

无蜂窝网络中的联邦学习用户调度与资源优化

王华华, 黄烨霞*, 李玲, 王嘉程

(重庆邮电大学通信与信息工程学院, 重庆 400065)

摘要: 在无蜂窝网络环境下实施联邦学习(FL)时,用户调度和资源分配策略对优化系统时间开销、提升用户可达速率以及加速 FL 收敛速率至关重要。为解决资源分配不均的问题,设计一种联合用户调度、CPU 处理频率和功率分配的优化方案。通过最大化系统的最小用户速率来实现资源的公平分配,并提升 FL 性能。联合优化问题被分解为用户调度和功率分配两个子问题。在用户调度方面,设计基于 k -means 聚类的贪婪调度算法,以综合评估用户的信道状态和数据“价值”,并将用户划分为不同的群组。随后,针对每个群组的资源占用情况,为组内用户制定个性化的 CPU 处理频率分配方案。最后,通过在各群组中独立执行用户调度,实现高效且精准的用户选择,并通过提前分组有效降低用户选择的复杂度。在功率分配方面,引入基于二分法的功率分配算法(BM-PA)。该算法不仅考虑了用户间的公平性,还针对资源受限用户进行了优先处理,以确保其能够获得更优质的资源分配。BM-PA 算法通过低复杂度的迭代优化过程,实现了功率分配的快速收敛,并在保证系统性能的同时,显著提升了资源的利用效率。合理的用户调度策略是功率分配子问题获得最优解的基础,采用交替迭代的方法允许在每个子问题中独立进行优化,同时考虑到另一个子问题的解。这种相互依赖的关系通过多轮迭代优化过程,确保功率资源被合理地分配给那些最需要或最有可能有效利用它们的用户,从而使系统整体性能得到提升,实现联合优化求解,显著提升系统整体性能。仿真实验结果表明,与基准算法相比,所提算法在下行可达速率方面,最佳平均提升幅度高达 103.34%,在上行可达速率方面,最佳提升幅度达到 102.78%。此外,相较于基准算法还能平均节省 67.44% 的 FL 任务训练时间,特别是在 FL 学习模型精度达到 90% 时,所提算法的时间开销最小。

关键词: 无蜂窝网络;联邦学习;用户调度; k -means 聚类;资源优化

中图分类号: TN929.5

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0069877

Federated Learning User Scheduling and Resource Optimization in Cell-Free Networks

WANG Huahua, HUANG Yexia*, LI Ling, WANG Jiacheng

(School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

【Abstract】 When implementing Federated Learning (FL) in a cell-free network environment, user scheduling and resource allocation strategies are crucial for optimizing system time overhead, improving user reachability, and accelerating FL convergence rate. To address the issue of uneven resource allocation, this study designs an optimization scheme that combines user scheduling, CPU processing frequency, and power allocation. This scheme aims to achieve fair resource allocation by maximizing the minimum user rate in the system, thus enhancing FL performance. The joint optimization problem is decomposed into two subproblems: user scheduling and power allocation. For user scheduling, this study proposes a greedy scheduling algorithm based on k -means clustering to comprehensively evaluate channel conditions and data "value" of users and categorize users into different groups. Subsequently, for the resource occupation situation, a personalized CPU processing frequency allocation plan is developed for users within each group based on their resource occupancy. Finally, by independently executing user scheduling within each group, user selection is performed efficiently and precisely, and the complexity of user selection is effectively reduced via early grouping. For power allocation, this study introduces a Bisection Method-based Power Allocation (BM-PA) algorithm. This algorithm not only considers fairness among users but also prioritizes resource-constrained users to ensure that they can obtain superior resource allocation. The BM-PA algorithm achieves fast convergence of power allocation using a low-complexity iterative optimization process, significantly improving the resource utilization efficiency without deteriorating the system performance. In this study, a reasonable user scheduling strategy serves as the foundation for obtaining optimal solutions for the power allocation subproblem. This study adopts an alternating iteration method that allows independent optimization in

收稿日期: 2024-05-21 修回日期: 2024-06-21

基金项目: 重庆市自然科学基金创新发展联合基金(中国星网)(CSTB2023NSCQ-LZX0114)。

通信作者 E-mail: * huangyexia2022@163.com

each subproblem while considering the solution of the other subproblem. Via multiple rounds of iterative optimization, this interdependent relationship ensures that power resources are reasonably allocated to users who need them the most or are most likely to effectively utilize them, thus enhancing the overall system performance. This study realizes joint optimization solutions that significantly improve overall system performance. Simulation results show that compared with the baseline algorithm, the proposed algorithm exhibits outstanding performance in terms of downlink achievable rates-the average improvement reaches up to 103.34% under optimal conditions. Additionally, the uplink achievable rates improve by up to 102.78%. Furthermore, the proposed algorithm can save 67.44% of the FL task training time on average compared to the baseline algorithm, particularly when the FL learning model accuracy reaches 90%, wherein the time overhead of the proposed algorithm is minimal.

【Key words】 cell-free network; Federated Learning (FL); user scheduling; k -means clustering; resource optimization

0 引言

近年来,人工智能技术在无线通信领域蓬勃发展,尤其是联邦学习(FL)技术^[1-2]。作为一种新兴的分布式机器学习框架,FL的核心优势在于能够在本地设备上完成模型训练,并仅将模型参数上传至中央服务器进行聚合,从而避免了原始数据的直接共享,能有效保护用户隐私^[3]。然而,FL任务中设备间频繁的模型参数交换对通信链路的稳定性和可靠性提出了较高要求。在此背景下,无蜂窝大规模多输入多输出(CF mMIMO)网络^[4-5]作为一种潜在的解决方案,受到了广泛关注。CF mMIMO网络通过在有限区域内部署大量网络接入点(AP),为FL提供了高质量的通信链路,确保了模型迭代过程中的通信稳定性。将CF mMIMO与联邦学习技术相结合,可以在不加重数据传输负担的前提下,充分利用分布式数据进行高效的模型训练,从而提高通信效率并降低训练过程中的总通信成本^[6]。然而,随着移动设备数量增加,“落伍者”效应出现的概率也增加,成为提升FL性能的挑战。因此,在众多用户设备(UE)中选择合适的子集并合理分配资源,是CF mMIMO网络中应用FL时需要解决的关键问题。

许多专家学者已经对无线网络中的用户调度问题进行了广泛研究。比如,XIAO等^[7]提出在线UE选择方案,以最小化训练时间并确保模型收敛性。丁青锋等^[8]将有效信道与干扰信道增益的占比作为排序指标进行AP选择,实现了系统的能效均衡。温依霖等^[9]提出了ChFL客户端选择方法,通过分析客户端本地模型损失值和训练时间,评估客户端模型质量,以指导客户端选择,提高全局模型精度和收敛速度。陶梅霞等^[10]考虑用户的计算能力和信道状态,将调度优化问题转化为一个一维背包问题,并利用动态规划算法搜索背包空间,算法复杂度较高。尽管上述用户调度策略在一定程度上提升了系统性能,但它们也存在偏向

性。例如,ChFL策略倾向于选择具有高质量数据和较高计算能力的客户端参与训练,这可能导致数据质量较低或计算能力较弱的客户端长期未被选中,从而影响公平性。

值得注意的是,仅依靠UE采样技术难以完全有效避免“落伍者”效应,还需要综合考虑无线网络中FL的资源调度与分配问题^[11-13]。为了解决功率控制问题,FAROOQ等^[11]提出一种将原始非凸问题转化为凸规划问题的方法,并结合Nesterov平滑技术和加速投影梯度法进行求解,但是该研究只考虑了上行链路的功率控制,未涉及下行链路。相关研究^[12-13]则深入探讨了联合带宽分配和设备调度问题,以最大化时间的收敛速率。NGO等^[14]提出一种基于回程链路功耗和频谱效率的AP选择方案,有效提升了总能量效率。然而,单纯地追求每轮迭代中UE数量的增加可能会损害FL模型的精度。为提高收敛速度,CHEN等^[12]提出概率用户选择方案,并利用神经网络建立用户本地FL模型参数之间的关系。虽然这些方法在缩短收敛时间和提高准确性方面取得了进展,但其算法复杂度可能限制了其广泛应用。此外,CHEN等^[15]将FL过程中的功率、无线带宽资源分配和用户选择问题构建为一个优化问题,通过最小化FL损失函数来衡量FL算法的性能。然而,这些算法在无蜂窝网络系统中的实际应用效果仍有待进一步验证。

在近期的研究中,学术界开始关注在无蜂窝网络环境中实现FL的有效网络架构问题^[16-18]。具体来讲,VU等^[16]对模型精度、发射功率、数据速率以及用户CPU的处理频率进行了联合优化,然而,该研究并未充分考虑到用户调度策略对系统整体效率的影响。姚媛媛等^[17]提出基于信道排序的较差用户优先接入策略,该策略优先考虑信道条件较差的用户,以实现系统总能耗的最小化。VU等^[19]建模了联合优化用户选择、功率控制和数据速率以最小化FL通信轮次的执行时间。另

一方面, GUO 等^[20]以最小化总发射功率为目标, 建立了联合功率分配和用户分组问题的数学模型。为了求解该问题, 本文提出一种基于广义 Benders 分解 (GBD) 的算法, 通过迭代求解主问题和原问题来逼近最优解。然而, 该算法采用基于时间槽的用户分组方法, 每个用户组仅能在单一时间槽内通信, 这可能导致时间资源的利用率不高。此外, 优化总发射功率可能会使得某些用户获得过多资源, 而其他用户则面临资源不足的问题。但本文基于时间槽的用户分组方式, 每个组只能使用一个时间槽, 这可能导致时间资源利用率低。且最小发射功率的优化可能会导致某些用户获得过多的资源, 而其他用户则资源不足。上述研究主要集中于选择具有有利链路条件的用户, 这种做法可能对信道质量较差的用户不够友好, 并且在资源分配的公平性方面考虑不足。

