异构网络中采用无线能量传输的层次化联邦学习最优资源管理

Rami Hamdi，*IEEE*会员，Ahmed Ben Said、Emna Baccour、Aiman Erbad，*IEEE*高级会员，Amr Mohamed，*IEEE*高级会员，Mounir Hamdi，*IEEE*会士，以及Mohsen Guizani，*IEEE*会士

arXiv:2305.01953v1 [cs.NI] 3 May 2023

摘要—**远程监测系统分析不同智能工业应用中的环境动态，例如职业健康与安全以及环境监测。具体而言，在工业物联网（IoT）系统中，设备数量庞大且对性能的要求很高，这给计算、网络和设备能量等资源带来了压力。在异构无线网络（HetNets）中，为资源受限设备分布式训练用于智能工业物联网应用的机器学习/深度学习（ML/DL）模型极具挑战性。层次化联邦学习（HFL）在多个层级执行训练，将任务卸载至附近的多接入边缘计算（MEC）单元。本文提出一种由无线能量传输（WET）支持的新型节能型分层联邦学习框架，面向采用大规模多输入多输出（MIMO）无线回传的异构网络而设计。我们将能源效率优化方法 формулировать 为混合整数非线性规划（MINLP）问题，在其中优化HFL设备关联并管理无线传输能量。然而，由于其复杂度较高，我们设计了一种启发式资源管理算法，即H2RMA，在满足能量、信道质量和精度约束的同时，保持较低的计算复杂度。我们还通过高效的设备调度方案降低网络能耗。最后，我们研究设备移动性及其对HFL性能的影响。大量实验结果证实，与训练损失和电网能耗成本相关的指标相比，所提出的资源管理方法在异构网络上的HFL中具有优异性能。**

索引词—**异构网络、分层联邦学习、设备关联、无线能量传输、能效。**

I. 引言

职业健康与安全（OHS）近来已引起政府机构的关注，以保护工人并最大限度地降低健康风险 [2]–[4]。更广泛地说，远程监测系统能够捕捉环境动态并监测个人，应用于各种场景，以帮助老年人、患者和工厂工人。分布式人工智能（AI） [5] 和通信技术使得利用相关工业物联网设备（智能手机、移动设备、可穿戴设备、智能摄像头和智能手表）持续监测大量工人的关键健康指标（KHIs）成为可能。随后，收集到的关键健康指标会传输至中央单元或附近的边缘中央单元（CU），由机器学习和深度学习算法进行处理，以辅助决策任务。

随着连接设备数量不断增加，集中式机器学习和深度学习（机器学习/深度学习）算法需要越来越多的内存资源和大量计算能力（即 CPU） [6]。在具有实时和低延迟要求的无线网络环境中，集中式学习方案会因造成延迟而变得低效，并可能损害设备间交换数据的隐私。因此，去中心化学习方法应运而生，其中尤以联邦学习（FL） [6], [7]为代表。联邦学习能够让多个设备共同训练一个共享模型，每个设备仅使用自己的私有本地数据 [7]。具体而言，各设备利用本地数据样本训练并更新所接收全局模型的权重，然后将新的模型权重发送至中央单元，以聚合成共享学习模型。最后，中央单元将聚合后的模型权重广播给联邦学习参与方，以更新其模型。该流程不断迭代，直至达到所需的模型性能。联邦学习解决了数据隐私方面的挑战，同时减少了传输至中央单元的数据量，从而提升带宽利用率。然而，面对数千台设备，需要解决多项挑战，包括通信瓶颈、能耗增加、拥塞以及数据偏斜（即数据非独立同分布） [10]。为应对这些问题，层次化联邦学习（HFL）应运而生 [8], [9]。层次化联邦学习能够将计算任务卸载至设备附近的多接入边缘计算（MEC），从而提高能效并改善物联网节点的响应时间。在这一新配置中，设备将本地更新发送至对应的移动边缘计算服务器，以进行学习和模型聚合。随后，移动边缘计算服务器将聚合后的模型传输至中央单元，以计算全局聚合。层次化联邦学习最显著的特点是能够应对并管理多样化的数据分布，例如数据的非独立同分布特性。通过谨慎地将设备分配至移动边缘计算服务器，可以加快收敛速度并取得更高的性能水平 [10]。我们强调，如果数据是独立同分布的，层次化联邦学习可能搭配不同方式……

本研究的初步版本 [1]已发表于2021年的IEEE全球通信大会论文集。
R. Hamdi任职于卡塔尔多哈的阿伯丁大学附属AFG学院（电子邮箱：rami.hamdi@afg-aberdeen.edu.qa）。A. B. Said和A. Mohamed任职于卡塔尔多哈的卡塔尔大学工程学院（电子邮箱：abensaid@qu.edu.qa；amrm@qu.edu.qa）。E. Baccour、A. Erbad和M. Hamdi任职于卡塔尔基金会旗下哈马德·本·哈利法大学科学与工程学院信息与计算技术部（卡塔尔多哈）（电子邮箱：ebaccourepbe- said@hbku.edu.qa；aerbad@ieee.org；mhamdi@hbku.edu.qa）。M. Guizani任职于阿联酋阿布扎比的穆罕默德·本·扎耶德人工智能大学（MBZUAI）机器学习系（电子邮箱：mguizani@ieee.org）。

对系统性能没有太大贡献。

在设备间干扰持续存在、受到时延约束、能耗增加且设备异构性显著的无线通信环境中，实现 HFL 面临挑战。因此，近期许多研究致力于让 HFL 适用于无线通信网络。具体而言，近期研究解决了资源分配问题，包括无线资源分配、边缘设备分配、设备选择和能量管理，旨在充分发挥系统潜力并优化其性能。另一方面，无线能量传输（WET）是一项很有前景的技术，旨在解决无线网络中的能源瓶颈 [11]。该系统通过利用电源发射的电磁辐射波，为无线设备输送能量。通过无线技术供电，可满足移动设备频繁充电的需求。因此，这项技术是设计节能型物联网无线网络的关键技术 [12]。

机器学习/深度学习分布式部署面临的挑战之一是能耗较高。因此，为了提高采用无线回传的异构网络中分层联邦学习处理过程的能效，我们提出引入能量采集 [13], , 由各移动边缘计算节点通过无线能量传输将采集到的能量传输给其关联设备。与 [8]–[10], [14]–[26], 中的现有方案相比，我们的工作独具特色，因为我们专注于设计一种由无线能量传输辅助的节能型分层联邦学习架构。据我们所知，我们是首个考虑在由无线能量传输支持的大规模 MIMO无线回传异构网络中实现联邦学习的研究。此外，我们通过开发一种高效的动态资源分配方案，考虑了一个模拟建筑工地工人的实际移动场景。我们的工作贡献如下：

- 我们提出了一种新颖的节能型分层联邦学习框架，该框架面向采用大规模多输入多输出无线回传的异构网络，并由无线能量传输提供支持。
- 我们将能源效率优化方法建模为混合整数非线性规划问题，通过优化分层联邦学习的设备关联并管理无线传输能量，实现能源效率优化。
- 我们通过解析方法推导出最优设备关联，并采用暴力搜索确定能量管理决策。
- 鉴于最优解的复杂度较高，我们提出了一种设备关联与能量管理算法，旨在提升所提出的分层联邦学习框架的能源效率。
- 一些能量需求较高的设备可能会在部分帧中停用。因此，我们提出了一种设备调度策略，在不降低学习准确率的前提下降低能耗。
- 我们设计了一种高效的动态设备关联方案，并考虑了一个代表建筑工地工人的移动场景。
- 最后，我们对所提出的系统进行了全面评估，以证明我们的资源管理方法具有高能效和高准确率性能， de-

设备关联，以及不同网络配置下的设备调度方案。

本文其余部分的组织如下：第二节探讨了研究单层联邦学习和分层联邦学习的相关工作。第三节介绍系统模型。第四节对优化问题进行了 formul。第五节研究资源管理。第六节研究移动性对联邦学习性能的影响。第七节给出评估结果。最后，第八节得出结论。

II. 相关工作

A. 联邦学习

已有大量研究探讨了无线网络中的联邦学习资源分配问题 [27]–[37]。在 [27], 中，通过提出高效的用户关联和调度方案，构建了一种能量采集联邦学习框架。在考虑有限能量预算的情况下，通过设计高效的资源管理方案，研究了联邦学习无线系统的能耗最小化问题。在 [29],中，提出了数据增强和自适应用户调度方案，以克服联邦学习中数据分布不均衡所带来的挑战。在 [30]中，通过设计高效的控制方案，对收敛时间进行了理论研究和优化。在 [31] 中，通过提出高效的用户调度策略，提升了联邦学习性能。此外， [32] 的作者设计了一种最优用户选择算法，以最小化联邦学习的收敛速度。在 [33], 中，提出了一种用户选择方案，以减小不可靠用户的影响并提高联邦学习训练的准确率。为解决联邦学习用户的电池限制问题，早期研究 [34]–[37]在执行联邦学习任务时引入了无线能量传输系统。在 [34], 中，作者假设无人机（UAV）具备无线能量传输能力，并从能效角度提出了一种合适的联邦学习解决方案。在 [35], 中，实现了一种异构移动架构，并提出了能够最小化联邦学习设备能耗的最优资源分配方案。在 [36], 中，提出了一种低复杂度波束成形策略，以最小化由无线能量传输支持的联邦学习系统中的均方误差。 [37]的作者构建了一种分析型联邦学习框架，在模型收敛与功耗之间建立权衡，为确保高效的学习任务提供了指导。

