软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software [doi: 10.13328/j.cnki.jos.007032]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于 Bregman 散度和差分隐私的个性化联邦学习方法*

张少波[1], 张激勇[1], 朱更明[1], 龙赛琴[2], 李哲涛[2]

[1](湖南科技大学 计算机科学与工程学院, 湖南 湘潭 411201)

[2](暨南大学 信息科学技术学院, 广东 广州 510632)

通信作者: 龙赛琴, saiqinlong@jnu.edu.cn

摘　要: 联邦学习因能解决数据孤岛问题而被广泛关注, 但也存在用户隐私泄露风险和非独立同分布数据下模型异构导致性能下降的问题. 针对该问题, 提出基于 Bregman 散度和差分隐私的个性化联邦学习方法 (FedBDP). 所提方法采用 Bregman 散度衡量本地参数与全局参数的差异, 并将其作为正则化项更新损失函数, 以减小模型差异来提升模型准确率. 同时, 采用自适应差分隐私技术对本地模型参数进行扰动, 通过定义衰减系数动态调整每轮差分隐私噪声的大小, 以合理分配隐私噪声大小并提升模型可用性. 理论分析表明 FedBDP 在强凸和非凸光滑函数下满足收敛条件. 实验结果验证该方法在满足差分隐私的前提下, FedBDP 模型在 MNIST 和 CIFAR10 数据集下能够保证模型准确率.

关键词: 隐私保护; 个性化联邦学习; 差分隐私; Bregman 散度

中图法分类号: TP306

中文引用格式: 张少波, 张激勇, 朱更明, 龙赛琴, 李哲涛. 基于Bregman散度和差分隐私的个性化联邦学习方法. 软件学报. http://www.jos.org.cn/1000-9825/7032.htm

英文引用格式: Zhang SB, Zhang JY, Zhu GM, Long SQ, Li ZT. Personalized Federated Learning Method Based on Bregman Divergence and Differential Privacy. Ruan Jian Xue Bao/Journal of Software (in Chinese). http://www.jos.org.cn/1000-9825/7032.htm

Personalized Federated Learning Method Based on Bregman Divergence and Differential Privacy

ZHANG Shao-Bo[1], ZHANG Ji-Yong[1], ZHU Geng-Ming[1], LONG Sai-Qin[2], LI Zhe-Tao[2]

[1](School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411201, China)

[2](College of Information Science and Technology, Jinan University, Guangzhou 510632, China)

Abstract: Federated learning has caught much attention because it can solve data islands. However, it also faces challenges such as the risk of privacy leakage and performance degradation due to model heterogeneity under non-independent and identically distributed data. To this end, this study proposes a personalized federated learning method based on Bregman divergence and differential privacy (FedBDP). This method employs Bregman divergence to measure the differences between local and global parameters and adopt it as a regularization term to update the loss function, thereby reducing model differences to improve model accuracy. Meanwhile, adaptive differential privacy technology is utilized to perturb local model parameters, and the attenuation coefficient is defined to dynamically adjust the level of the differential privacy noise in each round, and thus reasonably allocate the privacy noise level and improve the model availability. Theoretical analysis shows that FedBDP satisfies convergence conditions under both strongly convex and non-convex smooth functions. Experimental results demonstrate that the FedBDP method can guarantee accuracy in the MNIST and CIFAR10 datasets on the premise of satisfying differential privacy.

Key words: privacy protection; personalized federated learning; differential privacy; Bregman divergence

---

* 基金项目: 国家重点研发计划 (2021YFB3101201); 国家自然科学基金 (62272162, 62172159, 62172350, 62032020); 教育部人文社会科学研究规划基金 (22YJAZH155); 湖南省自然科学基金 (2023JJ30267)
收稿时间: 2022-10-07; 修改时间: 2023-02-11; 采用时间: 2023-08-22; jos 在线出版时间: 2023-12-27

1 引 言

随着数据驱动智能应用的快速发展, 机器学习在自然语言处理[1]、计算机视觉[2]、智能物联网[3]、金融管理[4]和医疗卫生[5]等众多行业中已被广泛应用, 但它也面临着无法为所有用户提供稳健且高效的服务以及数据难以充分共享的难题[6], 而联邦学习 (federated learning, FL)[7−9]作为一种极具潜力的解决方法应运而生, 它是一种数据访问受限的分布式机器学习框架. 在联邦学习中, 分布式客户端根据自己的私有数据训练机器学习模型, 并借助参数服务器仅共享训练得到的梯度, 以此协作训练全局联邦模型, 整个过程实现了训练数据去中心化. 联邦学习体现了集中收集和数据最小化原则, 避免用户将自己的数据暴露给企业或其他参与方, 可以有效减少传统集中式机器学习带来的许多系统性隐私风险和开销问题[10]. 虽然联邦学习能解决数据孤岛问题而被广泛应用[11−13], 但它仍然面临许多挑战[14,15].

一方面, 安全性和隐私保护是联邦学习的核心问题. 目前已有研究表明攻击者可以根据上传的模型参数逆向推演出用户的原始数据[16], 这使得单一的依靠联邦学习来提高用户隐私变得困难. 在联邦学习训练过程或共享过程中, 需要精心设计加密技术对模型进行保护, 提高联邦学习的隐私保护能力. 为保护参与方数据隐私, 目前研究已经提出一些基于差分隐私的联邦学习隐私保护方法[17,18]. 如 Geyer 等人[19]提出了一种在联邦优化中对客户端进行差分隐私保护的算法. 该算法旨在隐藏单个客户端在训练过程中的贡献, 同时平衡隐私损失和模型效益. 每个参与者只将自己的本地模型更新发送给服务器, 而服务器仅负责将这些模型更新进行汇总和整合, 因此不会泄露任何参与者的隐私信息. 但是, 恶意服务器可以通过拦截客户端发送的模型更新来获取原始的模型更新, 从而暴露参与者的隐私信息. 为提高客户端本地数据的隐私安全, Hao 等人[20]提出在每轮客户端上传的局部梯度中加入噪声扰动, 以保护各参与方的隐私, 防止受到服务器或其他参与者的潜在威胁. 同时 Abadi 等人[21]提出了一种基于差分隐私的随机梯度下降算法, 该方法通过动量会计来计算隐私开销成本. 虽然聚合过程中本地参数满足中心化差分隐私需求, 但差分隐私在一定程度上会影响联邦模型的可用性. 因此, 采用差分隐私技术解决联邦学习的隐私安全问题时, 需要在效用性和隐私性之间进行权衡.

另一方面, 联邦学习参与训练的各个设备在采集数据时, 由于地理、时间和人文等因素, 造成数据非独立同分布 (non-independent identical distribution, non-IID)[22], 例如数据分布不均匀、标签漂移、特征漂移和特征偏好漂移等. 而联邦学习在非独立同分布数据下, 不仅难以在全局聚合过程中健壮地学习并概括每一个参与训练的设备, 而且会导致模型收敛性下降, 以致模型准确率降低[23,24]. 为解决该问题, Jeong 等人[25]提出了一种联邦增强方法, 该方法在 FL 服务器中训练生成对抗模型来生成额外的数据, 其中少数类的一些数据样本被上传到服务器以训练模型. 将训练好的生成对抗模型被分发给每个客户端, 以生成额外的数据并增加其本地数据, 从而得到一个独立同分布数据集. 然而, 这种方法可能会改变局部数据分布, 导致会丢失与客户行为多样性相关的有价值信息. 为充分利用联邦学习客户端数据集的特征, Fallah 等人[26]提出了 Per-FedAvg 算法, 通过建立一个初始元模型, 使在一个梯度下降步骤后能更有效地更新本地模型, 但在元优化[27]过程中, 该方法需要计算海森矩阵, 造成其计算开销过大. 此外 Li 等人[28]也提出一种 FedProx 算法, 该算法通过对局部子问题引入一个近端项, 每个参与方本地更新模型时, 都会在本地模型更新后添加一个正则化项, 该正则化项通过计算本地模型与全局模型的欧氏距离, 以此衡量模型之间的差异. 但在处理异常值和高维数据时, 欧氏距离会产生无法准确地反映数据之间相似性的问题.

综上所述, 现有研究主要存在以下不足: (1) 直接分享梯度会造成隐私泄露; (2) 采用差分隐私对联邦学习本地模型进行扰动时, 需要权衡隐私性和实用性; (3) Non-IID 数据会造成本地模型异构, 全局模型泛化误差也会随之增大. 为解决上述问题, 本文提出一种基于 Bregman 散度和差分隐私的个性化联邦学习 (FedBDP) 方法, 其主要贡献如下.

(1) 提出 FedBDP 方法为客户端提供个性化训练. 每个客户端对 non-IID 数据进行本地采样训练, 通过计算本地参数与全局参数的 Bregman 散度来计算近端项更新本地模型, 以提高模型准确率.

