

算力网络环境下基于区块链的自适应联邦学习

刘天瑞^{1,2}, 王连海^{1,2}, 王启正^{1,2}, 徐淑奖^{1,2}, 张淑慧^{1,2}, 王英晓春^{1,2}

(1. 齐鲁工业大学(山东省科学院) 山东省计算中心(国家超级计算济南中心);

算力互联网与信息安全教育部重点实验室, 山东 济南 250014;

2. 山东省计算机网重点实验室, 山东省基础科学研究中心(计算机科学), 山东 济南 250014)

摘要: 算力网络旨在深度融合算力资源与网络资源, 实现对泛在离散部署的海量数据和异构资源的高效利用。为了应对算力网络中复杂的多中心计算协同和数据隐私安全的要求, 联邦学习在隐私保护和边缘算力资源利用等方面有着先天的优势。然而, 基于算力边缘服务器的高度异构和广泛分布, 算力网络环境下的联邦学习仍然存在一定的困难。一方面, 算力网络中海量边缘服务器之间存在的数据异质性会引起非独立同分布问题, 导致联邦学习中局部模型的更新偏离全局最优值。另一方面, 由于不同边缘服务器的数据质量存在差异, 生成低质量的局部模型会显著影响训练效果。为了解决上述问题, 提出了基于区块链的自适应联邦学习(AWFL-BC, adaptive weight federated learning-blockchain)方案。通过智能合约计算不同边缘服务器的数据分布距离, 生成相似度矩阵指导聚合。同时, 设计了一种自适应权重聚合算法, 以减轻由数据质量差异引起的模型性能和稳定性下降, 从而提升模型的准确率, 加速模型收敛。最后, 结合区块链技术加强了安全保障机制, 可有效防止投毒攻击与推理攻击。在3个公有标准数据集上进行的综合实验表明, 与最先进的方法相比, AWFL-BC实现了更高的模型准确率, 且模型收敛速度更快。

关键词: 联邦学习; 区块链; 自适应权重; 非独立同分布数据; 算力网络

中图分类号: TP393

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2024045

Blockchain based adaptive federated learning in computational power network

LIU Tianrui^{1,2}, WANG Lianhai^{1,2}, WANG Qizheng^{1,2}, XU Shujia^{1,2},
ZHANG Shuhui^{1,2}, WANG Yingxiaochun^{1,2}

1. Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center

收稿日期: 2024-02-29; 修回日期: 2024-04-26

通信作者: 王连海, wanglh@sdas.org

基金项目: 山东省重点研发计划项目(2021CXGC010107); 国家自然科学基金(62102209); 济南市“新高校20条”资助项目(202228017); 泰山学者工程资助(tsqn202312231)

Foundation Items: Shandong Provincial Key Research and Development Program (2021CXGC010107), The National Natural Science Foundation of China (62102209), the New 20 project of higher education of Jinan, China (202228017), Taishan Scholars Program (tsqn202312231)

引用格式: 刘天瑞, 王连海, 王启正, 等. 算力网络环境下基于区块链的自适应联邦学习[J]. 网络与信息安全学报, 2024, 10(3): 130-142.

Citation Format: LIU T R, WANG L H, WANG Q Z, et al. Blockchain based adaptive federated learning in computational power network[J]. Chinese Journal of Network and Information Security, 2024, 10(3): 130-142.

(National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan 250014, China

2. Shandong Provincial Key Laboratory of Computer Networks,

Shandong Fundamental Research Center for Computer Science, Jinan 250014, China

Abstract: The computational power network was aimed to deeply integrate arithmetic resources and network resources, achieving efficient utilization of massive data and heterogeneous resources in ubiquitous discrete deployment. To cope with the complicated multicenter computing collaboration and requirements for data privacy and security in computational power networks, federated learning was recognized for its inherent advantages in privacy protection and edge arithmetic resource utilization. However, federated learning in the computational power network environment faced certain difficulties due to the highly heterogeneous and widely distributed arithmetic edge servers. On one hand, the heterogeneity of data among the massive edge servers in the computational power network caused the Non-IID problem, leading to the deviation of the local model update from the global optimum in federated learning. On the other hand, the generation of low-quality local models could significantly affect the training effect due to the differences in data quality among different edge servers. To solve the above problems, a blockchain-based adaptive federated learning framework AWFL-BC (adaptive weight federated learning-blockchain) was proposed. Initially, the data distribution distances of different edge servers were calculated through smart contracts to generate a similarity matrix that guided aggregation. Concurrently, an adaptive weight aggregation algorithm was designed to alleviate the decrease in model performance and stability caused by differences in data quality, thereby improving the accuracy of the model and accelerating model convergence. Finally, the integration of blockchain technology strengthened the security mechanism, effectively preventing poisoning attacks and inference attacks. Comprehensive experiments on three public standard datasets show that AWFL-BC achieves higher model accuracy and faster model convergence compared to state-of-the-art methods.

Keywords: federated learning, blockchain, adaptive weight, non-iid data distribution, computational power network

1 引言

近年来,随着5G、人工智能、大数据^[1-2]等新技术向各领域渗透,数字经济快速发展,社会对数据资源存储、计算和应用的需求大幅提升。如何满足日益增长的算力需求已经成为当前互联网所面临的一大挑战。算力网络^[3]是云化网络发展演进的下一个阶段,旨在将云、边、端的大量异构算力资源通过虚拟化的方式整合起来,实现全局资源和算力的优化,使海量的应用能够根据需要实时地调用分布式计算资源。作为一项新兴技术,算力网络在支撑国家“东数西算”^[4],在建设网络强国、数字中国的战略规划中意义重大。

然而,基于算力网络数据的特殊性,每个算力边缘服务器(下文简称为服务器)由海量边缘设备组成,且包含大量的隐私信息。在算力网络的数据共享过程中,一旦设备的敏感数据被泄露,将会影响整个算力网络的安全信誉度,降低各服务器中心共享本地数据的积极性。为了解决“数据孤岛”问题并保护数据隐私,近些年来联邦学习(Federated learning, FL)被提出。联邦

学习作为一种分布式机器学习方法^[5],可以在不共享敏感数据的情况下进行模型训练,大大提高了数据隐私的保护程度,避免了敏感数据被黑客入侵攻击的风险,非常适用于算力网络中的数据共享情景。算力网络环境下联邦学习的示意如图1所示。

FL与传统的分布式学习不同,传统的分布式学习通常将客户端数据集集中在一个位置,然后以独立同分布(independently identically distributed, IID)^[6]形式分发给代理。而IID数据的假设并不适用于真实的算力网络环境。一般来说,每个服务器之间都存在一定的数据异质性,其数据分布不同,生成的数据数量也极不平衡。下面给出了算力网络中3种类型的非独立同分布(non-independently identically distributed, Non-IID)示例。

1) 数据分布不均衡:不同服务器上的数据样本数量差异较大,导致模型学习受限。例如,某些服务器的数据集中包含某一类样本更多,而其他服务器则没有或包含的该类样本数量很少。

