基于谱聚类的傅里叶个性化联邦学习研究

金 彤 陈思光*

(南京邮电大学物联网学院 南京 210003)

摘 要：为了缓解联邦学习中跨不同用户终端数据非独立同分布(non-IID)引起的负面影响，该文提出一种基于谱聚类的傅里叶个性化联邦学习算法。具体地，构建一个面向图像分类识别的云边端协同个性化联邦学习模型，提出在云端协同下通过谱聚类将用户终端划分为多个聚类域，以充分利用相似用户终端学到的知识提升模型性能。其次，设计边端协同的局部联邦学习方法，通过代理模型在用户终端对个性化局部模型执行恢复与再更新的操作，可有效恢复聚合过程中丢失的本地知识。进一步地，设计云边协同的傅里叶个性化联邦学习方法，即云服务器通过傅里叶变换将局部模型参数转换到频域空间上进行聚合，为每个边缘节点定制高质量的个性化局部模型，可使全局模型更适用于各个分布式用户终端。最后，实验结果表明，与现有相关算法相比，所提算法收敛速度更快，准确率提高了3%～13%。

关键词：边缘计算；联邦学习；谱聚类；傅里叶变换

中图分类号：TN919.2; TP393 **文献标识码**：A **文章编号**：1009-5896(2023)06-1981-09

DOI: 10.11999/JEIT220529

Fourier Personalized Federated Learning Mechanism Based on Spectral Clustering

JIN Tong CHEN Siguang

(*School of Internet of Things*, *Nanjing University of Posts and Telecommunications*, *Nanjing* 210003, *China*)

Abstract: To relieve the negative impacts caused by non-Independent and Identically Distributed (non-IID) data across different clients in federated learning, a spectral clustering-based Fourier personalized federated learning mechanism is proposed to overcome the performance drops from data heterogeneity. Specifically, a cloud-edge-end collaborative personalized federated learning model for image recognition is constructed, and in order to make full use of the knowledge learned by similar clients, the clients are divided into multiple clusters by spectral clustering under cloud-edge collaboration. Next, a local federated learning method based on edge-end collaboration is proposed, in which an agent model is used to perform the process of restoring and re-updating the personalized local model at the clients to restore the local knowledge loss during aggregation. Furthermore, a cloud-edge collaborative Fourier personalized federated learning method is proposed to adapt the global model to each distributed client. In this method, the cloud server converts the local model parameters to the frequency domain space for aggregation through Fourier transform, and customizes high-quality personalized local model for each edge node. Finally, the experimental results demonstrate that the proposed algorithm obtains competitive convergence speed compared with existing representative works and the accuracy is 3%～13% higher.

Key words: Edge computing; Federated learning; Spectral clustering; Fourier transform

收稿日期：2022-04-27；改回日期：2022-12-07；网络出版：2022-12-23

*通信作者： 陈思光 sgchen@njupt.edu.cn

基金项目：国家自然科学基金(61971235)，中国博士后科学基金(2018M630590)，江苏省“333高层次人才培养工程”，江苏博士后科研资助计划(2021K501C)，南邮“1311”人才计划和江苏研究生科研创新计划(KYCX22_1029)

Foundation Items: The National Natural Science Foundation of China (61971235), China Postdoctoral Science Foundation (2018M630590), 333 High-level Talents Training Project of Jiangsu Province, Jiangsu Planned Projects for Postdoctoral Research Funds (2021K501C), 1311 Talents Plan of NJUPT, The Jiangsu Postgraduate Scientific Research Innovation Plan (KYCX22_1029)

1 引言

近年来，物联网技术已经广泛应用于智慧医疗、智慧交通和智慧农业等实际场景。物联网设备产生的巨大数据量使得基于数据驱动的人工智能解决方案在图像分类领域得到广泛应用[1]。由于数据驱动的传统机器学习方法往往以集中式训练模型为主，即中央服务器从不同的物联网用户终端收集分散的数据以训练机器学习模型，如卷积神经网络[2]，这类集中式的机器学习方法在实际应用时往往会面临一系列挑战，如大量原始数据传输带来的高通信开销和在通信过程中产生的隐私泄露风险等。

为了解决上述问题，Google于2017年提出一种去中心化的机器学习框架称为联邦学习(Federated Learning, FL)[3]。联邦学习使得用户终端以去中心化的方式，在不交换本地数据且不需要集中存储原始训练数据的情况下共同协作训练一个全局模型。具体地，各个参与联邦学习的分布式用户终端利用本地数据训练本地模型，将梯度上传给中央服务器聚合以更新全局模型，然后中央服务器将更新后的全局模型发送给用户终端用于更新本地模型，不断迭代直至收敛。由于联邦学习能够在充分利用多个用户终端数据训练机器学习模型的同时保护用户隐私并减少通信开销，因此联邦学习已经成为机器学习领域的研究热点，并已被应用于医疗图像分类识别、目标检测等多个领域[4–6]。

由于在实际应用中，跨不同用户终端的训练数据往往是非独立同分布的，也称为数据异构[7,8]。当用户终端之间数据分布有很大不同时，若用户终端直接获取从其他用户终端中学习到的知识，则会大大降低用户终端模型的性能。因此，相关研究者提出了一系列解决方案，用于解决数据异构所带来的问题，当前的相关研究大致可分为两类：基于预处理的方法和基于聚合方式优化的方法。

(1)基于预处理的方法。针对数据异构使训练模型产生次优结果的问题，部分研究对用户终端训练数据进行预处理，以此降低数据异构带来的负面影响。例如文献[9]提出了一种差别意识联邦学习来解决跨用户终端医疗图像数据异构问题，主要利用循环生成对抗网络合成图像，合成图像在保留原始图片特征的同时与原始图片有一定区分度，最终将每个用户终端上的原始数据转换到同一图像空间上，从而可安全有效地减少来自不同客户端图像之间的差异。文献[10]提出一种基于群体的联邦学习算法，该算法将所有用户终端的训练数据分类到多个群体，再用每个数据群体训练一个模型，优化了预测死亡率和ICU住院时间的准确率。上述方法在一定程度上缓解了数据异构造成的负面影响，但是在对原始数据预处理的过程中，会对用户数据隐私造成威胁。

