计 算 机 研 究 与 发 展
Journal of Computer Research and Development
DOI: 10.7544/issn1000-1239.202330885
61(9): 2170−2184, 2024

面向非独立同分布数据的车联网多阶段联邦学习机制

唐晓岚　梁煜婷　陈文龙
（首都师范大学信息工程学院　北京　100048）
（tangxl@cnu.edu.cn）

Multi-Stage Federated Learning Mechanism with non-IID Data in Internet of Vehicles

Tang Xiaolan, Liang Yuting, and Chen Wenlong
（*College of Information Engineering*, *Capital Normal University*, *Beijing* 100048）

Abstract　The Internet of vehicles (IoV) plays an indispensable role in the construction of smart cities, where cars are not just a means of transportation but also a crucial medium for information collection and transmission in the era of big data. With the rapid growth in the volume of data collected from vehicles and the increased awareness of privacy protection, ensuring users' data security and preventing data breaches in IoV have become an urgent issue to address. Federated learning, as a "data-does-not-move, model-moves" approach, offers a feasible method for protecting user privacy while achieving excellent performance. However, because of the differences of devices, regions and individual habits, data collected from multiple vehicle typically exhibit non-independent and identically distributed (non-IID) characteristics. Traditional federated learning algorithms have slow model convergence when processing non-IID data. In response to this challenge, we propose a multi-stage federated learning algorithm with non-IID data in IoV, named FedWO. In Stage 1, FedWO utilizes the federated averaging algorithm to expedite the global model in reaching a basic level of accuracy. In Stage 2, FedWO employs weighted federated learning, where the weight of a vehicle in the global model is calculated based on its data characteristics. This aggregation results in an improved global model. Moreover, we design a transmission control strategy to reduce communication overhead caused by model transmission. The Stage 3 involves personalized computation, where each vehicle employs its own data for personalized learning, fine-tuning the local model to obtain a model more aligned with local data. We conducted experimental evaluations using a driving behavior dataset. The results demonstrate that, compared with traditional methods, FedWO preserves data privacy while improving the accuracy of algorithms in non-IID data scenarios.

Key words　Internet of vehicles (IoV); federated learning (FL); non-IID data; privacy preserving; transmission control

摘　要　车联网在智慧城市建设中扮演着不可或缺的角色，汽车不仅仅是交通工具，更是大数据时代信息采集和传输的重要载体. 随着车辆采集的数据量飞速增长和人们隐私保护意识的增强，如何在车联网环境中确保用户数据安全，防止数据泄露，成为亟待解决的难题. 联邦学习采用"数据不动模型动"的方式，

收稿日期：2023-11-01；**修回日期**：2024-05-15
基金项目：国家自然科学基金项目(61872252)；北京市优秀青年人才培育计划项目(BPHR202203118)；首都师范大学"人工智能赋能首都教育改革与发展"科研项目(RGZNJY2023-YB-14)
This work was supported by the National Natural Science Foundation of China (61872252), the Beijing Outstanding Youth Talent Development Program (BPHR202203118), and the Research Project on "Artificial Intelligence Empowering Capital Education Reform and Development" at Capital Normal University (RGZNJY2023-YB-14).
通信作者：陈文龙(chenwenlong@cnu.edu.cn)

为保护用户隐私和实现良好性能提供了可行方案. 然而，受限于采集设备、地域环境、个人习惯的差异，多台车辆采集的数据通常表现为非独立同分布（non-independent and identically distributed, non-IID）数据，而传统的联邦学习算法在 non-IID 数据环境中，其模型收敛速度较慢. 针对这一挑战，提出了一种面向 non-IID 数据的车联网多阶段联邦学习机制，称为 FedWO. 第 1 阶段采用联邦平均算法，使得全局模型快速达到一个基本的模型准确度；第 2 阶段采用联邦加权多方计算，依据各车辆的数据特性计算其在全局模型中的权重，聚合后得到性能更优的全局模型，同时采用传输控制策略，减少模型传输带来的通信开销；第 3 阶段为个性化计算阶段，车辆利用各自的数据进行个性化学习，微调本地模型获得与本地数据更匹配的模型. 实验采用了驾驶行为数据集进行实验评估，结果表明相较于传统方法，在 non-IID 数据场景下，FedWO 机制保护了数据隐私，同时提高了算法的准确度.

关键词 车联网；联邦学习；非独立同分布数据；隐私保护；传输控制

中图法分类号 TP393

近年来，智能网联汽车和无人驾驶技术得到了飞速发展，新型的智慧车辆已配备了高效的通信和计算设备. 在此背景下，智能汽车不仅仅是传统意义上的交通工具，更成为一个复杂的移动计算设备，在交通场景中担当着数据采集、处理和通信的重要角色. 这种以车辆为节点构建起来的巨大网络系统，被称为车联网[1]（Internet of vehicles, IoV）. 车联网已经不局限于车辆之间的信息传输，还能与交通基础设施、行人、网络服务等进行数据交换，实现资源共享和智能协同，从而提升出行的安全、效率和舒适性. 然而，随着车联网在数据收集、传输和处理方面的作用日益增强，数据安全和隐私保护问题也随之凸显. 汽车在行驶过程中会收集各种类型的敏感数据，例如车辆轨迹、驾驶员的驾驶行为、车内和车外环境信息等，这对车联网安全性提出了前所未有的挑战[2].

在以车辆节点为边缘设备的计算场景中[3]，联邦学习为车辆的隐私保护提供了新的解决方案. 不同于传统的集中式学习方法，联邦学习无需将大量数据上传到云端，而是通过移动的边缘设备在本地进行机器学习模型训练，再上传模型到云端完成全局聚合. 联邦学习减少了对中央服务器的依赖，并避免了大量数据传输的需要，减轻了网络带宽压力，还能够降低潜在的延迟问题. 更重要的是，联邦学习能够有效地解决“数据孤岛”问题[4]，各个参与方在不直接共享原始数据的情况下，协同建立机器学习模型. 这种分布式学习结构将每个参与方的机器学习能力与对集中存储所有数据的需求分开，旨在不泄露个人或敏感信息的前提下提升模型的整体性能. 因此，联邦学习不仅提高了数据处理和机器学习的效率，还为保护用户隐私和数据安全提供了有效的途径.

然而，在实际应用中，联邦学习中多个参与方所拥有的数据往往是具有异构性的，称为非独立同分布（non-independent and identically distributed, non-IID）数据[5]. non-IID 数据的存在意味着各个车辆的数据可能在分布、量级，甚至质量上具有差异，这对联邦学习构成了新的挑战，其模型聚合过程可能需要更多的通信轮次和迭代步骤才能实现收敛，并且其模型的性能也会由于数据的不均衡性受到一定程度的影响，这是因为每个节点训练的模型参数反映的是其本地数据的特性，这些本地模型在全局聚合时可能会因为数据分布的差异性而导致不一致，从而影响最终模型的性能. 特别是在某些极端情况下，如某些参与节点的数据量极少或数据质量不佳，这些因素都可能导致整体模型表现下降. 因此，针对非独立同分布数据，在确保用户隐私安全的前提下，如何在有限的通信轮次和较短的时间内提高模型的收敛效率，是一个亟待解决的问题.

本文提出了面向 non-IID 数据的车联网多阶段联邦学习机制，称为 FedWO. 在第 1 阶段实现联邦平均多方计算，各车辆开始收集数据，数据量较小，服务器使用联邦平均（federated averaging, FedAvg）算法进行模型聚合，使全局模型快速达到一个较稳定的状态. 第 2 阶段是联邦加权多方计算，考虑到各车辆拥有的不同数据特性，依据模型准确度、数据丰富程度和数据量为各车辆的本地模型计算权重，实现联邦加权的模型聚合. 与此同时，设计了传输控制策略，选择部分车辆来上传本地模型和下载全局模型，从而降低模型传输的通信开销. 第 3 阶段为个性化计算，车辆不再与服务器通信，各个车辆依据本地的数据集微调模型参数，使模型达到更高的准确度.

本文的贡献主要有 3 个方面：

1）提出了 3 阶段联邦学习机制，包括联邦平均多

方计算、联邦加权多方计算和个性化计算，充分考虑了收集过程中数据从无到有且各车辆数据分布不同的特点，解决了联邦学习面对 non-IID 数据时模型难收敛的问题.

2)针对模型传输开销大的问题，设计了模型上传和下载的参与车辆选择方案，与上一轮全局模型相近的本地模型不再上传，在全局聚合中占比大的车辆不再下发全局模型，从而降低通信和计算开销.

3)采用真实的驾驶行为数据集开展大量的实验，结果表明多阶段联邦学习机制在保护了用户数据隐私的前提下提升了模型精度.

1 相关工作

联邦学习作为新兴的边缘机器学习方法，已经被应用于金融、医学、计算机科学等领域，在保护用户数据隐私的同时，解决了“数据孤岛”的问题[6]. 在商业环境下，数据泄露会给供应商带来严重的经济损失，Lu 等人[7] 提出了一种安全数据共享架构，该架构基于区块链授权，采用了联邦学习中的隐私保护技术，将数据共享问题形式化为机器学习任务. 该架构虽对数据进行了隐私保护，但在如何提升参与方模型准确度方面的研究还不充分. 在联邦学习中，网络延迟和通信成本等问题也同样具有挑战. Luo 等人[8] 提出了一种低成本的采样算法来减少迭代次数，在保证模型收敛的前提下，实现成本最小化. He 等人[9] 将联邦学习应用于计算机视觉任务中，提出了一种 FedCV 联邦学习库和基准测试框架，以评估联邦学习在图像分类、图像分割以及目标检测任务中的表现. 由于多模态模型的盛行，相关研究和应用在学术界和工业界均得到了广泛关注. 文献 [10] 提出了一种联邦学习中基于 Tucker 分解的多源异构数据融合方法，该方法通过建立一个融合了异构空间维度特性的高阶张量，有效捕获异构数据中的高维特征，进而实现在联邦学习场景下多源异构数据的高效整合.