(2) 提出采用自适应差分隐私对本地模型梯度进行扰动. 通过高斯噪声的衰减系数动态调整噪声大小, 以制定合理的差分隐私需求实现对用户的隐私保护.

(3) 在 MNIST 和 CIFAR10 数据集上对 FedBDP 方法进行评估, 实验结果验证了该方法在满足差分隐私需求前提下保证了模型准确率.

2 相关定义

2.1 联邦学习

假设具有相同数据结构的 N 个客户端通过协作使用各自数据集 $\mathcal{D}_i$ 来训练机器学习模型, 其中 $i \in \{1,2,\ldots,N\}$. 在联邦学习模型训练过程中, 每一个客户端数据都不会离开该客户端, 即数据不离开数据拥有者. 服务器为学习在 N 个客户端数据上的模型 M_i, 它需要从 N 个客户端接收本地参数 w_i, 并对所有客户端本地模型参数聚合得到全局参数 $w=\sum_{i=1}^{N} p_i w_i$, 其中 $p_i=|\mathcal{D}_i|/|\mathcal{D}|$ 且 $\sum_{i=1}^{N} p_i=1$. 当客户端数据集大小相同时, $p_i=1/N$. 其 FL 优化问题可表示为:

$$w=\arg\min_{w\in\mathbb{R}^d}\sum_{i=1}^{N} p_i f_i(w,\mathcal{D}_i) \tag{1}$$

其中, $f_i(\cdot)$ 为第 i 个客户端的局部损失函数, 它主要由局部经验风险决定. 联邦学习主要是建立一个基于分布数据集的联邦学习模型, 当服务器与客户端经过一定的本地训练和更新后, w 能够收敛到全局期望最优解. 联邦学习的训练过程主要包括以下步骤.

(1) 本地训练: 所有客户端通过模型训练得到本地参数, 并将其发送到服务器.

(2) 模型聚合: 服务器在不学习本地数据的情况下, 对所有客户端上传的参数进行安全聚合.

(3) 参数广播: 服务器向所有客户端广播聚合后的全局参数.

(4) 模型更新: 所有客户端使用全局参数更新各自模型, 并测试更新后模型性能.

2.2 Bregman 散度

Bregman 散度[29]是一类用于度量两个概率分布之间差异的函数, 它已被广泛应用于机器学习、信息论、统计学等领域. 其定义如下.

定义 1 (Bregman 散度). 设 V 是一个凸集, $h: V\to R$ 是一个凸函数, $x,y\in V$ 是两个向量, Bregman 散度 $D_h(x\|y)$ 定义为:

$$D_h(x\|y)=h(x)-h(y)-\langle\nabla h(y),x-y\rangle \tag{2}$$

其中, $\langle\cdot,\cdot\rangle$ 表示内积, $\nabla h(y)$ 表示函数 h 在点 y 处的梯度. Bregman 散度的核心思想是采用函数 f 的泰勒展开式来近似描述向量 x 相对于向量 y 的差异, 由于 h 是凸函数, 因此可以保证 Bregman 散度是非负的. 它的优点是对于不同的凸函数都有定义, 并且可以通过梯度下降等方法进行优化. 同时, 它还可以通过选择不同的凸函数来表示不同的先验知识和偏好, 从而获得更好的优化结果.

2.3 差分隐私

差分隐私 (differential privacy, DP)[30]的思想是当敌手试图从数据库中查询个体信息时, 将个性信息混淆后使敌手从查询结果中不能辨别出个体级别的敏感性. 具体定义如下.

定义 2 (差分隐私). 假设仅有一个记录不同的两个数据集 $\mathcal{D}$ 和 $\mathcal{D}'$, 满足差分隐私的随机化机制 $M: X\to R$, 对于 $S\in Range(M)$ 存在:

$$\Pr(M(\mathcal{D}_i)\in S)\leqslant beta^{\varepsilon}\Pr(M(\mathcal{D}_i')\in S)+\delta \tag{3}$$

其中, ε 为隐私预算、δ 表示失败概率. 当 $\delta=0$ 时, 可得到严格的 ε-差分隐私. 差分隐私满足两个算法组合特性, 分别是序列组合性[31]和并行性[32].

定义 3 (组合性). 对于数据集 $\mathcal{D}$ 和 n 个随机优化算法 $\{M_i\}$, $i\in\{1,\ldots,n\}$. 如果 $M_i(\mathcal{D})$ 满足 ε_i 差分隐私, 则 $\{M_i\}$ 在 $\mathcal{D}$ 上的顺序序列组合满足 $\sum^{n}\varepsilon_i$ 差分隐私. 这表明多个算法同时作用在同一数据集上时, 总体隐私预算是各隐私预算之和.

定义 4 (并行性). 将数据集 $\mathcal{D}$ 分成 n 个互不相交的集合 $\{\mathcal{D}_i\}$ 且 $i\in\{1,\ldots,n\}$, 每个集合分别作用于一个随机算法 $\{M_i\}$. 如果 M_i 满足 ε_i 差分隐私, 则 $\{M_i\}$ 在 $\mathcal{D}$ 上的并行序列组合满足 $\max(\varepsilon_i)$ 差分隐私. 这表明多个算法作用于一个数据集的不相交子集时, 总体隐私预算是各隐私预算中的最大值.

定义 5 (敏感度). 假设仅有一个记录不同的两个数据集 $\mathcal{D}$ 和 $\mathcal{D}'$, 对于任意域函数 $s: D\to\mathbb{R}^d$, s 的敏感度为 s

在接受所有可能输入后得到的输出最大变化值, 则 s 的敏感度定义为:

$$\Delta s = \max_{\mathcal{D},\mathcal{D}'} \|s^{\mathcal{D}} - s^{\mathcal{D}'}\|_p \tag{4}$$

其中, $\|\cdot\|$ 表示向量的范数. 敏感度反映了函数 s 在一堆相邻数据集上输出结果的最大变化范围, 其敏感度越小, 实现差分隐私时需要向输出结果添加的噪声就越小.

定义 6 (高斯机制). 对于给定 $M: X \to \mathbb{R}^d$, 存在高斯机制 $M_G(\mathcal{D}, q, \varepsilon, \delta) = M_q(\mathcal{D}) + N(0, \sigma^2 I_d)$, 其中 N 满足高斯分布, 且高斯噪声需满足 $\sigma \geqslant c\Delta s/\varepsilon$, 其中 $c = \sqrt{2\ln(1.25/\delta)}$.

通过差分隐私对模型参数添加足够的噪声, 可以为联邦学习提供强有力的隐私保护, 但也面临一些新的挑战, 如 Jayaraman 等人[33]发现用于机器学习的差分隐私机制难以权衡实用性和隐私性, 即较低的模型准确率损失会导致较弱的隐私保护, 而较强隐私保护会导致更高的模型准确率损失.

3 FedBDP 模型与具体实现

3.1 FedBDP 模型框架

针对联邦学习中存在的隐私安全和 non-IID 数据下模型异构问题, 本文提出的 FedBDP 模型框架如图 1 所示. 该模型主要包括客户端和服务器两个实体, 客户端主要是具有一定处理能力的智能移动终端设备, 服务器主要是为客户端提供计算和应用服务的设备. 与一般的联邦学习不同, FedBDP 采用 Bregman 散度对客户端数据进行采样训练, 并添加自适应差分隐私来加强用户隐私保护, 其主要工作过程如下.