B. 层次化联邦学习

其他研究也探讨了 HFL [8]–[10], [14]–[26] 的资源分配优化问题。具体而言， [9]中的作者利用 HFL 为物联网系统设计了一种可靠的异常检测策略。此外，为了应对非独立同分布（non-i.i.d.）数据带来的挑战， [10] 中的作者设计了一种边缘设备分配方案，以提升联邦平均算法的性能。考虑到用户之间存在相同的非独立同分布数据约束，[19]中的作者针对异构物联网系统，提出了一种高效的 HFL 资源分配与用户分配方案。同时， [14] 的作者提出了一种新颖的 HFL 架构，旨在最小化所消耗的

表 I：主要符号列表。

符号	说明	符号	说明
N_{cu}	中央单元的天线数量	$\mathbf{H}_m$	中央单元与移动边缘计算 m
N_{mec}	每个移动边缘计算的天线数量	$\mathbf{g}_{m,k}$	设备 k 与移动边缘计算 m
K	设备数量	$\beta_{m,k}$	路径损耗
L	帧数	σ^2	噪声方差
M	移动边缘计算服务器数量	$\mathbf{W}_m$	移动边缘计算服务器 m处的检测矩阵
T	每帧时长	$E^{\text{dev}}_{m,k}(l)$	设备 k 所需能量在帧 l 分配至移动边缘计算 m
Λ_m	与多接入边缘计算关联的设备集合 m	E^{cir}	固定电路能耗
K_m	Λ_m 的基数	$E^{\text{ac}}_{m,k}(l)$	设备 k所需传输能量在第 l帧
C	类别数量	Q	本地联邦学习模型的大小
$\mathcal{D}_k$	设备 k的数据集	B_{max}	最大电池容量
$\boldsymbol{x}_k$	设备 k 的输入	$A_{m,\ k}(l)$	收集到的能量量分配给移动边缘计算 m 的设备 k 帧 l
$\boldsymbol{y}_k$	设备 k 的输出	$B_k(l)$	设备 k 在第 l 帧的电池电量
$\boldsymbol{S}$	设备的数据矩阵	$E^{\text{wet}}_m(l)$	从移动边缘计算 m 传输的能量在帧 l
$S_{c,k}$	.....的训练数据样本数量类别 c的设备 k	$\alpha_m(l)$	网格的权重
$\boldsymbol{w}_k(l)$	本地机器学习模型的权重向量设备 k 在帧 l	Δ	电网能耗总成本
$\cup_{j\in\Lambda_m}\mathcal{D}_j$	移动边缘计算 m可用的虚拟数据集	$\chi_{m,k}$	设备关联索引
$p^{(m)}(i)$	每个移动边缘计算 i 的类别概率 m的分布	$p(\cdot)$	全局类别分布

能量以及在具有挑战性的数据分布条件下的总训练时延。 [17],中的研究也将能源效率作为重点，作者提出了一种适用于分层联邦学习的功率控制和客户端关联解决方案。 [8] 中的研究旨在利用一种新颖的资源分配和设备关联算法，最小化全局成本。联邦学习系统中的另一个极其重要的指标是时延，这也是 [21] 和 [15]两项研究的优化目标。前者优化了多个服务器与设备之间的关联，而后者则设计了一种跨异构蜂窝网络的分层联邦学习框架。 [18] 中的作者提出了一种灵活的去中心化训练过程控制方法，以增强分层联邦学习的隐私性；而 [22] 中的作者则建议采用区块链来满足构成分层联邦学习系统的车联网的隐私要求。 [23],提出了一种基于聚类的新型分层联邦学习框架，以提高测试准确率。 [24] 中的作者通过提出一种混合数据划分算法，对分层联邦学习进行了研究。 [25],提出了一种高效的分层联邦学习算法，旨在确定最优的聚类结构。 [26], 设计了一种资源分配启发式算法，以在满足公平性约束的条件下保证分层联邦学习的收敛速度。最后，博弈论被用于 [16] 和 [20] ，以设计高效的分层联邦学习资源分配策略。

与文献中提出的方案不同，本文提出了一种新型分层联邦学习架构，该架构采用大规模MIMO无线回传，即移动边缘计算服务器与设备之间以及移动边缘计算服务器之间的链路均为无线链路。此外，我们考虑了一种体现建筑工地工人特点的实际移动场景。另外，大多数相关研究并未考虑在分层联邦学习中使用能量采集。因此，本文假设移动边缘计算服务器具备无线能量传输能力，且设备配备能量采集电池，从而提高能效。考虑到这些具有挑战性的约束，我们构建并研究了一种新型资源优化问题。

III. 系统模型与能耗模型

A. 层次化联邦学习模型

大规模多输入多输出基于同时复用大量天线，以满足多个用户在相同时间-频率资源中的需求，这有助于

提高能效、频谱效率和时延 [38]。此外，异构网络基于在多个维度上采用不同类型的小区进行网络密集化，从而能够高效利用无线资源并提升服务质量。将大规模多输入多输出和异构网络引入物联网无线网络，有利于优化联邦学习性能。因此，我们采用如图 1 所示的异构网络。我们的系统由一个配备 N_{cu} 根天线的大规模多输入多输出中央单元（中央单元）和 M 个移动边缘计算服务器组成，每个移动边缘计算服务器均配备 N_{mec} 根多输入多输出天线（$N_{\text{cu}} > MN_{\text{mec}}$）。不同的移动边缘计算服务器通过无线回传链路 [39]与中央单元通信。此外，每个移动边缘计算服务器 m 为多个单天线设备提供服务。设 Λ_m 为由移动边缘计算服务器 m服务的 K_m 个设备的集合。我们注意到，网络中的设备总数等于 K。我们研究一个时间区间内的系统，并将该区间划分为 L 个帧，每帧持续时间为T。表 I 给出了本文引入的不同关键符号。设备关联索引定义为：

$$\chi_{m,k} = \begin{cases} 1, & \text{if device } k \text{ is associated to MEC } m, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

每个设备 k 拥有一个数据集 $\mathcal{D}_k = \{\boldsymbol{x}_k, \boldsymbol{y}_k\}$，其中 $\boldsymbol{x}_k$ 和$\boldsymbol{y}_k$ 分别表示设备 k的输入和输出。这些设备在本地训练其联邦学习模型，并将获得的权重发送至各自关联的移动边缘计算服务器。随后，每个移动边缘计算服务器聚合其设备的权重，并将结果转发至中央单元。最后，中央单元聚合从所有移动边缘计算服务器接收的权重，并将更新后的联邦学习模型共享给所有设备。该过程不断重复，直至达到期望的准确率。设矩阵$\boldsymbol{S} \in \mathbb{N}^{C\times K}$ 描述所有设备之间的数据划分，其中 C 表示学习问题中的类别总数。具体而言，每个元素 $S_{c,\ k}$ 表示设备 k 所拥有且属于类别 c的数据样本数量。

令 $\boldsymbol{w}_k(l)$表示设备 k 在帧 l内训练的本地模型的联邦学习权重向量。在聚合了所有权重（首先是从设备接收的权重，然后是从移动边缘计算服务器收集的权重）之后，中央单元通过最小化特定的损失函数，对所得权重进行全局拟合。在本文中，我们选用交叉熵损失函数，其可表示如下：

$$\mathcal{L}\left(\boldsymbol{w}\right) = \sum_{c=1}^{C} -p(c)\mathbb{E}_{\boldsymbol{x}\sim q(\boldsymbol{x}|y=c)}\left[\log d_c(\boldsymbol{x};\boldsymbol{w})\right], \tag{2}$$

其中， $p(\cdot)$定义类别的全局分布， $q(\cdot \mid \cdot)$表示似然函数，而 $d_c(\boldsymbol{x};\boldsymbol{w})$表示在参数 $\boldsymbol{w}$下，输入 $\boldsymbol{x}$ 属于类别 c的概率。

收到相关设备发送的本地联邦学习模型权重后，每个移动边缘计算节点均按照 [10]中的公式执行聚合：

$$\boldsymbol{w}_m^{\text{mec}}(l) = \sum_{k\in\Lambda_m} \frac{|\ \mathcal{D}_k\ |}{|\cup_{j\in\Lambda_m}\mathcal{D}_j|}\boldsymbol{w}_k(l), \tag{3}$$

其中， $\cup_{j\in\Lambda_m}\mathcal{D}_j$ 是与移动边缘计算 m相关的不同设备所拥有的本地数据集的集合。因此，可以将其视为在移动边缘计算 m处训练的虚拟数据集。接下来步骤是对中央单元从所有移动边缘计算服务器接收到的权重进行聚合，如 [10] 中所示：

$$\boldsymbol{w}^{\text{cu}}(l) = \sum_{m=1}^{M} \frac{|\cup_{k\in\Lambda_m}\mathcal{D}_k|}{|\cup_{j=1:K}\mathcal{D}_j|}\boldsymbol{w}_m^{\text{mec}}(l) \tag{4}$$

最后，将聚合后的权重广播给所有设备，并重复此过程，直到达到目标收敛率。

由于学习过程中的分布差异和联邦化，本地联邦学习模型权重与聚合后的权重之间可能存在偏差。 [10]研究了这种散度，并将其量化如下：

$$\theta = \sum_{m=1}^{M} \frac{|\cup_{k\in\Lambda_m}\mathcal{D}_k|}{|\cup_{j=1:K}\mathcal{D}_j|} \times \left\|D^{(m)}\right\|_1, \tag{5}$$

其中， $D^{(m)} = \left\{\left\|p(i) - p^{(m)}(i)\right\|\right\}_{i=1}^{C}$， 而 $p^{(m)}(i)$表示每个移动边缘计算节点 m的分布中类别 i 的概率。