近年来，一些工作以车辆为边缘计算节点，通过联邦学习实现车联网的协同决策. Kong 等人[11] 将联邦学习应用于移动设备中，提出了一种车牌识别框架 FedLPR，提高了车牌的检测准确度，同时拥有可接受的通信成本. Liang 等人[12] 提出了一种半同步联邦学习协议 Semi-SynFed，根据车辆节点的计算能力，动态地调整服务器的等待时间，以异步的方式进行全局聚合. 由于车联网在传输过程中的不可靠性，Lu 等人[13] 提出了一种基于区块链的联邦学习框架，将模型上传到区块链中并进行 2 阶段的验证，以保证共享数据的可靠性.

联邦学习在自动驾驶领域也有广泛的应用. Tang 等人[14] 提出了一种基于联邦强化学习的驾驶控制算法 DFRL，在 Torcs 平台上进行了大量实验，证明该算法提高了驾驶控制的精度，但未详细考虑车辆计算资源有限的问题. Parekh 等人[15] 提出了一种联邦学习的梯度加密算法，构建了一个德国交通标识识别系统，与传统的联邦学习相比，精度提高了 2%. 为了解决车联网中信息共享所带来的风险，Qu 等人[16] 提出了一种基于信息融合和个性化隐私的 PDP-PFL 算法，通过加入噪声以及轻量级的网络结构对局部模型进行微调，实现对数据隐私的保护. 为了解决车辆节点在一段时间内只能执行 1 个任务的问题，Li 等人[17] 设计了一种考虑车辆选择和无线通信资源分配的任务驱动的车辆联邦学习算法，提高了车联网中多任务联邦学习的效率.

车辆在行驶过程中通常会产生 non-IID 数据，导致全局模型难以收敛. 近年来，一些学者在研究面向 non-IID 数据的联邦学习机制时，通过应用聚类算法来提高模型的准确度；通过引入元学习技术，来实现个性化模型的训练；此外，还有通过优化模型权重分配机制，进一步提升模型的精度. 这些方法[18-26] 的具体介绍如表 1 所示.

综上所述，现有的技术主要解决独立同分布(IID)数据的联邦学习问题，而车联网中车辆收集的数据通常具有 non-IID 特性，针对 non-IID 数据的联邦学习解决方案还没有被充分挖掘，且在车辆计算资源有限的情况下，如何充分利用计算资源，提高通信效率是需要考虑的. 因此，本文旨在解决针对车联网 non-IID 数据的联邦学习中传输效率以及模型优化的问题.

2 相关概念

2.1 联邦学习

联邦学习[27] 由 Google 在 2016 年提出，旨在创建一个保护个人数据隐私的多方计算机器学习框架，同时确保数据交换过程中的信息安全. 联邦学习的特点是在数据本地存储的基础上，实现“模型共享，数据私有”，从而解决“数据孤岛”问题和数据隐私保护问题. 联邦学习过程中每一轮训练可分为 4 个主要阶段[28]：1)模型下载；2)本地训练；3)模型上传；4)全局模型聚合. 使用联邦学习的车联网场景如图 1 所

Table 1 Related Work on Federated Learning with non-IID Data

表 1 面向 non-IID 数据的联邦学习相关工作

类型	作者（年份）	算法	方法	数据集
聚类	He 等人[18] (2023)	ASCFL	采用基于相似度的聚类策略，选择客户端参与训练，在数据集准确性和收敛速度之间实现动态平衡.	①CIFAR-10 ②EMNIST
	Shu 等人[19] (2022)	FMTL	基于模型聚类实现对 non-IID 数据的多任务学习.	①MNIST ②CIFAR-10 ③Caltech-101
	Tian 等人[20] (2022)	WSCC	基于权重相似性以及亲和传播，实现动态聚类.	①MNIST ②CIFAR-10 ③IMDB Movie Review
元学习	Dong 等人[21] (2023)	PADP-FedMeta	通过自适应隐私参数，实现个性化、自适应差分隐私联邦元学习机制.	①MNIST ②Synthetic
	Yang 等人[22] (2023)	G-FML	根据客户数据分布的相似性，自适应地将客户分组，并在每个分组中使用元学习获得个性化模型.	①Synthetic ②FEMNIST ③Shakespeare
	Li 等人[23] (2022)	FML-ST	使用元学习的个性化联邦学习方法，通过评估全局和本地模式图的差异，使每个客户端能够定制其模型.	①Citi-Bike Dataset (NYC,DC,CHI)
权重	Hu 等人[24] (2024)	FedMMD	通过 DCMT（dilated convolution meet transformer）模型进行特征提取，并使用 SKNQ (student-keuls-newman-Q) 方法和熵权法确定模型全局聚合权重，提高全局模型的学习精度和泛化能力.	①MNIST ②FMNIST ③CIFAR-10 ④CIFAR-100
	Kim 等人[25] (2021)	FLC	通过分析机器学习模型各层的权重来对客户端进行聚类，并在聚类后的客户端中进行联邦学习.	①MNIST
	Zhang 等人[26] (2021)	CSFedAvg	利用权重差异识别客户数据的 non-IID 程度，选择 non-IID 数据程度较低的客户端，以更高的频率来训练模型.	①MNIST ②CIFAR-10

示. 下面将详细描述每个阶段的具体操作：

1）模型下载. 当前车辆在执行第 t 轮本地训练前，从服务器下载最新的全局模型，并加载到本地模型中. 服务器可以位于路边单元（road side unit, RSU）或云端，本文使用边缘节点 RSU 作为服务器，减少数据上传和下载的开销.

2）本地训练. 当前车辆基于已加载的本地模型，使用本地数据开展本地模型训练，从而更新模型参数.

3）模型上传. 当前车辆完成第 t 轮训练后，对本轮训练得到的本地模型参数进行加密，并上传到服务器.

4）全局模型聚合. 服务器接收到第 t 轮所有参与方的模型参数后，开展全局模型聚合，得到新的全局模型作为下一轮训练的初始参数.

在完成上述 4 个阶段后，标志着一轮联邦学习的结束. 通过不断迭代上述过程，实现模型的收敛.

2.2 non-IID 数据

在车联网的联邦学习应用中，每台车辆的行驶路线、每位驾驶员的操作习惯等均不相同，因此导致多台车辆的数据呈现出 non-IID 的特征. non-IID 数据主要包括 5 个种类[29]：

1）特征分布偏斜. 参与方的差异性导致数据的特征分布不均衡. 例如，在手写识别领域，用户写字的笔画粗细、倾斜度等可能会有很大不同.

2）标签分布偏斜. 不同参与方的标签比例是不均衡的. 例如，在某一车辆中，某些标签在数据集中占有较高的比例，但在另一车辆中，这些标签并不会出现.

3）相同标签，不同特征. 尽管不同参与方的样本携带相同的标签，但由于参与方所处的文化背景和环境差异，这些样本的特征表示可能会有显著的变化. 例如，“停车”标志在不同地区可能设计风格不一，尽管它们都传达相同的“停车”这一指令，但它们的外观和呈现的特征可能因地区而异，从而存在不同的特征表示.

4）相同特征，不同标签. 对于具有相似特征的样本，其标签可能因个体差异或地域差异而有所不同. 例如，驾驶员头部左右转动以便查看路口处的路况，

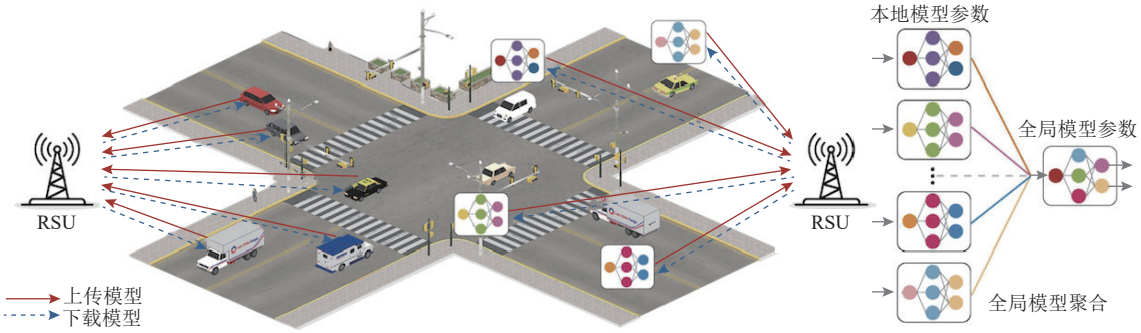


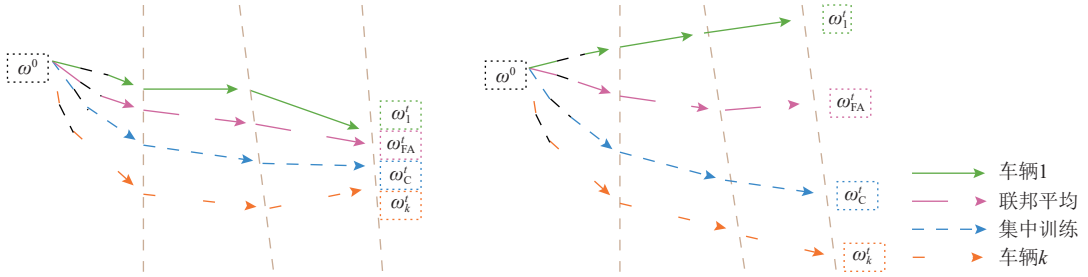
Fig. 1 Illustration of federated learning in IoV

图 1 车联网联邦学习示意图

相同的头部动作对于视力健全人士是安全的, 对于视力障碍人士可能不够安全.