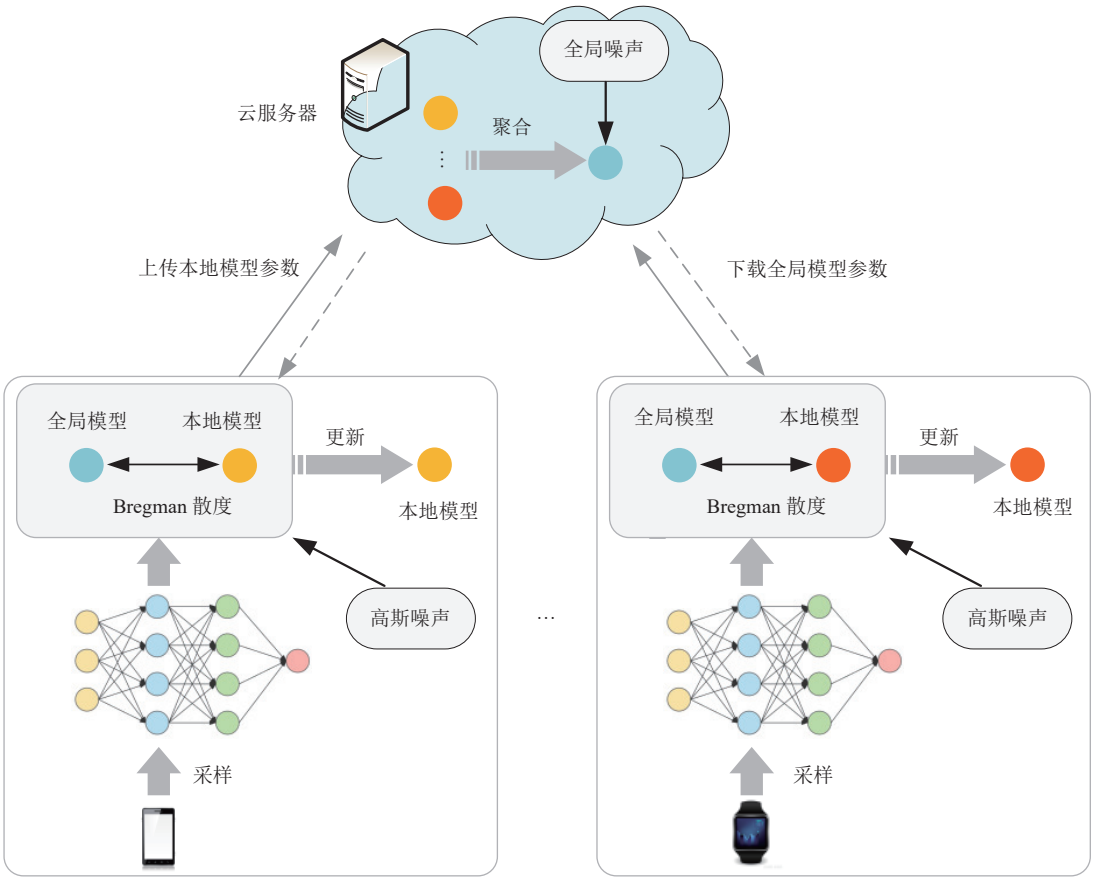


图 1 FedBDP 模型框架

(1) 客户端的本地模型对数据采样并进行个性化训练, 通过 Bregman 散度衡量本地模型与全局模型的差异, 将其作为惩罚项更新损失函数. 然后采用自适应差分隐私对本地模型参数进行扰动再将其上传至服务器.

(2) 服务器接收所有客户端模型参数, 执行聚合操作得到全局参数, 并向所有客户端下发全局参数.

(3) 客户端接收全局参数进行更新, 然后重复上述步骤直到模型收敛.

3.2 模型威胁

在 FedBDP 的威胁模型中, 本文考虑由 N 个客户端共同训练联邦模型的情景, 参与的客户端容易受到潜在攻击者的控制, 将其称之为内部攻击者. 该攻击者被认为是一个诚实而好奇的参与者, 他试图提取有关其他客户端数据的敏感信息[34,35]. 具体而言, 本文对内部攻击者做出以下假设.

(1) 攻击者只能观察到其他客户端上传到服务器的本地模型.

(2) 攻击者不会干扰协作学习以获得更有益的参数, 这意味着攻击者是半诚实的, 会按规则训练和更新模型.

(3) 攻击者可以分析先前的知识或生成模型来推断其他参与者的信息, 然而攻击者只能推断数据信息, 可能不知道实际攻击目标.

(4) 攻击者可能利用其他客户端上传的模型信息来反向推断其本地数据的敏感信息, 例如推断出特定用户的偏好或行为模式.

本文中主要符号定义及含义如表 1 所示.

表 1 本文中主要符号定义

参数	含义	参数	含义
$\mathcal{D}_i$	客户端 i 的本地数据集	f_i	客户端 i 的损失函数
w_i	客户端 i 本地模型参数	$D_h(w_i, w)$	通过 h 函数计算本地模型与全局模型的Bregman散度
N	客户端总数	T	全局迭代次数
C	梯度裁剪阈值	R	客户端本地模型迭代次数
w_t	第 t 轮迭代时的全局参数	κ	衰减速率
η	学习率	β	差分隐私噪声的衰减系数

3.3 FedBDP 具体实现

FedBDP 算法主要分为 Bregman 优化和自适应差分隐私两个部分, 其具体实现过程分别描述如下.

3.3.1 Bregman 优化

因联邦学习在 non-IID 数据下存在本地模型异构, 导致全局模型无法为各客户端的任务提供较好性能. 为解决该问题, FedBDP 重新定义联邦学习优化问题, 并个性化训练本地模型以提升模型性能.

假设 N 个客户端与服务器通信需要通过公式 (1) 求解找到一个全局模型 w, 其中函数 $f_i: \mathbb{R}^d \to \mathbb{R}$ 表示客户端 i 数据分布的经验损失:

$$f_i(w) = \mathbb{E}[f_i(w; \mathcal{D}_i)] \tag{5}$$

其中, w 表示全局参数, $f_i(w; \mathcal{D}_i)$ 表示损失函数. 客户端的数据 $\mathcal{D}_i$ 可能来自不同的环境、上下文和应用程序, 导致其分布是 non-IID. 与 FedProx 算法不同, 本文引入 Bregman 散度作为近端项解决 FL 优化问题, Bregman 散度相比欧氏距离更加灵活和敏感, 可以更好地区分不同的向量. 因此, 本文将 Bregman 散度作为正则化项更新各客户端的损失函数:

$$f_i(w_i) + \lambda D_h(w_i, w) \tag{6}$$

其中, w_i 表示本地模型, λ 表示控制 w_i 与 w 之间差异的正则化参数, 它可以根据模型的收敛情况进行调整, 以平衡全局模型参数的影响和客户端本地数据的重要性. 因此可将 FL 优化问题转化为:

$$\min_{w \in d} \left\{ F(w) = \frac{1}{N} \sum_{i=1}^{N} F_i(w) \right\} \tag{7}$$

$$F_i(w) = \min_{w_i \in \mathbb{R}^d} \{ f_i(w_i) + \lambda D_h(w_i, w) \} \tag{8}$$

通过公式 (7) 可以根据每个客户端与全局模型的差异进行个性化训练, 以此进行模型更新. 接下来需要对公式 (7) 进行最小化求解, 首先通过损失函数采样训练数据 $\mathcal{D}_i$:

$$\mathbb{E}[\nabla f_i(w_i, \mathcal{D}_i)] = \nabla f_i(w_i) \tag{9}$$

然后计算 Bregman 散度, 考虑玻尔兹曼香农熵[36] $h(x) = \sum_{j=1}^{J} x_j \ln(x_j) - x_j$, $x = (x_j)_{1 \leqslant j \leqslant J}$, 且 $y = (y_j)_{1 \leqslant j \leqslant J}$, 因而与 h 相关的 Bregman 距离可以定义为:

$$D_h(x, y) = \sum_{j=1}^{J} x_j \ln\left(\frac{x_j}{y_j}\right) - x_j + y_j, \text{如果} x \geqslant 0,\ y > 0 \tag{10}$$

因此, 本地模型和全局模型之间的 Bregman 散度可表示为:

$$D_h(w_i, w) = \sum_{j=1}^{d} \left(w_i^j \ln\left(\frac{w_i^j}{w_t^j}\right) - w_i^j + w_t^j \right) \tag{11}$$

其中, w_i 和 w_t 都是 d 维向量. 但本地模型和全局模型无法直接求得 Bregman 散度, 因为它们包含负数. 本文采用 Softmax 函数使其满足概率分布, 再求取本地模型与全局模型的差异. 通过 Bregman 散度改进损失函数, 它可以减少 FL 客户模型训练过程中模型异质性引起的收敛性下降的影响.

最后, 本文可以通过优化上述带正则化项的本地损失函数来更新客户端模型参数:

$$g_i(w_{i,r}^t, w_t) = f_i(w_{i,r}^t) + \lambda D_h(w_{i,r}^t, w_t) \tag{12}$$

$$w_{i,r+1}^t = w_{i,r}^t - \eta \nabla g_i \tag{13}$$

Bregman 优化过程具体如图 2 所示, 在 non-IID 数据下, 联邦学习的全局最优解与各个局部最优解之间的距离并不相等. 该情况下, 平均后的模型将会偏离全局最优解, 导致全局模型无法收敛到其真正的全局最优解. 通过 Bregman 散度对全局模型和局部模型的差异进行度量, 可以促使全局模型重新收敛到全局最优解.