B. 通信模型

我们假设，中央单元与移动边缘计算 m 之间的信道经历莱斯衰落，可表示为 $\mathbf{H}_m \in \mathbb{C}^{M_s\times N}$。另一方面，我们考虑移动边缘计算 m 与每个设备 $k \in \Lambda_m$之间服从高斯独立同分布（i.i.d.）的小尺度衰落信道，记为 $\mathbf{f}_{m,\ k}$。移动边缘计算 m 与每个设备 $k \in \Lambda_m$ 之间的接入链路由向量 $\mathbf{g}_{m,\ k} = \beta_{m,\ k}^{\frac{1}{2}}\mathbf{f}_{m,\ k} \in \mathbb{C}^{N_{\text{mec}}\times 1}$表示，其中$\beta_{m,\ k}$ 表示路径损耗。为此，信道矩阵定义为$\mathbf{G}_m = [\mathbf{g}_{m,\ k}]_{k=1:K_m} \in \mathbb{C}^{N_{\text{mec}}\times K_m}$。

在我们的系统中，我们假设每个移动边缘计算服务器都能获得所有设备的完美信道状态信息（CSI）。因此，所有移动边缘计算服务器都可以采用协作式迫零（ZF）方案，以抑制小区间和小区内干扰的影响，具体如 [40]所示。令$\mathbf{Z}_m = \mathbf{G}_m(\mathbf{G}_m^H\mathbf{G}_m)^{-1}$表示移动边缘计算服务器 m的解码矩阵。相应地，与移动边缘计算服务器 m连接的设备 k 的速率可表示为：

$$r_{m,k}^{\text{dev}}(l) = \lambda \cdot \log_2\left(1 + \frac{p_{m,k}^{\text{dev}}(l)}{\sigma^2\left[(\mathbf{G}_m^H\mathbf{G}_m)^{-1}\right]_{k,k}}\right), \tag{6}$$

其中， λ 表示带宽， σ^2 表示加性白高斯噪声（AWGN）的方差。

当移动边缘计算服务器从其相关设备接收到联邦学习权重后，会进行聚合，并将聚合后的权重转发至中央单元。在中央单元侧，对于多天线接收器，采用块对角化（BD）来抑制多设备干扰。该方法在 MIMO 系统中表现出优异的性能 [41]。因此，我们为每个移动边缘计算服务器m 分配一个检测矩阵 W_m，以区分从不同移动边缘计算服务器接收到的信号。根据 [41], W_m位于矩阵 H_m的零空间中，其中 $\mathrm{H}_m = [\mathrm{H}_1^T...\mathrm{H}_{m-1}^T\mathrm{H}_{m+1}^T...\mathrm{H}_M^T]^T$。移动边缘计算服务器 m 的解码矩阵表示为 $\mathrm{W}_m = \mathrm{V}_m^{(0)}\mathrm{V}_m^{(1)}$，其中 $\mathrm{V}_m^{(0)}$由 H_m的最后 $N_{\text{cu}} - J_m$ 个右奇异向量组成，$\mathrm{V}_m^{(1)}$表示 $\mathrm{H}_m\mathrm{V}_m^{(0)}$的前 J_m 个右奇异向量，$J_m = \quad {}_m \quad J_m = \quad {}_m$ rank(H)，以及 rank(H $\mathrm{V}_m^{(0)}$)。因此，

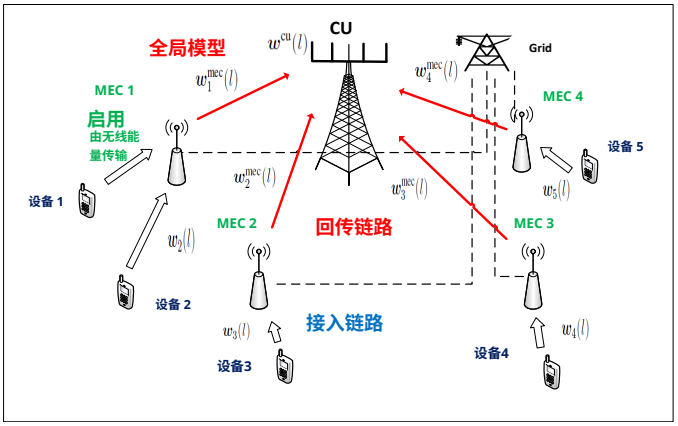

图 1：电池供电终端设备中支持无线能量传输的异构网络分层联邦学习。

移动边缘计算服务器 m的上行速率表示为 [42]:

$$r_m^{\text{mec}}(l) = \lambda \cdot \log_2\left(1 + \frac{p_m^{\text{mec}}(l)\,\|\mathbf{H}_m\mathbf{W}_m\|_F^2}{\sigma^2}\right). \quad (7)$$

C. 能耗模型

本节评估移动边缘计算服务器及其相关设备执行不同联邦学习操作所需的能量。首先，设备 k 连接到移动边缘计算服务器m ，在帧 l 内完成联邦学习任务所需的能量表示为:

$$E_{m,k}^{\text{dev}}(l) = E^{\text{cir}} + E_k^{\text{cmp}} + E_{m,k}^{\text{ac}}(l), \quad (8)$$

其中， E^{cir} 是滤波器、混频器和数模转换器消耗的固定电路能量。 E_k^{cmp} 是设备 k 训练其本地数据所需的能量，而 $E_{m,\,k}^{\text{ac}}(l)$ 是在帧l传输所得权重所需的能量。

设备 k 训练其数据所需的能量可按 [43]中的方法计算:

$$E_k^{\text{cmp}} = \varsigma\omega\vartheta^2 Q, \quad (9)$$

其中， ς 是取决于设备芯片的能量系数， ω 表示 CPU 周期数， ϑ 是 CPU 时钟频率，Q 表示本地模型的大小，而该大小在很大程度上取决于所部署的联邦学习算法。值得一提的是，所有设备、移动边缘计算服务器和中央单元中的本地模型和全局模型均具有相同的权重大小。
接下来，在帧 l，设备 k 将已训练本地模型的权重发送至其指定的移动边缘计算服务器 m ，所需能量为:

$$E_{m,k}^{\text{ac}}(l) = \frac{p_{m,k}^{\text{dev}}(l)Q}{r_{m,k}^{\text{dev}}(l)}. \quad (10)$$

其中，$p_{m\,,k}^{\text{dev}}(l)$ 是在帧 t连接到移动边缘计算 m 的设备 k 所需的发射功率。根据公式 (6)，$p_{m\,,k}^{\text{dev}}(l)$可表示为:

$$p_{m,k}^{\text{dev}}(l) = \sigma^2\left[(\mathbf{G}_m^H\mathbf{G}_m)^{-1}\right]_{k,k}\left(2^{\frac{r_{m,k}^{\text{dev}}(l)}{\lambda}} - 1\right). \quad (11)$$

因此，将 $p_{m\,,\,k}^{\text{dev}}(l)$ 替换为其在 (10) 式中的表达式，我们得到:

$$E_{m,k}^{\text{ac}}(l) = \frac{\sigma^2 Q\left[(\mathbf{G}_m^H\mathbf{G}_m)^{-1}\right]_{k,k}\left(2^{\frac{r_{m,k}^{\text{dev}}(l)}{\lambda}} - 1\right)}{r_{m,k}^{\text{dev}}(l)}. \quad (12)$$

我们的联邦学习系统中的每台设备都配备了电池，最大容量为 B_{max}。这些设备端电池也用于存储无线传输能耗。令 $A_{m,\,k}(l)$ 表示在时间帧 l传输至连接到移动边缘计算服务器 m 的设备 k的能量。因此，我们将与移动边缘计算服务器 m 关联的任意设备 k 的电池电量更新定义如下:

$$B_k(l+1) = \min\left(B_{\max}, B_k(l) - E_{m,k}^{\text{dev}}(l) + A_{m,k}(l)\right). \quad (13)$$

此外，我们将任意移动边缘计算 m在帧 l 使用无线能量传输为其关联设备供能所消耗的能量定义如下:

$$E_m^{\text{mec}}(l) = E^{\text{cir}} + E_m^{\text{wet}}(l) + E_m^{\text{bh}}(l), \quad (14)$$

其中， $E_m^{\text{wet}}(l)$表示帧 l时，移动边缘计算 m通过无线能量传输系统传输给其相关设备的能量。同时， $E_m^{\text{bh}}(l)$表示移动边缘计算 m 向中央单元传输聚合的联邦学习权重所消耗的能量。需要指出的是， E^{cir} 为固定电路能耗。
接下来，我们采用MIMO无线能量传输系统的最优波束成形器，计算帧 l时与移动边缘计算 m 相关联的每台设备 k 获得的能量。该能量表示为:

$$A_{m,k}(l) = E_m^{\text{wet}}(l)\beta_{m,k}\xi_{m,k}N_{\text{mec}}, \quad (15)$$

其中$\xi_{m,k}$ 是设备 k的波束成形系数。
帧 l 时，移动边缘计算 m 向中央单元传输聚合的联邦学习权重所需的能量为:

$$E_m^{\text{bh}}(l) = \frac{p_m^{\text{mec}}(l)Q}{r_m^{\text{mec}}(l)}. \quad (16)$$

发射功率 $p_m^{\text{mec}}(l)$可由式（7）推导得出:

$$p_m^{\text{mec}}(l) = \frac{\sigma^2}{\|\mathbf{H}_m\mathbf{W}_m\|_F^2}\left(2^{\frac{r_m^{\text{mec}}(l)}{\lambda}} - 1\right). \quad (17)$$

因此，将 $p_m^{\text{mec}}(l)$ 替换为其在式 (16) 中的表达式，可得:

$$E_m^{\text{bh}}(l) = \frac{\sigma^2 Q\left(2^{\frac{r_m^{\text{mec}}(l)}{\lambda}} - 1\right)}{\|\mathbf{H}_m\mathbf{W}_m\|_F^2\, r_m^{\text{mec}}(l)}. \quad (18)$$

我们假设每个移动边缘计算服务器 m 在时间帧 l 消耗的能量由参数 $\alpha_m(l)$ [44] 加权。因此，所有移动边缘计算服务器消耗的能量总成本计算如下:

$$\Delta = \sum_{l=1}^{L}\sum_{m=1}^{M}\alpha_m(l)E_m^{\text{mec}}(l). \quad (19)$$

