

支持多数不规则用户的隐私保护联邦学习框架

陈前昕^{1,2}, 毕仁万^{1,2}, 林劫¹, 金彪¹, 熊金波^{1,2}

(1. 福建师范大学计算机与网络空间安全学院, 福建 福州 350117;

2. 福建师范大学福建省网络安全与密码技术重点实验室, 福建 福州 350007)

摘要: 针对联邦学习存在处理大多数不规则用户易引起聚合效率降低, 以及采用明文通信导致参数隐私泄露的问题, 基于设计的安全除法协议构建针对不规则用户鲁棒的隐私保护联邦学习框架。该框架通过将模型相关计算外包给两台边缘服务器以减小采用同态加密产生的高额计算开销, 不仅允许模型及其相关信息以密文形式在边缘服务器上聚合, 还支持用户在本地进行模型可靠性计算以减小传统方法采用安全乘法协议造成的额外通信开销。在该框架的基础上, 为更精准评估模型的泛化性能, 用户完成本地模型参数更新后, 利用边缘服务器下发的验证集与本地持有的验证集联合计算模型损失值, 并结合损失值历史信息动态更新模型可靠性以作为模型权重。进一步, 在模型可靠性先验知识指导下进行模型权重缩放, 将密文模型与密文权重信息交由边缘服务器对全局模型参数进行聚合更新, 保证全局模型变化主要由高质量数据用户贡献, 提高收敛速度。通过 HybridArgument 模型进行安全性分析, 论证表明 PPRFL (privacy-preserving robust federated learning) 可以有效保护模型参数以及包括用户可靠性等中间交互参数的隐私。实验结果表明, 当联邦聚合任务中的所有参与方均为不规则用户时, PPRFL 方案准确率仍然能达到 92%, 收敛效率较 PPFDL (privacy-preserving federated deep learning with irregular users) 提高 1.4 倍; 当联邦聚合任务中 80% 用户持有的训练数据都为噪声数据时, PPRFL 方案准确率仍然能达到 89%, 收敛效率较 PPFDL 提高 2.3 倍。

关键词: 联邦学习; 隐私保护; 安全聚合; 大多数不规则用户; 安全除法协议

中图分类号: TP309

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2022011

Privacy-preserving federated learning framework with irregular-majority users

CHEN Qianxin^{1,2}, BI Renwan^{1,2}, LIN Jie¹, JIN Biao¹, XIONG Jinbo^{1,2}

1. College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350117, China

2. Fujian Provincial Key Laboratory of Network Security and Cryptology, Fujian Normal University, Fuzhou 350007, China

收稿日期: 2021-10-07; 修回日期: 2022-01-05

通信作者: 林劫, linjie891@163.com

基金项目: 国家自然科学基金(61872088, 61872090, U1905211); 福建省自然科学基金(2019J01276)

Foundation Items: The National Natural Science Foundation of China (61872088, 61872090, U1905211), The Natural Science Foundation of Fujian Province (2019J01276)

引用格式: 陈前昕, 毕仁万, 林劫, 等. 支持多数不规则用户的隐私保护联邦学习框架[J]. 网络与信息安全学报, 2022, 8(1): 139-150.

Citation Format: CHEN Q X, BI R W, LIN J, et al. Privacy-preserving federated learning framework with irregular-majority users[J]. Chinese Journal of Network and Information Security, 2022, 8(1): 139-150.

Abstract: In response to the existing problems that the federated learning might lead to the reduction of aggregation efficiency by handing the majority of irregular users and the leak of parameter privacy by adopting plaintext communication, a framework of privacy-preserving robust federated learning was proposed for ensuring the robustness of the irregular user based on the designed security division protocol. PPRFL could enable the model and its related information to aggregate in ciphertext on the edge server facilitate users to calculate the model reliability locally for reducing the additional communication overhead caused by the adoption of the security multiplication protocol in conventional methods, apart from lowering the high computational overhead resulted from homomorphic encryption with outsourcing computing to two edge servers. Based on this, user could calculate the loss value of the model through jointly using the verification sets issued by the edge server and that held locally after parameter updating of the local model. Then the model reliability could be dynamically updated as the model weight together with the historic information of the loss value. Further, the model weight was scaled under the guidance of prior knowledge, and the ciphertext model and ciphertext weight information are sent to the edge server to aggregate and update the global model parameters, ensuring that global model changes are contributed by high-quality data users, and improving the convergence speed. Through the security analysis of the Hybrid Argument model, the demonstration shows that PPRFL can effectively protect the privacy of model parameters and intermediate interaction parameters including user reliability. The experimental results show that the PPRFL scheme could still achieve the accuracy of 92% when all the participants in the federated aggregation task are irregular users, with the convergence efficiency 1.4 times higher than that of the PPFDL. Besides, the PPRFL scheme could still reach the accuracy of 89% when training data possessed by 80% of the users in the federated aggregation task were noise data, with the convergence efficiency 2.3 times higher than that of the PPFDL.

Keywords: federated learning, privacy-preserving, secure aggregation, irregular-majority users, security division protocol

0 引言

近年来, 各类数据驱动的深度学习 (DL, deep learning) 系统性能获得大幅提升, 人工智能实现了经济与公众效益的共赢^[1]。然而, 由第三方服务商收集用户数据将不可避免地导致个人隐私数据泄露^[2]。为解决数据孤岛难题并实现数据隐私共享, 联邦学习 (FL, federated learning) 技术应运而生^[3-4]。尽管联邦学习能缓解公众对隐私泄露的担忧, 但相关研究^[5-6]表明, 攻击者可以通过用户上传的明文梯度信息间接推理出用户本地部分数据集成员和对应标签的信息。此外, 明文梯度有可能遭受模型反演、模型推理等攻击^[7], 导致数据隐私泄露。

针对联邦学习梯度逆推带来的新挑战, 已有一些研究^[8-9]聚焦于提升联邦学习的隐私性, 并应用于各种场景^[10-12]。在保证隐私性的同时, 多种联邦学习方案进一步研究如何提高系统的训练效率。Kanagavelu 等^[13]为降低通信成本和提高可扩展性, 提出一种支持安全多方计算的两阶段联邦学习方案。董业等^[14]基于秘密共享技术和 Top-K 梯度选择算法, 提出高效安全的联邦学习框架。

Aono 等^[15]对梯度数据进行同态加密, 保证多个训练者的本地数据隐私安全。文献^[16]方案使用 Paillier 同态加密方法优化加密模型聚合过程。

然而, 这些隐私保护联邦学习方案假设每个用户均持有高质量的数据, 没有考虑系统中存在不规则用户的情况。实际上, 拥有高级专业知识或高级设备的用户通常会生成高质量的数据, 其他用户却可能由于噪声干扰、记录错误等产生低质量数据, 这些用户称为不规则用户^[17], 他们上传的模型参数可能会对全局模型精度造成负面影响。

