兼顾通信效率与效用的自适应高斯差分隐私个性化联邦学习

李 敏 肖 迪 陈律君

(重庆大学计算机学院 重庆 400044)

摘 要 近年来,由于联邦学习中的通信参数(或梯度)会给参与方本地敏感数据带来重大的隐私泄露风险,联邦学习隐私保护引起了广泛的关注.然而,梯度交换频繁、数据分布异构、参与方本地硬件资源受限等一系列不可避免的因素给联邦学习隐私保护增加了挑战难度.为了以一种统一的方式同时有效地解决数据隐私、模型效用、通信效率以及参与方数据非独立同分布等四个方面的问题,本文提出了一种新的兼顾通信效率与效用的自适应高斯差分隐私个性化联邦学习(Communication-efficient and Utility-aware Adaptive Gaussian Differential Privacy for Personalized Federated Learning, CUAG-PFL)方法.具体而言,本文提出一种动态层级压缩模型梯度的方案先为通信模型梯度每一层动态生成特定的压缩率,再根据压缩率构造对应的确定性二进制测量矩阵去除梯度冗余信息.随后,通过同时优化裁剪阈值、敏感度和噪声尺度等隐私相关参数来对压缩的模型梯度执行自适应高斯差分隐私操作.此外,本文对CUAG-PFL进行了严格的隐私分析.为了验证CUAG-PFL在隐私、效用、通信效率以及个性化四个方面的优势,本文在CIFAR-10和CIFAR-100两个真实联邦数据集上进行了大量实验模拟、对比和分析,结果表明CUAG-PFL能够提高参与方本地数据隐私性、通信效率和模型效用,同时解决了数据非独立同分布的问题.特别地,即使在隐私预算仅为0.92且上行通信量减少68.6%时,CUAG-PFL因隐私保护和梯度压缩所引起的模型效用损失仅为1.66%.

关键词 自适应高斯差分隐私;隐私—效用权衡;动态层级压缩;通信高效;个性化联邦学习;隐私计算

中图法分类号 TP309 **DOI号** 10.11897/SP.J.1016.2024.00924

Communication-efficient and Utility-Aware Adaptive Gaussian Differential Privacy for Personalized Federated Learning

LI Min XIAO Di CHEN Lü-Jun

(*College of Computer Science, Chongqing University, Chongqing* 400044)

Abstract In recent years, there has been an increasing focus on the privacy protection in the field of federated learning (FL). This widespread attention is mainly due to the fact that communication parameters (or gradients) during the process of collaborative learning among the central server and various participants can cause the significant risk of the privacy leakage. In other words, the communication process in the FL system poses a potential threat of exposing the sensitive data belonging to local participants, which has raised heightened concerns among researchers and practitioners. Furthermore, in addition to the challenge of the privacy protection in FL, a series of other unavoidable factors such as the frequent gradients exchange, the heterogeneous data distribution among local participants, and limited resources available on the local hardware need to be simultaneously taken into consideration. These factors obviously add diffi-

收稿日期:2023-06-30;在线发布日期:2024-01-16. 本课题得到国家重点研发项目(No. 2020YFB1805400)、国家自然科学基金(No. 62072063)、重庆市研究生科研创新项目(No. CYB22063)资助. **李 敏**,博士研究生,主要研究领域为联邦学习、隐私保护、压缩感知和差分隐私. E-mail:minli@cqu.edu.cn. **肖 迪(通信作者)**,博士,教授,中国计算机学会(CCF)高级会员,主要研究领域为隐私保护和数据安全. E-mail:dixiao@cqu.edu.cn. **陈律君**,博士研究生,主要研究领域为隐私保护、安全多方计算和联邦学习.

culties to the challenge of the privacy protection in FL. In order to effectively address four critical issues of data privacy, model utility, communication efficiency, and non-independently and identically distributed data among local participants in a unified manner, this paper proposes a novel Communication-efficient and Utility-aware Adaptive Gaussian Differential Privacy for Personalized FL method, called CUAG-PFL. Specifically, a dynamic layer-compression scheme for model gradients in the FL system is proposed. This scheme aims to improve the communication efficiency as much as possible and reduce the loss of the model utility caused by compression and reconstruction through dynamically customizing the compression rate for each layer of communication gradients, and then constructing the corresponding deterministic binary measurement matrix based on the compression rate. This designed deterministic binary measurement matrix can effectively remove the redundant information of model gradients that needs to be uploaded to the central server. Subsequently, the adaptive Gaussian differential privacy operation is performed on compressed model gradients of local participants. This operation involves optimizing the main privacy-related parameters such as the clipping threshold, the sensitivity, and the noise scale. By optimizing these parameters at the same time, this operation ensures that the privacy of the local data is preserved, while allowing each model of the corresponding local participant to have the satisfactory performance. In addition, the rigorous privacy analysis of the proposed CUAG-PFL is presented in this paper. In order to validate the superiority of the proposed CUAG-PFL in four critical aspects of data privacy, model utility, communication efficiency, and personalization, a large number of experimental simulations, comparisons, and analyses are conducted on two classic real-world federated datasets, i. e., CIFAR-10 and CIFAR-100. All experimental results and analyses show that the proposed CUAG-PFL can simultaneously improve the privacy of local sensitive data, the communication efficiency and the model utility, as well as address the problem of non-independently and identically distributed data among local participants in the FL system. In particular, it is worth emphasizing that even when the privacy budget is only 0.92 and the amount of the upstream communication is reduced by 68.6%, the loss of the model performance caused by both the privacy protection and the communication gradients compression is just 1.66% for the proposed CUAG-PFL.

Keywords　adaptive Gaussian differential privacy; privacy-utility trade-off; dynamic hierarchical compression; high-efficient communication; personalized federated learning; private computing

1　引　言

信息时代的发展使人工智能领域呈现出爆炸式增长,但是传统的机器学习过于依赖中央服务器对大规模移动设备数据的存储汇总和计算,忽略了与个人密切相关的数据隐私性和安全性.近期,各国政府陆续颁布和实施一系列法律法规来规范数据的商业使用以便为个人数据隐私提供法律层面的根本保障[1],例如《通用数据保护条例》(General Data Protection Regulation, GDPR)、《中华人民共和国数据安全法》和《中华人民共和国个人信息保护法》.受法规、监管和技术等的限制,各行各业数据无法共享和交换,信息化社会逐渐发展成数据孤岛状态,制约了大数据驱动的人工智能应用发展.联邦学习(Federated Learning, FL)[2]作为一种同时兼顾数据安全和数据融合的新分布式机器学习范式,通过将模型训练分发至数据来源的远程设备中以使得人工智能可惠及社会个人数据敏感的应用领域,其核心原理是数据可用不可见,即数据不动模型动.目前,FL 已被广泛应用于政务开放、医疗健康、交通运输和金融风控等公共领域[3-5].

尽管 FL 避免了直接暴露本地数据,对于数据隐私具备天然的保护作用,但是依然存在大量隐私泄漏的风险.例如,已有研究工作表明一个技术足够熟练的攻击者可以通过 FL 中通信的模型参数(或

梯度)来还原部分的敏感信息甚至推断掌握的数据是否源于某个特定的参与方[6-7].除此之外,切实将FL落地于现实世界中面临着其它诸多挑战[8],主要包括服务器和参与方之间模型参数(或梯度)传输引起的巨大通信开销,参与方本地设备在计算存储能力等方面的硬件资源限制以及参与方之间的本地数据往往是非独立同分布的(Non-Independent and Identically Distributed, Non-IID).

目前,主流的FL隐私保护研究方法是将经典的机器学习隐私保护技术融入FL中[9-13].相比于同态加密、秘密共享等传统密码学技术,差分隐私(Differential Privacy, DP)[14]因为强大的信息理论保障、简单灵活且易于部署的算法以及相对较少的系统开销更受相关领域学者与机构的青睐.尤其是针对大规模FL系统,只添加少量的噪声便可实现高度的隐私保护.除此之外,得益于DP后处理性,模型参数(或梯度)可被持续地保护.为了抵御不可信的服务器引发的重构攻击,能够为参与方数据提供更严格保障的本地差分隐私(Local Differential Privacy, LDP)[15]被广泛应用于FL中.然而,这类加噪的隐私保护机制通常需要在参与方数据隐私与模型效用之间进行权衡.目前,绝大多数基于LDP的FL研究工作都固定裁剪阈值、敏感度以及噪声尺度,即每轮全局通信时各个参与方为模型参数(或梯度)注入相同的噪声方差[16-23].这些研究工作在一定程度上为参与方数据提供LDP保证,但是存在模型效用低、容易受到梯度泄漏攻击等情况.为了保障参与方本地数据的隐私,一些前沿的研究工作采用固定隐私预算思想或动态裁剪阈值策略来设计基于LDP的FL方法[24-26],但是隐私预算和噪声标准差成反比的固有特性以及粗糙的裁剪阈值策略使得共享的全局模型性能下降.为了同时提高基于LDP的FL隐私性和效用性,这里本文引入一种可分析、易处理且高精确的高斯差分隐私(Gaussian Differential Privacy, GDP)[27-28].相比于基于高斯机制的传统DP,GDP只需添加更少量的噪声便能为参与方本地数据提供相同的隐私保障,从而减少对模型性能退化的影响[28].因此,本文设计了一种自适应GDP策略来弥补使用固定裁剪阈值、敏感度和噪声尺度等隐私相关参数策略的固有局限性.

遗憾的是,由于现实世界中各个参与方的本地数据收集方式和目的不同,因此本质上FL系统中参与方之间的本地数据是Non-IID.这种数据异构问题使得所有参与方难以协同训练一个可以与各方全部数据完美契合的共享全局模型,进而导致整个FL系统收敛缓慢且模型效用不佳.区别于传统FL,个性化FL通过联合所有参与方的共同知识和各方的本地数据特征来为每一个参与方定制一个更好地拟合其独特数据分布的个性化模型,以减轻参与方数据异构问题对整个FL系统性能退化的影响.在这种情况下,基于LDP的个性化FL研究工作被相继提出,主要典型代表包括DP-SCAFFOLD[25]、PPSGD[29]和PRIVATEFL[30].然而,这些研究工作都未考虑FL通信效率的问题.尤其是在DP-SCAFFOLD方法中,每一轮全局通信的服务器和参与方之间则需要同时传递模型参数更新和梯度,翻倍增加了通信开销.

综上所述,目前FL隐私保护的研究工作主要面临的三个核心挑战是隐私性、效用性和通信效率.具体而言,基于秘密共享等传统密码学技术的FL隐私保护方法通常会引入额外的通信开销以及计算成本,减缓模型收敛速度,而提高通信效率又可能会损失模型效用.尽管基于LDP的FL方法通信成本相对较低,但是往往通过牺牲模型性能来提高参与方数据的隐私性.由此可知,隐私性、效用性和通信效率三个挑战两两之间相互影响,所以如何权衡隐私性—效用性—通信效率则是FL隐私保护研究的核心工作.此外,参与方本地数据Non-IID这个现实条件加剧了以上三个方面的挑战.即使个性化FL能够一定程度上减轻这种数据异构问题对整个FL系统性能退化的影响,但是可能增加隐私暴露风险,甚至需要更多通信开销.因此,为了探索一种新的面向参与方本地数据是Non-IID的高效率、高效用FL隐私保护方法来促进FL实用化,本文提出了一种兼顾通信效率与效用的自适应GDP个性化FL(Communication-efficient and Utility-aware Adaptive GDP for Personalized FL, CUAG-PFL)方法.本文的主要贡献如下:

(1)提出CUAG-PFL以同时有效解决隐私性、效用性、通信效率和参与方数据Non-IID四个问题.

(2)提出了一种动态层级压缩模型梯度的方案,既可以最大限度保留有效梯度信息和去除冗余信息以提高通信效率和模型效用,同时还能够根据通信资源受限的FL系统具体需求动态调整每一轮全局通信的梯度总量.

(3)设计了一种新的确定性二进制测量矩阵来压缩通信梯度,相比于高斯随机矩阵等常用的测量矩阵减少了本地计算量,避免了额外的通信开销和

本地存储负担.

(4)从裁剪阈值、敏感度和噪声尺度三个方面同时优化隐私相关参数,本文进而提出一种自适应GDP策略同时提高参与方本地数据的隐私性和模型效用性,尤其是本文通过最大化压缩梯度所暗含信息的利用率来设计层级裁剪阈值操作.除此之外,本文对CUAG-PFL进行了严格的隐私分析.

(5)基于CIFAR-10和CIFAR-100两种真实数据集,在六种Non-IID数据分割方式下验证了CUAG-PFL的有效性.大量实验模拟、对比和分析证明了CUAG-PFL能够提高本地数据隐私性、通信效率和模型效用,同时解决了FL中数据Non-IID问题.

本文第2节从隐私—效用权衡、个性化和通信效率三个方面分别对相关研究工作进行概述,并且相应地阐述了本文的研究动机.第3节介绍了预备知识.第4节给出了个性化FL的问题定义和本文需要强调的威胁模型.第5节对本文提出的CUAG-PFL进行了系统化的描述.第6节对CUAG-PFL进行了隐私分析.第7节通过大量的对比实验验证了CUAG-PFL在隐私性、效用性、通信效率以及个性化四个方面的优势.最后,第8节总结了全文.

2　相关工作

本文主要关注FL中隐私保护、通信效率和Non-IID数据等核心实用化挑战,因此下面将从这几个方面简要地介绍相关研究工作并分析本文提出CUAG-PFL的研究动机.

2.1　联邦学习隐私保护

因为FL系统中服务器和参与方之间双向通信的模型参数(或梯度)本质上暗含可以在隐私攻击中被利用的敏感信息,显然只分离数据和模型无法提供合理的数据隐私保证[6-7].为了抵御隐私泄露的威胁,同态加密、秘密共享和DP已被广泛用于设计各种FL隐私保护方法.然而,通常认为,同态加密计算开销巨大且无法处理复杂运算,秘密共享通信开销大且计算效率低[11-13].DP[14]作为一种公认的隐私保护技术,只需通过添加噪声来扰动敏感数据便可保证对各个参与方本地任意数据的合理否认,甚至是对FL中任意参与方的合理否认.

2.2　基于本地差分隐私的联邦学习

目前,基于LDP的FL研究工作主要致力于解决诚实但好奇的服务器或参与方的存在所导致的本地数据隐私泄露问题.诚实但好奇的(即半诚实的)意指攻击者严格地遵循和执行FL协议,但是试图通过模型重构攻击或成员推理攻击观察捕获的模型参数(或梯度)来重构参与方敏感的本地数据或者推断某一个数据样本是否属于某一个参与方的本地数据集[6-7].其中,绝大多数研究工作都是固定裁剪阈值、敏感度和噪声尺度,从而整个FL系统中每一轮全局通信的模型参数(或梯度)所有层都被注入了恒定的噪声量[16-23].然而研究表明此类方法不仅模型效用低,而且易引起数据隐私泄漏[31],这促使本文分析了使用固定隐私相关参数策略的固有局限性.第一,因为不同敏感度的模型层所需添加的噪声量不同,显然对通信模型参数(或梯度)的所有层都注入相同的DP噪声量会导致聚合的全局模型效用降低.第二,随着参与方与服务器之间的通信次数愈来愈多,隐私预算逐渐增大,则参与方本地数据受到的隐私保护强度显著降低,从而容易受到梯度泄漏攻击.尽管研究工作UDP[24]和DP-SCAFFOLD[25]通过固定隐私预算或动态裁剪阈值策略来保障参与方本地数据的隐私,但是却忽略了提高模型效用.

那么如何同时提高基于LDP的FL系统隐私性和效用性?第一,区别于经典的LDP技术(如拉普拉斯机制和高斯机制),本文利用GDP为通信模型梯度提供可分析、易处理且高精确的隐私保障,从而相同隐私强度下模型效用性受噪声的影响程度减小.第二,为了最大限度地保留压缩梯度的有效信息和降低聚合梯度的重构误差,本文设计了一种自适应层级压缩梯度裁剪方案,进而通信模型各层敏感度不同.综上二者,本文提出的CUAG-PFL可同时自适应地优化裁剪阈值、敏感度以及噪声尺度三种GDP相关参数.