5)数量偏斜或不平衡. 这表现在数据集的规模差异上, 通常是由于不同传感器收集的数据量不均衡等导致的. 样本数量较少的参与方在模型训练中可能面临偏差或过拟合的风险.

由于现实中的复杂情况, 各车辆的数据集通常表现出 non-IID 数据的特征. 联邦学习常用的全局模型聚合方法为 FedAvg[30] 算法. 在面对 IID 数据时, 该算法的梯度下降方向相对一致, 从而能较快地达到模型的收敛. 然而, 在面对 non-IID 数据时, 算法的梯度下降方向可能表现出发散的特征. 图 2 展示了这一情况. 因此, 在处理 non-IID 数据时, 联邦学习能否快速收敛是存在挑战的. 鉴于此, 本文提出了一种多阶段联邦学习机制 FedWO, 旨在解决面对 non-IID 数据时联邦学习的收敛问题.



(a) IID数据 (b) non-IID数据

Fig. 2 Model convergence trends for IID data and non-IID data

图 2 IID 数据和 non-IID 数据的模型收敛趋势

3 面向 non-IID 数据的多阶段联邦学习机制

本文提出车联网的多阶段联邦学习机制 FedWO, 该机制在确保车辆隐私的基础上, 针对 non-IID 数据, 旨在提升联邦学习方法的性能. 在本文讨论的车联网场景中, 车辆从零开始收集数据, 一边收集数据, 一边开展联邦学习, 实现全局模型的快速收敛, 最终车辆的本地模型获得良好的性能. 本文所使用的主要符号及其意义如表 2 所示.

如图 1 所示, 在 1 轮联邦学习过程中, 车辆需要与 RSU 通信, 实现模型下载和上传. 考虑到车辆处于快速移动过程中, 为了避免车辆在未完成 1 轮联邦学习时离开 RSU 通信范围造成的资源浪费, RSU 选择特定的车辆参与联邦学习, 即当车辆在目前所处 RSU 通信范围内停留时间大于完成模型下载、本地训练以及模型上传的总时间时, 即

$$T_k^{t,\text{stay}} > T_k^{t,\text{download}} + T_k^{t,\text{train}} + T_k^{t,\text{upload}}, \tag{1}$$

其中 $T_k^{t,\text{stay}} = \dfrac{L_k^t}{sp_k^t}$, 车辆 v_k 被选择参与联邦学习, 否则不参与本轮联邦学习.

本文在模型传输过程中, 采用本地差分隐私(local differential privacy, LDP)技术来实现隐私保护, 在数据集中输入任何一对 x, x', 若随机机制 $\mathcal{M}$ 对任何输出 Y 都满足 ϵ-LDP, 则称 $\mathcal{M}$ 符合 ϵ-LDP. 机制 $\mathcal{M}$ 的隐私保证由隐私预算决定[31], 用 ϵ 来表示, 那么本地模型的隐私保护需要满足的条件[32] 是

$$Pr[\mathcal{M}(x) = Y] \leqslant e^{\epsilon} Pr[\mathcal{M}(x') = Y], \tag{2}$$

其中 ϵ 的值越小, 其隐私保护的程度就越高.

如图 3 所示, 本文提出的多阶段联邦学习机制

Table 2 Main Symbols in Our Paper

表 2 本文主要符号

符号	描述
$\mathcal{V}$	参与联邦学习的车辆集合, 车辆 $v_k \in \mathcal{V}$
L_k^t	第 t 轮车辆 v_k 距离 RSU 覆盖边缘的距离
sp_k^t	第 t 轮车辆 v_k 的平均速度
$T_k^{t,\text{stay}}$	第 t 轮车辆 v_k 在 RSU 通信范围内停留时间
$T_k^{t,\text{download}}$	第 t 轮车辆 v_k 下载模型的传输时间
$T_k^{t,\text{train}}$	第 t 轮车辆 v_k 本地训练的时间
$T_k^{t,\text{upload}}$	第 t 轮车辆 v_k 上传模型的传输时间
$\mathcal{M}$	用于本地差分隐私的随机机制
ω^t	第 t 轮全局模型参数
ω_k^t	第 t 轮车辆 v_k 的本地模型参数
η	本地模型学习率
$\nabla F_k(\omega^t)$	第 t 轮车辆 v_k 的本地模型损失函数
A_k^t	第 t 轮车辆 v_k 的本地模型准确度
A	全部参与车辆 $\mathcal{V}$ 中最大的本地模型准确度
DS_k^t	第 t 轮车辆 v_k 本地训练的数据丰富程度
DS	全部参与车辆综合的数据丰富程度
DQ_k^t	第 t 轮车辆 v_k 本地训练的数据量
DQ	全部参与车辆的总数据量
$diff_k^t$	第 t 轮车辆 v_k 本地模型与全局模型的差值

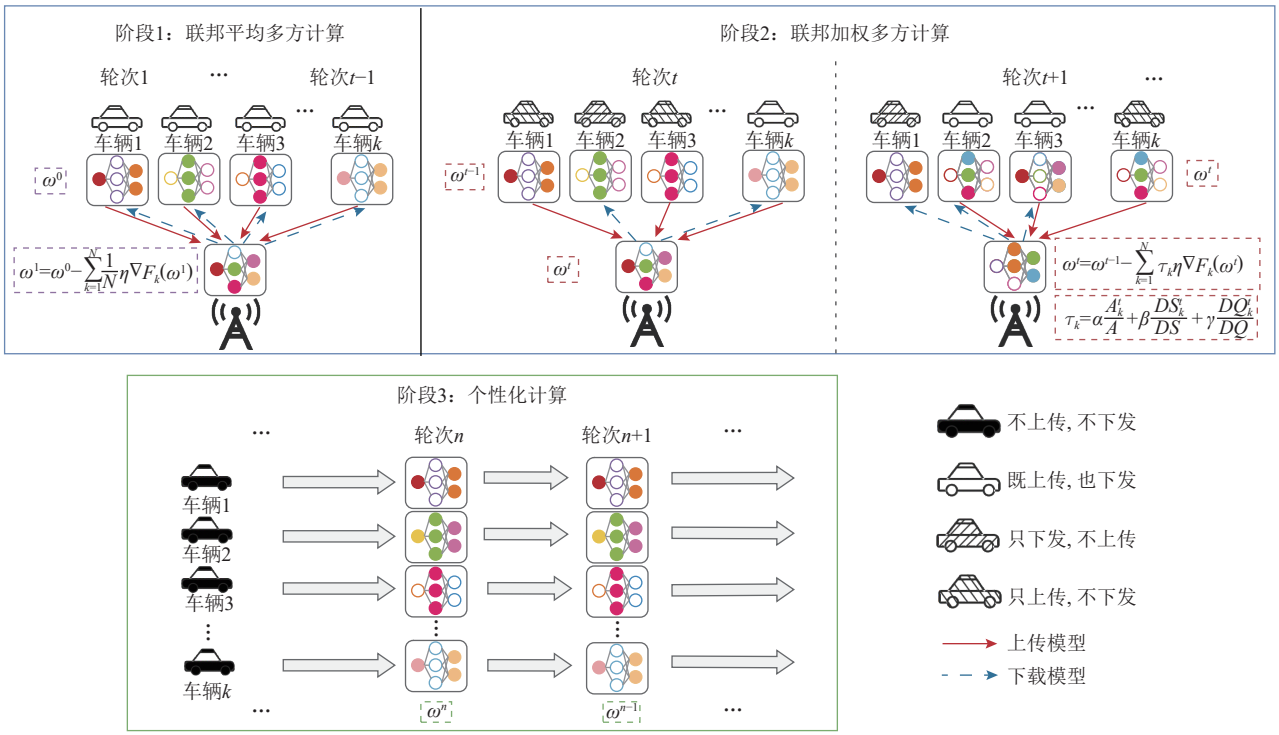


Fig. 3 Illustration of multi-stage federated learning

图 3 多阶段联邦学习示意图

FedWO 由 3 个阶段构成，即第 1 阶段联邦平均多方计算阶段、第 2 阶段联邦加权多方计算阶段、第 3 阶段个性化计算阶段，接下来将详细阐述各个阶段是如何工作的.

3.1 联邦平均多方计算阶段

在数据收集的起始阶段，当 RSU 尚未获得一个相对稳定的全局模型时，联邦学习需要大量的参与者来贡献数据. 在联邦学习中有 4 种较为典型的算法，分别是联邦平均(FedAvg)、联邦近似(federated proximal, FedProx[33])、基于随机控制平均的联邦学习(federated learning via stochastic controlled averaging, SCAFFOLD[34])和基于归一化平均的联邦学习(federated learning via normalized averaging, FedNova[35]). FedAvg 通过对各参与方的模型权重求平均来计算全局模型；FedProx 在 FedAvg 的基础上增加了一个正则化项，以减少由于数据异质性带来的不稳定性，但需要调整额外的超参数；SCAFFOLD 使用控制变量减少客户端与全局模型之间的差异，算法实现更复杂、计算和存储开销更大；FedNova 则利用 2 阶优化方法和自适应正则化技术提高收敛速度和性能，其实现相对复杂，需要精确控制不同客户端的更新步数.