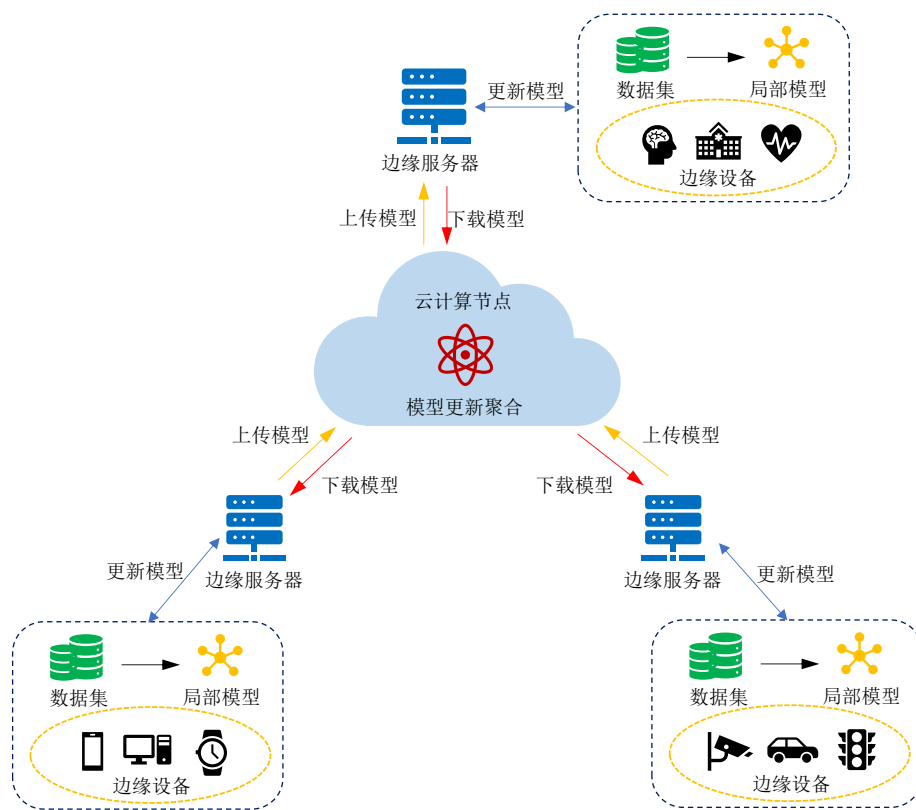


图1 算力网络环境下的联邦学习示意

Figure 1 Schematic diagram of federated learning in computing power network environment

2) 数据标签不一致性: 不同服务器上的数据标签存在一定差异, 这可能是数据集中存在标注错误或分类评判标准不一致导致的。例如, 对于同一类别的对象, 某些服务器可能将其标为正类, 而其他服务器将其标为负类。

3) 数据特征分布差异: 不同服务器上的数据特征分布不同, 这可能是服务器之间不同的样本采集环境、操作习惯等造成的。因此, 会导致训练出的模型在不同服务器上的性能差异较大。

虽然目前流行的FL算法——FedAvg算法^[7]在多种情况下都表现出了优异的性能, 然而, FedAvg无法处理不同服务器之间的Non-IID数据, 因为它直接对参与训练的客户端模型参数进行平均。因此, 数据分布的统计异质性会极大地降低模型的准确率和泛化性能。另外, 由于算力网络中接入了海量的边缘服务器, 其可信度和数据质量存在波动, 不同服务器训练的局部模型准确率差异较大, 这些因素都可能造成模型质量的下降。

本文提出了基于区块链的自适应联邦学习(AWFL-BC, adaptive weight federated learning-blockchain)方案, 解决算力网络环境下的联邦学习因数据异质性和数据质量差异引起的模型性能下降的问题, 并结合区块链技术, 实现隐私保护增强的安全联邦学习框架。主要贡献总结如下。

1) 考虑到现有算力网络中联邦学习参与方的数据分布和质量差异, 为算力网络系统提出AWFL-BC方案, 基于各参与方的数据分布生成相似度矩阵, 并根据区块链上对应的局部模型准确率, 自适应地调整模型聚合权重, 从而加速模型收敛, 缓解了FL中数据分布的不一致性和数据质量差异带来的性能下降的问题。

2) 结合区块链技术设计了一套安全保障机制, 通过智能合约计算和存储各参与方训练的局部模型哈希值及相应准确率, 防止恶意节点改变准确率信息, 干扰模型训练过程。并利用区块链的去中心化和难以篡改特性, 在模型聚合前验证

链上存储的局部模型的哈希值，若哈希值不同，则中止本轮聚合，有效防止投毒攻击与推理攻击，保护参与方数据的隐私和安全性。

3) 对 AWFL-BC 方案进行分析并且评估 AWFL-BC 在 3 个公共医疗数据集上的性能，实验表明，AWFL-BC 在 3 个数据集上都比最先进的方法取得更高的准确率和更快的收敛速度。

2 相关工作

谷歌在 2016 年提出了 FedAvg 算法，其关键思想是将与模型参数相关的交换取代直接的数据交换，能够有效解决“数据孤岛”问题。通过对联邦学习技术的全面研究，Li 等^[8]对 FL 相关的算法、问题和实例进行全面地总结并概述未来的方向 and 面临的挑战。然而，FL 的大多数研究工作是假设各方的数据是独立同分布的，即数据的分布不会发生变化。而在现实场景中，联邦学习的各参与方之间的统计异质性是主要挑战之一，因为每个参与方的数据分布不同，且在数据质量上存在差异，这将对模型的性能产生影响。图 2 所示为不同服务器具有不同的数据分布和模型准确率，直接聚合则只能生成性能次优的全局模型。

虽然 FedAvg 算法可以适应大多数的联邦学

习场景，但应用 FedAvg 算法会使 MNIST 数据集的模型准确率降低 11%，CIFAR-10 的模型准确率降低 51%，且算法存在不稳定性 and 收敛问题。因此，与假设的数据相比，Non-IID 数据会导致聚合模型偏离全局最优值，表现出性能下降。

有研究者提出了 FedAvg 算法的修改版本，例如 FedProx^[9]、FedBn^[10]和 FedNova^[11]，通过纠正参与方之间数据分布的异构性，来缓解 Non-IID 数据引起的问题。Zhao 等^[12]通过共享一部分数据子集来减轻 Non-IID 数据对 FL 的影响。然而，共享小部分数据的方法违背了 FL 在隐私保护方面的初衷。张泽辉等^[13]提出一种自适应模型聚合方案，保证了模型精度的同时，提高了训练效率，并基于混沌系统和同态加密的方法保护隐私。Geyer 等^[14]提出在中心服务器将高斯机制注入每一轮迭代后的权重矩阵中，并将单个客户的贡献和数据集隐藏在聚合过程中，在隐私保护和模型性能之间进行权衡。然而，当参与训练的客户端数量较少时，会导致模型精度的下降。Cai 等^[15]提出了将 DPCR 思想引入联邦学习的方法，在保护模型参数的同时，减轻因向梯度添加噪声造成模型准确性的下降。为了减轻 Non-IID 数据的影响，Zhou 等^[16]提出了将全局共享数据分发给各客户端的方法，但会有通信开销增加和

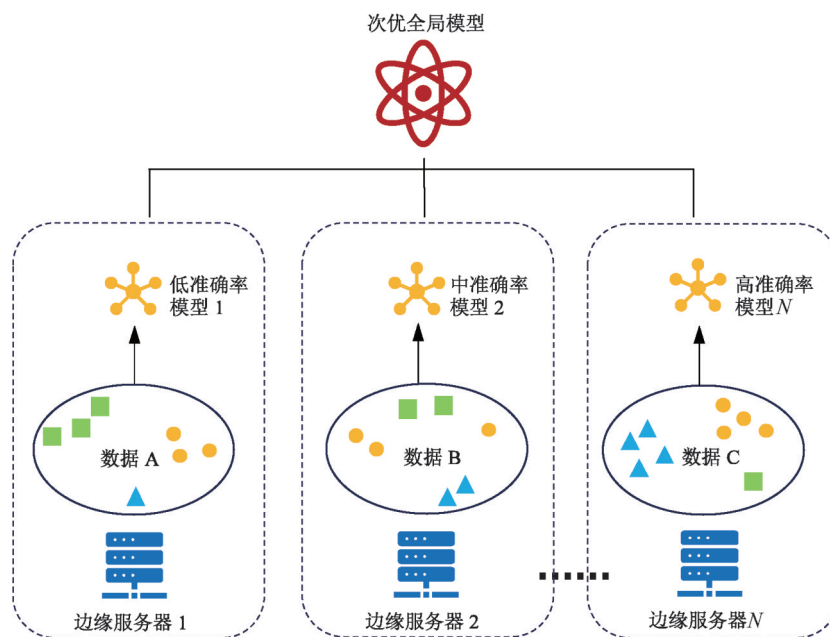


图 2 联邦学习中不同服务器的数据分布和数据质量差异

Figure 2 Differences in data distribution and quality among different devices in federated learning