隐私保护的深度学习协作系统 (SecProbe, privacy-preserving collaborative deep learning system)^[17]是第一种处理不规则用户的方案, 该方案采用差分隐私技术向深度神经网络 (DNN, deep neural network) 添加噪声, 并提出在线抽样选择用户的机制。已有研究^[6,18]表明目前用于个性化深度学习的差分隐私机制, 尚不能为复杂任务提供可接受的效用与隐私权衡。成艺等^[19]集成 Diffie-Hellman 密钥协商技术和 Shamir (t, n) 门限机制^[20]设计保护聚合结果的减少异质性数据影响的联邦学习方案 (SAHPP, a privacy-preserving

federated learning solution to reduce the impact of heterogeneous data) 方案, 该方案需要用户与其他多个用户进行密钥协商, 消耗大量的网络资源, 不适合大规模联合学习。Xu 等^[21]基于真值发现方法^[18,22]和边缘计算架构^[22-24]设计在线计算模型梯度可靠性、存在不规则用户的隐私保护联邦深度学习 (PPFDL, privacy-preserving federated deep learning with irregular users) 方案。然而, SAHPP 和 PPFDL 严重依赖于局部模型与中心聚合模型的距离, 在不规则用户占大多数参与方的场景下, 其聚合效率受到较大影响。

针对上述直接传输原始模型梯度导致模型隐私信息泄露以及大多数不规则用户存在导致现有联邦学习框架效率降低的严峻问题, 本文提出一种支持大多数不规则用户鲁棒的隐私保护联邦学习 (PPRFL, privacy-preserving robust federated learning) 框架, 主要贡献总结如下。

1) 设计一种动态的模型可靠性更新策略, 解决大多数不规则用户参与训练的精度提升缓慢问题, 并基于可靠性先验知识指导模型权重缩放, 实现全局模型聚合更新, 保证全局模型变化主要由高质量数据用户贡献, 进一步提高了联邦模型收敛的效率。

2) 设计高效的安全除法协议, 构建边缘服务器辅助的鲁棒隐私保护联邦学习框架 PPRFL, 规避上传和处理过程中的模型参数隐私泄露风险。

3) 安全性详细证明 PPRFL 框架可以有效保护模型参数的隐私性。实验结果表明, 当所有参与方都为不规则用户时, PPRFL 方案的准确率仍可达 92%, 相比 PPFDL 方案, 聚合效率提高 1.4 倍, 并且服务器间的计算与通信开销较低。

1 预备知识

1.1 联邦学习

联邦学习框架一般由中心服务器和持有隐私数据的用户构成^[25], 系统中的用户在接收初始模型后对本地参数进行训练。给定训练集 $\mathcal{D} = \{(x_i, y_i); i=1, 2, \dots, N\}$ 和模型参数 w , DNN 的输出可表示为 $y = f(x_i, w)$, 其中 f 为神经网络的输出函数, y 为模型输出结果。参数训练的目的为从函数集合中寻找一个函数来近似每个样本特

征向量 x_i 到标签 y_i 之间的真实映射关系。训练集上的目标损失函数如式(1)所示, $l_f(x, y, w)$ 为单个样本的损失, 如交叉熵损失 $l_f(x, y, w) = -y \ln f(x, w) - (1 - y) \ln(1 - f(x, w))$ 。交叉熵损失函数通常用于分类任务, 本文实验采用该损失函数作为用户端的损失函数。

$$L_f(\mathcal{D}, w) = \frac{1}{|\mathcal{D}|} \sum_{(x, y) \in \mathcal{D}} l_f(x, y, w) \quad (1)$$

在 j 次训练时, 采用随机梯度下降 (SGD, stochastic gradient descent) 方法更新模型参数, 直到目标函数 $L_f(\mathcal{D}, w)$ 达到最小值, 其中 η 为学习率, 如式(2)所示。在联邦学习训练过程中, 用户基于本地数据集使用 SGD 方法优化模型参数, 并对同一批数据进行多次训练。

$$w^{j+1} \leftarrow w^j - \eta \frac{1}{|\mathcal{D}|} \sum_{n=1}^N \frac{\partial L_f(\mathcal{D}^{(n)}, w^j)}{\partial w} \quad (2)$$

为避免服务器访问本地数据, 用户将训练后的模型发送给服务器。服务器收到用户 $i \in [1, n]$ 发送的模型参数 w_i^{j+1} 和数据集数量 $|\mathcal{D}_i|$ 后, 通常采用如式(3)进行全局模型更新^[26]。

$$w^{j+1} \leftarrow w^j - \eta \sum_{i=1}^n \frac{w_i^{j+1}}{|\mathcal{D}_i|} \quad (3)$$

1.2 Paillier 密码系统

Paillier 密码系统^[27]是一种部分同态加密方案 (PHE, partial homomorphic encryption), 包含密钥生成、加密和解密算法, 具有加法同态的性质^[28]。已知明文消息 m_1 和 m_2 , 使用同一公钥 pk 可加密为密文 $\text{Enc}_{pk}(m_1)$ 和 $\text{Enc}_{pk}(m_2)$, 满足 $\text{Enc}_{pk}(m_1) \cdot \text{Enc}_{pk}(m_2) = \text{Enc}_{pk}(m_1 + m_2)$ 。已知常数 t 和明文消息 m , 满足 $\text{Enc}_{pk}(m)^t = \text{Enc}_{pk}(t \cdot m)$ 。为避免明文模型与交互参数在通信和处理过程中的隐私泄露风险, 本文利用 Paillier 密码系统构建了边缘服务器辅助的隐私保护联邦学习框架。

2 系统模型与安全模型

2.1 系统模型

PPRFL 框架的系统模型如图 1 所示, 包含服

服务器端和用户端。用户首先利用本地数据集对模型进行更新,然后基于验证集损失值计算模型的可靠性,并将相关信息加密发送给服务器。两台非共谋的服务器 S_0 和 S_1 交互执行隐私保护模型聚合过程,并将加密全局模型返回给所有用户。服务器和用户反复执行这个流程,直到全局模型满足预定义的终止迭代条件。其中, S_0 作为接收用户加密数据的边缘服务器, S_1 提供解密密钥并协助 S_0 进行计算。这种架构广泛地应用于外包计算场景^[17,21-22],其优点在于外包计算给两台边缘服务器以减轻本地计算资源的压力。

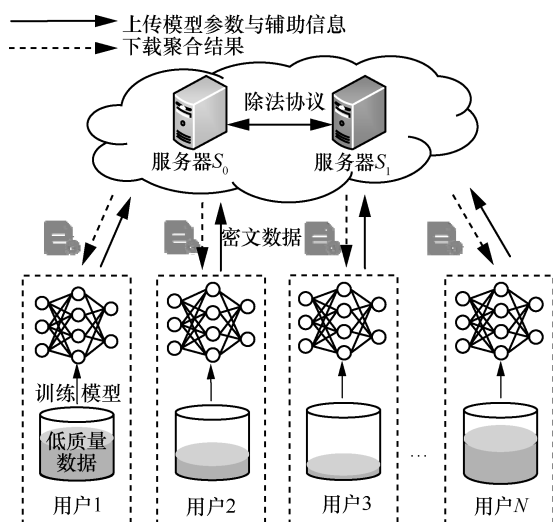


图1 PPRFL 方案系统模型
Figure 1 System model of the PPRFL scheme

2.2 安全模型

如同大多数外包计算场景的安全假设一样,两台服务器为不共谋的诚实但好奇服务器^[20-22]。服务器能正确地执行协议规范和相应操作,但可能利用在协议期间收集到的存储信息和消息流中的数据来推测用户的隐私数据^[29]。恶意用户通常会对图像进行标签翻转或植入 Pattern 实现投毒攻击。与恶意用户不同,不规则用户通常在生成和存储数据的过程中,由于记录错误、噪声数据干扰、使用陈旧数据等产生低质量数据。此外,用户会诚实地执行模型训练以及后续加密上传等操作,因此认为用户是可信的。

