中国科学 : 信息科学 2026 年 第 56 卷 第 10 期: 2711–2726
SCIENTIA SINICA Informationis
面向 6G 的通感算可重构无线网络专题 • 论文

流态天线辅助的联邦遗忘学习系统

陈一瑄[1], 杨照辉[1*], 徐浩[2], 曾成[3], 陈煜琪[1], Kai-Kit WONG[4,5], 张朝阳[1]

1. 浙江大学信息与电子工程学院, 杭州 310027, 中国
2. 东南大学移动通信国家重点实验室, 南京 210096, 中国
3. 南京理工大学电子工程与光电技术学院, 南京 210094, 中国
4. Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, UK
5. Department of Electronic Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do 17104, Republic of Korea
* 通信作者. E-mail: yang_zhaohui@zju.edu.cn

收稿日期: 2026–04–30; 修回日期: 2026–08–13; 接受日期: 2026–09–08; 网络出版日期: 2026–10–10

国家自然科学基金 (批准号: 62394292, 62394290) 资助项目

摘要 联邦遗忘学习 (federated unlearning, FU) 允许用户撤销其数据对联邦学习模型的贡献, 但多轮模型校准会产生较大的计算和通信开销. 为降低无线边缘网络中的 FU 完成时延, 本文研究流态天线 (fluid antenna, FA) 辅助的无线 FU 系统. 在空间相关 Rician 信道下, 建立联合优化 FA 激活端口集合, CPU 计算频率, 上行发射功率, 传输时延和带宽分配的混合离散连续优化问题. 针对该问题, 提出一种基于动态规划和嵌套二分搜索的联合优化算法. 该算法利用动态规划求解满足最小间距约束的最优端口集合, 并通过嵌套二分搜索完成通信资源分配. 仿真结果表明, 与固定天线, 平均带宽分配和随机端口选择方案相比, 所提方案能够有效降低 FU 完成时延. 当最大发射功率为 10 dBm 时, 所提方案相较固定天线方案将完成时延降低约 31.4%. Fashion-MNIST 实验进一步评价了遗忘后的模型效用和遗忘效果.

关键词 联邦遗忘学习, 流态天线, 端口选择, 资源分配, 时延最小化

1 引言

随着万物互联时代的到来, 边缘智能已成为支撑智慧城市、工业物联网及个性化服务的核心技术 [1,2]. 联邦学习 (federated learning, FL) 作为一种兼顾模型训练与数据隐私的分布式范式, 被广泛应用于智能手机、可穿戴设备等边缘网络中, 有效规避了原始数据上传带来的隐私泄露风险 [3~5]. 然而, 随着《通用数据保护条例》和《加州消费者隐私法案》等法规的深入实施, 用户被赋予了 “被遗忘权”, 即有权要求从已训练的模型中彻底移除其个人数据痕迹 [6,7]. 传统的联邦学习架构仅支持数据的增量聚合, 无法直接响应此类遗忘请求. 若简单删除本地数据并保留现有模型, 残留的参数仍可能泄露被遗忘用户的信息, 这促使了联邦遗忘学习 (federated unlearning, FU) 这一新兴领域的诞生 [8,9]. FU 旨在通过逆向工程或高效的重训练机制, 精确抵消特定用户数据对全局模型的影响, 从而在满足合规性要求的同时维持模型的效用.

引用格式: 陈一瑄, 杨照辉, 徐浩, 等. 流态天线辅助的联邦遗忘学习系统. 中国科学: 信息科学, 2026, 56: 2711–2726, doi: 10.1360/SSI-2026-0205
Chen Y X, Yang Z H, Xu H, et al. Fluid antenna assisted federated unlearning system. Sci Sin Inform, 2026, 56: 2711–2726, doi: 10.1360/SSI-2026-0205

© 2026《中国科学》杂志社

尽管 FU 在算法层面展现出巨大的潜力, 将其部署于带宽受限、能量敏感的无线边缘网络中仍面临严峻挑战. 与标准的 FL 不同, FU 通常需要剩余用户进行多轮复杂的校准迭代 (如 FedEraser 框架) 以中和被遗忘用户的影响, 这导致了极其高昂的通信开销 [10,11]. 在无线网络中, 任务的完成时延由本地计算时延与无线传输时延共同决定, 且二者与模型的本地精度存在复杂的耦合关系. 一方面, 为了减少通信轮次, 用户倾向于在本地进行高精度的计算, 但这会显著增加本地计算能耗与时间; 另一方面, 若降低本地精度以减少计算负载, 则需要更多的全局通信轮次来保证模型收敛, 这在带宽稀缺的无线信道中往往成为系统瓶颈 [12,13].

现有的研究在解决 FU 的通信效率问题时, 多集中于算法层面的优化, 如减少通信轮次或模型压缩, 往往忽略了物理层无线资源的联合优化, 或者假设理想的通信链路条件 [14,15]. 与此同时, 无线 FL 领域已有研究开始关注物理层资源配置问题. 相关工作分别从本地计算与无线通信资源的协同配置、用户选择与无线资源分配以及资源约束下的本地更新和全局聚合控制等方面优化 FL 训练性能 [16~18]. 然而, 这些研究主要面向常规 FL 训练过程和固定天线架构, 尚未考虑 FU 多轮模型校准场景下流态天线端口选择与计算、通信资源之间的联合优化. 传统的多输入多输出 (multiple-input multiple-output, MIMO) 技术虽然通过空间复用提升了频谱效率, 但在面对边缘设备尺寸受限、天线间距过小导致的互耦效应以及深衰落信道时, 其性能提升遭遇瓶颈 [19,20]. 为了突破物理尺寸对天线性能的限制, 流态天线系统 (fluid antenna system, FAS) 作为一种新兴的可重构天线技术应运而生 [21,22]. FAS 利用微流控技术, 允许天线在物理空间的一维或多维线上灵活切换其物理位置. 这种物理上的 “移动性” 赋予了系统强大的空间自由度, 使得用户能够通过简单的端口位置切换, 避开深衰落点, 从而获得极大的信道增益 [23~25]. 相比于传统的固定位置天线 (fixed-position antenna, FPA), FAS 在不增加射频链路数量的前提下, 以极低的硬件成本实现了类似大规模 MIMO 的空间分集效果. 最新的研究进一步探讨了 FAS 在 MIMO 系统中的信息论表征、中断容量及混合视距/非视距环境下的分集性能, 并针对近场通信场景提出了联合波束赋形与天线设计的方案, 为在极大规模天线阵列中利用近场效应优化边缘学习提供了理论支撑 [26~29].

针对上述问题, 本文研究流态天线 (fluid antenna, FA) 辅助的无线 FU 系统, 通过联合利用 FA 的空间可重构能力与无线资源优化降低 FU 总完成时延. 本文的主要贡献如下.

(1) 提出一种 FA 辅助的无线 FU 框架. 考虑由单天线基站和配置一维多端口 FA 的用户设备组成的边缘网络. 在具有空间相关性的 Rician 信道下, 用户从多个密集候选端口中选择满足最小激活间距约束的端口集合, 从而改善模型更新上传过程中的等效信道增益.

(2) 建立以 FU 总完成时延最小化为目标的混合离散连续优化问题. 在给定模型校准轮数的条件下, 联合优化用户的 FA 激活端口集合, CPU 计算频率, 上行发射功率, 传输时延和带宽分配.

(3) 提出一种基于动态规划和嵌套二分搜索的联合优化算法. 利用动态规划在最小端口间距约束下求解各用户的最优激活端口集合, 根据目标函数的单调性确定最优计算频率和发射功率, 并通过外层完成时延二分搜索与内层带宽二分搜索求解最优传输时延和带宽分配.

(4) 通过蒙特卡罗仿真和 Fashion-MNIST 数据集上的 FU 实验评价所提方案的性能. 结果表明, 与固定天线方案, 平均带宽分配方案和随机端口选择方案相比, 所提方案能够在不同发射功率、带宽、端口数量和信道条件下有效降低 FU 完成时延.

本文其余部分组织如下: 第 2 节介绍系统模型与联邦遗忘机制; 第 3 节给出联合优化问题及其求解算法; 第 4 节展示仿真结果并进行分析; 第 5 节总结全文并讨论未来研究方向.

2　系统模型

考虑一个由单天线基站 (base station, BS) 和 K 个用户设备组成的无线 FU 系统, 如图 1 所示. 设用户集合为 $\mathcal{K}=\{1,2,\ldots,K\}$. 假设某一用户 $k_u\in\mathcal{K}$ 请求行使 “被遗忘权”, 即从全局模型中移除其数据贡献. 根据 FU 的设定, 剩余用户集合 $\mathcal{K}'=\mathcal{K}\setminus\{k_u\}$ 需要协同 BS 采用 FedEraser 框架执行遗忘任务. 在该过程中, 用户端配备一维 FA, 通过动态调整天线激活端口位置来利用空间自由度提升通信性能.