隐私泄露的风险。Zheng 等^[17]在平均化共享参数时,引入了基于注意力的加权平均法,以针对异质数据。Yang 等^[18]提出了联合特征蒸馏的方法,来缓解联邦学习中的数据异构性问题,但共享数据中的部分特征仍在一定程度上有数据泄露的风险。侍江烽等^[19]提出计算局部模型准确率阈值的方法,并以模型筛选的方式来提升全局模型的性能,但会存在恶意节点谎报模型参数和准确率的问题,这会导致模型聚合稳定性下降。

以上对于联邦学习的研究都基于参与方节点是“诚实且可信”的假设,但在算力网络的海量边缘服务器中,基于庞大的服务器数量,其可信度和数据质量皆存在差异,模型的训练过程容易受到服务器单点故障或恶意攻击的影响,因此,亟须一种能够保护训练过程的方法,而基于区块链的联邦学习成为新的选择。

Wang 等^[20]讨论了现有联邦学习算法存在的问题和缺陷,从实际应用的角度分析了区块链如何与联邦学习相结合,并总结了未来的挑战和研究方向。应作斌等^[21]提出动态聚合权重的联邦学习框架,并通过注入噪声保护模型参数,降低不可信中心服务器带来的隐私泄露风险。朱建明等^[22]利用区块链记录联邦学习训练过程中的模型参数,通过奖惩机制鼓励节点参与模型参数验证,从而保证模型质量。Kim 等^[23]应用区块链技术来激励用户参与学习,并基于客户端模型的准确性和其参与训练的频率来筛选客户端,提高了模型的稳定性和收敛速度。然而,该方案没有考虑参与方的数据相似度,且海量服务器参与训练时所存储的局部模型会给区块链带来额外的通信和存储开销。Li 等^[24]提出使用区块链记录来自参与方的模型更新以及对这些更新进行评估。但是,同时维护多个区块链不利于模型的共享,区块的存储和同步会消耗巨大的硬件资源。表 1 对相关工作中提到的部分联邦学习算法进行了比较。

综上所述,现有的方法要么仅考虑了服务器端聚合的优化或专注于提升客户端训练效率,要么没有考虑对不同客户端数据质量及可信度的评估和服务器遭受恶意攻击的情形。目前针对算力网络环境下联邦学习的研究较少,而算力网络环

境下的联邦学习有一些独特的特点。本文在算力网络的环境下,研究从整个训练过程的角度来提升联邦学习准确性和安全性的新框架。

表 1 相关工作比较

Table 1 Comparison of related work

工作	隐私保护	考虑 Non-IID 情况	权重聚合方式
FedAvg ^[7]	否	否	数据集大小
Zheng ^[17]	否	是	数据分布
Shi ^[19]	否	是	模型质量
Wang ^[20]	否	是	数据分布
Zhu ^[22]	是	否	模型质量
AWFL-BC	是	是	自适应权重聚合

3 方案实现

3.1 系统架构

本文设计了基于区块链的自适应联邦学习框架,AWFL-BC 流程如图 3 所示,主要包括 3 个步骤。

步骤 1 区块链授权可信的云服务器作为聚合服务器,并将预训练好的模型分发给算力网络中的每个边缘服务器,作为联邦学习的参与方。

步骤 2 每个参与方设备基于本地数据训练模型并使用对称密钥加密算法 AES (Advanced Encryption Standard)^[25]和哈希加密算法^[26]将局部模型加密后,分别发送到聚合服务器和区块链。智能合约根据不同参与方的数据分布距离计算相似度矩阵,并对区块链中存储的对应局部模型的准确率进行归一化。

步骤 3 聚合服务器验证局部模型的哈希值,并基于相似度矩阵和归一化后的准确率,进行自适应加权聚合,之后聚合服务器更新模型到每个参与方,重复上述迭代训练,直至模型收敛或达到预先设定的迭代轮数。表 2 列出了 AWFL-BC 框架的相关符号和定义。

在 FL 的训练过程中,每个参与方 i 具有各自的损失函数,用于测量相对于当前模型参数的输出误差。FL 的训练目标是将损失函数最小化直至收敛。故 FL 的全局优化问题可以被公式化为:

$$\min_w F(w) \triangleq \sum_{i=1}^N p_i F_i(w_i; D_i) \quad (1)$$

表2 AWFL-BC框架的符号和定义
Table2 of AWFL-BC frame Notations and definitions

符号	定义
i	参与方
t	迭代轮数
w_i	局部模型参数
D_i	数据集
ϕ_i	BN层参数
ψ_i	其他层参数
E	相似度矩阵
l	批归一化层
μ_i	均值
σ_i	协方差矩阵
$d_{i,j}$	参与方分布距离
Acc_i	参与方准确率
α	自适应权重参数

其中, $p_i = |D_i| / \sum_{i=1}^N |D_i|$, $F(w_i, D_i)$ 是局部目标函数, N 是参与方的数量。

在第 t 轮迭代中, 聚合服务器会向参与方发送最新的模型参数 w_i^t , 并基于 w_i^t 计算关于本地数据的模型梯度 g_i^t , 即

$$g_i^t = \nabla F_i(w_i, D_i) \quad (2)$$

随后, 模型参数将按照式(3)进行更新:

$$w_i^{t+1} = w_i^t - \eta g_i^t \quad (3)$$

其中, η 是学习率。

假设有 N 个不同参与方 $\{C_1, C_2, \dots, C_N\}$, 每个参与方各自的模型表示为 $\{w_i\}_{i=1}^N$, 且每个参与方各自的数据集 $D_i = \{(x_{ij}, y_{ij})\}_{j=1}^m$, 其中, x_{ij} 是数据属性的向量, y_{ij} 是对应的类别标签, m 为样

本的数量。由此, 式(1)中给出的每个参与方 i 的局部目标函数可以进一步表示为:

$$F_i(w, D_i) \triangleq \frac{1}{m} \sum_{j=1}^m \ell(f(x_{ij}, w), y_{ij}) \quad (4)$$

其中, ℓ 是损失函数, $f(x_{ij}, w)$ 表示分类模型的输出向量。而对于模型权重参数 w^t 的更新, 将其优化为:

$$\tilde{w}^t = \sum_{i=1}^N (\alpha_i^{t-1} |w_i^t|) \quad (5)$$

其中, α_i^{t-1} 为自适应权重参数。

AWFL-BC 框架流程如图3所示, 整个过程在算法1中描述。

算法1 AWFL-BC框架

输入 w^0

1) 初始化全局权重参数 w^0 , 并将其分配给每个参与方;

2) 每个参与方计算各自的统计数据 (μ_i, σ_i) 并发送到聚合服务器;

3) 调用算法2生成 $E_{n \times n}$;

4) 各参与方使用本地数据训练模型并加密, 并将加密后的 w_i^t 推送到聚合服务器和区块链;

5) 聚合服务器从区块链获取 $E_{n \times n}$ 和4)生成的局部模型哈希值;

6) 聚合服务器对参与聚合的局部模型进行哈希操作, 并与区块链获取的哈希值作对比;

7) 验证哈希值后, 根据 $\begin{cases} \phi_i^{t+1} = \phi_i^{t*} \\ \psi_i^{t+1} = \sum_{j=1}^N \sigma_{ij} \psi_j^{t*} \end{cases}$

更新 $\{\theta_i^{t+1}\}_{i=1}^N$;

8) 调用自适应权重聚合的智能合约;