在数据收集起始阶段，场景中的数据从无到有，逐渐积累. 虽然 FedProx, SCAFFOLD, FedNova 在处理数据异质性方面有其独特的优势，但此时的主要矛盾不是各参与方的数据异质性，而是尽快建立一个基本稳定且有效的全局模型. 此时，简单高效的 FedAvg 更为适合，它不涉及额外的控制信息或参数(如 FedProx 的正则化项或 SCAFFOLD 的控制变量)，对于初期的模型探索和快速部署来说，这种简单性是一个显著优势. 因此，本文第 1 阶段使用 FedAvg 来聚合全局模型，所有车辆的本地模型拥有相同的权重，公式如下：

$$\omega^t = \omega^{t-1} - \sum_{k=1}^{N} \frac{1}{N} \eta \nabla F_k\left(\omega^t\right), \tag{3}$$

其中ω^t是第 t 轮全局模型参数，ω^{t-1}是第 $t-1$ 轮全局模型参数，N 是参与联邦学习的车辆数，η是本地模型学习率，$\nabla F_k(\omega^t)$是第 t 轮车辆 v_k的本地模型损失函数.

服务器端的全局模型性能将被视为进入第 2 阶段的重要指标. 当服务器的模型精度趋向稳定时，本文的联邦学习机制 FedWO 进入第 2 阶段.

3.2 联邦加权多方计算阶段

考虑到车辆数据的 non-IID 特性，若持续使用 FedAvg 算法，将导致全局模型难以实现优化训练精度的目标. 为了解决这一问题，本文在第 2 阶段全局模型聚合时对不同车辆的权重进行重新分配. 权重的取值受到 3 个因素的影响：1)模型的准确度. 具有

较高模型准确度的车辆在全局模型聚合中将被分配更高的权重. 2)数据集的丰富程度. 数据集更为丰富的车辆在全局模型中的权重更大. 3)数据集的大小. 车辆所收集到的数据量越大，在全局模型中的权重越大.

综合上述 3 个因素，在全局聚合过程中，为不同车辆的本地模型分配不同的权重，从而优化全局模型的泛化能力. 其全局模型聚合公式为

$$\omega^t=\omega^{t-1}-\sum_{k=1}^{N}\left(\alpha\frac{A_k^t}{A}+\beta\frac{DS_k^t}{DS}+\gamma\frac{DQ_k^t}{DQ}\right)\eta\nabla F_k\left(\omega^t\right),\quad(4)$$

其中A_k^t是第 t 轮车辆v_k的本地模型准确度，A是全部参与车辆中最大的本地模型准确度，DS_k^t是第 t 轮车辆v_k本地训练的数据丰富程度，DS是全部车辆综合的数据丰富程度，DQ_k^t是第 t 轮车辆v_k本地训练的数据量，DQ是全部车辆的总数据量，权重$0\leqslant\alpha,\beta,\gamma\leqslant1$且$\alpha+\beta+\gamma=1$.

在联邦加权多方计算阶段，为了优化计算和通信资源的利用率，本文提出一种传输控制策略，通过选择参与联邦学习的车辆来减少传输开销. 选择联邦学习的参与方需要从 2 个维度进行评估：一是车辆本地模型的上传；二是 RSU 的模型下发. 因此，该传输控制策略的核心是本地模型上传和全局模型下载的参与车辆选择，下面分别进行详细讨论.

1)车辆本地模型的选择性上传

若第 t 轮的本地模型参数与第 $t-1$ 轮的全局聚合模型之间存在较大差异，则意味着该车辆在本轮训练中学到了新的知识，这部分知识应被纳入全局模型中，因此，该车辆需上传本地模型参数. 相反，若在第 t 轮迭代中，车辆v_k的本地模型参数与第 $t-1$ 轮的全局模型参数差异不显著，那么该车辆v_k将不上传本轮次的本地模型参数，从而节省通信开销和计算开销. 本文使用 L2 范数来计算 2 个模型的差异，即

$$diff_k^t=\sqrt{(\omega_k^t-\omega^{t-1})^2}.\quad(5)$$

2)服务器端全局模型的选择性下发

若车辆v_k在第 t 轮中上传了本地模型，并且其在全局模型聚合中的权重较大，即$\alpha\frac{A_k^t}{A}+\beta\frac{DS_k^t}{DS}+\gamma\frac{DQ_k^t}{DQ}>\varphi$($\varphi$为超参数)，则认为该轮的全局模型与车辆$v_k$的本地模型相似的可能性较大，那么 RSU 不向该车辆分发第 t 轮的全局聚合模型，从而节省传输开销.

在该传输控制策略中，只有对全局模型可能产生显著影响的车辆才会上传本地模型，只有新的全局模型与本地模型差异较大的车辆才会下载新的全局模型，从而避免不必要的数据传输，减少车辆和 RSU 之间模型传输带来的通信开销，同时降低 RSU 对全局模型聚合的计算开销，以及车辆更新本地模型的计算开销，提升资源利用率. 未来通过车辆聚类、异步联邦等技术，有望进一步降低通信成本.

在模型的选择性上传和下载的过程中，为保证本地模型与全局模型的一致性，设置 2 项规则：1)在每一轮迭代中，每辆车至少参与模型参数的上传或下载其中之一，不允许车辆在同一轮次中完全不与服务器交互. 2)若在第 t 轮中，RSU 没有向车辆v_k分发全局模型，那么在第 t+1 轮，该车辆必须主动参与本地模型参数的上传. 上述规则确保了联邦加权多方计算过程的平稳进行，同时提升了车辆和 RSU 之间的传输效率.

3.3 个性化计算阶段

针对 non-IID 数据，传统的联邦计算方法，如联邦平均计算和联邦加权计算，可能面临一些挑战. 尤其是在经过这些计算阶段后，通过全局聚合得到一个通用模型，该模型可能无法捕捉到某一特定数据源(如某一辆车)不同于其他数据源的独特数据特征. 这种损失可能导致无法进一步优化特定的本地模型，有时甚至可能导致模型性能的下降. 为了解决这一问题，本文在获得通用模型后，利用本地数据特征对其进行微调，从而更好地适应本地的数据分布. 这种基于本地特性的微调可以有效地增强模型的性能，特别是在处理那些与总体数据分布存在差异的本地数据时. 本文将这个基于本地数据特性微调的阶段称为“个性化计算”阶段. 这一阶段的目标是确保模型能够准确地捕捉并利用每一个数据源的独特信息，从而实现在本地数据上获得最佳性能的模型. 为了解决这一问题，当车辆v_k使用全局模型训练的性能有所下降时，本文建议车辆不再参与联邦学习而进入个性化计算阶段.

个性化学习阶段的车辆本地模型架构如图 4 所示，共有 5 个卷积块，每个卷积块由卷积层和最大池化层组成，卷积层的激活函数是 ReLU，在 5 个卷积块之后有一个全局平均池化层，随之相连的是全连接层，全连接层的激活函数采用 Softmax. Softmax 函数通常用于分类任务，将网络输出转换为概率分布. 模型微调策略只重新训练全连接层，实验中的模型设置见 4.2 节. 在前 2 个阶段的联邦学习中，车辆已经对模型进行多轮训练，卷积层已经学到了足够的特征表示，而全连接层更关注于如何将这些特征映射到最终的输出. 此阶段的模型微调只训练全连接层，不仅可以进一步提升模型准确度，而且可以减少

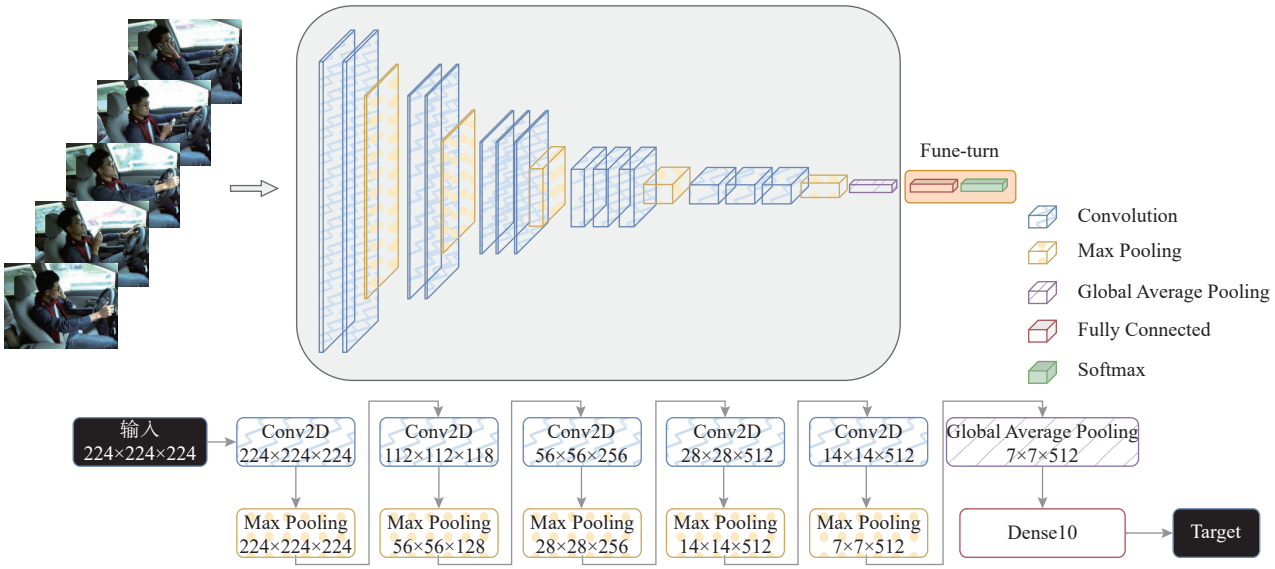


Fig. 4 Network fine-tuning structure diagram for the personalized computing phase

图 4 个性化计算阶段网络微调结构图

训练的参数, 节省计算资源和训练时间.

服务器和车辆的多阶段联邦学习算法分别如算法 1 和算法 2 所示.