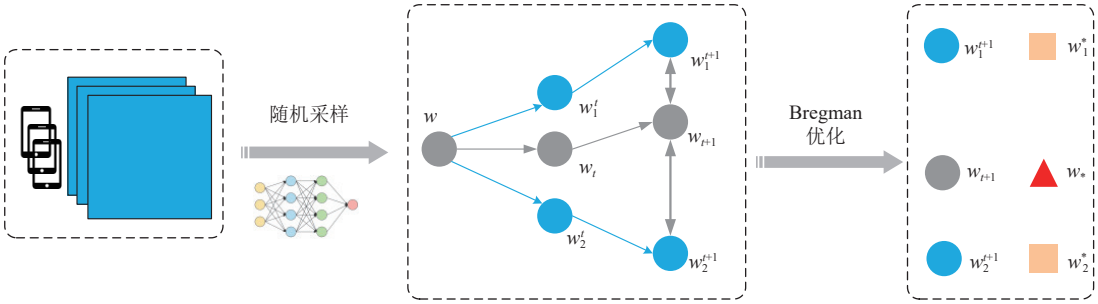


图 2 Bregman 优化过程

3.3.2 自适应差分隐私

为防止本地模型参数上传时造成的用户隐私泄露, FedBDP 方法引入高斯机制对每个客户端更新后的模型参数进行扰动, 并在每次迭代中添加高斯噪声来协同学习具有差分隐私保护的联邦模型.

在差分隐私梯度下降算法中, 通常采用裁剪技术确保 $\| w_i \| \leqslant C$, 其中 w_i 表示来自第 i 个客户端的模型参数, C 表示 w_i 边界的裁剪阈值. 因此客户端模型梯度裁剪可以表示为:

$$\bar{g}_i = \nabla g_i / \max\left(1, \frac{\|\nabla g_i\|_2}{C}\right) \tag{14}$$

假设本地训练中的采用批处理大小 B 对样本进行采样训练, 同时可以根据梯度裁剪阈值得到模型梯度的敏感度 C . 因此, 对本地模型 i 添加高斯噪声可以表示为:

$$\tilde{g}_i(w) = \frac{1}{B} \sum_{j} g_i(w; x_j) + \mathcal{N}(0, \beta\sigma^2 C^2 \mathbf{I}) \tag{15}$$

由于迭代次数的增加, 会造成过多的隐私成本累积. 随着训练轮数的增加, 模型逐渐趋于收敛, 各个客户端之间的模型参数差异也会变得越来越小. 在这种情况下, 攻击者将很难从模型参数中提取某个特定客户端的敏感信息. 对此, 本文设计了一个动态噪声系数, 在模型初始阶段添加较大的噪声, 而在模型趋于收敛时减少噪声扰动. 指数衰减函数的动态噪声系数可定义为:

$$\beta=\mathrm{e}^{-\kappa\frac{tr}{TR}} \tag{16}$$

其中, κ 表示衰减速率且 $\kappa\in(0,1)$, r 和 t 分别表示当前的本地和全局迭代轮数, R 和 T 分别为总的本地和全局迭代次数. 该公式利用指数衰减函数实现在训练初期添加较大噪声, 而在训练后期减少噪声的目的. 因此, 在保证模型收敛性的同时, FedBDP 方法还能够保护用户数据隐私, 为联邦学习场景中的数据安全提供有力保障.

通过添加自适应差分隐私, 客户端 i 的本地模型更新定义为:

$$w_{i,r+1}^{t}=w_{i,r}^{t}-\eta\tilde{g}_i \tag{17}$$

本地模型完成更新后, 需要将本地模型上传至服务器进行更新. 根据差分隐私并行性的特征, 在上传服务器过程中所有本地模型满足 (ε,δ)-差分隐私. 因此, 通过局部差分隐私对本地模型进行扰动, 恶意攻击者将难以从其上传的模型参数 w_i 中推断出客户端 i 上的用户敏感信息.

在 FedBDP 中, 各客户端 i 在 non-IID 数据下进行采样训练, 采用 Bregman 散度量化本地模型与全局模型的差异, 并更新模型参数 w_i. 通过 Bregman 优化减少了 non-IID 数据下模型异构造成性能降低的负面影响, 缩小本地模型之间的差异, 加速全局收敛并提高模型准确率. 同时对本地参数 w_i 进行裁剪, 并添加自适应差分隐私噪声对其进行扰动后发送至服务器进行全局聚合, 有效防止恶意攻击者分析模型参数以获取用户敏感信息. 其具体算法如算法 1 所示.

算法 1. FedBDP 算法.

输入: 正则化参数: λ, 学习率: η, 噪声系数: β, 梯度裁剪值 C;
输出: w_t.

1) 初始化模型参数 w_0;
2) for $t=0,\ldots,T-1$ do 全局迭代
3) $\quad w_{i,0}^{t}=w_t$
4) $\quad$ for $r=0,\ldots,R-1$ do 本地迭代
5) $\quad\quad$ Bregman 优化 $\nabla g_i(w_i^t,w)=\arg\min_{w\in d} f_i(w_i^t)+\lambda D_h(w_i^t,w)$
6) $\quad\quad$ 梯度裁剪 $\bar{g}_i=\nabla g_i/\max\left(1,\frac{\|\nabla g_i\|_2}{C}\right)$
7) $\quad\quad$ 添加自适应差分隐私 $\tilde{g}_i(w)=\frac{1}{B}\sum_j g_i(w;x_j)+\mathcal{N}\left(0,\beta\sigma^2C^2\mathbf{I}\right)$
8) $\quad\quad$ 梯度更新 $w_{i,r+1}^{t}=w_{i,r}^{t}-\eta\tilde{g}$
9) $\quad$ end for
10) $\quad$ 服务器接受 S 个客户端上传的本地模型 $w_{i,R}^{t}$ 并进行聚合 $w_t=\frac{1}{S}\sum_{i\in S} w_{i,R}^{t}$
11) end for
12) return w_t

4 理论分析

本节对 FedBDP 方法在非独立同分布数据下能否满足收敛条件以及高斯机制的差分隐私需求进行了理论分析, 分析的目的在于探究 FedBDP 在实际应用中的可行性. 对于收敛性和隐私性两个方面, 本节进行了详细的讨论和分析.

4.1 收敛性分析

在个性化训练中, 本文方法采用 Bregman 散度计算近端点来加速收敛. 为便于分析 FedBDP 的收敛性, 假设存

在以下条件.

(1) f_i 为 μ-强凸函数, 则对 $\forall w, w'$ 存在 $f_i(w) \geqslant f_i(w') + \nabla f_i(w')(w - w') + \frac{\mu}{2} \| w - w' \|^2$.

(2) f_i 为非凸 L-光滑函数, 则对 $\forall w, w'$ 存在 $\| \nabla f_i(w) - \nabla f_i(w') \| \leqslant L \| w - w' \|$.

通过以上假设获得 FedBDP 在强凸和非凸光滑函数下的收敛性, 如定理 1 和定理 2 所示.

定理 1 (非凸收敛性). 当 f_i 为非凸 L-光滑函数时, FedBDP 算法收敛性满足:

$$\frac{1}{T}\sum_{t=0}^{T-1}\|\nabla F(w_t)\|^2 \leqslant 2\sqrt{\frac{(F(w_0)-F(w_T))L}{S}} + \frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\left\|\nabla g_i\left(w_{i,r}^t, w_t\right) - \nabla F(w_t)\right\|^2 \tag{18}$$

证明: 根据非凸 L-光滑函数的性质, 可得:

$$f_i\left(w_{i,r+1}^t\right) - f_i\left(w_{i,r}^t\right) \leqslant \nabla f_i\left(w_{i,r}^t\right)^{\top}\left(w_{i,r+1}^t - w_{i,r}^t\right) + \frac{L}{2}\left\|w_{i,r+1}^t - w_{i,r}^t\right\|^2 \tag{19}$$

将公式 (19) 代入局部模型更新公式:

$$w_{i,r+1}^t = w_{i,r}^t - \eta \nabla g_i\left(w_{i,r}^t, w_t\right) \tag{20}$$

计算可得:

$$f_i\left(w_{i,r+1}^t\right) - f_i\left(w_{i,r}^t\right) \leqslant -\eta \nabla f_i\left(w_{i,r}^t\right)^{\top} \nabla g_i\left(w_{i,r}^t, w_t\right) + \frac{L\eta^2}{2}\left\|\nabla g_i\left(w_{i,r}^t, w_t\right)\right\|^2 \tag{21}$$

累加关于 r 的不等式, 从 0 到 R−1, 然后取平均:

$$\frac{1}{R}\sum_{r=0}^{R-1}\left[f_i\left(w_{i,r+1}^t\right) - f_i\left(w_{i,r}^t\right)\right] \leqslant -\frac{\eta}{R}\sum_{r=0}^{R-1}\nabla f_i\left(w_{i,r}^t\right)^{\top}\nabla g_i\left(w_{i,r}^t, w_t\right) + \frac{L\eta^2}{2R}\sum_{r=0}^{R-1}\left\|\nabla g_i\left(w_{i,r}^t, w_t\right)\right\|^2 \tag{22}$$