9) 根据式(5)更新 $\{w_i^{t+1}\}_{i=1}^N$ 并将其分发给各

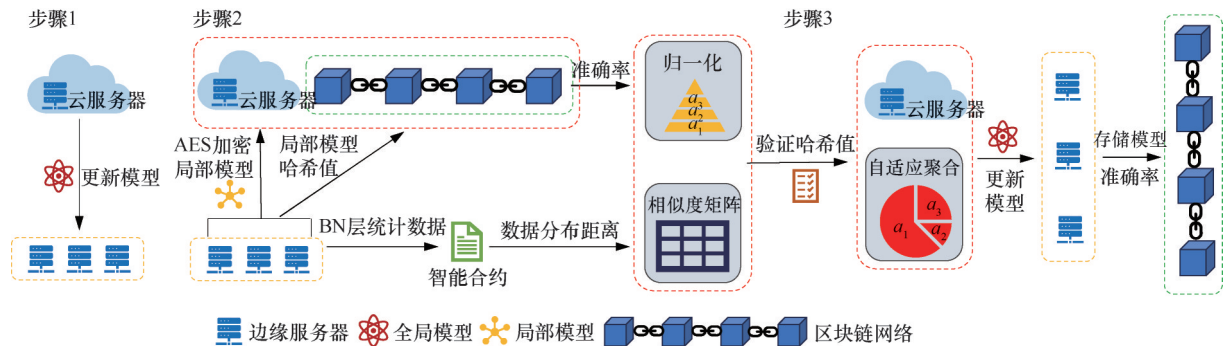


图3 AWFL-BC 流程

Figure 3 The illustration of AWFL-BC

参与方；

10) 重复 6)~9)，直到达到收敛或最大轮数。

3.2 相似度矩阵计算

本节将介绍相似度矩阵的计算方法。借鉴 Wang 等^[27]提出的权重矩阵计算方法与聚合更新策略，为了加强计算过程的安全性，将计算过程放到了智能合约中执行。

相似度矩阵 $\mathbf{E}$ 可基于各参与方数据的特征输出统计来计算，而神经网络模型的批归一化 (batch normalization, BN) 层参数可以用于计算移动平均值和移动方差，具有能潜在代表数据特征分布的特点，可以解决特征偏移问题。因此，将每个模型 w_i 的参数表示为 $\theta_i = \phi_i \cup \psi_i$ ，其中 ϕ_i 对应于每个参与方 BN 层的参数， ψ_i 是其他层的参数。使用 $\theta_i^t = \phi_i^t \cup \psi_i^t$ 表示第 t 轮中参与方 i 的模型参数，符号 $*$ 来表示更新的参数，所以，当第 i 个参与方使用本地数据更新 θ_i^t 后，更新后的模型参数可以表示为 $\theta_i^{t*} = \phi_i^{t*} \cup \psi_i^{t*}$ 。对于服务器的聚合更新策略，如式(6)所示。

$$\begin{cases} \phi_i^{t+1} = \phi_i^{t*} \\ \psi_i^{t+1} = \sum_{j=1}^N \varpi_{ij} \psi_j^{t*} \end{cases} \quad (6)$$

$\varpi_{ij} \in [0,1]$ 表示参与方 i 和参与方 j 之间的相似度， ϖ_{ij} 越大，两个参与方越相似。 $\sum_{j=1}^N \varpi_{ij}$ 中的 d_{ij} 构成了一个 $N \times N$ 的相似度矩阵 $\mathbf{E}_{n \times n}$ ，用于描述参与方之间的相似性。

使用上标符号 $l \in \{1,2,\dots,L\}$ 表示模型中的不同 BN 层。并保留每个参与方 BN 层中包含的均值和方差。 $z^{i,l}$ 表示第 i 个参与方在第 l 批归一化层的输入。将 $z^{i,l}$ 视为一个符合高斯分布的矩阵。对于第 i 个参与方的第 l 层，其分布为 $\gamma(\mu^{i,l}, \sigma^{i,l})$ 。因此，对于第 i 个参与方的 BN 层统计数据可以表示为

$$(\mu_i, \sigma_i) = [(\mu^{i,1}, \sigma^{i,1}), (\mu^{i,2}, \sigma^{i,2}), \dots, (\mu^{i,L}, \sigma^{i,L})] \quad (7)$$

采用 Wasserstein 距离来计算两个高斯分布之间的距离，以此推得两个参与方之间的相似度。如式(8)所示。

$$\begin{aligned} W_2^2(\gamma(\mu^{i,l}, \sigma^{i,l}), \gamma(\mu^{j,l}, \sigma^{j,l})) \\ = \|\mu^{i,l} - \mu^{j,l}\|^2 + \\ \text{tr}(\sigma^{i,l} + \sigma^{j,l} - 2((\sigma^{i,l})^{1/2} \sigma^{j,l} (\sigma^{i,l})^{1/2})^{1/2}) \end{aligned} \quad (8)$$

其中， $\|\mu^{i,l} - \mu^{j,l}\|^2$ 表示多元高斯分布 $\gamma(\mu^{i,l}, \sigma^{i,l})$,

$\gamma(\mu^{i,l}, \sigma^{i,l})$ 均值的二范数， tr 是矩阵的迹。为了减少计算开销，本文执行近似计算，因此，将 $\sigma^{i,l}$ 设为对角矩阵，即 $\sigma^{i,l} = \text{Diag}(r^{i,l})$ 。则 Wasserstein 距离的近似值为：

$$\begin{aligned} W_2^2(\gamma(\mu^{i,l}, \sigma^{i,l}), \gamma(\mu^{j,l}, \sigma^{j,l})) \\ = \|\mu^{i,l} - \mu^{j,l}\|^2 + \left\| \sqrt{r^{i,l}} - \sqrt{r^{j,l}} \right\|_2^2 \end{aligned} \quad (9)$$

所以，两个参与方 i, j 之间的距离为：

$$\begin{aligned} d_{ij} &= \sum_{l=1}^L W_2(\gamma(\mu^{i,l}, \sigma^{i,l}), \gamma(\mu^{j,l}, \sigma^{j,l})) \\ &= \sum_{l=1}^L \left(\|\mu^{i,l} - \mu^{j,l}\|^2 + \left\| \sqrt{r^{i,l}} - \sqrt{r^{j,l}} \right\|_2^2 \right)^{1/2} \end{aligned} \quad (10)$$

d_{ij} 越大意味着第 i 个参与方和第 j 个参与方之间的分布距离越大。因此， d_{ij} 越大，两个参与方越不相似；而 d_{ij} 越小，两个参与方越相似。相似度矩阵计算描述见算法 2。

算法 2 相似度矩阵计算智能合约

输入 (μ_i, σ_i)

- 1) for $ij = 1, 2, 3, \dots, N$ do
- 2) for $l = 1, 2, 3, \dots, L$ do
- 3) 计算各参与方之间高斯分布的距离；
- 4) 存储 d_{ij} 到 $\mathbf{E}_{n \times n}$ ；
- 5) end for
- 6) end for
- 7) return $\mathbf{E}_{n \times n}$

3.3 自适应权重聚合算法

自适应权重 (AW, adaptive weight) 聚合机制是基于每个联邦学习参与方的模型质量，动态调整模型权重的更新策略。在 AW 聚合算法中，根据区块链记录的上一轮迭代中各参与方模型的历史准确率，作为对参与方数据质量的评估依据，从而自适应地计算本轮迭代的聚合权重。同时，为了保证模型训练过程的稳定性，防止出现梯度爆炸等问题，引入了 sigmoid 激活函数，以此提升梯度下降优化算法的稳定性。sigmoid 函数的表达式如下：

$$\text{sig}(x) = \frac{1}{1 + \exp(-x)} \quad (11)$$

当第 t 轮训练完成后，聚合服务器会调用算法 3 的智能合约查询 Acc_i^{t-1} ，并通过 sigmoid 函数将 Acc_i^{t-1} 归一化后得到 α_i^t ，从而确定该轮次各模型参与聚合的权重。 α_i^t 的归一化计算公式如下。

$$\alpha_i^t = \text{sig} \left(\frac{\text{Acc}_i^t}{\sum_{i=1}^N \text{Acc}_i^t} \right) \quad (12)$$

其中, α_i^t 为第 i 个参与方第 t 轮的聚合权重。

当对各参与方自适应加权完成后, 聚合服务器会根据式(5)完成本轮模型参数的更新聚合。自适应权重聚合智能合约的流程如算法3所示。

算法3 自适应权重聚合智能合约

输入 学习任务 id, t

- 1) for $i = 1, 2, 3, \dots, N$ do
- 2) 查询各参与方上一轮的准确率 Acc_i^{t-1} ;
- 3) $\alpha_i^t \leftarrow \text{sig}(\text{Acc}_i^{t-1})$
- 4) end for
- 5) 将 $\{\alpha_i^t\}_{i=1}^N$ 代入式(5)更新 $\{w_i^{t+1}\}_{i=1}^N$;
- 6) return $\{w_i^{t+1}\}_{i=1}^N$