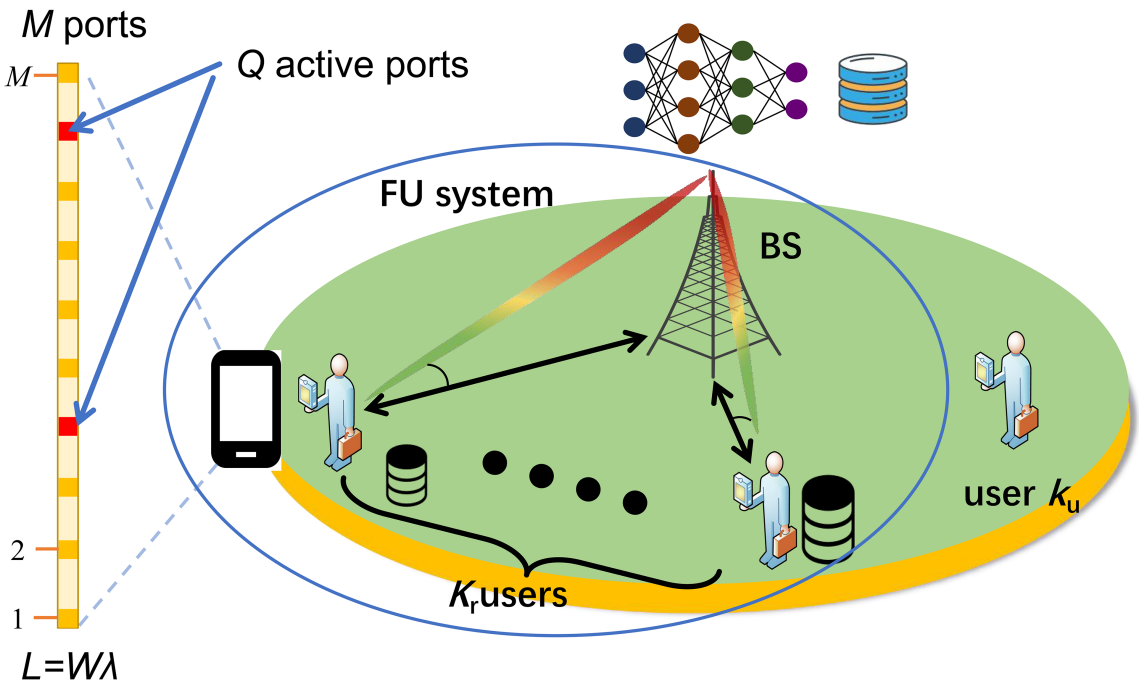

图 1　(网络版彩图) 流态天线辅助的联邦遗忘学习系统.

Figure 1　(Color online) Fluid antenna-assisted federated unlearning system.

2.1　流态天线系统模型

每个参与用户配置一个长度固定的一维多端口流态天线系统, 其物理孔径表示为

$$L = W\lambda, \tag{1}$$

其中 W 为归一化孔径, λ 为系统中心载频 f_c 对应的波长, 即

$$\lambda = \frac{c_0}{f_c}, \tag{2}$$

c_0 表示真空中的光速. 沿长度为 L 的固定孔径均匀设置 M 个候选端口. 候选端口是可重构辐射单元可以到达或切换的 M 个候选位置. 定义候选端口集合为 $\mathcal{M} = \{1, 2, \ldots, M\}$. 第 m 个候选端口的物理位置为

$$x_m = (m-1)\Delta_{\text{grid}}, \quad m \in \mathcal{M}, \tag{3}$$

其中 $\Delta_{\text{grid}} = \frac{L}{M-1}$ 表示相邻候选端口之间的空间采样间隔. 由于流态天线的特性, 当 $\Delta_{\text{grid}} < \frac{\lambda}{2}$ 时, 相邻候选端口的信道通常具有较强空间相关性. 该相关性将在后续信道模型中显式考虑.

假设用户 k 从 M 个候选端口中选择 Q 个端口同时激活, 其激活端口索引集合记为 $\mathcal{S}_k = \{r_{k,1}, r_{k,2}, \ldots, r_{k,Q}\}$, 其中 $1 \leqslant r_{k,1} < r_{k,2} < \cdots < r_{k,Q} \leqslant M$. 为了支持 Q 个端口同时工作, 假设用户侧配置 N_{RF} 条射频链, 并且 $Q \leqslant N_{\text{RF}}$. 此外, 为减小同时激活端口之间的互耦, 定义最小激活间距为 $\Delta_{\min}$. 可行端口集合表示为

$$\mathcal{C}(M, Q; \Delta_{\min}) = \left\{ \mathcal{S} \subseteq \mathcal{M} : \ |\mathcal{S}| = Q, |x_m - x_n| \geqslant \Delta_{\min}, \ \forall m, n \in \mathcal{S}, \ m \neq n \right\}. \tag{4}$$

因此, M 和 Q 应满足

$$1 \leqslant Q \leqslant \min\left\{ M, \ N_{\text{RF}}, \ 1 + \left\lfloor \frac{L}{\Delta_{\min}} \right\rfloor \right\}, \tag{5}$$

其中, M 主要由固定孔径内的空间采样分辨率决定, 而 Q 主要受射频链数量, 可同时工作的辐射单元数量, 互耦和硬件尺寸限制. 在实际系统中, 候选端口数通常远大于同时激活端口数, 即 $M \gg Q$.

令 ℓ_k 表示用户 k 到 BS 的大尺度信道功率增益, κ_k 表示 Rician 因子. 用户 k 在全部 M 个候选端口处的上行信道向量建模为

$$\boldsymbol{h}_k = \sqrt{\ell_k}\left(\sqrt{\frac{\kappa_k}{\kappa_k+1}}\,\boldsymbol{a}_k + \sqrt{\frac{1}{\kappa_k+1}}\,\boldsymbol{R}_k^{1/2}\boldsymbol{z}_k\right) \in \mathbb{C}^{M\times 1}, \tag{6}$$

其中, $\boldsymbol{z}_k \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_M)$ 为独立同分布的复高斯随机向量, $\boldsymbol{0}$ 为 M 维零向量, $\boldsymbol{I}_M$ 为 M 阶单位矩阵, $\boldsymbol{a}_k$ 为确定性视距 (line-of-sight, LoS) 分量对应的阵列响应向量, $\boldsymbol{R}_k$ 为散射分量的空间相关矩阵.

设用户 k 的主出发方向为 θ_k, 则阵列响应向量 $\boldsymbol{a}_k$ 的第 m 个元素可表示为

$$[\boldsymbol{a}_k]_m = \exp\left(-\mathrm{j}\frac{2\pi}{\lambda}\left(x_m - \frac{L}{2}\right)\sin\theta_k\right), \quad m \in \{1,\ldots,M\}. \tag{7}$$

在二维各向同性散射和理想全向端口假设下, 候选端口 m 与候选端口 n 之间的空间相关系数可采用零阶贝塞尔相关模型表示为

$$[\boldsymbol{R}_k]_{m,n} = J_0\left(\frac{2\pi\,|x_m - x_n|}{\lambda}\right), \quad m,n \in \{1,\ldots,M\}, \tag{8}$$

其中 $J_0(\cdot)$ 表示第一类零阶贝塞尔函数. 候选端口之间的空间相关性取决于端口间的物理距离及载波波长. 通常情况下, 相邻候选端口之间的距离越小, 其信道相关性越强. 令 $\boldsymbol{A}(\mathcal{S}_k) \in \{0,1\}^{Q\times M}$ 表示与用户 k 的激活端口集合 $\mathcal{S}_k$ 对应的选择矩阵. 矩阵 $\boldsymbol{A}(\mathcal{S}_k)$ 的每一行仅包含一个值为 1 的元素, 用于从全部 M 个候选端口中提取相应的激活端口. 因此, 用户 k 的激活端口信道向量可表示为

$$\widetilde{\boldsymbol{h}}_k(\mathcal{S}_k) = \boldsymbol{A}(\mathcal{S}_k)\,\boldsymbol{h}_k \in \mathbb{C}^{Q\times 1}. \tag{9}$$

假设用户采用最大比发射 (maximum-ratio transmission, MRT), 则其单位范数发射波束向量为

$$\boldsymbol{w}_k = \frac{\widetilde{\boldsymbol{h}}_k(\mathcal{S}_k)}{\left\|\widetilde{\boldsymbol{h}}_k(\mathcal{S}_k)\right\|_2}, \tag{10}$$

其中, 用户 k 的总发射功率按照波束向量 $\boldsymbol{w}_k$ 在 Q 个激活端口之间进行分配, 并满足 $\|\boldsymbol{w}_k\|_2^2 = 1$. 因此, 用户 k 在基站处的等效上行信道功率增益为

$$g_k(\mathcal{S}_k) = \left|\widetilde{\boldsymbol{h}}_k^{\mathrm{H}}(\mathcal{S}_k)\,\boldsymbol{w}_k\right|^2 = \left\|\widetilde{\boldsymbol{h}}_k(\mathcal{S}_k)\right\|_2^2 = \sum_{q=1}^{Q}\left|h_{k,r_{k,q}}\right|^2. \tag{11}$$

通过合理选择激活端口, 用户可以获得更高的等效上行信道增益, 进而提高模型参数的上行传输速率并降低联邦遗忘过程的通信时延.