2.3　基于本地差分隐私的个性化联邦学习

尽管解决Non-IID数据问题的个性化FL研究已经如火如荼,但是基于LDP的个性化FL研究工作较少.Noble等人[25]直接简单粗暴地将DP应用于一种流行的个性化FL框架SCAFFOLD[32],命名为DP-SCAFFOLD.显然,相比于第2.2节中涉及的非个性化FL方法,DP-SCAFFOLD能够在Non-IID数据的FL场景下提供较优的模型效用性.然而因为SCAFFOLD方法需要在每一轮全局通信中同时交互模型参数更新和梯度,所以显著增加了DP-SCAFFOLD的通信负担.Bietti等人[29]从个性化改善隐私—效用权衡的角度出发提出PPSGD,但是参与方需同时完成本地模型和全局模型的训练以至于计算开销太大.Yang等人[30]则从

Non-IID 数据本身出发，在本地模型训练期间同时更新本地数据变换层从而减少 DP 引入的额外异构性，进而提高面向 Non-IID 数据的 FL 模型效用，但是引入了额外的本地计算量.

显然，上述三种方法都不适合参与方硬件资源受限以及 FL 通信资源有限的应用，同时亦未考虑固定裁剪阈值、敏感度和噪声尺度等隐私相关参数所造成的模型效用损失. 因此，本文的目标是提出一个高效率高效用的基于 LDP 的个性化 FL 方法.

2.4 基于本地差分隐私的高效联邦学习

为了提高基于 LDP 的 FL 系统通信效率，Shin 等人[33]试图通过随机投影对梯度实现降维来显著降低通信成本. 尽管降维之后添加的噪声量减少了，但是由于参与方随机丢弃某些维度，从而导致服务器恢复梯度时引入较大的误差，进而可能损害模型效用. Liu 等人[34]则首次尝试减轻基于 LDP 的 FL 中维度依赖问题，其中参与方仅扰动依据指数机制选择的权重最高的 k 个维度. Cui 等人[35]利用拉普拉斯机制和稀疏向量技术提出一种自适应梯度阈值选择方法来进一步改善 FL 模型的学习性能. 然而，这两种方法的梯度维度选择引入了额外的本地计算量，甚至导致 FL 通信延迟效应. 为了在梯度降维方面简单高效，Kerkouche 等人[36]和 Farokhi[37]都证实了一种简捷且易操作的压缩感知(Compressed Sensing，CS)技术能够在不损失隐私的前提下提高基于 LDP 的 FL 模型效用，但是这只是一般性的框架，缺乏精细的隐私策略保障和隐私预算分配.

目前，绝大多数仅基于 CS 提高 FL 通信效率的非隐私保护方法都只是利用固定的采样率对通信的模型参数(或梯度)进行压缩. 近期，Miao 等人[38]则根据每一轮全局通信的模型损失提出了一种自适应采样率的方案用于压缩 FL 中通信的模型参数，却忽略了模型层与层之间的信息冗余度亦是不尽相同的. 此外，在基于 LDP 的 FL 系统中，通信梯度的压缩程度直接影响到后续添加自适应 DP 的噪声总量. 因此本文提出一种动态层级压缩模型梯度方案来尽可能地去除梯度冗余信息，不仅大幅度提高了通信效率，适用于通信资源受限的 FL 应用，同时与自适应 GDP 策略相辅相成来提高模型效用性.

3 预备知识

3.1 差分隐私

DP[14]的主旨是从输出中排除对任何潜在记录的对抗性推断，然而最初基于概率提出的用于发布隐私数据的 DP 是建立在服务器完全可信的假设性前提. 因此，对于服务器不受参与方信任的情况，则需通过各个参与方本地扰动敏感数据从而提供更强的隐私保护，这便是 LDP[15].

3.1.1 基于高斯机制的传统差分隐私

定义 1. (ε,δ)-DP[14-15]. 假设一个随机机制 $M:A\to Z$，输入域是 A(即所有可能的训练数据集)且输出域是 Z(即所有可能的训练模型)，当且仅当某个参与方任意两个数据集 $D,D'\in A$ 是相邻的(即数据集 D 和 D' 仅仅存在一个数据样本不同：$\|D-D'\|_0=1$)，并对输出的任何子集 $O\subseteq Z$ 满足

$$\Pr[M(D)\in O]\leqslant e^{\varepsilon}\cdot\Pr[M(D')\in O]+\delta \tag{1}$$

则称 M 满足 (ε,δ)-DP. 其中 $\varepsilon(\varepsilon\in[0,\infty))$是指 DP 的隐私预算，用于限制 M 的隐私保障级别. $\delta(\delta\in[0,1))$表示添加 M 后，D 和 D' 输出概率比不能以 e^{ε} 为界的事件概率. 固定 δ，较大的 ε 表示隐私泄露风险更大. 当 $\delta=0$ 时，M 被称为纯 ε-DP.

定义 2. 敏感度[14]. 假设 A 是输入域且 Z 是输出域，函数 $f(\cdot):A\to Z$ 的敏感度是更改单个输入条目时函数值变化的最大量，即

$$s_f=\max_{D,D'\subseteq A,\|D-D'\|_0=1}\|f(D)-f(D')\|_{\ell} \tag{2}$$

通常情况下，$\ell=1$ 或者 $\ell=2$.

定理 1. (ε,δ)-DP 中的高斯机制[14]. 已知 s_f 是本地训练过程的敏感度、ρ 是采样率且 T 是通信轮数，为了保证深度学习模型满足 (ε,δ)-DP，所需添加的高斯机制噪声标准差 ζ 应该满足

$$\zeta=\frac{s_f\sqrt{2\rho T\ln(1/\delta)}}{\varepsilon} \tag{3}$$

3.1.2 高斯差分隐私

基于定义 1，假设 P 和 Q 分别表示 $M(D)$ 和 $M(D')$ 的概率分布. 那么攻击者的目的是试图区分 P 和 Q，即等价执行以下假设检验问题：

$$H_0:\text{输出}\sim P\quad vs.\quad H_1:\text{输出}\sim Q \tag{4}$$

直观而言，如果假设检验问题难以解决，那么隐私便会被较好地保证. Dong 等人[27]提出使用 $\alpha(\alpha\in[0,1])$级别最佳似然比检验的类型 1(当 H_0 为真时错误拒绝 H_0 的概率)和类型 2(当 H_1 为真时错误接受 H_0 的概率)错误之间的权衡作为隐私保证的度量. 如果 $\bar{\phi}$ 是对 H_0 和 H_1 进行检验的一个拒绝规则，那么 $\bar{\phi}$ 的类型 1 和类型 2 错误分别是 $\mathbb{E}_P(\bar{\phi})$ 和 1

$-\mathbb{E}_Q(\bar{\phi})$.两个概率分布 P 和 Q 之间的权衡函数 $\bar{T}:[0,1]\to[0,1]$ 定义为

$$\bar{T}(P,Q)(\alpha)=\inf_{\bar{\phi}}\{1-\mathbb{E}_Q(\bar{\phi}):\mathbb{E}_P(\bar{\phi})\leqslant\alpha\} \quad (5)$$

对于一个固定的显著性级别 α,$\bar{T}(P,Q)(\alpha)$ 是检验此级别可以达到的最小类型 2 错误.最优检验由内曼—皮尔逊引理给出,可解释为最强大的攻击者.此外,较大的权衡函数意味着相关算法的隐私程度较高.当在两个高斯分布之间定义权衡函数时,则能够获取 GDP 保障.GDP 正式定义详见定义 3.

定义 3. μ-GDP[27-28]. 假设 Φ 是标准正态分布的累积分布函数,$G_\mu:=\bar{T}(\mathrm{N}(0,1),\mathrm{N}(\mu,1))\equiv\Phi(\Phi^{-1}(1-\alpha)-\mu)$ 且 $\mu\geqslant 0$.对于任意相邻数据集 D 和 D',如果 $\bar{T}(M(D),M(D'))\geqslant G_\mu$,那么随机机制 M 满足 μ-GDP.

尽管 Abadi 等人[14]利用矩计量法这一种复杂技术改进了深度学习中应用 (ε,δ)-DP 的隐私成本分析,但是 GDP[27-28]已经被证明拥有一个可分析、易处理且高精确的隐私会计.此外,GDP 能转换为 (ε,δ)-DP(见定理 3),同时依然可提供比矩计量法更严格的隐私会计.基于中心极限定理,GDP 利用单个隐私参数 μ 便可量化深度学习中任意随机机制的隐私(见定理 4).与隐私预算 ε 类似,较小的 μ 表示隐私保障程度较强.

定理 2. GDP 中的高斯机制[27-28]. 考虑隐私发布数据集 D 的单变量统计 $f(D)$ 问题,定义 $f(\cdot)$ 的敏感度是 $s_f=sup_{D,D'}|f(D)-f(D')|$,其中上确界 sup 是针对所有相邻数据集.那么,GDP 的高斯机制 $M(D)=f(D)+\xi$ 满足 μ-GDP,其中 $\xi\sim\mathrm{N}(0,s_f^2/\mu^2)$.

定理 3. μ-GDP $\to$ (ε,δ)-DP[28]. 如果一个机制当且仅当对于所有 $\varepsilon\geqslant 0$ 的 $(\varepsilon,\delta(\varepsilon))$-DP 满足

$$\delta(\varepsilon)=\Phi\left[-\frac{\varepsilon}{\mu}+\frac{\mu}{2}\right]-e^{\varepsilon}\Phi\left[-\frac{\varepsilon}{\mu}-\frac{\mu}{2}\right] \quad (6)$$

则此机制是 μ-GDP,其中 Φ 是标准正态高斯分布的累积分布函数.

定理 4. GDP 的隐私中心极限定理[27]. 当 ρ 是采样率,T 是通信轮数,σ 是噪声尺度,为了保证深度学习模型满足 μ-GDP,$\mu=\rho\sqrt{T(e^{1/\sigma^2}-1)}$.

3.2　压缩感知

CS[39-40]作为一种开拓性信号采样理论,因为同时完成信号采样和压缩以及采样率远低于经典的奈奎斯特/香农准则这两个极具前途的特性而备受关注.数学表达式为 $Y=\boldsymbol{\phi X}$,其中 $\boldsymbol{X}(\boldsymbol{X}\in\mathbb{R}^{a\times 1})$是一个稀疏或可压缩的信号,$\boldsymbol{Y}(\boldsymbol{Y}\in\mathbb{R}^{b\times 1})$是对应的线性测量值且 $b<a$,$\boldsymbol{\phi}(\boldsymbol{\phi}\in\mathbb{R}^{b\times a})$是测量矩阵.高斯随机矩阵和伯努利矩阵已经被证明对于任意给定的稀疏基(即字典)都满足互相干性和零空间特性,因此被广泛作为 CS 的测量矩阵[41].

利用传统的 CS 理论对信号采样时,多维信号需在采样前通过向量化来转换为一维信号,这显然扩大了测量矩阵的维度,继而增加信号采样的计算复杂度.除此之外,由于模型梯度维度同时包括多维和一维,因此对任意信号中的每一列直接进行并行采样和重构的并行压缩感知(Parallel Compressed Sensing, PCS)[42]更适用于压缩模型梯度.

3.3　个性化联邦学习

与专注于协同训练通用的全局模型以探索整个系统全局最优的传统 FL 不同,个性化 FL 目的是根据各方本地数据的统计特征,联合共享全局模型为每一个参与方定制能够正确覆盖其本地数据分布的个性化模型.假设一个 FL 系统是由一个服务器和 N 个参与方构成,各个参与方之间的本地数据是 Non-IID.每一个参与方拥有 $|D_n|(n=1,2,\cdots,N)$ 个含敏感信息的本地私有数据样本,并且具有一定的嵌入式计算能力来定制本地个性化模型.

现有的大部分个性化 FL 主流方法架构如图 1 所示,其中每一轮全局通信包括四个基本步骤:(1)各个参与方利用本地数据从共享的全局模型 W 中训练一个能够更好地拟合其独特数据分布的个性化模型;(2)参与方根据具体的个性化 FL 方法将需要聚合的模型上传至服务器,例如 FedRep[43]方法中的参与方仅上传模型的全局表示而将暗含分类或者预测等重要特征信息(即模型的头部表示)留存在本地,PFLDyn(Proto)[44]方法中的服务器和参与方之间只通信元模型;(3)服务器聚合各方上传的模型来

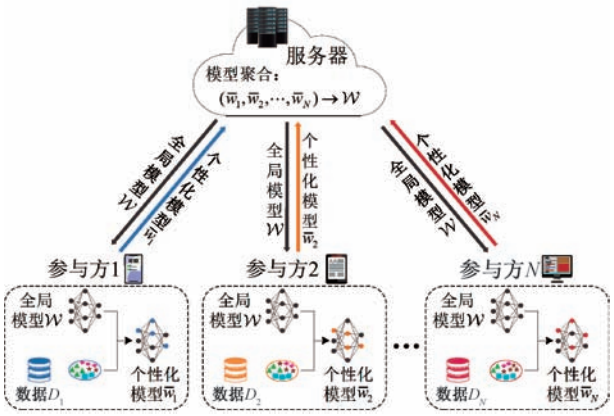

图 1　个性化 FL 示意图

更新全局模型;(4)服务器将新的全局模型下发给各个参与方,以用于下一轮全局通信时本地个性化模型的更新.参与方和服务器之间的双向通信过程需不断重复,直至其定制的个性化模型收敛为止.

4 问题定义和威胁模型

4.1 问题定义

根据研究工作 PFLDyn(Proto)[44]方法的核心思想,本文基于元学习设计了一种个性化 FL 方法.概括而言,联邦系统中的各个参与方协同训练一个共享的全局元模型,同时每一个参与方在本地训练过程中利用私有数据对全局元模型进行变换来定制个性化模型,以适应其本地数据的统计特征.

如图 1 所示,以参与方 $n(n\in[1,N])$为例,这种从全局元模型 W 到参与方个性化模型 $\bar{w}_n$ 的变换过程 $T_n(\mathbb{R}^{d\times1}\to\mathbb{R}^{d\times1})$可以建模为 $\bar{w}_n=T_n(W)$.其中,$\bar{w}_n$ 是将输入 $x_n^i\in\mathbb{R}^{c\times1}(i\in[1,|D_n|])$映射到与真实标签 y_n^i 尽可能相似的预测标签 $\bar{w}_n(x_n^i)\in\boldsymbol{Y}$.整体个性化 FL 的目标数学优化形式可以定义如下:

$$\underset{W\in\mathbb{R}^{d\times1}}{\operatorname{argmin}}\left[F(W)\triangleq\frac{1}{N}\sum_{n\in[1,N]}f_n(\bar{w}_n)\right] \tag{7}$$

其中误差函数 f_n 是对参与方本地数据的预期风险,且 $f_n(\bar{w}_n):=\mathbb{E}_{\{x_n,y_n\}}[\ell(\bar{w}_n(x_n),y_n)]$.$\ell:\boldsymbol{Y}\times\boldsymbol{Y}\to\mathbb{R}$ 是一个损失函数,意指对预测标签 $\bar{w}_n(x_n)$ 和真实标签 y_n 之间距离的惩罚.