IV. 问题表述

本研究旨在为异构网络设计一种节能型分层联邦学习框架，其中设备由无线能量传输供能。具体而言，我们

旨在通过优化设备关联和无线能量，在联邦学习权重与中心权重之间的散度约束下，最大限度地降低不同移动边缘计算服务器消耗的电网能源总成本。由于联邦学习权重与中心权重之间的偏差不应超过给定阈值，因此在推导设备关联时，应考虑不同设备的本地数据分布。基于此，优化问题被表述为式 (20a)。约束 (20a) 确保中心权重与联邦学习权重之间的偏差不超过阈值$\theta_{\max}$。约束 (20b) 保证网络中的每台设备仅关联到一个移动边缘计算服务器。式 (20c) 确保每帧设备消耗的能量受电池可用能量的限制。式 (20d) 确保存储的能量不超过电池的最大容量。最后，约束 (20e) 和 (20f) 保证通过无线链路传输的能量为正，且设备关联决策为二元变量。

$$\min_{\substack{\{\chi_{m,k},E_m^{\text{wet}}(l)\}\\ m=1,\ldots,M,k=1,\ldots,K,l=1,\ldots,L}} \Delta=\sum_{l=1}^{L}\sum_{m=1}^{M}\alpha_m(l)E_m^{\text{mec}}(l) \tag{20}$$

$$\text{s. t. } \theta\le\theta_{\max}, \tag{20a}$$

$$\textstyle\sum_{m=1}^{M}\chi_{m,k}=1,\ \ \forall k=1,\ldots,K, \tag{20b}$$

$$\chi_{m,k}\textstyle\sum_{i=1}^{l}E_{m,k}^{\text{dev}}(i)\le B_k(0)+\sum_{i=1}^{l}A_{m,k}(l),$$
$$\forall m=1,\ldots,M,k=1,\ldots,K,l=1,\ldots,L, \tag{20c}$$

$$\textstyle\sum_{i=1}^{l}A_{m,k}(l)-\chi_{m,k}\sum_{i=1}^{l-1}E_{m,k}^{\text{dev}}(i)\le B_{\max},$$
$$\forall m=1,\ldots,M,k=1,\ldots,K,l=2,\ldots,L, \tag{20d}$$

$$E_m^{\text{wet}}(l)\ge 0,\ \ \forall m=1,\ldots,M,l=1,\ldots,L, \tag{20e}$$

$$\chi_{m,k}\in\{0,1\},\ \ \forall m=1,\ldots,M,k=1,\ldots,K. \tag{20f}$$

此外，我们研究了通过纳入设备调度而扩展的公式问题 (20)。为了在不损害联邦学习训练损失的前提下提高网络的能源效率，一些获得大量无线传输能量的设备可能会在某些通信轮次中不参与调度。这使优化问题变得更具挑战性。此外，我们还将公式问题 (20) 扩展到移动场景，并假设采用动态设备关联，即关联索引会随时间变化，如$\chi_{m,\ k}(l)$。

V. 资源管理

A. 最优解

本节将推导问题 (20) 的最优解。首先可以看出，由于其组合性质以及约束 (20a) 的非线性，我们的优化问题属于混合整数非线性规划（MINLP）。因此，我们提出使用以下定理来求出每个移动边缘计算在每个时帧传输的最优能量，从而简化该问题。

定理 1. 对于一个移动边缘计算 m 及其关联设备集合 Λ_m，从移动边缘计算 m传输至其关联设备的最优传输能量 $E_m^{\text{wet}*}(l)$ 为：

$$E_m^{\text{wet}*}(l)=\max_{k\in\Lambda_m}\frac{E_{m,k}^{\text{dev}}(l)-B_k(l)}{\beta_{m,k}\xi_{m,k}N_{\text{mec}}}. \tag{21}$$

证明。在帧 l，与移动边缘计算 m 关联的每个设备 k，除了自身电池供电外，还可通过无线能量传输（WET）获取收集到的能量。k 的总能量表示如下：

$$E_{m,k}^{\text{dev}}(l)=B_k(l)+A_{m,k}(l). \tag{22}$$

因此，通过将 $A_{m,k}(l)$ 替换为式 (15) 中给出的表达式，我们可以推导出由移动边缘计算 m传输至设备 k的能量 $E_m^{\text{wet}}(l)$：

$$E_m^{\text{wet}}(l)=\frac{E_{m,k}^{\text{dev}}(l)-B_k(l)}{\beta_{m,k}\xi_{m,k}N_{\text{mec}}}. \tag{23}$$

每个移动边缘计算节点根据最高能量需求为其关联的设备提供服务。更具体地说，移动边缘计算节点会找出能量需求最高的设备，并为其提供服务，同时为其他具有相同能量需求的设备提供服务。

□

利用公式 (21) 中给出的能量传输表达式 $E_m^{\text{wet}}(l)$，优化问题 (20) 可简化为一个整数非线性问题。该简化问题旨在寻找最优设备关联，是一个组合问题，且已知为 NP-hard [45]。不过，可以使用暴力搜索（广度优先搜索）求解。尽管推导最优解仍然十分复杂，但所得结果可作为后文提出的启发式方法的基准。还值得一提的是，忽略约束 (20a) 后，该优化问题将变为整数线性问题，可使用 CVX 等数值工具求解。

B. 启发式方法

由于最优解的组合复杂度可能会妨碍其在线部署，我们提出一种低复杂度的启发式算法来解决设备关联和能量管理问题。与优化问题类似，我们的算法旨在最小化所消耗电网能源的成本，同时满足联邦学习散度约束。

令 R_m 表示 MEC m 的范围，即与其关联的最远设备的路径损耗。根据式（16），可以看出，路径损耗越高的设备通过 WET 接收到的能量越少。换言之，分配给 MEC m 且路径损耗低于 R_m 的任何设备，即 $\beta_{m,\ k}\le R_m$，都可能受益于更高的传输能量。因此，每个设备 k 都会尝试寻找范围大于其路径损耗的 MEC 集合。随后，它选择散度最小的 MEC。该设备关联过程逐一进行。如果某个设备无法找到具有合适范围的 MEC，则将其关联到距离最近的 MEC。我们提出的设备关联启发式算法不仅降低了能耗，还减小了联邦学习与中心权重之间的散度，有助于最小化训练损失。最后，对于每个帧 l，我们可以使用公式计算每个 MEC 通过 WET 传输的能量

（21）。算法1展示了我们提出的 HFL启发式资源管理算法（H2RMA）的不同步骤。此外，所提出的算法如图2所示，我们还在图3中给出了设备关联过程的示例。从图3可以看出，设备3位于 MEC 1 和 MEC 2 的范围内。因此，它选择与散度最小的 MEC 关联。对于设备4，它位于所有 MEC 的范围之外。因此，它选择与能量需求最小的 MEC 关联。随后，根据设备4更新 MEC 4 的范围。

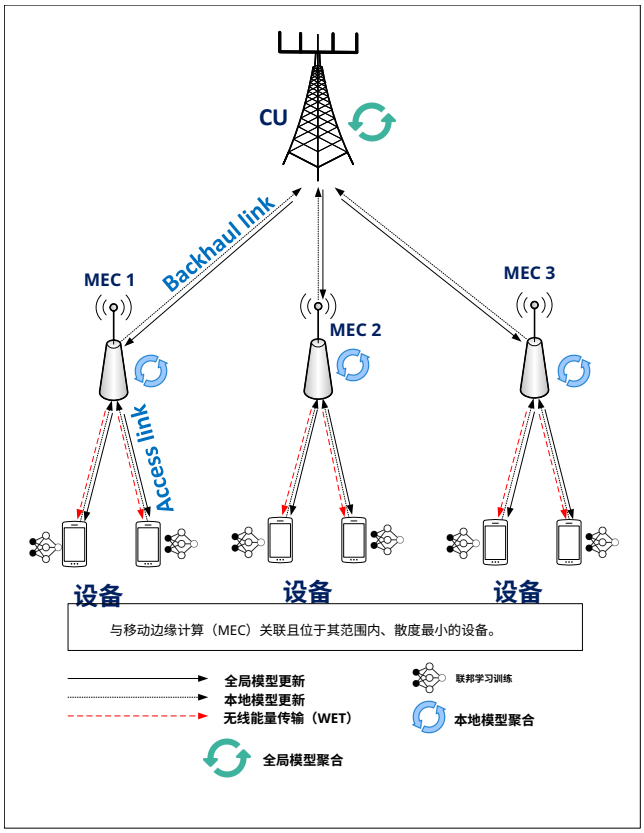

图2：所提出的算法 H2RMA 的示意图。

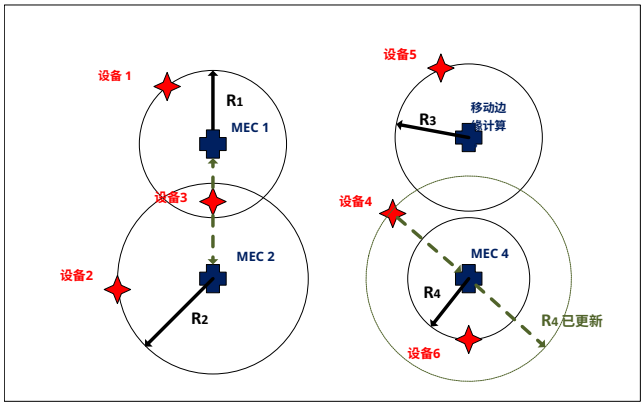

图3：设备关联过程示例。

我们研究了 H2RMA 的复杂度。第 (3-20) 行展示了设备关联过程，其中使用了两个 **for**，复杂度为 $O(MK)$。在第 (21-29) 行的 for 循环中，执行功率控制，其复杂度为 $O(L(2K+M))$。因此，H2RMA 的复杂度可表示为：