此外,模型引入一个概率多项式时间的敌手 $\mathcal{A}$,该敌手拥有下述攻击能力和约束。① $\mathcal{A}$ 可能攻击 S_1 以猜测来自用户密文对应的明文值,也可

能通过执行交互协议猜测来自另一方服务器密文对应的明文值。② $\mathcal{A}$ 可能攻击 S_2 ,通过执行交互协议猜测来自 S_1 密文对应的明文值。③ $\mathcal{A}$ 不能同时攻击 S_1 和 S_2 。若服务器和 $\mathcal{A}$ 不能获得隐私信息,则意味着本文方案在此安全模型下是安全的。

基于上述威胁模型,与文献[19,21]类似,本文有如下隐私需求。① 用户模型参数。敌手可以通过本地模型和全局模型参数对用户进行成员推理等攻击恢复出用户的敏感信息,为保护用户数据隐私,模型参数应该以密文形式提交到服务器。② 模型可靠性。模型可靠性可以认为是用户数据质量的评价,为使学习过程公平无歧视,这个信息应对除用户外的其他参与方保密。③ 模型聚合结果。全局模型的聚合结果可以视为用户使用大量数据资源生成的知识产权,除了参与训练的用户外应对其他参与方保密。

3 PPRFL 方案

PPRFL 包含用户端和服务器端两个部分,用户端上的参与方在本地训练得到局部模型,基于组合验证集计算损失值,并结合历史信息,采用动态更新以及权重缩放方式得到模型可靠性;在服务器端上,服务器收到加密参数后先利用加同态性质进行聚合,然后交互执行安全除法协议得到全局模型。参与联邦学习的用户有着相似的收集数据场景,类似于隐私保护联邦学习方案^[11,15,30],本文考虑每个用户持有的高质量数据都满足独立同分布 (IID, independent and identically distributed)。

3.1 模型可靠性更新

本节通过验证集损失分析来介绍处理大多数不规则用户的动态更新策略。在低质量数据集的联邦学习场景中,直接使用验证集损失值作为模型权重的计算方法将导致权重区分度失衡。这种失衡表现为经不同噪声比例数据训练出来的模型,在同一损失值下模型精度差异大;或是在相同的模型精度上,损失值差异大。基于以上观察,本文设计组合验证集计算用户模型损失值,并结合损失值历史信息辅助评估当前模型可靠性的动态更新方法。

具体来说,在第 E 轮迭代中,用户 i 结合本

地验证集和服务端验证集共同计算损失值 loss_i^E 。由于用户模型可能对本地数据拟合效果好，产生较低损失值的可能性较大，直接使用本地验证集损失值作为模型可靠性的计算依据是不公平的。而采用服务器维护的小部分高质量验证集^[18,31-32]损失情况作为计算依据，将不能得到本地模型对噪声数据的拟合情况。

为了更精准地评估模型的泛化性能，采用本地验证集和服务端验证集共同计算损失，计算方法如式(4)所示，其中 $|\mathcal{D}_v|$ 表示本地验证集数量， $|\mathcal{D}_s|$ 表示服务器验证集数量。

$$\text{loss}_i^E = \frac{|\mathcal{D}_v|}{|\mathcal{D}_v| + |\mathcal{D}_s|} \sum_{(x,y) \in \mathcal{D}_v} l_f(x,y,w) + \frac{|\mathcal{D}_s|}{|\mathcal{D}_v| + |\mathcal{D}_s|} \sum_{(x,y) \in \mathcal{D}_s} l_f(x,y,w) \quad (4)$$

已有研究^[33-35]表明历史行为可以反映一个实体在当下活动中的可靠程度。受此启发，本文考虑结合损失值的历史信息来动态更新当前用户模型的可靠性。在得到 loss_i^E 后，利用第 $E-1$ 轮的模型可靠性 u_i^{E-1} ，计算得到当前轮次的模型可靠性 u_i^E ，如式(5)所示。

$$u_i^E = \text{loss}_i^E + \gamma u_i^{E-1} \quad (5)$$

其中， $\gamma = \frac{1}{2} + \frac{\ln E}{10}$ ， γ 随着 E 的增加而缓慢增加，并且当 $E=1$ 时， $u_i^{E-1} = 0$ 。

此外，随着全局训练迭代轮次增加，用户模型损失值差异逐渐减小，不规则用户模型占全局模型权重逐渐提高。文献^[18,22]表明不同用户利用高质量 IID 样本训练得到的梯度，在梯度下降方向上有着很高的相似概率，这些高质量 IID 数据产生的梯度更加接近于使全局损失最小化的最优梯度^[19]。由于不规则用户采用噪声数据对模型进行训练，产生的梯度会偏离高质量数据产生的梯度方向。因此，大量不规则用户参与聚合将会减缓全局模型收敛速度。为解决上述问题，本文基于可靠性先验知识，采用式(6)方式对模型可靠性进行缩放。在固定训练轮次 E 下，将基于损失值的模型可靠性 u_i 映射到函数上，使 u_i 越小 τ_i 越大。随着训练轮次 E 的增加，激励小值 u_i 对应的 τ_i ，抑制大值 u_i 对应的 τ_i 。通过这样缩放可靠性

的方式，减小不规则用户模型在全局模型的占比权重，保证全局梯度方向主要由高质量数据用户模型贡献。

$$\tau_i = \begin{cases} \left(\frac{1}{\ln 2} \right)^{u_i^E}, & E=1 \\ \left(\frac{1}{\ln E} \right)^{u_i^E}, & E \geq 2 \end{cases} \quad (6)$$

PPRFL 框架在用户端的执行算法如算法 1 所示，在用户 i 得到模型可靠性 τ_i 之后，结合模型参数 $\mathcal{W}_i$ 计算 $\hat{w}_i \leftarrow w_i \cdot \tau_i$ ，并将加密参数 $\text{Enc}_{\text{pk}}(\hat{w}_i)$ 和 $\text{Enc}_{\text{pk}}(\tau_i)$ 上传到服务器 S_0 。

算法 1 PPRFL 框架在用户端的执行算法

初始化 用户持有公-私钥对 (pk,sk)

输入 用户持有 $\text{Enc}_{\text{pk}}(\bar{w})$

输出 用户返回 $\text{Enc}_{\text{pk}}(\hat{w}_i)$ 和 $\text{Enc}_{\text{pk}}(\tau_i)$

- 1) 对于每一个用户 $i \in [1, n]$
- 2) 从服务器端下载初始化加密模型，使用 sk 解密得到模型参数 $\bar{w}$
- 3) 在每一轮中执行 SGD，以式(2)方式更新模型参数 $\mathcal{W}_i$
- 4) 以式(4)方式在组合验证集计算损失值 loss_i^E
- 5) 以式(5)方式计算模型可靠性 $u_i^E = \text{loss}_i^E + \gamma u_i^{E-1}$
- 6) 以式(6)方式计算缩放后的模型可靠性 τ_i
- 7) 采用 $\hat{w}_i \leftarrow w_i \cdot \tau_i$ 更新模型参数
- 8) 将 $\hat{w}_i$ 和 τ_i 加密为 $\text{Enc}_{\text{pk}}(\hat{w}_i)$ 和 $\text{Enc}_{\text{pk}}(\tau_i)$ ，并上传到服务器 S_0
- 9) 进入下一轮迭代

