计算机科学
COMPUTER SCIENCE

基于差分隐私的联邦学习方案

孙敏, 丁希宁, 成倩

引用本文

孙敏, 丁希宁, 成倩. 基于差分隐私的联邦学习方案[J]. 计算机科学, 2024, 51(6A): 230600211-6.

SUN Min, DING Xining, CHENG Qian. Federated Learning Scheme Based on Differential Privacy[J]. Computer Science, 2024, 51(6A): 230600211-6.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.230600211

基于差分隐私的联邦学习方案

孙 敏 丁希宁 成 倩
山西大学计算机与信息技术学院 太原 030000

摘 要 联邦学习的特点之一是进行训练的服务器并不直接接触数据,因此联邦学习本身就具有保护数据安全的特性。但是研究表明,联邦学习在本地数据训练和中心模型聚合等方面均存在隐私泄露的问题。差分隐私是一种加噪技术,通过加入适当噪声达到攻击者区分不出用户信息的目的。文中研究了一种基于本地和中心差分隐私的混合加噪算法(LCDP-FL),该算法能根据各个客户端不同权重、不同隐私需求,为这些客户端提供本地或混合差分隐私保护。而且我们证明该算法能够在尽可能减少计算开支的同时,为用户提供他们所需的隐私保障。在 MNIST 数据集和 CIFAR-10 数据集上对该算法进行了测试,并与本地差分隐私(LDP-FL)和中心差分隐私(CDP-FL)等算法进行对比,结果显示该混合算法在精确度、损失率和隐私安全方面均有改进,其算法性能最优。

关键词:联邦学习;差分隐私;隐私保护;混合加噪;梯度下降
中图分类号 TP393

Federated Learning Scheme Based on Differential Privacy

SUN Min,DING Xining and CHENG Qian
College of Computer and Information Technology,Shanxi University,Taiyuan 030000,China

Abstract One of the characteristics of federated learning is that the server being trained does not directly contact the data,so federated learning itself has the characteristics of protecting data security. However,research shows that federated learning has privacy leakage problems in local data training and central model aggregation. Differential privacy is a noise augmentation technique that adds appropriate noise to prevent an attacker from distinguishing user information. We study a hybrid noise adding algorithm based on local and central differential privacy(LCDP-FL),which can provide local or hybrid differential privacy protection for each client according to its different weights and privacy requirements. It's shown that the algorithm can provide users with the privacy they need with minimal computational overhead. The algorithm is tested on the MNIST dataset and CIFAR-10 dataset,and compared with local differential privacy(LDP-FL) and central differential privacy(CDP-FL) algorithms,and the results show that the hybrid algorithm has improved accuracy,loss rate and privacy security,and its algorithm performance is the best.

Keywords Federated learning,Differential privacy,Privacy protection,Hybrid noise,Gradient descent

1 引言

近年来,机器学习在金融、医疗、城市规划、自动驾驶等多个领域发挥了重要作用,其性能和隐私性也逐渐受到关注。传统的机器学习可能会在数据的收集过程中泄露敏感数据,这将直接威胁用户隐私安全。为此联邦学习[1]应运而生。联邦学习是一种数据保留在本地,多个客户端协同训练模型而不分享数据,最终对多个客户端训练的模型进行聚合的学习机制。尽管联邦学习避免了将数据直接暴露给第三方,但是其中依然存在大量隐私泄露的风险。比如:1)我们不能保证所有的客户机参与方都是可靠的,不可靠的参与方增加了隐私泄露的风险;2)联邦学习需要把本地训练好的模型参数上传至中心服务器,攻击者可以通过模型参数反推出用户信息;3)训练完成的模型也面临着隐私泄露的风险。

针对联邦学习所面临的隐私风险,目前有两种解决思路,分别是加密和扰动。加密最常用的方法是利用同态加密[1]和秘密分享[1]技术为联邦训练过程中数据的传输提供隐私保护,这两种技术是常用的密码学工具,但因为其计算代价高,会带来通信开销大的问题。扰动就是通过差分隐私等技术在本地数据或模型训练过程产生的梯度数据上添加噪声,使发布的模型在保持可用性的同时得到保护。这些方法的应用及优缺点如表 1 所列。

针对以上不足,本文主要工作如下:

