arXiv:2303.05048v1 [cs.LG] 9 Mar 2023

面向大规模物联网网络协同智能的半联邦学习

Wanli Ni、Jingheng Zheng 和 Hui Tian

摘要——在大规模物联网网络中实施现有的联邦学习面临诸多严峻挑战，例如数据不平衡、统计异构以及设备多样性。为此，我们提出一种半联邦学习（SemiFL）框架，为实现智能物联网提供一种潜在解决方案。通过无缝整合集中式范式与联邦式范式，我们的半联邦学习框架展现出良好的可扩展性，即使存在计算受限传感器，也能支持大量物联网设备。此外，与传统学习方法相比，所提出的半联邦学习能够更好地利用分布式数据和计算资源，这得益于边缘服务器与本地设备之间的协同模型训练。仿真结果表明，我们的半联邦学习框架适用于大规模物联网网络。代码可在 https://github.com/niwanli/SemiFL_物联网**中获取**。

索引词—协同智能、数据与设备异构性、物联网（IoT）、半联邦学习。

I. 引言

借助机器学习驱动的智能解决方案，未来无线网络的通信范式正从物联网（IoT）转向互联智能 [1]。尽管大规模物联网设备的无处不在的连接有利于为各种智能物联网服务提供海量实时数据，但如何实现大规模物联网网络中的协同智能愿景仍然是一个悬而未决的问题 [2]。具体而言，通过收集大规模物联网设备的数据样本，采用集中式学习（CL）的边缘平台可以利用其强大的计算资源来训练高性能模型，但这也带来了无线边缘环境中通信成本高和隐私泄露风险高的问题 [3]。

联邦学习（FL）范式不发送原始数据，而是允许边缘设备在本地处理数据，然后仅将模型参数（例如权重或梯度）传输至边缘平台 [2]–[4]。然而，联邦学习中的所有边缘设备都应具备足够的计算资源来进行本地模型训练，而在实际中这并不总是成立[5]。也就是说，这些计算能力受限设备所拥有的数据集难以参与联邦学习模型训练

这一过程可能导致性能下降以及资源利用效率低下 [5]–[7]。为此，我们提出了一种新颖的半联邦学习框架，该框架包含两个计算层，通过同时考虑数据上传和梯度更新，兼具集中式学习和联邦学习的优势。

本研究部分受到中国国家自然科学基金项目（批准号：62071068）的资助，部分受到北京邮电大学（北京邮电大学）优秀博士生基金项目（批准号：CX2022301）的资助。（通讯作者：*Hui Tian*。）

Wanli Ni、Jingheng Zheng 和 Hui Tian 任职于中国北京 100876 的北京邮电大学网络与交换技术国家重点实验室（电子邮箱：{charleswall、zhengjh、tianhui}@bupt.edu.cn）。

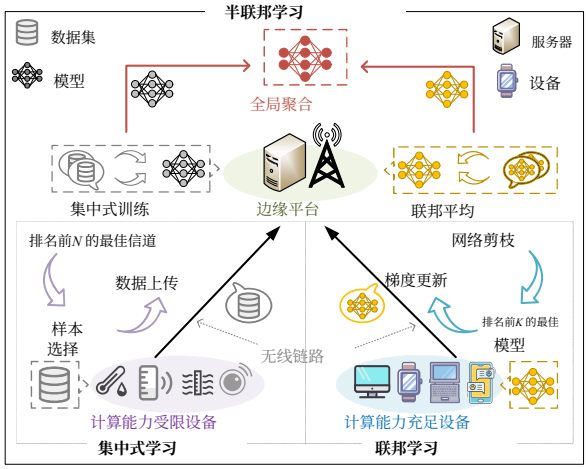

图1. 所提出的半联邦学习框架示意图。边缘平台由一个基站和一台边缘计算服务器组成。计算能力有限的设备被指定为集中式学习用户，用于上传原始数据进行集中式训练；而其他计算能力充足设备则被指定为联邦学习用户，用于更新模型梯度以进行联邦平均。

