隐私，将密文发送至服务器；③服务器聚合所有节点的掩蔽梯度，抵消随机掩蔽得到正确的全局梯度。本文使用 ADMM 算法的对偶变量生成可相互抵消的随机掩蔽，无须密钥协商。并利用伪随机数生成器（pseudo-random generators, PRG）^[28] 对 ADMM 算法中节点的对偶变量 $\boldsymbol{\lambda}_{ij}^0$ 进行初始化，在每轮交互中使用 $\boldsymbol{\lambda}_{ij}^t$ 作为掩蔽值。

4 具体方案

针对 SDFL 节点的隐私保护，本文设计了基于随机掩蔽的隐私保护机制，保证节点在半诚实安全模型下的隐私安全。利用种子的机密性和伪随机数生成器的伪随机性初始化 ADMM 的对偶变量，并用其掩蔽本地梯度和损失值。对偶变量随 ADMM 算法的迭代而更新，并对模型聚合的相关参数进行新一轮掩蔽，结合半分布式系统的通信限制，确保任何半诚实参与方无法通过中间参数推理节点的本地数据。针对数据异构问题，本文设计了双重加权聚合机制，保证在数据非独立同分布时模型仍可快速收敛。簇内节点根据本地损失动态分配权重进行簇内模型聚合，CS 按照簇规模抽取节点参与全局聚合，并根据簇内损失均值差异化分配全局聚合权重，构建跨簇的全局模型。具体包含以下 4 个步骤。

(1) 系统初始化

CS 发布任务并广播初始化全局模型 θ_0 ，公共参数 $\text{pp} = \{\rho, \tau, \mathcal{S}, \mathbf{z}^0 = \mathbf{0}\}$ 。节点 P_i 生成随机种子 δ_{ij} 并初始化对偶变量 $\boldsymbol{\lambda}_{ij}^0 = \text{PRG}(\delta_{ij})$ 。

(2) 本地模型更新

节点 P_i 根据随机梯度下降算法计算第 t 轮梯度 $\nabla f_i(\theta^t)$, 并更新本地模型为 $\theta_i^{t+1} = \theta_i^t - \eta \nabla f_i(\theta^t)$ 。簇内节点 P_i 只与其邻居节点单向通信, 并聚合内邻模型为 $\bar{\theta}_i^t$ 。

(3) 簇内模型聚合

在第 t 轮, 节点 $P_i \in C_l$, 根据其本地模型 θ_i^t 利用式 (1) 更新本地共识变量 $\mathbf{x}_i^t$, 对偶变量 λ_{ij}^{t-1} 将 $\mathbf{x}_i^t$ 和损失 $f_i(\theta^t)$ 分别掩蔽为 $[\mathbf{x}_i^t] = \mathbf{x}_i^t + \frac{1}{\rho} \lambda_{ij}^{t-1}$ 和 $[f_{ij}(\theta^t)] = f_i(\theta^t) + \frac{1}{\rho} \lambda_{ij}^{t-1}$, 并将 $[f_{ij}(\theta^t)]$, $[\mathbf{x}_i^t]$ 及其内邻列表 $[N_i^-]$ 发送给辅助节点 AP_l 。从而, AP_l 可获知簇内连通情况。 AP_l 根据式 (2) 为 P_i 计算 $[\mathbf{x}_{ki}^t]$ 的均值 $\mathbf{z}_i^t$ 并发送回 P_i , $k \in [N_i^-]$, 计算 P_i 内邻损失均值 $f_i^{\text{avg}}(\theta^t) = \frac{1}{|N_i^-|} \sum_{k \in [N_i^-]} [f_{ki}(\theta^t)]$ 并广播至 C_l 中所有节点。 P_i 利用式 (3) 更新 λ_{ij}^t 。第一轮迭代后, 节点 P_i 的内邻对偶变量之和为 0, 掩蔽互相抵消 (证明过程如定理 1 所示)。节点 P_i 直接获得 P_j 的明文内邻损失均值 $f_j^{\text{avg}}(\theta^t) = [f_j^{\text{avg}}(\theta^t)]$, 据此计算其对 P_j 的簇内聚合权重参数:

$$w_{ij}^t = \begin{cases} \frac{\mathcal{N}}{1 + q(f_j^{\text{avg}}(\theta^t) - f_i(\theta^t))}, & f_i(\theta^t) \leq f_j^{\text{avg}}(\theta^t) \\ \frac{1 - \mathcal{N}}{1 + q(f_i(\theta^t) - f_j^{\text{avg}}(\theta^t))}, & f_i(\theta^t) > f_j^{\text{avg}}(\theta^t) \end{cases}$$

其中, $\mathcal{N} \in [0.5, 1]$, 表示使损失低的节点初始权重更高的超参数; q 为大于 1 的实数, 用于表示增加权重随损失变化的敏感度。 P_i 用 λ_{ij}^t 掩蔽 w_{ij}^t 及 $w_{ij}^t \cdot \theta_i^t$, 得到 $[w_{ij}^t] = w_{ij}^t + \frac{1}{\rho} \lambda_{ij}^t$ 和 $[w_{ij}^t \cdot \theta_i^t] = w_{ij}^t \cdot \theta_i^t + \frac{1}{\rho} \lambda_{ij}^t$, 并将 $[w_{ij}^t]$ 及 $[w_{ij}^t \cdot \theta_i^t]$ 发送至外邻 P_j 。同时, P_i 接收来自其内邻 P_k 的掩蔽信息并聚合内邻模型为:

$$\bar{\theta}_i^t = \frac{\sum_{k \in N_i^-(t)} \left(w_{ki}^t \cdot \theta_k^t + \frac{1}{\rho} \lambda_{ki}^t \right)}{\sum_{k \in N_i^-(t)} \left(w_{ki}^t + \frac{1}{\rho} \lambda_{ki}^t \right)} = \sum_{k \in N_i^-(t)} \mathbf{M}_{ik} \theta_k^t$$

在簇内模型聚合阶段, P_i 的内邻 P_k 对于节

点 P_i 的聚合权重为 $\mathbf{M}_{ik} = \frac{w_{ki}^t}{\sum_{k \in N_i^-(t)} w_{ki}^t}$, 且 P_i 的内邻

权重之和为 1, 即 $\sum_{k \in N_i^-(t)} \mathbf{M}_{ik} = 1$ 。

模型聚合过程为 $\theta^{t+1} = \mathbf{M}(t)\theta^t - \eta \nabla f(\theta^t)$, 其中 η 为学习率, $\theta^t = [\theta_1^t, \theta_2^t, \dots, \theta_n^t]$, $\nabla f(\theta^t) = [\nabla f_1(\theta^t), \nabla f_2(\theta^t), \dots, \nabla f_n(\theta^t)]$, 并且

$$\mathbf{M}(t) = \begin{bmatrix} 0 & M_{12} & \cdots & M_{1n} \\ M_{21} & 0 & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & 0 \end{bmatrix}$$

当 $\text{mod}(t, \tau) \neq 0$ 时, P_i 根据内邻模型聚合值更新其本地模型 $\theta_i^{t+1} = \bar{\theta}_i^t - \eta \nabla f_i(\theta^t)$; 否则, CS 执行全局模型聚合。