算法 1. 服务器的多阶段联邦学习算法.

输入: 参与联邦学习的车辆集合$\mathcal{V}$, 训练轮次 T, 阈值φ;

输出: 全局模型参数ω^t.

① 初始化全局模型参数ω^0, $t=0$, $\mathcal{V}^*=\mathcal{V}$;

② while $t \leqslant T$ do

③ $t=t+1$;

④ if t 属于“阶段 1”do

⑤ 将ω^{t-1}下发给所有参与车辆$\mathcal{V}$;

⑥ 接收所有参与车辆的本地模型参数ω_k^t;

⑦ $\omega^t=\omega^{t-1}-\sum_{k=1}^{N}\frac{1}{N}\eta\nabla F_k\left(\omega^t\right)$; /*联邦平均*/

⑧ else if t 属于“阶段 2”do

⑨ 将ω^{t-1}下发给选择出的参与车辆 $\mathcal{V}^*$;

⑩ 接收参与车辆的本地模型参数ω_k^t;

⑪ $\omega^t=\omega^{t-1}-\sum_{k=1}^{N}\left(\alpha\frac{A_k^t}{A}+\beta\frac{DS_k^t}{DS}+\gamma\frac{DQ_k^t}{DQ}\right)\cdot \eta\nabla F_k(\omega^t)$; /*联邦加权*/

⑫ $\mathcal{V}^*=\mathcal{V}$;

⑬ for $\forall v_k\in\mathcal{V}$ do

⑭ if $\alpha\frac{A_k^t}{A}+\beta\frac{DS_k^t}{DS}+\gamma\frac{DQ_k^t}{DQ}>\varphi$ do

⑮ $\mathcal{V}^*=\mathcal{V}^*-\{v_k\}$; /*下发车辆*/

⑯ end if

⑰ end for

⑱ end if

⑲ end while

算法 2. 车辆的多阶段联邦学习算法.

输入: 参与联邦学习的车辆$v_k\in\mathcal{V}$, 训练轮次 T, 阈值δ, 学习率η;

输出: 本地模型参数ω_k^T.

① $t=0$;

② while $t \leqslant T$ do

③ $t=t+1$;

④ if t 属于“阶段 1”do

⑤ 接收服务器的全局模型参数ω^{t-1};

⑥ $\omega_k^t=\omega^{t-1}-\eta\nabla F_k(\omega^t)$; /*本地训练*/

⑦ 上传本地模型参数ω_k^t给服务器;

⑧ else if t 属于“阶段 2” do

⑨ if 服务器向v_k发送了全局模型参 数ω^{t-1} do

⑩ 接收服务器的全局模型参数 ω^{t-1};

⑪ $\omega_k^t=\omega^{t-1}-\eta\nabla F_k(\omega^t)$; /*本地训练*/

⑫ if $\sqrt{(\omega_k^t-\omega^{t-1})^2}>\delta$ do

⑬ 上传本地模型参数ω_k^t给服务器; /*上传车辆*/

⑭ end if

⑮ else

⑯ $\omega_k^t=\omega_k^{t-1}-\eta\nabla F_k(\omega^t)$; /*本地训练*/

⑰ 上传本地模型参数ω_k^t给服务器;

⑱ end if

⑲ else /* 阶段 3, 个性化计算 */

⑳ $\omega_k^t=\omega_k^{t-1}-\eta\nabla F_k\left(\omega_k^t\right)$;

㉑ end if

㉒ end while

㉓ 返回本地模型参数ω_k^T.

4 实验结果与分析

4.1 数据集

本文采用2016年在Kaggle上发布的公开数据集[36]，该数据集是由车内摄像头采集到的驾驶员状态图像，展现了驾驶过程中出现的驾驶分心情况，其中包括10种类别，如表3所示，每种类别的行为示例见图5.

本文先将数据集按驾驶员进行划分，再按采集时间将数据划分到多个轮次(epoch)，模拟在车辆行驶过程中，车内传感器不断收集数据和进行联邦学习. 以驾驶员People81为例，每个轮次累计的图片数量如图6所示，第1轮采集到了伸手到后面(C7)以及整理发型和妆容(C8)类别的动作图像；随着时间的推移，不断采集到更多种类的行为数据，第4轮已经收集到右手打电话(C2)、操作收音机(C5)、伸手到后面(C7)以及整理发型和妆容(C8)这4种类别的图像.

Table 3 Driver Behavior Category

表3 司机行为类别

类别	司机的行为
C0	安全驾驶
C1	用右手发短信
C2	用右手打电话
C3	用左手发短信
C4	用左手打电话
C5	操作收音机
C6	喝东西
C7	伸手到后面
C8	整理发型和妆容
C9	与乘客交谈

Fig. 5 Examples of dataset classification

图5 数据集分类示例

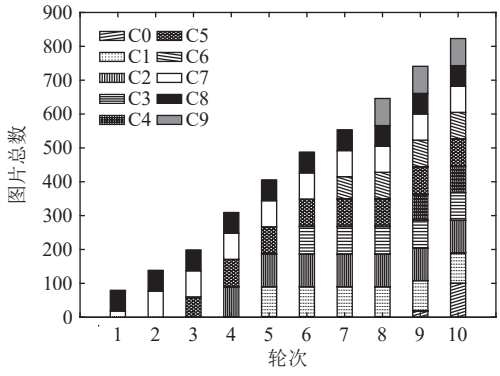


Fig. 6 Data collection process of People81

图6 People81的数据采集过程

本实验选择了5位驾驶员的数据，分别是People26, People35, People42, People72, People81，对每一个参与方每一类动作的图片数量进行统计，其结果如图7所示. 由图7可见，People72数据量最小，People26数据量最大. 对于大部分人，每一种行为的数据量基本一致，经统计，People42呈现均匀分布，每一类图片数量都是59张；People72的数据最不均衡，最多的一类C0有63张，而C7种类的图片只有2张. 这反映出车辆所收集到的数据是non-IID数据. 本文将数据集划分为训练集(70%)和测试集(30%).

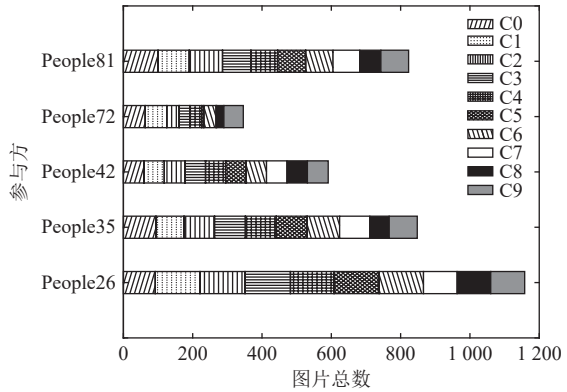


Fig. 7 Behavior images statistics of participants

图7 参与方行为的图片统计

4.2 实验设置

本文评估了 3 种经典的神经网络模型在驾驶行为分析任务中的模型性能，包括 VGG 16, Inception, ResNet 50，评估指标包括模型的准确度和损失. 准确度反映了模型正确预测样本的比例，在分类任务中，高准确度意味着模型能够正确分类更多的样本；损失反映了模型预测值与真实值之间的差异，损失越低表示模型的预测越接近真实标签，在训练深度学习模型时的目标是最小化损失函数. 经过实验评估，选择 VGG 16 模型作为多阶段联邦学习的基准模型，模型结构如图 4 所示，输入图片大小为 224×224，经过卷积层、最大池化层、全局平均池化层和全连接层，最终输出一个大小为 10 的预测向量，这对应于本文所采用数据集中的 10 种驾驶行为.

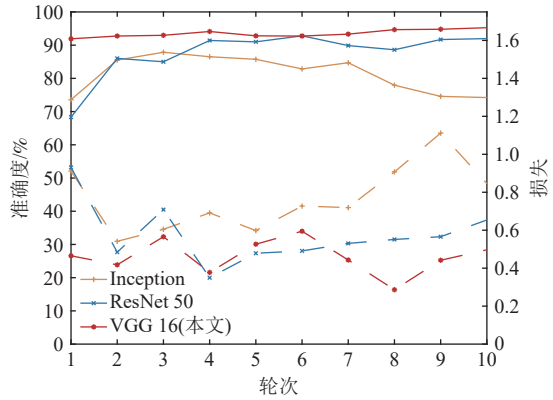
为验证本文提出的多阶段联邦学习机制 FedWO 的性能，一共进行 10 轮次训练. 经实验分析，设置第 1~3 轮次为第 1 阶段，开展联邦平均多方计算；第 4~7 轮次为第 2 阶段，开展联邦加权多方计算；第 8~10 轮次为第 3 阶段，开展个性化计算. 实验选择了 4 种对比方法，包括 Only, FedA, FedAO, FedW. Only 算法只有本地训练、没有联邦学习，即车辆执行 10 轮本地训练；FedA 是联邦平均算法，即车辆执行 10 轮联邦平均；FedAO 是联邦平均与个性化结合算法，即第 1~7 轮采用联邦平均，第 8~10 轮采用本地训练；FedW 是联邦加权算法，即第 1~3 轮采用联邦平均，第 4~10 轮采用式(4)的联邦加权.

在本研究中，超参数δ和φ的合适取值是通过实验分析获得的. 在模型比较的初步实验中，根据多参与方模型差值的分布规律，设置δ=0.4. 在实验中观察到当φ=0.3 时，模型性能与通信效率之间达到平衡. 此外，在式(4)中，α, β, γ的取值影响着模型聚合. 在本实验中，算法准确度、数据丰富程度以及数据量大小同等重要，因此设置$\alpha=\beta=\gamma=\frac{1}{3}$. 在某些情况下，可以忽略特定因素的影响，例如：当各参与方的数据集大小相同时，可将数据集大小的参数γ设置为 0，忽略该参数对全局模型的影响. 同理，若本地模型准确度对于全局聚合影响较大，则适当调大α.