II. 框架设计

如图所示。1，我们提出了一种半联邦学习框架，根据设备的硬件能力，将本地设备分为计算能力受限设备和计算能力充足设备 [5]。前者被指定为集中式学习用户，上传原始数据以进行集中式训练；后者被指定为联邦学习用户，更新模型梯度以进行联邦平均。由一个基站和一台边缘服务器组成的边缘平台负责利用从集中式学习用户收集的原始数据以及联邦学习用户计算出的本地梯度来更新全局模型 [6]。通过协调联邦平均和集中式训练，我们的半联邦学习框架具有以下优势：i）物联网网络中的所有边缘设备，无论其计算能力如何，都能够参与学习过程，从而提高分布式数据的可用性。ii）即使不同设备间数据的统计特征存在异质性，也可以利用从集中式学习用户收集的大量数据重构出适合学习的数据集。iii）由于数据集传输以及本地设备与边缘平台之间的协同计算，边缘网络在计算和存储方面的资源利用率能够得到显著提升。

如图2所示，半联邦学习的整体学习过程包括三个主要阶段：初始化、模型训练和反馈。

- **阶段 1（初始化）**：边缘平台首先请求所有设备报告其数据量和分布、信道条件、计算资源等信息。基于所接收的信息，边缘平台生成

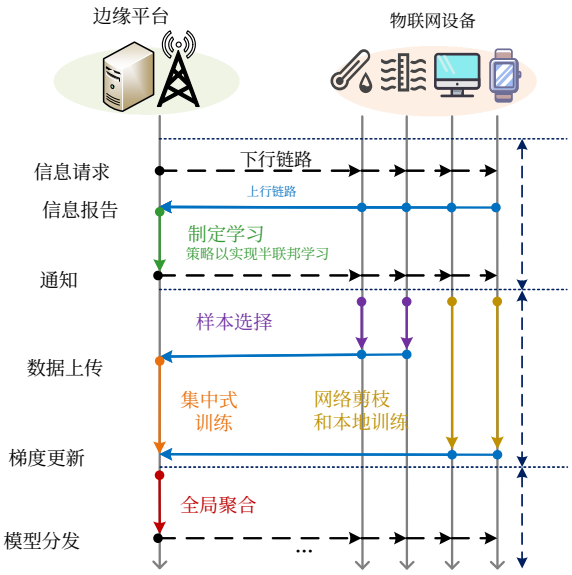

图2. 所提出的半联邦学习框架的学习过程在一轮中由三个阶段组成：初始化、模型训练和反馈。黑色虚线（蓝色实线）表示边缘平台与物联网设备之间的下行链路（上行链路）通信。

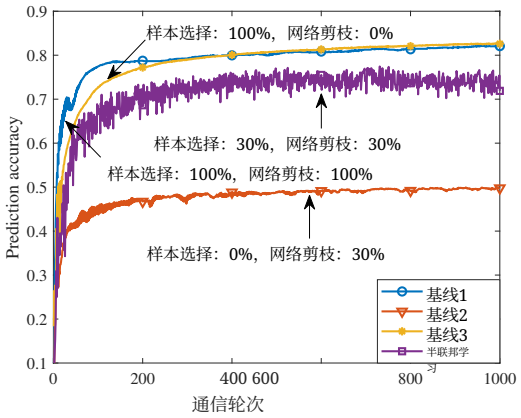

图3. Fashion-MNIST数据集上的学习性能。

一种学习策略，并将其通知所有设备，以用于后续的模型训练阶段。该学习策略的基本组成部分包括设备分类 [5],样本选择 [8], 和网络剪枝 [9]。

- **阶段 2（模型训练）**：根据协同学习策略，计算能力受限且信道质量排名前 N 的集中式学习用户，会将部分数据样本上传至边缘平台，由边缘平台代为进行集中式训练[5]。与此同时，所有联邦学习用户都会进行网络剪枝和本地训练，其中模型排名前 K 的用户会将其梯度上传至边缘平台 [10]。
- **阶段 3（反馈）**：边缘平台利用集中式训练和联邦平均得到的梯度进行全局模型聚合，然后将更新后的模型分发给所有设备，以供下一轮使用。