2.2　联邦遗忘学习模型

系统采用 FedEraser [10] 框架进行 FU 训练. 为了消除被遗忘用户 k_u 数据的影响, 剩余用户 $\mathcal{K}'$ 需进行多轮校准迭代, 将训练好的模型参数更新通过上行链路传输到 BS. 为了专注于资源分配优化, 本文设定联邦遗忘过程包含固定的 J 轮全局校准迭代, 且每轮全局迭代中用户执行固定的 E 轮本地校准计算. BS 汇总全局预测模型参数, 并通过下行链路将这些参数广播给所有用户. 与功率和带宽受限的用户设备相比, 基站通常具有更大的发射功率和更充足的下行带宽. 本文考虑下行链路处于非瓶颈状态, 即下行全局模型广播时延远小于用户的本地计算时延与上行模型上传时延. 因此, 本文忽略下行传输时延, 重点研究用户侧本地计算和上行通信造成的联邦遗忘完成时延.

设用户 $k \in \mathcal{K}'$ 完成一次本地校准迭代所需的计算量为 C_k (cycles), CPU 计算频率为 f_k (cycles/s). 则用户 k 完成 E 轮本地更新所需的本地计算时延为

$$\tau_k^{\mathrm{comp}} = \frac{EC_k}{f_k}. \tag{12}$$

相比于本地计算时延和传输时延, FA 的 CSI 获取和端口选择时延相对较小, 可以忽略不计. 假设上行采用正交频分多址 (orthogonal frequency division multiple access, OFDMA). 系统总上行带宽为 B, 位于中心载频 f_c 附近. 由于系统总带宽满足 $B \ll f_c$, 本文采用窄带信道假设, 各用户分配的上行子带均位于中心载频 f_c 附近. 因此, 不同子带对应的波长差异可以忽略, 流态天线的端口位置, 空间相关矩阵和阵列响应统一按照中心波长 $\lambda = c_0/f_c$ 建模. 令 p_k 表示用户 k 在 Q 个激活端口上的总发射功率, N_0 表示噪声功率谱密度. 根据香农公式, 用户 k 的上行链路传输速率 R_k 为

$$R_k = b_k \log_2 \left(1 + \frac{p_k g_k(\mathcal{S}_k)}{N_0 b_k}\right). \tag{13}$$

由于上行链路总带宽的限制, 我们有 $\sum_{k\in\mathcal{K}'} b_k \leqslant B$. 假设每个用户每轮上传的模型更新数据量大小为 s bits, 则传输时延 t_k 满足

$$t_k R_k \geqslant s. \tag{14}$$

设 T_k 为用户 k 完成 J 轮全局校准所需的总时间. 由于每轮全局迭代包含 “本地计算” 和 “模型上传” 两个串行阶段, 用户 k 单轮耗时为 $\tau_k^{\mathrm{comp}} + t_k$. 因此, 用户 k 完成 FU 的总时延为

$$T_k = J \cdot (\tau_k^{\mathrm{comp}} + t_k) = J\left(\frac{EC_k}{f_k} + t_k\right). \tag{15}$$

FU 的完成时间由最慢用户决定, 因此系统总完成时延定义为 $T = \max_{k\in\mathcal{K}'} T_k$. 为便于后续求解, 在联合优化问题中将 T 作为最大用户时延的上界变量, 并通过约束 (16b) 实现上述最大值关系.

2.3　联合优化问题

在上述系统模型基础上, 本文通过联合优化流态天线的激活端口集合, 用户 CPU 计算频率, 上行发射功率, 上行传输时延和带宽分配, 最小化联邦遗忘总完成时延. 将所有参与用户的激活端口集合表示为 $\boldsymbol{\mathcal{S}} = \{\mathcal{S}_k : k \in \mathcal{K}'\}$, CPU 计算频率向量表示为 $\boldsymbol{f} = (f_k : k \in \mathcal{K}')$, 上行发射功率向量表示为 $\boldsymbol{p} = (p_k : k \in \mathcal{K}')$, 上行传输时延向量表示为 $\boldsymbol{t} = (t_k : k \in \mathcal{K}')$, 带宽分配向量表示为 $\boldsymbol{b} = (b_k : k \in \mathcal{K}')$. 构建优化问题 ($\mathcal{P}1$) 如下:

$$\mathcal{P}1: \quad \min_{\boldsymbol{\mathcal{S}},\boldsymbol{f},\boldsymbol{p},\boldsymbol{t},\boldsymbol{b},T} \quad T \tag{16}$$

$$\text{s.t.} \quad \sum_{k\in\mathcal{K}'} b_k \leqslant B, \tag{16a}$$

$$J\left(\frac{EC_k}{f_k} + t_k\right) \leqslant T, \qquad \forall k \in \mathcal{K}', \tag{16b}$$

$$s \leqslant t_k b_k \log_2 \left(1 + \frac{p_k g_k(\mathcal{S}_k)}{N_0 b_k}\right), \qquad \forall k \in \mathcal{K}', \tag{16c}$$

$$0 < f_k \leqslant f_k^{\max}, \qquad \forall k \in \mathcal{K}', \tag{16d}$$

$$0 \leqslant p_k \leqslant p_k^{\max}, \qquad \forall k \in \mathcal{K}', \tag{16e}$$

$$t_k > 0,\ b_k > 0, \qquad \forall k \in \mathcal{K}', \tag{16f}$$

$$\mathcal{S}_k \in \mathcal{C}(M, Q; \Delta_{\min}), \qquad \forall k \in \mathcal{K}', \tag{16g}$$

其中, 约束 (16a) 为系统上行带宽约束, 保证分配给所有参与用户的带宽之和不超过系统可用的上行总带宽 B. 约束 (16b) 为用户完成时延约束, 保证辅助变量 T 不小于任意参与用户完成 J 轮本地模型校准和上行模型上传所需的总时延. 约束 (16c) 为模型更新数据的上行传输约束. 约束 (16d) 为用户计算频率约束, 保证用户 k 的 CPU 计算频率为正, 且不超过其最大可用计算频率 $f_k^{\max}$. 约束 (16e) 为用户上行发射功率约束, 保证用户 k 的发射功率非负, 且不超过其最大允许发射功率 $p_k^{\max}$. 约束 (16f) 规定上行传输时延 t_k 和分配带宽 b_k 的正值约束. 约束 (16g) 为流态天线的端口选择约束.

问题 ($\mathcal{P}1$) 是一个混合离散 – 连续非线性优化问题. 离散的激活端口集合通过等效信道增益与发射功率, 带宽及传输时延等连续变量耦合, 同时各用户又受共享带宽约束和最大完成时延目标的共同影响. 因此, 原问题变量之间耦合紧密, 难以直接利用标准凸优化方法求解, 需要进一步利用等效信道增益的单调性对端口选择和连续资源分配进行分解.

3　算法设计

本节首先利用目标函数和约束关于计算频率, 发射功率及等效信道增益的单调性, 将原问题分解为端口选择子问题和连续资源分配子问题, 分别采用动态规划和双层二分搜索求解上述两个子问题.

3.1　流态天线端口选择

不同用户的端口选择之间不存在耦合关系, 并且系统完成时延关于等效信道增益 $g_k(\mathcal{S}_k)$ 单调递减. 因此, 每个用户均应独立选择使其等效信道增益最大的可行端口集合. 在给定计算频率 $\boldsymbol{f}$, 传输功率 $\boldsymbol{p}$ 和带宽 $\boldsymbol{b}$ 的情况下, 用户 k 的端口选择子问题可表示为

$$\mathcal{P}2: \quad \max_{\mathcal{S}_k} \quad g_k(\mathcal{S}_k) = \sum_{m\in\mathcal{S}_k} |h_{k,m}|^2 \tag{17}$$

$$\text{s.t.} \quad \mathcal{S}_k \in \mathcal{C}(M, Q; \Delta_{\min}). \tag{17a}$$

由于候选端口均匀设置在一维物理孔径上, 可以利用端口索引之间的顺序关系, 通过动态规划求解端口选择问题 $\mathcal{P}2$. 定义两个同时激活端口之间所需满足的最小索引间隔为

$$D_{\min} = \max\left\{1, \left\lceil \frac{\Delta_{\min}}{\Delta_{\text{grid}}} \right\rceil \right\}. \tag{18}$$

由于候选端口均匀分布, 当端口 m 被激活时, 与其相邻的前一个激活端口的索引不得大于 $m - D_{\min}$. 定义用户 k 的第 m 个候选端口的权重为

$$v_{k,m} = |h_{k,m}|^2. \tag{19}$$

令 $F_k(m, q)$ 表示从前 m 个候选端口中选择 q 个满足最小间距约束的激活端口时, 能够获得的最大等效信道增益. 对于候选端口 m, 存在以下两种情况.

第一种情况是不选择端口 m. 此时, 最优信道增益为 $F_k(m-1, q)$.

第二种情况是选择端口 m. 由于端口之间需要满足最小激活间距, 剩余的 $q-1$ 个激活端口只能从前 $m-D_{\min}$ 个候选端口中选择. 此时, 能够获得的最大信道增益为 $v_{k,m} + F_k(m - D_{\min}, q-1)$.