1)提出一种基于本地和中心差分隐私相结合的联邦学习方法,结合两者优势结合解决联邦学习训练过程中的隐私泄露问题。

2)设计一种随机为用户选择差分隐私加噪方式的算法,在保证联邦模型训练隐私性的基础上,最大化数据的利用率并减少隐私预算。

基金项目:山西省基础研究计划(20210302123455,201701D121052)
This work was supported by the Shanxi Province Basic Research Program,China(20210302123455,201701D121052).
通信作者:孙敏(minsun@sxu.edu.cn)

3)在 MNIST 和 CIFAR-10 这 2 个真实的数据集上，分别从精确率、损失率和运行时间 3 个方面与 LDP-FL，CDP-FL 和 Paillier-FL 3 种算法进行对比实验，实验结果表明本文方法的效果更优。

表 1 研究方案对比

Table 1 Comparison of research plans

技术	方案	概述	缺点
差分隐私	文献[2]	设计用户级别的差分隐私联邦学习算法	需要可信的第三方服务器
	文献[3]	利用指数机制思想设计差分隐私参数扰动	性能损失大
同态加密	文献[4]	基于同态加密思想设计联邦学习参数保护方法	计算代价高
	文献[5]	由终端自行生成密钥，联邦模型训练时结合加法同态技术对参数进行保护	
	文献[6]	引入第三方进行私钥分配后再使用同态加密技术	
秘密共享	文献[7]	结合秘密分享技术对参与方传递的参数进行保护	通信开销大

2 背景知识

2.1 联邦学习

2017 年，Google 公司首次提出联邦学习的概念。联邦学习保证数据不出本地，本地客户机通过训练本地数据得到模型并发送给中心服务器，最后由中心服务器完成模型的聚合。理想状况下，联邦学习得到的共享模型与数据集中在中心服务器上训练所得的模型相比，效果相近或更好。联邦学习的数据不出本地理念对数据隐私有天然的保护作用，但是其中依然存在大量隐私泄露的风险。根据参与各方数据源分布的情况不同，联邦学习可以被分为 3 类：横向联邦学习、纵向联邦学习、联邦迁移学习。联邦学习如图 1 所示。本文用 FedAvg 算法作为联邦聚合算法，采用加权平均的方式进行模型聚合。该算法的优势在于通信消耗小，并且可以处理不同的数据集。

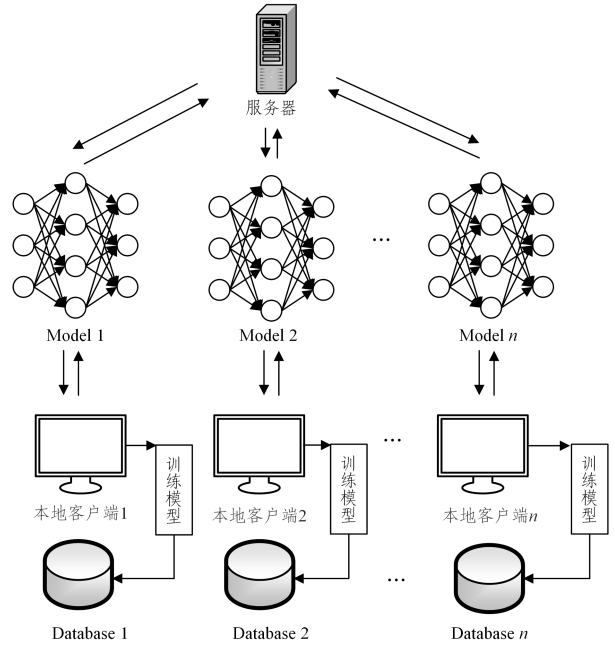

图 1 联邦学习示例图

Fig. 1 Example diagram of federated learning

2.2 差分隐私

差分隐私[8]最早于 2008 年由 Dwork 提出，通过严格的数学证明，使用随机应答方法确保数据集在输出信息时受单条记录的影响始终低于某个阈值，从而使第三方无法根据输出的变化判断单条记录的更改或增删，被认为是目前基于扰动的隐私保护方法中安全级别最高的方法。现给定两个相邻数据集 D 和 D'，它们有且仅有一条数据不一样。那么对于一个随机化算法 A，使之满足 $\frac{\Pr\{A(D)=O\}}{\Pr\{A(D')=O\}}\leqslant e^{\varepsilon}$，则称该算法 A 是满足差分隐私的。根据使用差分隐私加噪的位置不同，差分隐私被分为本地差分隐私和中心差分隐私。二者都可以防御后门攻击，但是前者不能抵抗属性推理攻击，而后者在消耗一些数据性能成本的前提下可以抵抗属性推理攻击。