公式(7)这个通用目标取决于从全局元模型到本地个性化模型的变换函数 $T_n(n\in[1,N])$.目前,元学习领域内已存在许多变换函数,如模型无关元学习(Model-Agnostic Meta-Learning, MAML)[45]和原型(Prototypical, Proto)[46].其中,MAML 无需额外的参数进行本地个性化模型定制,而 Proto 是将全局元模型用作特征表示来训练本地数据定制个性化模型.尽管基于 MAML 的个性化 FL 方法在大多数情况下表现良好,但是基于 Proto 的个性化 FL 方法在需要拟合新的参与方本地数据分布时更加有效,即泛化能力强[44].因此,本文主要关注 Proto.具体而言,Proto 变换函数使用全局元模型 W 和参与方本地数据来构造类表示.例如参与方 n 的某个数据样本 $x_n^i(i\in[1,|D_n|])$的表示可被定义为 $r_W(x_n^i)$.对于每一个类标签 $k\in\boldsymbol{Y}$,通过均值化这个类实例的表示来获取对应的类表示,即

$$U_W^{n,k}=\frac{1}{|V_n^k|}\sum_{x_n^i\in V_n^k}r_W(x_n^i) \tag{8}$$

其中,$V_n^k=\{(x_n^i,y_n^i):y_n^i=k,(x_n^i,y_n^i)\in D_n\}$ 则是参与方 n 本地标签为 k 的私有训练数据集.定制的参与方模型将测试数据标记为最近邻的类表示,即 $\underset{k\in\boldsymbol{Y}}{\operatorname{argmin}}\,\mathrm{d}(U_W^{n,k},r_W(x_n^j))$,$x_n^j$ 是参与方 n 本地测试数据样本,且 $\mathrm{d}(\cdot,\cdot)$ 是距离函数.

显然,公式(7)的目的是求解所有参与方的元损失函数均值.然而,每一个参与方无法访问全局损失,这便直接导致本地元损失和全局损失之间的错位,即 $\min_W f_n\circ T_n(W)\neq\min_W F(W)(n\in[1,N])$.因此,参与方元损失表现出的本地偏差必然会引起全局收敛的问题.为了避免此问题,Acar 等人[44]明确使本地目标去偏并将本地损失定向到全局目标.即先将本地目标 $f_n\circ T_n$ 去偏为一阶函数,再引入二次正则表达式,那么本地目标是

$$\min_W f_n\circ T_n(W)-\langle\nabla f_n\circ T_n(W^*),W\rangle+\frac{\alpha}{2}\|W-W^*\|^2 \tag{9}$$

其中 α 是超参数且 W^* 是公式(9)的一个静止点.

对于解 $\tilde{W}$,一阶条件可以表示为 $\nabla f_n\circ T_n(\tilde{W})-\nabla f_n\circ T_n(W^*)+\alpha(\tilde{W}-W^*)=\mathbf{0}$.由此可见,$W^*$ 满足此条件,目标不再偏向于参与方最小值.然而,获取最佳模型 W^* 并不严格可行.为了降低服务器和各个参与方之间的通信负担,本文选择全局通信第 t 轮中参与方本地元模型 w_n^t 而不是全局元模型 W^t 来作为最佳模型 W^* 的代理.那么,公式(9)可以重写为以下形式:

$$\min_W f_n\circ T_n(W)-\langle\nabla f_n\circ T_n(w_n^t),W\rangle+\langle\frac{1}{N}\sum_{m\in[1,N]}\nabla f_m\circ T_m(w_m^t),W\rangle+\frac{\alpha}{2}\|W-W^t\|^2 \tag{10}$$

除了参与方数据 Non-IID 这个问题,本文同时关注参与方本地数据的隐私性、个性化模型的效用性以及 FL 系统中通信效率这三个问题.因此本文的目标是在解决公式(10)这个问题的同时需要对各个参与方本地元模型 $w_n^t(n\in[1,N])$进行压缩和加噪.以参与方 n 为例,根据参与方 n 和服务器之间的上行通信资源来决定每一轮全局通信中可以传输的上行通信总量 Γ_n^t.其次参与方 n 压缩本地元模型参数(或梯度)以使得需要上传的模型参数(或梯度)总量不超过 Γ_n^t.此外,在压缩后的本地元模型参数(或梯度)上传之前,对其进行加噪,以使其满足

μ-GDP. 为了保障参与方本地个性化模型的效用,根据定理 2,每一轮全局通信时需要上传的模型参数(或梯度)被添加的高斯噪声方差 ξ(即 $\xi \sim \mathrm{N}(0, s_f^2/\mu^2)$)应该在固定隐私参数 μ 的前提下严格估计相应的敏感度 s_f,以减少 s_f 过大或过小引起的模型效用损失.

4.2　威胁模型

假设攻击者是诚实但好奇的服务器或参与方,这种设置在以前的研究工作中已经被广泛地形式化和实例化[17-25,29-31]. 具体而言,攻击者严格地遵循和执行 FL 协议,并且试图使用各种攻击方法来从捕获的模型参数(或梯度)中重构参与方本地隐私数据或推断某一个数据样本是否属于某一个参与方的本地数据集[47]. 例如,攻击者利用梯度深度泄漏攻击方法观察捕获的模型参数 θ 来恢复 FL 系统中参与方 n 本地数据集 D_n 的攻击目标[7]是

$$\underset{D'_n,\forall n\in[1,N]}{\operatorname{argmin}} \mathbb{E}\left[\left|\sum_{\bar{x}_n^i\in D'_n}\nabla \mathrm{L}(\bar{x}_n^i;\theta)-\sum_{x_n^i\in D_n}\nabla \mathrm{L}(x_n^i;\theta)\right|\right] \tag{11}$$

其中 $x_n^i(i\in[1,|D_n|])$是真实数据集 D_n 中的数据样本,D'_n 是由合成数据样本 $\bar{x}_n^i(i\in[1,|D'_n|])$组成的重构数据集. $\mathrm{L}(\cdot)$ 是基于距离的损失指标.

攻击者通过最小化重构数据集 D'_n 中所有数据样本对应的梯度 $\sum_{\bar{x}_n^i\in D'_n}\nabla \mathrm{L}(\bar{x}_n^i;\theta)$ 与真实数据集 D_n 中所有数据样本对应的梯度 $\sum_{x_n^i\in D_n}\nabla \mathrm{L}(x_n^i;\theta)$ 之间的平均绝对误差来确保合成数据集近似真实数据集. 此外,Gu 等人[48]提到即使被破坏的参与方偏离 FL 协议,亦不会获得任何额外的隐私推理收益,因为参与方上传的任何信息只可能损害模型效用而非影响 FL 隐私协议. 因此,本文的威胁模型是合理且现实的.

5　CUAG-PFL

为了同时有效解决隐私性、效用性、通信效率以及参与方数据 Non-IID 这四个方面的挑战,本文提出一种兼顾通信效率与效用的自适应 GDP 个性化 FL(CUAG-PFL)方法. 对于每一个挑战,解决方法的核心思路如下:

隐私性:(1)本文利用 GDP 为各个参与方上传的模型梯度提供可分析、高精确且严格的隐私保障;(2)本文通过固定隐私预算来自适应调整每一轮全局通信时参与方上传的模型梯度所需噪声尺度,从而保证每一个参与方在整个联邦系统协同训练过程中的本地数据隐私保护程度始终不变(详见第 6 节).

效用性:(1)本文利用 CS 技术在不牺牲隐私的情况下提高基于 LDP 的个性化联邦系统模型效用;(2)本文提出一种优化的确定性二进制测量矩阵来高精度地重构用于下一轮全局通信过程中协助参与方本地训练的聚合压缩梯度,减轻梯度压缩所引起的模型性能退化(详见第 5.3 节);(3)为了保留各个参与方所有的压缩梯度值来减少服务器聚合梯度的重构误差,同时避免因裁剪阈值过大或过小所导致的模型效用降低,本文直接将参与方上传的每一层压缩梯度 ℓ_2-范数设置为该层的裁剪阈值,严格估计压缩梯度每一层的敏感度(详见第 5.4 节).

高效率:(1)因为互联网环境中上行链路速度比下行链路速度慢[49],所以对于提高整个联邦系统的通信效率而言,降低上行通信开销更具有重要意义. 本文利用 CS 技术来节省 FL 系统上行链路的通信传输量;(2)本文根据 ℓ_2-范数来评估全局模型梯度每一层的重要性[50],并且在每一轮全局通信开始时为参与方上传的模型梯度每一层设置不同的压缩率来加强冗余梯度层的维度压缩,大幅度提高了通信效率;(3)相比于常用的高斯随机测量矩阵和伯努利测量矩阵,本文优化的确定性二进制测量矩阵避免了每一轮全局通信时测量矩阵传输引起的额外通信开销和本地存储负担.

个性化:受研究工作 PFLDyn(Proto)[44]启发,联邦系统中所有参与方协同训练一个良好的全局元模型,同时各个参与方基于私有数据从共享的全局元模型中衍生一个性能更优的本地个性化模型.

5.1　整体架构

CUAG-PFL 方法整体架构如图 2 所示,其中以参与方 1 为例描绘了当参与方在参加某一轮全局通信时的本地训练过程以及对本地元模型梯度上传之前进行的操作. 假设一个应用 CUAG-PFL 方法的联邦系统包含一个服务器和 N 个参与方,服务器在每一轮全局通信初始时将当前的全局元模型下发给每一个参与方,并且从全部参与方中随机选取一部分来参加此轮全局通信.

参与方在参加当前全局通信时本地训练的具体过程包括以下三个主要步骤:(1)初始化:将服务器下发的当前全局元模型设置为本地元模型;(2)定制个性化模型:在每一轮本地训练中,参与方随机从私有数据中选取两个等批次量的子数据集,其中某一

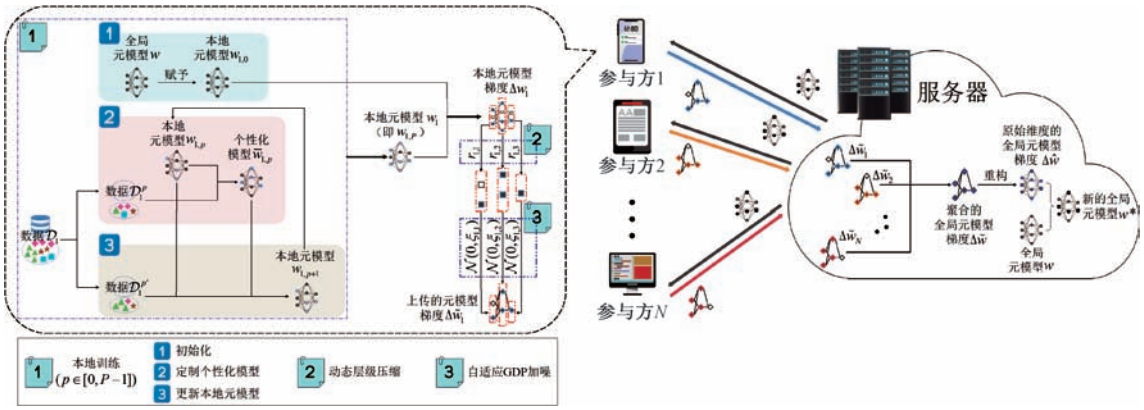

图 2 CUAG-PFL 方法整体架构

个被用于从本地元模型中定制个性化模型来拟合本地数据分布；(3)更新本地元模型：参与方随机从私有数据中选取的另一个子数据集被用于结合当前个性化模型来更新本地元模型. 当本地训练完成后，参与方对本地元模型梯度每一层依次执行动态压缩以及自适应加噪操作，将满足 GDP 的元模型压缩梯度上传至服务器. 对于所有上传的参与方本地元模型梯度，服务器通过重构聚合梯度获取原始维度的全局元模型梯度来更新全局元模型. 当参与方的个性化模型收敛时，此参与方停止与服务器之间的双向通信并且退出整个 FL 系统的协同训练.

5.2 方法描述

算法 1 详细地概述了 CUAG-PFL 方法的完整过程，下面对其进行具体描述. 在参与方进行协同训练之前，服务器需初始化全局元模型 w^0、全局元损失梯度 g^0 和每个参与方级别元损失梯度 $g_n^0(n\in[1,N])$，其中 $g_n^0=g^0=\mathbf{0}$，并且将全局元模型 w^0 下发给 FL 系统中的全部 N 个参与方. 各个参与方根据本地计算效率和网络通信速度与服务器重复地双向交替模型参数的更新，直至其本地个性化模型收敛为止. 此外，考虑到一个 FL 系统的网络通信资源无法承载大规模的参与方都同时参与每一轮的全局通信，因此每一轮通信回合中的服务器都根据预先设置的参与率 λ 从这 N 个参与方中随机选取一部分参与方参加此轮通信，例如参加第 t 轮全局通信的参与方集合是 Q_t.

以第 t 轮通信回合为例，当参与方 $n\in Q_t$ 时，首先设置本地模型 $w_{n,0}^{t+1}$ 为当前全局元模型 w^t，再开始进行本地模型训练. 假设此时参与方 n 正执行第 p 轮的本地训练，为了实现个性化模型和元损失之间的无偏梯度估计，参与方 n 先从本地含标签和敏感信息的数据集 D_n 中随机选取两个等批次量的子数据集 D_n^p 和 $D_n^{p'}$，其次利用子数据集 D_n^p 根据当前参与方 n 的本地元模型来定制属于参与方 n 的个性化模型 $\bar{w}_{n,p}^{t+1}=\tilde{T}_n(w_{n,p}^{t+1},D_n^p)$，其中这里 $\tilde{T}_n(\cdot)$ 是指 Proto 变换. 除此之外，当前参与方 n 的本地元模型 $w_{n,p}^{t+1}$ 的正则化器 $\mathfrak{R}_n^t(w_{n,p}^{t+1})$ 取决于服务器元模型 w^t 和当前参与方级别元损失梯度 g_n^t. 随后，第二个批量子数据集 $D_n^{p'}$ 被用于执行一步随机梯度下降(Stochastic Gradient Descent，SGD)来更新正则化经验损失，如下所示：

$$w_{n,p+1}^{t+1}=w_{n,p}^{t+1}-\beta(\nabla\tilde{f}_n(\bar{w}_{n,p}^{t+1},D_n^{p'})+\nabla\mathfrak{R}_n^t(w_{n,p}^{t+1})) \tag{12}$$

其中 $\tilde{f}_n(\bar{w}_{n,p}^{t+1},D_n^{p'})$ 是指在子数据集 $D_n^{p'}$ 上对定制的参与方 n 本地模型 $\bar{w}_{n,p}^{t+1}$ 的经验损失，β 是学习率.

当执行完 P 次 SGD 更新后，参与方 n 的本地模型训练过程全部结束，可以设置参与方 n 新的元模型为 $w_n^{t+1}=w_{n,P}^{t+1}$. 除此之外，被用于参与方 n 下一轮本地训练过程的元损失梯度更新方式取决于公式(10)，具体形式是 $g_n^{t+1}=g_n^t-\alpha(w_n^{t+1}-w^t)\approx\nabla\tilde{f}_n\circ\tilde{T}_n(w_n^{t+1})$. 此时，参与方 n 的元损失梯度 g_n^{t+1} 保留在本地而元模型梯度 $\Delta w_n^{t+1}=w_n^{t+1}-w^t$ 需要上传至服务器.

为了大幅度地减少 FL 上行链路的通信开销，并提高参与方本地数据隐私保护程度和模型效用，本文设计了一种动态层级压缩梯度方案来压缩模型梯度 Δw_n^{t+1}(详见第 5.3 节)，同时提出一种自适应 GDP 策略对压缩的模型梯度中每一层自适应地加噪(详见第 5.4 节). 随后，参与方将加噪后的压缩梯度上传至服务器.

如果第 t 轮全局通信中的某一个参与方 $n\notin Q_t$ 时，其本地元模型直接更新为当前服务器元模型 w^t，同时参与方级别的元损失梯度保持不变. 当服务器收到 Q_t 集合中所有参与方上传的满足 DP 的压缩模型梯度 $\Delta\tilde{w}_n^{t+1}(n\in Q_t)$ 时，服务器将其先聚

合再均值化,随后利用 CS 重构算法[40]来恢复原始维度的全局模型梯度 $\Delta\hat{w}^{t+1}$,最后服务器根据全局模型梯度 $\Delta\hat{w}^{t+1}$ 依次更新全局元损失梯度和全局元模型.当参与方的本地个性化模型收敛时,此参与方和服务器之间的双向通信终止.

算法 1.　CUAG-PFL.