3.2 模型聚合更新

本节介绍 PPRFL 框架在服务器端进行模型聚合更新的执行算法，具体步骤如下。

初始化 服务器 S_0 持有公钥 pk，服务器 S_1 持有公-私钥对 (pk,sk)

输入 用户发送 $\text{Enc}_{\text{pk}}(\hat{w}_i)$ 、 $\text{Enc}_{\text{pk}}(\tau_i)$ 到 S_0

输出 S_0 返回 $\text{Enc}_{\text{pk}}(\bar{w})$

- 1) S_0 接收 $\text{Enc}_{\text{pk}}(\hat{w}_i)$ 和 $\text{Enc}_{\text{pk}}(\tau_i)$
- 2) 对于每一个 $i \in [1, n]$

- 3) S_0 计算 $\text{Enc}_{\text{pk}}(\tau) \leftarrow \prod_{i=1}^n \text{Enc}_{\text{pk}}(\tau_i)$
- 4) S_0 计算 $\text{Enc}_{\text{pk}}(\hat{w}) \leftarrow \prod_{i=1}^n \text{Enc}_{\text{pk}}(\hat{w}_i)$
- 5) S_1 辅助 S_0 执行安全除法协议, 以式(7)方式计算密文下的全局模型参数 $\text{Enc}_{\text{pk}}(\bar{w}) \leftarrow \text{Enc}_{\text{pk}}\left(\frac{\hat{w}}{\tau}\right)$
- 6) S_0 下发 $\text{Enc}_{\text{pk}}(\bar{w})$ 给所有用户
- 7) 返回 1) 进入下一轮迭代

其中安全除法协议的构造详见附录。

S_0 首先聚合用户的加密模型可靠性 $\text{Enc}_{\text{pk}}(\tau_i)$ 和加密模型参数 $\text{Enc}_{\text{pk}}(\hat{w}_i)$, 进而得到含有权重的加密模型参数 $\text{Enc}_{\text{pk}}(\hat{w})$ 。接下来, S_0 和 S_1 协同执行安全除法协议得到归一化的加密全局模型参数 $\text{Enc}_{\text{pk}}(\bar{w})$, 如式(7)所示。易见, 某一用户的可靠性越高, 该用户的模型参数在聚合后将占有越大的比例。最后, S_0 下发 $\text{Enc}_{\text{pk}}(\bar{w})$ 给所有用户, 用户可以使用 sk 解密得到全局模型参数 $\bar{w}$ 。达到终止条件后, 算法停止迭代。

$$\bar{w} = \frac{\sum_{i=1}^n \tau_i w_i}{\sum_{i=1}^n \tau_i} \quad (7)$$

4 理论分析

4.1 安全性分析

命题 1 PPRFL 的安全性是服务器 S_0 和 S_1 无法获知任意用户的模型信息和协议交互过程中产生的明文结果。

证明 本文通过 Hybrid Argument^[8]模型来证明命题 1。给定安全参数 λ 、 n 个用户以及两个非共谋服务器 S_0 和 S_1 , 采用 δ_{real} 表示 PPRFL 框架中 S_0 和 S_1 的真实视图。另外, 证明存在一个概率多项式时间敌手 $\mathcal{A}$, 对于用户、 S_0 和 S_1 而言, $\mathcal{A}$ 的输出 δ_{sim} 与 δ_{real} 在计算上是不可区分的。

假设 1 初始化一个与真实视图 δ_{real} 同分布的变量。

假设 2 在这个过程中, 本文改变每个用户的行为, 每个用户在 Paillier 密码系统下使用公钥

pk_i 加密随机数 η_i^a 而不是模型参数 w_i^a 。 S_0 和 S_1 不共谋的假设以及 Paillier 密码系统的安全性可以确保假设 2 与假设 1 是不可区分的。

假设 3 模拟 S_0 计算 $\text{Enc}_{\text{pk}}(\lambda) \leftarrow \prod_{i=1}^n \text{Enc}_{\text{pk}}(\lambda_i)$

和 $\text{Enc}_{\text{pk}}(\hat{\eta}) \leftarrow \prod_{i=1}^n \text{Enc}_{\text{pk}}(\hat{\eta}_i)$ 而不是 $\text{Enc}_{\text{pk}}(\tau) \leftarrow$

$\prod_{i=1}^n \text{Enc}_{\text{pk}}(\tau_i)$ 和 $\text{Enc}_{\text{pk}}(\hat{w}) \leftarrow \prod_{i=1}^n \text{Enc}_{\text{pk}}(\hat{w}_i)$ 。 S_0 和 S_1

不共谋的假设以及 Paillier 密码系统的安全性可以确保 δ_{sim} 与 δ_{real} 在计算上是不可区分的。因此, 假设 3 与假设 2 是不可区分的。

假设 4 使用 $\text{Enc}_{\text{pk}}(\hat{\eta}^a)$ 和 $\text{Enc}_{\text{pk}}(\lambda)$ 而不是 $\text{Enc}_{\text{pk}}(\tau)$ 和 $\text{Enc}_{\text{pk}}(\hat{w})$ 作为安全除法协议的输入。

由于附录的命题 2 证明了安全除法协议的安全性, 可知 S_0 和 S_1 在计算上无法区分 δ_{sim} 与 δ_{real} 。

因此, 假设 4 与假设 3 也是不可区分的。

如上所示, 利用 Paillier 密码系统和除法协议的安全性, 可以得到概率多项式时间敌手 $\mathcal{A}$ 的输出 δ_{sim} , 在计算上与 δ_{real} 不可区分。而该不可区分性保证 S_0 和 S_1 在这些交互中无法从 δ_{sim} 获得关于任何 δ_{real} 的信息。

证毕。

4.2 功能性分析

表 1 比较了 SecProbe^[17]、SAHPP^[19]、PPFDL^[21] 以及本文 PPRFL 框架的功能性, 4 种方案均能保护用户上传的模型参数信息。运行在半可信服务器环境下的 SecProbe 需要在线评估模型效用并进行模型聚合, 这会造成可靠性隐私和聚合模型参数隐私泄露的问题, 并且 SecProbe 采用的差分隐私方法不支持用户中途退出。SAHPP、PPFDL 以及本文方案采用同态加密体系, 对中间过程参数和聚合结果都进行保护, 服务器无法得知用户可靠性和聚合参数结果。此外, SecProbe、SAHPP 和 PPFDL 在设计之初考虑的是低质量数据不超过全局总数据量一半的情况, 在噪声数据占比过多时将造成聚合效率降低的问题。而本文的 PPRFL 框架依托组合验证集上的损失值, 并采用动态更新和缩放方式计算模型可靠性, 使本文方案在大多数不规则用户环境下仍能保持较高的模型收敛效率。

表1 不同方案的功能性比较
Table 1 Comparison of functionality among different schemes

方案	保护模型参数	保护用户可靠性	保护聚合参数	用户退出鲁棒性	模型收敛效率
ecProbe ^[17]	√	×	×	×	×
SAHPP ^[19]	√	√	√	√	×
PPFDL ^[21]	√	√	√	√	×
PPRFL	√	√	√	√	√