2.2.1 本地差分隐私

部分用户可能不愿意相信本地数据收集器，他们可能希望数据在被送到中心服务器之前就被加噪。本地化差分隐私算法在数据被送去参加训练前可以帮助这些用户在本地进行加噪，并把加噪之后的数据送至中心服务器。因为给数据加噪的行为发生在本地客户机，所以称该方式为本地化差分隐私。这样数据收集器也不会准确获得用户隐私数据，用户也不需要依赖于可信的第三方服务器。本地差分隐私示意图如图 2 所示。

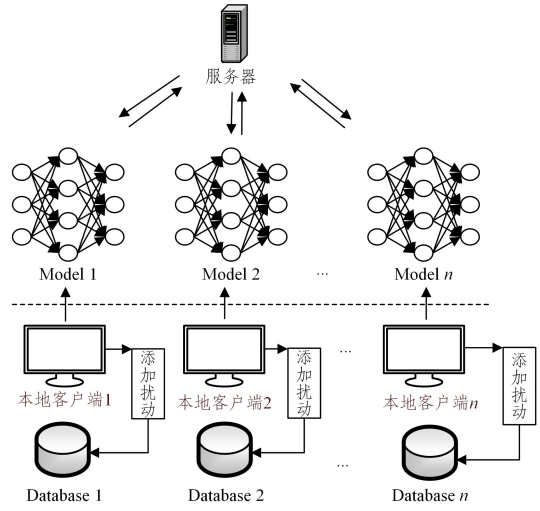

图 2 本地差分隐私示例图

Fig. 2 Example of local differential privacy

2.2.2 中心差分隐私

在联邦学习的过程中，中心服务器端收集客户端更新的梯度数据。现在有很多文献证明，通过梯度数据可以反推出用户信息，中心化差分隐私通过中心服务器为梯度数据加噪来达到保护用户信息的目的，因为加噪的行为发生在中心服务器，所以称之为中心化差分隐私。中心化差分隐私的局限在于需要依靠一个可信赖的中心服务器。中心差分隐私示意图如图 3 所示。

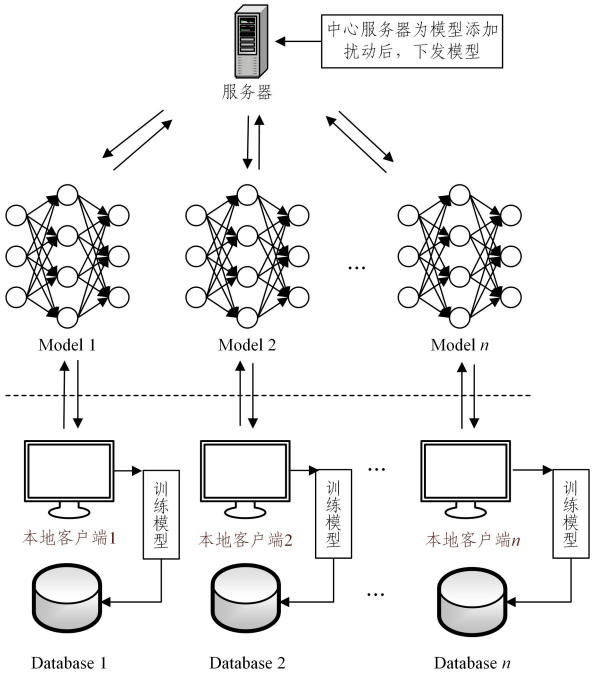

图 3　中心差分隐私示例图

Fig. 3　Example diagram of center differential privacy

3　问题定义

为抵御第 1 章中提到的 3 类隐私泄露风险，本文提出如下方案思路。1）敌手可依据某客户端每轮上传的模型数据，推测出本地数据集。对此，本文利用本地差分隐私技术，利用本地客户机对数据进行加噪，从而避免局部信息的泄露。2）中心服务器的工作是与参与者通信、交换和存储参数。在中心服务器给客户机下发模型时，可能造成训练好的模型参数的泄露。对此，本文基于中心差分隐私的思想，令中心服务器对本地训练得到的局部模型添加一定扰动，将扰动后的模型下发至客户机，使得每一轮的聚合过程满足差分隐私，进而达到保护模型数据的目的。