除了对训练数据进行预处理外，部分研究者在进行模型聚合之前对用户终端进行聚类，充分利用相似用户终端之间的数据，提升模型性能。例如，文献[11]提出聚类联邦学习算法(Clustered Federated Learning, CFL)，利用联邦学习损失的几何特性判断用户终端的数据分布相似性，将用户终端聚类成多个具有联合可训练数据分布的聚类域，以充分利用相似用户终端之间的数据。实验表明，CFL可以在分类准确度方面取得比传统FL更大的改善。类似地，文献[12]在联邦学习中引入层次聚类，根据用户终端的相似性划分用户终端集群，并对共同的全局模型进行更新，该算法比FL具有更快的收敛速度和更高的准确率。

(2)基于聚合方式优化的方法。在解决由于数据异构带来的问题中，除了上述在聚合操作之前进行相应的预处理外，部分研究主要对联邦学习中的聚合方式进行优化[13–16]。例如Ek等人[13]在联邦学习聚合阶段，通过欧几里得距离计算用户终端模型中发散的神经元，这些发散的神经元被作为新的神经元添加到聚合模型中，该方法灵活地调整模型架构以适应任务，有效地改善了数据异构带来的影响。文献[14]提出了一种新的联邦学习聚合方法，该方法保持局部模型批量归一化参数不与全局模型同步，从而减轻非独立同分布数据中的特征转移，有效改善了数据异构下模型的收敛速度。

用户终端之间数据的非独立同分布，会导致一些用户终端仅根据本地数据训练的模型比采用联邦学习后的表现更好，为了解决上述问题，部分研究者提出了一些个性化的联邦学习模型，促使全局联邦模型适应于各个分布式用户终端。例如文献[17]通过一种关注消息传递的机制，在云服务器上为每个用户终端维护一个个性化的云模型，显著提高了联邦协作的有效性。类似地，文献[18]通过计算用户终端可以从另一个用户终端模型中获益的多少，为每个用户终端计算最优加权组合，从而针对用户终端定制更好的模型。上述个性化联邦学习方案主要以协作的方式先构建一个全局模型，然后利用用户终端的私有数据为每个用户终端定制个性化模型，但是在数据异构的情况下直接进行聚合会降低全局模型性能。

基于上述挑战，本文提出了一种基于谱聚类的傅里叶个性化联邦学习算法，主要贡献总结如下：

(1)构建了一个面向图像分类识别的云边端协

同个性化联邦学习模型，提出了在云端协同下利用知识迁移对用户终端进行相似性判断，并通过谱聚类将用户终端划分为多个聚类域，使用户终端在数据异构的情况下可以充分利用相似用户终端学到的知识，提升模型性能。

(2)设计了边端协同的局部联邦学习方法，即在用户终端和相应的边缘节点之间进行局部的联邦学习，并在此基础上使用代理模型策略，在用户终端对个性化局部模型执行恢复与再更新的过程，恢复聚合过程中丢失的本地知识，可提高下一次迭代更新的效果。

(3)提出了云服务器与边缘节点协同训练的傅里叶个性化联邦学习方法，即云服务器通过傅里叶变换将局部模型参数转换到频域空间上，保留局部模型的高频部分，聚合低频部分，为每个边缘节点定制高质量的个性化局部模型，使全局模型更适用于各个分布式用户终端，提升模型的准确率和收敛速度。

通过大量的仿真与分析，验证了本文提出的算法能较好地解决数据异构带来的负面影响，并与其他经典的联邦学习算法对比，本文算法在准确率和收敛速度上有较大的性能优势。

2　系统模型

本节构建了一个面向图像分类识别的云边端协同个性化联邦学习模型，如图1所示，该模型由用户层、边缘层和云层组成，图1呈现了这些层之间关系，具体每层的功能定义如下：

(1)用户层。此层由多个用户终端组成，如各医疗机构的专用计算机。每个用户终端在本地存储私有的图像数据，由于隐私保护的限制，用户终端之间无法直接进行数据交流。用户终端被聚类为若干个聚类域，每个用户终端主要有两个功能。(a)上传本地识别模型：用户终端使用本地私有的数据对本地分类识别模型进行训练，本地模型训练结束后，在同一个聚类域内的用户终端将其本地识别模型上传到同一边缘节点进行聚合。(b)下载个性化局部模型：用户终端下载边缘节点接收到的来自云层的个性化局部模型至本地，利用代理模型对本地知识进行恢复，再对本地识别模型进行更新。

(2)边缘层。此层由多个边缘节点组成，云层对用户终端划分聚类域后，为每个聚类域内分配一个边缘节点。每个边缘节点主要包含两个功能。(a)上传局部模型：边缘节点对接收到的来自其覆盖范围的用户终端发送的本地模型进行聚合，然后将边缘节点内聚合后的局部模型上传到云层进行聚合。(b)下载个性化局部模型：边缘节点下载云层定制的个性化局部模型，并将个性化局部模型发送给其覆盖范围内的用户终端进行模型更新。

(3)云层。云层由具有强大计算能力的云服务器构成，云服务器主要包括两个功能。(a)划分聚类域：云服务器对用户终端之间的相似性进行评价并将它们聚类为若干个聚类域。(b)定制个性化模型：对所有边缘节点上传的局部模型进行聚合，并为每个边缘节点定制相应的个性化局部模型，然后发送回边缘节点。