III. 实验结果

在实验中，我们考虑一个由 4 台计算能力受限设备和 6 台计算能力充足设备组成的物联网网络。我们通过以下方式评估半联邦学习框架的学习性能：在 Fashion-MNIST 数据集上训练多层感知机。作为对比，我们考虑以下基线。*a)* 基线*1*（集中式学习）：所有设备都将其数据集发送到边缘平台，以进行集中式模型训练。*b)* 基线*2*（联邦学习）：所有计算能力充足设备都以联邦方式训练模型。*c)* 基线*3*（未进行样本选择和网络剪枝的半联邦学习）：所有集中式学习用户都将其全部数据集上传至服务器，所有联邦学习用户都将其完整模型更新上传至服务器 [11]。

图 3 展示了不同基准方法在非独立同分布设置下的学习性能。从图中可以得出两个关键观察结果：i）不采用样本选择和网络剪枝的半联邦学习框架（即基线*3*）优于联邦学习（即基线*2*），且性能与集中式学习（即基线*1*）相近；由于需要传输所有本地数据集，集中式学习的通信开销更高。ii）尽管样本选择和网络剪枝可以进一步降低半联邦学习的通信开销，但也会降低可实现的预测精度。更多仿真结果和分析请参见 `https://github.com/niwanli/SemiFL`_物联网。

IV. 结论

本文针对物联网网络提出了一种新颖的半联邦学习（SemiFL）概念，将集中式学习（CL）与联邦学习（FL）整合到一个协调统一的框架中。通过支持异步数据上传和部分模型更新，所提出的半联邦学习能够缓解数据异质性和设备异质性带来的问题，这些问题会阻碍大规模物联网网络中的学习性能。此外，由于计算架构从单侧计算转变为双侧计算，与传统的集中式学习和联邦学习范式相比，网络边缘分布的资源能够得到更充分的利用。未来，我们将致力于开发高效方案，以缓解数据上传型集中式学习用户面临的隐私和安全问题。参考文献

[1] K. Yang, Y. Shi 等，“基于可重构智能表面的智能物联网联邦机器学习”，*IEEE Netw.*，第34卷，第5期，第16–22页，2020年9月。[2] W. Ni, Y. Liu 等，“多RIS辅助系统中的联邦学习”，*IEEEInternetThings J.*，第9卷，第12期，第9608–9624页，2022年6月。[3] V.-D. Nguyen, S. K. Sharma 等，“用于无线物联网网络资源分配的高效联邦学习算法”，*IEEE Internet Things J.*，第8卷，第5期，第3394–3409页，2021年3月。[4] Z. Zhang, S. Ma 等，“用于无人机互联网图像自动识别的稳健半监督联邦学习”，*IEEE Internet Things J.*，2022年，提前在线发表，doi:10.1109/JIOT.2022.3151945。[5] A. M. Elbir, S. Coleri 等，“混合联邦学习与集中式学习”，载于*Proc. IEEE EUSIPCO*，爱尔兰都柏林，2021年8月，第1–5页。[6] W. Ni, Y. Liu 等，“融合空中联邦学习与非正交多址接入：可重构智能表面能发挥什么作用？”，*IEEE Trans. Wireless Commun.*，第21卷，第12期，第10 083–10 099页，2022年12月。[7] N. Huang, M. Dai 等，“采用混合本地与集中式训练的无线联邦学习：时延最小化设计”，*J. Sel. Topics Signal Proc.*，2022年，提前在线发表，doi: 10.1109/JSTSP.2022.3223498。[8] A. Li, L. Zhang 等，“联邦学习的样本级数据选择”，载于*IEEE INFOCOM*，加拿大温哥华，2021年5月，第1–10页。[9] J. Ren, W. Ni 等，“面向网络边缘联邦学习的通信-学习权衡”，*IEEE Commun. Lett.*，2022年，提前在线发表，doi: 10.1109/LCOMM.2022.3174295。[10] M. M. Amiri, D. G¨und¨uz 等，“无线边缘联邦学习中具备更新感知能力的设备调度的收敛性”，*IEEE Trans. Wireless Commun.*，第20卷，第6期，第3643–3658页，2021年6月。[11] W. Ni, Y. Liu 等，“智能透射与反射可重构智能表面集成非正交多址接入与空中联邦学习：框架、分析与优化”，*IEEE Internet Things J.*，第9卷，第18期，第17 136–17 156页，2022年9月。