注：“√”表示满足对应功能，“×”表示不满足。

5 性能分析

本节从模型收敛速度角度出发，详细分析实验结果以评估 PPRFL 框架的可应用性。实验采用 Python 环境并使用 MNIST 数据集^[37]，该数据集包含 55 000 个训练集样本、5 000 个验证集样本以及 10 000 个测试集样本。由于方案需要用到密码原语，使用 Paillier 库模拟同态加密过程，并利用 Matplotlib 数据绘图包输出实验结果。为生成用户数据，本文将 MNIST 训练集划分成等量的 i 份，对于用户 i 持有的数据集 D_i 有 $\sum |D_i| = 55\,000$ 。并且将 MNIST 验证集划分成等量的 $i+1$ 份，其中服务器持有一份验证集。

实验设备配置如下：服务器以及可信第三方，由两台服务器模拟，其配置分别为 Intel Core i5-10400 CPU @ 2.90GHz 和 64 GB RAM，512SSD、2TB 机械硬盘；所有用户程序运行的设备配置为 Intel Core i7-6700 CPU@ 3.40GHz。此外，类似于文献[19,21]方案，用户在本地上运行相同的卷积神经网络以获得训练后的模型参数，其中采用的卷积神经网络分别包含 20 个特征图和 50 个特征图的两个 5×5 卷积层、一个平均池化层和分别包含 256、10 个神经元的完全连接层，每层使用 ReLu 激活函数，并设置和前文工作^[19,21]一致的学习速率 0.01 和选择的最小批量大小 128。

在功能性分析中，SecProbe^[17]会导致模型可靠性和聚合参数隐私泄露并且无法应对用户中途退出的情况。因此在接下来的分析中，本文只与安全性较强的 SAHPP^[19]和 PPFDL^[21]方案进行对比。

5.1 通信开销

根据 2.1 节系统模型的描述可知，用户与 S_0 、 S_0 与 S_1 之间存在通信过程，会产生相应的通信开

销。表 2 给出了 3 种方案在用户端与服务器端通信开销方面的对比，其中用户与 S_0 之间的通信包括 n 个用户上传和下载容量为 m 的模型参数。

表2 不同方案的通信复杂度对比
Table 2 Comparison of communication complexity among different schemes

方案	用户与第三方	用户与服务器 S_0	服务器 S_0 与 S_1
SAHPP ^[19]	$O(n)$	$O(n^2 + nm)$	—
PPFDL ^[21]	—	$O(nm)$	$O(\kappa(n+m))$
PPRFL	—	$O(nm)$	$O(m)$

本文 PPRFL 在服务器端需要 S_0 和 S_1 进行 m 轮交互以计算安全除法协议，其通信复杂度为 $O(m)$ ，这意味着服务器端的通信开销取决于模型参数数量，与用户数量无关。相比之下，PPFDL 通信复杂度为 $O(\kappa(n+m))$ ，与用户数量呈线性增长关系，其中 κ 为单轮聚合需要反复执行协议的次数。因此，本文 PPRFL 不仅适用于对通信能力有限的用户，并且服务器间的通信复杂度优于 PPFDL。

5.2 计算开销

由于用户端的计算能力有限，在考虑模型精度和通信开销的同时应注意用户端的计算开销。图 2 给出在上传相同的模型参数数量时，不同方案的用户端计算开销。SAHPP^[19]不仅需要密钥协商交互计算，并且需要计算本地模型与全局模型的空间距离，因此 SAHPP 方案中的用户端计算开销较高。PPFDL^[21]仅需进行模型训练与模型参数加密，计算开销相比较低。PPRFL 比 PPFDL 方案需要多执行一次前向计算和缩放模型参数计算，计算开销略高于 PPFDL，但大大增强了模型对于不规则用户比例的鲁棒性。

本文 PPRFL 框架的计算开销主要由加密乘法和除法协议构成，计算复杂度为 $O(n+nm)$ 。而 PPFDL 方案的计算开销主要由乘法协议构成，计算复杂度为 $O(2\kappa nm)$ 。如图 3 所示，在上传相同数量的梯度后，PPFDL 方案的服务器端计算开销远高于 PPRFL。SAHPP 计算开销由计算复杂度为 $O(nm)$ 的加密乘法构成，因此 PPRFL 与 SAHPP 的服务器端计算开销接近。但 PPFDL 在后续中引入安全传输技术 (STT, secure transformation technique)^[36]，使其计算复杂度降低为 $O(2\kappa(n+m))$ 。为进一步比较计算开销，本文将 STT 引入

PPRFL, 使 PPRFL 计算复杂度降低到 $O(n+m)$ 。如表 3 所示, PPRFL 的计算复杂度相比 PPFDL 较低。

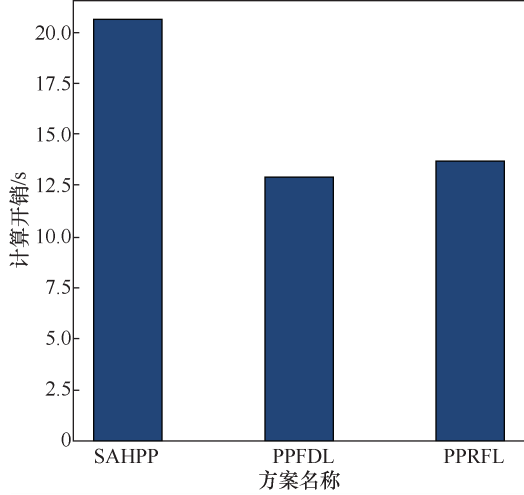


图 2 不同方案的客户端计算开销

Figure 2 Comparison of client's computation costs among different schemes

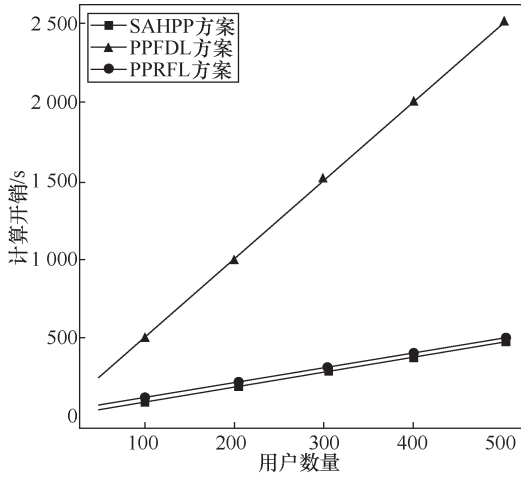


图 3 不同方案的服务器端计算开销

Figure 3 Comparison of server's computation costs among different schemes

表 3 不同方案的计算复杂度对比
Table 3 Comparison of computational complexity among different schemes

方案	采用 STT 前	采用 STT 后
SAHPP ^[19]	$O(nm)$	
PPFDL ^[21]	$O(2\kappa nm)$	$O(2\kappa(n+m))$
PPRFL	$O(n+nm)$	$O(n+m)$

5.3 收敛效率

影响模型收敛效率的因素包括不规则用户比例和不规则用户持有数据的噪声比例以及参与的

用户数量。为模拟该场景, 本文设置系统中不规则用户的比例为 P_1 , 其中 $P_1 \in [0,1]$ 。不规则用户持有的随机噪声数据量占本地总数据量的比例为 P_2 ($P_2 \in [0,1]$)。在 $P_1 \cdot P_2 \leq 50\%$ 时, SAHPP^[19] 和 PPFDL^[21] 均能有效处理不规则用户模型, 但在 $P_1 \cdot P_2 \geq 50\%$ 情况下, 过多不规则用户将使这些方案的聚合效率大大降低。而本文提出的可靠性动态更新以及权重缩放方法, 可以有效应对 $P_1 \cdot P_2 \geq 50\%$ 的情况。此外, 本文设置的收敛效率计算方式如式(8)所示, 其中 Eff 为收敛效率, $\text{Epoch}_{\text{PPFDL}}$ 和 $\text{Epoch}_{\text{PPRFL}}$ 分别为 PPFDL 和 PPRFL 收敛到预设准确度需要的通信轮数。

$$\text{Eff} = \frac{\text{Epoch}_{\text{PPFDL}}}{\text{Epoch}_{\text{PPRFL}}} \quad (8)$$