因此, 可以得到动态规划递推关系

$$F_k(m, q) = \max\left\{F_k(m-1, q), v_{k,m} + F_k(\max\{0, m - D_{\min}\}, q-1)\right\}. \tag{20}$$

动态规划的边界条件为

$$F_k(m, 0) = 0, \qquad m \geqslant 0, \tag{21}$$

以及

$$F_k(m, q) = -\infty, \qquad m \leqslant 0, \quad q > 0. \tag{22}$$

完成全部递推计算后, 通过从状态 $F_k(M, Q)$ 进行反向回溯, 可以得到用户 k 的最优激活端口集合 $\mathcal{S}_k^*$. 相应的最优等效信道增益为

$$g_k^* = F_k(M, Q) = \sum_{m\in\mathcal{S}_k^*} |h_{k,m}|^2. \tag{23}$$

3.2　联合资源分配

给定 FA 激活端口集合 $\mathcal{S}_k^*$ 和最优等效信道增益 g_k^*, 优化变量为计算频率 $\boldsymbol{f}$, 发射功率 $\boldsymbol{p}$, 传输时延 $\boldsymbol{t}$ 和带宽 $\boldsymbol{b}$. 为了最小化 FU 过程的总时延, 我们需要充分利用用户的本地计算资源和无线传输资源. 根据约束条件 (16b) 和 (16c) 的结构可知, 为了获得最小的完成时间 T, 用户应当在本地计算和模型上传阶段均采用最大可用资源. 具体而言, 计算频率 f_k 应取其上限 $f_k^{\max}$, 发射功率 p_k 应取其上限 $p_k^{\max}$. 令 $f_k = f_k^{\max}$ 和 $p_k = p_k^{\max}$, 我们可以构建优化子问题 ($\mathcal{P}3$) 如下:

$$\mathcal{P}3: \quad \min_{\boldsymbol{t},\boldsymbol{b},T} \quad T \tag{24}$$

$$\text{s.t.} \sum_{k\in\mathcal{K}'} b_k \leqslant B, \tag{24a}$$

$$J\left(\frac{EC_k}{f_k^{\max}} + t_k\right) \leqslant T, \qquad \forall k \in \mathcal{K}', \tag{24b}$$

$$s \leqslant t_k b_k \log_2\left(1 + \frac{p_k^{\max} g_k^*}{N_0 b_k}\right), \qquad \forall k \in \mathcal{K}', \tag{24c}$$

$$t_k > 0,\ b_k > 0, \qquad \forall k \in \mathcal{K}'. \tag{24d}$$

对于给定的激活端口集合 $\mathcal{S}_k$, 发射功率 p_k 和带宽 b_k, 增大 t_k 只会增大用户完成时延. 因此, 存在一个最优解使传输约束 (24c) 取等号, 即

$$t_k^* = \frac{s}{b_k \log_2\left(1 + \frac{p_k^{\max} g_k^*}{N_0 b_k}\right)}. \tag{25}$$

我们令 $a_k = \frac{EC_k}{f_k^{\max}}$, $\gamma_k = \frac{p_k^{\max} g_k^*}{N_0}$. 其中, a_k 表示用户 k 采用最大 CPU 计算频率时的单轮本地计算时延, γ_k 表示由最大发射功率和最优端口集合共同确定的等效上行信道参数. 用户 k 在带宽 b_k 下的上行传输速率可以写为

$$R_k(b_k) = b_k \log_2\left(1 + \frac{\gamma_k}{b_k}\right). \tag{26}$$

根据式 (25), 用户 k 的最小上行传输时延为

$$t_k(b_k) = \frac{s}{R_k(b_k)}. \tag{27}$$

因此, 子问题 $\mathcal{P}3$ 可以等价转化为以下带宽分配问题 $\mathcal{P}4$:

$$\mathcal{P}4: \quad \min_{\boldsymbol{b},T} \quad T \tag{28}$$

$$\text{s.t.} \quad \sum_{k\in\mathcal{K}'} b_k \leqslant B, \tag{28a}$$

$$J\left[a_k + \frac{s}{R_k(b_k)}\right] \leqslant T, \qquad \forall k \in \mathcal{K}', \tag{28b}$$

$$b_k > 0, \qquad \forall k \in \mathcal{K}'. \tag{28c}$$

对于给定的候选总完成时延 $\widehat{T}$, 用户 k 所允许的最大上行传输时延为

$$d_k(\widehat{T}) = \frac{\widehat{T}}{J} - a_k. \tag{29}$$

为了使候选总完成时延 $\widehat{T}$ 可行, 必须有 $d_k(\widehat{T}) > 0$. 同时, 需要为用户 k 分配足够的带宽, 使其在 $d_k(\widehat{T})$ 内完成大小为 s 的模型更新上传, 即

$$d_k(\widehat{T}) b_k \log_2\left(1 + \frac{\gamma_k}{b_k}\right) \geqslant s. \tag{30}$$

定义满足式 (30) 的最小带宽为 $b_k^{\min}(\widehat{T})$. 该带宽是以下方程在区间 $(0, B]$ 内的解:

$$d_k(\widehat{T}) b_k \log_2\left(1+\frac{\gamma_k}{b_k}\right) - s = 0. \tag{31}$$

对于固定的 $\gamma_k > 0$, 速率函数 $R_k(b_k)$ 关于 b_k 的导数为

$$\frac{\mathrm{d}R_k(b_k)}{\mathrm{d}b_k} = \frac{1}{\ln 2}\left[\ln\left(1+\frac{\gamma_k}{b_k}\right) - \frac{\gamma_k}{b_k+\gamma_k}\right] > 0. \tag{32}$$

因此, $R_k(b_k)$ 关于 b_k 严格单调递增, 方程 (31) 至多存在一个解. 当 $d_k(\widehat{T})R_k(B) \geqslant s$ 时, 方程 (31) 在区间 $(0, B]$ 内存在唯一解, 可以通过内层二分搜索求得. 对于给定的 $\widehat{T}$, 如果所有用户的最小所需带宽均存在, 并且满足

$$\sum_{k\in\mathcal{K}'} b_k^{\min}(\widehat{T}) \leqslant B, \tag{33}$$

则 $\widehat{T}$ 是问题 $\mathcal{P}4$ 的一个可行总完成时延. 随着 $\widehat{T}$ 增大, 用户可用于上行传输的时间 $d_k(\widehat{T})$ 增大, 相应的最小所需带宽 $b_k^{\min}(\widehat{T})$ 单调减小. 式 (33) 关于 $\widehat{T}$ 具有单调可行性, 可以采用外层二分搜索获得最优总完成时延.

令 $K_{\mathrm{r}} = |\mathcal{K}'|$ 表示参与模型校准的用户数量. 由于

$$\lim_{b_k\to\infty} b_k \log_2\left(1+\frac{\gamma_k}{b_k}\right) = \frac{\gamma_k}{\ln 2}, \tag{34}$$

用户 k 的上行传输时延满足理论下界

$$t_k \geqslant \frac{s\ln 2}{\gamma_k}. \tag{35}$$

外层二分搜索的初始下界设置为

$$T_{\mathrm{L}} = J \max_{k\in\mathcal{K}'}\left\{a_k + \frac{s\ln 2}{\gamma_k}\right\}. \tag{36}$$

假设参与校准的用户平均分配带宽, 即 $b_k = \frac{B}{K_{\mathrm{r}}}$. 相应的初始上界可以设置为

$$T_{\mathrm{U}} = J \max_{k\in\mathcal{K}'}\left\{a_k + \frac{s}{\frac{B}{K_{\mathrm{r}}}\log_2\left(1+\frac{\gamma_k K_{\mathrm{r}}}{B}\right)}\right\}. \tag{37}$$

在每次外层二分迭代中, 令 $\widehat{T} = \frac{T_{\mathrm{L}}+T_{\mathrm{U}}}{2}$. 对于该候选完成时延, 首先使用内层二分搜索计算所有用户的最小所需带宽 $b_k^{\min}(\widehat{T})$. 如果 $\sum_{k\in\mathcal{K}'} b_k^{\min}(\widehat{T}) \leqslant B$, 则说明 $\widehat{T}$ 可行, 应令 $T_{\mathrm{U}} \leftarrow \widehat{T}$. 否则, 说明 $\widehat{T}$ 不可行, 应令 $T_{\mathrm{L}} \leftarrow \widehat{T}$. 当上下界满足 $T_{\mathrm{U}} - T_{\mathrm{L}} \leqslant \epsilon_T$ 时, 外层二分搜索终止, 并取 $T^* = T_{\mathrm{U}}$. 重新计算 T^* 对应的最小所需带宽, 可以得到最优带宽分配 $\boldsymbol{b}^*$.

综上, 优化问题 ($\mathcal{P}1$) 的求解过程总结在算法 1 中.