$$C^{\text{H2RMA}} = O(MK + L(2K + M)). \tag{24}$$

因此，H2RMA 可以在多项式时间内执行。

算法 1 HFL启发式资源管理算法（H2RMA）

1: 初始化: 2: $\chi_{i_k,k} \leftarrow 0$, $m = 1..M$, $k = 1..K$, 3: $R_m \leftarrow 0$, $m = 1..M$4: 对于 $k = 1 : K$ 执行 5: 对初始化 $i_k \leftarrow 0$6: 对初始化 $\theta_{\min} \leftarrow \theta_{\max}$7: 对于 $m = 1 : M$ 执行 8: 如果 $\beta_{m,k} \leq R_m$ 则，验证设备 k 是否在移动边缘计算服务器 m的范围内。 9: 计算 θ10: 如果 $\theta \leq \theta_{\min}$ 则 11: $i_k \leftarrow m$，选择散度最小的移动边缘计算服务器。 12: $\theta_{\min} \leftarrow \theta$13: 结束如果 14: 结束如果 15: 结束循环 16: 如果 $i_k = 0$ 则，如果设备 k 不在任何移动边缘计算服务器的范围内。 17: $i_k \leftarrow \text{argmin}_{m=1..M} \beta_{m,k}$，与最近的移动边缘计算服务器 $m*$关联。 18: $R_{m*} \leftarrow \beta_{m*,k}$，更新移动边缘计算服务器 $m*$的范围。 19: 结束如果 20: $\chi_{i_k,k} \leftarrow 1$21: 结束循环 22: 对于 $l = 1 : L$ 执行 23: 计算每个设备所需的能量。 24: 使用(21)计算 $E_m^{\text{wet}}(l)$ $m = 1..M$。 25: 将权重 $\boldsymbol{w}_k(l)$ 发送至相关移动边缘计算服务器，并使用(4)进行聚合。 26: $B_k(l+1) \leftarrow \min\left(B_{\max}, B_k(l) - E_m^{\text{dev}}, {}_k(l) + A_{m,k}(l)\right)$

电池更新27: 发送权重 $\boldsymbol{w}_m^{\text{mec}}(l)$ 到中央单元，使用 (5) 计算聚合。28: 将聚合后的权重发送给所有设备。29: 不同设备更新其联邦学习训练模型。30: **结束循环**。

C. 设备调度

我们现在研究所提出的分层联邦学习框架中的设备调度。事实上，在给定的帧 l 中，某些设备可能面临较差的信道条件，需要消耗大量无线传输能耗，却无法为分层联邦学习过程做出多少贡献。这会导致系统产生不必要的能耗，而这些能耗是可以避免的。在这种情况下，为了节省电网能源，一种切实可行的方法是根据可用能量和时变特性，在某些帧中停用部分设备。

信道。此外，每一帧中调度的设备都应满足最大散度约束，以免降低联邦学习的性能。

将设备调度纳入优化问题 (20) 后，该问题属于 MINLP，这增加了求得最优解的复杂度。可以通过广度优先搜索获得最优设备调度方案。然而，其较高的计算复杂度使该方法难以实现。因此，我们提出了一种高效的启发式设备调度方案，以节省不必要的电网能耗成本。我们首先假设所有设备均被选中。在当前帧中需要停用的设备，是指所需能量较高且超过某一阈值 E_{th} 的设备，其集合定义为：

$$\Omega = \{k \in \{1, \ldots, K\} | E^{\text{ac}}_{m,k}(l) > E_{\text{th}}\}. \tag{25}$$

我们逐个迭代地停用 Ω 中的设备，并从所需传输能耗较高的设备开始。在每次迭代中，仅当满足最大散度约束时才停用该设备，从而避免降低 FL 性能。

VI. 移动性影响

本节通过应用所提出的分层联邦学习框架，研究职业健康与安全的一个实际案例，以最小化建筑工地工人的健康风险。因此，我们选择了一个移动场景 [46] ，用于反映建筑工地上配备智能设备的工人。工人在工地上的动态交互使建筑工作成为风险最高的活动之一。工人的移动性可以通过隐马尔可夫模型 (HMMs) [46], 进行建模，该模型是一种强大的统计工具，用于表示观测结果上的概率分布。根据工人的速度和转向角，我们可以将其分为三种状态，即静止、正常和高风险。因此，可以根据这两个参数，即速度和转向角的取值，定义隐马尔可夫模型的隐状态。这些取值列于表 II 中。此外，工人的步长采用伽马分布建模，而转向角采用冯·米塞斯分布建模。将网络划分为长度为 μ 的微小区。正常工人从一个微小区移动到另一个微小区所需的平均帧数为：

$$\psi_{\text{nml}} = \lceil \frac{\mu}{TV_{\text{nml}}} \rceil, \tag{26}$$

其中，$\lceil \cdot \rceil$ 表示向上取整函数，V_{nml} 表示正常工人的平均速度。高风险工人从一个微小区移动到另一个微小区所经过的平均帧数为：

$$\psi_{\text{rsk}} = \lceil \frac{\mu}{TV_{\text{rsk}}} \rceil, \tag{27}$$

其中，V_{rsk} 是高风险工人的平均速度。显然，高风险工人切换微小区的频率比正常工人更高。此外，有些工人经过网络中的某些地点（如办公室、会议室、休息区.....）时，会在一段时间内保持静止。

考虑到这种移动场景，移动设备会四处移动，逐渐远离其关联的移动边缘计算服务器。因此，

表 II：隐状态的取值。

状态	速度（步/分钟）	转向角（弧度）
静止	0	角度 $< \pi/2$
正常	$0 <$ 步 ≤ 84	$\pi/2 \leq$ 角度 $< \pi$
高风险	步数 > 84	角度 $\geq \pi$

它们需要更多的无线传输能耗。事实上，在设备移动时采用固定设备关联方案会增加能耗成本。因此，我们需要设计一种自适应设备关联方案，并将设备的移动性考虑在内。

算法 2 动态 HFL 设备关联（DHDA）

对于 $l = 1 : L$，**执行**
  如果 l 模 $\psi_{\text{rsk}} = 0$ 或 l 模 $\psi_{\text{nrm}} = 0$ **则**，
    更新 Γ 发生变化的移动设备集合在当前帧期间的微小区。
    对于 $k \in \Gamma$ **执行**
      $i_k \leftarrow 0$，初始化
      $\theta_{\min} \leftarrow \theta_{\max}$，初始化
      对于 $m = 1 : M$ **执行**
        如果 $\beta_{m,\ k} \leq R_m$ **则**，验证移动设备是否设备 k 位于移动边缘计算 m的范围内。
          计算 θ
          如果 $\theta \leq \theta_{\min}$，**则**
            $i_k \leftarrow m$，选择散度最小的移动边缘计算。
            $\theta_{\min} \leftarrow \theta$
          结束如果
        结束 if
      结束 for
      如果 $i_k = 0$ **则**，如果移动设备 k 不在任何移动边缘计算的范围内。
        $i_k \leftarrow \text{argmin}_{m=1..M}\, \beta_{m,\ k}$，关联至距离最近的移动边缘计算 m_*。
        $R_{m*} \leftarrow \beta_{m*,\ k}$，更新移动边缘计算的范围移动边缘计算 m_*
      结束 if
      $\chi_{i_k,\ k} \leftarrow 1$
    结束 for
  结束 if
结束 for

我们提出一种启发式动态设备关联策略，用于由无线能量传输支持的异构网络中的分层联邦学习。首先，我们执行第 V 节中介绍的原始设备关联算法 H2RM。随后，正常工人每隔 ψ_{nml} 帧、高风险工人每隔 ψ_{rsk} 帧更换其关联的移动边缘计算服务器。移动设备 k 按如下方式更新其新关联的移动边缘计算服务器：首先找出范围大于其路径损耗的移动边缘计算服务器集合。我们已将移动边缘计算服务器 m 的范围 R_m 定义为与其关联的最远设备的路径损耗。因此，关联到这些移动边缘计算服务器之一不会产生额外的能量开销。该移动设备将选择散度 θ 最小的移动边缘计算服务器。如果该设备未能找到范围合适的移动边缘计算服务器

，则选择距离最近的移动边缘计算服务器。所提出的动态 HFL 设备关联算法（ DHDA ）如算法 2 所示。

为了计算 DHDA 的复杂度，我们需要将 H2RMA 第一个 **for** 循环的计算复杂度乘以我们重复移动设备 $\lceil \frac{L}{\psi_{\text{rsk}}} \rceil$ 关联的次数。因此，DHDA 的复杂度为：

$$C^{\text{DHDA}} = O(\lceil \frac{L}{\psi_{\text{rsk}}} \rceil MK + L(2K + M)). \tag{28}$$

VII. 数值结果

在本节中，我们研究所提出的分层联邦学习框架在不同网络配置下的性能。仿真设置如下：我们将中央单元放置在圆形宏小区的中心，并在所考虑的覆盖范围内均匀分布移动边缘计算服务器和设备。首先，我们考虑设备需要识别 0 到 9 之间的数字，这可以通过在 MNIST 数据集上训练分类模型来模拟。在仿真中，我们使用深度学习模型在 MNIST 数据集上进行训练。该模型由两个使用 3x3 滤波器的卷积层组成，并选择修正线性单元 (ReLU) 作为激活函数。为缩小生成的特征图，我们添加了一个最大池化层，其输出会被展平并馈送至后续的全连接层。最后一层用于生成 10 个类别的概率。我们指出，使用了丢弃法 [47] 来减少过拟合并对模型进行正则化。此外，我们在训练过程中纳入类别权重信息，以缓解类别不平衡问题；并使用简单的随机梯度下降 (SGD) 训练模型，在 20 个训练轮次内优化交叉熵损失。接下来，我们研究 CIFAR-10（加拿大高级研究所）数据集。它是一个知名图像数据集，常用于验证机器学习和计算机视觉算法的性能。该数据集包含 60000 张大小为 32x32 的彩色图像，分为 10 个类别；每个类别包含 6000 张图像。类别标签为：马、青蛙、狗、猫、鹿、鸟、汽车、飞机、卡车和船。各类别之间没有重叠，即类别互斥。我们在表 III 中汇总了本评估部分使用的仿真参数。