(4) 全局模型聚合

簇内为一个 DFL 共识系统, 随着簇内模型聚合轮次的增加, 簇内节点的模型会趋于一致^[2]。因此, 节点不必在每一轮都参与全局聚合, 每 τ 轮簇内模型聚合之后进行一次全局模型聚合即可。CS 根据簇大小按采样率 $\mathcal{S}$ 随机选取 C_l 中的 $|\alpha_l| = \lceil |C_l| \cdot \mathcal{S} \rceil$ 个节点参与全局聚合。被选节点 P_i 上传 $\bar{\theta}_i^t$ 至 CS。CS 计算 C_l 的采样模型均值 $\theta_{C_l}^t = \frac{1}{|\alpha_l|} \sum_{P_i \in \alpha_l} \bar{\theta}_i^t$ 作为 C_l 的模型参数。 AP_l 计算 C_l 的损失均值 $f_{C_l}^{\text{avg}}(t) = \frac{1}{|C_l|} \sum_{P_i \in C_l} f_i^{\text{avg}}(t)$ 并发送给 CS。CS 根据簇内损失均值越小, 对全局模型贡献越大的原则为 C_l 计算权重参数 $S_l = \frac{1}{f_{C_l}^{\text{avg}}(\theta^t)}$, 归一化权重 $\frac{S_l}{\sum_{l=1}^L S_l}$ 并聚合全局模型为

$$\theta^t = \sum_{l=1}^L \frac{S_l}{\sum_{l=1}^L S_l} \theta_{C_l}^t, \text{ 随后广播至所有节点。} P_i \text{ 根据}$$

全局模型更新本地模型为 $\theta_i^{t+1} = \theta^t - \eta \nabla f_i(\theta^t)$ 。

5 理论分析

5.1 正确性

定理 1 基于随机掩蔽的隐私保护机制的正确性依赖于 ADMM 算法对偶变量的零和性质,

即： $\forall t \geq 1, \forall i \in [n], \sum_{k \in [N_i^-]} \lambda_{ki}^t = 0$ 。

证明 在簇内聚合阶段， AP_i 获得由 P_i 的内邻节点 P_k 发送的掩蔽后的参数 $\llbracket \mathbf{x}_{ki}^t \rrbracket = \mathbf{x}_k^t + \frac{1}{\rho} \lambda_{ki}^{t-1}$ 和 $\llbracket f_{ki}(\theta^t) \rrbracket = f_k(\theta^t) + \frac{1}{\rho} \lambda_{ki}^{t-1}$ ，聚合 $\llbracket \mathbf{x}_{ki}^t \rrbracket$ 和 $\llbracket f_{ki}(\theta^t) \rrbracket$ 。以聚合 $\llbracket f_{ki}(\theta^t) \rrbracket$ 为例，聚合后为：

$$\sum_{k \in [N_i^-]} \llbracket f_{ki}(\theta^t) \rrbracket = \left(f_k(\theta^t) + \frac{1}{\rho} \lambda_{ki}^{t-1} \right) = \sum_{k \in [N_i^-]} f_k(\theta^t) + \frac{1}{\rho} \sum_{k \in [N_i^-]} \lambda_{ki}^{t-1}$$

其中，任何 $t > 1$ ，满足：

$$\begin{aligned} \sum_{k \in [N_i^-]} \lambda_{ki}^t &= \sum_{k \in [N_i^-]} \left(\lambda_{ki}^{t-1} + \rho(\mathbf{x}_k^t - \mathbf{z}_k^t) \right) = \\ \lambda_{ki}^{t-1} + \rho \left(\mathbf{x}_k^t - \frac{1}{|[N_i^-]|} \sum_{k \in [N_i^-]} \left(\mathbf{x}_k^t + \frac{1}{\rho} \lambda_{ki}^{t-1} \right) \right) &= \\ \sum_{k \in [N_i^-]} \left(\lambda_{ki}^{t-1} + \rho \mathbf{x}_k^t - \frac{\rho}{|[N_i^-]|} \sum_{k \in [N_i^-]} \left(\mathbf{x}_k^t + \frac{1}{\rho} \lambda_{ki}^{t-1} \right) \right) &= \\ \sum_{k \in [N_i^-]} \left(\lambda_{ki}^{t-1} - \frac{1}{|[N_i^-]|} \sum_{k \in [N_i^-]} \lambda_{ki}^{t-1} \right) &= 0 \end{aligned}$$

所以， $\sum_{k \in [N_i^-]} f_k(\theta^t) = \sum_{k \in [N_i^-]} \llbracket f_{ki}(\theta^t) \rrbracket$ 成立，

同理等式 $\sum_{k \in [N_i^-]} \mathbf{x}_{ki}^t = \sum_{k \in [N_i^-]} \llbracket \mathbf{x}_{ki}^t \rrbracket$ 成立，证明结束。

5.2 收敛性

本文收敛性分析使用了以下基本假设。

假设1 (μ -强凸) 所有节点的本地损失函数 $\{f_i(\theta^t)\}_{i=1}^n$ 都是 μ -强凸的，即存在 $\mu > 0$ ，使得对于所有本地模型 $\theta_1, \theta_2 \in \mathbf{R}^d$ 以及所有节点 $P_i \in [n]$ 满足 $(\nabla f_i(\theta_1) - \nabla f_i(\theta_2))^T(\theta_1 - \theta_2) \geq \mu \|\theta_1 - \theta_2\|^2$ 。

假设2 (β -Lipschitz 平滑) 所有节点的本地损失函数 $\{f_i(\theta^t)\}_{i=1}^n$ 都满足 β -Lipschitz 平滑，即存在一个有限的 β ，使得对于所有本地模型 $\theta_1, \theta_2 \in \mathbf{R}^d$ 及所有客户端节点 $P_i \in [n]$ 都满足 $\|\nabla f_i(\theta_1) - \nabla f_i(\theta_2)\| \leq \beta \|\theta_1 - \theta_2\|^2$ 。

假设3 (局部差异有界) 对于所有参与节点 P_i ，本地损失与全局损失在模型 θ 之间的差异

由 A 和 B 限定，即 $A \|\nabla f(\theta)\| \leq \|\nabla f_i(\theta)\| \leq B \|\nabla f(\theta)\|$ 。

假设3将局部差异性视为一种度量标准，用于量化各参与节点在相同训练配置（如学习率、批大小、训练轮次等）下的数据异构性。在所有节点本地数据样本相同的情况下， $A=B=1$ 。

定理2 行随机加权邻接矩阵 M 的最大奇异值 $\sigma_{\max}$ 满足不等式：

$$0 \leq \sigma_{\max}^2 - 1 \leq \delta_{\text{DS}}(M)$$

对于给定的行随机矩阵 M ，其与双随机矩阵的平均偏差 $\delta_{\text{DS}}(M)$ 是列和与1的平均差值，即

$$\delta_{\text{DS}}(M) = \frac{1}{n} \sum_{j=1}^n \left| \sum_{i=1}^n M_{ij} - 1 \right|$$

行随机矩阵与双随机的最大偏差定义为 $\Delta_{\text{DS}}(M) = \max_{j \in [n]} \left| \sum_{i=1}^n M_{ij} - 1 \right|$ ，

因此，若 M 是双随机矩阵，则 $\delta_{\text{DS}}(M) = 0$ 和 $\Delta_{\text{DS}}(M) = 0$ 。 M 偏离双随机的程度越大，簇内达成共识就越慢， M 的最大奇异值 $\sigma_{\max}$ 反映偏离双随机的程度。