首先我们应该解决的问题是用户加噪方式的选取。如果所有用户的加噪方式都一样，不仅需要消耗的隐私预算大，而且攻击者会根据加噪方式采取相应的攻击手段。本文设置了两组用户：只本地加噪用户和中心本地混合加噪用户。本文提出的用户选择算法 opt-user 可以为用户分配不同的加噪方式。

尽管本地差分隐私可以使数据收集器无法获得准确数据，但是如果攻击者发动足够多的攻击，仍可以获得相对准确的用户隐私，并且本地差分隐私算法不能抵御属性推理攻击，但中心差分隐私算法可以[9]。因此我们选择本地与中心混合加噪方式来更好地保护用户隐私。

对于本地差分隐私算法和中心差分隐私的加噪方式，在本地我们选择随机响应技术[10]，在中心我们选择用拉普拉斯函数[11]加噪。需要应用隐私扰动的位置与时间均不相同。在本地客户机中，数据在被训练成模型发送到服务器之前需要进行扰动，而在中心服务器中，服务器先收集所有梯度数据，完成聚合之后，对聚合模型数据加噪，然后下发给各个客户机进行下一次训练。同时，本地与中心差分隐私加噪对象的不同也导致其“相邻数据库”的含义存在差异。在本地差分隐私算法中，D 表示单个用户的数据，D' 表示同一用户的数据，只是确定该用户的数据在不在数据库中。在中心差分隐私算法中，D 代表所有用户的数据，D' 代表所有用户的数据，只有用户的某个值可能发生改变。

该算法运行一次的具体流程为，中心服务器给客户机下发训练模型，各个客户机接收训练模型后，统计本地数据量并上传至中心服务器，服务器根据各个客户机数据量所占数据总量的比重为各个客户机分配权重。利用用户选择算法为各个客户机分配加噪方式，之后客户机利用加噪数据训练模型。模型训练好后被送至中心服务器，中心服务器按权重聚合模型，并根据客户机的加噪方式选择添加或不添加中心差分隐私噪声得到本次训练的最终模型，将最终模型下发给各个客户机，以此类推，开始下一轮的训练。算法总体框架图如图 4 所示。

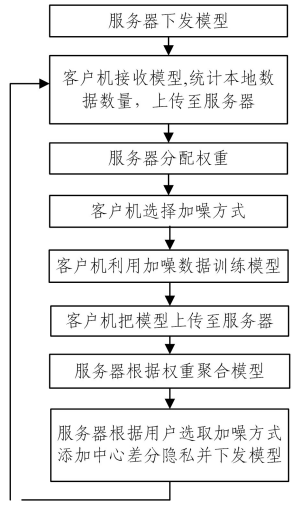

图 4　算法总体框架图

Fig. 4　Overall framework of the proposed algorithm

3.1　用户选择算法设计

所有客户机随机生成 flag，flag 的值为 0 或 1，0 代表不希望添加中心差分隐私，1 代表希望添加中心差分隐私并把数字发送给中心服务器。中心服务器在接收到客户机的 flag 后，会根据以下规律生成 newflag。如果收到的 flag 为 1，则判定客户机为中心本地混合加噪，如果收到的 flag 为 0，则该 flag 有 50% 的概率不变，newflag 也为 0，判定该客户机只本地加噪，有 1% 的概率变为 1，newflag 为 1，判定客户机为中心本地混合加噪。生成的 newflag 保存在服务器中，并不发送给客户机，这样各客户机也并不知道相互之间的加噪方式是什么，不仅减少了部分隐私预算，而且减小了恶意客户端通过加噪方式攻击数据成功的概率。

算法 1　用户选择加噪方式算法

输入：flag

输出：newflag

1. 选择用户：以采样率 q 从 N 个用户中选择 k 个参与方；
2. 初始化：用户随机生成自己的标签 flag 值为 0 或 1，并发送给服务器；
3. 服务器收到 flag=1，加噪方式为混合，newflag=1；
4. 服务器收到 flag=0，50% 概率随机扰动成 1；
5. 扰动后 newflag=0，则只进行本地加噪；
6. 扰动后 newflag=1，则进行混合加噪；
7. 输出结果：newflag。