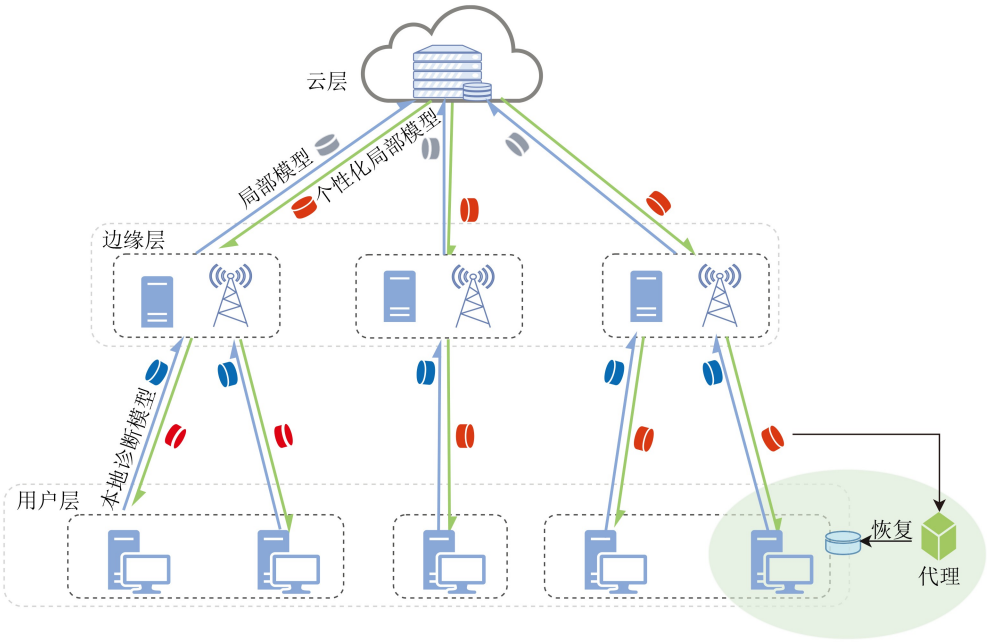


图 1 云边端协同个性化联邦学习模型

3 个性化联邦学习

本文提出了一种基于谱聚类的傅里叶个性化联邦学习算法，即为解决用户终端之间数据非独立同分布时联邦学习产生次优结果的问题，利用知识迁移对用户终端之间数据分布的相似性进行评价，并将用户终端聚类为多个可联合训练的聚类域。同时，为了保证用户终端模型的个性化并增强模型性能，基于云边端协同考量，在云服务器中采用傅里叶聚合方式为每个边缘节点定制个性化的局部模型，并在用户终端对从边缘节点中下载的个性化局部模型进行本地知识恢复，最终使得每个用户终端都能得到一个高质量的个性化模型。

3.1 个性化联邦学习框架

假设共有N个用户终端，i表示第i个用户终端，即$i\in\{1,2,\cdots,N\}$。每个用户终端i都拥有一组相同的无标签公共数据集$D_p=\{b_r\,|1\le r\le R\}$，其中$R$为公共数据集中样本的个数，以及本地私有的有标签数据集$D_i=\{x_{ij},y_{ij}\,|1\le j\le I_i\}$，其中$I_i$表示$D_i$中样本的个数，$x_{ij}$,$y_{ij}$分别表示$D_i$中第$j$个样本的数据与标签。本文提出的个性化联邦学习框架如图2所示。

为了解决用户终端之间数据非独立同分布时联邦学习产生次优结果的问题，在图2框架中，用户终端i首先从云服务器下载初始训练模型f到本地，然后利用本地私有数据D_i对本地模型进行训练更新，训练结束后，每个用户终端对公共无标签数据集D_p中的样本进行结果预测，再将预测结果发送到云服务器进行用户终端之间相似度计算，然后将用户终端聚类为若干个可联合训练的聚类域。同一个聚类域内的用户终端将本地模型上传给同一边缘节点进行聚合。

为了保证用户终端模型的个性化并增强模型性能，边缘节点将局部模型上传到云服务器后，云服务器使用傅里叶聚合为每个边缘节点定制个性化的局部模型，边缘节点接收到个性化局部模型后发送给用户终端的代理模型进行本地知识的恢复，最后更新本地模型。详细的训练过程设计将在3.2节进行描述。

3.2 基于谱聚类的傅里叶个性化联邦学习

本节提出的个性化联邦学习算法主要包括3个部分内容：云端协同的聚类域划分、边端协同的局部联邦学习和云边协同的傅里叶个性化联邦学习。

(1)云端协同的聚类域划分。在这一阶段，用户终端i从云服务器下载训练模型f到本地，然后利用本地私有的训练数据集D_i训练更新本地模型f_i。

接着，本地模型f_i对公共无标签数据集D_p中的样本进行结果预测，f_i对D_p中每一个样本b_r的预测结果输出为一个$E\times 1$的向量$\boldsymbol{p}_{ir}$，其中E为公共数据集D_p的样本类别总数，则用户终端i的本地模型f_i对公共数据集D_p的输出矩阵为$\boldsymbol{P}_i=(\boldsymbol{p}_{i1},\boldsymbol{p}_{i2},\cdots,\boldsymbol{p}_{iR})$，即本文将本地模型$f_i$学习到的知识，迁移表现为其