综上所述, 目前在无蜂窝网络系统中执行联邦学习任务的用户调度与资源分配策略的研究存在不足。现有策略往往侧重于信道状态良好的用户, 而忽略了对信道状态较差用户的合理资源分配, 导致部分用户过度占用网络资源, 进而降低了网络整体服务质量和联邦学习任务的性能。鉴于以上分析, 本文探讨了无蜂窝网络系统中执行 FL 任务的资源调度与分配问题, 并提出一种综合考虑用户选择、CPU 处理频率和功率分配的优化框架。

本文的主要贡献包括:

1) 在 FL 框架下, 建立了针对无蜂窝网络系统的时间消耗模型, 并基于 k -means 聚类方法对用户进行分组调度, 同时分配合适的 CPU 处理频率。该策略综合考虑用户的信道质量和数据重要性, 旨在提高系统资源分配的效率和公平性。

2) 基于上述调度策略, 提出 BM-PA 功率分配算法, 在保持低复杂度的同时, 实现对已关联用户上行和下行链路的功率分配优化, 从而提升系统速率, 并减少 FL 学习的总时耗。

3) 通过与其他基准算法的对比, 所提出的算法最大可提升上行和下行可达速率分别达到 102.78% 和 103.34%, 并减少了 67.44% 的 FL 学习时间开销。这为无蜂窝网络系统中的 FL 任务执行提供了有效的优化策略。

1 系统模型

本文考虑的是采用时分多址 (TDMA) 模式的无蜂窝网络系统, 其中大小为 M 的 AP 集 $\mathcal{M} \triangleq \{1, 2, \dots, M\}$ 为大小为 K 的 UE 集 $\mathcal{K} \triangleq \{1, 2, \dots, K\}$ 提供服务。AP 和 UE 均匀随机地分布在一个 $D \times D \text{ m}^2$ 的区域内, 且均为单天线。用户通过无线链路与 AP 连接, 而 AP 则通过高容量的回程链路连接到 CPU。因此, 本文不考虑 CPU 与所有 AP 之间的传输时延。具体的系统模型如图 1 所示。

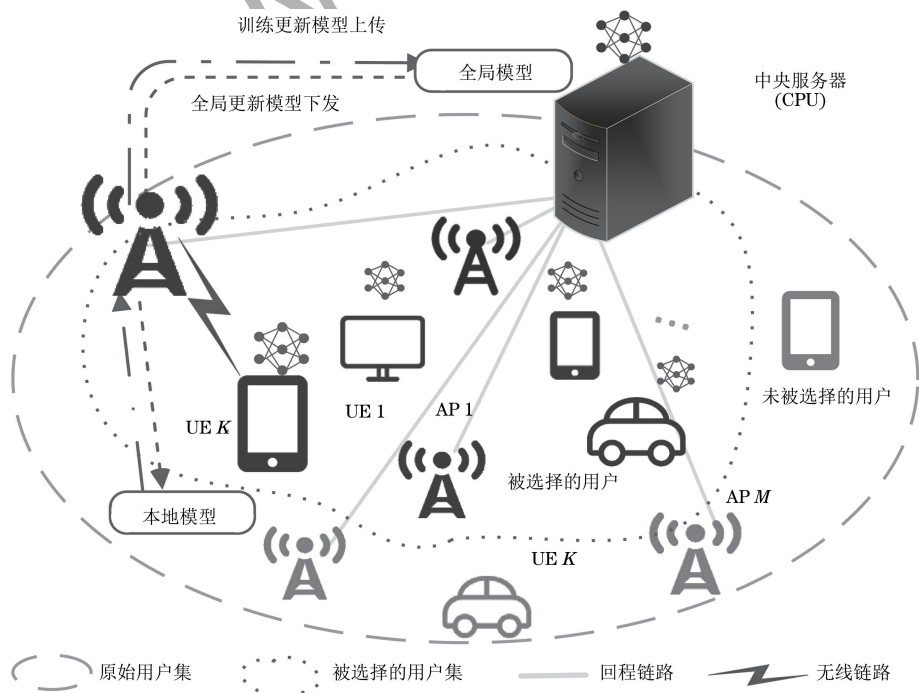


图 1 基于 FL 的 CF mMIMO 无线网络模型

Fig.1 CF mMIMO wireless network model based on FL

在一个 FL 训练任务中, CPU 充当中央服务器, UE 充当客户端, 而 AP 则在 CPU 和 UE 之间作为中继传递模型的更新参数。定义一个二元关联变量 a_k , 指示 UE k 是否被关联, 假设 $\mathbf{U}$ 为参与 FL 过程的用户集。则有, 当 $k \in \mathbf{U}$ 时 $a_k = 1$, 否则 $a_k = 0$ 。

在无蜂窝网络系统中, 被关联的 UE 每一个通信回合 FL 训练学习任务包含以下 4 个基本步骤^[21]:

步骤 1(S1) 模型下载。 CPU 执行用户调度算法以确定本轮的调度策略, 并将该策略和全局更新模型参数以广播形式下发给所有 AP。然后, AP 将全局更新模型参数转发给需要关联的 UE。

步骤 2(S2) 模型更新。 关联 UE 接收到全局模型参数, 并利用本地数据集计算并更新本地模型参数。

步骤 3(S3) 模型上传。 所有关联 UE 向 AP 上传其更新的本地训练模型参数, 并报告其数据集大小、存储容量、CPU 处理频率等状态信息, 以支撑服务器进行用户选择。最后, AP 将从 UE 获取的所有数据上传至服务器。

步骤 4(S4) 模型聚合。 CPU 在接收到所有 AP 发来的数据后, 采用 FedAvg 算法^[22] 更新全局模型参数。

以上步骤不断迭代执行, 直到达到最大迭代次数时终止整个 FL 任务。每轮 FL 的训练都在一个大尺度相干时间内完成。所有关联 UE 以同步的方式执行 S2 和 S3。聚合工作在接收到所有关联 UE 的本地更新模型参数后才进行。

$$R_{d,k}(a_k, \eta) = \frac{\tau_c - \tau_p}{\tau_c} B^D \log \left(1 + \frac{p_d \left(\sum_{k \in \mathbf{K}} \sum_{m \in \mathbf{M}} a_k \sqrt{\eta_{mk}} \sigma_{mk}^2 \right)^2}{p_d \sum_{k' \in \mathbf{K} \setminus k} a_{k'} A_{kk'}^2 |\varphi_{k'}^H \varphi_k|^2 + p_d \sum_{k \in \mathbf{K}} a_k B_k + 1} \right) \quad (4)$$

式中: B^D 为下行传输带宽; $A_{kk'} = \sum_{m=1}^M \sqrt{\eta_{mk'}} \sigma_{mk'}^2 (\beta_{mk} / \beta_{mk'})$; $B_k = \sum_{m=1}^M \eta_{mk} \sigma_{mk}^2 \beta_{mk}$ 。

1.2 上行数据传输模型

在上行数据传输阶段, 令 $s_{u,k}$ 表示 UE k 的本地更新模型参数编码符号, 则 UE k 处需要上传的基带信号表示为:

$$x_{u,k} = \sqrt{p_u \zeta_k} s_{u,k} \quad (5)$$

式中: p_u 是 UE k 处的最大归一化发射功率; ζ_k 是上行功率控制因子。

UE 的发送信号也需要满足平均发射功率约束: $\mathbb{E}\{|x_{u,k}|^2\} \leq p_u$, 即 $0 \leq \zeta_k \leq 1, \forall k$ 。则 AP m

1.1 下行数据传输模型

在上传载荷数据之前, UE 首先向 AP 发送长度为 $\tau_p \ll \tau_c$ 的导频序列, τ_c 表示每个相干块能够传送的样本数量。每个 AP 利用接收到的导频序列采用最小均方误差估计 (MMSE)^[23] 方法对 $\mathbf{g}_{mk}$ 进行估计得到估计值 $\bar{\mathbf{g}}_{mk} \sim \text{CN}(0, \sigma_{mk}^2)$, 其中 $\mathbf{g}_{mk} = \sqrt{\beta_{mk}} h_{mk}$, 表示从 UE k 到 AP m 的信道, σ_{mk}^2 的具体表达式^[5]如式(1)所示:

$$\sigma_{mk}^2 = \frac{\tau_p p_p (\beta_{mk})^2}{\tau_p p_p \sum_{i \in \mathbf{K}} \beta_{mi} |\varphi_k^H \varphi_i|^2 + 1} \quad (1)$$

式中: β_{mk} 和 h_{mk} 分别代表信道大尺度和小尺度衰落系数。

接着, AP 使用共轭波束成型技术向用户发送数据信号。令 $s_{d,k}$ 表示 CPU 需要下发的全局模型参数编码符号, 则在 AP m 处的发送信号可用式(2)表示:

$$x_{d,m} = \sqrt{p_d} \sum_{k \in \mathbf{K}} \sqrt{\eta_{mk}} (\bar{\mathbf{g}}_{mk}) \times s_{d,k} \quad (2)$$

式中: p_d 是每个 AP 处的最大归一化发射功率; $0 \leq \eta_{mk} \leq 1$ 是下行功率控制因子。AP m 处的发射功率需要满足平均归一化功率约束 $\mathbb{E}\{|x_{d,m}|^2\} \leq p_d$, 即 $\sum_{k \in \mathbf{K}} \sigma_{mk}^2 \eta_{mk} \leq 1, \forall m$, 则在 UE k 处的接收信号可由式(3)给出:

$$y_{d,k} = \sqrt{p_d} \sum_{m \in \mathbf{M}} \sqrt{\eta_{mk}} \mathbf{g}_{mk} (\bar{\mathbf{g}}_{mk}) \times s_{d,k} + \sqrt{p_d} \sum_{m \in \mathbf{M}} \sum_{k' \in \mathbf{K} \setminus k} \sqrt{\eta_{mk'}} \mathbf{g}_{mk'} (\bar{\mathbf{g}}_{mk'}) \times s_{d,k'} + n_k \quad (3)$$