3.2　本地差分隐私算法设计

为了解决诚实且好奇的中心参数服务器或参与方的存在

导致联邦学习中用户本地数据隐私泄露的问题，本文提出了一种基于本地差分隐私的方法，框架如图 1 所示。该方法由一个中心参数服务器和 N 个联邦学习客户机参与方组成，每个联邦学习参与方拥有一个由中心参数服务器下发的初始模型和本地的训练数据集。本地差分隐私方法的核心思想是对本地数据依据随机响应机制进行扰动，本地数据的图像转化为区间[−1,1]之间的 4 位小数，把[−1,1]平均分成 20 个区间，把各个区间内的数字四舍五入成一位小数，这样就得到了 21 个数字，按照 $P(R'|R)=\frac{1}{k-1+e^{\varepsilon}}\begin{cases}e^{\varepsilon}, & \text{if } R'=R\\ 1, & \text{if } R'\neq R\end{cases}$ 这个公式对这 21 个值进行扰动。具体来说，首先根据算法 1 选择用户加噪方式。然后被选中的客户端统计数据量上传至服务器，服务器根据 $\frac{w_i}{w_1+w_2+\cdots+w_n}$ 产生权重并保存在服务器。最后由中心服务器生成初始模型再分发给所选择的本地客户机，客户机接收到初始模型后利用本地加噪的数据集对初始模型进行训练。在每个客户机本地训练开始前，引入本地化差分隐私机制对原始数据进行扰动，利用扰动后的数据训练模型，避免恶意客户机泄露用户隐私。

算法 2 本地差分隐私算法

输入：输入数据集 D，本地模型迭代次数 E，本地差分隐私机制隐私参数 ε

输出：加噪后数据集 NEWD

1. Normalize 归一化数据集；
2. 把数据集平均分成 20 个区间，并向上取整得到 D[i]；
3. 对 D[i]中 21 个数字添加扰动；
4. 得到扰动后的数据集。

3.3 中心差分隐私算法

为了解决中心服务器下发模型时造成的模型数据泄露，本文采用基于中心差分隐私的方法。中心差分隐私方法的核心思想是对要下发的模型加噪，从而达到保护数据的效果。具体来说，由本地客户机利用数据训练模型，再发送给中心服务器，中心服务器接收到模型后利用拉普拉斯噪声对模型进行加噪。最后把加噪的模型下发给各个客户机。

算法 3 中心差分隐私算法

输入：加噪数据集 NEWD，联邦学习参与方数量 N，联邦学习采样率 q，联邦学习交流轮次 T

输出：训练完的模型 M

1. for t：1 to T do；
2. 以采样率 q 从 N 个用户中选择 k 个参与方；
3. 遍历从 N 中选择的 k 个参与方的训练模型 M_i；
4. 按权重聚合模型 $M\leftarrow\sum_{i=1}^{k}\frac{1}{w_i}M_i$；
5. 对聚合模型加噪 $M\leftarrow M+Lap(\chi|\beta)$；
6. 下发模型 M。

算法 4 Non_IID 数据处理算法

输入：要训练的数据集

输出：Non_IId 数据集

1. label_index 按类别排序获得数据样本 id 列表；
2. 设置长尾分布函数，输入上一步所得列表 label_index，数据的类别 num_class，数据不平衡的程度 imb_factor，根据 imb_type 获得每个类别数据量，longtail_setting=(label_index,num_class,imb_factor,imb_type)；
3. 通过 dirichlet 函数生成 Non_IId 数据集。

4 算法安全性分析

4.1 本地差分隐私算法安全性分析

本节对本地差分隐私方法的隐私安全性进行分析，采用 Peter 等[8]提出的对多个数据进行扰动的方法。P 代表数据受到扰动后会改变的概率，$1-P$ 是不会改变的概率。

扰动公式：

$$D_i'=\begin{cases}D-D_i, & \frac{P}{20}\\ D_i, & 1-P\end{cases}$$

取 $P=\frac{1}{e^{\varepsilon}+20}$。需证明 $\frac{\Pr(\widetilde{D}=D'|D=D_i)}{\Pr(\widetilde{D}=D'|D=D-D_i)}\leqslant e^{\varepsilon}$。证：

$$\Pr(D'=D_i|D=D_i)=1-20P$$

$$\Pr(D'=D_i|D=D-D_i)=20\times\frac{1}{20}P=P$$

$$\frac{\Pr(\widetilde{D}=D'|D=D_i)}{\Pr(\widetilde{D}=D'|D=D-D_i)}=\frac{e^{\varepsilon}}{e^{\varepsilon}+20}/\left(20\times\frac{1}{20}\times\frac{1}{e^{\varepsilon}+20}\right)=e^{\varepsilon}$$