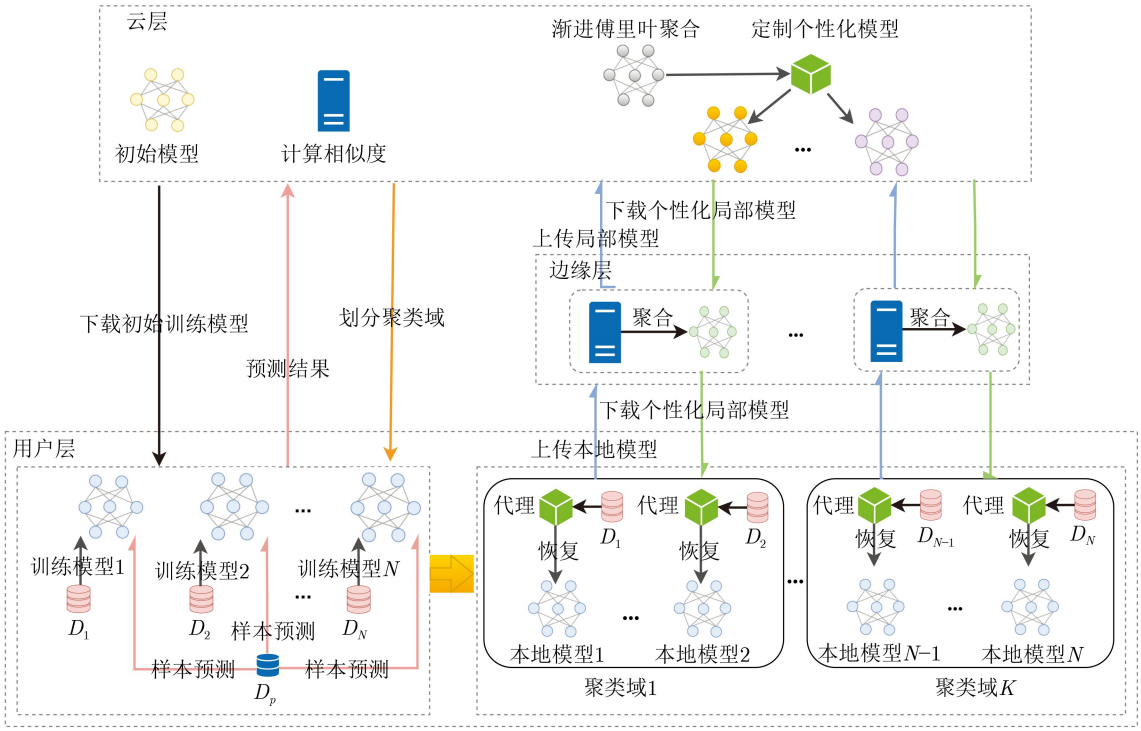


图 2 个性化联邦学习框架

对公共数据集D_p的预测结果。用户终端i将矩阵$\boldsymbol{P}_i$上传至云服务器，云服务器利用余弦相似度计算用户终端i和$j(1 \le i,j \le N)$之间的相似度S_{ij}，并构造对应的相似矩阵$\boldsymbol{S}_{N\times N}$。相似度$S_{ij}$的计算公式为

$$S_{ij}=\sum_{r=1}^{R}\frac{\boldsymbol{p}_{ir}\cdot\boldsymbol{p}_{jr}}{\|\boldsymbol{p}_{ir}\|\cdot\|\boldsymbol{p}_{jr}\|} \tag{1}$$

可以进一步计算得到度矩阵$\boldsymbol{Q}_{N\times N}$，其中

$$Q_{ij}=\begin{cases}\sum\limits_{j=1}^{N}S_{ij}, i=j\\ 0, i\ne j\end{cases} \tag{2}$$

由相似矩阵$\boldsymbol{S}_{N\times N}$和度矩阵$\boldsymbol{Q}_{N\times N}$可以进一步得到拉普拉斯矩阵$\boldsymbol{L}_{N\times N}$

$$\boldsymbol{L}_{N\times N}=\boldsymbol{Q}_{N\times N}-\boldsymbol{S}_{N\times N} \tag{3}$$

再将拉普拉斯矩阵$\boldsymbol{L}_{N\times N}$标准化为

$$\boldsymbol{L}'=\boldsymbol{Q}_{N\times N}^{-\frac{1}{2}}\boldsymbol{L}_{N\times N}\boldsymbol{Q}_{N\times N}^{-\frac{1}{2}} \tag{4}$$

假设聚类样本为c维，则选取矩阵$\boldsymbol{L}'$中c个最小的特征值对应的特征向量组成矩阵并按行标准化为特征矩阵$\boldsymbol{H}_{N\times c}$。将特征矩阵$\boldsymbol{H}_{N\times c}$中每一行视为一个$c$维的样本，则共有$N$个样本$a_1,a_2,\cdots,a_N$，组成样本集$A=\{a_1,a_2,\cdots,a_N\}$。

从样本集A中预选k_0个样本作为初始聚类中心$Z=\{z_1,z_2,\cdots,z_{k_0}\}$，则此时聚类域的个数$K=k_0$。计算样本集$A$中样本$a_i$到各个聚类中心的距离，并将$a_i$划分到距离其最小的聚类中心所对应的聚类域中，即为

$$a_i\in\text{cluster}_k, k=\mathop{\arg\min}_{k\in\{1,2,\cdots,k_0\}}\|a_i-z_k\| \tag{5}$$

其中，cluster_k表示第k个聚类中心对应的聚类域。

聚类域初步划分好后，若有聚类域中元素数目小于下限$N_{\min}$则丢弃该聚类域，并将该聚类域中的样本重新分配给距离剩下聚类域中心最小的聚类域，同时$K=K-1$。进一步地，对聚类结果的合理性进行判断，并迭代执行合并或分裂操作，具体步骤如下：

(a)若当前$K\le k_0/2$，则说明当前聚类域太少，执行分裂操作。计算聚类域cluster_k中样本各维度方差中的最大值δ，若$\delta>\delta_{\max}$，其中$\delta_{\max}$为聚类域中方差的上限，且cluster_k中包含的样本数量$N_k\ge 2N_{\min}$，则将cluster_k分裂为两个子聚类域，两个子聚类域的中心分别为z_k^+和z_k^-

$$z_k^+=z_k+\delta;\ \ z_k^-=z_k-\delta \tag{6}$$

同时聚类域个数$K=K+1$；否则不进行分裂操作。

(b)若当前$K\ge 2k_0$，则说明当前聚类域太多，执行合并操作。计算当前所有聚类域中心的两两距离，用矩阵$\boldsymbol{M}_{K\times K}$表示。若$M_{ij}<M_{\min}$，其中$M_{\min}$表示聚类域中心点之间距离的下限，则对这两个聚类域$\text{cluster}_i$, cluster_j进行合并操作合并为一个新的聚类域$\text{cluster}_{\text{new}}$，且该聚类域的中心为