表 III：我们实验中异构网络上分层联邦学习环境的配置。

符号说明		值
N_{cu}	天线数量在 CU	128
M	移动边缘计算服务器的数量	8
N_{mec}	每个设备的天线数量 MEC	16
K	设备数量	20
L	帧数	50
C	类别数量	10
$B_{\max}$	最大电池容量	1000 焦耳
$B_k(0)$	初始电池电量	200 焦耳
ν	路径损耗指数	3.7
λ	带宽	20 MHz
ϑ	CPU时钟频率	10^9
ω	CPU周期数	40
ς	能耗系数	10^{-27}
$\theta_{\max}$	最大散度	0.5
	小区半径	200 米
	每条射频链的电路功率	30 dBm
	噪声功率谱密度	-174 dBm/Hz

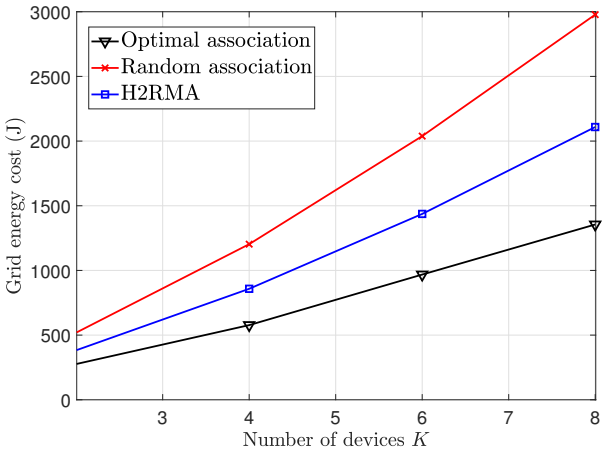

图4：随着设备数量变化的电网能耗成本K：最优设备关联、随机关联和 H2RMA 的比较（$M = 2$， $B_k(0) = 10$ J）。

A. 无调度的固定设备关联

我们首先研究所提出的分层联邦学习框架在静态网络和随时间保持固定设备关联的情况下的性能。在图4中，我们展示了设备数量变化时的电网能耗成本。求解最优设备关联方案具有较高的计算复杂度。因此，我们仅使用两个移动边缘计算服务器进行仿真，以建立最优解作为基准。与最优解相比，可以看出，H2RMA 在能耗成本方面表现良好，同时具有低得多的计算复杂度。在图5中，我们展示了随着设备数量 K 的变化，我们的启发式算法与随机设备关联方案相比所产生的电网能耗成本。使用大量

对于设备而言，H2RMA 凭借恰当的设备关联方法，显著降低了电网能耗成本。图6展示了改变移动边缘计算服务器数量时，我们的启发式算法的性能。随着移动边缘计算服务器数量增加，电网能源成本会降低，因为系统中的不同设备周围有更多移动边缘计算服务器可供选择，因此它们更有可能关联到距离较近的移动边缘计算服务器。此外，我们可以清楚地看到，对于不同数量的移动边缘计算服务器，我们的 H2RMA 算法均优于随机关联方案。

图7展示了 H2RMA 与随机关联方案相比的准确率性能，

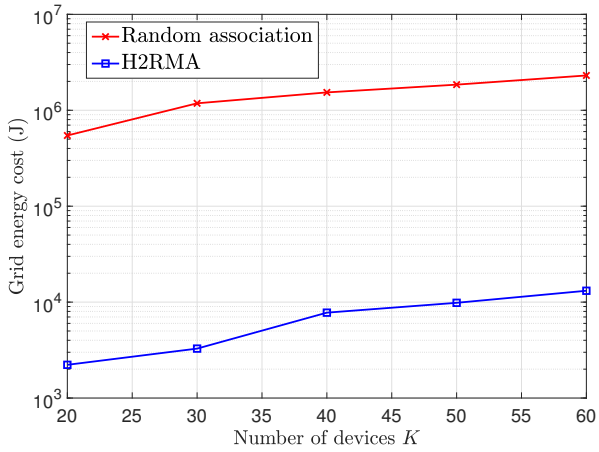

图5：设备数量变化时的电网能耗成本K：H2RMA与随机设备关联方案的比较（$M = 8$， $B_k(0) = 200$ J）。

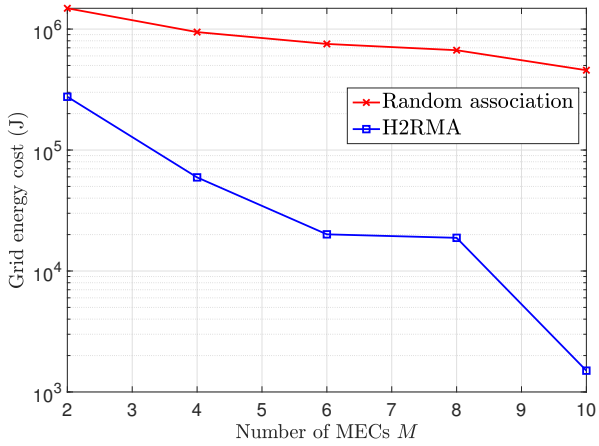

图6：移动边缘计算服务器数量变化时的电网能耗成本M：H2RMA与随机设备关联的比较（$K = 20$， $B_k(0) = 200$ J）。

测试使用的是 MNIST 数据集。可以看出，我们的算法收敛速度更快，这得益于其高效的设备关联过程，该过程能够最大限度地减少散度。图8评估了 H2RMA 在 CIFAR-10 数据集上的准确率性能。与 MNIST 数据集类似，我们提出的设备关联方案表现良好，在测试准确率方面优于随机设备关联方案。

B. 固定设备关联与调度

现在，我们在图 9 中研究所提出的设备调度对系统所消耗电网能量的影响。显然，在一定数量的帧内停用一些需要大量无线传输能耗的设备，可以节省大量能量。此外，与随机方案相比，所提出的设备调度方案表现非常出色，因为仅停用了距离较远的

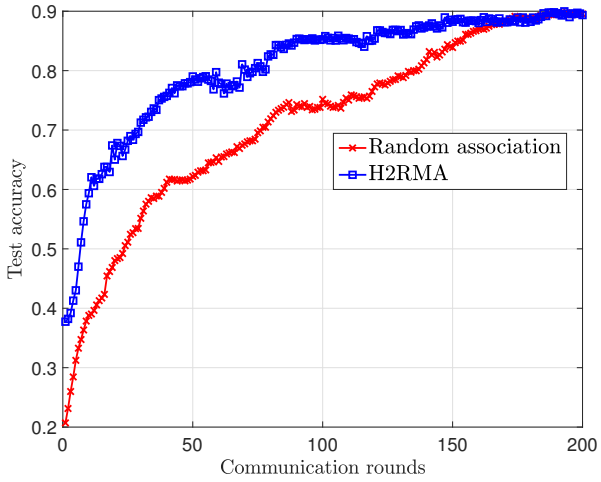

图7：测试准确率：H2RMA与随机设备关联方案在使用MNIST时的比较（K =15， $M = 3$， $C = 3$）。

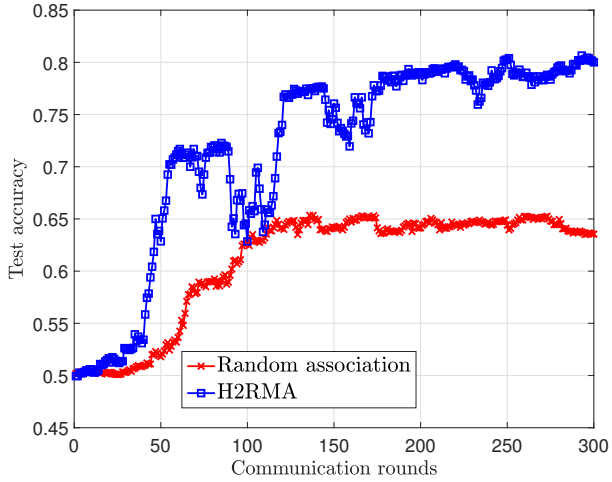

图8：测试准确率：H2RMA与随机设备关联方案在使用CIFAR-10时的比较（$K = 15$， $M = 3$， $C = 2$）。

设备，同时满足最大散度约束。

图10展示了在移动边缘计算服务器数量变化时，我们的设备调度启发式算法的性能。具体而言，得益于高效的设备调度方案，随着移动边缘计算服务器数量增加，我们能够节省大量电网能源。

图 11 展示了在 MNIST 数据集上，我们的启发式设备调度算法在测试准确率方面的表现。可以看出，随机停用部分设备会导致测试准确率下降。然而，所提出的启发式设备调度方案仅调度满足散度约束的设备子集，因此能够提高收敛速度。