实验使用 Linux 操作系统，GPU 内存 6 GB，程序开发工具为 Python 3.7, Tensorflow 2.0.0, Keras 2.3.1, CUDA 11.2. Tensorflow 具有出色的分布式处理能力，适用于联邦学习场景中在多个设备或节点上并行处理数据和训练模型. Keras 作为 Tensorflow 的一个高级 API，提供了用户友好的界面，使得模型构建和测试更便捷. CUDA 提供通用并行计算架构，使得 Tensorflow 和 Keras 能够利用 GPU 实现加速，显著提高了模型训练和数据处理的速度，从而能够解决复杂的计算问题. 本文选择 Tensorflow, Keras, CUDA，是因为它们为联邦学习提供了高效的分布式处理和 GPU 加速，这对于处理复杂的数据集和加快模型训练至关重要.

4.3 实验结果分析

本文进行了多次重复独立实验来消除实验结果的随机性，首先对比了不同神经网络模型的性能优劣，对比模型分别为 VGG 16, Inception, ResNet 50，结果如图 8 所示. 实线代表准确度，虚线代表损失函数值. 以 People81 为例，在 10 轮次后，使用 VGG 16 进行本地训练得到的模型准确度达到 95.25%，而使用 Inception 和 ResNet 50 模型的准确度分别为 74.22% 和 91.97%，相较于 VGG 16 分别降低了 21.03 个百分点和 3.28 个百分点；从模型收敛性来看，VGG 16 模型的损失值为 0.49，而 Inception 和 ResNet 50 的损失值分别为 0.84 和 0.65，可见在 10 个轮次内 VGG 16 模型收敛性更好. 因此，本文方法采用 VGG 16 模型，以实现较高的模型准确度和收敛速度.

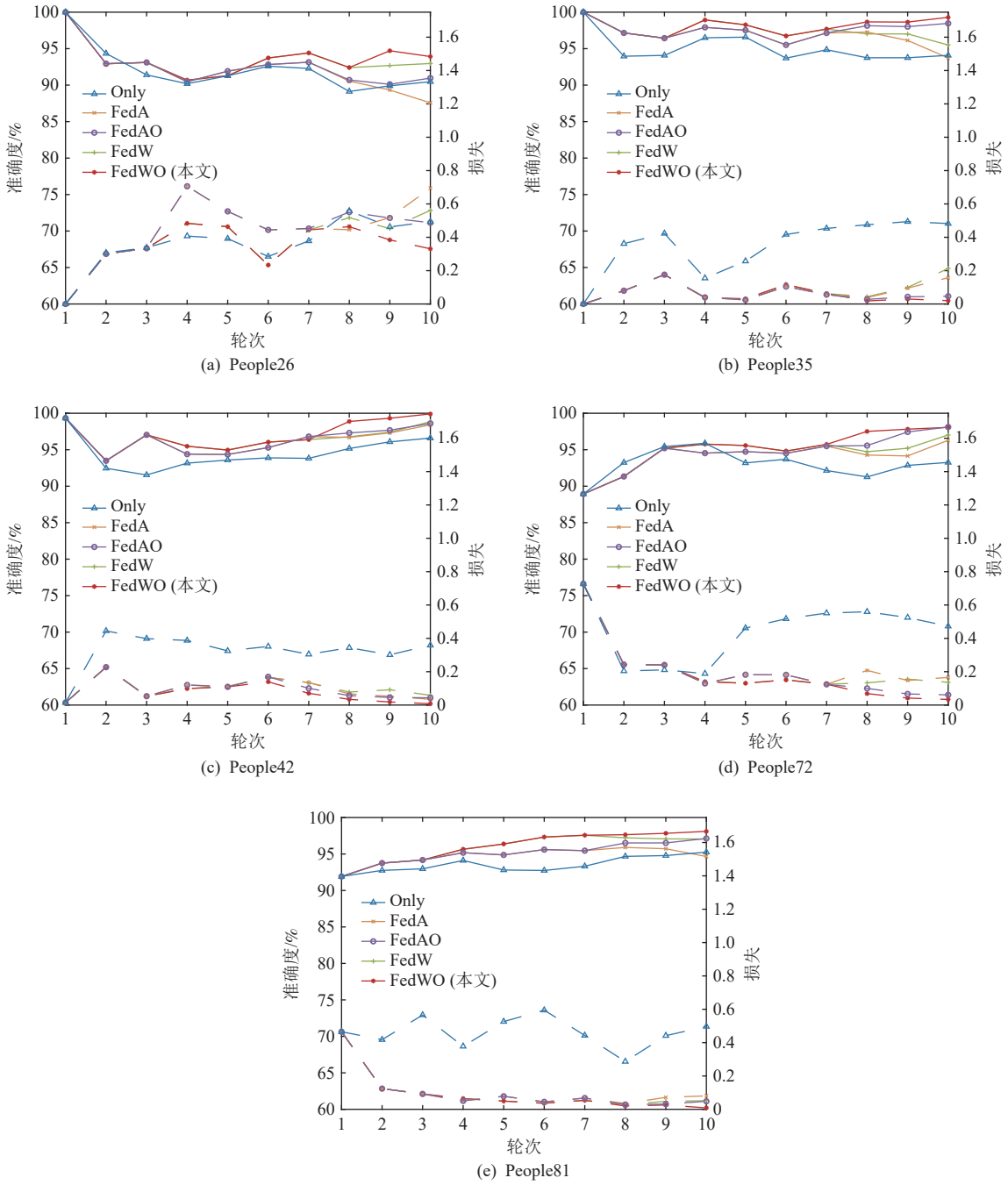


注：实线代表准确度；虚线代表损失.

Fig. 8 Accuracy and loss obtained by participant People81 using different models

图 8 参与方 People81 使用不同模型获得的准确度和损失

接下来对比了 Only, FedA, FedAO, FedW, FedWO 算法性能的优劣，结果如图 9 所示. People26 和 People35 在第 1 轮的损失函数值分别为 0.000 07 和 0.000 175，近乎于 0，准确度也达到了 100%. 分析其数据集可发现，第 1 轮他们分别只收集到 1 类数据，模型分类简单，因此达到了 100% 的准确度. 而 People42 和 People81 在第 1 轮收集到 2 类数据，模型分类准确度有所降低，但也在 90% 以上. 而 People72 在第 1 轮的准确度只有 88.93%，损失函数值为 0.72，也较高，分析发现，



注：实线表示准确度；虚线表示损失.

Fig. 9 Accuracy and loss of vehicles local model

图 9 车辆本地模型的准确度及损失

People72 在第 1 轮有 2 类数据，且数据量只有 45 张，数据集小以及数据种类增多导致了 People72 的分类准确度低.

对所有人而言，相较其他算法，本地训练算法 Only 的准确度是最低的，损失函数值也是最高的，Only 算法相较于其他算法不收敛，对于 People42 和 People72 结果尤其明显. 但对于 People26, People35, People81 来说，在最后 1 轮时，FedA 算法的准确度是低于 Only 算法的，这也证明了，针对 non-IID 数据，FedA 算法性能不够好.

与其他算法相比，本文提出的 FedWO 算法总体上呈现最优的效果，同时联邦加权算法 FedW 分类结果优于 FedA，虽然 People26 在第 5 轮时其性能有所下降，但随着轮次的增加，其性能是有所提升的. 对每一种算法在训练结束时(即第 10 轮)的分类准确度进行统计分析，如表 4 所示. 由表 4 可见，FedWO

拥有最高的准确度；与 FedA 相比，FedWO 算法在 People26, People35, People42, People72, People81 参与方的准确度分别提升了 6.33 个百分点、5.57 个百分点、1.5 个百分点、1.78 个百分点、3.45 个百分点，损失函数值降低了 0.36, 0.13, 0.02, 0.13, 0.07；与 Only 相比，FedWO 准确度分别提升了 3.44 个百分点、5.21 个百分点、3.31 个百分点、4.81 个百分点、2.85 个百分点，损失函数值降低了 0.16, 0.46, 0.34, 0.44, 0.48. 从 5 个参与方的算法损失值来看，在 10 轮训练结束后，FedWO 均展现了最低的损失函数值，这表明 FedWO 在有限的训练轮次中实现了更好的模型收敛. 综上可见，本文提出的多阶段联邦学习机制在面向 non-IID 数据时具有更好的性能.

Table 4 Comparison of Algorithm Accuracy

表 4 算法准确度对比 %

参与方	Only	FedA	FedAO	FedW	FedWO
People26	90.47	87.58	90.96	92.96	**93.91**
People35	94.08	93.72	98.46	95.50	**99.29**
People42	96.59	98.40	98.59	98.84	**99.90**
People72	93.25	96.28	98.11	97.02	**98.06**
People81	95.25	94.65	97.12	97.04	**98.10**

注：黑体数值为最优值.

在 FedWO 机制中，通过控制车辆上传本地模型以及服务器下发全局模型，实现对通信资源的有效利用. 为验证相关性能，本实验对模型平均传输次数进行统计. 对比算法有 4 个，即全部上传和下发的 FedWO、支持选择上传的 FedWO(up)、支持选择下发的 FedWO(down) 以及支持选择上传和选择下发的 FedWO(up+down)，其功能如表 5 所示. 选择上传表示车辆可以选择是否向服务器上传本地模型，全部上传表示全部参与车辆都要向服务器发送本地模型，选择下发表示服务器选择是否向车辆发送全局聚合模型，全部下发表示服务器将全局模型下发给全部参与车辆.