因此 $P=\frac{1}{e^{\varepsilon}+20}$ 时，该方案满足本地差分隐私。

4.2 中心差分隐私算法安全性分析

本节对中心差分隐私方法的隐私安全性进行分析，对于采样率为 q、迭代轮次为 T 的中心差分隐私方法，取 $\beta\geqslant\frac{\Delta f}{\varepsilon}$，证明如下：

$$\begin{aligned}lap(x|\beta)&=\frac{1}{2\beta}e^{\frac{-|x|}{\beta}}\frac{\Pr(A(D)=t)}{\Pr(A(D')=t)}\\&=\frac{\Pr(F(D)+x=t)}{\Pr(F(D')+x=t)}\\&=\frac{\Pr(x=t-F(D))}{\Pr(x=t-F(D'))}\\&=\frac{\frac{1}{2\beta}\exp\left[-\frac{|t-F(D)|}{\beta}\right]}{\frac{1}{2\beta}\exp\left[-\frac{|t-F(D')|}{\beta}\right]}\\&=\exp\left[\frac{|t-F(D')|-|t-F(D)|}{\beta}\right]\end{aligned}$$

利用三角不等式：$|a|-|b|\leqslant|a-b|$

$\therefore|t-F(D')|-|t-F(D)|\leqslant|F(D)-F(D')|$

$$\therefore\exp\left[\frac{|t-F(D')|-|t-F(D)|}{\beta}\right]\leqslant\exp\left[\frac{|F(D)-F(D')|}{\beta}\right]$$

$\because|F(D)-F(D')|\leqslant\Delta f$

$$\therefore\exp\left[\frac{|F(D)-F(D')|}{\beta}\right]\leqslant\exp\left(\frac{\Delta f}{\beta}\right)$$

$$\therefore\frac{\Delta f}{\beta}\leqslant\varepsilon$$

所以 $\beta\geqslant\frac{\Delta f}{\varepsilon}$ 时，原式成立，满足差分隐私。

5 实验分析

本章使用卷积神经网络(Convolutional Neural Networks，CNN)在手写数字数据集 MNIST 和 CIFAR-10 上评估 LC-FL 算法。从不同参数下的隐私预算对模型成功率的

影响以及 LC-FL 与现有方法的比较两个方面验证了所提出的 LC-FL 算法的合理性。

5.1　实验设置

5.1.1　实验环境

本节对本文所提 LC-FL 方法的有效性进行评估，并设计对比实验。所使用的实验平台操作系统为 Windows 10 (64 位)，开发环境为 Pycharm，编程语言为 Python 3.8，CPU 为 Intel(R) Core(TM) i7-10700 CPU @ 2.90 GHz 2.90 GHz，内存为 16 GB。实验使用 Pytorch1.7.1 训练深度学习模型，采用卷积神经网络(CNN)构建本文所提 LC-FL 方法，设置 2 个卷积层分别有 16 和 32 个特征，并使用一个 5×5、步长为 2 的卷积核，以及一个输入张量为 7×7×32、输出张量为 10 的全连接层，采用梯度下降进行模型训练时所选择的批次大小为 64，参与方本地训练迭代次数为 10 次。本文实验中联邦学习的客户端数目为 100 个，默认参与训练比例为 0.5。IID 和 Non_IID 训练 100 轮。训练集被分为 10 个类，对于 IID 数据实验的设置，每个客户端随机分配 10 个均匀分布的类。对于 Non_IID 数据实验的设置，每个客户端分配 10 个按指数分布的类，其中数据最多的类的数据量是数据最少的类的数据量的 100 倍。

5.1.2　数据集

实验采用 2 种数据集，分别是 MINST 数据集和 CIFAR-10 数据集。其中，MINST 数据集包含 10 种手写数字识别的灰度图像数据，有 60000 个训练图像和 10000 个测试图像，每个灰度图像包含 28 像素×28 像素；CIFAR-10 是一个用于识别普适物体的小型数据集，一共包含 10 个类别的 RGB 彩色图片，每张图像尺寸为 3×32 像素×32 像素。数据集中一共有 50000 张训练图片和 10000 张测试图片。