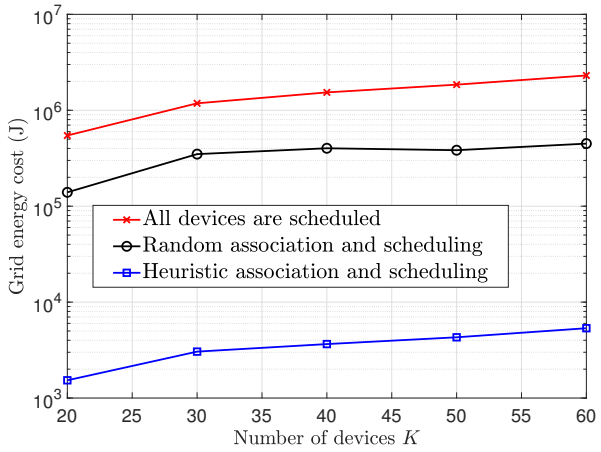

图 9：设备数量K变化时的电网能源成本：所提出的调度启发式算法、随机设备关联和所有设备的系统调度之间的比较（$M = 8,\ B_k(0) = 200$ J）。

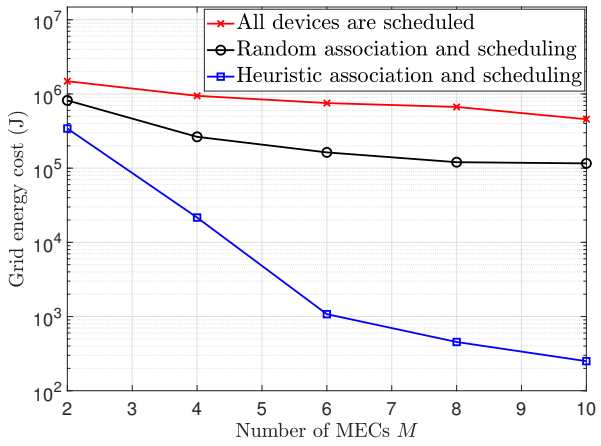

图 10：随着移动边缘计算服务器数量变化的电网能耗成本：对比所提出的调度启发式算法、随机设备关联和所有设备的系统调度（$K = 20,\ B_k(0) = 200$ J）。

C. 动态设备关联

最后，我们为设备设置了第 V 节所述的移动场景。本次仿真所用的移动参数如表 II 所示。图 12 评估了在移动场景下，我们的分层联邦学习框架所消耗的电网能源成本。实际上，所提出的动态设备关联算法能够显著节省能量，因为它可以高效地调整移动设备的关联。此外，随着网络中移动设备数量增加，动态设备关联与固定设备关联之间的性能差距也会扩大。

图 13 展示了在考虑移动场景的情况下，所提出的 HFL 框架在收敛速度方面的性能。所提出的动态设备关联方案确保移动设备遵守最大散度

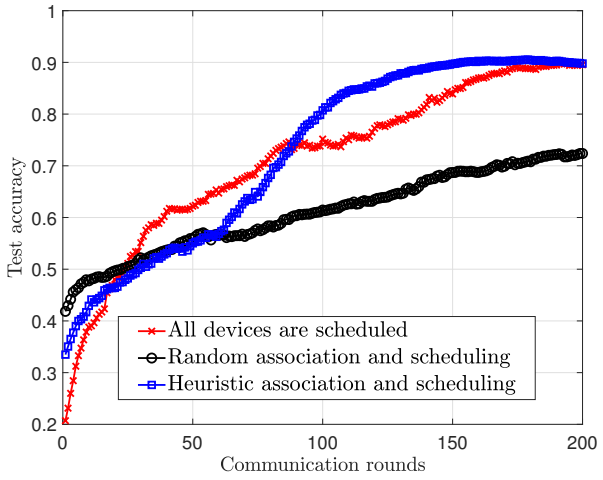

图 11：考虑 MNIST 时的测试准确率：所提出的调度启发式算法、随机设备关联和所有设备的系统调度之间的比较（$K = 15, M = 3, C = 3$）。

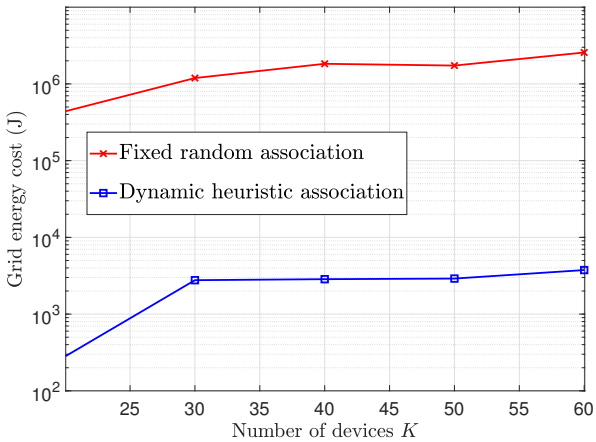

图 12：考虑移动场景时，随设备数量变化的电网能耗成本K（$M = 8,\ B_k(0) = 200$ J）。

约束，因此比固定设备关联实现了更快的收敛速度。

总之，我们通过展示所提出的资源管理方法的能源效率，以及在不同网络配置下使用两个知名数据集（即 MNIST 和 CIFAR-10）所达到的高准确率，评估了我们的分层联邦学习框架。此外，我们还展示了在网络移动场景下所提出的动态设备关联的效率。

VIII. 结论

本研究探讨了一种由无线能量传输（WET）支持、采用大规模MIMO无线回传的异构网络（HetNets）上的节能型分层联邦学习（HFL）。构建了一个优化问题，该问题涉及受能量可用性和最大散度约束的电网能耗成本。此外，还.....

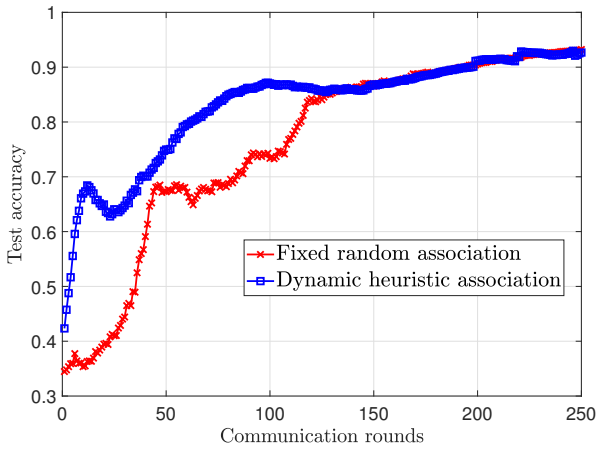

图 13：考虑移动场景和 MNIST 数据集时的测试准确率（$K = 15,\ M = 3,\ C = 3$）。

已推导出最优能量管理方案。为实现这一目标，我们将主要目标表述为一个设备关联优化问题，并使用广度优先搜索对其进行求解。由于获取最优解具有较高的复杂度，我们提出了一种高效的设备关联启发式算法，即 H2RMA，其计算复杂度较低。此外，我们设计了一种设备调度策略，以节省电网能源，同时不会降低学习准确率。我们还通过开发一种高效的动态设备关联方案，研究了移动性对联邦学习性能的影响。在不同网络配置下进行评估时，大量实验结果证实了我们的分层联邦学习框架在准确率、收敛速度和电网能耗方面的有效性。

作为未来工作，我们计划为支持无线能量传输的异构网络中的分层联邦学习开发一种基于强化学习的高效聚类策略。

致谢

本研究获得了卡塔尔国家研究基金（卡塔尔基金会成员）第十三（13）轮国家优先研究计划标准项目（NPRP-S）资助，项目编号为 NPRP13S-0205-200265。本文所述研究成果仅代表作者观点。