式中: n_k 是 UE k 的加性噪声。UE k 的下行可达速率如式(4)所示:

处的接收信号可由式(6)表示:

$$y_{u,m} = \sqrt{p_u} \sum_{k \in \mathbf{K}} \sqrt{\zeta_k} \mathbf{g}_{mk} s_{u,k} + n_{u,m} \quad (6)$$

式中: $n_{u,m}$ 是 AP m 处的加性噪声。为了检测从第 k 个用户发送的符号 $s_{u,k}$, AP 将接收信号 $y_{u,m}$ 与其信道估计 $\bar{\mathbf{g}}_{mk}$ (在上行信道估计时获得) 的共轭相乘, 然后通过回程链路发送到 CPU, 则 CPU 处的接收信号可由式(7)表示:

$$y_{u,k} = \sqrt{p_u} \sum_{m \in \mathbf{M}} \sqrt{\zeta_k} \mathbf{g}_{mk} (\bar{\mathbf{g}}_{mk}) \times s_{u,k} + \sqrt{p_u} \sum_{m \in \mathbf{M}} \sum_{k' \in \mathbf{K} \setminus k} \sqrt{\zeta_{k'}} \mathbf{g}_{mk'} (\bar{\mathbf{g}}_{mk'}) \times s_{u,k'} + \sum_{m \in \mathbf{M}} (\bar{\mathbf{g}}_{mk}) \times n_{u,m} \quad (7)$$

因此, UE k 可实现的上行传输速率为:

$$R_{u,k}(a_k, \zeta) = \frac{\tau_c - \tau_p}{\tau_c} B^U \ln \left(1 + \frac{p_u a_k \zeta_k \left(\sum_{m \in \mathcal{M}} \sigma_{mk}^2 \right)^2}{p_u \sum_{k' \in \mathcal{K} \setminus k} a_{k'} P_{kk'}^2 |\varphi_k^H \varphi_{k'}|^2 + p_u \sum_{k \in \mathcal{K}} a_k V_k + W_k} \right) \quad (8)$$

式中: B^U 为上行传输带宽; $P_{kk'} = \sum_{m=1}^M \zeta_k' \sigma_{mk}^2 (\beta_{mk'} / \beta_{mk})$; $V_k = \sum_{m=1}^M \zeta_k \sigma_{mk}^2 \beta_{mk}$; $W_k = \sum_{m=1}^M \sigma_{mk}^2$ 。

1.3 时间消耗模型

对系统时间消耗模型进行建模。在有用户调度策略的情况下,假设全局训练模型参数和本地更新模型参数的数据大小分别为 S_d 和 S_u (bit),则下行和上行的传输时延模型分别为:

$$t_{d,k}(a_k, \eta) = \frac{a_k S_d}{R_{d,k}(a_k, \eta)} \quad (9)$$

$$t_{u,k}(a_k, \zeta) = \frac{a_k S_u}{R_{u,k}(a_k, \zeta)} \quad (10)$$

式中: L 表示 UE 本地模型参数更新需要的迭代总次数; c 表示 UE k 处理一个数据样本需要的周期数; f_k 表示用户 k 的 CPU 处理频率 (周期/s); D_k 代表 UE k 的数据集大小。UE k 计算更新的时延如式(11)所示:

$$t_{c,k}(a_k) = \frac{a_k L D_k c}{f_k} \quad (11)$$

在每个通信回合中,根据 AP 所处的传输模式可将其分为上行传输阶段和下行传输阶段。在下行传输阶段,所耗费的时间取决于最后完成 S1 和 S2 的 UE 所需的时间,而在上行传输阶段,时延则取决于最后完成 S3 的 UE 所需的时间。因此,一个通信回合总时间消耗 T_o 可由式(12)、式(13)和式(14)求得:

$$T_d(a, \eta) = \max_{k \in \mathcal{K}} \left\{ \frac{a_k S_d}{R_{d,k}(a_k, \eta)} + \frac{a_k L D_k c}{f_k} \right\} \quad (12)$$

$$T_u(a, \zeta) = \max_{k \in \mathcal{K}} \frac{a_k S_u}{R_{u,k}(a_k, \zeta)} \quad (13)$$

$$T_o(a, \eta, \zeta) = T_d(a, \eta) + T_u(a, \zeta) \quad (14)$$

2 算法设计

在基于联邦学习框架下的无蜂窝网络系统中,本文旨在最大化用户的最小可达速率。针对联邦学习时间约束,设计一个综合考虑联合用户调度、CPU 频率分配以及上下行功率分配的优化问题。此问题的关键在于寻找一个最佳的 UE 子集,并为该子集制定合理的资源分配方案,通过相应的资源分配策略尽可能地实现每个 UE 被列入子集的概率相当,从而在确保网络质量和联邦学习性能的同时,尽可能地体现公平性。

整个联合优化问题可建模为如式(15)所示的数学表达式:

$$\begin{aligned} & \max_{a, \eta, \zeta, f} \min_{k \in \mathcal{U}} (R_{d,k}(a, \eta) + R_{u,k}(a, \zeta)) \\ & \text{s. t. } C_1: \frac{a_k S_d}{R_{d,k}(a, \eta)} + \frac{a_k L D_k c}{f_k} + \frac{a_k S_u}{R_{u,k}(a, \zeta)} < T_{\text{round}} \\ & C_2: a_k = \begin{cases} 1, & k \in \mathcal{U} \\ 0, & k \notin \mathcal{U} \end{cases} \\ & C_3: f_{\min} \leq f_k \leq f_{\max} \\ & C_4: 0 \leq \zeta_k \leq 1 \\ & C_5: \sum_{k \in \mathcal{K}} \sigma_{mk}^2 \eta_{mk} \leq 1 \\ & C_6: 0 \leq \eta_{mk} \end{aligned} \quad (15)$$

式中: T_{round} 代表一个 FL 通信回合允许的最大执行时间,超过最大执行时间的轮次将被视为无效的通信回合; $f_{\min}$ 和 $f_{\max}$ 分别是用户最小和最大的 CPU 处理频率。

式(15)具有非凸随机性、混合整数、二元约束和变量间紧耦合等特点。为了解决这一优化问题,将其分解为用户调度子问题和功率分配子问题,并采用交替优化的方式来解决。在式(15)中,UE 选择变量在 a 和用户 CPU 处理频率 f 在执行 FL 过程之前被优化,并在一个相干时间内保持不变,而功率控制因子 η 和 ζ 则在每个 FL 通信回合中根据 a 和 f 结果进行优化,经过多轮迭代优化最终得到最优的结果 η^* 和 ζ^* 。

2.1 基于 k -means 聚类的用户调度算法

本节对支持联邦学习的无蜂窝网络系统中的用户调度策略进行了详细阐述。该策略分为分组和选择两个环节。首先,根据链路状态和持有训练数据集的质量,采用 k -means 聚类^[24]将所有用户分为“优”、“良”和“差”三组。然后,在每组用户中以最大化平均可达速率为目标进行用户调度和 CPU 处理频率分配。这一策略的关键在于确保即使在链路质量较差的情况下,用户也有机会参与联邦学习过程,并得到优化的 CPU 处理频率分配,从而增加下次参与联邦学习的机会。

首先定义一个调度二元变量 $b_{g,k} \in \{0, 1\}$:

$$b_{g,k} = \begin{cases} 1, & k \in \Omega_g \\ 0, & k \notin \Omega_g \end{cases} \quad (16)$$

式中: Ω_g 代表第 g 组中参与 FL 的用户子集,并且有 $\mathcal{U} \triangleq \{\Omega_g, g=1, 2, \dots, G\}$, G 代表总的分组数目

(本文设置为 3)。具有比例公平性的用户分组与调度可以概括为:

$$\begin{aligned} & \max_{b_{g,k}, f_k} \sum_{k=1}^K \sum_{g=1}^G b_{g,k} \ln \left(\frac{\bar{r}_{g,k}}{\sum_i b_{g,i}} \right) \\ & \text{s. t. } C_1: \frac{b_{g,k} S_d}{R_{d,k}(b_{g,k})} + \frac{b_{g,k} L D_k c}{f_k} + \frac{b_{g,k} S_u}{R_{u,k}(b_{g,k})} < \\ & T_{\text{round}}, \forall g \\ & C_2: b_{g,k} = \begin{cases} 1, & k \in \Omega_g \\ 0, & k \notin \Omega_g \end{cases} \\ & C_3: f_{\min} \leq f_k \leq f_{\max} \end{aligned} \quad (17)$$

式中: $\bar{r}_{g,k}$ 是用户的平均吞吐量, 即 $\bar{r}_{g,k} = \ln(1 + \bar{S}_{g,k}^{\text{sinr}})$, $\bar{S}_{g,k}^{\text{sinr}}$ 是用户 k 被分配到组 $\Omega_{m,g}$ 时的平均信号与干扰加噪声比(SINR), 两个不同方向的 $\bar{S}_{g,k}^{\text{sinr}}$ 与式(4)和式(8)具有相同的表达形式。

1) 用户分组。采用 k -means 聚类, 并根据用户信道链路状态和持有训练数据的质量进行分类。信道链路的好坏用 g_{mk} 来衡量, 并引入指标 q_k 来衡量 UE k 持有数据的“重要度”:

$$q_k = \frac{D_k}{D} \cdot \frac{1}{\text{Var}(\mathcal{D}_k)} \quad (18)$$

式中: D_k 和 D 分别表示 UE k 数据集大小和系统总数据集大小; $\text{Var}(\cdot)$ 代表方差大小。用户分组具体算法步骤如算法 1 所示。