图 4 给出 $P_2 = 80\%$ 、不同不规则用户比例 P_1 , 3 种方案收敛效率的变化情况。可以观察到, 系统中持有高质量数据的用户越多, 全局模型的准确率就越高, 达到收敛需要的轮次也越低。在图 5(c) 所示 $P_1 \cdot P_2 = 64\%$ 时, SAHPP、PPFDL 和 PPRFL 方案收敛到准确率 0.92 需要的通信轮次分别为 22、21 和 10, PPRFL 收敛效率较 PPRFL 提高 2 倍。这是由于 PPRFL 能在众多不规则用户模型中识别出精度较高的模型, 赋予其更高的权重保证聚合结果主要来源于这些模型, 从而加快收敛速度。

当全局用户都为不规则用户时, 所有模型准确率较低, 各个方案达到收敛需要的轮次更多。如图 4(d) 所示 $P_1 = 100\%$ 时, SAHPP、PPFDL 和 PPRFL 方案收敛到预设准确率需要的通信轮次分别为 25、25 和 17, PPRFL 收敛效率较 PPRFL 提高 1.4 倍。基于中心梯度距离计算的 SAHPP、PPFDL 无法识别出哪些模型具有相对较好的效果, 方案效率降低。PPRFL 则能在这些低质量模型中提高相对优质模型的影响力, 减少需要的收敛轮次。

图 5 给出在相同 P_1 ($P_1 = 80\%$)、不同 P_2 值时, 3 种方案的模型精度随着迭代次数变化曲线。随着用户持有数据的噪声比 P_2 的提高, 全局模型收敛需要的轮次更多。在图 5(d) 全局总噪声数据占比达 $P_1 \cdot P_2 = 80\%$ 的情况下, SAHPP、PPFDL 和 PPRFL 方案收敛到准确率 0.89 需要的轮次分别为 32、30 和 13, PPRFL 收敛效率较 PPRFL 提高 2.3 倍。该场景只有少量的用户持有高质量数据,

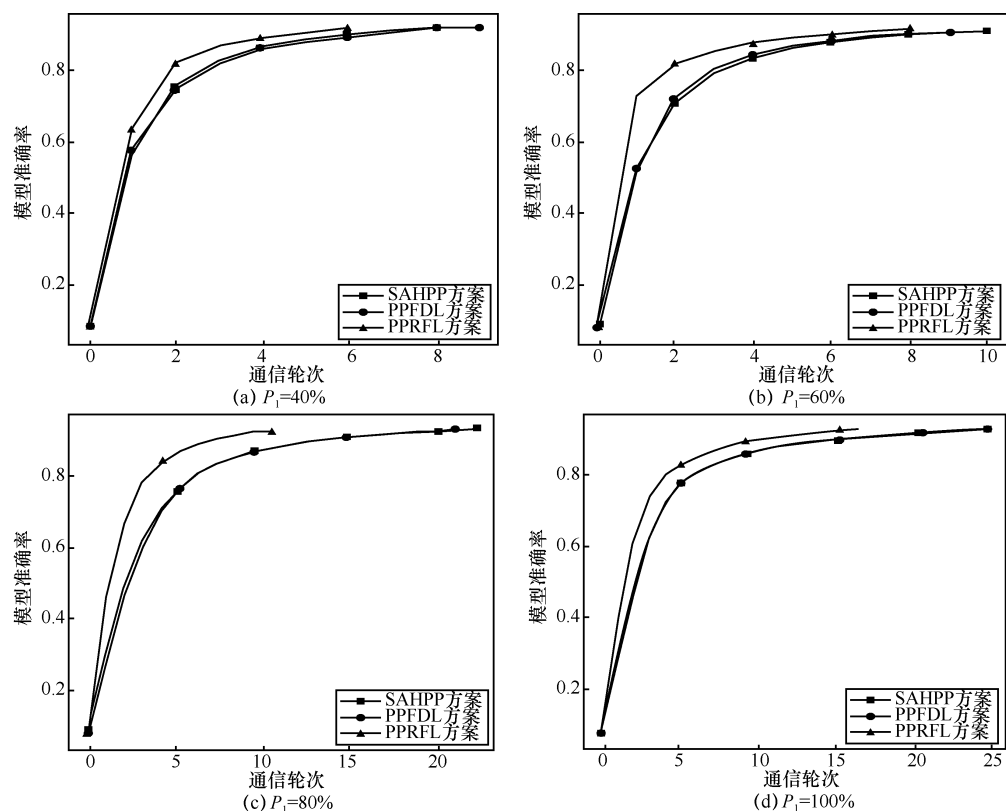


图 4 不同用户比例下模型准确率随着通信轮次的变化曲线

Figure 4 Curves showing the variations of model accuracy with number of communication rounds under different client ratios

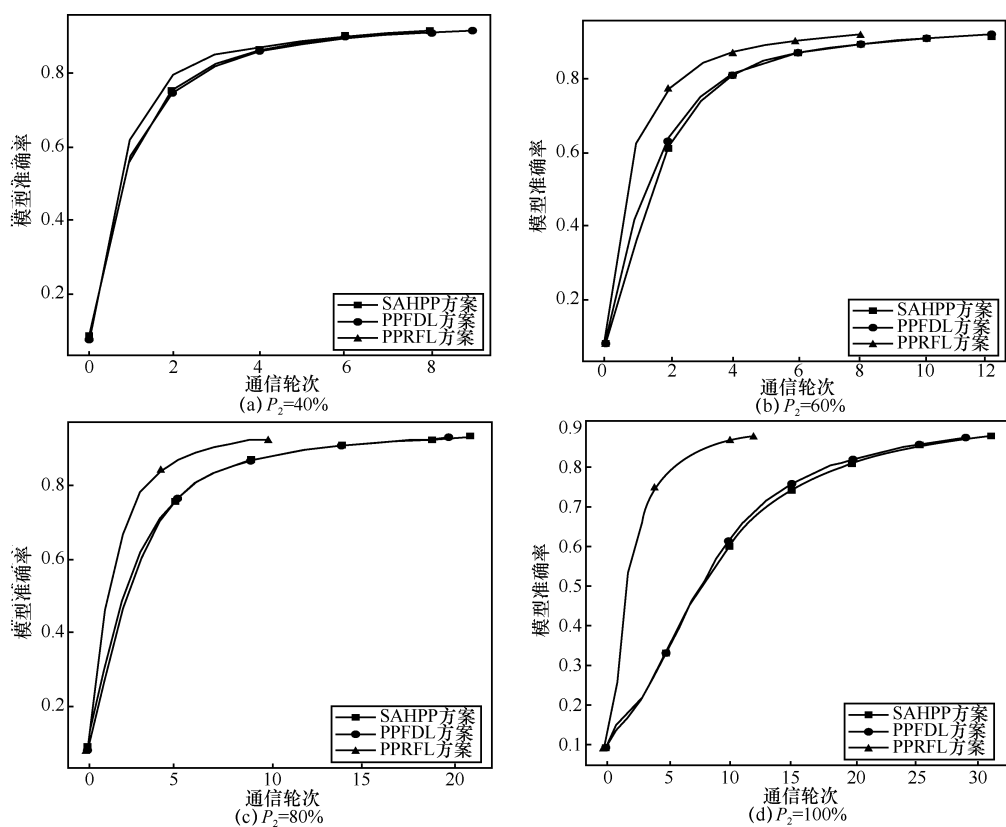


图 5 不同噪声比例下模型准确率随着通信轮次的变化曲线

Figure 5 Curves showing the variations of model accuracy with number of communication rounds under different noise ratios