两边同时乘以 $\frac{S}{N}$ 并对所有客户端求和:

$$\frac{S}{N}\sum_{i=1}^{N}\frac{1}{R}\sum_{r=0}^{R-1}\left[f_i\left(w_{i,r+1}^t\right) - f_i\left(w_{i,r}^t\right)\right] \leqslant -\frac{S\eta}{N}\sum_{i=1}^{N}\frac{1}{R}\sum_{r=0}^{R-1}\nabla f_i\left(w_{i,r}^t\right)^{\top}\nabla g_i\left(w_{i,r}^t, w_t\right) + \frac{SL\eta^2}{2N}\sum_{i=1}^{N}\frac{1}{R}\sum_{r=0}^{R-1}\left\|\nabla g_i\left(w_{i,r}^t, w_t\right)\right\|^2 \tag{23}$$

结合全局模型更新公式, 并利用性质 (2) 以及 Jensen 不等式, 可以证明:

$$F(w_{t+1}) - F(w_t) \leqslant -\frac{\eta}{2}\frac{S}{N}\|\nabla F(w_t)\|^2 + \frac{L\eta^2}{2}\frac{S}{N}\|\nabla F(w_t)\|^2 + \frac{\eta}{2}\frac{S}{N}\sum_{i=1}^{N}\left\|\nabla g_i\left(w_{i,r}^t, w_t\right) - \nabla F(w_t)\right\|^2 \tag{24}$$

根据公式 (24), 可以发现在每次全局模型更新后, 优化目标值会降低. 为求解收敛上界, 对公式 (24) 进行 T 次累加:

$$\sum_{t=0}^{T-1}[F(w_{t+1}) - F(w_t)] \leqslant -\frac{\eta}{2}\frac{S}{N}\sum_{t=0}^{T-1}\|\nabla F(w_t)\|^2 + \frac{L\eta^2}{2}\frac{S}{N}\sum_{t=0}^{T-1}\|\nabla F(w_t)\|^2 + \frac{\eta}{2}\frac{S}{N}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\left\|\nabla g_i\left(w_{i,r}^t, w_t\right) - \nabla F(w_t)\right\|^2 \tag{25}$$

根据梯度下降的性质, 可得:

$$F(w_0) - F(w_T) \leqslant \frac{\eta}{2}\frac{S}{N}\sum_{t=0}^{T-1}\|\nabla F(w_t)\|^2 + \frac{L\eta^2}{2}\frac{S}{N}\sum_{t=0}^{T-1}\|\nabla F(w_t)\|^2 + \frac{\eta}{2}\frac{S}{N}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\left\|\nabla g_i\left(w_{i,r}^t, w_t\right) - \nabla F(w_t)\right\|^2 \tag{26}$$

为使公式 (26) 左边非负, 需要满足以下条件:

$$\frac{\eta}{2}\frac{S}{N}\sum_{t=0}^{T-1}\|\nabla F(w_t)\|^2 + \frac{L\eta^2}{2}\frac{S}{N}\sum_{t=0}^{T-1}\|\nabla F(w_t)\|^2 + \frac{\eta}{2}\frac{S}{N}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\left\|\nabla g_i\left(w_{i,r}^t, w_t\right) - \nabla F(w_t)\right\|^2 \geqslant F(w_0) - F(w_T) \tag{27}$$

通过调整学习率 η, 关注梯度的平方范数项, 可以得到:

$$\frac{1}{T}\sum_{t=0}^{T-1}\|\nabla F(w_t)\|^2 \leqslant \frac{2(F(w_0) - F(w_T))}{\eta S} + \frac{2L\eta}{S}\sum_{t=0}^{T-1}\|\nabla F(w_t)\|^2 + \frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\left\|\nabla g_i\left(w_{i,r}^t, w_t\right) - \nabla F(w_t)\right\|^2 \tag{28}$$

定理 2 (强凸收敛性). 当 f_i 为 μ-强凸函数, FedBDP 算法收敛性满足:

$$F(w_T) - F(w^*) \leqslant F(w_0) - F(w^*) - \frac{\eta}{2}\frac{S}{N}\sum_{t=0}^{T-1}\|\nabla F(w_t)\|^2 + \frac{L\eta^2}{2}\frac{S}{N}\sum_{t=0}^{T-1}\mathbb{E}\left[\left\|w_{i,r}^t - w_{i,r+1}^t\right\|^2\right] \tag{29}$$

证明: 对于每个客户端 i, 根据局部模型的更新不等式和强凸性质, 可以得到:

$$f_i\left(w_{i,r+1}^t\right) \leqslant f_i\left(w_{i,r}^t\right) - \eta\left\langle \nabla f_i\left(w_{i,r}^t\right), w_{i,r}^t - w_{i,r+1}^t \right\rangle + \frac{L\eta^2}{2}\left\| w_{i,r}^t - w_{i,r+1}^t \right\|^2 - \frac{\mu\eta^2}{2}\left\| w_{i,r}^t - w_{i,r+1}^t \right\|^2 \tag{30}$$

对于每一轮迭代 t, 随机选择 S 个客户端进行聚合. 对这些局部模型更新不等式求期望, 然后乘以 $\frac{S}{N}$. 同时需要分析期望:

$$\mathbb{E}\left[\left\langle \nabla f_i\left(w_{i,r}^t\right), w_{i,r}^t - w_{i,r+1}^t \right\rangle\right] \tag{31}$$

为计算公式 (31), 将 $\nabla f_i(w_{i,r}^t)$ 替换为 $\nabla F(w_t)$, 并利用了随机选择 S 个客户端的机制. 因此, 该期望可以表示为:

$$\mathbb{E}\left[\left\langle \nabla F(w_t), w_{i,r}^t - w_{i,r+1}^t \right\rangle\right] = \left\langle \nabla F(w_t), \mathbb{E}\left[w_{i,r}^t - w_{i,r+1}^t\right] \right\rangle \tag{32}$$

将公式 (32) 代入公式 (30) 得到:

$$\frac{S}{N}\mathbb{E}\left[f_i\left(w_{i,r+1}^t\right)\right] \leqslant \frac{S}{N}\mathbb{E}\left[f_i\left(w_{i,r}^t\right)\right] - \eta\frac{S}{N}\left\langle \nabla F(w_t), \mathbb{E}\left[w_{i,r}^t - w_{i,r+1}^t\right] \right\rangle + \frac{L\eta^2}{2}\frac{S}{N}\mathbb{E}\left[\left\| w_{i,r}^t - w_{i,r+1}^t \right\|^2\right] - \frac{\mu\eta^2}{2}\frac{S}{N}\mathbb{E}\left[\left\| w_{i,r}^t - w_{i,r+1}^t \right\|^2\right] \tag{33}$$

为了进一步分析收敛性, 需要考虑在整个训练过程中的所有迭代轮次. 首先, 对每个迭代 t 的局部更新不等式求和, 并将其累加到全局模型更新中. 然后, 计算可得:

$$\sum_{t=0}^{T-1}\left[F(w_{t+1}) - F(w_t)\right] \leqslant -\frac{\eta}{2}\frac{S}{N}\sum_{t=0}^{T-1}\left\|\nabla F(w_t)\right\|^2 + \frac{L\eta^2}{2}\frac{S}{N}\sum_{t=0}^{T-1}\mathbb{E}\left[\left\| w_{i,r}^t - w_{i,r+1}^t \right\|^2\right] - \frac{\mu\eta^2}{2}\frac{S}{N}\sum_{t=0}^{T-1}\mathbb{E}\left[\left\| w_{i,r}^t - w_{i,r+1}^t \right\|^2\right] \tag{34}$$

对上式两边求和并重新排列, 得到:

$$F(w_T) - F(w_0) \leqslant -\frac{\eta}{2}\frac{S}{N}\sum_{t=0}^{T-1}\left\|\nabla F(w_t)\right\|^2 + \frac{L\eta^2}{2}\frac{S}{N}\sum_{t=0}^{T-1}\mathbb{E}\left[\left\| w_{i,r}^t - w_{i,r+1}^t \right\|^2\right] - \frac{\mu\eta^2}{2}\frac{S}{N}\sum_{t=0}^{T-1}\mathbb{E}\left[\left\| w_{i,r}^t - w_{i,r+1}^t \right\|^2\right] \tag{35}$$

因 $\frac{\mu\eta^2}{2}\frac{S}{N}\sum_{t=0}^{T-1}\mathbb{E}\left[\left\| w_{i,r}^t - w_{i,r+1}^t \right\|^2\right]$ 项是正数, 可进一步计算得:

$$F(w_T) - F(w^*) \leqslant F(w_0) - F(w^*) - \frac{\eta}{2}\frac{S}{N}\sum_{t=0}^{T-1}\left\|\nabla F(w_t)\right\|^2 + \frac{L\eta^2}{2}\frac{S}{N}\sum_{t=0}^{T-1}\mathbb{E}\left[\left\| w_{i,r}^t - w_{i,r+1}^t \right\|^2\right] \tag{36}$$

4.2 隐私性分析

因差分隐私能有效保护数据隐私且具有全局敏感特征, FedBDP 方法利用差分隐私中的高斯机制对联邦学习模型进行噪声扰动, 且选择松弛的高斯机制 (ε, δ)-差分隐私.