证明 令 M 的谱为 $S(M) = \{\varsigma\}$ ， ς 是 M 的特征值。 $s(M) = \max \{|\varsigma| \mid \varsigma \in S(M)\}$ 为 M 的谱半径。 M 的谱半径是 $\|M\|$ 的下界，即 $s(M) \leq \|M\|$ ，其中 $\|M\| = \sqrt{\lambda_{\max}(M^T M)} = \sigma_{\max}(M)$ 。因此， $\sigma_{\max}(M) = \|M\| \geq s(M) = 1$ ，即 $\sigma_{\max}$ 的下界是1。对于上界，首先假设 M 是行随机矩阵。那么

$$\sigma_1^2 - 1 \leq \max_{j \in [n]} \sum_{i=1}^n (M^T M)_{ij} - 1 =$$

$$\max_{j \in [n]} \sum_{i=1}^n \sum_{k=1}^n M_{kj} M_{ki} - 1 =$$

$$\max_{j \in [n]} \sum_{k=1}^n M_{kj} \left(\sum_{i=1}^n M_{ki} \right) - 1 =$$

$$\max_{j \in [n]} \sum_{k=1}^n M_{kj} - 1 \leq \max_{j \in [n]} \left| \sum_{k=1}^n M_{kj} - 1 \right|$$

定理3 设损失函数 $f_i(\theta^t)$ 满足假设1和假设2，且假设 θ^t 不是一个平稳解，则连续两次迭代之间全局损失函数的期望减少量可以表示为：

$$f(\theta^{t+1}) \leq f(\theta^t) - \eta \mathbb{E}_{i_t} \left[\left(\frac{2 - B\beta\eta}{2B} \right) \|\nabla f_i(\theta^t)\|^2 \right]$$

其中，期望 $\mathbb{E}_{i_t}$ 指的是参与节点 $P_i \in [n]$ 在全局模型聚合中的加权策略。

证明 根据假设 2 中 $f(\theta^t)$ 的 β -Lipschitz 平滑性以及泰勒展开, 可得:

$$f(\theta^{t+1}) \leq f(\theta^t) + \langle \nabla f(\theta^t), f(\theta^{t+1}) - f(\theta^t) \rangle + \frac{\beta}{2} \|f(\theta^{t+1}) - f(\theta^t)\|^2$$

不等式右侧最后两项分别有如下界限:

① $\|f(\theta^{t+1}) - f(\theta^t)\|^2$ 有界: 据 θ^{t+1} 的聚合定义知 $\|f(\theta^{t+1}) - f(\theta^t)\| = \mathbb{E}_{i|t}[\|f_i(\theta^{t+1}) - f_i(\theta^t)\|]$

通过应用 SGD 优化, 上述方程式右侧期望中的每一项可以表示为:

$$f_i(\theta^{t+1}) = f(\theta^t) - \eta \nabla f_i(\theta^t)$$

因此, 根据柯西-施瓦茨不等式可得:

$$\begin{aligned} \|f(\theta^{t+1}) - f(\theta^t)\|^2 &= (\mathbb{E}_{i|t}[\|f_i(\theta^{t+1}) - f_i(\theta^t)\|])^2 = \\ &\eta^2 (\mathbb{E}_{i|t}[\|\nabla f_i(\theta^t)\|])^2 \leq \eta^2 (\mathbb{E}_{i|t}[\|\nabla f_i(\theta^t)\|^2]) \end{aligned}$$

② $\langle \nabla f(\theta^t), f(\theta^{t+1}) - f(\theta^t) \rangle$ 有界: 根据 θ^{t+1} 的全局聚合定义可得:

$$\begin{aligned} \langle \nabla f(\theta^t), f(\theta^{t+1}) - f(\theta^t) \rangle &= \\ -\eta \mathbb{E}_{i|t}[\langle \nabla f(\theta^t), \nabla f_i(\theta^t) \rangle] \end{aligned}$$

上述期望项可表示为:

$$\begin{aligned} \mathbb{E}_{i|t}[\langle \nabla f(\theta^t), \nabla f_i(\theta^t) \rangle] &= \\ \mathbb{E}_{i|t} \left[\frac{\langle \nabla f(\theta^t), \nabla f_i(\theta^t) \rangle}{\|\nabla f(\theta^t)\| \cdot \|\nabla f_i(\theta^t)\|} \cdot \|\nabla f(\theta^t)\| \cdot \|\nabla f_i(\theta^t)\| \right] &\geq \\ \mathbb{E}_{i|t} \left[\frac{\langle \nabla f(\theta^t), \nabla f_i(\theta^t) \rangle}{\|\nabla f(\theta^t)\| \cdot \|\nabla f_i(\theta^t)\|} \cdot \frac{\|\nabla f_i(\theta^t)\|^2}{B} \right] \end{aligned}$$

假设 3 使得上述不等式成立, 局部差异性上界为 B 。结合①和②可得:

$$\begin{aligned} \langle \nabla f(\theta^t), f(\theta^{t+1}) - f(\theta^t) \rangle + \frac{\beta}{2} \|f(\theta^{t+1}) - f(\theta^t)\|^2 &\leq \\ -\eta \mathbb{E}_{i|t} \left[\frac{\langle \nabla f(\theta^t), \nabla f_i(\theta^t) \rangle}{\|\nabla f(\theta^t)\| \cdot \|\nabla f_i(\theta^t)\|} \cdot \frac{\|\nabla f_i(\theta^t)\|^2}{B} \right] + \\ \frac{\beta \eta^2}{2} \mathbb{E}_{i|t}[\|\nabla f_i(\theta^t)\|^2] &\leq \\ -\eta \mathbb{E}_{i|t} \left[\frac{\|\nabla f_i(\theta^t)\|^2}{B} \right] + \frac{\beta \eta^2}{2} \mathbb{E}_{i|t}[\|\nabla f_i(\theta^t)\|^2] &= \\ -\eta \mathbb{E}_{i|t} \left[\left(\frac{2 - B\beta\eta}{2B} \right) \|\nabla f_i(\theta^t)\|^2 \right], &\text{证明结束。} \end{aligned}$$

定理 4 节点 P_i 的内邻共识变量以线性速率

$\frac{\rho}{2+\rho}$ 收敛至共识。

证明 已知共识变量的最优解为 z_i^*

$\frac{1}{|N_i^-|} \sum_{k \in N_i^-} \theta_k^t$, 定义第 t 轮中共识变量与最优解的差

异为 $\Delta z_i^t = z_i^t - z_i^*$ 。

$$\Delta z_i^t = z_i^t - z_i^* =$$

$$\frac{1}{|N_i^-|} \sum_{k \in N_i^-} \left(x_k^t + \frac{1}{\rho} \lambda_{ki}^{t-1} \right) - z^* =$$

$$\frac{1}{|N_i^-|} \sum_{k \in N_i^-} x_k^t - \frac{1}{|N_i^-|} \sum_{k \in N_i^-} \theta_k^t =$$

$$\frac{1}{|N_i^-|} \sum_{k \in N_i^-} \left(\frac{1}{2+\rho} \left(2\theta_k^t - \sum_{i \in N_k^+} \lambda_{ki}^{t-1} + \rho z_k^{t-1} \right) - \theta_k^t \right) =$$

$$\frac{1}{|N_i^-|} \sum_{k \in N_i^-} \left(\frac{\rho}{2+\rho} z_k^{t-1} - \frac{\rho}{2+\rho} \theta_k^t \right) =$$

$$\frac{\rho}{2+\rho} \left(z_i^{t-1} - \frac{1}{|N_i^-|} \sum_{k \in N_i^-} \theta_k^t \right) =$$

$$\frac{\rho}{2+\rho} (z_i^{t-1} - z_i^*) \Rightarrow \|\Delta z_i^t\| \leq \|\Delta z_i^{t-1}\|$$