PPRFL通过利用损失值历史信息更新可靠性和可靠性权重缩放等方法识别出这类用户,给予这些用户更多权重值计数,大大提高前几轮聚合的模型准确率,达到更高的收敛效率。

通常来说,参与联邦聚合的用户越多,模型的收敛速度和准确度将越高。但在该环境下,不规则用户随着参与的用户增加而增多,全局模型收敛速度受影响,收敛速度变化不规律。如图6所示,在相同 P_1 、 P_2 时达到收敛准确率0.92,PPRFL较SAHPP、PPFDL需要的收敛轮次少。这是由于在SAHPP和PPFDL中,中心梯度受不规则用户影响,高质量数据的用户梯度与平均梯度空间距离远,这类用户模型权重占比减小,收敛效率降低。而PPRFL则考虑到更多不规则用户的“质量分数”,在一定程度使聚合结果倾向于高质量数据的用户产生的模型。

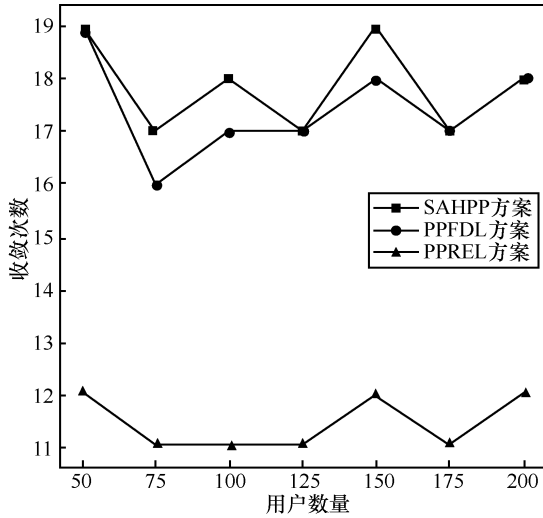


图6 相同噪声比下收敛次数随着用户数量的变化曲线
Figure 6 Curves showing the variations of convergence counts with number of clients under the same noise ratios

6 结束语

针对现有联邦学习方案受大多数不规则用户影响产生聚合效率降低、采用明文通信导致隐私信息泄露的问题,本文提出了一种支持多数不规则用户的鲁棒隐私保护联邦学习框架 PPRFL。PPRFL在用户端利用组合验证集的损失值动态计算和缩放模型可靠性;在服务器端结合设计的除法协议进行安全参数聚合。经过理论分析和实验

验证,PPRFL可以保证模型精度并提高联邦聚合的收敛效率,而且计算和通信开销均优于PPFDL。未来的工作将研究纵向联邦学习架构下处理不规则用户的方案,寻找进一步减小计算和通信开销的方法。

附录

本节详细介绍基于PHE的安全两方除法协议细节。PPFDL^[21]提出了一种由两台不共谋服务器协作执行的安全除法协议SecDiv,但混淆电路以及不经意传输进行通信需要大量额外的计算开销。为了克服这一局限性,本文提出一种更加高效的安全除法协议,可信第三方为 S_1 生成PHE的公私钥对 (pk, sk) ,并广播 pk 给 S_0 ,由 S_0 和 S_1 协作完成加密除法运算,具体构造如下。

输入 S_0 持有密文 $Enc_{pk}(x)$ 、 $Enc_{pk}(y)$ 和公钥 pk ;
 S_1 持有公-私钥对 (pk, sk)

输出 S_0 返回密文 $Enc_{pk}(x/y)$

- 1) S_0 随机选择 h_1 和 h_2
- 2) S_0 计算 $c_0 \leftarrow Enc_{pk}(x)^{h_1}$ 和 $c_1 \leftarrow Enc_{pk}(y)^{h_2}$, 并发送 c_0 和 c_1 给 S_1
- 3) S_1 计算 $d_0 \leftarrow Dec_{sk}(c_0)$ 、 $d_1 \leftarrow Dec_{sk}(c_1)$ 和 $c_2 \leftarrow Enc_{pk}(d_0/d_1)$, 并发送 c_2 给 S_0
- 4) S_0 计算 $Enc_{pk}(x/y) \leftarrow c_2^{h_2/h_1}$

安全除法协议的安全性证明如命题2所述。

命题2 基于PHE的安全除法协议中, S_0 和 S_1 均无法获取明文值 x 、 y 和 x/y 。

证明 协议安全性依赖于PHE的安全性, S_0 仅知道密文 $Enc_{pk}(x)$ 、 $Enc_{pk}(y)$ 和 $Enc_{pk}(x/y)$, 不知道私钥 sk , 无法解密恢复出明文值 x 、 y 和 x/y 。依赖于随机独立分布的计算安全, S_1 仅知道 $h_1 \cdot x$ 、 $h_2 \cdot y$ 和 $h_1 \cdot x/h_2 \cdot y$, 无法获取 h_1 和 h_2 , 不能推测出 x 、 y 和 x/y 。

证毕。

本文安全除法协议与SecDiv协议^[21]的计算开销对比如表4所示。记SecDiv构造混淆电路的计算时间为 T_{OT} 、运行不经意传输的计算时间为 T_{GC} , 一次密文乘密文计算时间记为CMult, 一次明文乘密文计算时间记为Mult, 一次加密计算时间记为Enc, 一次解密计算时间记为Dec。本文提出的安全除法协议仅比SecDiv多进行一次明文乘法而不需要执行开销更高的 T_{OT} 和 T_{GC} 。

表4 不同除法协议的计算开销对比
Table 4 Comparison of computational overhead among different protocol

协议	T_{OT}	T_{GC}	CMult	Mult	Enc	Dec
SecDiv ^[21]	1	1	3	—	4	2
本文协议	—	—	3	1	3	2

在通信开销方面, SecDiv 需要 2 次通信传输同态密文, 且需要执行通信开销较高的混淆电路和不经意传输操作, 而本文提出的安全除法协议仅需要 2 次通信传输同态密文, 通信开销远低于 SecDiv 协议。

参考文献:

- [1] 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.
- [2] YANG Q, LIU Y, CHEN T J, et al. Federated machine learning[J]. ACM Transactions on Intelligent Systems and Technology, 2019, 10(2): 1-19.
- [3] 梁应敞, 谭俊杰, Dusit Niyato. 智能无线通信技术研究概况[J]. 通信学报, 2020, 41(7): 1-17.
LIANG Y C, TAN J J, NIYATO D. Overview on intelligent wireless communication technology[J]. Journal on Communications, 2020, 41(7): 1-17.
- [4] 杨强. AI 与数据隐私保护: 联邦学习的破解之道[J]. 信息安全研究, 2019, 5(11): 961-965.
YANG Q. AI and data privacy protection: the way to federated learning[J]. Journal of Information Security Research, 2019, 5(11): 961-965.
- [5] 谭清尹, 曾颖明, 韩叶, 等. 神经网络后门攻击研究[J]. 网络与信息安全学报, 2021, 7(3): 46-58.
TAN Q Y, ZENG Y M, HAN Y, et al. Survey on backdoor attacks targeted on neural network[J]. Chinese Journal of Network and Information Security, 2021, 7(3): 46-58.
- [6] MELIS L, SONG C Z, DE CRISTOFARO E, et al. Exploiting unintended feature leakage in collaborative learning[C]//Proceedings of 2019 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2019: 691-706.
- [7] 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. Survey of federated learning research[J]. Chinese Journal of Network and Information Security, 2021, 7(5): 77-92.
- [8] HITAJ B, ATENIESE G, PEREZ-CRUZ F. Deep models under the GAN: information leakage from collaborative deep learning[C]//Proceedings of CCS '17: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. 2017: 603-618.
- [9] MOTHUKURI V, PARIZI R M, POURIYEH S, et al. A survey on security and privacy of federated learning[J]. Future Generation Computer Systems, 2021, 115: 619-640.
- [10] WAGH S, GUPTA D, CHANDRAN N. SecureNN: 3-party secure computation for neural network training[J]. Proceedings on Privacy Enhancing Technologies, 2019, 2019(3): 26-49.
- [11] XU R H, BARACALDO N, ZHOU Y, et al. HybridAlpha: an efficient approach for privacy-preserving federated learning[C]//Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security. 2019: 13-23.
- [12] 方晨, 郭渊博, 王一丰, 等. 基于区块链和联邦学习的边缘计算隐私保护方法[J]. 通信学报, 2021, 42(11): 28-40.
FANG C, GUO Y B, WANG Y F, et al. Edge computing privacy protection method based on blockchain and federated learning[J]. Journal on Communications, 2021, 42(11): 28-40.
- [13] KANAGAVELU R, LI Z X, SAMSUDIN J, et al. Two-phase multi-party computation enabled privacy-preserving federated learning[C]//Proceedings of 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID). 2020: 410-419.
- [14] 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.
- [15] PHONG L T, AONO Y, HAYASHI T, et al. Privacy-preserving deep learning via additively homomorphic encryption[J]. IEEE Transactions on Information Forensics and Security, 2018, 13(5): 1333-1345.
- [16] 张泽辉, 富瑶, 高铁杠. 支持数据隐私保护的联邦深度神经网络模型研究[J]. 自动化学报, 2020.
ZHANG Z H, FU Y, GAO T G. Research on federated deep neural network model for data privacy protection[J]. Acta Automatica Sinica, 2020.
- [17] ZHAO L C, WANG Q, ZOU Q, et al. Privacy-preserving collaborative deep learning with unreliable participants[J]. IEEE Transactions on Information Forensics and Security, 2020, 15: 1486-1500.
- [18] JAYARAMAN B, EVANS D. Evaluating differentially private machine learning in practice[C]//Proceedings of the 28th USENIX Security Symposium. 2019: 1895-1912.
- [19] 成艺. 联合学习环境下保护隐私的数据聚合技术研究[D]. 成都: 电子科技大学, 2020: 17-45.
CHENG Y. Research on data aggregation technology based on privacy-preserving in federated learning[D]. Chengdu: University of Electronic Science and Technology of China, 2020: 17-45.
- [20] SHAMIR A. How to share a secret[J]. Communications of the ACM, 1979, 22(11): 612-613.
- [21] XU G W, LI H W, ZHANG Y, et al. Privacy-preserving federated deep learning with irregular users[J]. IEEE Transactions on Dependable and Secure Computing, 2020, (99): 1.
- [22] ZHENG Y F, DUAN H Y, WANG C. Learning the truth privately and confidently: encrypted confidence-aware truth discovery in mobile crowdsensing[J]. IEEE Transactions on Information Forensics and Security, 2018, 13(10): 2475-2489.
- [23] TIAN Y L, LI T, XIONG J B, et al. A blockchain-based machine learning framework for edge services in IoT[J]. IEEE Transactions on Industrial Informatics, 2022, 18(3): 1918-1929.
- [24] XIONG J B, BI R W, ZHAO M F, et al. Edge-assisted privacy-preserving raw data sharing framework for connected autonomous

- vehicles[J]. IEEE Wireless Communications, 2020, 27(3): 24-30.
- [25] BONAWITZ K, EICHNER H, GRIESKAMP W, et al. Towards federated learning at scale: system design[J]. arXiv preprint arXiv:1902.01046, 2019.
- [26] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data[C]//Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. 2017: 1273-1282.
- [27] PAILLIER P. Public-key cryptosystems based on composite degree residuosity classes[C]// Proceedings of the International Conference on the Theory and Applications of Cryptographic Techniques. 1999: 223-238.
- [28] ACAR A, AKSU H, ULUAGAC A S, et al. A survey on homomorphic encryption schemes[J]. ACM Computing Surveys, 2019, 51(4): 1-35.
- [29] CANETTI R, FEIGE U, GOLDREICH O, et al. Adaptively secure multi-party computation[C]//Proceedings of the twenty-eighth annual ACM Symposium on Theory of Computing. 1996: 639-648.
- [30] HENDERSON M, THOMSON B, WILLIAMS J D. The third dialog state tracking challenge[C]//Proceedings of 2014 IEEE Spoken Language Technology Workshop. 2014: 324-329.
- [31] BUDZIANOWSKI P, WEN T H, TSENG B H, et al. MultiWOZ - A large-scale multi-domain wizard-of-oz dataset for task-oriented dialogue modelling[C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018: 5016-5026.
- [32] SHOKRI R, SHMATIKOV V. Privacy-preserving deep learning[C]//Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. 2015: 1310-1321.
- [33] KANG J W, XIONG Z H, NIYATO D, et al. Toward secure blockchain-enabled Internet of vehicles: optimizing consensus management using reputation and contract theory[J]. IEEE Transactions on Vehicular Technology, 2019, 68(3): 2906-2920.
- [34] KANG J W, XIONG Z H, NIYATO D, et al. Incentive mechanism for reliable federated learning: a joint optimization approach to combining reputation and contract theory[J]. IEEE Internet of Things Journal, 2019, 6(6): 10700-10714.
- [35] KANG J W, XIONG Z H, NIYATO D, et al. Reliable federated learning for mobile networks[J]. IEEE Wireless Communications, 2020, 27(2): 72-80.
- [36] CATALANO D, FIORE D. Using linearly-homomorphic encryption to evaluate degree-2 functions on encrypted data[C]//Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. 2015: 1518-1529.
- [37] 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]



陈前昕（1996—），男，福建泉州人，福建师范大学硕士生，主要研究方向为安全深度学习、隐私保护技术。



毕仁万（1996—），男，湖南常德人，福建师范大学博士生，主要研究方向为安全深度学习、安全多方计算。



林劼（1972—），男，福建三明人，博士，福建师范大学教授，主要研究方向为机器学习、生物信息学。



金彪（1985—），男，安徽六安人，博士，福建师范大学高级实验师，主要研究方向为信息安全、数据挖掘。



熊金波（1981—），男，湖南益阳人，博士，福建师范大学教授，主要研究方向为安全深度学习、移动群智感知、隐私保护技术。