算法 1 基于改进 k -means 的用户分组算法

输入 用户数目 K ; 信道向量 g_{mk} ; UE k 的训练数据集 D_k 及其大小 $\mathcal{D}_k$; 总训练数据集大小 D ; 迭代精度 ϵ ; 总分组个数 G

输出 用户聚类组集 $\Omega \triangleq \{\Omega_1, \Omega_2, \Omega_3\}$

步骤 1 初始化总分组个数 $G=3, i=1, j=1, U=\emptyset, \Omega_{g=1,2,3}=\emptyset$;

步骤 2 随机选择 G 个初始聚类中心 $C^0 \triangleq \{c_1^0, c_2^0, c_3^0\}$;

步骤 3 循环执行以下步骤

步骤 4 for $i=1, 2, \dots, K$ do

步骤 5 计算 $q_k g_{mk}$ 与每个质心 $c_g^i (g=1, 2, 3)$ 之间的相关系数: $d_{k,g}^i = (q_k g_{mk}) c_g^i / (|q_k g_{mk}| |c_g^i|)$;

步骤 6 选择 $g^* = \arg\max_{g \in \{1, 2, 3\}} d_{k,g}^i$ 并更新 $\Omega_{g^*} = \Omega_{g^*} \cup \{k\}$;

步骤 7 end for

步骤 8 更新 $i=i+1$;

步骤 9 for $g=1, 2, 3$ do

步骤 10 更新质心 $c_g^i = \frac{1}{\Omega_g} \sum_{k \in \Omega_g} q_k g_{mk}$;

步骤 11 end for

步骤 12 直到 $\|C^i - C^{i-1}\| \leq \epsilon$ 。

用户分组具体执行过程如图 2 所示。

本文用户分组算法的关键是采用了 $q_k g_{mk}$ 作为

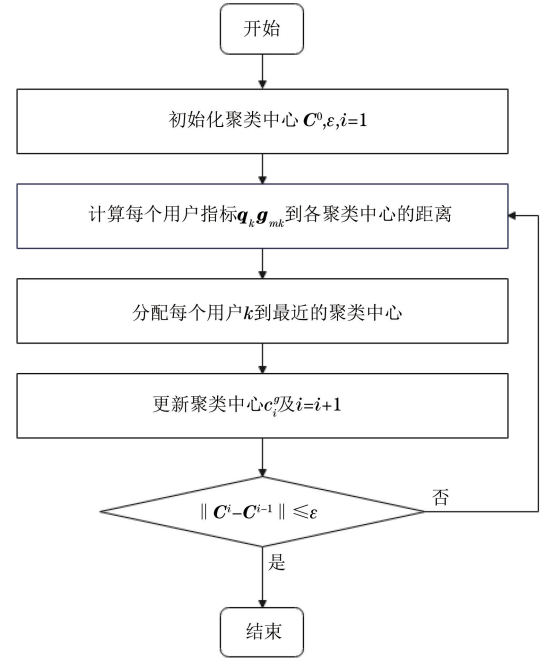


图 2 基于改进 k -means 的用户分组算法流程

Fig.2 Procedure of user grouping algorithm based on improved k -means

聚类的准则, 该指标反映了用户 k 对此次联邦学习和通信任务的贡献大小。由式(18)可知, q_k 越大代表用户 k 用于本次联邦学习任务训练的数据集越大, 并且数据异质性小, 这有利于加快联邦学习的收敛。而 g_{mk} 越大说明用户 k 此时的链路状态越好, 意味着用户能够以更低的时延和更高的可靠性完成通信任务, 这对于联邦学习中的模型更新和聚合至关重要。将这两个指标的乘积作为用户聚类的度量, 能够同时考虑用户对模型训练的数据贡献和通信质量的贡献, 从而选出既能够提高学习性能又能够高效通信的用户组合。

假设分别用 Ω_1, Ω_2 和 Ω_3 代表“优”、“良”和“差”三类用户的用户集。为了改善“差”用户的信道状态, 根据聚类结果做出如式(19)所示的 CPU 频率处理分配策略:

$$f_k = \begin{cases} f_{\min}, & k \in \Omega_1 \\ f_k^*, & k \in \Omega_2 \\ f_{\max}, & k \in \Omega_3 \end{cases} \quad (19)$$

式中: $f_k^* = \arg\min_{k \in \Omega_2} \left(\frac{S_d}{R_{d,k}} + \frac{L D_k c}{f} + \frac{S_u}{R_{u,k}} \right)$ 。

在无线通信中, CPU 处理频率的合理分配对于确保通信质量和系统效率至关重要。对于拥有优质通信资源的“优”, 可以考虑牺牲一定的计算速率以降低系统能耗。对于“差”用户集的用户, 分配更高的 CPU 频率意味着更快的数据处理速度, 减少通信延迟, 能在一定程度上弥补信道质量不佳的弊端,

提升用户体验。而对于“良”用户集中的用户,本文在保证通信质量的前提下,尽可能地采用较低的 CPU 处理频率以节省能源开销。式(19)的 CPU 处理频率分配方案旨在平衡用户通信资源和计算资源的占用,进而提升系统效率。此外,该分配方案还能有效改善“差”用户的资源紧张状况,增加其参与 FL 的概率。

2) 用户调度。根据用户分组和 CPU 处理频率分配方案,用户调度式(17)可转换为求解式(20):

$$\max_{b_{g,k}} \sum_{k=1}^K \sum_{g=1}^G b_{g,k} \ln \left(\frac{r_{g,k}}{\sum_{i \in \Omega_g} b_{g,i}} \right) \quad (20)$$

s. t. C_2

将 $U_{m,g}$ 定义为第 m 个 AP 所有关联用户中被划分为编号为 g 的用户集,则具体用户调度方案的求解算法如算法 2 所示。

算法 2 基于聚类贪婪算法的用户调度策略

输入 AP 和 UE 数目为 M 和 K ; 噪声归一化下行、上行及导频功率为 p_d 、 p_u 和 p_p ; 上下行功率控制因子为 ζ 和 η ; 上下行传输模型大小为 S_d 和 S_u ; 本地更新迭代次数为 L ; 处理一个样本所需周期为 c

输出 用户调度策略 α 和 CPU 处理频率分配方案 f

步骤 1 初始化 $U \triangleq \{U_{m,g}, m=1,2,\dots,M, g=1,2,3\} = \emptyset, b_{g,k}=0, a_k=0$;

步骤 2 for $m=1,2,\dots,M$ do

步骤 3 执行算法 1 得到用户聚类组 $\Omega_g \triangleq \{\Omega_g, g=1,2,3\}$;

步骤 4 for $g=1,2,3$ do

步骤 5 for $k \in \Omega_g$ do

步骤 6 计算 UE k 被关联后 $U_{m,g}$ 用户集的平均速率,如果平均速率有提升则更新 $b_{g,k}=1, U_{m,g} = U_{m,g} \cup \{k\}$ 并设置 $a_k=1$;

步骤 7 end for

步骤 8 end for

步骤 9 得到与 AP m 关联的用户集 $U_m \triangleq \{U_{m,g}, g=1,2,3\}$, 并根据式(20)为关联用户分配 CPU 处理频率 f_k ;

步骤 10 end for

2.2 基于二分法的功率分配算法

为了确保系统的公平性,本文以最大化系统中所有用户的最小可达速率为目标,对上行和下行的发射功率进行联合优化调整。针对这一复杂的随机非凸问题,将其解耦为两个独立的子问题,并提出一种基于二分法^[25-26]的功率分配算法 BM-PA。该算法首先将联合优化问题分解为上行和下行两个子问题,随后通过二分法不断迭代求解每个子问题以逐步逼近最优解。通过优化对发射功率的分配,以改善聚类标签为“差”用户组(即严重阻碍系统通信

效率的“短板”用户子集)中参与用户的信道链路状态,从而逐步实现整个系统中各用户资源分配的均衡化,并确保各用户能够获得相近的参与联邦学习的机会。通过持续提升“短板”用户子集的性能,进而优化系统整体性能。

在确定的用户调度策略 α 和用户 CPU 处理频率分配下,功率分配问题可简化为如下优化模型:

$$\min_{\eta, \zeta} \left(\max_{k \in U} \frac{S_d}{R_{d,k}(\eta)} + \max_{k \in U} \frac{S_u}{R_{u,k}(\zeta)} \right) \quad (21)$$

s. t. $C_1: \sum_{k \in K} \sigma_{mk}^2 \eta_{mk} \leq 1$
 $C_2: 0 \leq \eta_{mk}$
 $C_3: 0 \leq \zeta_k \leq 1$