其中, $\|\Delta z_i^t\|$ 表示 z_i^t 与 z_i^* 的差距, 由于 $\frac{\rho}{2+\rho} < 1$,

因此节点的内邻共识变量实现了线性收敛。

5.3 隐私性

中心服务器进行全局聚合时只能收到节点的簇内聚合均值, 无法对单个本地模型发起隐私推理攻击。在本文中半诚实参与方的恶意行为主要包括参与节点多轮交互利用累积信息推理, 以及辅助节点与参与节点合谋发起推理攻击两类。

定理 5 由于 PRG 的安全性以及 ADMM 的零和约束, 半诚实的实体无法通过多轮交互推理出节点的任何数据隐私。

证明 存在一个概率多项式时间攻击者 $\mathcal{A}$, 在联邦学习的多轮交互中 (如 T 轮), 逐步收集

以下关于节点 P_i 的信息: ① $\llbracket w_{ij}^t \rrbracket = w_{ij}^t + \frac{1}{\rho} \lambda_{ij}^t$,

② $\llbracket w_{ij}^t \cdot \theta_i^t \rrbracket = w_{ij}^t \cdot \theta_i^t + \frac{1}{\rho} \lambda_{ij}^t$, 试图推断目标节点

P_i 的隐私信息 θ_i^t , 其中 $\lambda_{ij}^t = \lambda_{ij}^{t-1} + \rho(x_i^t - z_i^t)$ 。

假设攻击者 $\mathcal{A}$ 能够通过多轮交互信息, 以不可忽略的概率 $1 - \text{negl}(\kappa)$ 成功恢复 P_i 的隐私信

息, 即 $\Pr \left[\mathcal{A} \left(\left\{ \left[w_{ij}^t \right], \left[w_{ij}^t \cdot \theta_i^t \right], \lambda_{ij}^t \right\}_{t=1}^T = \theta_i^t \right) \right] \geq 1 - \text{negl}(\kappa)$, 其中 κ 为安全参数。攻击者 $\mathcal{A}$ 成功恢复 P_i 的隐私信息需要满足3个假设: ①区分真实掩蔽模型 $\left[w_{ij}^t \cdot \theta_i^t \right]$ 和随机噪声; ②推导出PRG的随机种子 δ_{ij} ; ③成功绕过ADMM的零和约束。

若攻击者能区分 $\left[w_{ij}^t \cdot \theta_i^t \right]$ 和相同轮数的随机数 PRG(δ_{sim}^t), 则存在区分器 $\mathcal{D}$ 使得 $|\Pr \left[\mathcal{D} \left(\left[w_{ij}^t \right] \right) = 1 \right] - \Pr \left[\mathcal{D} \left(\text{PRG}(\delta_{\text{sim}}^t) \right) = 1 \right]| \geq 1 - \text{negl}(\kappa)$ 成立, 直接违背PRG的伪随机性, 因此假设①不成立。首轮掩蔽值 λ_{ij}^0 由PRG生成, 若攻击者能从首轮掩蔽值 λ_{ij}^0 推断出种子 δ_{ij} , 则直接违背PRG的前向安全性, 与PRG定义相矛盾, 因此假设②不成立。

若攻击者能直接从掩蔽向量 $\left[w_{ij}^t \cdot \theta_i^t \right]$ 及 $\left[w_{ij}^t \right]$ 中推断出第 t 轮掩蔽值 λ_{ij}^t , 则需要获取 z_i^t 以及 x_i^t , 与方案交互限制相矛盾。此外每轮ADMM对偶变量由 P_i 自行更新, 且满足零和约束 $\sum_{i \in [N_j^-]} \lambda_{ij}^t = 0$, P_i 的外邻 P_j 的内邻数为 $|N_j^-|$ 。若攻击者试图通过累计 T 轮零和方程 $\sum_{i \in [N_j^-]} \lambda_{ij}^t = 0$ 解出 λ_{ij}^t , 则需求解 T 个包含 $|N_j^-|$ 个变量的方程组, 但仅有一个约束条件, 因此解空间维度为 $|N_j^-| - 1$, 存在无限多解, 与绕过ADMM的零和约束矛盾, 因此假设③不成立。

综上, 在PRG安全且ADMM零和约束成立的条件下, 不存在有效的多轮交互隐私推理攻击。

定理6 由于拓扑结构的单向性及PRG的随机性, 半诚实的实体无法通过合谋攻击推理出任何节点的隐私数据。

证明 假设辅助节点AP联合 P_i 的所有外邻 $[N_i^+]$ 通过相互共享信息 $\left\{ \left[w_{ij}^t \right], \left[w_{ij}^t \cdot \theta_i^t \right], \lambda_{ij}^t \right\}_{i=1}^T$ 推理 P_i 的隐私信息 θ_i^t , 其中 $P_j \in [N_i^+]$ 。

由于拓扑结构的单向性, P_i 发送给其每个外邻的参数均由对偶变量 λ_{ij}^t 进行掩蔽, 并且首轮对偶变量由PRG根据不同的种子生成, 导致发送给不同的外邻每一轮对偶变量之间无关联。此外, AP联合 $P_j \in [N_i^+]$ 推理 θ_i^t 时, 由于遵循

PRG的随机性, 被掩蔽的参数与随机数在计算上无法区分。因此, 合谋的参与方之间不能为彼此提供有效信息, 不存在有效的合谋攻击。

6 性能评估

6.1 实验设置

1) 硬件环境: 实验在8核AMD处理器、16 GB内存和40 GB磁盘空间的系统上进行, 使用一台配置为RTX2080Ti GPU、62 GB内存和100 GB磁盘空间的云服务器进行基于卷积神经网络 (convolutional neural network, CNN) 模型的联邦学习任务。

2) 数据集: 采用MNIST、Fashion-MNIST和CIFAR-10基准数据集, 这3个数据集均包含60 000张训练图像和10 000张测试图像。MNIST数据集是一个包含10类手写数字图像的分类数据集, Fashion-MNIST数据集由70 000张服饰产品图像组成, 共分为10类。CIFAR-10数据集包含60 000张32×32像素的彩色图像, 分为10类。相比于MNIST和Fashion-MNIST, CIFAR-10的图像尺寸较小, 颜色更为复杂, 每个像素由3个通道 (红、绿、蓝) 组成, 因此更具真实性和挑战性。

3) 参数设置: 本文将50个节点随机分为3个簇, 全局模型聚合频率 $\tau = 2$ 。使用迪利克雷分布对数据集进行非独立同分布划分, 当迪利克雷分布参数 $\alpha=1$ 时, 客户端之间的数据分布为独立同分布; 当 $\alpha < 1$ 时, 表现为非独立同分布, α 越小非独立同分布的程度越强。鉴于本文聚焦于数据隐私保护与加权聚合策略, 而非CNN模型架构的优化, 因此未对CNN的超参数配置进行调整。针对MNIST和Fashion-MNIST数据集, 使用具有两个5×5的卷积层和两个2×2的最大池化层的CNN架构, 卷积层分别为32通道和64通道。而对于CIFAR-10数据集, 使用更复杂的CNN模型进行训练, 该模型包含4个5×5卷积层。