待本轮模型聚合完成后, 聚合服务器将更新后的模型参数发至参与聚合的各个参与方, 从而完成所有局部模型的更新。

4 方案分析

4.1 安全性分析

(1) 机密性

大部分情况下, 云服务器被视为安全可信的实体作为联邦学习的模型聚合服务器, 而当云服务器遭受恶意攻击时, 可能会泄露模型参数。因此, 出于机密性的考虑, AWFL-BC 方案不会公开参与方的模型中间参数和本地数据, 每个联邦学习的参与方使用各自的数据集进行训练后, 仅需上传更新参数。为了保证这些中间参数的机密性, AWFL-BC 采用 AES 加密算法实现这一目标。待局部模型训练完成后, 参与方利用 AES 加密中间参数后上传聚合服务器, 并将局部模型的哈希值存储到区块链, 以便在聚合阶段进行验证。这一机制有效保证局部模型的安全性, 并为模型安全聚合提供了条件。整个过程都是在密文的形式下进行的, 不会泄露参与方的敏感信息, 保证了参与方的本地数据和模型信息的机密性。

(2) 防止投毒攻击

投毒攻击^[28]的目的是通过篡改和操纵部分

训练数据或模型, 破坏联邦学习的可信度和效果。常见的攻击方式有模型投毒攻击和数据投毒攻击。对于模型投毒攻击, AWFL-BC 的各参与方在局部模型训练完成后, 会将局部模型的哈希值存储到区块链, 在模型聚合前, 聚合服务器首先对各局部模型进行一次哈希操作, 并与区块链上存储的哈希值作对比, 验证成功后再执行聚合。区块链能够实现数据的安全可信多方存储, 将局部模型的哈希值存储到区块链后, 如果有攻击者对模型投毒, 会导致哈希值与区块链中记录的值不符, 便无法通过聚合前的验证, 中止模型聚合。对于数据投毒攻击, 则通过 AWFL-BC 的自适应权重聚合算法实现这一目标, AW 聚合机制依据区块链记录的局部模型准确率对参与方数据质量进行评估, 如果参与方遭受数据投毒, 则训练出的局部模型的准确率会显著降低, 在之后的迭代轮次中, AW 聚合机制会自适应地降低其参与聚合的权重, 以此达到防御的目的。

(3) 防止推理攻击

推理攻击^[29]是指攻击者企图获取敏感数据信息的行为, 可能包括某个样本的具体数值或统计特征。首先, 在数据层面, 各参与方的样本数据在模型训练时不出本地, 仅通过交换局部模型参数协同训练全局模型, 保证了样本数据的安全性。此外, 在模型层面, 攻击者可以通过获取模型梯度来还原训练数据, 对于推理攻击, AWFL-BC 使用了 AES 算法对局部模型进行加密, 保证了模型参数在传输过程中的安全性。即使攻击者成功截获了局部模型, 也只会获得其密文, 而没有密钥的攻击者无法将其解密为明文, 得到模型梯度信息, 即攻击不成功, 因此, 保证了参与方的数据和模型安全。

(4) 效率

AWFL-BC 方案的各参与方在每轮迭代完成后, 分别采用 AES 对称加密算法和哈希加密算法对局部模型进行加密, 共同保证局部模型的传输过程安全和存储安全。此外, 该方案仅对模型哈希值进行存储, 可以有效地减轻区块链的存储压力, 大大降低了区块链的通信和存储负担, 从而提高了系统的效率。

5 实验设计

5.1 实验设置

(1) 实验环境

通过部署在 AMD Ryzen 7735H CPU、操作系统为 Windows 11 的计算机上进行单机模拟和实验仿真，并基于 Pytorch 深度学习框架，Python3.9、CUDA Toolkit 12.1 来评估 AWFL-BC 的性能。另外，使用以太坊来模拟区块链环境，采用 Solidity 语言编写智能合约，并使用 Web3 库来实现 Python 与以太坊的交互。

(2) 实验数据

本文评估了 AWFL-BC 方案在 3 个医疗数据集上的性能。3 个数据集的统计信息如表 3 所示。MedMnist^[30]数据集常应用于医学图像领域，是基于 MNIST 数据集的扩展。MedMnist 包含了不同类型的医学图像，例如 X 射线影像、放射学图像等，它包含 12 个 2D 数据集和 6 个 3D 数据集。从 12 个 2D 数据集中选择了 3 个关于腹部 CT 图像的数据集：OrganAMNIST、OrganCMNIST、OrganSMNIST。

表 3 3 个数据集的统计信息

数据集	类型	类别	实例数
OrganAMNIST	图像	11	58 850
OrganCMNIST	图像	11	23 660
OrganSMNIST	图像	11	25 221

5.2 实验结果分析

为了模拟 AW 算法中的问题情况，本文使用 Yurochkin 等^[31]的方法，通过调整 α 参数来对数据集进行划分，并构建了 $\alpha=0.1$ 和 $\alpha=0.01$ 两种 Non-IID 情况来模拟不同程度的数据异质性，以此来考察数据异质性对模型性能的影响。 α 参数可以用于控制贝叶斯非参数模型中的潜在分布形状，较小的 α 值可以增加数据的多样性和异质性。训练过程使用 CNN 模型的 LeNet-5^[32]，主要参数设定如下：参与方 $N=20$ ，训练轮次 $\text{epoch}=50$ ，学习率 $\text{lr}=0.01$ ，动量 $\text{momentum}=0.3$ ，批量大小 $\text{batchsize}=32$ 。