输入:全局通信的总轮数 T,参与方总数目 N,每一轮全局通信中参与方的参与率 λ,本地训练的总轮数 P,含敏感信息和标签的本地数据 $D_n(n\in[1,N])$,超参数 α,参与方训练本地模型时的学习率 β,批量大小 B 和隐私参数 μ

输出:全局元模型 W

1. 服务器初始化全局元模型 w^0 并将 w^0 下发给全部 N 个参与方,同时将全局元损失梯度 g^0 和参与方级别元损失梯度 $g_n^0(n\in[1,N])$都初始化为零,即 $g_n^0=g^0=\mathbf{0}$,
2. FOR 全局通信轮数 $t=0,1,\cdots,T-1$ DO
3. 服务器随机选取 λN 个参与方组成参加此一轮全局通信的参与方集合 $Q_t(Q_t\subseteq\{1,2,\cdots,N\})$;
4. IF 全局通信轮数 $t=0$ THEN
5. 通信梯度所有层压缩率集 $r^t=$ 算法 $2(w^0)$;
6. ELSE
7. 通信梯度所有层压缩率集 $r^t=$ 算法 $2(\Delta\hat{w}^{t+1})$;
8. END IF
9. FOR 参与方 $n\in Q_t$ DO
10. 设置本地模型 $w_{n,0}^{t+1}=w^t$;
11. FOR 本地训练轮数 $p=0,1,\cdots,P-1$ DO
12. 从本地数据集 D_n 中随机选取两个等批量大小的子数据集 D_n^p 和 $D_n^{p'}$;
13. 定制个性化模型 $\bar{w}_{n,p}^{t+1}=\widetilde{T}_n(w_{n,p}^{t+1},D_n^p)$;
14. 求解本地元模型的正则化器 $\Re_n(w_{n,p}^{t+1})=-\langle w_{n,p}^{t+1},g_n^t\rangle+\frac{\alpha}{2}\|w_{n,p}^{t+1}-w^t\|$;
15. 更新参与方本地元模型 $w_{n,p+1}^{t+1}=w_{n,p}^{t+1}-\beta(\nabla\widetilde{f}_n(\bar{w}_{n,p}^{t+1},D_n^{p'})+\nabla\Re_n(w_{n,p}^{t+1}))$;
16. END FOR
17. 设置参与方 n 新的元模型 $w_n^{t+1}=w_{n,P}^{t+1}$;
18. 更新元损失梯度 $g_n^{t+1}=g_n^t-\alpha(w_n^{t+1}-w^t)$;
19. 计算参与方元模型梯度 $\Delta w_n^{t+1}=w_n^{t+1}-w^t$;
20. 生成测量矩阵集 $\widetilde{\phi}^t=$ 算法 $3(r^t,\Delta w_n^{t+1})$;
21. 参与方 n 对元模型梯度 Δw_n^{t+1} 依次执行动态层级压缩和自适应 GDP 加噪操作获取 $\Delta\tilde{w}_n^{t+1}=$ 算法 $4(\Delta w_n^{t+1},\widetilde{\phi}^t,B,t,\lambda,P,|D_n|,\mu)$,并将 $\Delta\tilde{w}_n^{t+1}$ 上传至服务器
22. END FOR
23. FOR $n\notin Q_t$ DO
24. 更新本地元模型 $w_n^{t+1}=w^t$,并且参与方级别的元损失梯度 $g_n^{t+1}=g_n^t$;
25. END FOR
26. 服务器均值化所有参与方的 $\Delta\tilde{w}_n^{t+1}(n\in Q_t)$,即 $\Delta\tilde{w}^{t+1}=\frac{1}{\lambda N}\sum_{n\in Q_t}\Delta\tilde{w}_n^{t+1}$;
27. 利用 CS 重构算法恢复原始维度的全局梯度 $\Delta\hat{w}^{t+1}=Rec(\Delta\tilde{w}^{t+1},\widetilde{\phi}^t)$,其中 $Rec(\cdot)$ 指 CS 重构算法;
28. 服务器更新全局元损失梯度 $g^{t+1}=g^t-\alpha\Delta\hat{w}^{t+1}$;
29. 服务器更新全局元模型 $w^{t+1}=w^t+\Delta\hat{w}^{t+1}-\frac{1}{\alpha}g^{t+1}$;
30. 当服务器元模型 w^{t+1} 收敛时,整个 FL 协同训练过程结束,最终的共享全局元模型 $W=w^{t+1}$;
31. END FOR

5.3　动态层级压缩梯度

目前,已有研究工作表明 ℓ_2-范数可以用于评估神经网络模型每一层的重要性,并且指出 ℓ_2-范数值较小的模型层通常情况下对神经网络模型效用损失影响较小[50].因此,本文这里根据梯度 ℓ_2-范数比来决定对应层的动态压缩率.值得注意的是,第一轮全局通信中参与方上传的模型梯度各层的压缩率是由服务器初始化的全局元模型 w^0 决定而非取决于前一轮恢复原始维度后的聚合元模型梯度.之所以选择依据全局元模型的梯度范数比来设置每一层的压缩率,是因为参与全局通信的参与方在定制个性化模型和更新本地元模型时都涉及到全局元模型.

动态层级压缩率集的生成过程则如算法 2 所示.根据 FL 应用场景通信资源限制的具体情形,预先设置最高压缩率 $r_{\max}$ 和最低压缩率 $r_{\min}$ 分别控制模型梯度层的通信效率上限和因梯度压缩所导致的模型性能退化下限.具体而言,通过赋予梯度压缩比值大的层较高的压缩率来保证在一定程度上提高通信效率时降低对重要梯度的压缩程度,进而减少梯度重构误差对模型重要层的效用损失.与此同时,在尽可能地减少模型效用损失的前提下对梯度压缩比值较小的层进行大幅度压缩以提高 FL 通信效率.基于上述原理,服务器通过计算每一层的梯度范数占总模型梯度范数的权重比来动态地生成这一层所对应的压缩率 $r_l\in(r_{\min},r_{\max}](l\in[1,L])$,其中 L 是模型梯度的总层数.

算法 2.　动态层级压缩率集的生成.

输入:初始化的全局元模型 w^0 或者前一轮恢复原始维度的聚合模型梯度 Δw(为了方便,下面统一用 w 表示),最高压缩率 $r_{\max}$ 和最低压缩率 $r_{\min}$

输出:模型梯度所有层压缩率集 $r=\{r_1,r_2,\cdots,r_L\}$

1. FOR 层数 $l=1,2,\cdots,L$ DO

2. 计算当前层 l 的梯度范数占总的模型梯度范数比，即 $weight_l = \| w_l \|_2 / \| w \|_2$；
3. IF $weight_l < r_{\min}$ THEN
4. $r_l = \text{round}(r_{\max} - \text{round}(weight_l, 2), 1)$；
5. ELIF $weight_l \geqslant r_{\min}$ & $\text{round}(weight_l, 1) < r_{\max}$ THEN
6. $r_l = \text{round}(weight_l, 1)$；
7. ELSE
8. $r_l = r_{\max}$；
9. END IF
10. END FOR

当计算出通信模型梯度所有层的压缩率集 $r^t = \{r_1, r_2, \cdots, r_L\}$ 时，每一个参与方可以为当前通信模型梯度每一层生成对应的测量矩阵 ϕ_l. 目前，将 CS 技术用于提高 FL 系统通信效率的大多数研究工作都直接利用高斯随机矩阵压缩模型梯度. 然而，服务器和各个参与方之间传输随机测量矩阵所引起的额外通信开销无法被忽略，并且过多的不确定性会导致参与方硬件成本高、存储容量大. 为了解决上述问题的同时减轻梯度压缩引起的联邦系统模型性能退化，本文设计了一个确定性二进制测量矩阵，其生成过程如算法 3 所示，并且数学化形式可以表达如下：

$$\boldsymbol{\phi} = \begin{bmatrix} \overbrace{[1\cdots1]}^{z+1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \overbrace{[1\cdots1]}^{z} \end{bmatrix}, s.t.\ z = \lfloor 1/r \rfloor \tag{13}$$

其中 r 是压缩率，符号 $\lfloor h \rfloor$ 则是将元素 h 舍入到最接近负无穷大的整数.

理论上，若优化的测量矩阵能够最大限度地保留有效梯度信息和去除冗余梯度信息，即聚合的压缩梯度可以被高精度地重构以便用于下一轮全局通信过程中参与方本地训练，那么此测量矩阵 $\boldsymbol{\phi}$ 与信号任意稀疏基 $\boldsymbol{\psi}$ 都应不相干[39]，其相干性的数学表达式如下：

$$v(\boldsymbol{\phi}, \boldsymbol{\psi}) = \sqrt{a} \max_{i,j} \frac{|\langle \boldsymbol{\phi}_i, \boldsymbol{\psi}_j \rangle|}{\| \boldsymbol{\phi}_i \|_2 \| \boldsymbol{\psi}_j \|_2} \tag{14}$$

如果 $\boldsymbol{\phi}$ 尺寸大小是 $b \times a (b < a)$ 且 $\boldsymbol{\psi} \in \mathbb{R}^{a \times a}$，那么 $i \in \{1, 2, \cdots, b\}$ 且 $j \in \{1, 2, \cdots, a\}$. 此外，当 $v(\boldsymbol{\phi}, \boldsymbol{\psi})$ 值越小，聚合梯度的重构质量越高.

这里，本文验证高斯随机矩阵 $\boldsymbol{\phi}_G$、伯努利矩阵 $\boldsymbol{\phi}_B$ 和优化的确定性二进制测量矩阵 $\boldsymbol{\phi}$ 分别与两个常用的信号稀疏基（即离散余弦变换矩阵 $\boldsymbol{\psi}_{DCT}$ 和离散傅里叶变换矩阵 $\boldsymbol{\psi}_{DFT}$）之间的相干性. 假设 $a = 512$，压缩率 $r = \lfloor b/a \rfloor$. 如图 3 所示，高斯随机矩阵 $\boldsymbol{\phi}_G$ 或伯努利矩阵 $\boldsymbol{\phi}_B$ 和信号稀疏基之间的相干性都在某一区间内上下波动，而本文设计的确定性二进制测量矩阵 $\boldsymbol{\phi}$ 与 $\boldsymbol{\psi}_{DCT}$ 或 $\boldsymbol{\psi}_{DFT}$ 之间的相干性随着 r 增大而减小. 当 $r > 0.1$ 时，$v(\boldsymbol{\phi}_G, \boldsymbol{\psi})$ 和 $v(\boldsymbol{\phi}_B, \boldsymbol{\psi})$ 都大于 $v(\boldsymbol{\phi}, \boldsymbol{\psi})$. 尤其是当 $r \geqslant 0.5$ 时，$v(\boldsymbol{\phi}_G, \boldsymbol{\psi})$ 和 $v(\boldsymbol{\phi}_B, \boldsymbol{\psi})$ 至少是 $v(\boldsymbol{\phi}, \boldsymbol{\psi})$ 的两倍，并且 $v(\boldsymbol{\phi}, \boldsymbol{\psi}) \leqslant 2$. 除此之外，为了进一步验证本文设计的确定性二进制测量矩阵 $\boldsymbol{\phi}$ 能够在相同情况下提高 FL 系统通信效率的同时减少模型性能退化，本文将高斯随机矩阵 $\boldsymbol{\phi}_G$、伯努利矩阵 $\boldsymbol{\phi}_B$ 以及 $\boldsymbol{\phi}$ 应用于 CUAG-PFL 方法中，具体分析详见第 7.4.4 节.

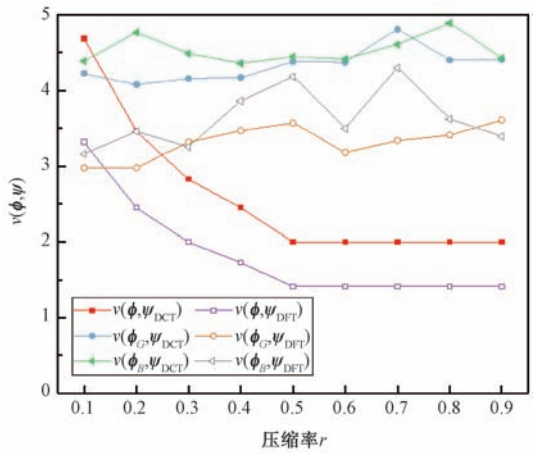

图 3 确定性二进制测量矩阵 $\boldsymbol{\phi}$、高斯随机矩阵 $\boldsymbol{\phi}_G$ 和伯努利矩阵 $\boldsymbol{\phi}_B$ 分别与两个信号稀疏基（即离散余弦变换矩阵 $\boldsymbol{\psi}_{DCT}$ 和离散傅里叶变换矩阵 $\boldsymbol{\psi}_{DFT}$）之间的相干性比较

算法 3. 确定性二进制测量矩阵集的生成.

输入：梯度所有层压缩率集 $r = \{r_1, r_2, \cdots, r_L\}$，梯度 Δw

输出：梯度所有层对应的测量矩阵集 $\tilde{\boldsymbol{\phi}} = \{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \cdots, \boldsymbol{\phi}_L\}$

1. 将梯度 Δw 所有层都转换成向量形式；
2. FOR 层数 $l = 1, 2, \cdots, L$ DO
3. 假设当前层 l 的向量维度是 I，那么此层测量矩阵 $\boldsymbol{\phi}_l$ 的大小是 $(\lfloor r_l \cdot I \rfloor, I)$；
4. 计算包含 $\lfloor 1/r_l \rfloor + 1$ 个值为 1 的行数是 $count = I - \lfloor 1/r_l \rfloor \cdot r_l \cdot I$；
5. IF $count > 0$ THEN
6. FOR 测量矩阵行数 $j \in [0, count - 1]$ DO
7. 此行中列数 $[j \cdot (\lfloor 1/r_l \rfloor + 1), (j+1) \cdot (\lfloor 1/r_l \rfloor + 1) - 1]$ 的元素设置为 1，其余列元素设置为 0；
8. END FOR

9.　　FOR 测量矩阵行数 $j \in [count, \lfloor r_l \cdot I \rfloor - 1]$ DO

10.　　此行中列数 $[I - \lfloor 1/r_l \rfloor \cdot \lfloor r_l \cdot I \rfloor + j \cdot \lfloor 1/r_l \rfloor, I - \lfloor 1/r_l \rfloor \cdot \lfloor r_l \cdot I \rfloor + (j+1) \cdot \lfloor 1/r_l \rfloor - 1]$ 的元素设置为1,其余列元素设置为0;

11.　　END FOR

12.　　ELSE

13.　　FOR 测量矩阵行数 $j \in [0, \lfloor r_l \cdot I \rfloor - 1]$ DO

14.　　此行中列数 $[j \cdot \lfloor 1/r_l \rfloor, (j+1) \cdot \lfloor 1/r_l \rfloor - 1]$ 的元素设置为1,其余列元素设置为0;

15.　　END FOR

16.　　END IF

17.　　获取此层的测量矩阵 $\boldsymbol{\phi}_l$;

18. END FOR

5.4　自适应 GDP 加噪

随后,本文先利用生成的测量矩阵集 $\tilde{\boldsymbol{\phi}}$ 对需要上传的模型梯度进行动态层级压缩,再对压缩梯度执行自适应 GDP 加噪操作,如算法 4 所示. 根据预先设置的隐私参数 μ,可以直接计算当前第 t 轮全局通信时参与方 n 通信模型梯度所需的噪声尺度 $\sigma_n^t = \sqrt{1/\ln(\mu^2 \mid D_n \mid^2/(4\lambda^2 B^2 tP)+1)}$,其中 B 是指参与方本地训练时所选取的数据批量大小,即 $B = \mid D_n^p \mid = \mid D_n^{p'} \mid$. 噪声尺度 σ 的计算推导详见第 6 节.