4) 对比方案: TT-HF^[2] (two timescale hybrid federated) 是两阶段聚合的SDFL方案, 其中簇内D2D交互是双向通信, 基于节点间的邻接关系度量簇内聚合权重。Parasnis等^[5]提出簇内单向D2D交互的SDFL方案, 基于节点的邻接关系建立列随机矩阵进行簇内加权聚合。Zhao等^[15]提出的DRFL (dynamic reweighting strategy for fair federated learning) 是客户端服务器架构, 其

中根据客户端损失值计算聚合权重, 损失越低则权重越高。

6.2 收敛性能

本节在不同的 Non-IID 程度的数据上评估模型收敛速率和准确率, 并与其他 3 个方案进行对比。如图 2~4 所示, 随着数据 Non-IID 程度的增长, 4 种方案的收敛性能均有所降低, 但本文方案仍保持较好的收敛性能。如图 2(d) 所示, 当 $\alpha=0.4$ 时, 为使模型准确率达到 0.8, 本文方案、TT-HF、Parasnis 及 DRFL 方案分别需要训练 20 轮、40 轮、44 轮和 30 轮。

在复杂程度更高的 Fashion-MNIST 和 CIFAR-10 数据集下, 本文方案依然能以较快的速率收敛并保持较高的准确度。如图 3(d) 所示, 本文方案、TT-HF、Parasnis 及 DRFL 方案分别需要 24 轮、38 轮、48 轮和 30 轮模型准确率才能达到 0.6。如图 4(d) 所示, 经过 30 轮训练, 本文方案、TT-HF、Parasnis 及 DRFL 方案的准确率分别为 0.68、0.55、0.51 和 0.63。

TT-HF 和 Parasnis 方案具有几乎相同的收敛趋势, 其原因在于它们仅在通信形式上有所不同, 本质上都基于邻接关系设置权重, 并没有考虑不同本地模型对全局模型的各种贡献。相反, 本文方案不仅考虑了邻接关系, 还基于节点模型的损失值自适应地更新了簇内聚合权重, 并且在簇间聚合时根据各簇的贡献对不同簇进行加权。

6.3 计算开销

本文方案的计算开销包括系统初始化、簇内模型聚合和全局模型聚合阶段的计算开销。在初始化阶段, 仅需生成一些初始参数, 其计算开销可忽略不计。

首先分析不同条件下簇内模型聚合计算开销。如图 5 所示, 实验评估将相同数量的节点划分为不同簇规模, 对簇内模型聚合计算开销和单个节点计算开销的影响。在此实验中, 平均每个节点拥有的样本量为 500 个。对于图 5(a), 当不同数量的节点被划分为相同数量的簇时, 节点总数越多,

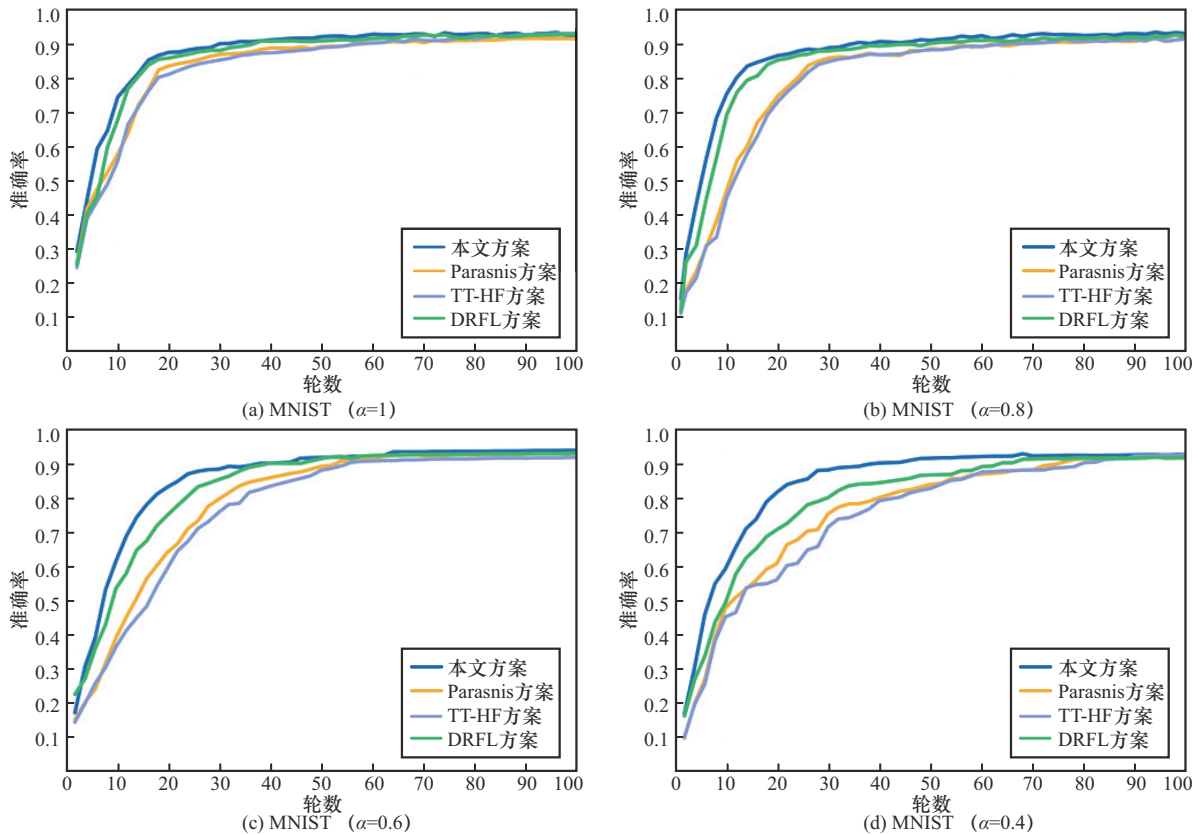


图2 不同 Non-IID 程度的 MNIST 数据集上的收敛性能

Figure 2 Comparison of the test accuracy on MNIST datasets under different Non-IID degrees