$$z_{\text{new}}=\frac{1}{N_i+N_j}(N_iz_i+N_jz_j) \tag{7}$$

其中，N_i,N_j分别表示聚类域cluster_i,cluster_j中样本点的个数，同时聚类域个数$K=K-1$；否则不进行合并操作。

根据上述流程将N个用户终端聚类成K个互不相交的聚类域

$$\bigcup_{k=1}^{K}\text{cluster}_k=\{1,2,\cdots,N\} \tag{8}$$

(2)边端协同的局部联邦学习。在上述云端协同的聚类域划分的基础上，采用Fedavg算法的思想，在同一个聚类域cluster_k内的用户终端$i\in\text{cluster}_k$将本地模型参数f_i上传给同一边缘节点B_k进行平均聚合操作，边缘节点B_k获得局部模型θ_k，即

$$\theta_k=\sum_{i\in\text{cluster}_k}\frac{I_i}{\sum\limits_{j\in\text{cluster}_k}I_j}f_i, k\in\{1,2,\cdots,K\} \tag{9}$$

其中，I_i表示第i个用户终端中数据样本的个数。

边缘节点B_k将θ_k上传至云服务器，经过云边协同的傅里叶个性化联邦学习为每个边缘节点定制个性化局部模型θ_k后，将边缘节点B_k上的局部模型发送到其覆盖范围内的用户终端。用户终端i接收到局部模型参数θ_k后，若直接更新本地模型可能会丢失本地模型学到的知识，降低下一次迭代的优化效率。因此，本文使用代理模型，将接收到的局部模型发送给本地代理向本地模型学习，恢复聚合过程中丢失的知识，再更新本地模型。

假设用户终端i上的代理模型为d_i，代理模型d_i更新为从对应边缘节点下载的局部模型θ_k。代理模型d_i对x_{ij}的预测概率分布为$q(x_{ij})$，本地模型f_i对x_{ij}的预测概率分布为$p(x_{ij})$。将两者预测概率分布之间的KL (Kullback-Leibler)散度作为代理模型d_i向本地模型f_i学习恢复本地知识的损失函数，表示为

$$L_{\text{KL}}=\sum_{j=1}^{I_i}p(x_{ij})\ln\frac{p(x_{ij})}{q(x_{ij})} \tag{10}$$

其中，x_{ij}为用户终端i本地私有数据集D_i的第j个训练样本，I_i为D_i的样本总数。

同时，代理模型利用本地数据进行训练更新，交叉熵损失函数为

$$L_{\mathrm{CE}} = -\sum_{j=1}^{I_i} y_{ij} \ln \hat{y}_{ij} \tag{11}$$

其中，y_{ij}为用户终端i本地私有数据集D_i的第j个训练样本标签，$\hat{y}_{ij}$为预测结果。

则代理模型在训练过程中的总损失函数表示为

$$L_d = L_{\mathrm{CE}} + L_{\mathrm{KL}} \tag{12}$$

当代理模型d_i与本地模型f_i性能相近时，即$\phi_{\mathrm{val}}(d_i) \geq \lambda_1 \phi_{\mathrm{val}}(f_i)$，其中$\phi_{\mathrm{val}}$为对本地验证集的预测准确率，$\lambda_1$为超参数，用代理模型更新本地模型。

(3)云边协同的傅里叶个性化联邦学习。为了保证每个用户终端模型的个性化并增强模型性能，边缘节点B_k将局部模型参数θ_k上传给云服务器后，云服务器利用快速傅里叶变换将局部模型参数转换到频域空间上。文献[19]表明，经过快速傅里叶变换后，模型参数的低频部分表示模型的基础知识，则本文对局部模型参数的低频部分进行平均聚合，以共享来自不同聚类域之间的知识，对于高频部分即包含每个局部模型特定知识的部分进行保留，从而为每个边缘节点定制一个个性化的局部模型。

将局部模型θ_k的卷积层参数$\boldsymbol{w}_k \in R^{O\times C\times s_1\times s_2}$转换为2维矩阵$\boldsymbol{w}'_k \in R^{s_1O\times s_2C}$，$O$和$C$分别为输出通道数和输入通道数，$s_1$和$s_2$为卷积核的空间形状，对$\boldsymbol{w}'_k$使用快速傅里叶变换$F$，由式(13)、式(14)可以得到振幅图$F^{\mathrm{A}}$和相位图$F^{\mathrm{P}}$。

$$F(w'_k) = \sum_{x=0}^{s_1O-1}\sum_{y=0}^{s_2C-1} \boldsymbol{w}'_k(x,y)\mathrm{e}^{-\mathrm{j}2\pi(\frac{x}{s_1O}m+\frac{y}{s_2C}n)},\quad j^2=-1 \tag{13}$$

$$F(\boldsymbol{w}'_k) = F^{\mathrm{A}}\mathrm{e}^{\mathrm{j}F^{\mathrm{P}}} \tag{14}$$

其中，m和n为给定参数。

为了提取低频分量进行聚合，本文使用一个低频掩码G，除中心区域外值为0

$$G = \mathbb{Z}_{(m,n)\in[-gs_1O:gs_2O,-gs_2C:gs_2C]} \tag{15}$$

其中，$\mathbb{Z}$为指示函数，$g \in (0, 0.5)$表示低频阈值。

通过平均低频分量，保留高频分量，则第k个局部模型聚合后在频域空间上为$\hat{F}^{\mathrm{A}}(\boldsymbol{w}'_k)$

$$\hat{F}^{\mathrm{A}}(\boldsymbol{w}'_k) = (1-G)\circ F^{\mathrm{A}}(\boldsymbol{w}'_k) + \frac{1}{K}\sum_{k=1}^{K} G\circ F^{\mathrm{A}}(\boldsymbol{w}'_k) \tag{16}$$

最后利用傅里叶逆变换F^{-1}，将振幅映射和相位映射转换为参数形式$\hat{\boldsymbol{w}}_k$

$$\hat{\boldsymbol{w}}_k = F^{-1}([\hat{F}^{\mathrm{A}}(\boldsymbol{w}'_k), F^{\mathrm{P}}(\boldsymbol{w}'_k)]) \tag{17}$$