两个不同方向的功率分配方案可以独立进行优化求解。因此,上述资源分配子问题式(21)可以进一步解耦为式(22)和式(23)两个问题:

$$Q1 \quad \min_{\eta} \max_{k \in U} \frac{S_d}{R_{d,k}(\eta)} \quad (22)$$

s. t. $C_1: \sum_{k \in K} \sigma_{mk}^2 \eta_{mk} \leq 1$
 $C_2: 0 \leq \eta_{mk}$

$$Q2 \quad \min_{\zeta} \max_{k \in U} \frac{S_u}{R_{u,k}(\zeta)} \quad (23)$$

s. t. $C_1: 0 \leq \zeta_k \leq 1$

这两个问题属于同一类型的问题,可采用相同的方法求解,本文采用二分法求解。

1) 下行功率分配子问题 Q1。本文求解 $\min_{\eta} \max_{k \in U} S_d/R_{d,k}$ 相当于求解 $\max_{\theta, \vartheta} \min_{\eta} R_{d,k}$ 。引入松弛变量 $\theta_{k'k}$ 和 ϑ_m , 于是将式(22)改写如下:

$$\max_{\theta_{k'k}, \vartheta_m} \min_{\eta} \left(\frac{\left(\sum_{k \in U} \sum_{m \in \mathcal{M}} v_{mk} \gamma_{mk} \right)^2}{\sum_{k' \in U \setminus k} |\varphi_k^H \varphi_{k'}|^2 \theta_{k'k}^2 + \sum_{k \in U} \sum_{m \in \mathcal{M}} \beta_{mk} \vartheta_m^2 + \frac{1}{p_d}} \right) \quad (24)$$

s. t. $C_1: \sum_{k' \in U} \gamma_{mk} v_{mk'} \leq \vartheta_m^2$
 $C_2: \sum_{m \in \mathcal{M}} \gamma_{mk} v_{mk'} \frac{\beta_{mk}}{\beta_{mk'}} \leq \theta_{k'k}, \forall k' \neq k$
 $C_3: 0 \leq \vartheta_m \leq 1$
 $C_4: 0 \leq v_{mk}$

式中: $v_{mk} \triangleq \sqrt{\eta_{mk}}$; $\gamma_{mk} \triangleq \sigma_{mk}^2$ 。

2) 上行功率分配子问题 Q2。类似于 Q1, 问题式(23)等价于式(25):

$$\max_{\zeta_k} \min_{k \in U} \left(\frac{\zeta_k \left(\sum_{m \in \mathcal{M}} \sigma_{mk}^2 \gamma_{mk} \right)^2}{\sum_{k' \in U \setminus k} P_{kk'}^2 |\varphi_k^H \varphi_{k'}|^2 + \sum_{k \in U} V_k + \frac{W_k}{p_u}} \right) \quad (25)$$

s. t. C_1

式(24)和式(25)可以通过二分搜索在每一步中

求解一系列凸可行性问题,从而有效地求解。对于下行功率分配相当于求解以下凸问题:

$$\begin{cases} \|Y_k\| \leq \frac{1}{\sqrt{t}} \sum_{m \in \mathcal{M}} \gamma_{mk} v_{mk} \\ \sum_{k \in \mathcal{U}} \gamma_{mk} v_{mk}^2 \leq \vartheta_m^2 \\ \sum_{m \in \mathcal{M}} \gamma_{mk'} \frac{\beta_{mk}}{\beta_{mk'}} v_{mk'} \leq \theta_{k'k}, \forall k' \neq k \\ 0 \leq \vartheta_m \leq 1 \\ v_{mk} \geq 0 \end{cases} \quad (26)$$

式中: $Y_k \triangleq [Y_1^T I_{-k} Y_2^T 1/\sqrt{p_d}]^T$; $Y_1 \triangleq [\varphi_1^H \varphi_k \theta_{1k}, \dots, \varphi_K^H \varphi_k \theta_{Kk}]^H$; I_{-k} 是 $K \times K$ 的单位矩阵去掉第 k 列后大小为 $K \times (K-1)$ 的矩阵, 并且 $Y_2 \triangleq [\sqrt{\beta_{1k}} \vartheta_1, \dots, \sqrt{\beta_{Mk}} \vartheta_M]^T$ 。

求解上行功率分配相当于求解如下凸问题:

$$\begin{cases} \|Z_k\| \leq \frac{1}{\sqrt{t}} \zeta_k \sum_{m \in \mathcal{M}} \gamma_{mk} \\ 0 \leq \zeta_k \leq 1 \end{cases} \quad (27)$$

式中: $Z_k \triangleq [Z_1^T I_{-k} Z_2^T Z_3^T]^T$, $Z_1 \triangleq [\varphi_1^H \varphi_k \zeta_k g_{1k}, \dots, \varphi_K^H \varphi_k \zeta_k g_{Kk}]^T$, $g_{k'k} = \sum_{m \in \mathcal{M}} \gamma_{mk} \beta_{mk'} / \beta_{mk}$, $Z_2 \triangleq [\sqrt{\beta_{1k}} \gamma_{1k} \zeta_k, \dots, \sqrt{\beta_{Mk}} \gamma_{Mk} \zeta_k]^T$, $Z_3 \triangleq [\sqrt{\gamma_{1k}/p_u}, \dots, \sqrt{\gamma_{Mk}/p_u}]^T$ 。

具体的功率分配算法总结如算法 3 所示。

算法 3 基于二分搜索的功率分配算法

输入 AP 和 UE 数目 M 和 K ; 噪声归一化下行、上行及导频功率 p_d 、 p_u 和 p_p ; 用户调度策略 $\mathbf{a}$ 和 CPU 功率分配方案 $\mathbf{f}$; 模型迭代精度 ε

输出 最优的下行(上行)功率分配方案 $\boldsymbol{\eta}^*$ ($\boldsymbol{\zeta}^*$)

步骤 1 初始化 $t_{\min}, t_{\max}$;

步骤 2 循环执行如下步骤

步骤 3 更新 $t = \frac{t_{\min} + t_{\max}}{2}$;

步骤 4 求解凸问题式(26)或者式(27);

步骤 5 if 凸问题式(26)或者式(27)具有可行解集

$\mathcal{F}$

步骤 6 更新 $t_{\min} = t$;

步骤 7 else if 凸问题式(26)或者式(27)不可解

步骤 8 更新 $t_{\max} = t$;

步骤 9 end if

步骤 10 直到 $(t_{\max} - t_{\min}) \leq \varepsilon$

步骤 11 得到最优解 $\boldsymbol{\eta}^* = \underset{\boldsymbol{\eta} \in \mathcal{F}}{\operatorname{argmin}} \left(\max_{k \in \mathcal{U}} \frac{S_d}{R_{d,k}(\boldsymbol{\eta})} \right)$ 或

者 $\boldsymbol{\zeta}^* = \underset{\boldsymbol{\zeta} \in \mathcal{F}}{\operatorname{argmin}} \left(\max_{k \in \mathcal{U}} \frac{S_u}{R_{u,k}(\boldsymbol{\zeta})} \right)$ 。

基于二分搜索的功率分配算法的具体求解过程如图 3 所示。

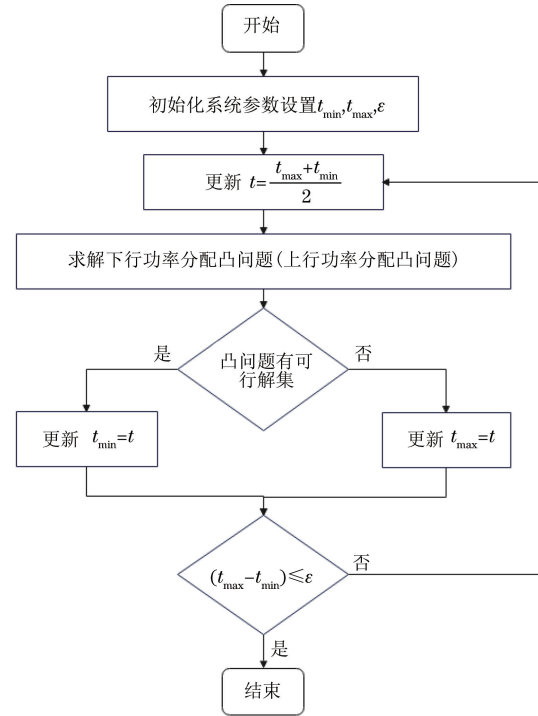


图 3 基于二分搜索的功率分配算法流程
Fig.3 Procedure of power allocation algorithm based on binary search

2.3 基于交替迭代优化的联合优化算法

本文采用交替迭代的优化方法解决联合优化问题。具体而言,首先,在大尺度相干时间内(即一个通信回合 r 内),求解式(17)获得最优调度方案 $\mathbf{a}^*$ 和用户 CPU 处理频率分配方案 $\mathbf{f}^*$ 。对于给定的用户选择策略 $\mathbf{a}^*$,资源分配子问题式(21)通过算法 3(BM-PA)进行若干次迭代求解,以获得最优解 $\boldsymbol{\zeta}^*$ 和 $\boldsymbol{\eta}^*$ 。随后,将 $\boldsymbol{\zeta}^*$ 、 $\boldsymbol{\eta}^*$ 作为算法 3 功率控制因子的值,重新执行用户调度算法,并持续更新 $\mathbf{a}^{r+1} = \mathbf{a}^*$ 、 $\mathbf{f}^{r+1} = \mathbf{f}^*$ 、 $\boldsymbol{\zeta}^{r+1} = \boldsymbol{\zeta}^*$ 、 $\boldsymbol{\eta}^{r+1} = \boldsymbol{\eta}^*$,直至达到最大迭代次数。

整个联合优化问题的求解算法步骤如算法 4 所示。

算法 4 基于交替迭代优化的联合优化算法

输入 AP 和 UE 数目 M 和 K ; 噪声归一化下行、上行及导频功率 p_d 、 p_u 和 p_p ; UE k 的训练数据集 D_k 及其大小 D_k ; 总训练数据集大小 D ; 上下行传输模型大小 S_d 和 S_u ; 本地模型更新总迭代次数 L ; 处理一个样本所需周期 c ; 模型迭代精度 ε ; FL 训练总迭代次数 R

输出 用户调度策略、CPU 处理频率分配方案以及上下行功率分配方案

步骤 1 初始化 $r=1$;

步骤 2 循环执行如下步骤

步骤 3 执行算法 3 得到用户本轮最优用户调度策略 $\mathbf{a}^*$ 及 CPU 处理频率分配方案 $\mathbf{f}^*$;