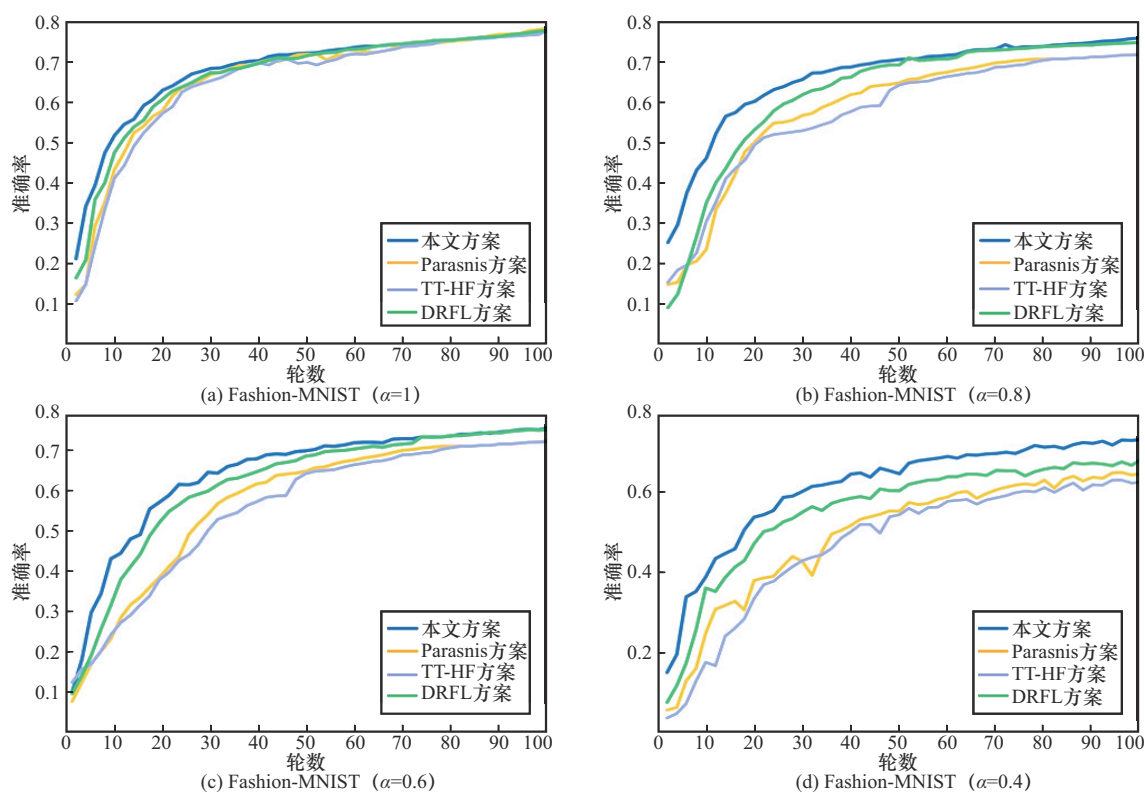


图3 不同Non-IID程度的Fashion-MNIST数据集上的收敛性能

Figure 3 Comparison of the test accuracy on Fashion-MNIST datasets under different Non-IID degrees

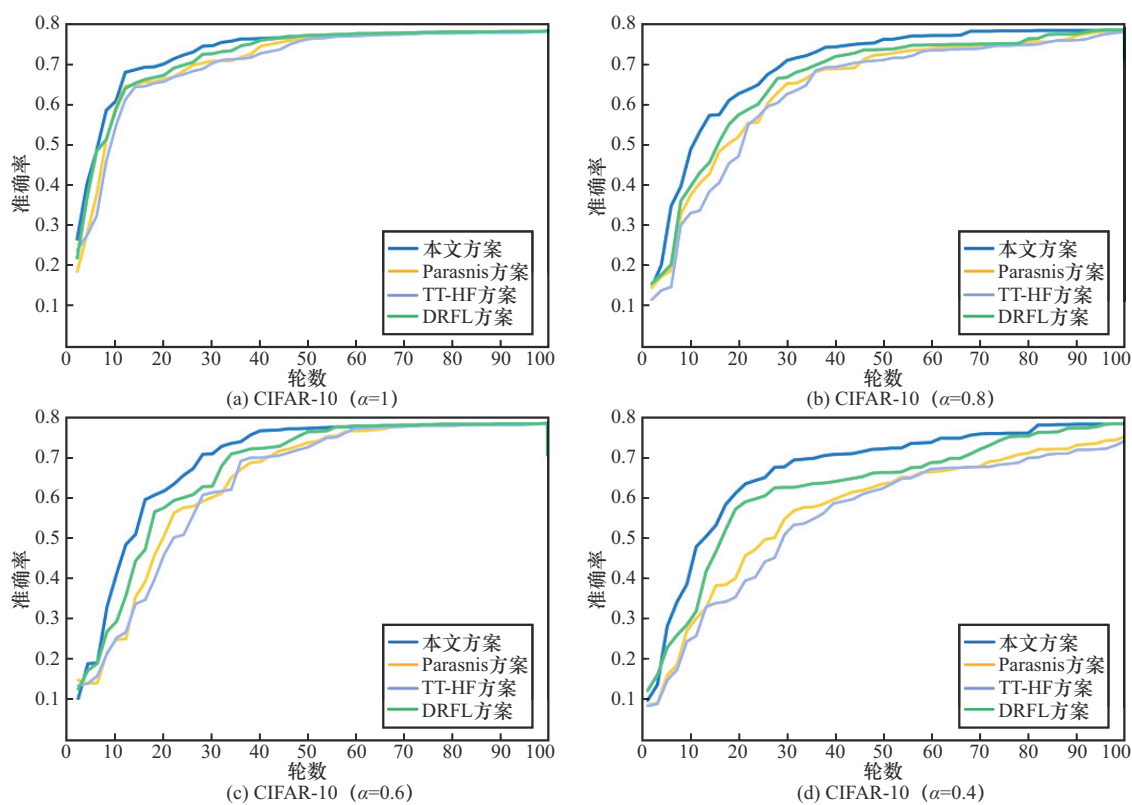
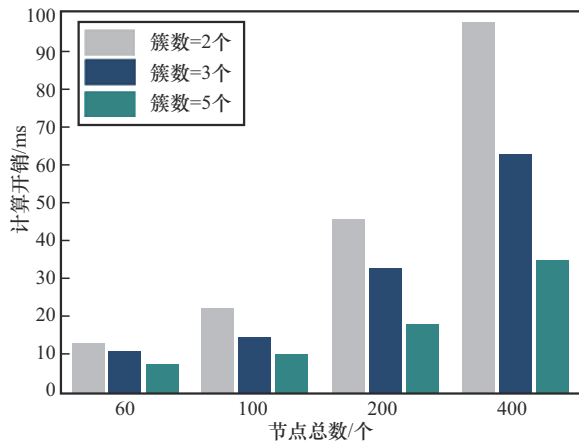


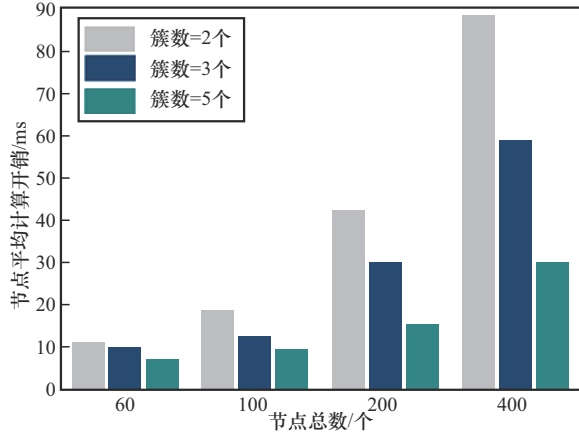
图4 不同Non-IID程度的CIFAR-10数据集上的收敛性能

Figure 4 Comparison of the test accuracy on CIFAR-10 datasets under different Non-IID degrees

簇的规模越大, 簇内模型聚合的计算开销也越高。例如, 当簇数为3个时, 当节点总数分别为60个、100个、200个和400个时, 簇内聚合的计算开销分别为 10.669 ms、14.439 ms、32.683 ms 和 63.007 ms。将相同数量的节点划分为不同数量的簇时, 簇数越多, 其规模越小, 簇内模型聚合的计算开销也越小。此外, 由图 5(b)可知, 节点平均计算开销的变化规律与簇内模型聚合计算开销一致, 但略低于簇内模型聚合的开销。因为在簇内模型聚合时, 所有节点同时进行簇内模型聚合操作, 此外辅助节点还需要计算损失均值。