根据上述流程云服务器为每个边缘节点定制相应的个性化局部模型θ_k，边缘节点再将个性化局部模型θ_k发送给其覆盖范围内的用户终端，进行边端协同的局部联邦学习，最终得到个性化本地模型$f_i(1 \leq i \leq N)$。

为更好地理解本文提出的基于谱聚类的傅里叶个性化联邦学习，将上述过程简化凝练为算法1。

4 仿真与性能评估

本节通过仿真实验来评估基于谱聚类的傅里叶个性化联邦学习算法的有效性，并将本文所提出的方案与其他经典基准方案进行对比，以突出本文方案的性能优势。

本仿真样本集为EMNIST中By_Class 数据集，将该数据集中训练图像随机排序并随机选取3万张图像，其中前2万张作为训练数据集，后1万张作为测试数据集。

根据分布参数为$\alpha = 1$的狄利克雷分布将训练数据集非独立同分布地分配给N个用户终端，再通过图像旋转的方式模拟K个聚类域，即对用户终端图像进行K–1次旋转处理，第$r(1 \leq r \leq K-1)$次选取的用户终端为$\{i \mid \lfloor N/K \rfloor \times (r-1)+1 \leq \lfloor N/K \rfloor \times r,$

算法1 基于谱聚类的傅里叶个性化联邦学习算法

输入：$D_i(1 \leq i \leq N)$, D_p, f和通信轮数V.

输出：$f_i(1 \leq i \leq N)$.

1. BEGIN
2. 用户终端i从云服务器下载初始训练模型f到本地;
3. 用户终端i利用本地私有数据D_i训练更新本地模型f_i;
4. 用户终端i将D_p的预测结果矩阵$\boldsymbol{P}_i = (p_{i1}, p_{i2}, \cdots, p_{iR})$上传至云服务器;
5. 云服务器利用式(1)为用户终端计算相似矩阵$\boldsymbol{S}_{N\times N}$;
6. 云服务器利用式(2)—式(8)，通过谱聚类将N个用户终端聚类成K个互不相交的聚类域$\bigcup_{k=1}^{K}\mathrm{cluster}_k = \{1, 2, \cdots, N\}$;
7. 设置迭代次数$v = 1$;
8. 本轮迭代开始:
9. 用户终端$i \in \mathrm{cluster}_k$将本地模型参数f_i上传给同一边缘节点B_k，聚合为局部模型θ_k;
10. 边缘节点将局部模型θ_k上传给云服务器;
11. 云服务器利用式 (13)—式(17) ，采用傅里叶聚合为每个边缘节点B_k定制相应的个性化局部模型θ_k;
12. 用户终端i下载个性化局部模型θ_k到本地代理模型d_i，利用式(10)—式(12)恢复聚合过程中丢失的知识，将聚合到的知识传输到本地模型;
13. $v = v + 1$;
14. 若$v < V$，重复步骤7—步骤13，否则跳出迭代;
15. 得到个性化本地模型参数$f_i(1 \leq i \leq N)$.
16. END

$i \in \mathbb{N}^+\}$，旋转的度数为$(360°/K) \times r$。例如当N=10，K=2时，将训练数据集分配给10个用户终端后，对前5个用户终端上的图像旋转180°，则N个用户终端聚类为两个聚类域分别为$\text{cluster}_1 = \{i|1 \le i \le 5, i \in \mathbb{N}^+\}$和$\text{cluster}_2 = \{i|6 \le i \le 10, i \in \mathbb{N}^+\}$。

用户终端利用本地图像使用上文提出的基于谱聚类的傅里叶个性化联邦学习算法联合训练各自的5层卷积神经网络，包括两层卷积层、两层池化层和1层线性层。为了进一步地评判上述算法，本文采用准确率(Accuracy, Ac)和平均更新范数(Average Update Norm, AUN)作为评判算法有效性的标准。设T_i为用户终端i正确预测的样本个数，P_i为用户终端i的测试样本总数，准确率的计算为式(18)。设$\boldsymbol{dW}_i$为用户终端i在一轮通信中的参数更新向量，平均更新范数的计算为式(19)

$$\text{Ac} = \frac{1}{N}\sum_{i=1}^{N}\frac{T_i}{P_i} \tag{18}$$

$$\text{AUN} = \left\| \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{dW}_i \right\|_2 \tag{19}$$

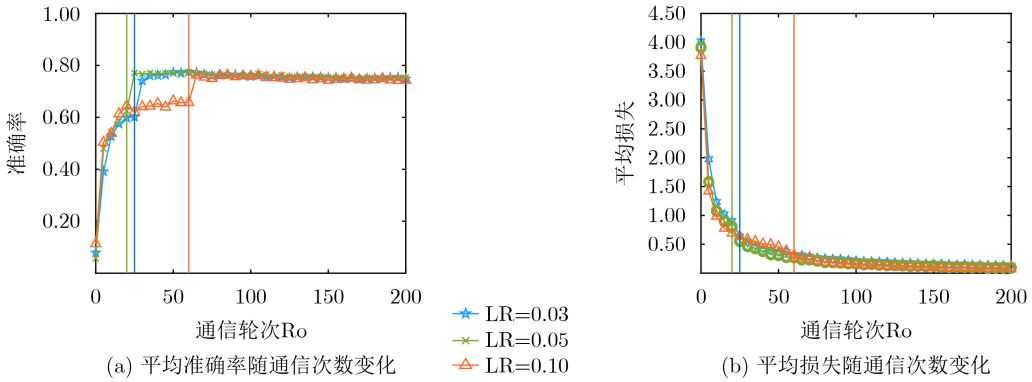
图3(a)为N=10个用户终端在聚类域个数K=2的设置下，基于谱聚类的傅里叶个性化联邦学习算法训练样本在不同学习率(Learning Rate, LR)下Ac随着通信轮数(Rounds, Ro)的变化曲线。从该图可以看出，在不同LR的设置下，随着Ro的增加，Ac不断增加并在100通信轮次内开始收敛趋向于同一值，表明本文训练模型设计的有效性。同时可以发现，Ro<20阶段LR越大曲线上升得越快，但最终LR=0.05却比LR=0.1收敛得更快，这是由于LR会影响聚类发生所在的通信轮次，当LR=0.05时，在Ro =20附近发生聚类，因此可以更有效地利用同一聚类域内不同用户终端的局部信息，从而使模型准确性迅速增长并更快地收敛，同样地LR=0.03时，Ro =25附近发生聚类，LR=0.1时，Ro =60附近发生聚类。