3.3　收敛性和复杂度分析

对于端口选择子问题, 式 (20) 分别考虑选择和不选择当前端口两种情况, 并结合式 (21) 和 (22) 的边界条件遍历全部动态规划状态. 因此, 通过 $F_k(M, Q)$ 反向回溯得到的 $\mathcal{S}_k^*$ 是问题 $\mathcal{P}2$ 的全局最优解, 且该过程经过有限次状态更新后终止. 对于资源分配子问题, 由式 (32) 可知, $R_k(b_k)$ 关于 b_k 严格单调递增, 因而式 (31) 在可行区间内存在唯一解, 内层二分搜索可以收敛至 $b_k^{\min}(\widehat{T})$. 同时, 随着 $\widehat{T}$ 增大, 各用户的最小所需带宽单调减小, 故式 (33) 关于 $\widehat{T}$ 具有单调可行性, 外层二分搜索能够收敛至最优完成时延. 内层和外层二分搜索的最大迭代次数分别为

$$I_b = \left\lceil \log_2 \frac{B}{\epsilon_b} \right\rceil, \qquad I_T = \left\lceil \log_2 \frac{T_{\mathrm{U}} - T_{\mathrm{L}}}{\epsilon_T} \right\rceil. \tag{38}$$

因此, 所提算法能够在精度 ϵ_b 和 ϵ_T 下获得问题 $\mathcal{P}1$ 的全局最优解.

每个用户的动态规划表包含 MQ 个状态, 故全部用户的端口选择复杂度为 $\mathcal{O}(K_{\mathrm{r}}MQ)$. 每次外层搜索需要对所有用户执行内层带宽搜索, 其复杂度为 $\mathcal{O}(K_{\mathrm{r}}I_b)$. 因此, 所提算法的总计算复杂度为 $\mathcal{O}\left(K_{\mathrm{r}}MQ + K_{\mathrm{r}}I_T I_b\right)$.

Algorithm 1 Joint optimization of FA port selection and resource allocation.

Input: Candidate-port channels $\{\boldsymbol{h}_k\}_{k\in\mathcal{K}'}$, M, Q, $\Delta_{\min}$, Δ_{grid}, B, s, J, E, N_0, $\{C_k, f_k^{\max}, p_k^{\max}\}_{k\in\mathcal{K}'}$, and accuracy thresholds ϵ_T and ϵ_b.

Output: Optimal solutions $\boldsymbol{\mathcal{S}}^*$, $\boldsymbol{f}^*$, $\boldsymbol{p}^*$, $\boldsymbol{t}^*$, $\boldsymbol{b}^*$, and T^*.

Calculate $D_{\min}$ according to (18);
for $k \in \mathcal{K}'$ **do**
  Calculate the port weights $\{v_{k,m}\}_{m=1}^{M}$ according to (19);
  Initialize the DP states according to (21) and (22), and update them using (20);
  Obtain $\mathcal{S}_k^*$ by backtracking from $F_k(M,Q)$ and calculate g_k^* according to (23);
  Set $f_k^* = f_k^{\max}$ and $p_k^* = p_k^{\max}$;
  Calculate $a_k = EC_k/f_k^{\max}$ and $\gamma_k = p_k^{\max} g_k^*/N_0$;
end for
Initialize T_{L} and T_{U} according to (36) and (37), respectively;
Set $b_k^{(\text{U})} = B/K_{\text{r}}$, $\forall k \in \mathcal{K}'$;
repeat
  Set $T_{\text{mid}} = (T_{\text{L}} + T_{\text{U}})/2$, $B_{\text{sum}} = 0$, and $\chi = 1$;
  for $k \in \mathcal{K}'$ **do**
    Calculate $d_k(T_{\text{mid}})$ according to (29);
    if $d_k(T_{\text{mid}}) \leqslant 0$ or $d_k(T_{\text{mid}})R_k(B) < s$ **then**
      Set $\chi = 0$ and **break**;
    end if
    Obtain $b_k^{\min}(T_{\text{mid}})$ by applying bisection search to (31) over $(0, B]$ with accuracy ϵ_b;
    Update $B_{\text{sum}} \leftarrow B_{\text{sum}} + b_k^{\min}(T_{\text{mid}})$;
  end for
  if $\chi = 1$ and $B_{\text{sum}} \leqslant B$ **then**
    Set $T_{\text{U}} \leftarrow T_{\text{mid}}$ and $\boldsymbol{b}^{(\text{U})} \leftarrow \boldsymbol{b}^{\min}(T_{\text{mid}})$;
  else
    Set $T_{\text{L}} \leftarrow T_{\text{mid}}$;
  end if
until $T_{\text{U}} - T_{\text{L}} \leqslant \epsilon_T$;
Set $T^* = T_{\text{U}}$ and $\boldsymbol{b}^* = \boldsymbol{b}^{(\text{U})}$;
for $k \in \mathcal{K}'$ **do**
  Calculate t_k^* according to (27);
end for
return $\boldsymbol{\mathcal{S}}^*$, $\boldsymbol{f}^*$, $\boldsymbol{p}^*$, $\boldsymbol{t}^*$, $\boldsymbol{b}^*$, and T^*.

4　仿真分析

为验证所提流态天线辅助联邦遗忘方案的有效性, 本文分别从通信完成时延和联邦遗忘效果两个方面开展仿真实验. 除非另有说明, 系统中共有 $K = 10$ 个用户, 其中一个用户提出遗忘请求, 其余 $K_{\text{r}} = 9$ 个用户参与模型校准. 载波频率设置为 30 GHz, 系统总带宽为 20 MHz, 流态天线归一化孔径为 $W = 2$, 候选端口数为 $M = 41$, 激活端口数为 $Q = 3$, 最小激活端口间距为 0.5λ. 用户最大发射功率为 23 dBm, 噪声功率谱密度为 -167 dBm/Hz, 联邦遗忘校准轮数为 10, 每轮本地迭代次数为 5, 模型更新数据量为 5 Mbit.

用户到基站的距离在 50∼200 m 内随机生成, 路径损耗指数为 2.2, 阴影衰落标准差为 4 dB, 默认 Rician 因子为 5 dB. 用户每轮本地计算量在 0.8×10^8∼1.2×10^8 CPU cycles 之间随机生成, 最大 CPU 计算频率在 1.5∼2.5 GHz 之间随机取值. 通信实验均进行 1000 次独立 Monte Carlo 仿真, 并对所得结果取平均值.

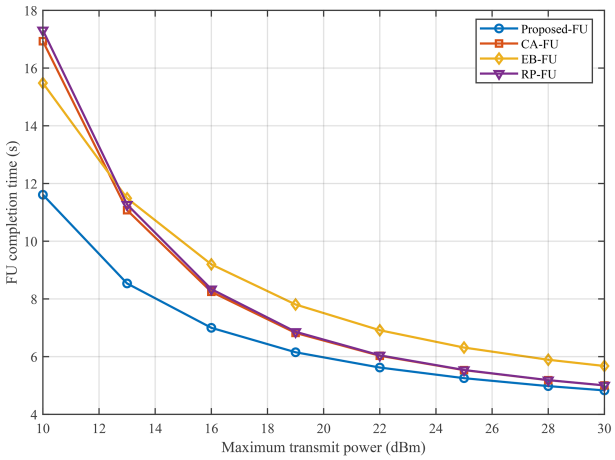

图 2　(网络版彩图) 最大上行发射功率与 FU 完成时延的关系.

Figure 2　(Color online) FU completion time versus maximum uplink transmit power.

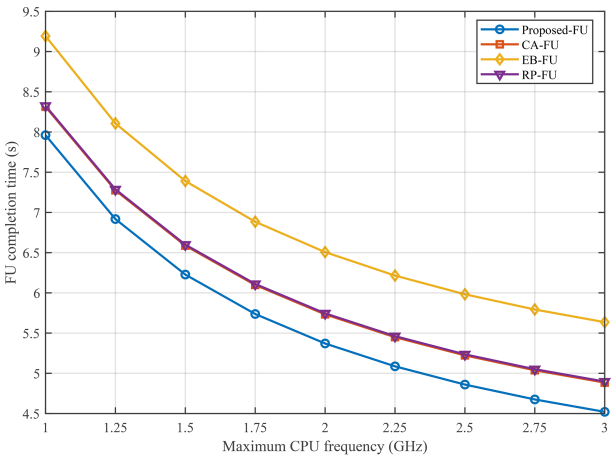

图 3　(网络版彩图) 最大 CPU 计算频率与 FU 完成时延的关系.

Figure 3　(Color online) FU completion time versus maximum CPU frequency.

通信性能实验采用以下四种方案进行比较.

• **Proposed-FU:** 采用本文提出的端口动态规划和最优带宽分配方法, 联合优化流态天线激活端口和通信, 计算资源.

• **CA-FU:** 固定选择 Q 个均匀分布端口来模拟传统天线, 但仍对带宽, 发射功率和计算频率进行优化.

• **EB-FU:** 采用与本文相同的流态天线端口选择方法, 但将系统带宽平均分配给所有参与用户.

• **RP-FU:** 随机选择满足最小间距约束的流态天线端口, 并采用最优带宽分配.