除此之外,在基于 LDP 的 FL 系统中,每一轮全局通信过程中不同参与方本地模型梯度的敏感度可能不同,并且任意一个参与方在不同轮数的全局通信过程中的本地模型梯度敏感度可能也不相同. 同理,参与方本地模型梯度每一层的敏感度亦可能不同[31]. 然而,目前绝大多数已有的相关研究工作都预设置固定的裁剪阈值 C 来估计敏感度 S. 假设存在一个固定裁剪阈值 C,则需要对算法 4 第 3 步获取的当前层 l 所对应的压缩梯度 $\Delta\hat{w}_l$ 进行梯度裁剪,即 $\Delta\bar{w}_l = \Delta\hat{w}_l \cdot \min\{1, C/\parallel \Delta\hat{w}_l \parallel_2\}$. 下面将考虑两种情况:(1)如果任意参与方本地模型各层的压缩梯度范数都小于预先设置的裁剪阈值 C,那么 $\Delta\bar{w}_l = \Delta\hat{w}_l$,此时 C 是在任何给定的梯度层、参与方和全局通信轮数情况下的压缩梯度敏感度的松散估计,而各层压缩梯度范数实际上是对相应层注入噪声量的敏感度严格估计;(2)如果任意参与方本地模型每一层的压缩梯度范数大于裁剪阈值 C,那么 $\Delta\bar{w}_l = \Delta\hat{w}_l \cdot (C/\parallel \Delta\hat{w}_l \parallel_2)$,但是裁剪后的 $\Delta\bar{w}_l$ 会影响聚合的压缩梯度高精度重构.

因此,为了严格估计参与方本地模型各层压缩梯度注入噪声量的敏感度以及减轻梯度压缩引起的 FL 模型性能退化,鉴于各个参与方已经对模型梯度各层都进行了较好的动态压缩,本文这里直接跟踪参与方本地模型各层压缩梯度的 ℓ_2-范数. 即令 $C_l = \parallel \Delta\hat{w}_l \parallel_2$,则 $\Delta\bar{w}_l = \Delta\hat{w}_l \cdot \min\{1, C/\parallel \Delta\hat{w}_l \parallel_2\} = \Delta\hat{w}_l$. 根据第 6 节,当前层 l 的压缩梯度注入噪声量的敏感度是 $S_l = 2C_l/B$. 相应地,当前层 l 的噪声方差是 $\xi_l = (2C_l \cdot \sigma/B)^2$. 综上所述,通信的模型梯度各层的噪声方差因裁剪阈值(或对应的敏感度)和噪声尺度的变化而变化,呈现出自适应 GDP 的现象.

算法 4.　动态层级压缩和自适应 GDP 加噪.

输入:本地模型梯度 Δw,梯度所有层对应的测量矩阵集 $\tilde{\boldsymbol{\phi}} = \{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \cdots, \boldsymbol{\phi}_L\}$,批量大小 B,当前全局通信轮数 t,参与率 λ,本地训练的总轮数 P,参与方本地数据总量 $\mid D \mid$ 和隐私参数 μ

输出:满足 GDP 的压缩模型梯度 $\Delta\tilde{w}$

1. 计算噪声尺度

$\sigma = \sqrt{1/\ln(\mu^2 \mid D \mid^2/(4\lambda^2 B^2 tP)+1)}$;

2. FOR 层数 $l = 1, 2, \cdots, L$ DO

3.　　当前层 l 对应的压缩梯度 $\Delta\hat{w}_l = \boldsymbol{\phi}_l \cdot \Delta w_l$;

4.　　计算当前层 l 的裁剪阈值 $C_l = \parallel \Delta\hat{w}_l \parallel_2$;

5.　　计算当前层 l 的噪声方差 $\xi_l = (2C_l \cdot \sigma/B)^2$;

6.　　添加噪声,即 $\Delta\tilde{w}_l = \Delta\hat{w}_l + \mathrm{N}(0, \xi_l)$;

7. END FOR

6　隐私分析

在基于 LDP 的 FL 中,定义 1 中的相邻数据集 D 和 D' 存在两个不同层面的定义:(1)样本级 LDP:如果 D' 是通过数据集 D 添加或删除某一个数据样本构成,那么 D 和 D' 被定义为样本级相邻的;(2)参与方级 LDP:如果 D' 是通过数据集 D 添加或删除某一个参与方所有数据样本构成,那么 D 和 D' 则被定义为参与方级相邻的. 一般情况下,样本级 LDP 是隐私协议的标准要求,通常比参与方级 LDP 弱[36]. 根据第 4.2 节,假设攻击者妄图推断 FL 系统中某一个参与方本地数据集共性,则本文的隐私目标是提出一个高效率高效用的参与方级 LDP 的 FL 方法.

除此之外,为了保证在 FL 系统协同训练过程中,各个参与方的本地数据隐私保护程度始终不变,即固定隐私预算,每轮全局通信时将动态调整通信梯度所需要添加的高斯噪声方差. 因此,对于满足 μ-GDP 的 CUAG-PFL 方法,始终有定理 5 成立.

定理 5. 为了保证在任何一轮全局通信过程中,

参加此轮通信的参与方 $n(n\in[1,N])$的本地模型渐进满足参与方级 μ_n -GDP,则此参与方通信模型梯度所需要添加的高斯噪声方差应该满足:

$$\xi=4C^2/\left(B^2\ln\left(\frac{\mu_n^2\mid D_n\mid^2}{4\lambda^2B^2tP}+1\right)\right) \tag{15}$$

其中,D_n 是参与方 n 的含敏感信息和标签的本地数据集,C 是裁剪阈值,B 是参与方本地训练过程中设置的数据批量大小,P 则是本地训练的总轮数.此外,λ 是每一轮全局通信中参与方的参与率,t 是指当前全局通信轮数.

证明. 通过跟踪本文 CUAG-PFL 方法中每一个经历交互通信的参与方本地模型更新来分析参与方级的隐私.在任何一轮全局通信过程中,Q_t 中的各个参与方需要对本地模型梯度进行参与方级的梯度裁剪来保证其上传至服务器的模型梯度不超过阈值 C ,所以就 ℓ_2-范数而言,每轮全局通信中添加或删除一个参加此轮通信的参与方的模型梯度将改变共享的聚合梯度的上限是 C .显然,梯度裁剪的目的是限制共享的聚合梯度对某一个参与方通信模型梯度的 ℓ_2-敏感度.具体以 Q_t 中的参与方 n 为例,先假设参与方本地训练的批量大小等于训练样本总数 $\mid D_n\mid$,则本地模型训练过程被定义为 $s_n^{D_n}\triangleq w_n=\frac{1}{\mid D_n\mid}\sum_{q=1}^{\mid D_n\mid}\min_{w_n\in W}f_n(w_n,D_{n,q})$,其中 $D_{n,q}$ 是 D_n 的第 q 个数据样本.此外,假设数据集 D_n 某个样本级的相邻数据集是 D'_n ,那么这种情况下通信模型梯度的敏感度是 $\Delta s_n=\max_{D_n,D'_n}\| s_n^{D_n}-s_n^{D'_n}\|=\frac{2C}{\mid D_n\mid}$.通常情况下,各个参与方将本地训练的批量大小 B 设置成远小于训练样本总数 $\mid D_n\mid$ 的值,即 $B<\mid D_n\mid$,此时参与方级的敏感度是 $\Delta s_n=2C/B$.为了保障各个参与方的本地数据隐私,CUAG-PFL 中每一个参与方的通信模型梯度所需要添加的噪声标准差是 $2C\sigma/B$.因此,根据定理 2,参与方 n 的本地模型更新满足 $1/\sigma$ -GDP.

当参与方进行本地模型训练时,P 中的每一轮用于本地模型更新的训练数据样本被选中的概率都是 $2B/\mid D_n\mid(n\in Q_t)$.同时已知每一个参与方在每一轮全局通信中被服务器选中的概率是 λ ,那么每一轮全局通信中任意一个参与方的本地数据被采样的概率是 $2\lambda B/\mid D_n\mid$.假设服务器随机选择参与方时遵循无差别原则,并且用于更新参与方本地模型的数据样本都是被均匀地随机选择.那么根据定理 4,当前第 t 轮全局通信时 Q_t 中的参与方 n 的本地模型满足参与方级 μ_n -GDP,其中隐私参数 $\mu_n=\frac{2\lambda B}{\mid D_n\mid}\sqrt{tP(e^{1/\sigma^2}-1)}$.因此参与方 n 的通信模型梯度需要添加的高斯噪声尺度是

$$\sigma=\sqrt{1/\ln\left(\frac{\mu_n^2\mid D_n\mid^2}{4\lambda^2B^2tP}+1\right)} \tag{16}$$

相应地,参与方 n 的通信模型梯度需要添加的高斯噪声方差应该满足:

$$\xi=4C^2/\left(B^2\ln\left(\frac{\mu_n^2\mid D_n\mid^2}{4\lambda^2B^2tP}+1\right)\right) \tag{17}$$

此时,参加任何一轮全局通信过程的参与方 n 的通信模型渐进地满足参与方级 μ_n -GDP,从而保证了参与方的本地私有数据在整个 FL 协同训练过程中不会被诚实但好奇的其它参与方或者服务器所窃取.

证毕.

随着当前全局通信轮数 t 的增加,更新参与方本地元模型时使用的 SGD 方法的固有特性会导致参与方元模型梯度的 ℓ_2-范数逐渐减小.假设参与方 n 元模型梯度第 l 层是 Δw_l 且对应的压缩梯度是 $\Delta\hat{w}_l$.根据 CS 的约束等距性(Restricted Isometry Property, RIP)[51],Δw_l 和 $\Delta\hat{w}_l$ 满足

$$(1-\eta)\|\Delta w_l\|_2\leqslant\|\Delta\hat{w}_l\|_2\leqslant(1+\eta)\|\Delta w_l\|_2 \tag{18}$$

其中,$\eta\in(0,1)$.

当原始梯度 Δw_l 的 ℓ_2-范数减小时,对应的压缩梯度 ℓ_2-范数 $\|\Delta\hat{w}_l\|_2$ 才相应减小.根据公式(16)和(17),高斯噪声方差 ξ 主要是由裁剪阈值 C 和噪声尺度 σ 这两个隐私相关参数来决定的.固定公式(16)中的 μ_n、$\mid D_n\mid$、λ、B 和 P,σ 随着 t 的增加而增大.因此,在 FL 后期模型收敛阶段,即 t 接近全局通信的总轮数 T 时,压缩梯度范数 $\|\Delta\hat{w}_l\|_2$ 可能降低到很小的数值,等价于裁剪阈值 C 较小.但是因为此时的噪声尺度 σ 值较大,因此 $\Delta\hat{w}_l$ 所添加的噪声量不会和 C 一样都非常小.综合而言,相比于固定裁剪阈值 C 和噪声尺度 σ 的方法,CUAG-PFL 方法中的隐私参数 μ 固定不变,即各个参与方的本地数据隐私保护程度始终不变.而相比于固定裁剪阈值 C 和隐私参数 μ 的方法,CUAG-PFL 能够确保协同训练的后期模型收敛阶段的噪声量较小,进而避免影响模型效用.

7 实验结果与分析

本文在 CIFAR-10 和 CIFAR-100 两个真实的

Non-IID 联邦数据集上进行大量实验,并且通过与六种最相关的方法(即 FedAvg[8]、PriFedSync[21]、UDP[24]、DP-SCAFFOLD[25]、SCAFFOLD[32] 和 PFLDyn(Proto)[44])进行比较来验证本文提出的 CUAG-PFL 方法在隐私性、效用性、通信效率和个性化四个方面的优势.

7.1　实验环境

实验环境是 Python 3.7.11 开发,运行环境 CPU 为 Intel(R) Xeon(R) Gold 5117 @ 2.00 GHz,GPU 为 NVIDIA Tesla V100 SXM2,内存是 32 GB×4,实验操作系统是 Ubuntu 18.04.2 (OS).

7.2　实验设置

7.2.1　联邦数据集(Non-IID)

本文使用两个流行的基准数据集(即 CIFAR-10 和 CIFAR-100)进行实验仿真,并且按照研究工作 PFLDyn(Proto)[44] 中的配置构造 Non-IID 联邦数据集.基本目标是对参与方数据进行采样以合成参与方之间不同程度的统计数据异构性和任务多样性,具体分割方式为以下两种.

(1)显示类诱导多样性(Active Class Induced Diversity, ACID).在此设置中,首先为类列表大小较小的每一个参与方分配一个固定大小的类列表.选择类之后,根据相应的类列表从实际的训练/测试数据集拆分中随机无替换地选取数据样本来构建每一个参与方的训练/测试数据集.例如,在 CIFAR-100 数据集中,假设一个 FL 系统存在 100 个参与方,其中每一个参与方的类标签固定为 5.因为 CIFAR-100 数据集中总共是 100 个类,并且每一个参与方具有的数据类别总数被限制为 5 个,因此这个 FL 系统会出现许多参与方之间都有严格不同的数据类别,即彼此之间的数据类别重叠概率极小.此外,本文还通过增加和减少每一个参与方本地数据类别的总个数来指出本文方法和其它方法在处理不同级别的任务多样性时的性能.

(2)匿名类诱导多样性(Anonymous Label Induced Diversity, ALID).与 ACID 类似,ALID 先为每一个参与方选择固定数量的类并且构建参与方数据集.然后对于每一个参与方,随机排列类索引,以便一个参与方中的类索引与另一个参与方中的类索引无任何关系.显然在这种情况下,即使所有参与方都具有相同的类标签,参与方数据集联合分布之间的距离也非常大.进行这项联邦数据集分割方式的动机源于参与方可能不希望透露类信息,但是仍然渴望从联邦训练中受益.

7.2.2　模型架构

CIFAR-10 和 CIFAR-100 两个联邦数据集对应的卷积神经网络(Convolutional Neural Networks, CNNs)构造类似于研究工作 PFLDyn(Proto)[44],此模型包括两个卷积层、两个最大池化层、两个全连接层和一个 softmax 激活层.值得注意的是,尽管这种模型架构对于 CIFAR-10 和 CIFAR-100 而言并不是最前沿的 CNNs,但是也足以证明 CUAG-PFL 和其它方法的相对性能.为了方便,下面统一用 CNNs 表示本节所述的模型架构.

7.2.3　对比方法

这里简单概括六个和 CUAG-PFL 比较的方法以及阐述对应的原因.

(1)FedAvg[8].这是 FL 领域内最早被提出的方法,通常是绝大部分 FL 相关研究的一种比较基线方法,主要思想是服务器和各个参与方协同训练一个全局共享模型.

(2)UDP[24].这是基于高斯机制的参与方级 LDP 的 FL 中的典型方法,其中包含了自适应裁剪阈值策略.除此之外,UDP 是通过固定隐私预算来确保每一轮全局通信的参与方本地数据隐私,因此每轮全局通信的噪声是自适应变化的.UDP 和 FedAvg 一致,都是非个性化 FL 方法.

(3)PFLDyn(Proto)[44].作为启发本文提出的 CUAG-PFL 的一种最新个性化 FL 方法,可以被认为是 CUAG-PFL 在非压缩且无隐私情况下个性化 FL 模型效用的上限.

(4)PriFedSync[21].据我们所知,PriFedSync 是一种将 GDP 应用于 FL 中的最新通用框架,但是固定了噪声尺度和裁剪阈值.此外,PriFedSync 可以应用于各种个性化 FL 框架中,因此这里本文采用的基础框架是 PFLDyn(Proto),便于和 CUAG-PFL 自适应 GDP 策略形成鲜明的对比.

(5)DP-SCAFFOLD[25].这是一种基于高斯机制的个性化 FL 领域中具有代表性的新颖方法,将每一次本地训练中梯度范数的中位数作为裁剪阈值,但是每一轮全局通信中所有参与方上传的模型梯度所添加的噪声是固定的.

(6)SCAFFOLD[32].作为 DP-SCAFFOLD 在非隐私保护情况下对应的个性化 FL 模型效用上限.