图3(b)为N=10个用户终端在聚类域个数K=2的设置下，基于谱聚类的傅里叶个性化联邦学习算法训练样本在不同LR下的平均损失(Loss)随着通信轮数的变化曲线。由该图可以看出在不同LR的设置下，曲线均在Ro<100内开始收敛并趋向0，表明本文的训练模型设计的有效性，且在趋向收敛前阶段斜率迅速变大，这是由于聚类的发生，使同一聚类域内用户终端学到的知识得到交互，模型因此降低损失更快地趋向收敛。

本文模型与现有的3种联邦学习模型：聚类联邦学习 (Clustered Federated Learning, CFL)[11]，PRR-FL (Personalized Retrogress Resilient-Federated Learning)[19]，联邦平均(Federated averaging, Fedavg)[3]进行对比。

图4绘制了当N=10时，Ac与聚类域个数K之间的关系。当K=1时所有用户终端视为一个聚类域，此时只进行局部联邦平均，并没有进行傅里叶个性化联邦学习，因此在该情况下本文算法与Fedavg和CFL的准确率相同。从整体来看，在聚类域个数K>1的情况下，本文提出算法的准确性比其他3种更高，凸显了本文算法的优越性。本文算法与CFL在K>1的情况下，准确性均优于其他两种方案，且随着K值的增大，用户终端数据之间异构性变大，对准确性的影响反而相对较小，其优势主要在于将用户终端根据相似度划分为多个聚类域，因此可以更有效地利用用户终端的局部信息。此外，本文算法融入傅里叶个性化模型定制模块以及本地代理恢复模型，使用户终端学习到其他用户终端知识的同时不丢失本地学到的知识，因此本文算法的准确性比CFL更高。

图5描绘了当K=2时，Ac与用户终端个数N之间的关系。从图中可以看出，本文提出算法的准确性比其他3种更高，进一步地体现了本文算法的优势。当聚类域个数K一定时，随着用户终端的个数N增加需要旋转的图像增多，则用户终端之间的异构性增加，从该图中可以发现，当异构性增加时单



(a) 平均准确率随通信次数变化　　(b) 平均损失随通信次数变化

图 3 收敛性分析

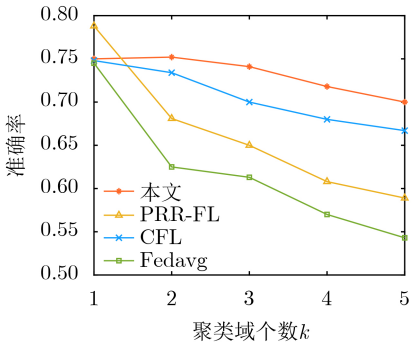


图 4 不同模型准确率与聚类域个数关系对比

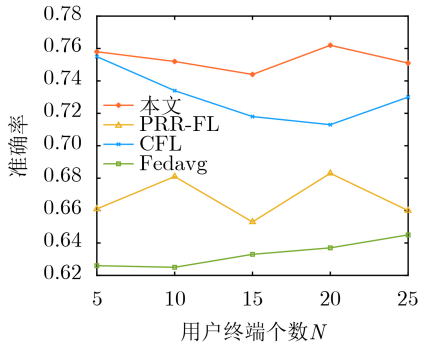


图 5 不同模型准确率与用户终端个数关系对比

独使用聚类算法的CFL与单独使用傅里叶个性化聚合算法的PRR-FL在准确率上的表现波动较大，而本文提出的将聚类与傅里叶个性化聚合相结合的基于谱聚类的傅里叶个性化联邦学习算法与经典算法Fedavg在面对用户终端数量增加时的稳定性更好，但总体准确率本文算法比Fedavg更高。

图6(a)、图6(b)分别描绘了聚类域个数为K=2和K=4时，AUN随着Ro的变化曲线。可以看出在Ro<20阶段，4种模型的AUN变化相似，但是当Ro=25， Ro=30时本文AUN突然增加，随后本文模型的AUN比其他对比模型更快收敛趋向于0，这是由于激增的点对应了发生聚类的通信轮次，聚类后使得用户终端更高效地利用同一聚类域内的局部信息。虽然CFL也有上述的聚类后的变化，但是显然本文模型比CFL收敛得更快，主要原因是本文加入傅里叶个性化模型的定制以及本地更新时使用代理模型对知识的恢复，使得各个用户终端上的模型快速训练至收敛。