联邦遗忘性能实验采用 Fashion-MNIST 数据集, 将训练数据按照 Dirichlet 参数 0.5 非独立同分布地划分给 10 个用户. 原始联邦学习执行 20 轮全局训练, 每个用户每轮执行 2 个本地训练周期. 本文方案使用保留用户的历史更新进行模型校准, 并与未执行遗忘的 Original-FL 模型和删除目标用户后从头训练的 FedRetrain 模型进行比较. 所有结果均为 3 个随机种子实验的平均结果.

图 2 给出了用户最大发射功率与 FU 完成时延之间的关系. 随着最大发射功率由 10 dBm 增加至 30 dBm, 所有方案的完成时延均显著下降. 这是因为发射功率增加提高了上行信噪比和模型更新传输速率, 从而缩短了上行通信时延. 在整个功率范围内, Proposed-FU 均取得了最低的完成时延, 且其优势在功率受限场景下尤为明显. 当最大发射功率为 10 dBm 时, Proposed-FU 的完成时延约为 11.61 s, 相较于 CA-FU, EB-FU 和 RP-FU 分别降低约 31.4%, 25.0% 和 32.9%. 这是因为在低功率条件下, 上行通信是系统完成时延的主要瓶颈, 通过选择信道增益较大的流态天线端口可以显著改善等效信道质量. 随着发射功率继续增大, 不同端口选择方案之间的差距逐渐缩小. 当功率达到 30 dBm 时, Proposed-FU 的完成时延约为 4.83 s. 此时上行通信时延已经较小, 本地计算时延在总完成时延中占据更高比例, 继续增加发射功率或改善信道所能获得的收益有限.

图 3 给出了用户最大 CPU 计算频率与 FU 完成时延之间的关系. 随着最大 CPU 频率由 1 GHz 增加至 3 GHz, 所有方案的完成时延均持续下降. 这是因为提高 CPU 频率能够直接缩短用户的本地模型校准时间. Proposed-FU 的完成时延由约 7.96 s 下降至 4.52 s. 当最大 CPU 频率为 3 GHz 时, 其相较于 CA-FU, EB-FU 和 RP-FU 分别降低约 7.4%, 19.8% 和 7.7%. 随着 CPU 频率增加, 曲线斜率逐渐减小, 表明在计算能力较强时, 系统瓶颈逐渐从本地计算转移至上行模型传输. Proposed-FU 在不同 CPU 频率下始终具有最低时延, 说明通信资源优化和计算资源提升具有互补作用.

图 4 给出了激活端口数 Q 对 FU 完成时延的影响. 当 Q 从 1 增加至 4 时, Proposed-FU 的完成时延由约 6.09 s 下降至 5.41 s, 说明增加激活端口可以获得更高的合并信道增益. 但是, 时延下降幅度随 Q 增加而逐渐减

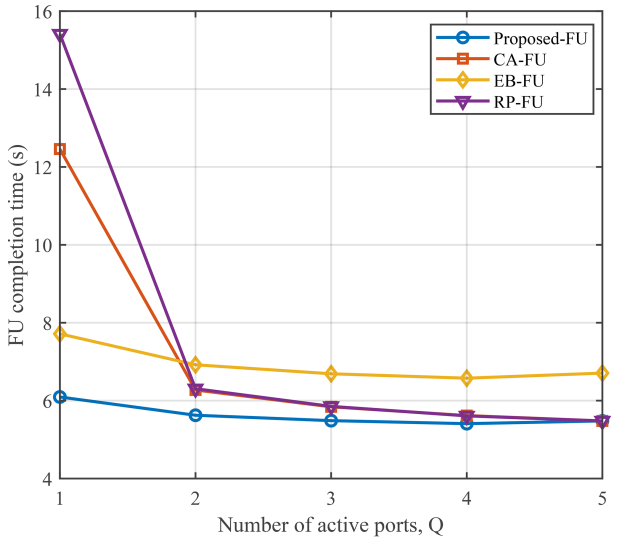

图 4　(网络版彩图) 激活端口数与 FU 完成时延的关系.

Figure 4　(Color online) FU completion time versus the number of active ports.

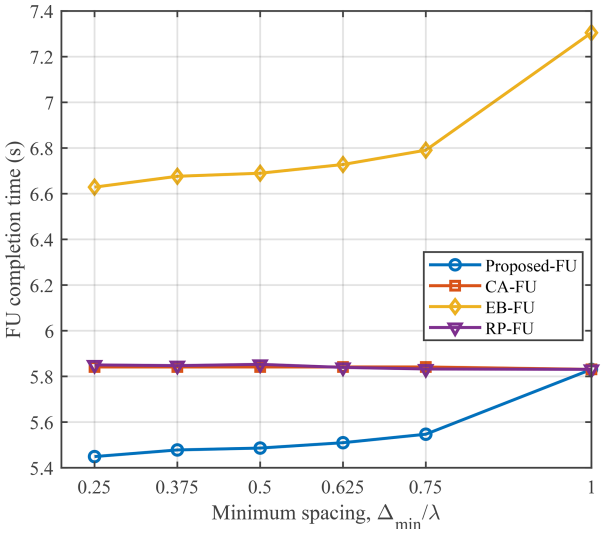

图 5　(网络版彩图) 最小端口间距与 FU 完成时延的关系.

Figure 5　(Color online) FU completion time versus minimum port spacing.

小, 表明激活端口数量存在明显的边际收益递减现象. 当 $Q = 1$ 时, Proposed-FU 相较于 CA-FU 和 RP-FU 表现出显著优势. 这是因为在单端口激活条件下, 端口位置直接决定用户的瞬时信道质量, 最优端口选择能够充分利用不同候选位置间的信道差异. 当 $Q = 5$ 时, Proposed-FU, CA-FU 和 RP-FU 的结果趋于一致. 因为在给定孔径和最小间距约束下, 5 个端口已经达到可激活端口数上限, 满足间距约束的端口组合基本唯一, 端口选择自由度随之消失. 因此, 实际部署中候选端口数 M 应明显大于激活端口数 Q, 同时不宜令 Q 过于接近给定孔径内的最大可行端口数.

图 5 进一步给出了最小端口间距的影响. 随着最小间距由 0.25λ 增大至 λ, Proposed-FU 的完成时延由约 5.45 s 增加至 5.83 s. 这是因为间距约束增大会减少可行端口组合, 使用户难以同时激活多个信道质量较好的端口. 当最小间距达到 λ 时, 可行组合空间大幅缩小, Proposed-FU, CA-FU 和 RP-FU 的性能趋于一致. 该结果说明, 最小间距约束不仅影响天线互耦和空间相关性, 也会直接决定流态天线端口优化所能获得的选择增益.

图 6 给出了归一化流态天线孔径 W 与 FU 完成时延之间的关系. 随着 W 由 1 增加至 4, Proposed-FU 的完成时延由约 5.83 s 下降至 5.36 s, 而 CA-FU 和 RP-FU 基本保持在 5.84 s 附近. 这说明增大物理孔径能够降低不同候选位置之间的空间相关性, 并扩展满足最小间距约束的端口组合, 从而为最优端口选择提供更大的空间自由度. 当 $W = 1$ 时, Proposed-FU, CA-FU 和 RP-FU 的性能基本一致. 这是因为较小的物理孔径和最小间距约束共同限制了可行端口组合. 随着孔径增大, Proposed-FU 与固定端口和随机端口方案之间的差距逐渐扩大, 验证了流态天线空间可重构能力在具有足够物理孔径时能够有效降低系统时延.

图 7 给出了 Rician 因子对系统性能的影响. 在 Rayleigh 衰落和较小 Rician 因子条件下, Proposed-FU 相较其他端口选择方案具有较明显的时延优势. 随着 Rician 因子增大, Proposed-FU 的完成时延略有增加, 而 CA-FU 和 RP-FU 的完成时延逐渐下降, 三者之间的差距随之缩小. 较大的 Rician 因子意味着 LoS 分量占据主导地位, 不同候选端口处的信道幅度差异相对减小. 因此, 通过比较瞬时信道质量进行端口选择所能获得的增益降低. 该结果同时表明, 本文所考虑的空间相关 Rician 模型能够合理体现流态天线的适用条件. 在具有一定散射分量的位置相关信道中, 端口选择收益更加明显. 在强 LoS 且不同端口幅度接近的场景下, 其收益会有所减弱, 但不会立即完全消失.

图 8 给出了系统总带宽与 FU 完成时延之间的关系. 随着总带宽由 5 MHz 增加至 50 MHz, 所有方案的完成时延均明显下降, 但下降速度逐渐减缓. 这是因为增加带宽能够提高用户的上行传输速率, 而当通信时延降

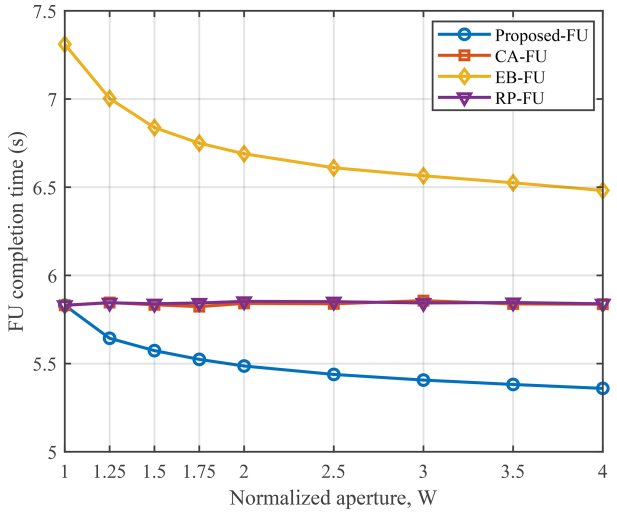

图 6　(网络版彩图) 归一化流态天线孔径与 FU 完成时延的关系.