7.2.4　超参数设置

为了便于和 FedAvg、UDP、PFLDyn(Proto)、PriFedSync、SCAFFOLD 和 DP-SCAFFOLD 进行公平比较,本文仿真的 FL 系统是由一个中心服务

器和 100 个参与方构成，即 $N=100$. 每一轮有效通信回合初始时，中心服务器随机选择 10 个参与方参加此轮协同训练，即 $\lambda=0.1$. 本文采用第 7.2.1 节中的两个联邦数据集分割方式 ACID 和 ALID，其中每一个参与方的类标签数目设定为三种（即 3、5、7），即一共实验仿真了六种 Non-IID 的 FL 场景. 除此之外，本文设置了相同的超参数，包括：全局通信总轮数 $T=300$，每个参与方本地训练轮数 $P=25$，本地训练的批量大小 $B=50$，本地训练的学习率 $\beta=0.1$，参数 $\alpha=0.1$. 并且为了避免发散，跨全局通信轮的学习速率衰减被设置为 0.997，权重衰减设置为 0.001. 对于基于 LDP 的 FL 方法，固定裁剪阈值设置为 1，PriFedSync 固定噪声设置为 0.45 而 DP-SCAFFOLD 固定噪声设置为 1.0，UDP 的固定隐私预算 $\varepsilon=10(\delta=10^{-5})$. 默认情况下，CUAG-PFL 中隐私参数 $\mu=0.25$，最高压缩率 $r_{\max}=0.5$ 且最低压缩率 $r_{\min}=0.2$.

7.2.5 评估标准

对于个性化 FL 方法，即 CUAG-PFL、PFLDyn（Proto）、PriFedSync、SCAFFOLD 和 DP-SCAFFOLD，衡量指标是各个参与方利用各自个性化模型测试本地数据性能的均值. 至于 FedAvg 和 UDP 两种非个性化 FL 方法，则是直接利用共享的全局模型测试各个参与方本地数据的准确率均值作为评价指标. 显然，准确率越小意味着实用价值越低. 值得注意的是，为了公平公正，无论何种 Non-IID 联邦数据集情形，这七种方法中每一轮被选择的参与方都是完全一致的.

7.3 实验对比与分析

下面本节将依次从隐私性、效用性、通信效率和个性化四个方面对 CUAG-PFL 和其它六种相关研究方法进行相应的实验对比与分析.

隐私性：根据定理 3，μ-GDP 可以转换为 (ε,δ)-DP. 因此，为了公平公正地将 CUAG-PFL 与其它三种基于 LDP 的 FL 方法比较，本文固定 $\delta=10^{-5}$，将满足 μ-GDP 的 CUAG-PFL 和 PriFedSync 两种方法中隐私参数 μ 转换成 (ε,δ)-DP 中相应的隐私预算 ε，从而直观地分析这四种方法对参与方数据隐私保障的强度. 正如第 7.2.3 节中所述，CUAG-PFL 和 UDP 都是固定 ε 来确保任何一个参与方在参加联邦系统协同训练过程中的数据隐私保障强度始终不变，而 PriFedSync 和 DP-SCAFFOLD 却是固定噪声尺度 σ. 结合定理 1、定理 4 以及第 6 节，PriFedSync 和 DP-SCAFFOLD 的 ε 随着全局通信轮数的增加而增大，且 CUAG-PFL 和 UDP 的 σ 随着全局通信轮数的增加而增大，分别如图 4(a)和 4(b)所示. 根据第 7.2.4 节中 σ 的设置，PriFedSync 的 ε 变换范围是[0.41, 10.24]，DP-SCAFFOLD 的 $\varepsilon\in[1.52,26.28]$. 为了证明 CUAG-PFL 即使是在强隐私性的情况下亦具有较优的模型效用，这里取 $\mu=0.25$，即 $\varepsilon=0.92$，大约是 PriFedSync 和 DP-SCAFFOLD 两个方法中 ε 最小值的均值. 对于非个性化 FL 方法 UDP，则固定 ε 的值为个性化 FL 方法 PriFedSync 中 ε 最大值取整，即 $\varepsilon=10$. 除此之外，以 ACID3、5、7 这三种分割 CIFAR-10 联邦数据集的场景为例，表 1 给出了上述四种基于 LDP 的 FL 方法在模型性能达到目标准确率时需要消耗的隐私预算 ε. 其中，每一种场景下的三个目标准确率分别对应 UDP、DP-SCAFFOLD 和 PriFedSync 在全局通信总轮数 $T=300$ 时的最高准确率. 显然，CUAG-PFL 和 PriFedSync 这两种个性化 FL 方法在模型效用方面更具有显著优势. 然而，随着目标准确率的增加，PriFedSync 消耗的隐私预算 ε 逐渐增大而本文提出的 CUAG-PFL 方法所对应的隐私预算 ε 一直保持较小的值不变. 尤其是在 ACID7 这种数据分割方式下，当目标准确率是 82.98%时，对于所需要消耗的隐私预算 ε，PriFedSync 大约是 CUAG-PFL 的 11 倍. 根据定义 1，较小的 ε 表示隐私性更强. 因此，相比于其它三种基于 LDP 的 FL 方法，CUAG-PFL 对参与方本地数据隐私保护程度更强.

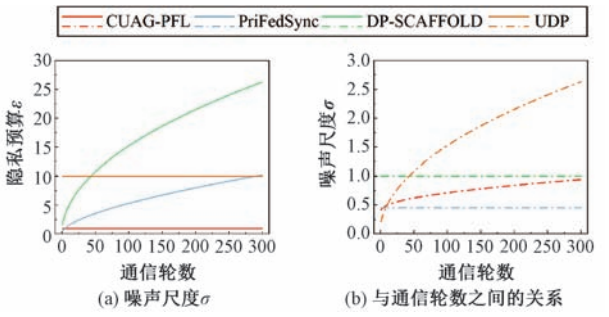

图 4 CUAG-PFL、PriFedSync[21]、DP-SCAFFOLD[25]和 UDP[24]中的隐私预算 ε

效用性：从图 5 和图 6 中可以明显看出，对于参与方数据 Non-IID 的场景，非个性化 FL 方法（即 FedAvg 和 UDP）训练的共享全局模型准确率较差. 相反，基于前沿的个性化 FL 方法 PFLDyn(Proto)改进的隐私保护方法 CUAG-PFL 和 PriFedSync 在效用性方面都比其它基于 LDP 的 FL 方法更优. 尤

表 1　对于 CIFAR-10 联邦数据集,ACID3、5、7 三种场景的 CUAG-PFL、PriFedSync[21]、DP-SCAFFOLD[25] 和 UDP[24] 在目标准确率下的隐私预算 ε

FL 场景	目标准确率	CUAG-PFL	PriFedSync[21]	DP-SCAFFOLD[25]	UDP[24]
	0.5528	$\varepsilon = 0.92$	$\varepsilon = 0.41$	$\varepsilon = 5.03$	$\varepsilon = 10$
ACID3	0.7010	$\varepsilon = 0.92$	$\varepsilon = 1.99$	$\varepsilon = 23.31$	\
	0.9016	$\varepsilon = 0.92$	$\varepsilon = 10.15$	\	\
	0.5981	$\varepsilon = 0.92$	$\varepsilon = 1.93$	$\varepsilon = 15.77$	$\varepsilon = 10$
ACID5	0.6286	$\varepsilon = 0.92$	$\varepsilon = 2.28$	\	$\varepsilon = 10$
	0.8592	$\varepsilon = 0.92$	$\varepsilon = 10.15$	\	\
	0.5480	$\varepsilon = 0.92$	$\varepsilon = 2.23$	$\varepsilon = 15.25$	$\varepsilon = 10$
ACID7	0.6632	$\varepsilon = 0.92$	$\varepsilon = 3.66$	\	$\varepsilon = 10$
	0.8298	$\varepsilon = 0.92$	$\varepsilon = 10.21$	\	\

注:其中每一种场景中的三个目标准确率分别是 UDP、DP-SCAFFOLD 和 PriFedSync 在全局通信总轮数 $T = 300$ 时的最高准确率

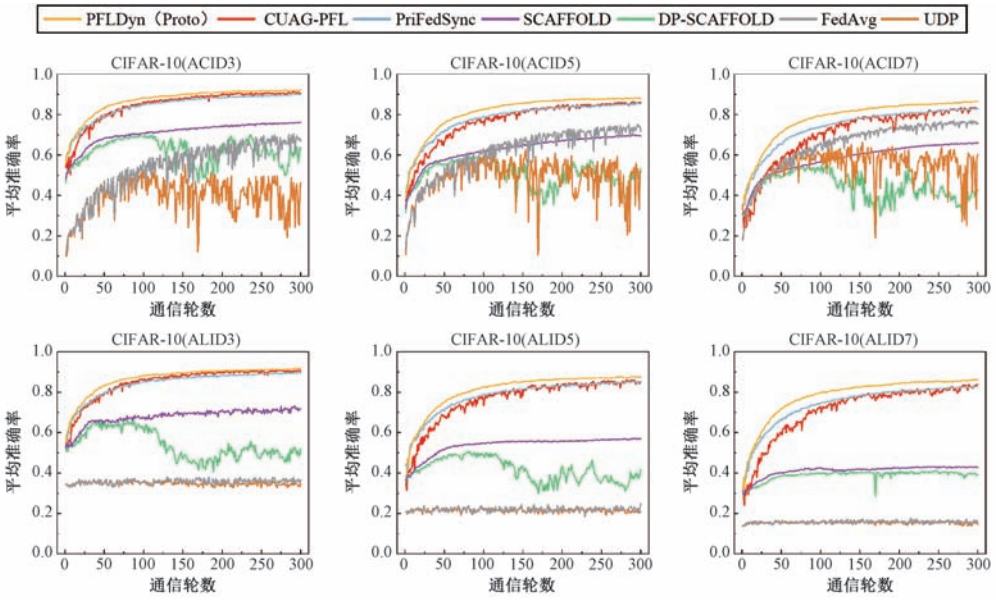

图 5　对于 CIFAR-10 联邦数据集,六种 Non-IID 数据分割方式下 CUAG-PFL 与 FedAvg[8]、PriFedSync[21]、UDP[24]、DP-SCAFFOLD[25]、SCAFFOLD[32]、PFLDyn(Proto)[44] 六种对比方法的平均准确率曲线

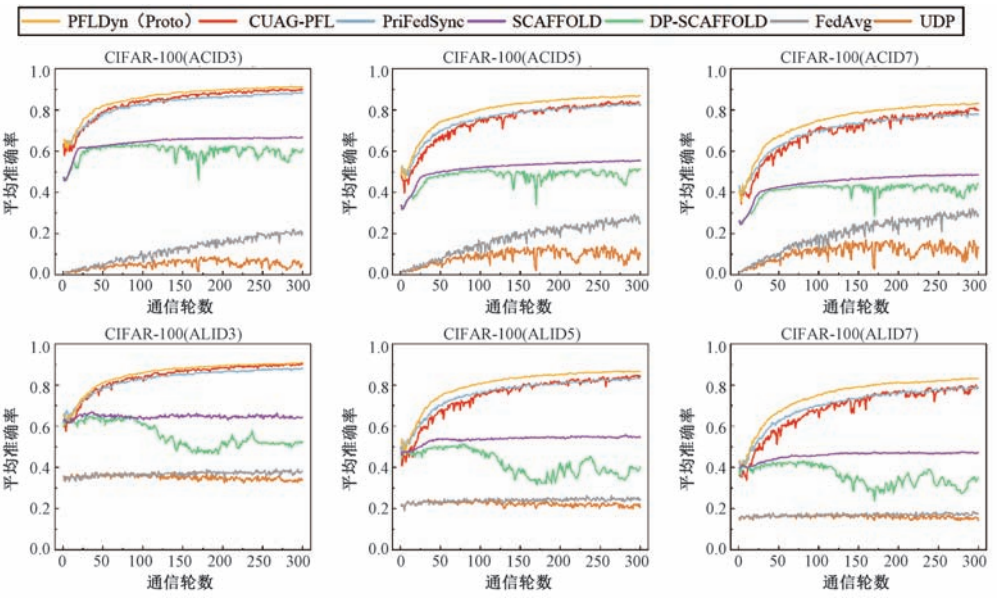

图 6　对于 CIFAR-100 联邦数据集,六种 Non-IID 数据分割方式下 CUAG-PFL 与 FedAvg[8]、PriFedSync[21]、UDP[24]、DP-SCAFFOLD[25]、SCAFFOLD[32]、PFLDyn(Proto)[44] 六种对比方法的平均准确率曲线

其是在 ACID3 和 ALID3 这两种 Non-IID 数据分割方式下，CUAG-PFL 和 PriFedSync 定制的个性化模型平均效用几乎都接近二者对应的模型效用上限对应的个性化 FL 方法 PFLDyn(Proto). 例如对于 CIFAR-100 联邦数据集，在 ALID3 这种 Non-IID 数据分割方式下，CUAG-PFL 与 PFLDyn(Proto) 之间的模型准确率差值仅为 0.47%.

通信效率：在 FedAvg、UDP、PFLDyn(Proto) 和 PriFedSync 四种方法中，各个参与方和服务器之间双向通信的模型参数(或梯度)都是原始维度，对应于图 7 中的无压缩情形. 而在 SCAFFOLD 和 DP-SCAFFOLD 这两种方法中，每一轮全局通信的服务器和参与方之间需要同时传输模型参数更新和梯度，翻倍增加了通信开销，即全局通信量是图 7 中无压缩情形下全局通信量的两倍. 对于本文提出的 CUAG-PFL 方法，每一个参与方都是传输压缩梯度至服务器. 根据第 7.2.4 节中所述，需要上传的模型梯度每一层最高压缩率仅为 0.5，因此 CUAG-PFL 大幅度地减少了通信开销，尤其是提高了上行通信效率. 如图 7 所示，对于 CIFAR-10 联邦数据集，相比于四种无压缩的方法(即 FedAvg、UDP、PFLDyn(Proto)和 PriFedSync)，CUAG-PFL 在 ALID7 这种数据分割方式下的上行通信量减少了 69.33%. 此外，因为 CUAG-PFL 中的各个参与方对上传的本地模型梯度都进行动态层级压缩，所以即使固定最高压缩率和最低压缩率，对于任何一个联邦数据集，CUAG-PFL 在不同的 Non-IID 数据分割方式下的通信开销都不完全相同.

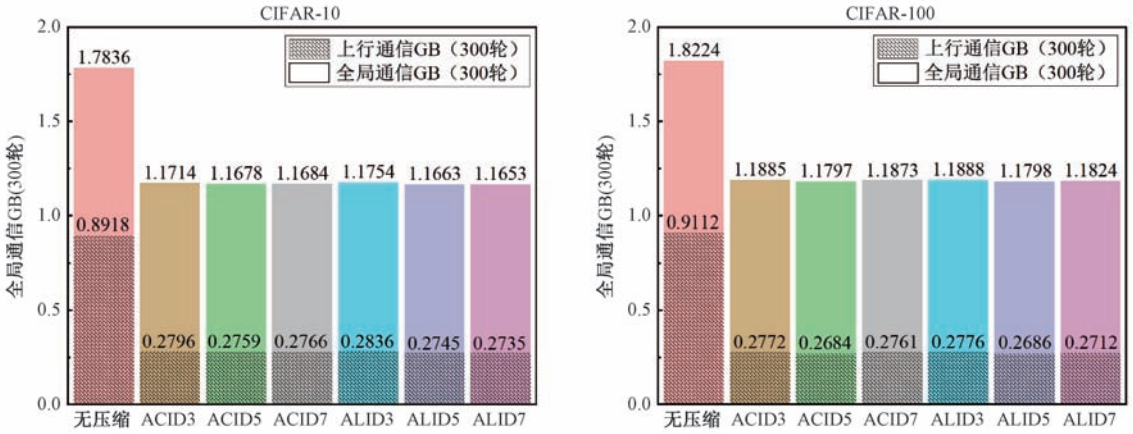

图 7 对于 CIFAR-10 和 CIFAR-100 两种联邦数据集，六种 Non-IID 数据分割方式下 CUAG-PFL 与无压缩的方法在通信 300 轮时的上行通信开销和全局通信开销

个性化：正如第 7.2.3 节中所述，CUAG-PFL 与 PriFedSync、PFLDyn(Proto)、DP-SCAFFOLD、SCAFFOLD 这四种方法都能够为参与方定制一个本地个性化模型.