对于两个相邻数据集 $\mathcal{D}$ 和 $\mathcal{D}'$ 、输出 o 和随机函数 M, 该随机函数造成的隐私损失 $c_M(o, \mathcal{D}, \mathcal{D}')$ 可定义为:

$$c_M(o, \mathcal{D}, \mathcal{D}') = \ln\frac{\Pr[M(\mathcal{D}) = o]}{\Pr[M(\mathcal{D}') = o]} \tag{37}$$

随机函数 M 满足 (ε, δ)-差分隐私的充分条件为其隐私损失 $c_M(o, \mathcal{D}, \mathcal{D}')$ 满足 $\Pr[c_M(o, \mathcal{D}, \mathcal{D}') > \varepsilon] \leqslant \delta$. 那么在高斯机制中隐私损失可定义为:

$$\ln\frac{\mathrm{e}^{(-1/2\sigma^2)x^2}}{\mathrm{e}^{(-1/2\sigma^2)(x+\Delta s)^2}} \tag{38}$$

又由于概率恒正, 因此可得:

$$\left|\ln\frac{\mathrm{e}^{(-1/2\sigma^2)x^2}}{\mathrm{e}^{(-1/2\sigma^2)(x+\Delta s)^2}}\right| = \left|\frac{1}{2\sigma^2}\left(2x\Delta s + (\Delta s)^2\right)\right| \tag{39}$$

其中, $\left|\frac{1}{2\sigma^2}\left(2x\Delta s + (\Delta s)^2\right)\right| < \varepsilon$, $x < \sigma^2\varepsilon/\Delta s - \Delta s/2$. 为确保隐私损失以 ε 为界, 且概率至少为 $1-\delta$, 则存在:

$$\left[\Pr|x| \geqslant \sigma^2\varepsilon/\Delta s - \Delta s/2 < \delta\right] \Leftrightarrow \Pr\left[x \geqslant \sigma^2\varepsilon/\Delta s - \Delta s/2\right] < \delta/2 \tag{40}$$

设 $t = \sigma^2\varepsilon/\Delta f - \Delta f/2$, 可将公式 (40) 转化为 $\Pr[x > t] \leqslant \frac{\sigma}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2\sigma^2}$. 又由于:

$$\frac{\sigma}{\sqrt{2\pi}}\frac{1}{t}\mathrm{e}^{t^2/2\sigma^2} < \delta/2 \Leftrightarrow \frac{t}{\sigma}\mathrm{e}^{t^2/2\sigma^2} > 2/\sqrt{2\pi}\delta \Leftrightarrow \ln\left(\frac{t}{\sigma}\right) + \frac{t^2}{2\sigma^2} > \ln\left(\frac{2}{\sqrt{2\pi}\delta}\right) \tag{41}$$

根据文献 [30], 计算可得 $c \geqslant \sqrt{2\ln(1.25/\delta)}$. 因此, 当 $c = \sqrt{2\ln(1.25/\delta)}$ 时 FedBDP 方法能够满足差分隐私需求, 同时根据差分隐私的组合性可知 FedBDP 满足全局 (ε,δ) -DP.

5 实 验

为验证 FedBDP 方法的有效性, 本节主要通过实验验证在相关参数变化时对 FedBDP 模型准确率的影响, 并将 FedBDP 方法与 FedProx[28]、FedAvg[9]和 DP-FedAvg[21]方法进行实验对比.

实验硬件配置为 Intel(R) Core(TM) i7-10750H CPU, GTX 2060 GPU, 16 GB RAM, 操作系统为 Windows 10, 并使用 PyTorch 1.7.0 深度学习框架训练机器学习模型和差分隐私噪声添加. 实验数据集采用目前被联邦学习算法广泛使用的 MNIST 和 CIFAR10 数据集. 由于数据大小的限制, 本实验将完整的数据集分发到 100 个客户端. 此外, 本文利用狄利克雷分布策略, 实现对各客户端随机分配不同大小的本地数据集和标签数量的目标, 确保每个客户端数据具有较强的数据统计异质性, 从而满足非独立同分布 (non-IID) 的条件, 为建立模型异构环境提供基础. 狄利克雷分布方法能够产生多样化的数据子集, 其中每个子集都具有独特的特征, 进而为每个客户端的模型性能带来不同程度挑战. 通过这种方式可以模拟现实世界中数据分布不均匀的情况, 本文将狄利克雷分布值设置为 0.1, 为模型训练提供更为复杂的条件.

对于 MNIST 数据, 本文采用具有 Softmax 激活函数和交叉熵损失函数的 l_2 -正则化多项逻辑回归模型进行训练, 并将梯度裁剪阈值设置为 0.2. 对于 CIFAR10 数据, 采用两层深度神经网络和交叉熵损失函数, 其隐藏层大小为 100, 并对不同层分别使用 ReLU 和 Softmax 函数激活, 并将梯度裁剪阈值设置为 0.1. 在模型训练过程中, 所有数据集被随机分割, 其中有 75% 用于训练集和 25% 用于测试集. 同时为确保隐私保护敏感度取得更低值, 实验将参数 δ 、 λ 、B 和 η 分别设置为 0.01、0.1、10 和 0.005.

5.1 相关参数对 FedBDP 模型的影响

(1) 客户端选择数量 S

本文通过研究不同 S 值对 FedBDP 模型测试准确率的影响, 在 MNIST 和 CIFAR10 数据集下对 FedBDP 进行实验评估. 在实验过程中, 选择 4 种不同的客户端选择数量 S, 分别为 10、30、50 和 100. 从图 3(a) 和图 3(b) 中可以看出, 随着客户端数量的逐渐增加, 模型的准确率也相应提高. 然而, 这种提高并非无代价, 联邦学习过程中的整体开销也会随着客户端数量的增加而变大. 为在准确率和开销之间寻求平衡, 本文在后续实验中选择 S=10 个客户端的数量作为模型训练的基准, 以在保证一定程度准确率的同时, 减少联邦学习过程中的计算和通信开销.

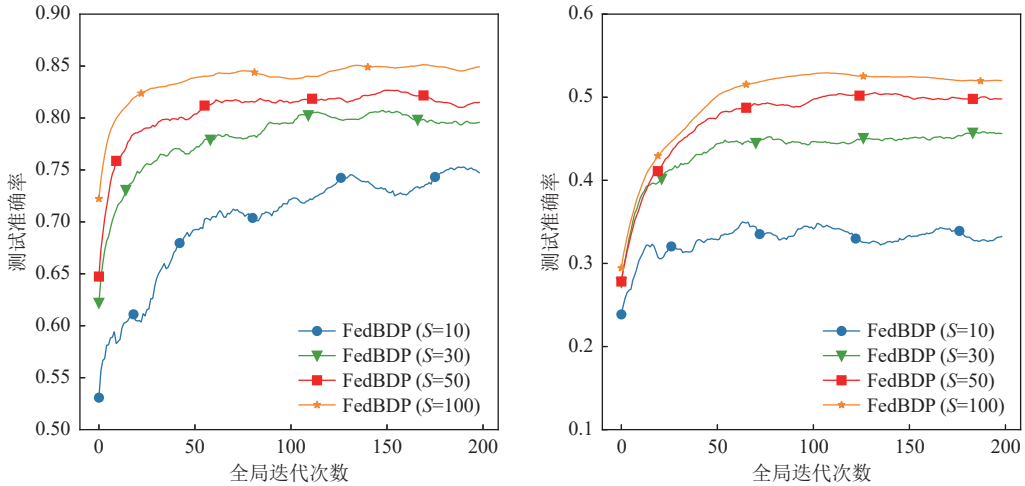


图 3 S 对 FedBDP 模型准确率的影响

(2) 衰减速率 κ

在 FedBDP 方法中, 通过定义一个衰减函数来控制每轮本地迭代所需添加的差分隐私噪声大小. 在实验过程中, 本文选择 4 种不同的衰减速率 κ, 分别为 0.1、0.3、0.6 和 0.9. 从图 4(a) 和图 4(b) 可知, 随着衰减速率的增加, 模型的测试准确率也随之提高. 这主要是因为衰减函数使得差分隐私噪声的大小逐渐降低, 从而提高了模型的可用性. 在迭代过程中这一特性也有助于加速模型的收敛速度, 以提高训练效率.