步骤 4 更新 $\mathbf{a}^{r+1} = \mathbf{a}^*$, $\mathbf{f}^{r+1} = \mathbf{f}^*$;

步骤 5 把 a^* 和 f^* 作为算法 3 的输入求解得到上下行最优功率分配策略 ζ^* 和 η^* ;

步骤 6 更新 $\eta^{r+1} = \eta^*$, $\zeta^{r+1} = \zeta^*$;

步骤 7 更新 $r = r + 1$;

步骤 8 直到 $r = R$ 。

基于交替迭代优化的联合优化算法的具体求解流程如图 4 所示。

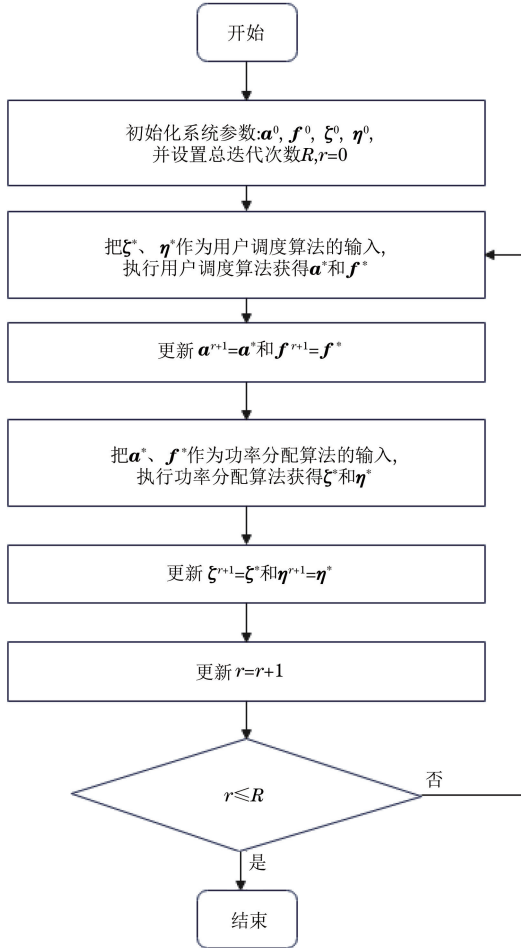


图 4 基于交替迭代优化的联合优化算法流程

Fig.4 Procedure of joint optimization algorithm based on alternating iterative optimization

2.4 计算复杂度分析

算法 1 采用 k -means 进行用户分组的算法复杂度为 $O(K)$, M 个 AP 关联 UE 先进行分组然后进行调度的总算法复杂度为 $O(MK)$ 。算法 3 在单个方向上采用二分搜索算法求解, 该算法最大的求解复杂度为 $O(\log_a(t_{\max} - t_{\min}))$, 假使上行和下行的功率分配设置相同的 $t_{\max}$ 和 $t_{\min}$, 则两个方向执行 BM-PA 算法的总复杂度为 $O(2 \log_a(t_{\max} - t_{\min}))$ 。算法 4 需要执行 R 次迭代, 总复杂度为 $O(R(MK + 2 \log_a(t_{\max} - t_{\min})))$ 。而文献[5]在进行功率分配算法之前还需要额外使用算法复杂度为 $O(MK)$ 贪婪算法进行导频的分配, 其中 M 为 AP 总数目, 然后

执行算法复杂度同样为 $O(2 \log_a(t_{\max} - t_{\min}))$ 的功率分配算法, 经过 R 次迭代后其算法总体复杂度为 $O(R(MK + 2 \log_a(t_{\max} - t_{\min})))$ 。此外文献[14]需要通过求解一系列的二阶锥规划问题获得功率分配和用户选择方案, 其中一个二阶锥规划问题的求解所需迭代次数为 $\sqrt{3(K+M)+MK}$, 每次求解的复杂度大约为 $O(n((M+1)K^3 + 2nK^2 + 2M^2))$, n 表示需要优化的总变量的个数, R 次迭代后其算法复杂度为 $O(nR \sqrt{3(K+M)+MK} ((M+1)K^3 + 2nK^2 + 2M^2))$ 。相比之下, 本文所提算法的复杂度较低。

3 仿真结果与分析

3.1 仿真参数设置与对比基准

为了验证本文所提算法的有效性, 本文利用 MATLAB 机器学习工具箱进行了 FL 算法的手写数字识别仿真, 并对所提出的用户调度算法和 BM-PA 功率分配算法的有效性进行了验证。在仿真实验中, 本文假设有 100 个 AP 和 20 个 UE 随机均匀分布在一个 $500 \text{ m} \times 500 \text{ m}$ 的特定区域内。网络系统的参数配置为: 下行和上行的系统带宽 B^D 和 B^U 分别为 20 MHz 和 1 MHz; 相干时间块长度 $\tau_c = 200$ (sample); 导频序列长度 $\tau_p = 20$ (sample); 全局模型和本地模型大小 $S_d = S_u = 5 \text{ MB}$; 上行和导频发射功率 $p_u = p_p = 0.2 \text{ W}$; 下行发射功率 $p_d = 1 \text{ W}$; 噪声归一化功率 $\sigma_0^2 = -92 \text{ dBm}$; 用户本地数据集大小 $D_k \in [1, 40] \times 10^4$ (sample); 本地训练模型更新需要总迭代次数 $L = 5$; $c \in [10, 40]$ (cycles/sample); 用户的 CPU 处理频率 $f_k \in [100, 2000] \text{ MHz}$, 模型迭代精度 $\epsilon = 0.001$ 。参数具体配置如表 1 所示。

表 1 仿真参数配置

Table 1 Simulation parameter configuration

参数	物理含义	参数值
$p_p, p_d, p_u/\text{W}$	导频、下行和上行发射功率	0.2, 1, 0.2
$\tau_c, \tau_p/\text{sample}$	相干时间和导频序列长度	200, 20
$S_d, S_u/\text{MB}$	下发和上传数据量的大小	5
$B^U, B^D/\text{MHz}$	系统上、下行传输带宽	1, 20
σ_0^2/dBm	噪声功率	-92
D_k	第 k 个 UE 训练数据集大小	$110^4 \sim 4010^4$
f_k/MHz	用户 k 的 CPU 处理频率	100~2000
$L/\text{次}$	UE 本地模型更新所需迭代次数	5
ϵ	模型迭代精度	0.001

本文将算法 4 (PROPOSED) 与以下 4 种基线方案进行对比验证。

基准 1(THUS EQPA):只调度每轮完成 FL 训练任务的时间小于某一阈值的用户,功率分配采用等功率分配算法。

基准 2(NOUS MMPA):未进行用户调度,采用文献[5]中改进的 Max-Min 功率分配方案。

基准 3(PRBS EEPA):采用文献[14]提出的基于接收功率(PRBS)的用户调度算法和能效最大化的功率分配方案。

基准 4(RAW):未采取任何优化策略的网络。

3.2 结果分析

图 5 展示了执行 100 次 FL 迭代后,针对 4 种资源优化方案与未优化网络的系统下行可达速率及其累积分布函数曲线。如图 5(a)所示,本文提出的方案(PROPOSED)在提升系统下行吞吐量方面表现最佳,相较于未优化的基准 4,下行可达速率最大提升 112.6%。相反,未采用合理的用户调

度方案的 NOUS MMPA 算法对速率提升效果不佳。本文算法通过选择合适用户并公平分配资源,降低系统资源分配的不均,使得下行可达速率比 NOUS MMPA 算法提升了 103.34%。如图 5(a)所示,相较于采用等功率算法的 THUS EQPA 算法,本文提出的算法最大额外提升下行吞吐量可达速率为 80.64%,平均提升 76.49%。本文提出的资源优化算法相较于 PRBS EEPA 算法在提升系统下行吞吐量上更具优势,平均提升 34.41%的下行吞吐量。图 5(b)显示,使用本文提出的算法进行资源优化后,系统下行吞吐量平均集中在 70~75 Mb/s,而经 NOUS MMPA 算法和 PRBS EEPA 算法优化后系统的下行吞吐量分别集中在 35~40 Mb/s 和 50~58 Mb/s 左右。这一结果得益于本文算法平衡了各用户之间的资源,提高了整体系统的吞吐量。

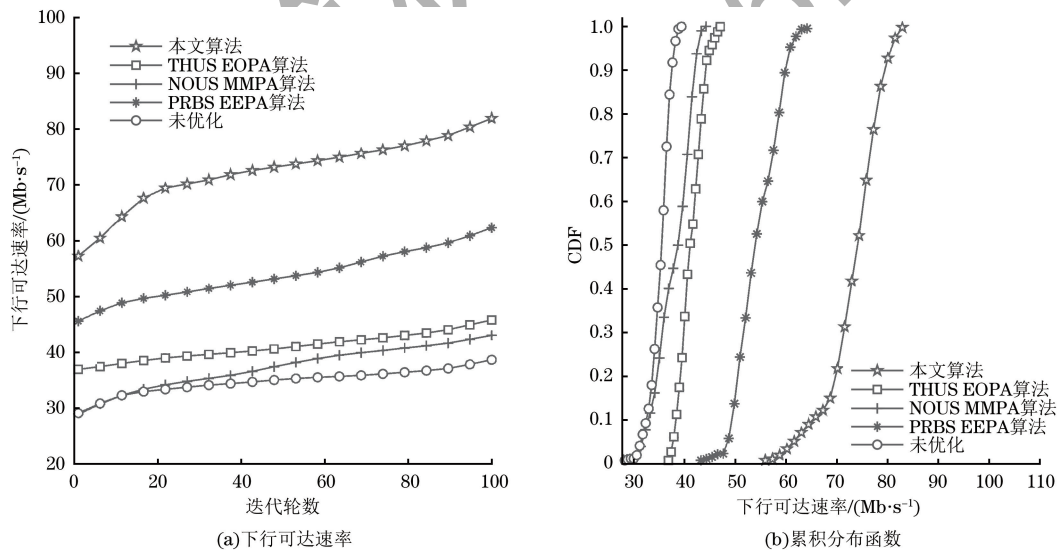


图 5 不同优化方案实现的下行可达速率及其累积分布函数曲线

Fig. 5 The downlink reachable rates and their cumulative distribution function curves achieved by different optimization schemes