(1) 准确率结果分析

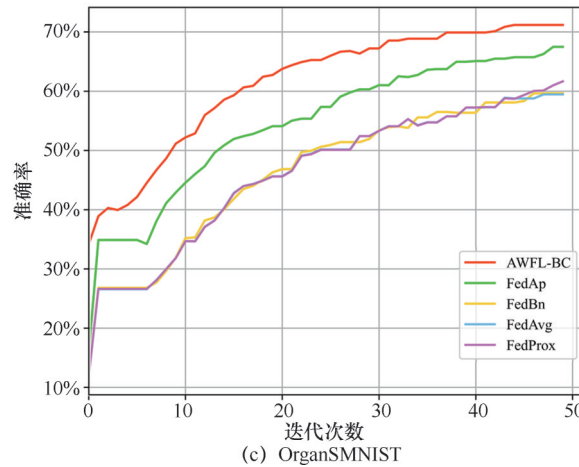
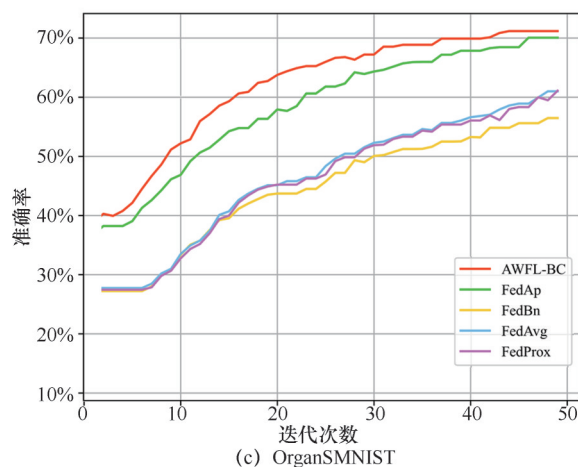
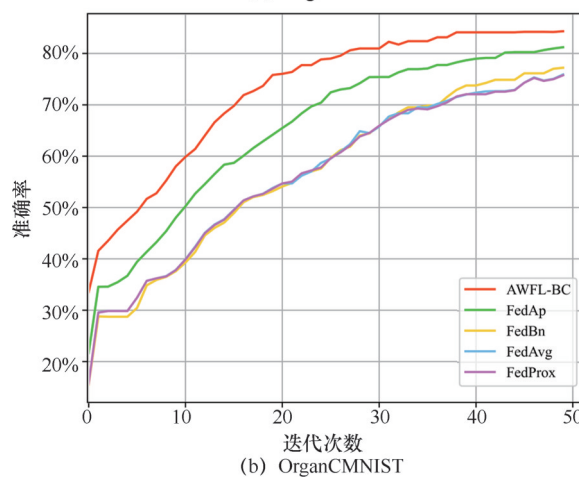
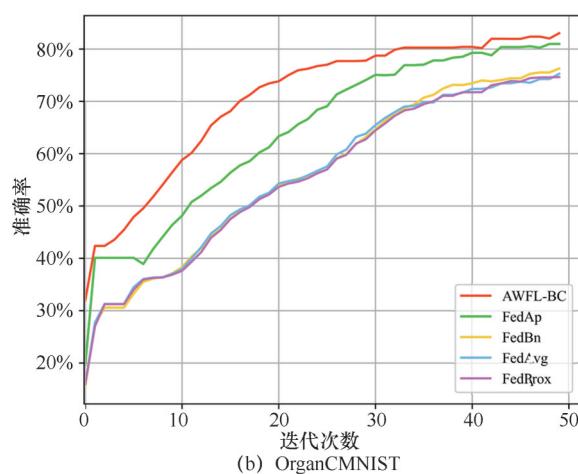
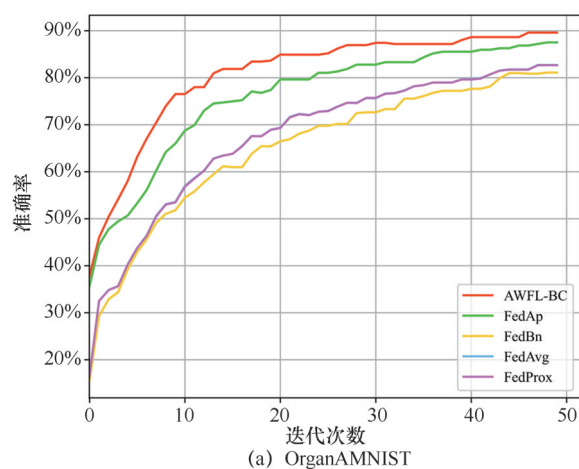
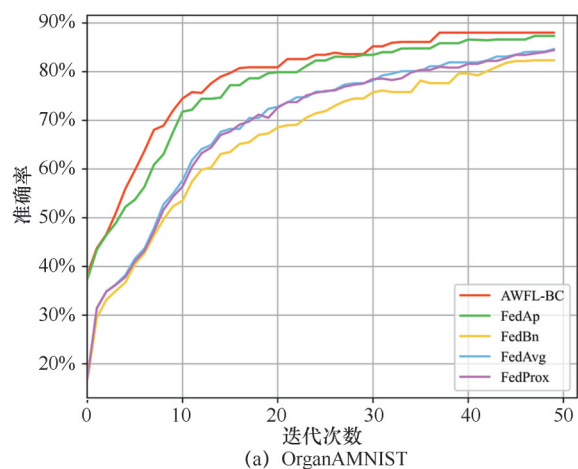
将本文提出的 AW 算法与常用的 FL 算法进行比较，其中包括 FedAvg^[7]算法、FedProx^[9]算法、FedBn^[10]算法、FedAp^[27]算法。

图 4 所示为迭代训练 50 轮的模型准确率情况，观察可知，在 $\alpha=0.1$ 的情况下，在 3 个数据集上都取得了最高的平均测试准确率，且超过 FedAvg 算法约 7%，FedAp 算法约 1.3%，当 $\alpha=0.01$ 时，平均准确率超过 FedAvg 算法约 9%，FedAp 算法约 2.98%。因为 FedAp 算法考虑参与方之间的相似度，取得了第二好的结果，但是它没有考虑局部模型的质量。而在图 4(c)中可以看出，FedBn 算法的结果欠佳，这可能是 FedBn 算法没有考虑不同参与方之间的数据相似性，且不擅长处理标签转移。而 FedProx 算法的准确性非常接近 FedAvg 算法，这表明 FedProx 算法中小的正则化项系数在联邦学习中效果并不明显，而过大的正则化项系数，会导致 FedProx 算法收敛缓慢或欠拟合。所以，相对其他 4 个算法的性能表现，本文算法 AWFL-BC 均有一定的提升。

在 MedMNIST 3 个数据集上的准确率实验结果如表 4 所示。将 α 参数缩小为 0.01 时，实验结果如图 5 所示。观察可知，不考虑 Non-IID 情况的 FedAvg 算法准确率有所下降，FedBn 算法和 FedProx 算法也都受到不同程度的影响。考虑参与方相似度的 FedAp 算法取得了第二好的准确率，而 AW 算法的准确率并没有下降，这表明 AW 算法相比其他算法，对 Non-IID 情况下的数据有更好的适应性。

最后，将 epoch 增加到 100，以此对比不同训练轮次下联邦学习的效果，实验结果如图 6、图 7 所示。观察可知，在训练过程的第 60 轮左右，模型趋向于收敛，而在 3 个不同的数据集上，AW 算法的收敛速度更快。随着训练轮次的增加，AW 算法与 FedAp 算法的收敛曲线逐渐重合，当训练轮次在第 100 轮时，AW 算法与其他 4 个联邦学习算法相比，仍取得了较高的准确率，证明了其在 Non-IID 数据场景下的有效性。

综上所述，虽然 Non-IID 数据会影响联邦学习的效果，而本文提出的方案减轻了 Non-IID 数据引起的性能下降，在 3 个数据集上都取得了更

图4 3个数据集下的准确率($\alpha=0.1$, epoch=50)Figure 4 The accuracy on three datasets
($\alpha=0.1$, epoch=50)

好的平均准确率, 且模型收敛速度更快, 说明了AW算法框架的有效性。

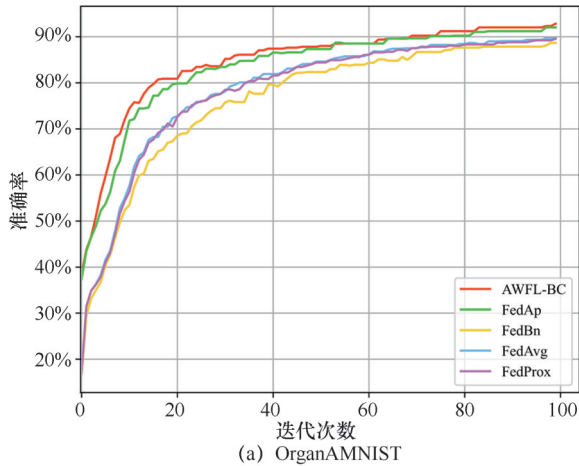
(2) 耗时对比

本文对比了AW算法和其他常用算法每轮训练的平均耗时。设置了50轮迭代训练, 在 $\alpha=$

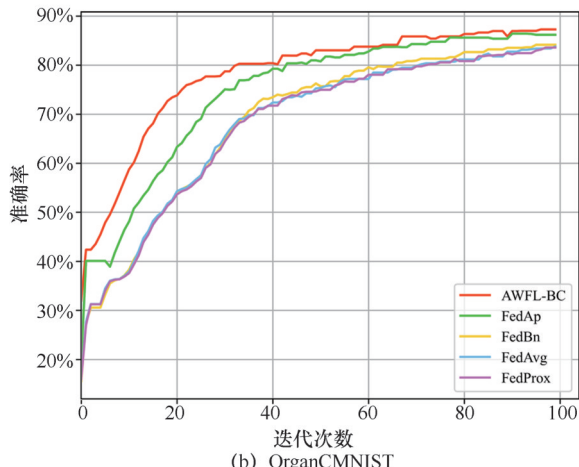
图5 3个数据集下的准确率($\alpha=0.01$, epoch=50)Figure 5 The accuracy on three datasets
($\alpha=0.01$, epoch=50)