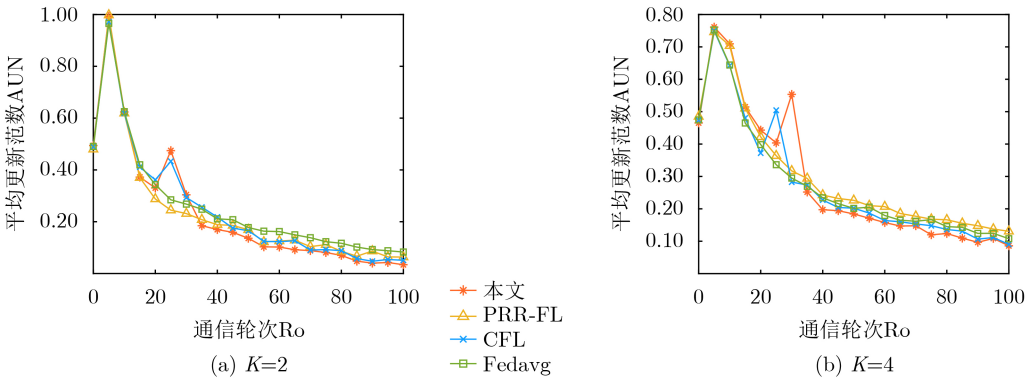


图 6 不同模型平均更新范数与通信轮数之间关系对比

5 结束语

为了降低数据异构对联邦学习的负面影响，本文提出了一种基于谱聚类的个性化联邦学习算法。具体地，构建了一个面向图像分类识别的云边端协同个性化联邦学习模型，提出了将用户终端划分为多个聚类域，使相似用户终端之间充分交互从而提升模型性能。进一步地，设计了边端协同的局部联邦学习方法，以提高迭代更新的效果，减少通信开销。此外，设计了云边协同的傅里叶个性化联邦学习方法，为每个边缘节点定制高质量的个性化局部模型。最后，与现有算法的对比结果表明，本文算法可以显著提高图像分类识别的准确性，且收敛速度更快。

参 考 文 献

[1] MOHAMMADI M, AL-FUQAHA A, GUIZANI M, *et al*. Semisupervised deep reinforcement learning in support of IoT and smart city services[J]. *IEEE Internet of Things Journal*, 2018, 5(2): 624–635. doi: 10.1109/jiot.2017.2712560.

[2] LI Chengxi, LI Gang, and VARSHNEY P K. Federated learning with soft clustering[J]. *IEEE Internet of Things Journal*, 2022, 9(10): 7773–7782. doi: 10.1109/JIOT.2021.3113927.

[3] MCMAHAN B, MOORE E, RAMAGE D, *et al*. Communication-efficient learning of deep networks from decentralized data[C]. The 20th International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, USA, 2017: 1273–1282.

[4] FEKI I, AMMAR S, KESSENTINI Y, *et al*. Federated learning for COVID-19 screening from chest X-ray images[J]. *Applied Soft Computing*, 2021, 106: 107330. doi: 10.1016/j.asoc.2021.107330.

[5] SILVA S, GUTMAN B A, ROMERO E, *et al*. Federated learning in distributed medical databases: Meta-analysis of large-scale subcortical brain data[C]. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI), Venice, Italy, 2019: 270–274. doi: 10.1109/ISBI.2019.8759317.

[6] LIU Yang, HUANG Anbu, LUO Yun, *et al*. FedVision: An online visual object detection platform powered by federated learning[C]. The AAAI Conference on Artificial Intelligence, New York, USA, 2020: 13172–13179. doi: 10.1609/aaai.v34i08.7021.

[7] ZHAO Zhongyuan, FENG Chenyuan, HONG Wei, *et al*. Federated learning with non-IID data in wireless networks[J]. *IEEE Transactions on Wireless Communications*, 2022, 21(3): 1927–1942. doi: 10.1109/TWC.2021.3108197.

[8] KULKARNI V, KULKARNI M, and PANT A. Survey of personalization techniques for federated learning[C]. The 4th World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), London, UK, 2020: 794–797. doi: 10.1109/WorldS450073.2020.9210355.

[9] YAN Zengqiang, WICAKSANA J, WANG Zhiwei, *et al*. Variation-aware federated learning with multi-source decentralized medical image data[J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 25(7): 2615–2628. doi: 10.1109/JBHI.2020.3040015.

[10] HUANG Li, SHEA A L, QIAN Huining, *et al*. Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records[J]. *Journal of Biomedical Informatics*, 2019, 99: 103291. doi: 10.1016/j.jbi.2019.103291.

[11] SATTLER F, MÜLLER K R, and SAMEK W. Clustered federated learning: Model-agnostic distributed multitask optimization under privacy constraints[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(8): 3710–3722. doi: 10.1109/TNNLS.2020.3015958.

[12] BRIGGS C, FAN Zhong, and ANDRAS P. Federated learning with hierarchical clustering of local updates to improve training on non-IID data[C]. 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 2020: 1–9. doi: 10.1109/IJCNN48605.2020.9207469.

[13] EK S, PORTET F, LALANDA P, *et al*. Artifact: A federated learning aggregation algorithm for pervasive computing: Evaluation and comparison[C]. 2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), Kassel, Germany, 2021: 448–449. doi: 10.1109/PerComWorkshops51409.2021.9431080.

[14] LI Xiaoxiao, JIANG Meirui, ZHANG Xiaofei, *et al*. FedBN: Federated learning on non-IID features via local batch normalization[C]. 9th International Conference on Learning Representations (ICLR), Vienna, Austria, 2021: 1–27.

[15] CHEN Hongyou and CHAO Weilun. FedBE: Making Bayesian model ensemble applicable to federated learning[C]. 9th International Conference on Learning Representations (ICLR), Vienna, Austria, 2020: 1–21.

[16] YE Dongdong, YU Rong, PAN Miao, *et al*. Federated learning in vehicular edge computing: A selective model aggregation approach[J]. *IEEE Access*, 2020, 8: 23920–23935. doi: 10.1109/ACCESS.2020.2968399.

[17] HUANG Yutao, CHU Lingyang, ZHOU Zirui, *et al*. Personalized cross-silo federated learning on non-IID data[J/OL]. The AAAI Conference on Artificial Intelligence, 2021: 7865–7873. doi: 10.1609/aaai.v35i9.16960.

[18] ZHANG M, SAPRA K, FIDLER S, *et al*. Personalized federated learning with first order model optimization[C]. 9th International Conference on Learning Representations (ICLR), Vienna, Austria, 2020: 1–17.

[19] CHEN Zhen, ZHU Meilu, YANG Chen, *et al*. Personalized retrogress-resilient framework for real-world medical federated learning[C]. The 24th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Strasbourg, France, 2021: 347–356. doi: 10.1007/978-3-030-87199-4_33.

金　彤：女，博士生，研究方向为联邦学习与边缘智能等.
陈思光：男，博士，教授，研究方向为边缘智能、智慧物联网等.

责任编辑：马秀强