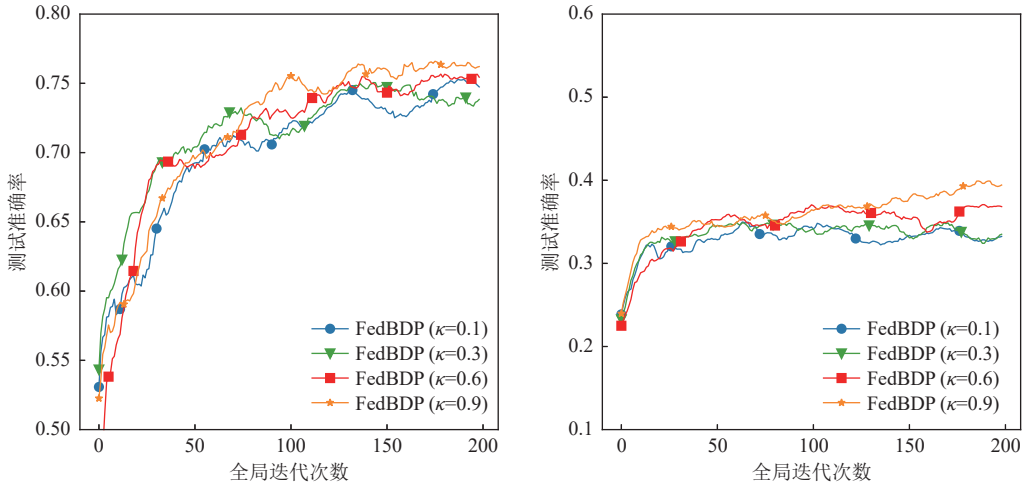


图 4 κ 对 FedBDP 模型准确率的影响

(3) 隐私预算 ε

差分隐私应用于机器学习时, 必须考虑隐私预算与模型准确率之间的均衡性. 为使模型准确率和隐私保护取得最佳水平, 在 MNIST 数据集下, 对 ε 为 3、5、10 和 15 进行实验. 在 CIFAR10 数据集下, 对 ε 为 10、15、20 和 30 进行实验. 由图 5 可知, 随着隐私预算 ε 值变大, FedBDP 中模型准确率也随之增加, 这正好与差分隐私中隐私预算和模型准确率成反比的性质相吻合, 即隐私预算越小, 隐私保护越强, 但模型准确率越低; 隐私预算越大, 模型准确率得到提高, 但隐私保护变弱.

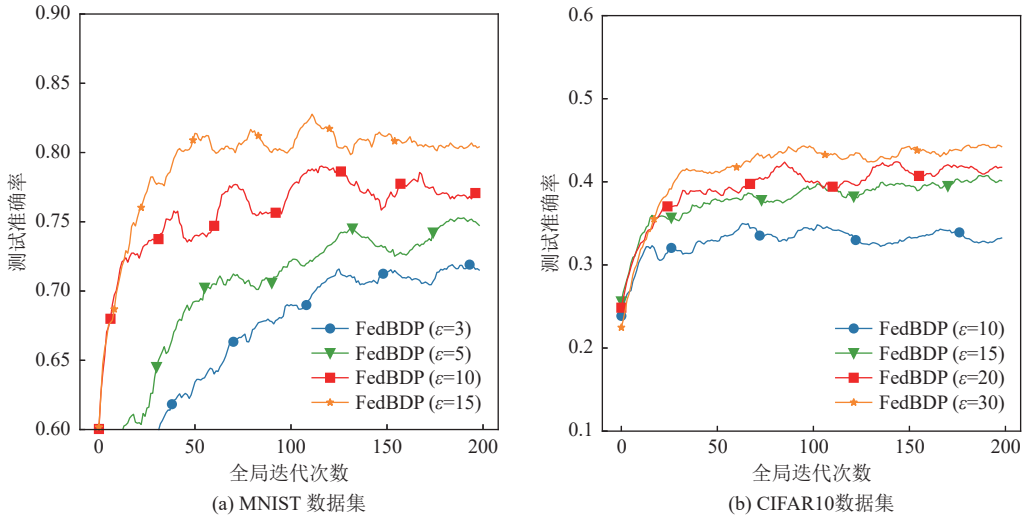


图 5 ε 对 FedBDP 模型准确率的影响

5.2 实验对比

为了验证在 non-IID 数据下, Bregman 优化和自适应差分隐私对联邦学习的有效性, 本文将 FedBDP 与 FedAvg、FedProx 进行实验对比. 同时, 在隐私预算相同的情况下与 DP-FedAvg 算法进行实验比较.

由图 6(a) 可知, 在 MNIST 数据集下, FedProx 和 FedAvg 的模型测试准确率分别达到了 87.92% 和 81.20%. 当 $\varepsilon = 5$ 时, DP-FedAvg 的模型测试准确率达到 70.63%, 而 FedBDP 的模型测试准确率达到 75.80%, FedBDP 相比于 DP-FedAvg 略高 5.17%. 由图 6(b) 可知, 在 CIFAR10 数据集下, FedProx 和 FedAvg 的模型测试准确率分别达到了 50.70 % 和 48.91%. 当 $\varepsilon = 10$ 时, DP-FedAvg 的模型测试准确率达到 29.98%, FedBDP 的模型测试准确率达到 34.23%, FedBDP 相比于 DP-FedAvg 略高 4.25%. 由此可见, FedBDP 方法在采用高斯机制实现差分隐私对本地模型进行保护的前提下, 能够有效提高模型准确率. 因此, 本文算法适用于对模型精度要求高且需要隐私保护的联邦学习应用场景.

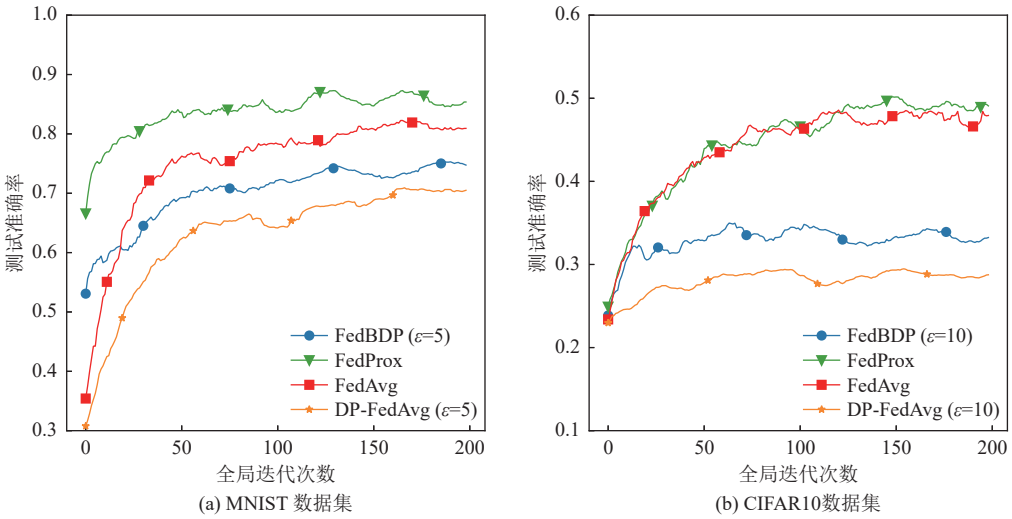


图 6 模型准确率对比

6 结束语

针对模型异构和隐私安全问题, 本文提出一种基于 Bregman 散度和差分隐私的个性化联邦学习方法. 该方法在非独立同分布数据下, 利用 Bregman 散度重新定义联邦学习中的损失函数, 并行优化各客户端数据对其进行个性化训练, 有效提高模型准确率. 同时在客户端上传本地模型参数时, 定义自适应差分隐私合理分配噪声大小, 以减小噪声扰动对模型准确率的影响. 通过理论分析了 FedBDP 在强凸和非凸光滑函数下的收敛性与隐私性. 实验结果表明 FedBDP 方法在实现隐私保护的前提下, 确保了模型准确率. 然而, 该方法中高斯机制的差分隐私预算分配容易造成过大的隐私开销, 下一步将考虑设计效率更高的隐私预算分配算法, 有效减少 FedBDP 模型的开销成本.