综合而言，如表 2 所示，显然只有 CUAG-PFL、PriFedSync 和 DP-SCAFFOLD 是能够保护参与方数据隐私的三种个性化 FL 方法. 然而，在效用性方面，DP-SCAFFOLD 无法较好地解决参与方数据 Non-IID 问题. 相比于 PriFedSync，CUAG-PFL 中参与方的 ε 始终固定为 0.92，即本地数据隐私保护程度一直处于较强的状态. 此外，CUAG-PFL 显著提高了联邦系统的通信效率. 以 ACID3 方式分割 CIFAR-10 联邦数据集为例，CUAG-PFL 即使在 ε 仅为 0.92 且上行通信量减少 68.6%的情况下，其与 PFLDyn(Proto) 之间的模型准确率只相差 1.66%，即因隐私保护和梯度压缩所引起的模型效用损失仅为 1.66%. 由此说明，本文提出的 CUAG-PFL 能减少通信开销和提高本地数据隐私保护程度，并且为各个参与方定制一个较好拟合其本地数据分布的个性化模型.

表 2 CUAG-PFL 与其它六种方法的总结与比较，其中效用性和通信效率(上行通信开销)以 ACID3 方式分割 CIFAR-10 联邦数据集为例

FL 方法	隐私性 ($\delta = 10^{-5}$)	效用性	通信效率 (上行通信开销)	个性化
CUAG-PFL	$\varepsilon = 0.92$	0.9046	0.2796 GB	√
PriFedSync[21]	$\varepsilon \in [0.41, 10.24]$	0.8986	0.8918 GB	√
PFLDyn(Proto)[44]	×	0.9212	0.8918 GB	√
DP-SCAFFOLD[25]	$\varepsilon \in [1.52, 26.28]$	0.6275	1.7836 GB	√
SCAFFOLD[32]	×	0.7624	1.7836 GB	√
UDP[24]	$\varepsilon = 10$	0.4624	0.8918 GB	×
FedAvg[8]	×	0.6686	0.8918 GB	×

7.4 进一步研究与分析

7.4.1 隐私参数 μ 对模型效用的影响

为了进一步表明 CUAG-PFL 同时具有较强的

隐私性和效用性,这里设置六个不同的隐私参数 μ,即{0.1, 0.15, 0.2, 0.5, 1, 2}.根据定理 3,固定 $\delta=10^{-5}$ 时,这六个隐私参数 μ 值对应的隐私预算 ε 值是{0.34, 0.53, 0.73, 1.99, 4.38, 10}.

对于 CIFAR-100 联邦数据集在六种 Non-IID 数据分割方式下的情形,表 3 和表 4 分别给出了应用不同模型的 CUAG-PFL 在这六个不同 μ 值的设置下所对应的模型效用.表 3 的模型架构 CNNs 见第 7.2.2 节,表 4 的模型架构则是具有代表性的 VGG16[52].从表 3 和表 4 中可以明显看出,无论在哪种 Non-IID 数据分割方式情况下,六个 μ 对应的模型准确率之间的差值都非常小.尤其是当模型架构是 VGG16 且数据分割方式是 ALID3 时,隐私参数 $\mu=0.1$(即 $\varepsilon=0.34$)对应的模型准确率与这六个隐私参数 μ 对应的最优模型准确率之间的差值仅仅是 0.03%.由此说明,本文提出的 CUAG-PFL 方法能够同时提高参与方数据隐私性和模型效用.

表 3　对于 CIFAR-100 联邦数据集,六种 Non-IID 数据分割方式下不同隐私参数 μ 对 CUAG-PFL 模型(模型架构是 CNNs)效用性的影响

FL 场景	$\mu=0.1$	$\mu=0.15$	$\mu=0.2$	$\mu=0.5$	$\mu=1$	$\mu=2$
ACID3	0.8860	0.8984	0.8977	0.8965	0.8862	0.9038
ACID5	0.8308	0.8512	0.8425	0.8485	0.8451	0.8363
ACID7	0.7881	0.7851	0.7974	0.7785	0.7892	0.7933
ALID3	0.8958	0.9015	0.9022	0.9053	0.9031	0.8986
ALID5	0.8414	0.8334	0.8460	0.8550	0.8489	0.8538
ALID7	0.7813	0.7998	0.8004	0.8016	0.8008	0.7916

表 4　对于 CIFAR-100 联邦数据集,六种 Non-IID 数据分割方式下不同隐私参数 μ 对 CUAG-PFL 模型(模型架构是 VGG16[52])效用性的影响

FL 场景	$\mu=0.1$	$\mu=0.15$	$\mu=0.2$	$\mu=0.5$	$\mu=1$	$\mu=2$
ACID3	0.8921	0.8974	0.8913	0.8930	0.9066	0.8911
ACID5	0.8241	0.8259	0.8326	0.8376	0.8368	0.8331
ACID7	0.8057	0.8047	0.8037	0.8170	0.8106	0.8112
ALID3	0.9164	0.9151	0.9167	0.9133	0.9135	0.9156
ALID5	0.8486	0.8476	0.8471	0.8462	0.8611	0.8584
ALID7	0.8147	0.8174	0.8123	0.8239	0.8242	0.8367

7.4.2　自适应 GDP 策略对模型效用的影响

在本文提出的 CUAG-PFL 中,自适应 GDP 的现象主要是通过同时自适应裁剪阈值(或者对应的敏感度)和噪声尺度来呈现的.正如第 6 节中所述,当隐私参数 $\mu=0.25$(即噪声尺度 σ 自适应)时,参与方本地数据隐私保护程度始终不变.然而,当噪声尺度 σ 固定时,根据定理 4,参与方本地数据隐私保护程度随着全局通信轮数的增加而降低,即隐私参数 μ 减小.此外,某一轮全局通信中的 σ 越大,则 μ 越小,参与方本地数据隐私保护程度越强.为了验证本文提出的自适应裁剪阈值策略能够减少噪声尺度太大所引起的模型准确率降低,这里考虑以下四种情况:(1)同时固定 σ 和裁剪阈值 C;(2)固定 σ 和自适应裁剪阈值 C;(3)自适应 σ 和固定裁剪阈值 C;(4)同时自适应 σ 和 C.以 CIFAR-10 联邦数据集为例,将以上四种情况应用到本文提出的 CUAG-PFL 中,表 5 给出了六种 Non-IID 数据分割方式下 σ 和 C 是否固定对模型效用的影响.

如表 5 所示,无论 σ 是否固定,显然裁剪阈值 C 自适应时所对应的模型准确率都更高.此外,当 $\sigma=1.0$ 或者 σ 自适应时,裁剪阈值 C 固定对模型效用的损失更大.尤其是当 $\sigma=1.0$ 时,在 ALID5 这种 Non-IID 数据分割方式下,自适应 C 和固定 $C=1.5$ 所对应的模型准确率之差最大,其中前者比后者高 14.16%.如果固定 C,当 σ 越小,模型性能明显相对越优.然而,如果自适应 C,那么当 σ 从 0.45 增加到 1.0 时,模型效用损失远小于固定 C 的情况.由此可见,本文提出的自适应裁剪阈值策略对模型效用有益.

表 5　对于 CIFAR-10 联邦数据集,六种 Non-IID 数据分割方式下噪声尺度 σ 和裁剪阈值 C 是否固定对模型效用性的影响

FL 场景	$\sigma=0.45$				$\sigma=0.65$				$\sigma=1.0$				$\mu=0.25$(即 σ 自适应)			
	$C=0.5$	$C=1.0$	$C=1.5$	C 自适应	$C=0.5$	$C=1.0$	$C=1.5$	C 自适应	$C=0.5$	$C=1.0$	$C=1.5$	C 自适应	$C=0.5$	$C=1.0$	$C=1.5$	C 自适应
ACID3	0.9051	0.9069	0.9059	0.9189	0.8982	0.9024	0.8961	0.9101	0.8491	0.8543	0.8554	0.9045	0.8787	0.8828	0.8823	0.9046
ACID5	0.8490	0.8442	0.8452	0.8666	0.8459	0.8441	0.8377	0.8615	0.7441	0.7491	0.7423	0.8506	0.8314	0.8288	0.8332	0.8566
ACID7	0.8132	0.8182	0.8178	0.8291	0.7990	0.8069	0.7991	0.8263	0.7484	0.7553	0.7527	0.8240	0.7942	0.7899	0.7907	0.8282
ALID3	0.8980	0.9022	0.8879	0.9067	0.8974	0.9009	0.8874	0.9036	0.8703	0.8613	0.8626	0.9018	0.8839	0.8760	0.8807	0.9026
ALID5	0.8510	0.8500	0.8552	0.8625	0.8452	0.8389	0.8454	0.8580	0.7099	0.7156	0.7081	0.8497	0.8095	0.8193	0.8189	0.8517
ALID7	0.8216	0.8222	0.8178	0.8467	0.7981	0.8027	0.7961	0.8466	0.7448	0.7485	0.7364	0.8336	0.7945	0.8033	0.7964	0.8381

7.4.3　动态层级压缩率 vs. 固定压缩率

本节讨论动态层级压缩梯度方案对于 FL 通信效率和模型效用的影响.如表 6 所示,这里给出了 ACID3、5、7 这三种分割 CIFAR-10 联邦数据集的场景下动态层级压缩率($r_{\max}=0.5$ 且 $r_{\min}=0.2$)和三种固定压缩率 $r=0.5$、0.4、0.3 时 300 轮通信的

上行链路通信总量和对应的模型精度.

无论模型架构是选择第 7.2.2 节中规模较小的模型 CNNs 还是具有代表性且规模较大的 VGG16，显然三种分割 CIFAR-10 联邦数据集场景下的动态层级压缩率所对应的上行链路通信总量都不同. 当模型架构是 VGG16 时，CUAG-PFL 中的梯度压缩率是否固定在上行通信效率和模型效用两个方面的差异化更加显著. 尤其是当数据分割方式是 ACID5，动态层级压缩率所需的上行通信量仅占固定压缩率为 0.5 时所需的上行通信量的 60%，同时前者的模型准确率比后者高 3.44%. 由此说明本文设计的动态层级压缩梯度方案不仅大幅度提高通信效率，适应不同 FL 场景动态变化，而且模型效用性也较强.

表 6 对于 CIFAR-10 联邦数据集，ACID3、5、7 三种场景的动态层级压缩率和固定压缩率对 CUAG-PFL 上行通信效率和模型效用性的影响

模型架构	压缩率	ACID3		ACID5		ACID7	
		上行通信量/GB	准确率	上行通信量/GB	准确率	上行通信量/GB	准确率
CNNs	$r \in (0.2, 0.5]$	0.2796	0.9046	0.2760	0.8566	0.2766	0.8282
	$r = 0.5$	0.4459	0.9092	0.4459	0.8583	0.4459	0.8277
	$r = 0.4$	0.3567	0.9055	0.3567	0.8505	0.3567	0.8124
	$r = 0.3$	0.2675	0.9001	0.2675	0.8462	0.2675	0.8063
VGG16	$r \in (0.2, 0.5]$	45.1137	0.9173	45.0496	0.8763	45.0874	0.8441
	$r = 0.5$	75.0468	0.8939	75.0468	0.8419	75.0468	0.8109
	$r = 0.4$	60.0375	0.8803	60.0375	0.8243	60.0375	0.7899
	$r = 0.3$	45.0281	0.8637	45.0281	0.8087	45.0281	0.7731

7.4.4 不同测量矩阵对模型效用的影响

本文将高斯随机矩阵 $\boldsymbol{\phi}_G$、伯努利矩阵 $\boldsymbol{\phi}_B$ 以及确定性二进制测量矩阵 $\boldsymbol{\phi}$ 依次作为 CUAG-PFL 中动态层级压缩梯度的测量矩阵来验证不同测量矩阵对定制的参与方个性化模型平均准确率的影响. 如表 7 所示，对于 CIFAR-10 和 CIFAR-100 两种联邦数据集在六种 Non-IID 数据分割方式下的共 12 种不同情形，测量矩阵是 $\boldsymbol{\phi}$ 的 CUAG-PFL 在每一种参与方数据 Non-IID 情形下的模型准确率都是最高的. 尤其是对于 CIFAR-100 联邦数据集，在 ACID3 这种 Non-IID 数据分割方式下，相比于常用的高斯随机矩阵 $\boldsymbol{\phi}_G$ 和伯努利矩阵 $\boldsymbol{\phi}_B$，$\boldsymbol{\phi}$ 能够减少的模型效用损失分别为 4.32% 和 4.41%. 因此，本文设计的确定性二进制测量矩阵能够最大限度地保留有效梯度信息，并且有益于减少梯度压缩所引起的模型性能退化.

表 7 本文设计的确定性二进制测量矩阵 $\boldsymbol{\phi}$、高斯随机矩阵 $\boldsymbol{\phi}_G$ 和伯努利矩阵 $\boldsymbol{\phi}_B$ 对 CUAG-PFL 模型效用性的影响

测量矩阵	CIFAR-10						CIFAR-100					
	ACID3	ACID5	ACID7	ALID3	ALID5	ALID7	ACID3	ACID5	ACID7	ALID3	ALID5	ALID7
$\boldsymbol{\phi}$	0.9046	0.8566	0.8282	0.9026	0.8517	0.8381	0.8963	0.8238	0.7998	0.9037	0.8426	0.7866
$\boldsymbol{\phi}_G$	0.8820	0.8219	0.7874	0.8722	0.8098	0.7993	0.8531	0.8003	0.7724	0.8617	0.8174	0.7471
$\boldsymbol{\phi}_B$	0.8710	0.8158	0.7938	0.8699	0.8139	0.7947	0.8522	0.7916	0.7602	0.8686	0.8120	0.7534

7.4.5 异构 DP 对模型效用的影响

在现实场景中，除了参与方之间的本地数据是 Non-IID 之外，各个参与方对于本地数据隐私保护程度的需求亦可能完全不同，即基于 LDP 的联邦系统还存在异构 DP 问题. 本小节简单讨论了异构 DP 对 CUAG-PFL 效用性的影响. 这里设置 100 个参与方的隐私参数 μ，最大值是 4 且最小值是 0.1. 先等间隔地划分区间[0.1, 4]来获取 100 个 μ 值，再将它们随机分配给各个参与方. 基于此异构 DP 的设置，对于 CIFAR-10 和 CIFAR-100 两种联邦数据集在六种 Non-IID 数据分割方式下的情形，CUAG-PFL 定制的个性化模型平均效用如表 8 所示. 结合表 3 和表 4 可知，对所有参与方强制执行相同的数据隐私保护级别，可能会导致模型效用产生不必要的降级. 为了对联邦系统中的各个参与方数据进行更加细粒度的隐私保护，未来的工作可以进一步优化 CUAG-PFL 来解决异构 DP 问题.

表 8 对于 CIFAR-10 和 CIFAR-100 两种联邦数据集，六种 Non-IID 数据分割方式下异构 DP($\mu_n \in [0.1, 4]$ 且 $n \in [1, 100]$)对 CUAG-PFL 模型效用性的影响

FL 场景	ACID3	ACID5	ACID7	ALID3	ALID5	ALID7
CIFAR-10	0.9087	0.8398	0.8289	0.8989	0.8476	0.8391
CIFAR-100	0.9066	0.8418	0.7830	0.8995	0.8386	0.7966

8　结束语

本文提出的 CUAG-PFL 方法能够同时有效地解决隐私性、效用性、通信效率和参与方数据 Non-IID 四个问题.具体而言,先对每一轮全局通信中需要上传的模型梯度每一层动态生成其特定的压缩率,再利用新设计的确定性二进制测量矩阵对模型梯度压缩,最大限度保留有效梯度信息,从而提高通信效率和模型效用.随后对压缩的模型梯度执行自适应 GDP 操作,通过优化裁剪阈值、敏感度和噪声尺度等隐私相关参数来提高参与方数据的隐私性和个性化模型的准确率.除此之外,本文将动态层级压缩方案和自适应 GDP 策略应用于新的个性化 FL 框架 PFLDyn(Proto)中来解决参与方数据 Non-IID 情况下的通信效率和隐私保护问题.大量实验对比和分析证明了 CUAG-PFL 是一种高效率、高效用的个性化 FL 隐私保护方法.然而,CUAG-PFL 仅强调自适应压缩单个模型梯度各层的内部冗余,并且忽略了现实场景中参与方之间可能存在不同的隐私需求.未来的工作将探索相邻全局通信中参与方模型梯度之间的冗余,并且设计更加精细的隐私策略来满足各个参与方的隐私期望.