4 种算法的平均传输次数如图 10 所示. 在 FedWO 中，每一轮次都有 2 次数据传输，分别是上传与下发，前 7 轮训练共传输 14 次，后 3 轮是个性化计算，不需要上传和下发，而本文所提出的传输控制策略明显降低了传输开销. 对于 People26 与 People35, 3 种降低开销方法的平均传输次数是一样的，分别是 11 次和 9 次，都明显低于 FedWO 的 14 次. People42 的 FedWO(up) 的平均传输次数略高于 FedWO(down) 与 FedWO(up+down)，可见控制模型下发对其更有效. 对于 People72, FedWO(up+down) 明显低于其他 3 种方法，平均传输次数为 7.2 次；对于 People81, FedWO (up+down) 的传输开销略高于 FedWO(down). 总体来看，FedWO(up+down) 表现出较好的性能，验证了本文所设计的传输控制策略能够降低通信资源消耗.

Table 5 Transmission Control Policy

表 5 传输控制策略

算法	选择上传	全部上传	选择下发	全部下发
FedWO		√		√
FedWO (up)	√			√
FedWO (down)		√	√	
FedWO (up+down)	√		√	

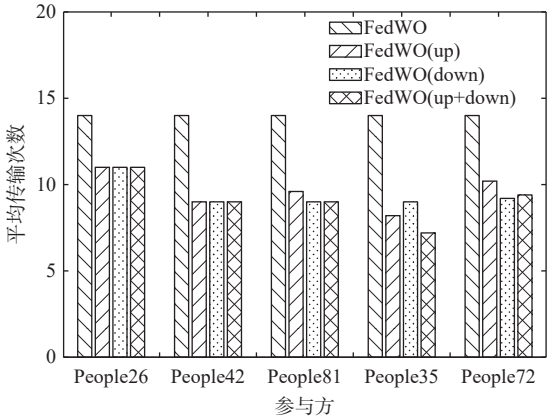
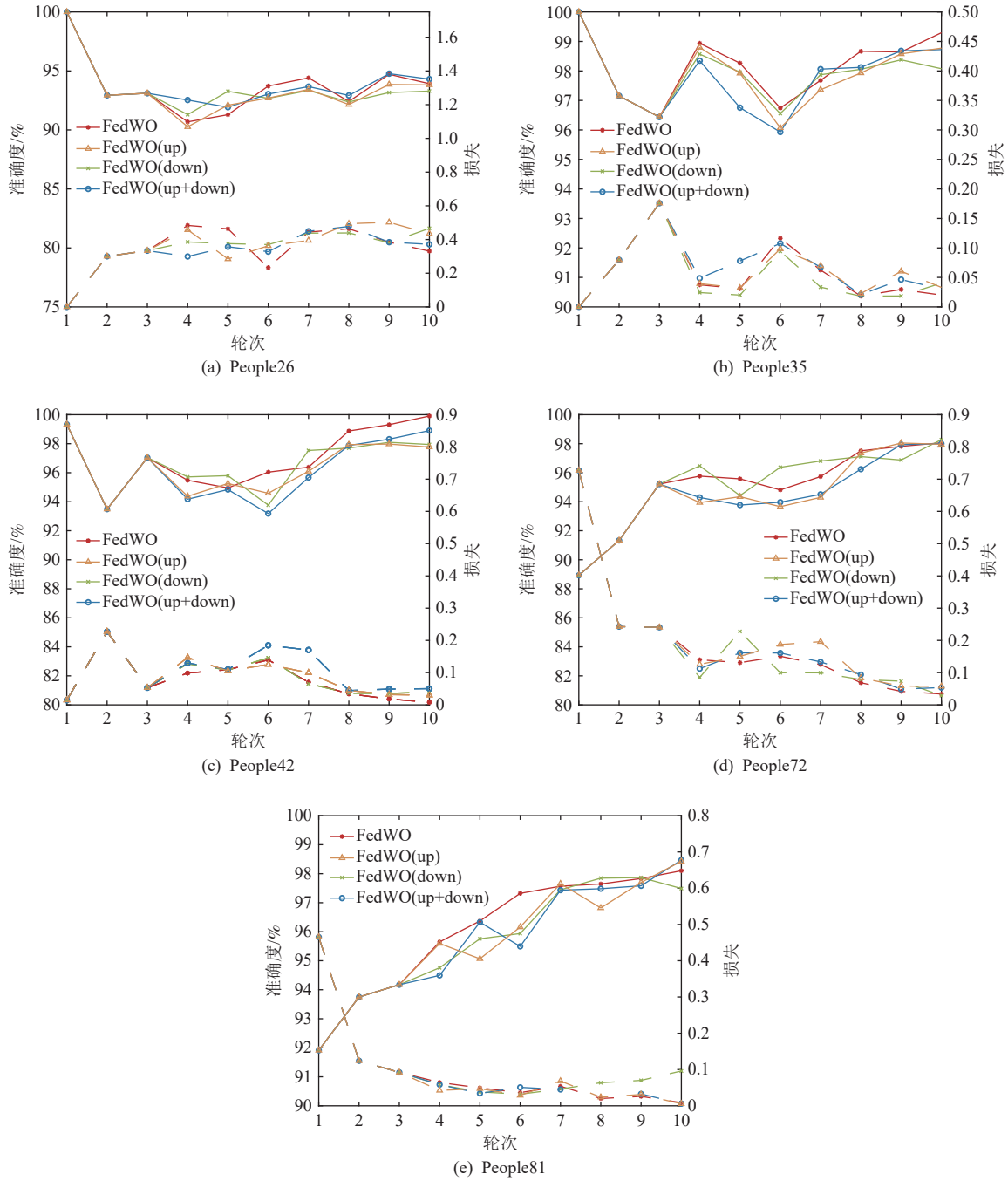


Fig. 10 Average number of transmissions

图 10 平均传输次数

执行传输控制策略的前提是不能对模型性能产生显著影响，FedWO, FedWO(up), FedWO(down), FedWO (up+down) 的分类准确度及损失变化如图 11 所示. 除了 People72 在第 1 轮的准确度是 88.93% 以外，各算法的准确度都在 90% 以上. 不同的传输控制策略在不同轮次中表现出不同的性能. FedWO 和 FedWO (up+down) 在大部分轮次中性能较好，尤其是在后几轮. 而 FedWO(up) 和 FedWO(down) 在中间的轮次中波动较大，但最终都趋于稳定. 在 10 轮训练结束后，FedWO(up+down) 对 People26 和 People81 的分类准确度最高，分别达到了 94.3% 和 98.4%. 而对 People35 和 People42 来说，依然是 FedWO 的准确度较高，比 FedWO(up+down) 分别高了 0.6 个百分点和 1 个百分点. 对于 People72, FedWO 和 FedWO(up+down) 的准确度基本一致，分别是 98.06% 和 98%. 从 5 个参与方的损失函数值变化来看，在有限的 10 轮次训练中，整体模型收敛趋势一致，在第 10 轮结束后，损失值均趋向于 0. 而对 People81 来说，FedWO(up)在最后的 3 轮次中，损失值有所升高，其原因可能是在进行



注：实线代表准确度；虚线代表损失.

Fig. 11 Accuracy and loss of local model with transmission control

图 11 具有传输控制的本地模型准确度和损失

个性化训练时，出现了模型过拟合的情况. 总体来说，与 FedWO 相比，具有传输控制方法的准确度变化不大，模型收敛情况基本一致. 因此，FedWO(up+down)能够有效地减少开销，且不会对参与方模型准确度造成显著影响，验证了 FedWO 的有效性.

本文还比较了不同数量的参与方对联邦学习算法性能的影响，结果见表 6. FedWO(5) 表示 5 个参与方的联邦学习，FedWO(4) w/o People26 表示不包含 People26 的 4 个参与方(即 People35, People42, People72, People81)的联邦学习，以及 FedWO(4) w/o People72 表示不包含 People72 的 4 个参与方(即 People26, People35, People42, People81)的联邦学习. 数据显示，增加 People26 后，People35, People42, People81 的模型准确度分别上升了 0.52 个百分点、1.4 个百分点、0.06 个百分点；增加 People72 后，People35 和 People42 的模型准确度略微上升，分别上升了 0.29 个百分点

和0.37个百分点，而People81的准确度下降了1.4个百分点. 在第7轮次结束后，People72与People81拥有相同的3类数据，数据种类少、关联性大，People72对People81可能存在消极影响，导致增加People72后，People81的模型准确度有所下降.

Table 6 Impact of Different Participants on Models' Accuracy

表6 不同参与方对模型准确度的影响

参与方	FedWO(4) w/o People26	FedWO(4) w/o People72	FedWO(5)
People35	98.77	99.00	99.29
People42	98.50	99.53	99.90
People81	98.04	99.50	98.10

综合表6和图6可见，联邦学习的参与方数量及其数据质量会影响算法的准确度. People26具有较为丰富的数据集，其参与联邦学习能够优化全局模型，使其他参与方的模型准确度有所上升. 相比之下，People72的数据量和数据种类较少，其加入联邦学习对其他参与方的影响有好有坏. 这表明，拥有高质量数据的参与方加入联邦学习，将对整体模型的准确度产生积极影响. 反之，如果某参与方的数据量和质量较低，则可能对模型准确度产生负面影响.

5 总 结

本文提出了一种面向non-IID数据的多阶段联邦学习机制FedWO，旨在解决联邦学习参与方本地数据不均衡的情况下，联邦学习中模型不收敛的问题. 具体来说，第1阶段使用联邦平均算法(FedAvg)进行联邦学习，其目的是更快地达到一个全局模型参数基准；第2阶段考虑每台车辆本地模型的精度、数据丰富程度和数据量，使用联邦加权算法进行多方计算，同时加入了传输控制策略，以减少在联邦学习中上传和下载模型所带来的开销；为了使本地模型达到更高的准确度，第3阶段采用个性化计算，各个参与车辆使用本地数据再次微调本地模型参数，使得模型更优. 使用驾驶员状态数据集的实验表明，与其他算法相比，本文提出的多阶段联邦学习机制针对non-IID数据具有更高的模型准确度，同时降低了传输开销.