参考文献

[1] R. Hamdi, A. B. Said, A. Erbad, A. Mohamed, M. Hamdi, 和 M. Guizani，“无线能量传输赋能的异构网络层次化联邦学习”，载于 *IEEE* 全球通信大会（*GLOBE- COM*）论文集，西班牙马德里，2021年12月，第1-6页。[2] D. Roberts，“风险管理与电气安全：职业健康安全管理体系的实施”，*IEEE*《工业应用杂志》，第21卷，第3期，第67-74页，2015年2月。[3] A. Ranjan, Y. Zhao, H. B. Sahu, 和 P. Misra，“极端工业环境中的健康感知机遇与挑战：地下矿井的视角”，*IEEE Access*，第7卷，第139181-139195页，2019年9月。[4] J. A. Palazon, J. Gozalvez, J. L. Maestre, 和 J. R. Gisbert，“改善工业环境中健康与安全工作条件的无线解决方案”，载于 *IEEE*国际电子健康网络、应用与服务会议（*Healthcom*）论文集，葡萄牙里斯本，2013年10月，第544-548页。[5] E. Baccour, N. Mhaisen, A. A. Abdellatif, A. Erbad, A. Mohamed, M. Hamdi, 和 M. Guizani，“面向物联网应用的普适人工智能：资源高效的分布式人工智能”，*IEEE*《通信调查与教程》，第24卷，第4期，第2366-2418页，2022年8月。[6] M. Chen, D. Gündüz, K. Huang, W. Saad, M. Bennis, A. V. Feljan, and H. V. Poor，“无线网络中的分布式学习：近期进展与未来挑战”，IEEE J. Sel. Areas Commun.，第39卷，第12期，第3579-3605页，2021年12月。[7] W. Y. B Lim, N. C. Luong, D. T. Hoang, Y. Jiao, Y. C. Liang, Q. Yang, D. Niyato, and C. Miao，“移动边缘网络中的联邦学习：全面综述”，IEEE《通信调查与教程》，第22卷，第3期，第2031-2063页，2020年4月。[8] S. Luo, X. Chen, Q. Wu, Z. Zhou, and S. Yu，“HFEL：面向成本高效层次化联邦边缘学习的联合边缘关联与资源分配”，IEEE Trans. Wireless Commun.，第19卷，第10期，第6535-6548页，2020年6月。[9] X. Wang, S. Garg, H. Lin, J. Hu, G. Kaddoum, M. J. Piran, and M. S. Hossain，“利用层次化联邦学习实现工业物联网中的精确异常检测”，IEEE Internet Things J.，第9卷，第10期，第7110-7119页，2022年5月。[10] N. Mhaisen, A. Awad, A. Mohamed, A. Erbad, and M. Guizani，“基于统计特性和网络拓扑约束的层次化联邦学习最优用户-边缘分配”，IEEE Trans. Netw. Sci. Eng.，第9卷，第1期，第55-66页，2022年2月。[11] G. Yang, C. K. Ho, R. Zhang, and Y. L. Guan，“无线能量传输供能的大规模多输入多输出系统吞吐量优化”，IEEE J. Sel. Areas Commun.，第33卷，第8期，第1640-1650页，2015年8月。[12] J. Huang, Y. Zhou, Z. Ning, and H. Gharavi，“无线能量传输与能量采集：当前状况与未来展望”，IEEE Wireless Commun.，第26卷，第4期，第163-169页，2019年8月。[13] R. Hamdi, E. Driouch, and W. Ajib，“混合能源大规模多输入多输出系统中的能量管理”，IEEE Trans. Veh. Technol.，第66卷，第11期，第10183-10193页，2017年11月。[14] L. Liu, J. Zhang, S. H. Song, and K. B. Letaief，“客户端-边缘-云层次化联邦学习”，载于 IEEE Int. Conf. Commun. (ICC) 论文集，线上会议，2020年6月，第1-6页。[15] M. S. H. Abad, E. Ozfatura, D. Gunduz, and O. Ercetin，“异构蜂窝网络中的层次化联邦学习”，载于 IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP) 论文集，西班牙巴塞罗那，2020年5月，第1-5页。[16] W. Y. B. Lim, J. S. Ng, Z. Xiong, D. Niyato, C. Miao, and D. I. Kim，“自组织层次化联邦学习网络中的动态边缘关联与资源分配”，IEEE J. Sel. Areas Commun.，第39卷，第12期，第3640-3653页，2021年12月。[17] B. Xu, W. Xia, W. Wen, P. Liu, H. Zhao, and H. Zhu，“无线网络上的自适应层次化联邦学习”，IEEE Trans. Veh. Technol.，第71卷，第2期，第2070-2083页，2022年2月。[18] A. Wainakh, A. S. Guinea, T. Grube, and M. Mühlhäuser，“通过层次化联邦学习增强隐私保护”，载于 IEEE Euro. Symp. Security Privacy Workshops (EuroS&PW) 论文集，意大利热那亚，2020年9月，第1-4页。[19] A. A. Abdellatif, N. Mhaisen, A. Mohamed, A. Erbad, M. Guizani, Z. Dawy, and W. Nasreddine，“面向数据不平衡的物联网异构系统的通信高效层次化联邦学习”，Future Generation Computer Syst.，第128卷，第406-419页，2022年3月。[20] W. Y. B. Lim, J. S. Ng, Z. Xiong, J. Jin, Y. Zhang, D. Niyato, C. Leung, and C. Miao，“去中心化边缘智能：层次化联邦学习的动态资源分配框架”，IEEE Trans. Parallel Distrib. Syst.，第33卷，第3期，第536-550页，2022年3月。[21] S. R. Pandey, M. N. Nguyen, T. N. Dang, N. H. Tran, K. Thar, Z. Han, and C. S. Hong，“面向联邦分析的边缘辅助民主化学习”，IEEE Internet Things J.，第9卷，第1期，第572-588页，2022年1月。[22] H. Chai, S. Leng, Y. Chen, and K. Zhang，“一种用于车联网知识共享的区块链赋能层次化联邦学习算法”，IEEE Trans. Intell. Transp. Syst.，第22卷，第7期，第3975-3986页，2021年7月。[23] C. Briggs, Z. Fan, and P. Andras，“通过对本地更新进行层次聚类，改进非独立同分布数据上的联邦学习训练”，载于 Int. Joint Conf. Neural Netw. (IJCNN) 论文集，英国格拉斯哥，2020年7月，第1-9页。[24] L. Su and V. K. N. Lau，“面向多类型传感器间混合数据划分的层次化联邦学习”，IEEE Internet Things J.，第8卷，第13期，第10922-10939页，2021年7月。[25] Z. Wang, H. Xu, J. Liu, H. Huang, C. Qiao, and Y. Zhao，“边缘计算中采用层次化聚合的资源高效联邦学习”，载于 IEEE Conf. Comput. Commun. (INFOCOM) 论文集，线上会议，2021年5月，第1-9页。[26] J. Feng, L. Liu, Q. Pei, and K. Li，“无线边缘网络中高效层次化联邦学习的极小极大成本优化”，IEEE Trans. Parallel Distrib. Syst.，第33卷，第11期，第2687-2700页，2022年11月。

[27] R. Hamdi、M. Chen、A. B. Said、M. Qaraqe 和 H. V. Poor，“能量采集无线网络中的联邦学习”，*IEEE* 物联网期刊，第 9 卷，第 1 期，第 663-675 页，2022年1月。 [28] Z. Yang、M. Chen、W. Saad、C. S. Hong 和 M. Shikh-Bahaei，“无线通信网络中的节能联邦学习”，*IEEE* 无线通信汇刊，第 20 卷，第 3 期，第 1935-1949 页，2021年3月。 [29] M. Duan、D. Liu、X. Chen、R. Liu、Y. Tan 和 L. Liang，“移动系统中基于全局不平衡数据的自平衡联邦学习”，*IEEE* 并行与分布式系统汇刊，第 32 卷，第 1 期，第 59-71 页，2021年1月。 [30] S. Wang、T. Tuor、T. Salonidis、K. K. Leung、C. Makaya、T. He 和 K. Chan，“资源受限边缘计算系统中的自适应联邦学习”，*IEEE* 精选通信领域期刊，第 37 卷，第 6 期，第 1205-1221 页，2019年6月。 [31] H. H. Yang、Z. Liu、T. Q. Quek 和 H. V. Poor，“无线网络中联邦学习的调度策略”，*IEEE* 通信汇刊，第 68 卷，第 1 期，第 317-333 页，2020年1月。 [32] M. Chen、Z. Yang、W. Saad、C. Yin、H. V. Poor 和 S. Cui，“无线网络中联邦学习的联合学习与通信框架”，*IEEE* 无线通信汇刊，第 20 卷，第 1 期，第 269-283 页，2021年1月。 [33] J. Kang、Z. Xiong、D. Niyato、Y. Zou、Y. Zhang 和 M. Guizani，“移动网络中的可靠联邦学习”，*IEEE* 无线通信，第 27 卷，第 2 期，第 72-80 页，2020 年 4 月。 [34] Q. V. Pham 、 M. Le 、 T. Huynh-The、Z. Han 和 W. J. Hwang，“无人机赋能的无线供能通信中的节能联邦学习”，*IEEE* 车辆技术汇刊，第 71 卷，第 5 期，第 4977-4990 页，2022年2月。 [35] J. Feng、W. Zhang、Q. Pei、J. Wu 和 X. Lin，“无线供能联邦边缘学习系统中的异构计算与资源分配”，*IEEE* 通信汇刊，第 70 卷，第 5 期，第 3220-3233 页，2022年3月。[36] M. Arslan、C. Cruz 和 D. Ginhac，“无线能量传输网络中用于同时信息传输和联邦学习的无电池海量接入”，载于 *IEEE* 计算机通信会议研讨会论文集（*INFOCOM WKSHPS*），纽约，美国，2022年5月，第 1-2 页。

[37] Q. Zeng、Y. Du 和 K. Huang，《无线供电联邦边缘学习：收敛与能量传输之间的最优权衡》，*IEEE* 无线通信汇刊，第 21 卷，第 1 期，第 680–695 页，2022 年 1 月。[38] T. E. Bogale 和 L. B. Le，《面向 5G 无线异构网络的大规模 MIMO 与毫米波》，*IEEE* 车辆技术杂志，第 11 卷，第 1 期，第 64–75 页，2016 年 3 月。[39]R. Hamdi、E. Driouch 和 W. Ajib，《一种适用于采用大规模 MIMO 无线回传的异构网络的新型高效传输技术》，*IEEE* 车辆技术汇刊，第 69 卷，第 1 期，第 663–675 页，2020 年 1 月。[40] O. Somekh、O. Simeone、Y. Bar-Ness、A. M. Haimovich 和 S. Shamai，《蜂窝下行信道中的协作式多小区迫零波束成形》，*IEEE* 信息论汇刊，第 55 卷，第 7 期，第 3206–3219 页，2009 年 7 月。[41] W. Ni 和 X. Dong，《面向大规模多用户 MIMO 系统的混合块对角化》，*IEEE* 通信汇刊，第 64 卷，第 1 期，第 201–211 页，2016 年 1 月。[42] T. X. Tran 和 K. C. Teh，《采用块对角化的大规模多用户多输入多输出系统的性能分析》，*IET* 通信，第 10 卷，第 7 期，第 832–838 页，2016 年 5 月。[43] M. Chen、Z. Yang、W. Saad、C. Yin、H. V. Poor 和 S. Cui，《面向无线网络联邦学习的联合学习与通信框架》，*IEEE* 无线通信汇刊，第 20 卷，第 1 期，第 269–283 页，2021年。[44] Y. Che、L. Duan 和 R. Zhang，《大规模绿色蜂窝网络中的基站动态运行》，*IEEE* 选定通信领域期刊，第 34 卷，第 12 期，第 3127–3141 页，2016 年 12 月。[45] Z. Q. Luo 和 S. Zhang，《动态频谱管理：复杂度与对偶性》，*IEEE* 选定信号处理主题期刊，第 2 卷，第 1 期，第 57–73 页，2008 年 2 月。[46] M. Arslan、C. Cruz 和 D. Ginhac，《利用语义轨迹上的隐马尔可夫模型理解停留地点内的工人移动性》，发表于 *IEEE* 新兴技术国际会议（*ICET*）论文集，巴基斯坦伊斯兰堡，2018 年 11 月，第 1–6 页。[47] N. Srivastava、G. Hinton、A. Krizhevsky、I. Sutskever 和 R. Salakhutdi- nov，《Dropout：一种防止神经网络过拟合的简单方法》，机器学习研究期刊，第 15 卷，第 1 期，第 1929–1958 页，2014 年 6 月。