Figure 6　(Color online) FU completion time versus normalized fluid antenna aperture.

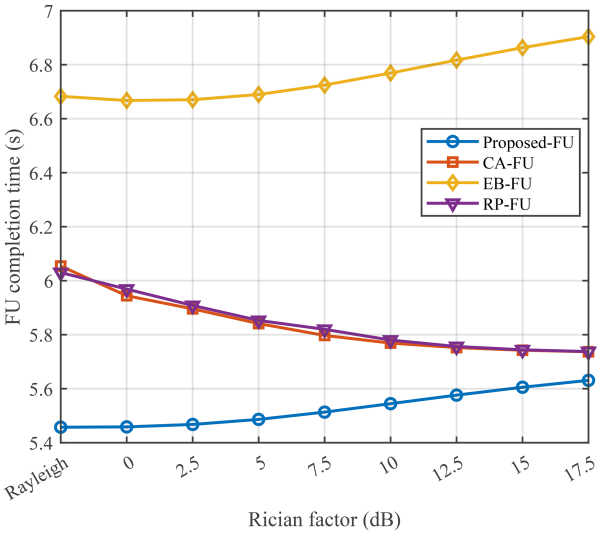

图 7　(网络版彩图) Rician 因子与 FU 完成时延的关系.

Figure 7　(Color online) FU completion time versus Rician factor.

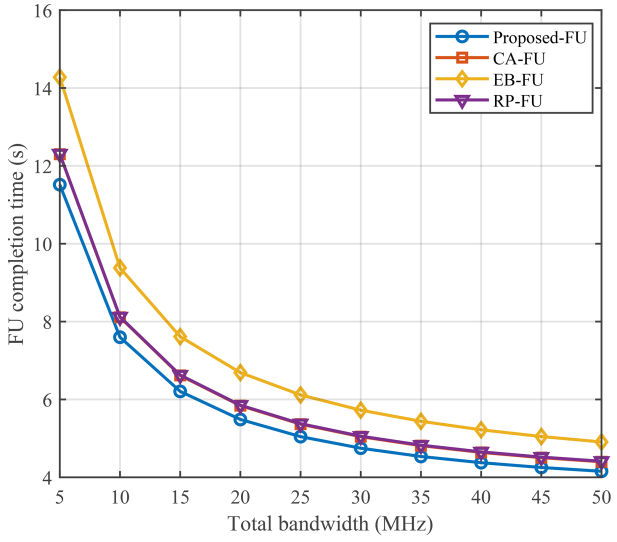

图 8　(网络版彩图) 系统总带宽与 FU 完成时延的关系.

Figure 8　(Color online) FU completion time versus total system bandwidth.

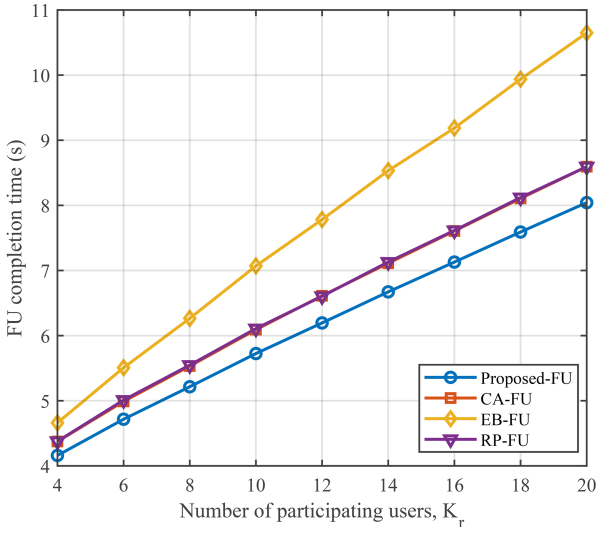

图 9　(网络版彩图) 参与用户数与 FU 完成时延的关系.

Figure 9　(Color online) FU completion time versus the number of participating users.

低到一定程度后, 本地计算时延成为主要瓶颈, 继续增加带宽所带来的边际收益逐渐减小. 在默认带宽 20 MHz 下, Proposed-FU 的完成时延约为 5.49 s, 相较于 CA-FU, EB-FU 和 RP-FU 分别降低约 6.1%, 18.0% 和 6.3%. EB-FU 在整个带宽范围内均表现出较大的完成时延, 说明同步 FU 系统的性能由最慢用户决定, 平均带宽分配无法根据用户信道质量和计算能力差异调整资源, 容易造成带宽利用效率下降.

图 9 给出了参与校准用户数 K_{r} 对完成时延的影响. 随着参与用户数由 4 增加至 20, 所有方案的完成时延近似单调增加. 这是因为更多用户需要共享有限的系统带宽, 同时同步聚合机制要求等待最慢用户完成本地计算和模型上传, 从而增加了系统瓶颈用户出现的概率. 当 $K_{\mathrm{r}} = 20$ 时, Proposed-FU 的完成时延约为 8.04 s, 而

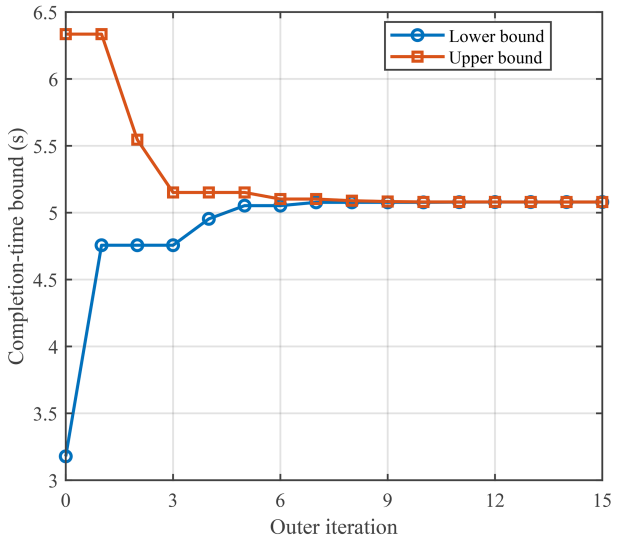

图 10　(网络版彩图) 完成时延上下界间隙的收敛过程.

Figure 10　(Color online) Convergence of the completion-time bound gap.

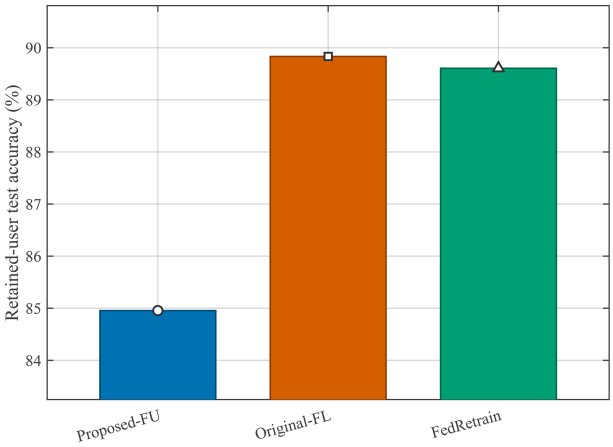

图 11　(网络版彩图) 不同联邦遗忘方案的剩余用户测试集准确率.

Figure 11　(Color online) Retained user test accuracy of different federated unlearning schemes.

CA-FU, EB-FU 和 RP-FU 分别约为 8.59, 10.65 和 8.59 s. Proposed-FU 在用户数量增加时仍保持稳定优势, 说明联合端口选择和带宽优化能够有效缓解多用户场景中的带宽竞争问题.

图 10 给出了外层二分搜索过程中完成时延上下界及其间隙的变化情况. 可以看到, 随着迭代次数增加, 下界逐渐增大, 上界逐渐减小, 并最终收敛至约 5.08 s. 上下界在约 6 次迭代后已经在图示尺度下基本重合.

图 11 比较了不同方案在剩余用户测试集上的模型准确率. Original-FL, Proposed-FU 和 FedRetrain 的平均准确率分别为 89.83%, 84.96% 和 89.61%. Proposed-FU 相较 FedRetrain 下降约 4.65 个百分点, 但仍保留了 FedRetrain 约 94.8% 的剩余用户测试精度. 这说明所提方法在无需完全重新训练模型的情况下, 能够保留大部分模型效用, 但模型校准过程仍会造成一定精度损失.