References:

[1] Naseem U, Razzak I, Khan SK, Prasad M. A comprehensive survey on word representation models: From classical to state-of-the-art word representation language models. ACM Trans. on Asian and Low-resource Language Information Processing, 2021, 20(5): 74. [doi: 10.1145/3434237]

[2] Toldo M, Maracani A, Michieli U, Zanuttigh P. Unsupervised domain adaptation in semantic segmentation: A review. Technologies, 2020, 8(2): 35. [doi: 10.3390/technologies8020035]

[3] López KL, Gagné C, Gardner MA. Demand-side management using deep learning for smart charging of electric vehicles. IEEE Trans. on

Smart Grid, 2019, 10(3): 2683–2691. [doi: 10.1109/TSG.2018.2808247]

[4] Lin WY, Hu YH, Tsai CF. Machine learning in financial crisis prediction: A survey. IEEE Trans. on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2012, 42(4): 421–436. [doi: 10.1109/TSMCC.2011.2170420]

[5] Waring J, Lindvall C, Umeton R. Automated machine learning: Review of the state-of-the-art and opportunities for healthcare. Artificial Intelligence in Medicine, 2020, 104: 101822. [doi: 10.1016/j.artmed.2020.101822]

[6] Pouyanfar S, Sadiq S, Yan YL, Tian HM, Tao YD, Reyes MP, Shyu ML, Chen SC, Iyengar SS. A survey on deep learning: Algorithms, techniques, and applications. ACM Computing Surveys, 2018, 51(5): 92. [doi: 10.1145/3234150]

[7] Yang Q, Liu Y, Chen TJ, Tong YX. Federated machine learning: Concept and applications. ACM Trans. on Intelligent Systems and Technology, 2019, 10(2): 12. [doi: 10.1145/3298981]

[8] Bonawitz KA, Eichner H, Grieskamp W, *et al*. Towards federated learning at scale: System design. In: Proc. of the 2nd Conf. on Machine Learning and Systems (MLSys 2019). Stanford, 2019.

[9] McMahan B, Moore E, Ramage D, Hampson S, Arcas BAY. Communication-efficient learning of deep networks from decentralized data. In: Proc. of the 20th Int'l Conf. on Artificial Intelligence and Statistics. Fort Lauderdale: AISTATS, 2017. 1273–1282.

[10] Li T, Sahu AK, Talwalkar A, Smith V. Federated learning: Challenges, methods, and future directions. IEEE Signal Processing Magazine, 2020, 37(3): 50–60. [doi: 10.1109/MSP.2020.2975749]

[11] Nishio T, Yonetani R. Client selection for federated learning with heterogeneous resources in mobile edge. In: Proc. of the 2019 IEEE Int'l Conf. on Communications. Shanghai: IEEE, 2019. 1–7. [doi: 10.1109/ICC.2019.8761315]

[12] Tran NH, Bao W, Zomaya A, Nguyen MNH, Hong CS. Federated learning over wireless networks: Optimization model design and analysis. In: Proc. of the 2019 IEEE Conf. on Computer Communications. Paris: IEEE, 2019. 1387–1395. [doi: 10.1109/INFOCOM.2019.8737464]

[13] Zhu JM, Zhang QN, Gao S, Ding QY, Yuan LP. Privacy preserving and trustworthy federated learning model based on blockchain. Chinese Journal of Computers, 2021, 44(12): 2464–2484 (in Chinese with English abstract). [doi: 10.11897/SP.J.1016.2001.02464]

[14] Kairouz P, McMahan HB, Avent B, *et al*. Advances and open problems in federated learning. Foundations and Trends® in Machine Learning, 2021, 14(2): 1–210. [doi: 10.1561/2200000083]

[15] Smith V, Chiang CK, Sanjabi M, Talwalkar A. Federated multi-task learning. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 4427–4437.

[16] Fredrikson M, Jha S, Ristenpart T. Model inversion attacks that exploit confidence information and basic countermeasures. In: Proc. of the 22nd ACM SIGSAC Conf. on Computer and Communications Security. Denver: ACM, 2015. 1322–1333. [doi: 10.1145/2810103.2813677]

[17] Liu YX, Chen H, Liu YH, Li CP. Privacy-preserving techniques in federated learning. Ruan Jian Xue Bao/Journal of Software, 2022, 33(3): 1057–1092 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6446.htm [doi: 10.13328/j.cnki.jos.006446]

[18] Yin XF, Zhu YM, Hu JK. A comprehensive survey of privacy-preserving federated learning: A taxonomy, review, and future directions. ACM Computing Surveys, 2021, 54(6): 131. [doi: 10.1145/3460427]

[19] Geyer RC, Klein T, Nabi M. Differentially private federated learning: A client level perspective. arXiv:1712.07557, 2017.

[20] Hao M, Li HW, Xu GW, Liu S, Yang HM. Towards efficient and privacy-preserving federated deep learning. In: Proc. of the 2019 IEEE Int'l Conf. on Communications (ICC). Shanghai: IEEE, 2019. 1–6. [doi: 10.1109/ICC.2019.8761267]

[21] Abadi M, Chu A, Goodfellow I, Mcmahan HB, Mironov I, Talwar K, Zhang L. Deep learning with differential privacy. In: Proc. of the 2016 ACM SIGSAC Conf. on Computer and Communications Security. Vienna: ACM, 2016. 308–318. [doi: 10.1145/2976749.2978318]

[22] Tan AZ, Yu H, Cui LZ, Yang Q. Towards personalized federated learning. IEEE Trans. on Neural Networks and Learning Systems, 2022, 1–17. [doi: 10.1109/TNNLS.2022.3160699]

[23] Zhao Y, Li M, Lai LZ, Suda N, Civin D, Chandra V. Federated learning with non-IID data. arXiv:1806.00582, 2018.

[24] Zhang XW, Hong MY, Dhople S, Yin W, Liu Y. FedPD: A federated learning framework with adaptivity to non-IID data. IEEE Trans. on Signal Processing, 2021, 69: 6055–6070. [doi: 10.1109/TSP.2021.3115952]

[25] Jeong E, Oh S, Kim H, Park J, Bennis M, Kim SL. Communication-efficient on-device machine learning: Federated distillation and augmentation under non-IID private data. arXiv:1811.11479, 2018.

[26] Fallah A, Mokhtari A, Ozdaglar A. Personalized federated learning: A meta-learning approach. arXiv:2002.07948, 2020.

[27] Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks. In: Proc. of the 34th Int'l Conf. on Machine Learning. Sydney: JMLR.org, 2017. 1126–1135.

[28] Li T, Sahu AK, Zaheer M, Sanjabi M, Talwalkar A, Smith V. Federated optimization in heterogeneous networks. In: Proc. of the 3rd Conf. on Machine Learning and Systems (MLSys 2020). Austin. 2020.

[29] Bregman LM. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. USSR Computational Mathematics and Mathematical Physics, 1967, 7(3): 200–217. [doi: 10.1016/0041-5553(67)90040-7]

[30] Dwork C, Roth A. The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science, 2013, 9(3–4): 211–407. [doi: 10.1561/0400000042]

[31] McSherry F, Talwar K. Mechanism design via differential privacy. In: Proc. of the 48th Annual IEEE Symp. on Foundations of Computer Science. Providence: IEEE, 2007. 94–103. [doi: 10.1109/FOCS.2007.66]

[32] McSherry FD. Privacy integrated queries: An extensible platform for privacy-preserving data analysis. In: Proc. of the 2009 ACM SIGMOD Int'l Conf. on Management of Data. Providence: ACM, 2009. 19–30. [doi: 10.1145/1559845.1559850]

[33] Jayaraman B, Evans D. Evaluating differentially private machine learning in practice. In: Proc. of the 28th USENIX Conf. on Security Symp. Santa Clara: USENIX Association, 2019. 1895–1912.

[34] Zhu LG, Liu ZJ, Han S. Deep leakage from gradients. In: Proc. of the 33rd Int'l Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2019. 1323.

[35] Song CZ, Ristenpart T, Shmatikov V. Machine learning models that remember too much. In: Proc. of the 2017 ACM SIGSAC Conf. on Computer and Communications Security. Dallas: ACM, 2017. 587–601. [doi: 10.1145/3133956.3134077]

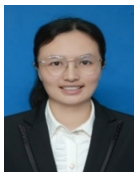
[36] Bauschke HH, Dao MN, Lindstrom SB. Regularizing with Bregman—Moreau envelopes. SIAM Journal on Optimization, 2018, 28(4): 3208–3228. [doi: 10.1137/17M1130745]

附中文参考文献:

[13] 朱建明, 张沁楠, 高胜, 丁庆洋, 袁丽萍. 基于区块链的隐私保护可信联邦学习模型. 计算机学报, 2021, 44(12): 2464–2484. [doi: 10.11897/SP.J.1016.2001.02464]

[17] 刘艺璇, 陈红, 刘宇涵, 李翠平. 联邦学习中的隐私保护技术. 软件学报, 2022, 33(3): 1057–1092. http://www.jos.org.cn/1000-9825/6446.htm [doi: 10.13328/j.cnki.jos.006446]

张少波(1979－), 男, 博士, 教授, CCF 高级会员, 主要研究领域为人工智能隐私保护与安全, 大数据隐私保护.



龙赛琴(1986－), 女, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为人工智能, 云计算, 边缘计算, 大数据.

张激勇(1998－), 男, 硕士, 主要研究领域为联邦学习, 差分隐私, 机器学习.

李哲涛(1980－), 男, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为人工智能, 物联网, 网络空间安全.

朱更明(1967－), 男, 教授, 主要研究领域为人工智能隐私保护, 信息安全, 机器视觉.