参　考　文　献

[1] Zhao Jing-Xin, Yue Xing-Hui, Feng Chong-Peng, Zhang Jing, Li Yin, Wang Na, Ren Jia-Dong, Zhang Hao-Xing, Wu Gao-Fei, Zhu Xiao-Yan, Zhang Yu-Qing. Survey of data privacy security based on general data protection regulation. Journal of Computer Research and Development, 2022, 59(10): 2130-2163 (in Chinese)
(赵景欣, 岳星辉, 冯崇朋, 张静, 李印, 王娜, 任家东, 张昊星, 伍高飞, 朱笑岩, 张玉清. 基于通用数据保护条例的数据隐私安全综述. 计算机研究与发展, 2022, 59(10): 2130-2163)

[2] Li T, Sahu A K, Talwalkar A, Smith V. Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 2020, 37(3): 50-60

[3] Zhang T, Gao L, He C, Zhang M, Krishnamachari B, Avestimehr A S. Federated learning for the internet of things: Applications, challenges, and opportunities. IEEE Internet of Things Magazine, 2022, 5(1): 24-29

[4] Alazab M, RM S P, M P, Maddikunta P K R, Gadekallu T R, Pham Q-V. Federated learning for cybersecurity: Concepts, challenges, and future directions. IEEE Transactions on Industrial Informatics, 2022, 18(5): 3501-3509

[5] Rasha A H, Li T, Huang W, Gu J, Li C. Federated learning in smart cities: Privacy and security survey. Information Sciences, 2023, 632: 833-857

[6] Melis L, Song C, Cristofaro E D, Shmatikov V. Exploiting unintended feature leakage in collaborative learning//Proceedings of the 40th IEEE Symposium on Security and Privacy (SP). San Francisco, USA, 2019: 691-706

[7] Zhu L, Liu Z, Han S. Deep leakage from gradients//Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS). Vancouver, Canada, 2019: 14774-14784

[8] McMahan B, Moore E, Ramage D, Hampson S, y Arcas B A. Communication-efficient learning of deep networks from decentralized data//Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS). Fort Lauderdale, USA, 2017, 54: 1273-1282

[9] Ji Shou-Ling, Du Tian-Yu, Li Jin-Feng, Shen Chao, Li Bo. Security and privacy of machine learning models: A survey. Journal of Software, 2021, 32(01): 41-67 (in Chinese)
(纪守领, 杜天宇, 李进锋, 沈超, 李博. 机器学习模型安全与隐私研究综述. 软件学报, 2021, 32(01): 41-67)

[10] Zhou Chun-Yi, Chen Da-Wei, Wang Shang, Fu An-Min, Gao Yan-Song. Research and challenge of distributed deep learning privacy and security attack. Journal of Computer Research and Development, 2021, 58(05): 927-943 (in Chinese)
(周纯毅, 陈大卫, 王尚, 付安民, 高艳松. 分布式深度学习隐私与安全攻击研究进展与挑战. 计算机研究与发展, 2021, 58(05): 927-943)

[11] Qian Wen-Jun, Shen Qing-Ni, Wu Peng-Fei, Dong Chun-Tao, Wu Zhong-Hai. Research progress on privacy-preserving techniques in big data computing environment. Chinese Journal of Computers, 2022, 45(4): 669-701 (in Chinese)
(钱文君, 沈晴霓, 吴鹏飞, 董春涛, 吴中海. 大数据计算环境下的隐私保护技术研究进展. 计算机学报, 2022, 45(4): 669-701)

[12] Xiao Xiong, Tang Zhuo, Xiao Bin, Li Ken-Li. A survey on privacy and security issues in federated learning. Chinese Journal of Computers, 2023, 46(5): 1019-1044 (in Chinese)
(肖雄, 唐卓, 肖斌, 李肯立. 联邦学习的隐私保护与安全防御研究综述. 计算机学报, 2023, 46(5): 1019-1044)

[13] Yin X, Zhu Y, Hu J. A comprehensive survey of privacy-preserving federated learning: A taxonomy, review, and future directions. ACM Computing Surveys, 2021, 54(6): 1-36

[14] Abadi M, Chu A, Goodfellow I, McMahan H B, Mironov I, Talwar K, Zhang L. Deep learning with differential privacy//Proceedings of the 23rd ACM SIGSAC Conference on Computer and Communications Security (CCS). Vienna, Austria, 2016: 308-318

[15] Cormode G, Jha S, Kulkarni T, Li N, Srivastava D, Wang T. Privacy at scale: Local differential privacy in practice//

Proceedings of the 44th International Conference on Management of Data (SIGMOD). Houston, USA, 2018: 1655-1658

[16] Wang N, Xiao X, Yang Y, Zhao J, Hui S C, Shin H, Shin J, Yu G. Collecting and analyzing multidimensional data with local differential privacy//Proceedings of the 35th International Conference on Data Engineering (ICDE). Macao, China, 2019: 638-649

[17] Wang Y, Tong Y, Shi D. Federated latent dirichlet allocation: a local differential privacy based framework//Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI). New York, USA, 2020, 34(04): 6283-6290

[18] Truex S, Liu L, Chow K H, Gursoy M E, Wei W. LDP-Fed: Federated learning with local differential privacy// Proceedings of the 3rd ACM International Workshop on Edge Systems, Analytics and Networking (EdgeSys). Heraklion, Greece, 2020: 61-66

[19] Lu Y, Huang X, Dai Y, Maharjan S, Zhang Y. Differentially private asynchronous federated learning for mobile edge computing in urban informatics. IEEE Transactions on Industrial Informatics, 2020, 16(3): 2134-2143

[20] Sonee A, Rini S, Huang Y C. Wireless federated learning with limited communication and differential privacy// Proceedings of the 64th IEEE Global Communications Conference (GLOBECOM). Madrid, Spain, 2021: 1-6

[21] Zheng Q, Chen S, Long Q, Su W. Federated f-differential privacy//Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS). San Diego, USA, 2021, 130: 2251-2259

[22] Wu M, Ye D, Ding J, Guo Y, Yu R, Pan M. Incentivizing differentially private federated learning: A multidimensional contract approach. IEEE Internet of Things Journal, 2021, 8 (13): 10639-10651

[23] Cui L, Qu Y, Xie G, Zeng D, Li R, Shen S, Yu S. Security and privacy-enhanced federated learning for anomaly detection in IoT infrastructures. IEEE Transactions on Industrial Informatics, 2022, 18(5): 3492-3500

[24] Wei K, Li J, Ding M, Ma C, Su H, Zhang B, Poor H V. User-level privacy-preserving federated learning: Analysis and performance optimization. IEEE Transactions on Mobile Computing, 2022, 21(9): 3388-3401

[25] Noble M, Bellet A, Dieuleveut A. Differentially private federated learning on heterogeneous data//Proceedings of the 25th International Conference on Artificial Intelligence and Statistics (AISTATS). Valencia, Spain, 2022, 151: 10110-10145

[26] Andrew G, Thakkar O, McMahan B, Ramaswamy S. Differentially private learning with adaptive clipping// Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS). Montreal, Canada, 2021, 34: 17455-17466

[27] Dong J, Roth A, Su W J. Gaussian differential privacy. Journal of the Royal Statistical Society Series B: Statistical Methodology. 2022, 84(1): 3-37

[28] Bu Z, Dong J, Long Q, Su W. Deep learning with gaussian differential privacy. Harvard Data Science Review, 2020, 2020(23): 1-32

[29] Bietti A, Wei C, Dudik M, Langford J, Wu S. Personalization improves privacy-accuracy tradeoffs in federated learning//Proceedings of the 39th International Conference on Machine Learning (ICML). Baltimore, USA, 2022, 162: 1945-1962

[30] Yang Y, Hui B, Yuan H, Gong N, Cao Y. PrivateFL: Accurate, differentially private federated learning via personalized data transformation//Proceedings of the 32nd International Conference on USENIX Security Symposium. Anaheim, USA, 2023: 1-17

[31] Wei W, Liu L. Gradient leakage attack resilient deep learning. IEEE Transactions on Information Forensics and Security, 2022, 17: 303-316

[32] Karimireddy S P, Kale S, Mohri M, Reddi S, Stich S, Suresh A T. SCAFFOLD: Stochastic controlled averaging for federated learning//Proceedings of the 37th International Conference on Machine Learning (ICML). Online, 2020, 119: 5132-5143

[33] Shin H, Kim S, Shin J, Xiao X. Privacy enhanced matrix factorization for recommendation with local differential privacy. IEEE Transactions on Knowledge and Data Engineering, 2018, 30(9): 1770-1782

[34] Liu R, Cao Y, Yoshikawa M, Chen H. Fedsel: Federated sgd under local differential privacy with top-k dimension selection//Proceedings of the 25th International Conference on Database Systems for Advanced Applications (DASFAA). Jeju, Republic of Korea, 2020: 485-501

[35] Cui L, Ma J, Zhou Y, Yu S. Boosting accuracy of differentially private federated learning in industrial IoT with sparse responses. IEEE Transactions on Industrial Informatics, 2023, 19(1): 910-920

[36] Kerkouche R, Ács G, Castelluccia C, Genevès P. Compression boosts differentially private federated learning//Proceedings of the 6th IEEE European Symposium on Security and Privacy (EuroS&P). Vienna, Austria, 2021: 304-318

[37] Farokhi F. Gradient sparsification can improve performance of differentially-private convex machine learning//Proceedings of the 60th IEEE Conference on Decision and Control (CDC). Austin, USA, 2021: 1695-1700

[38] Miao Y, Chen S. Efficient privacy-preserving federated learning against inference attacks for IoT//Proceedings of the 21st IEEE Wireless Communications and Networking Conference (WCNC). Glasgow, UK, 2023: 1-6

[39] Donoho D L. Compressed sensing. IEEE Transactions on Information Theory. 2006, 52(4): 1289-1306

[40] Dai Qiong-Hai, Fu Chang-Jun, Ji Xiang-Yang. Research on compressed sensing. Chinese Journal of Computers, 2011, 34(3): 3425-3434 (in Chinese)

(戴琼海, 付长军, 季向阳. 压缩感知研究. 计算机学报, 2011, 34(3): 3425-3434)

[41] Ding X, Chen W, Wassell I J. Joint sensing matrix and sparsifying dictionary optimization for tensor compressive sensing. IEEE Transactions on Signal Process. 2017, 65(14): 3632-3646

[42] Fang H, Vorobyov S A, Jiang H, Taheri O. Permutation meets parallel compressed sensing: How to relax restricted isometry property for 2D sparse signals. IEEE Transactions on Signal Process. 2014, 62(1): 196-210

[43] Collins L, Hassani H, Mokhtari A, Shakkottai S. Exploiting shared representations for personalized federated learning// Proceedings of the 38th International Conference on Machine Learning (ICML). Online, 2021, 139: 2089-2099

[44] Acar D A E, Zhao Y, Zhu R, Matas R, Mattina M, Whatmough P, Saligrama V. Debiasing model updates for improving personalized federated training//Proceedings of the 38th International Conference on Machine Learning (ICML). Online, 2021, 139: 21-31

[45] Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks//Proceedings of the 34th International Conference on Machine Learning (ICML). Sydney, Australia, 2017, 70: 1126-1135

[46] Snell J, Swersky K, Zemel R. Prototypical networks for few-shot learning//Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS). Long Beach, USA, 2017: 4080-4090

[47] Niu Jun, Ma Xiao-Ji, Chen Ying, Zhang Ge, He Zhi-Peng, Hou Zhe-Xian, Zhu Xiao-Yan, Wu Gao-Fei, Chen Kai, Zhang Yu-Qing. A survey on membership inference attacks and defenses in machine learning. Journal of Cyber Security, 2022, 7(06): 1-30 (in Chinese)
(牛俊, 马骁骥, 陈颖, 张歌, 何志鹏, 侯哲贤, 朱笑岩, 伍高飞, 陈恺, 张玉清. 机器学习中成员推理攻击和防御研究综述. 信息安全学报, 2022, 7(06): 1-30)

[48] Gu X, Li M, Xiong L. PRECAD: Privacy-preserving and robust federated learning via crypto-aided differential privacy. arXiv preprint arXiv:2110.11578, 2021

[49] Li C, Li G, Varshney P K. Communication-efficient federated learning based on compressed sensing. IEEE Internet of Things Journal. 2021, 8(20): 15531-15541

[50] He Y, Kang G, Dong X, Fu Y, Yang Y. Soft filter pruning for accelerating deep convolutional neural networks// Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI). Stockholm, Sweden, 2018: 2234-2240

[51] Candès E J. The restricted isometry property and its implications for compressed sensing. Comptes Rendus Mathematique. 2008, 346(9-10): 589-592

[52] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014

LI Min, Ph. D. candidate. Her main research interests include federated learning, privacy preserving, compressed sensing and differential privacy.

XIAO Di, Ph. D., professor. His main research interests include privacy preserving and data security.

CHEN Lü-Jun, Ph. D. candidate. Her main research interests include privacy preserving, secure multi-party computation and federated learning.

Background

Federated learning (FL) is a forward-looking distributed machine learning paradigm, where remote devices or isolated data centers can collaboratively train a shared global model for prediction, classification, and other functions while ensuring the localization of private data. However, there still exists a significant risk of privacy leakage when directly communicating model parameters (gradients) among the server and participants. As a lightweight data privacy protection mechanism, local differential privacy (LDP) is widely used in FL. However, existing LDP-based FL research works face many problems and challenges, such as weakened trade-offs between the data privacy and the model utility, lower communication efficiency, and difficulties in real-world applications of FL.

Specifically, although LDP-based FL can theoretically provide strict privacy protection for the sensitive data of local participants, the majority of relevant research works adopt a strategy of fixing various privacy-related parameters, which leads to excessive noise, increases the impact on the degradation of model utility, and is also susceptible to gradient leakage attacks. In addition, although compressed sensing (CS) has been proved to improve both the model utility and the communication efficiency of LDP-based FL without compromising privacy, the existing researches on CS-based FL mainly rely on fixed compression rates to compress each layer of the communication model. Obviously, an optimal balance between the communication efficiency and the model utility has not yet been achieved.

In response to the challenges mentioned above, and to effectively address the four issues of data privacy, model utility, communication efficiency, and non-independently and identically distributed (non-IID) data among participants, this paper proposes a novel communication-efficient and utility-aware adaptive Gaussian differential privacy for personalized FL method, called CUAG-PFL. In this paper, the specific compression ratio for each layer of uploaded gradients is dynamically generated, and then a new designed deterministic binary measurement matrix is used to compress the corresponding layer of model gradients. It not only preserves the most effective information of model gradients as much as possible, but also improves both the communication efficiency and the model utility. Subsequently, an adaptive Gaussian differential privacy (GDP) operation is performed on compressed gradients. Both the local data privacy and the model utility are improved by optimizing privacy-related parameters such as the clipping threshold, the noise scale, and the sensitivity. Additionally, the dynamic hierarchical compression and the adaptive GDP mechanism are applied to the popular personalized FL framework PFLDyn (Proto) to address the non-IID data issue. Numerous experimental comparisons and analyses have proven that CUAG-PFL is a communication-efficient and utility-aware personalized FL privacy protection method, which promotes the practicality of FL.

This work was supported in part by the National Key R&D Program of China under Grant 2020YFB1805400, the National Natural Science Foundation of China under Grant 62072063, and the Project Supported by Graduate Student Research and Innovation Foundation of Chongqing, China under Grant CYB22063.