5.1.3　评价指标

1)全局准确率。经过多次迭代后，联邦模型的全局准确率是衡量算法有效性的关键指标。通过对比相同条件下不同算法的全局准确率，可以直观地判断算法的性能。

2)性能损失。性能损失是衡量联邦模型性能的指标，通过性能估计机制进行计算。

3)运行时间。算法的运行时间是衡量通信开销的重要指标。运行时间越长，则通信开销越大。

5.2　实验结果与对比分析

本节探究 LC_FL 算法的有效性，将 LC-FL 算法与基于本地差分隐私的联邦学习算法 LDP-FL、基于中心差分隐私的联邦学习算法 CDP-FL 和基于同态加密的联邦学习算法 Paillier-FL 在 MINST 数据集和 CIFAR-10 上的结果进行比较。

5.2.1　精确度

本联邦学习迭代轮次 $T=20$，本地隐私预算 $\varepsilon_1=5$，全局隐私预算 $\varepsilon_2=10$，总共隐私预算 $\varepsilon=15$。LC-FL 在 MINST 与 CIFAR-10 IID 数据集上的精确度分别为 94.3%与 44.5%，在 Non_IID 数据集上的精确度分别为 91.9%与 42.3%，比 LDP-FL，CDP-FL 和 Paillier-FL 算法的精确度略高，但是 LC-FL 算法能带来更高的安全性，因此其性能更好。精确度对比图如图 5－图 8 所示。

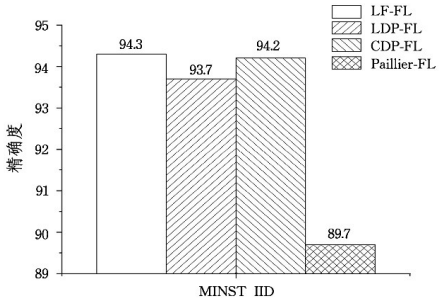

图 5　MINST IID 数据精确度对比图

Fig. 5　MINST IID data accuracy comparison

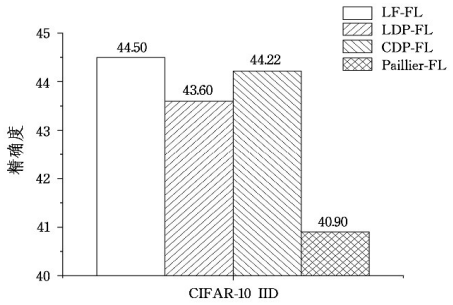

图 6　CIFAR-10 IID 数据精确度对比图

Fig. 6　CIFAR-10 IID data accuracy comparison

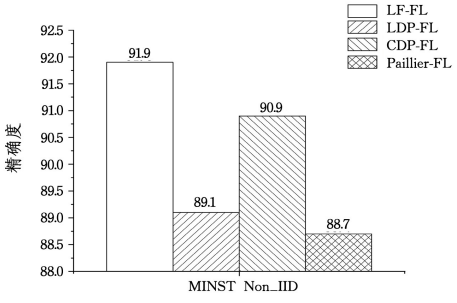

图 7　MINST Non_IID 数据精确度对比图

Fig. 7　MINST Non_IID data accuracy comparison

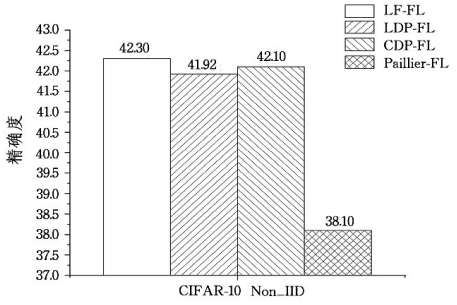

图 8　CIFAR-10 Non_IID 数据精确度对比图

Fig. 8　CIFAR-10 Non_IID data accuracy comparison

5.2.2　损失率

差分隐私的特性会对数据的可用性造成影响。在 LC-FL 算法的损失率与 LDP-FL，CDP-FL 几乎相同的前提下，LC-FL 算法的准确率比这两个算法高，而且 LC-FL 算法的损失率明显低于 Paillier-FL，并且能为算法提供隐私保证，因此可以证明 LC-FL 算法性能更好。损失率结果如表 2 和表 3 所列。