随着人工智能技术的飞速发展，模型的解释性在联邦学习的应用中扮演着重要角色，尤其在敏感领域，如交通、金融、医疗和法律等，用户和专业人士需要清晰地了解模型的决策逻辑，以培养对模型的信任并合理地应用模型输出. 如何提升联邦学习模型的可解释性，以及优化联邦学习的传输控制以适应不同类型的网络环境，将是未来的研究方向.

作者贡献声明：唐晓岚提出了算法思路和实验方案；梁煜婷负责完成实验并撰写论文；陈文龙提出指导意见并修改论文.

参 考 文 献

[1] Chai Haoye, Leng Supeng, Chen Yijin, et al. A hierarchical blockchain-enabled federated learning algorithm for knowledge sharing in Internet of vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22(7): 3975−3986

[2] Wang Ge, Xu Fangmin, Zhang Hengsheng, et al. Joint resource management for mobility supported federated learning in Internet of vehicles[J]. Future Generation Computer Systems, 2022, 129: 199−211

[3] Ren Ju, Zhang Deyu, He Shiwen, et al. A survey on end-edge-cloud orchestrated network computing paradigms: Transparent computing, mobile edge computing, fog computing, and cloudlet[J]. ACM Computing Surveys, 2019, 52(6): 1−36

[4] Zhang Wei, Li Xiang, Ma Hui, et al. Federated learning for machinery fault diagnosis with dynamic validation and self-supervision[J]. Knowledge-Based Systems, 2021, 213: 106679

[5] Cao Longbing. Non-IID recommender systems: A review and framework of recommendation paradigm shifting[J]. Engineering, 2016, 2(2): 212−224

[6] Kaissis G A, Makowski M R, Rückert D, et al. Secure, privacy-preserving and federated machine learning in medical imaging[J]. Nature Machine Intelligence, 2020, 2(6): 305−311

[7] Lu Yunlong, Huang Xiaohong, Dai Yueyue, et al. Blockchain and federated learning for privacy-preserved data sharing in industrial IoT[J]. IEEE Transactions on Industrial Informatics, 2019, 16(6): 4177−4186

[8] Luo Bing, Li Xiang, Wang Shiqiang, et al. Cost-effective federated learning in mobile edge networks[J]. IEEE Journal on Selected Areas in Communications, 2021, 39(12): 3606−3621

[9] He Chaoyang, Shah A D, Tang Zhenheng, et al. FedCV: A federated learning framework for diverse computer vision tasks[J]. arXiv preprint, arXiv: 2111.11066, 2021

[10] Mo Huiling, Zheng Haifeng, Gao Min, et al. Multi-source heterogeneous data fusion based on federated learning[J]. Journal of Computer Research and Development, 2022, 59(2): 478−487 (in Chinese)
(莫慧凌, 郑海峰, 高敏, 等. 基于联邦学习的多源异构数据融合算法[J]. 计算机研究与发展, 2022, 59(2): 478−487)

[11] Kong Xiangjie, Wang Kailai, Hou Mingliang, et al. A federated learning-based license plate recognition scheme for 5G-enabled Internet of vehicles[J]. IEEE Transactions on Industrial Informatics, 2021, 17(12): 8523−8530

[12] Liang Feiyuan, Yang Qinglin, Liu Ruiqi, et al. Semi-synchronous federated learning protocol with dynamic aggregation in Internet of

vehicles[J]. IEEE Transactions on Vehicular Technology, 2022, 71(5): 4677−4691

[13] Lu Yunlong, Huang Xiaohong, Zhang Ke, et al. Blockchain empowered asynchronous federated learning for secure data sharing in Internet of vehicles[J]. IEEE Transactions on Vehicular Technology, 2020, 69(4): 4298−4311

[14] Tang Xiaolan, Liang Yuting, Wang Guan, et al. Assisted driving system based on federated reinforcement learning[J]. Displays, 2023, 80: 102547

[15] Parekh R, Patel N, Gupta R, et al. Gefl: Gradient encryption-aided privacy preserved federated learning for autonomous vehicles[J]. IEEE Access, 2023, 11: 1825−1839

[16] Qu Zhiguo, Tang Yang, Muhammad G, et al. Privacy protection in intelligent vehicle networking: A novel federated learning algorithm based on information fusion[J]. Information Fusion, 2023, 98: 101824

[17] Li Zejun, Wu Hao, Lu Yunlong. Coalition based utility and efficiency optimization for multi-task federated learning in Internet of vehicles[J]. Future Generation Computer Systems, 2023, 140: 196−208

[18] He Jingyi, Gong Biyao, Yang Jiadi, et al. ASCFL: Accurate and speedy semi-supervised clustering federated learning[J]. Tsinghua Science and Technology, 2023, 28(5): 1−15

[19] Shu Jiangang, Yang Tingting, Liao Xinying, et al. Clustered federated multitask learning on non-IID data with enhanced privacy[J]. IEEE Internet of Things Journal, 2022, 10(4): 3453−3467

[20] Tian Pu, Liao Weixian, Yu Wei, et al. WSCC: A weight-similarity-based client clustering approach for non-IID federated learning[J]. IEEE Internet of Things Journal, 2022, 9(20): 20243−20256

[21] Dong Fang, Ge Xinghua, Li Qinya, et al. PADP-FedMeta: A personalized and adaptive differentially private federated meta learning mechanism for AIoT[J]. Journal of Systems Architecture, 2023, 134: 102754

[22] Yang Lei, Huang Jiaming, Lin Wanyu, et al. Personalized federated learning on non-IID data via group-based meta-learning[J]. ACM Transactions on Knowledge Discovery from Data, 2023, 17(4): 1−20

[23] Li Wenzhu, Wang Shuang. Federated meta-learning for spatial-temporal prediction[J]. Neural Computing and Applications, 2022, 34(13): 10355−10374

[24] Hu Kai, Li Yaogen, Zhang Shuai, et al. FedMMD: A federated weighting algorithm considering non-IID and local model deviation[J]. Expert Systems with Applications, 2024, 237: 121463

[25] Kim H, Kim Y, Park H. Reducing model cost based on the weights of each layer for federated learning clustering[C]//Proc of the 12th Int Conf on Ubiquitous and Future Networks (ICUFN). Piscataway, NJ: IEEE, 2021: 405−408

[26] Zhang Wenyu, Wang Xiumin, Zhou Pan, et al. Client selection for federated learning with non-IID data in mobile edge computing[J]. IEEE Access, 2021, 9: 24462−24474

[27] McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[C]//Proc of the 20th Artificial Intelligence and Statistics. New York: PMLR, 2017: 1273−1282

[28] Zhong Zhengyi, Bao Weidong, Wang Ji, et al. A hierarchically heterogeneous federated learning method for cloud-edge-end system[J]. Journal of Computer Research and Development, 2022, 59(11): 2408−2422(in Chinese)
(钟正仪, 包卫东, 王吉, 等. 一种面向云边端系统的分层异构联邦学习方法[J]. 计算机研究与发展, 2022, 59(11): 2408−2422)

[29] Kairouz P, McMahan H B, Avent B, et al. Advances and open problems in federated learning[J]. Foundations and Trends® in Machine Learning, 2021, 14(1/2): 1−210

[30] Li Yifei, Guo Yijia, Alazab M, et al. Joint optimal quantization and aggregation of federated learning scheme in VANETs[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(10): 19852−19863

[31] Van Tilborg H C A, Jajodia S. Encyclopedia of Cryptography and Security[M]. Berlin: Springer, 2014

[32] Sun Lichao, Qian Jianwei, Chen Xun. LDP-FL: Practical private aggregation in federated learning with local differential privacy[J]. arXiv preprint, arXiv: 2007.15789, 2020

[33] Li Tian, Sahu A K, Zaheer M, et al. Federated optimization in heterogeneous networks[J]. Proceedings of the Machine Learning and Systems, 2020, 2: 429−450

[34] Karimireddy S P, Kale S, Mohri M, et al. SCAFFOLD: Stochastic controlled averaging for federated learning[C]//Proc of the 37th Int Conf on Machine Learning. New York: PMLR, 2020: 5132−5143

[35] Wang Jianyu, Liu Qinghua, Liang Hao, et al. Tackling the objective inconsistency problem in heterogeneous federated optimization[J]. Advances in Neural Information Processing Systems, 2020, 33: 7611−7623

[36] Google. StateFarm[EB/OL]. 2016[2023-02-01].https://www.kaggle.com/c/state-farm-distracted-driver-detection

Tang Xiaolan, born in 1987. PhD, associate professor. Member of CCF. Her main research interests include Internet of vehicles, smart education, urban computing, and driving behavior analysis.

唐晓岚, 1987 年生. 博士, 副教授. CCF 会员. 主要研究方向为车联网、智慧教育、城市计算、驾驶行为分析.

Liang Yuting, born in 2000. Master candidate. Her main research interests include Internet of vehicles and federated learning.

梁煜婷, 2000 年生. 硕士研究生. 主要研究方向为车联网、联邦学习.

Chen Wenlong, born in 1976. PhD, professor. Member of CCF. His main research interests include network protocol, Internet architecture, high performance router, and wireless sensor networks.

陈文龙, 1976 年生. 博士, 教授. CCF 会员. 主要研究方向为网络协议、互联网架构、高性能路由器、无线传感器网络.