(a) 簇内模型聚合计算开销



(b) 节点平均计算开销

图 5 不同节点总数的簇内模型聚合的计算开销

Figure 5 Computational cost of intra-cluster aggregation model for different total number of nodes

如图 6 所示, 本文评估了在相同簇规模下, 节点被分配不同数量的训练样本对簇内模型聚合计算开销的影响。当固定节点的平均样本量时, 簇内模型聚合的计算开销随簇规模的增大而增加。

例如, 当节点平均样本量为 500 个时, 簇内模型聚合的计算开销在簇规模为 30 个节点、50 个节点、70 个节点和 90 个节点时分别为 9.044 ms、12.408 ms、21.302 ms 和 23.302 ms。这是簇规模增大导致节点需要聚合更多的内邻模型, 且辅助节点需要为更多节点计算损失均值等参数。而当簇规模固定时, 簇内模型聚合的计算开销不随节点平均样本量的增加而改变。因为样本量大小不影响簇内模型聚合的计算过程, 只会对本地训练产生影响。

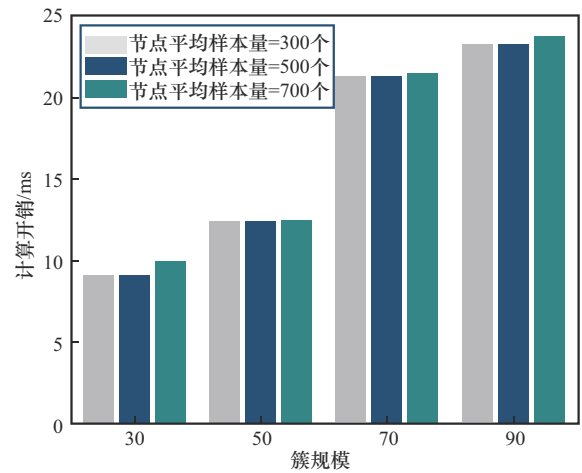


图 6 同样本量的节点在不同簇规模的计算开销

Figure 6 Computational cost for the same cluster size at different average node data sizes

跨簇的全局模型聚合发生在不同的簇之间, 并受到簇的大小、簇的数量和采样率的影响。图 7 评估了不同的簇规模和采样率下的计算开销。

由图 7(a)可知 (假设总共有 3 个簇), 计算开销不仅随着采样率的增加而增加, 还随着簇规模的增大而增加。例如, 当采样率为 0.2 时, 全局聚合的计算开销在簇规模为 30 个节点、50 个节点、70 个节点和 90 个节点时分别为 17.617 ms、28.436 ms、38.559 ms 和 49.466 ms。这是因为簇越大或采样率越高, 服务器将聚合更多的本地模型。由 7(b)可知, 当平均簇大小固定为 30 时, 全局模型聚合的计算开销也随着簇数量的增加而增加。因为服务器需要为更多的簇分配聚合权重, 且聚合的本地模型数量随着簇数量的增加而增加。此外, 如表 2 所示, 本文方案的计算开销比基于 Paillier 同态加密方案低一个数量级。

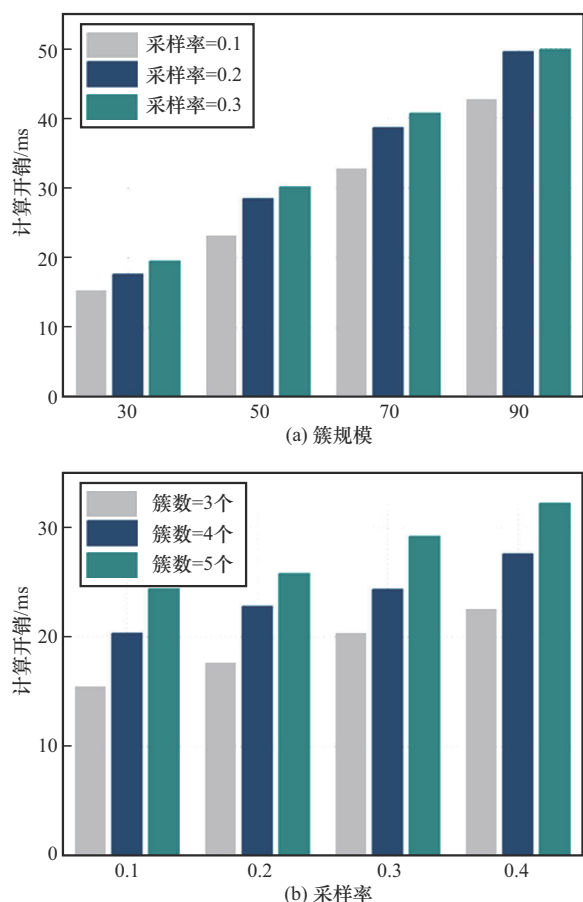


图7 不同采样率在不同簇规模下的计算开销

Figure 7 Comparison of computational cost for the same cluster size at different average node data sizes

表2 不同模型参数下的计算开销

Table 2 Computation costs for different model sizes

模型参数	基于 Paillier 同态加密方案/s	本文方案/ms
10 000	19.6	13.7
30 000	58.8	42.3
50 000	98.1	59.4
70 000	137.2	67.4

6.4 通信开销

由于本文方案主要通信开销为模型参数交互，因此评估不同模型参数大小对簇内模型聚合阶段及全局聚合阶段的通信开销的影响，具体见表3。

本文方案通信开销主要产生在簇内模型聚合阶段及全局聚合阶段，在5个大小为30的簇，整个簇共45条通信链路且全局聚合采样率为0.2的条件下评估了本文方案在不同模型参数下的通信开销。在簇内模型聚合阶段主要的通信开销为

节点之间交互掩蔽的模型参数，及节点发送与模型参数大小几乎相同的本地变量 $\llbracket x_{ij}^t \rrbracket$ 至辅助节点。在采样率为0.2的全局聚合阶段，每个簇仅有6个节点需向服务器上传内邻聚合模型。其他参数交互造成的通信开销小于10 kB，可忽略不计，因此本文方案主要的通信开销正比于模型参数大小。基于Paillier同态加密方案的通信开销是本文方案的约2.5倍。

表3 不同模型参数下的通信开销

Table 3 Communication overhead for different model sizes

模型参数	基于 Paillier 同态加密方案/MB	本文方案/MB
10 000	10.103	3.949
30 000	27.782	11.746
50 000	48.963	19.602
70 000	77.243	27.399

7 结束语

本文提出了一种面向异构数据的半分布式联邦学习安全聚合方案，利用ADMM对偶变量的零和性质实现了本地数据的隐私保护。基于全局模型在不同节点样本上的损失以及不同簇对全局模型的贡献，设计的加权聚合策略提高了全局模型在非独立同分布数据集下的收敛率和准确率。虽然本文方案实现了隐私保护，但需为不同的外邻计算不同的对偶变量，且需要交互大小与本地模型等同的本地共识变量，引入了额外的通信开销。因此，未来的研究工作将着重于减少对偶变量的数量及优化通信开销，以期达成更高的效率。另外，鉴于当前隐私保护策略仅限于固定拓扑结构，设计一种适用于动态拓扑结构的轻量级隐私保护机制，也是未来工作的研究重点。