图 6 展示了执行 100 次 FL 迭代后,针对 4 种不同资源优化方案及未优化网络的系统上行可达速率和吞吐量的累积分布函数曲线。根据图 6(a)的结果,本文提出的方案(PROPOSED)在提升系统上行吞吐量方面表现最佳。具体数据显示,相较于未进行优化的网络基准 4,上行可达速率最大可提升 115.15%,而其他优化方案 THUS EQPA 算法、NOUS MMPA 算法和 PRBS EEPA 算法的提升分别为 55.10%、102.78%和 27.02%。相较于 NOUS MMPA 算法中的改进 Max-Min 功率分配算法,本文提出的算法额外平均提升上行可达速率为 69.75%。而与 THUS EQPA 算法采用的门限截止的用户调度算法相比,本文的调度策略额外平均提

升速率 48.17%,展现了其优越性。此外,相较于基于接收功率和能效最大化的资源优化方案,本文提出的方案在提升系统上行容量方面效果也较优,平均可额外提升上行可达速率 25.33%。图 6(b)的结果显示,使用本文提出的算法优化后,系统上行可达速率显著高于其他 3 个基准,平均集中在 2.3~2.5 Mb/s。综上所述,图 6 的结果进一步验证了本文提出的算法在上行链路上的优越性,能够显著提高系统的吞吐量。

图 7 显示了执行 100 次 FL 迭代后,4 种不同资源优化方案及未优化网络的每个通信回合所需的时间消耗情况。根据图 7 的结果显示,本文提出的算法 4(PROPOSED)在节省时间消耗方面表现最佳,

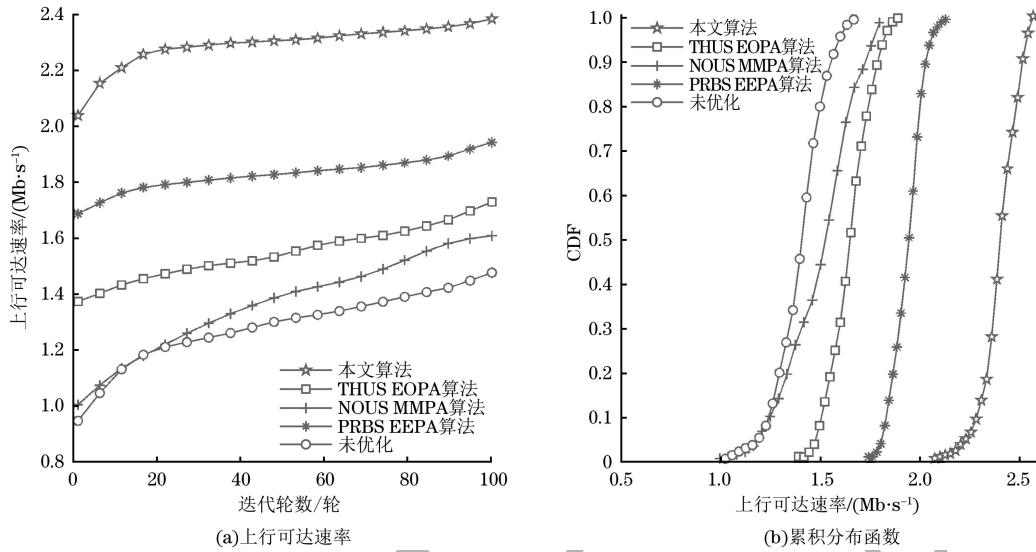


图 6 不同优化方案实现的上行可达速率及其累积分布函数曲线

Fig. 6 The uplink reachable rates and their cumulative distribution function curves achieved by different optimization schemes

能够有效地加快 FL 的收敛速率。从图中可以发现,随着迭代次数增加,系统时间消耗呈下降趋势,这是因为随着 FL 的进行,用户的本地模型逐渐接近最优全局模型,算法优化效果增强。数值结果显示,相较于无优化网络基准 4,本文提出的算法 4 在其基础上平均节省 80.77% 的时间开销;相较于 NOUS MMPA 算法采用改进的 Max-Min 功率分配算法,本文提出的算法平均节省时间消耗为 67.44%;进一步对比基准 3 中采用的 PRBS EEPA 算法,本文提出的算法在其基础上额外能够平均节省 42.08% 的时间开销。这表明了所提算法在降低系统时延方面的有效性。本文的算法设计将有限的资源尽可能公平、合理地分配给系统中被调度的用户,有效地削弱了系统的异构性。通过改善“差”用户(即“落伍者”)的资源,有效缓解了网络中“落伍者”严重减缓整个 FL 进程的影响,因此在时间开销上具有优越性。与 THUS EQPA 算法采用的门限截止调度相比,本文提出的算法在 FL 训练前期时间开销稍高。这是因为基于门限截止的用户选择过于极端,只选择满足时间约束的用户,因此在时间消耗上效果最佳。然而,这种策略会导致参与 FL 的用户较少,训练数据集丰富度和大小受损,从而影响 FL 模型的性能。为了达到一定的 FL 模型精度,需要更多的通信回合。尽管本文所提算法在 FL 训练初期时间开销略高于 THUS EQPA 算法,但它能够保持较好的 FL 性能。随着 FL 进入后期,本文提出的算法的时间开销逐渐接近甚至优于 THUS EQPA 算法。

为了方便度量本文算法在节省系统时间开销的效果,本文定义一个能够反映时间节省量的指标:

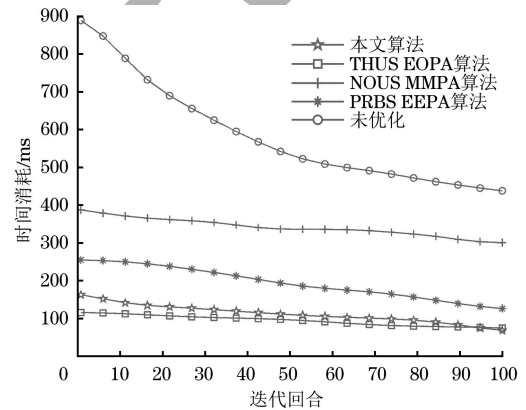


图 7 不同方案一个 FL 通信回合时间消耗情况

Fig. 7 Time consumption of one FL communication round under different schemes

(长耗时-短耗时)/长耗时 $\times 100\%$ 。

图 8 和表 2 展示了不同优化方案在 FL 模型精度分别达到 85% 和 90% 时,所提出的算法 4 与其他 4 个基准方法之间的时间消耗情况。根据图 8 的结果,无论是模型精度达到 85% 还是 90%,所提出的算法所需的时间消耗都明显少于基准 2、3 和 4。当模型精度为 85% 时,提出的算法仅需约 1.8 s,相比于 NOUS MMPA 算法、PRBS EEPA 算法以及 RAW,分别节省了约 75.03%、59.20% 和 91.60% 的时间。表 2 的数据显示,未进行优化的网络基准 4 无法实现 90% 的模型精度要求。相较于 THUS EQPA 算法,在模型精度达到 85% 时,本文提出的算法时间消耗略高,但在达到更高的模型精度 90% 时,其优越性得以体现,时间开销节省了约 9.38%。相较于基准 2、3 和 4,提出的方案在模型精度达到 90% 时仍然具有优势,所需时间约为 9.1 s,相比于 NOUS MMPA 算法和 PRBS EEPA

算法,分别节省了约 69.57%和 43.39%的时间。

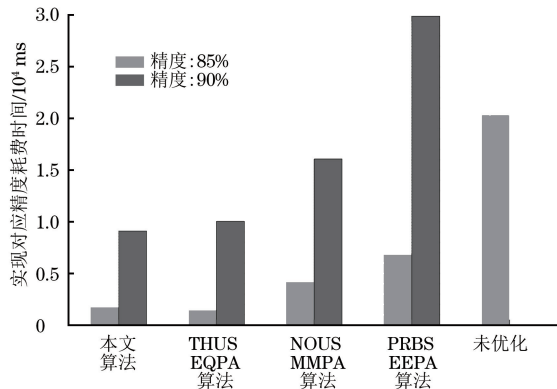


图 8 不同算法 FL 模型精度达到 85%和 90%的时间消耗
Fig.8 The FL model accuracy of different algorithms achieve 85% and 90% of time consumption

图 9 展示了 3 种不同的用户调度方案下各用户参与 FL 训练任务的情况。相较于 THUS EQPA 算法完全不给“差”用户参与 FL 训练任务机会的基于门限截止的调度方案,本文所提算法和PRBSEEPa