0.01情况下的3个数据集进行了测试, 实验结果如图8所示。

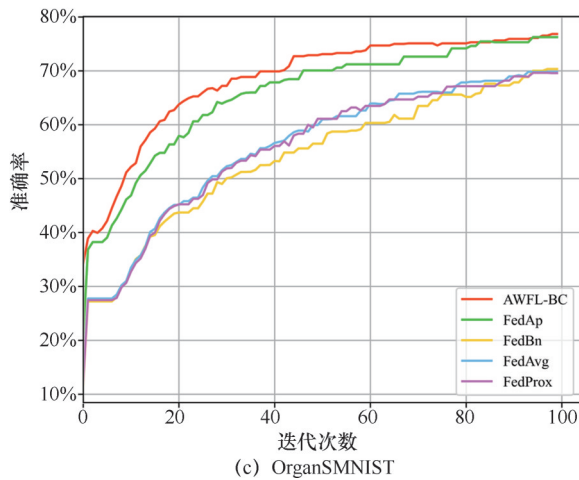
由实验结果可知, FedAvg算法和FedBn算法的执行时间最短, 因为传统的联邦学习通常只需要本地训练和模型聚合两个步骤, 耗时虽低但



(a) OrganAMNIST



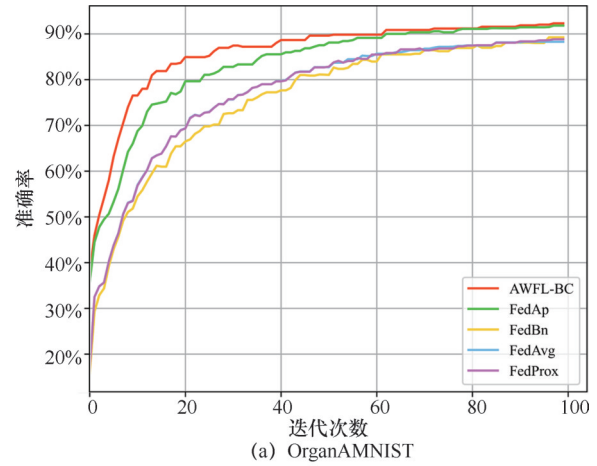
(b) OrganCMNIST



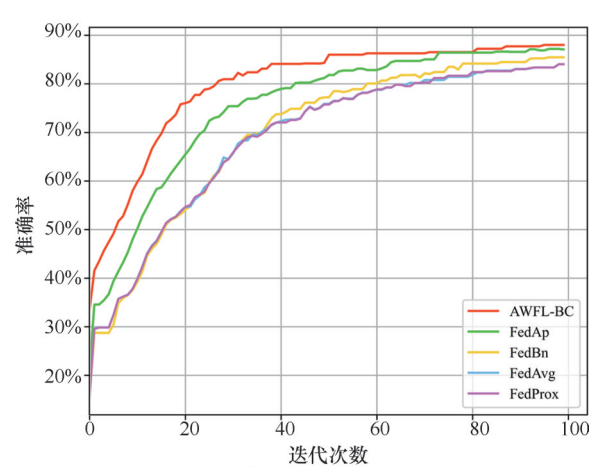
(c) OrganSMNIST

图6 3个数据集下的准确率($\alpha=0.01, \text{epoch}=100$)Figure 6 The accuracy on three datasets
($\alpha=0.01, \text{epoch}=100$)

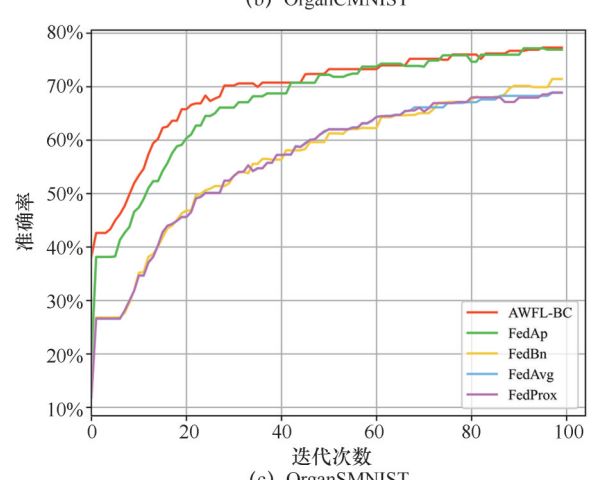
无法有效解决 Non-IID 引起的问题。本文提出的 AWFL-BC 方案基于相似度矩阵和自适应权重聚合算法,提升了 Non-IID 数据下联邦学习的准确率和收敛速度,且利用了区块链的去中心化和防



(a) OrganAMNIST



(b) OrganCMNIST



(c) OrganSMNIST

图7 3个数据集下的准确率($\alpha=0.01, \text{epoch}=100$)Figure 7 The accuracy on three datasets
($\alpha=0.01, \text{epoch}=100$)

篡改特性,一定程度上保护了模型隐私和训练过程的安全性。

本文模拟使用 20 个区块链节点测试了性能, AWFL-BC 方案的一轮迭代训练中,智能合约执

行的平均耗时仅为169.76 ms。说明本文提出的框架在隐私和模型性能之间实现了良好的平衡。

表4 准确率实验结果

Table 4 The experimental result of accuracy

Non-IID 划分	算法	MedMNIST 数据集		
		Organ-A	Organ-C	Organ-S
$\alpha=0.1$	AWFL-BC	88.01%	83.02%	71.18%
	FedAp	87.33%	80.97%	70.06%
	FedBn	82.33%	76.26%	56.48%
	FedProx	84.41%	74.65%	61.09%
	FedAvg	84.62%	75.26%	61.00%
$\alpha=0.01$	AWFL-BC	89.62%	84.32%	71.18%
	FedAp	87.52%	81.20%	67.47%
	FedBn	81.09%	77.21%	59.63%
	FedProx	82.67%	75.75%	61.66%
	FedAvg	82.72%	75.89%	59.45%

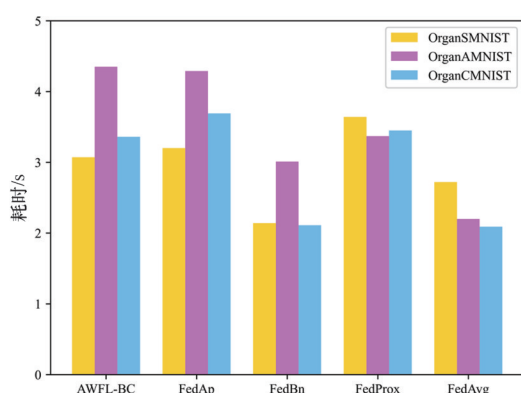


图8 AWFL-BC、FedAp、FedBn、FedProx 和 FedAvg 耗时对比

Figure 8 Comparison of time consumption between AWFL-BC, FedAp, FedBn, FedProx, and FedAvg

6 结束语

针对算力网络环境下的多中心数据共享过程中,不同边缘服务器存在数据质量差异和非独立同分布数据,导致的联邦学习中模型性能下降问题。提出了AWFL-BC方案。该方案通过计算不同参与方的数据分布距离,生成相似度矩阵指导聚合,同时,设计了一种自适应权重聚合算法,以平衡各参与方数据质量差异引起的模型性能下降问题,并结合区块链技术,提高模型训练过程的隐私性和安全性。实验表明,AWFL-BC方案在模型精度和收敛速度上都优于现有方案。未

来,计划将AWFL-BC和可信执行环境相结合,进一步提升联邦学习的隐私性和安全性。

参考文献:

- [1] RODRIGUEZ I, MOGENSEN R S, SCHJØRRING A, et al. 5G swarm production: advanced industrial manufacturing concepts enabled by wireless automation[J]. IEEE Communications Magazine, 2021, 59(1): 48-54.
- [2] LIAO W X, LUO C Q, SALINAS S, et al. Efficient secure outsourcing of large-scale convex separable programming for big data[J]. IEEE Transactions on Big Data, 2019, 5(3): 368-378.
- [3] SUN Y K, LEI B, LIU J L, et al. Computing power network: a survey [EB]. arXiv preprint arXiv: 2210.06080, 2022.
- [4] 梁芳, 佟恬, 马贺荣, 等. 东数西算下算力网络发展分析[J]. 信息技术与政策, 2022(11): 79-83.
- [5] LIANG F, TONG T, MA H R, et al. Analysis of the development of the computing power network under east-data-west-computing project[J]. Information and Communications Technology and Policy, 2022(11): 79-83.
- [6] SAVAZZI S, NICOLI M, BENNIS M, et al. Opportunities of federated learning in connected, cooperative, and automated industrial systems[J]. IEEE Communications Magazine, 2021, 59(2): 16-21.
- [7] VAHIDIAN S, MORAFAB M, LIN B. Personalized federated learning by structured and unstructured pruning under data heterogeneity [C]//Proceedings of the 2021 IEEE 41st International Conference on Distributed Computing Systems Workshops (ICDCSW). Piscataway: IEEE Press, 2021: 27-34.
- [8] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data[EB]. arXiv Preprint, 2016, arXiv:1602.05629.
- [9] LI T, SAHU A K, TALWALKAR A, et al. Federated learning: challenges, methods, and future directions[J]. IEEE Signal Processing Magazine, 2020, 37(3): 50-60.
- [10] YUAN X T, LI P. On convergence of FedProx: local dissimilarity invariant bounds, non-smoothness and beyond[EB]. arXiv Preprint arXiv: 2206.05187, 2022.
- [11] LI X X, JIANG M R, ZHANG X F, et al. FedBn: federated learning on non-IID features via local batch normalization[EB]. arXiv Preprint arXiv: 2102.07623, 2021.
- [12] WANG J Y, LIU Q H, LIANG H, et al. Tackling the objective inconsistency problem in heterogeneous federated optimization[EB]. arXiv Preprint, 2020, arXiv: 2007.07481.
- [13] ZHAO Y, LI M, LAI L Z, et al. Federated learning with non-IID data[EB]. arXiv Preprint, 2018, arXiv: 1806.00582.
- [14] 张泽辉, 李庆丹, 富瑶, 等. 面向非独立同分布数据的自适应联邦深度学习算法[J]. 自动化学报, 2023, 49(12): 2493-2506.
- [15] ZHANG Z H, LI Q D, FU Y, et al. Adaptive federated deep learning with non-IID data[J]. Acta Automatica Sinica, 2023, 49(12): 2493-2506.
- [16] GEYER R C, KLEIN T, NABI M. Differentially private federated learning: a client level perspective[EB]. arXiv Preprint arXiv: 1712.07557, 2017.
- [17] CAI J P, LIU X M, YE Q Q, et al. A federated learning framework based on differentially private continuous data release[J]. IEEE Transactions on Dependable and Secure Computing, 2024, (99): 1-16.
- [18] ZHOU F, CONG G J. On the convergence properties of a K-step averaging stochastic gradient descent algorithm for nonconvex optimi-

- zation[EB]. arXiv Preprint arXiv: 1708.01012, 2017.
- [17] ZHENG Q Q, CHEN S X, LONG Q, et al. Federated f-differential privacy[EB]. arXiv Preprint arXiv: 2102.11158, 2021.
- [18] YANG Z Q, ZHANG Y G, ZHENG Y, et al. FedFed: feature distillation against data heterogeneity in federated learning[EB]. arXiv Ppreprint arXiv: 2310.05077, 2023.
- [19] 侍江峰, 冯宝, 陈业航, 等. 基于自适应聚合权重联邦学习的肺结节 CT 图像分类[J]. 激光与光电子学进展, 2023, 60(22): 3788/LOP223027.
- SHI J F, FENG B, CHEN Y H, et al. Lung nodule CT image classification based on adaptive aggregate weight federated learning[J]. Laser & Optoelectronics Progress, 2023, 60(22): 3788/LOP223027.
- [20] WANG Z L, HU Q. Blockchain-based federated learning: a comprehensive survey[EB]. arXiv Preprint arXiv: 2110.02182, 2021.
- [21] 应作斌, 方一晨, 张怡文. 动态聚合权重的隐私保护联邦学习框架[J]. 网络与信息安全学报, 2022, 8(5): 56-65.
- YING Z B, FANG Y C, ZHANG Y W. Privacy-preserving federated learning framework with dynamic weight aggregation[J]. Chinese Journal of Network and Information Security, 2022, 8(5): 56-65.
- [22] 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[J]. 计算机学报, 2021, 44(12): 2464-2484.
- ZHU J M, ZHANG Q N, GAO S, et al. Privacy preserving and trustworthy federated learning model based on blockchain[J]. Chinese Journal of Computers, 2021, 44(12): 2464-2484.
- [23] KIM Y J, HONG C S. Blockchain-based node-aware dynamic weighting methods for improving federated learning performance [C]//Proceedings of the 2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS). Piscataway: IEEE Press, 2019: 1-4.
- [24] LI Y Z, CHEN C, LIU N, et al. A blockchain-based decentralized federated learning framework with committee consensus[J]. IEEE Network, 2021, 35(1): 234-241.
- [25] DAEMEN J, RIJMEN V. REIJNDAEL: The advanced encryption standard[J]. Dr. Dobbs' Journal: Software Tools for the Professional Programmer, 2001, 26(3): 137-139.
- [26] CORON J S, DODIS Y, MALINAUD C, et al. Merkle-Damgård revisited: How to construct a hash function[C]//Advances in Cryptology - CRYPTO 2005: 25th Annual International Cryptology Conference, Santa Barbara, California, USA, August 14-18, 2005. Proceedings 25. Springer Berlin Heidelberg, 2005: 430-448.
- [27] LU W, WANG J D, CHEN Y Q, et al. Personalized federated learning with adaptive batchnorm for healthcare[J]. IEEE Transactions on Big Data, 2022, PP(99): 1.
- [28] MUÑOZ-GONZÁLEZ L, BIGGIO B, DEMONTIS A, et al. Towards poisoning of deep learning algorithms with back-gradient optimization[C]//Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security. New York: ACM, 2017: 27-38.
- [29] MELIS L, SONG C Z, DE CRISTOFARO E, et al. Exploiting unintended feature leakage in collaborative learning[C]//Proceedings of the 2019 IEEE Symposium on Security and Privacy (SP). Piscataway: IEEE Press, 2019: 691-706.
- [30] YANG J C, SHI R, WEI D L, et al. MedMNIST v2 - a large-scale lightweight benchmark for 2D and 3D biomedical image classification[J]. Scientific Data, 2023, 10: 41.
- [31] YUROCHKIN M, AGARWAL M, GHOSH S, et al. Bayesian non-parametric federated learning of neural networks[EB]. arXiv Preprint, arXiv: 1905.12022, 2019.
- [32] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[作者简介]



刘天瑞 (1995-), 男, 山东济南人, 齐鲁工业大学 (山东省科学院) 硕士生, 主要研究方向为区块链技术、联邦学习。



王连海 (1969-), 男, 山东潍坊人, 齐鲁工业大学 (山东省科学院) / 山东省计算中心 (国家超级计算济南中心) 研究员, 主要研究方向为数据安全、区块链、计算机取证。



王启正 (1993-), 男, 山东济南人, 齐鲁工业大学 (山东省科学院) / 山东省计算中心 (国家超级计算济南中心) 研究员, 主要研究方向为隐私保护、人工智能。



徐淑奖 (1979-), 男, 山东临沂人, 齐鲁工业大学 (山东省科学院) / 山东省计算中心 (国家超级计算济南中心) 研究员, 主要研究方向为数据安全、区块链。



张淑慧 (1981-), 女, 山东济宁人, 齐鲁工业大学 (山东省科学院) / 山东省计算中心 (国家超级计算济南中心) 研究员, 主要研究方向为数据安全、区块链、计算机取证。



王英晓春 (1998-), 女, 山东青岛人, 齐鲁工业大学 (山东省科学院) 硕士生, 主要研究方向为数据安全、区块链应用。