参考文献：

- [1] 周传鑫,孙奕,汪德刚,等. 联邦学习研究综述[J]. 网络与信息安全学报,2021,7(5):77-92.
ZHOU C X, SUN Y, WANG D G, et al. A survey of federated learning research[J]. Chinese Journal of Network and Information Security, 2021, 7(5): 77-92.
- [2] LIN F, HOSSEINALIPOUR S, AZAM S, et al. Semi-decentralized federated learning with cooperative D2D local model aggregations[J]. IEEE Journal on Selected Areas in Communications, 2021, 39(12):

- 3851-3869.
- [3] SATHYA S S, VEPAKOMMA P, RASKAR R, et al. A review of homomorphic encryption libraries for secure computation[C]// CoRR. 2018, 1-12.
 - [4] BOYLE E, GILBOA N, ISHAI Y, et al. Function secret sharing: Improvements and extensions[C]//ACM Conference on Computer and Communications. 2016: 1292-1303.
 - [5] PARASNIS R, HOSSEINALIPOUR S, CHU Y, et al. Energy-efficient connectivity-aware learning over time-varying D2D networks[J]. IEEE Journal of Selected Topics in Signal Processing, 2024, 18(2): 242-258.
 - [6] WANG B, FANG J, LI H, et al. Confederated learning: federated learning with decentralized edge servers[J]. IEEE Transactions on Signal Processing, 2023, 71: 248-263.
 - [7] HARD K, PARTRIDGE, MATHEWS R, et al. Jointly learning from decentralized (federated) and centralized data to mitigate distribution shift[C]//NeurIPS Workshop on Distribution Shifts: Connecting Methods and Applications. 2021: 1 - 8.
 - [8] BELLET A, KERMARREC A M, LAVOIE E. D-cliques: compensating for data heterogeneity with topology in decentralized federated learning[C]//International Symposium on Reliable Distributed Systems, 2022, 1-11.
 - [9] YEMINI M, SAHA R, OZFATURA E, et al. Semi-decentralized federated learning with collaborative relaying[C]//2022 IEEE International Symposium on Information Theory. Piscataway: IEEE Press, 2022: 1471-1476.
 - [10] SUN Y, SHAO J, MAO Y, et al. Semi-decentralized federated edge learning with data and device heterogeneity[J]. IEEE Transactions on Network and Service Management, 2023, 20(2): 1487-1501.
 - [11] LI X, HUANG K, YANG W, et al. On the convergence of FedAvg on non-IID data[C]//International Conference on Learning Representations. 2020: 1-26.
 - [12] CHEN S, SHEN C, ZHANG L, et al. Dynamic aggregation for heterogeneous quantization in federated learning[J]. IEEE Transactions on Wireless Communications, 2021, 20(10): 6804-6819.
 - [13] WANG J, LIU Q, LIANG H, et al. A novel framework for the analysis and design of heterogeneous federated learning[J]. IEEE Transactions on Signal Processing, 2021, 69: 5234-5249.
 - [14] CAO X, FANG X, LIU J, et al. FLTrust: Byzantine-robust federated learning via trust bootstrapping[C]//Network and Distributed System Security Symposium. 2021.
 - [15] ZHAO Z, JOSHI G, A dynamic reweighting strategy for fair federated learning[C]//ICASSP IEEE International Conference on Acoustics, Speech and Signal Processing. 2022: 8772-8776.
 - [16] MA Q, XU Y, XU H, et al. FedSA: A semi-asynchronous federated learning mechanism in heterogeneous edge computing[J]. IEEE Journal on Selected Areas in Communications, 2021, 39(12): 3654-3672.
 - [17] ZHAO J, ZHU H, WANG F, et al. PVD-FL: A privacy-preserving and verifiable decentralized federated learning framework[J]. IEEE Transactions on Information Forensics and Security, 2022, 17: 2059-2073.
 - [18] LIU J, LI Y, ZHAO M, et al. EPFFL: enhancing privacy and fairness in federated learning for distributed e-healthcare data sharing services[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 1-14.
 - [19] CHEN Q, WANG Z, WANG H, et al. FedDual: Pair-wise gossip helps federated learning in large decentralized networks[J]. IEEE Transactions on Information Forensics and Security, 2022, 18: 335-350.
 - [20] XU M, ZOU Z, CHENG Y, et al. SPDL: Blockchain-secured and privacy-preserving decentralized learning[J]. IEEE Transactions on Computers, 2023, 72(2): 548-558.
 - [21] CHEN S, WANG Y, YU D, et al. Privacy enhanced decentralized federated learning at dynamic edge[J]. IEEE Transactions on Computers, 2023, 72(8): 2165-2180.
 - [22] GAO W, ZHAO Z, WEI M, et al. Decentralized online bandit federated learning over unbalanced directed networks[J]. IEEE Transactions on Network Science and Engineering, 2024, 1-14.
 - [23] BELL J H, BONAWITZ K A, GASCÓN A, et al. Secure single-server aggregation with (poly) logarithmic overhead[C]//ACM Conference on Computer and Communications. 2020: 1253-1269.
 - [24] 郑诚波, 闫皓楠, 傅彩利, 等. 面向智能物联网的双层级联邦安全学习架构[J]. 网络与信息安全学报, 2024, 10(6): 71-80.
 - ZHENG C B, YAN H N, FU C L, et al. Double layer federated security learning architecture for artificial intelligence of things[J]. Chinese Journal of Network and Information Security, 2024, 10(6): 71-80.
 - [25] 董业, 侯炜, 陈小军, 等. 基于秘密分享和梯度选择的高效安全联邦学习[J]. 计算机研究与发展, 2020, 57(10): 2241-2250.
 - DONG Y, HOU W, CHEN X J, et al. Efficient and secure federated learning based on secret sharing and gradients selection[J]. Journal of Computer Research and Development, 2020, 57(10): 2241 -2250.
 - [26] SHRESTHA A K, KHAN F A, SHAIKH M A, et al. Enhancing scalability and reliability in semi-decentralized federated learning with blockchain: Trust penalization and asynchronous functionality[C]//2023 IEEE 14th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). Piscataway: IEEE Press, 2023: 0230-0236.
 - [27] JEON B, FERDOUS S M, RAHMAN M R, et al. Privacy-preserving decentralized aggregation for federated learning[C]//IEEE Conference on Computer Communications Workshops. Piscataway: IEEE Press, 2021: 1-6.
 - [28] BLUM M, MICALI S. How to generate cryptographically strong sequences of pseudo-random bits[C]//Providing Sound Foundations for Cryptography. 2019: 227-240.

[作者简介]



黄梅 (1998-), 女, 陕西安康人, 青岛科技大学硕士生, 主要研究方向为联邦学习隐私保护。



王玲玲（1982-），女，四川巴中人，博士，青岛科技大学副教授、硕士生导师，主要研究方向为应用密码学及联邦学习隐私安全。



刘宇飞（2001-），女，山东聊城人，青岛科技大学硕士生，主要研究方向为联邦学习隐私保护。



张政胤（2000-），男，山东枣庄人，青岛科技大学硕士生，主要研究方向为联邦学习隐私保护。



孙艺桐（2000-），女，山东日照人，青岛科技大学硕士生，主要研究方向为联邦学习隐私保护。