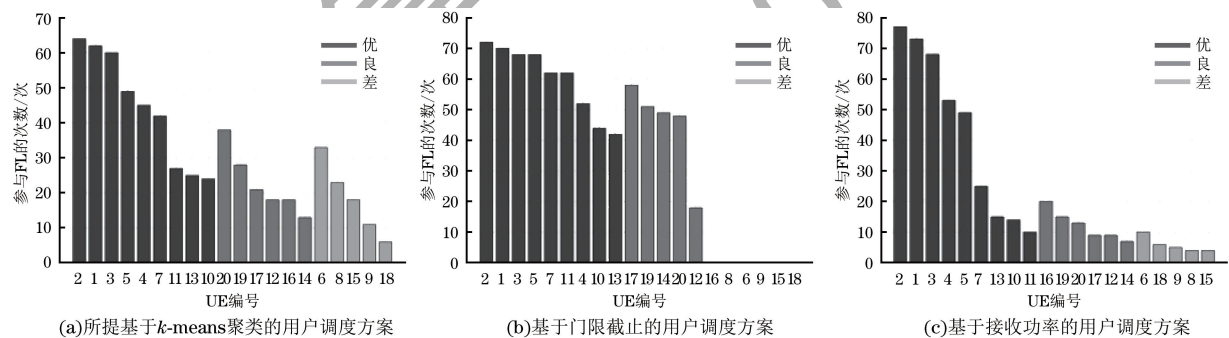


图 9 不同调度方案用户参与 FL 训练的状况

Fig.9 The status of users in FL training under different scheduling schemes

4 结束语

本文将联邦学习引入无蜂窝网络系统,旨在通过设计一个联合优化问题来解决网络资源调度和分配时存在的不公平的问题,以最大化系统的最小速率。该问题涉及用户调度、CPU 处理频率和上下行功率分配,其目标在于通过公平的资源调度和分配方案来提升整体网络速率。在时间约束下,本文建立了一个混合整数随机非凸问题模型。采用交替优化变量的方法,提出基于 k -means 聚类的贪婪用户调度算法,用于选择合理的用户参与 FL 任务并为其分配 CPU 处理频率。对于功率分配问题,采用基于二分法的功率分配算法,以较低的算法复杂度为资源紧缺的用户分配更优质的资源,有效削弱了系统用户资源占用两极化和不公平引起的速率提升受限问题。实验结果表明,相较于基准方案,本文提出方法提升了系统整体速率,并在联邦学习优化上

表 2 不同优化算法模型精度达到 85%和 90%的时间消耗

Table 2 The model accuracy of different optimization algorithms achieve 85% and 90% of time consumption 单位:ms

优化算法	模型精度达到 85%时耗时	模型精度达到 90%时耗时
本文算法	1 705.6	9 097.6
THUS EQPA 算法	1 447.8	1 039.8
NOUS MMPA 算法	6 831.2	29 894.4
PRBS EEPA 算法	4 180.7	6 071.5
未优化	20 311.7	NaN

算法让“差”用户也有参与 FL 训练任务的可能。此外,相较于 PRBS EEPA 算法,本文算法通过在每个组中都进行调度并为“差”用户改善资源分配的方法,使得“差”用户参与 FL 学习任务的机会增加。图 9 结果显示,相较于 PRBS EEPA 算法,本文“优”、“良”、“差”三类用户参与 FL 训练的机会差异度更小,更加平均,体现了本文算法的公平性。

取得了明显效果,显著减少了 FL 训练时间,加快了模型的收敛速度。下一步研究将考虑用户的移动性,并配合使用更加复杂的机器学习技术,如深度强化学习等来进一步提高系统资源分配和调度的动态性和智能化。

参考文献

- [1] LI T, SAHU A K, TALWALKAR A, et al. Federated learning: challenges, methods, and future directions [J]. IEEE Signal Processing Magazine, 2020, 37(3): 50-60.
- [2] TRAN N H, BAO W, ZOMAYA A, et al. Federated learning over wireless networks: optimization model design and analysis [C] // Proceedings of IEEE Conference on Computer Communications. Washington D. C., USA: IEEE Press, 2019: 1387-1395.
- [3] 熊世强,何道敬,王振东,等. 联邦学习及其安全与隐私保护研究综述[J]. 计算机工程, 2024, 50(5): 1-15.
XIONG S Q, HE D J, WANG Z D, et al. Review of federated learning and its security and privacy protection [J]. Computer Engineering, 2024, 50(5): 1-15. (in Chinese)
- [4] ELHOUSHY S, IBRAHIM M, HAMOUDA W. Cell-free massive MIMO: a survey [J]. IEEE Communications Surveys

- & Tutorials, 2022, 24(1): 492-523.
- [5] NGO H Q, ASHIKHMIM A, YANG H, et al. Cell-free massive MIMO versus small cells[J]. *IEEE Transactions on Wireless Communications*, 2017, 16(3): 1834-1850.
- [6] MU Y C, GARG N, RATNARAJAH T. Federated learning in massive MIMO 6G networks: convergence analysis and communication-efficient design [J]. *IEEE Transactions on Network Science and Engineering*, 2022, 9(6): 4220-4234.
- [7] XIA W C, QUEK T Q S, GUO K, et al. Multi-armed bandit-based client scheduling for federated learning [J]. *IEEE Transactions on Wireless Communications*, 2020, 19(11): 7108-7123.
- [8] 丁青峰, 徐梦引, 石辉, 等. 基于 ADC 精度与 AP 选择的蜂窝系统能效均衡方法[J]. *计算机工程*, 2022, 48(11): 192-200.
DING Q F, XU M Y, SHI H, et al. Energy efficiency equalization method of cell-free system based on ADC precision and AP selection[J]. *Computer Engineering*, 2022, 48(11): 192-200. (in Chinese)
- [9] 温依霖, 赵乃良, 曾艳, 等. 基于本地模型质量的客户端选择方法[J]. *计算机工程*, 2023, 49(6): 131-143.
WEN Y L, ZHAO N L, ZENG Y, et al. Client selection method based on local model quality [J]. *Computer Engineering*, 2023, 49(6): 131-143. (in Chinese)
- [10] 陶梅霞, 王栋, 孙瑞, 等. 联邦学习中基于时分多址接入的用户调度策略[J]. *通信学报*, 2021, 42(6): 16-29.
TAO M X, WANG D, SUN R, et al. TDMA-based user scheduling policies for federated learning [J]. *Journal on Communications*, 2021, 42(6): 16-29. (in Chinese)
- [11] FAROOQ M, NGO H Q, NAM TRAN L. A low-complexity approach for max-min fairness in uplink cell-free massive MIMO[C]//*Proceedings of the 93rd IEEE Vehicular Technology Conference*. Washington D. C., USA: IEEE Press, 2021: 321-328.
- [12] 姜兴, 刘耀, 李晓峰, 等. 一种毫米波宽带宽波束双极化微带天线[J]. *电讯技术*, 2023, 63(9): 1433-1438.
JIANG X, LIU Y, LI X F, et al. A broadband and wide beam dual polarization microstrip antenna in millimeter wave band[J]. *Telecommunication Engineering*, 2023, 63(9): 1433-1438. (in Chinese)
- [13] SHI W Q, ZHOU S, NIU Z S. Device scheduling with fast convergence for wireless federated learning[C]//*Proceedings of IEEE International Conference on Communications*. Washington D. C., USA: IEEE Press, 2020: 431-444.
- [14] NGO H Q, TRAN L N, DUONG T Q, et al. On the total energy efficiency of cell-free massive MIMO [J]. *IEEE Transactions on Green Communications and Networking*, 2018, 2(1): 25-39.
- [15] CHEN M Z, YANG Z H, SAAD W, et al. A joint learning and communications framework for federated learning over wireless networks [J]. *IEEE Transactions on Wireless Communications*, 2021, 20(1): 269-283.
- [16] VU T T, NGO D T, TRAN N H, et al. Cell-free massive MIMO for wireless federated learning[J]. *IEEE Transactions on Wireless Communications*, 2020, 19(10): 6377-6392.
- [17] 姚媛媛, 刘忆秋, 黄赛, 等. 无蜂窝大规模 MIMO 网络下基于联邦学习的用户接入策略及能耗优化[J]. *通信学报*, 2023, 44(10): 112-123.
YAO Y Y, LIU Y Q, HUANG S, et al. Federated learning-based user access strategy and energy consumption optimization in cell-free massive MIMO network[J]. *Journal on Communications*, 2023, 44(10): 112-123. (in Chinese)
- [18] VU T T, NGO D T, NGO H Q, et al. Joint resource allocation to minimize execution time of federated learning in cell-free massive MIMO [J]. *IEEE Internet of Things Journal*, 2022, 9(21): 21736-21750.
- [19] VU T T, NGO D T, NGO H Q, et al. Straggler effect mitigation for federated learning in cell-free massive MIMO[C]//*Proceedings of IEEE International Conference on Communications*. Washington D. C., USA: IEEE Press, 2021: 1-6.
- [20] GUO F Q, LU H C, GU Z J. Joint power and user grouping optimization in cell-free massive MIMO systems[J]. *IEEE Transactions on Wireless Communications*, 2022, 21(2): 991-1006.
- [21] ALEDHARI M, RAZZAK R, PARIZI R M, et al. Federated learning: a survey on enabling technologies, protocols, and applications [J]. *IEEE Access*, 2020, 8: 140699-140725.
- [22] SUN T, LI D, WANG B. Decentralized federated averaging [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(4): 4289-4301.
- [23] PREMKUMAR M, RAJAKUMAR S, SUBRAJA R. Signal processing algorithms for mean square error analysis in MIMO wireless transceivers[J]. *Ingénierie Des Systèmes d'Information*, 2023, 28(6): 1695-1700.
- [24] IKOTUN A M, EZUGWU A E, ABUALIGAH L, et al. k -means clustering algorithms: a comprehensive review, variants analysis, and advances in the era of big data[J]. *Information Sciences*, 2023, 622: 178-210.
- [25] RAHIM N A, AMIR A, AMIR A, et al. Application of bisection search method for maximum power extraction[C]//*Proceedings of International Conference on Intelligent and Advanced System*. Washington D. C., USA: IEEE Press, 2018: 445-453.
- [26] KOOLIVAND M, BAHONAR M H, FAZEL M S. Improving energy efficiency of massive MIMO relay systems using power bisection allocation for cell-edge users [C]//*Proceedings of the 27th Iranian Conference on Electrical Engineering*. Washington D. C., USA: IEEE Press, 2019: 1470-1475.