表 2 IID 数据损失率对比表

Table 2 IID data loss rate comparison

算法	MINST		CIFAR-10	
	精确度	损失率	精确度	损失率
LC-FL	94.3	1.5	44.50	1.5
LDP-FL	93.7	1.4	43.60	1.4
CDP-FL	94.2	1.5	44.22	1.5
Paillier-FL	89.7	2.1	40.90	2.3

表 3 Non_IID 数据损失率对比

Table 3 Non_IID data loss rate comparison

算法	MINST		CIFAR-10	
	精确度	损失率	精确度	损失率
LC-FL	91.9	1.9	42.30	2.8
LDP-FL	89.1	2.1	41.92	3.0
CDP-FL	90.9	2.3	42.10	3.1
Paillier-FL	88.7	2.7	38.10	3.5

5.2.3 时间

在时间效率上，LC-FL 算法消耗的时间与 LDP-FL，CDP-FL 相近，这是增加差分隐私的缘故，但时间消耗远远低于 Paillier-FL。LC-FL 可以为用户提供更高的安全性，可以利用训练时间换取用户隐私的安全，并且训练时间也是可接受的。

结束语 本文提出一种基于本地和中心差分隐私的联邦学习算法 LC-FL，在为各个客户端提供安全保证的前提下，还能提供较高的精确度和较低的损失率。核心思想是将该机制作用在联邦学习参数的传递过程中，增加联邦模型训练的隐私性。在真实的数据集上通过实验验证了所提 LC-FL 方法的有效性。未来的工作将集中在如何减少差分隐私带来的可用性损失上，以及隐私保护联邦学习在应用方面的拓展。如研究在医疗和物联网环境下如何在保证隐私安全的同时提高联邦模型的准确率。

参 考 文 献

[1] LIU Y X, CHEN H, LIU Y H, et al. Privacy-preserving techniques in federated learning[J]. Journal of Software, 2022, 33(3): 1057-1092.

[2] GEYER R C, KLEIN T, NABI M. Differentially private federated learning: a client level perspective[J]. arXiv: 1712.07557, 2017.

[3] TRUEX S, LIU L, CHOW K H, et al. LDP-Fed: federated learning with local differential privacy[C]// Proceedings of the Third ACM International Workshop on Edge Systems, Analytics and Networking. New York: ACM Press, 2020: 61-66.

[4] LIU X Y, LI H W, XU G W, et al. Privacy-enhanced federated learning against poisoning adversaries[J]. IEEE Transactions on Information Forensics and Security, 2021, 16: 4574-4588.

[5] PHONG L T, AONO Y, HAYASHI T, et al. Privacy-preserving deep learning via additively homomorphic encryption[C]// Proceedings of IEEE Transactions on Information Forensics and Security. Piscataway: IEEE Press, 2019: 1333-1345.

[6] OU W, ZENG J, GUO Z, et al. A homomorphic-encryption-based vertical federated learning scheme for rick management [J]. Computer Science and Information Systems, 2020, 17(3): 819-834.

[7] TANG L T, WANG D, ZHANG L F, et al. Federated learning scheme based on secure multi-party computation and differential privacy[J]. Computer Science, 2022, 49(9): 297-305.

[8] KAIROUZ P, BONAWITZ K, RAMAGE D. Discrete distribution estimation under local privacy[C]// International Conference on Machine Learning. PMLR, 2016: 2436-2444.

[9] NASERI M, HAYES J, DE CRISTOFARO E. Toward robustness and privacy in federated learning: Experimenting with local and central differential privacy[C]// Proceedings of the 33rd International Conference on International Conference on Machine Learning(ICML'16). Volume 48. 2016: 2436-2444.

[10] KAIROUZ P, BONAWITZ K, RAMAGE D. Discrete Distribution Estimation under Local Privacy[C]// Proceedings of the 33rd International Conference on International Conference on Machine Learning(ICML'16). 2016: 2436-2444.

[11] HUANG J W. Federated learning data privacy security technology based on differential privacy[J]. Communication Technology, 2022, 55(12): 1618-1625.

SUN Min, born in 1966, master, professor. Her main research interests include computer network and information security.