图 12 和 13 进一步比较了不同方案在被遗忘用户数据上的准确率和交叉熵损失. Original-FL 尚未执行遗忘, 其准确率最高, 约为 91.06%, 损失最低, 约为 0.246, 说明原模型仍然充分保留了对该用户数据的拟合信息. 执行所提遗忘方法后, 被遗忘用户数据上的准确率下降至约 79.05%, 损失增加至约 0.537, 表明模型对被遗忘用户数据的预测能力明显减弱. FedRetrain 在被遗忘用户数据上的准确率和损失分别约为 86.05% 和 0.378. Proposed-FU 产生了比 FedRetrain 更低的准确率和更高的损失, 说明其对目标用户数据的抑制程度更强. 该差异也表明 Proposed-FU 没有完全复现 FedRetrain 的模型行为, 存在一定程度的过度遗忘现象. 结合图 11 可知, 所提方法以部分剩余用户精度损失为代价, 增强了对目标用户数据的遗忘效果, 体现了模型效用与遗忘强度之间的权衡.

5　总结与展望

本文研究了流态天线辅助无线联邦遗忘系统中的完成时延最小化问题. 针对固定物理孔径内候选端口信道的空间相关性, 建立了基于空间相关 Rician 信道的系统模型, 并构建了流态天线激活端口、CPU 计算频率、上行发射功率、传输时延和带宽分配的联合优化问题. 利用完成时延关于计算频率、发射功率和等效信道增益的单调性, 将原问题分解为端口选择与资源分配两个子问题, 分别采用动态规划和双层二分搜索进行求解. 仿真结果表明, 所提方案在不同发射功率、系统带宽、用户数量和流态天线配置下均能有效降低 FU 完成时延, 尤其适用于功率和带宽受限场景. Fashion-MNIST 实验进一步表明, 所提方法能够削弱模型对被遗忘用户数据的预测能力, 并保留大部分剩余用户测试精度.

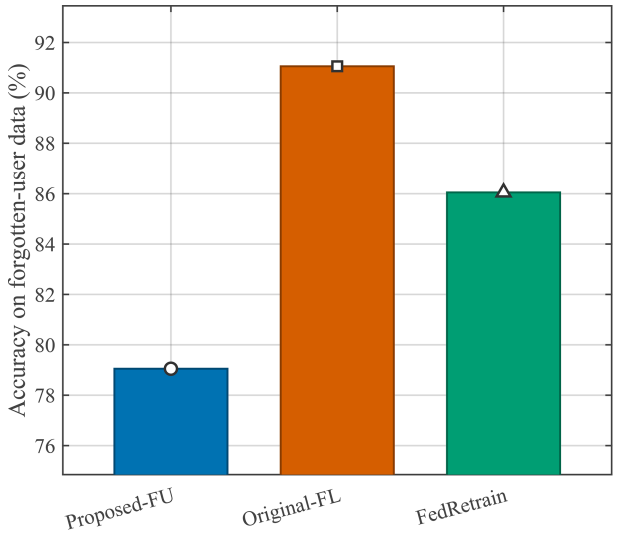

图 12　(网络版彩图) 被遗忘用户数据上的准确率.

Figure 12　(Color online) Accuracy on the forgotten user's data.

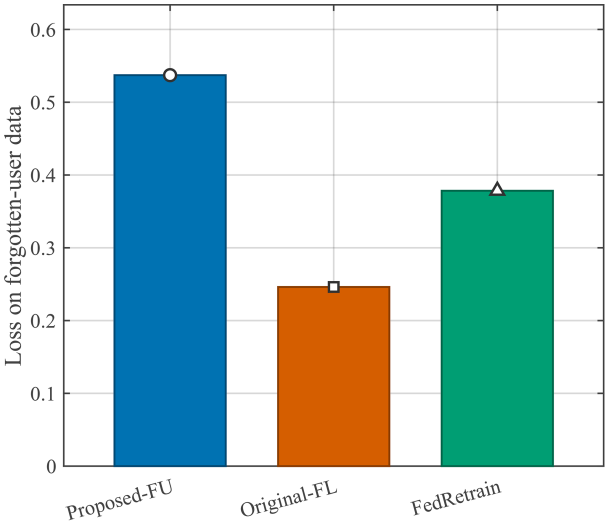

图 13　(网络版彩图) 被遗忘用户数据上的损失.

Figure 13　(Color online) Loss on the forgotten user's data.

本文仍存在一定局限性. 当前模型假设仅有一个用户提出遗忘请求, 并忽略下行传输, 候选端口 CSI 获取和端口切换开销, 同时采用固定校准轮数, 准静态窄带信道和完美 CSI. 未来可将遗忘用户扩展为集合 $\mathcal{U} \subseteq \mathcal{K}$, 研究多用户遗忘请求下校准轮数, 遗忘效果与通信资源的联合优化, 进一步考虑不完美 CSI、用户移动、宽带信道及流态天线实际开销, 建立更加符合实际部署条件的系统模型. 此外, 可结合成员推理攻击、参数距离和遗忘认证等指标评价隐私残留风险, 并研究自适应校准方法, 以减小过度遗忘及其造成的模型精度损失.

参考文献

1 Zhou Z Y, Liu P J, Feng J H, et al. Computation resource allocation and task assignment optimization in UAV-enabled mobile edge computing. IEEE Trans Mobile Comput, 2020, 19: 2577–2591

2 Shi W, Cao J, Zhang Q, et al. Edge computing: vision and challenges. IEEE Int Things J, 2016, 3: 637–646

3 McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data. In: Proceedings of AISTATS, 2017. 54: 1273–1282

4 Yang Q, Liu Y, Chen T, et al. Federated machine learning: concept and applications. ACM Trans Intell Syst Technol, 2019, 10: 1–19

5 Xiao X, Tang Z, Xiao B, et al. A survey on privacy and security issues in federated learning. Chin J Comput, 2023, 46: 1019–1044 [肖雄, 唐卓, 肖斌, 等. 联邦学习的隐私保护与安全防御研究综述. 计算机学报, 2023, 46: 1019–1044]

6 European Parliament and Council of the European Union. EU Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data. Official Journal of the European Union. 2016

7 State of California. California Consumer Privacy Act (CCPA). California Civil Code. 2018

8 Cao Y Z, Yang J F. Towards making systems forget with machine unlearning. In: Proceedings of IEEE Symp Security Privacy, 2015. 463–480

9 Bourtoule L, Chandrasekaran V, Choquette-Choo C A, et al. Machine unlearning. In: Proceedings of IEEE Symp Security Privacy, 2021. 141–159

10 Liu G Y, Ma X Q, Yang Y, et al. FedEraser: enabling efficient client-level data removal from federated learning models. In: Proceedings of IEEE IWQoS, 2021. 1–10

11 Zhao Y, Wang P, Qi H, et al. Federated unlearning with momentum degradation. IEEE Int Things J, 2024, 11: 8860–8870

12 Yang Z, Chen M, Saad W, et al. Energy efficient federated learning over wireless communication networks. IEEE Trans Wireless Commun, 2021, 20: 1935–1949

13 Yang H H, Liu Z, Quek T Q S, et al. Scheduling policies for federated learning in wireless networks. IEEE Trans Commun, 2020, 68: 317–333

14 Mohammadi Amiri M, Gunduz D. Machine learning at the wireless edge: distributed stochastic gradient descent over-the-air. IEEE Trans Signal Process, 2020, 68: 2155–2169

15 Du Z, Wu C, Yoshinaga T, et al. Federated learning for vehicular Internet of Things: recent advances and open issues. IEEE Open J Comput Soc, 2020, 1: 45–61

16 Tran N H, Bao W, Zomaya A Y, et al. Federated learning over wireless networks: optimization model design and analysis. In: Proceedings of IEEE INFOCOM, 2019. 1387–1395

17 Chen M, Poor H V, Saad W, et al. Convergence time optimization for federated learning over wireless networks. IEEE Trans Wireless Commun, 2021, 20: 2457–2471

18 Wang S, Tuor T, Salonidis T, et al. Adaptive federated learning in resource constrained edge computing systems. IEEE J Sel Areas Commun, 2019, 37: 1205–1221

19 Wang Z, Zhang J, Du H, et al. Extremely large-scale MIMO: fundamentals, challenges, solutions, and future directions. IEEE Wireless Commun, 2024, 31: 117–124

20 Zhu L, Ma W, Zhang R. Movable antennas for wireless communication: opportunities and challenges. IEEE Commun Mag, 2024, 62: 114–120

21 Wong K K, Shojaeifard A, Tong K F, et al. Fluid antenna systems. IEEE Trans Wireless Commun, 2021, 20: 1950–1962

22 Psomas C, Smith P J, Suraweera H A, et al. Continuous fluid antenna systems: modeling and analysis. IEEE Commun Lett, 2023, 27: 3370–3374

23 Wong K K, Tong K F, Chae C B. Fluid antenna system-part II: research opportunities. IEEE Commun Lett, 2023, 27: 1924–1928

24 Khammassi M, Kammoun A, Alouini M S. A new analytical approximation of the fluid antenna system channel. IEEE Trans Wireless Commun, 2023, 22: 8843–8858

25 Li F, Zhang J, Li T, et al. Dynamic-weighted joint beamforming design for communication and sensing in fluid antenna-assisted ISAC system. J Commun, 2025, 46: 101–114 [李飞, 张杰, 李汀, 等. 流体天线辅助的 ISAC 系统通感动态加权联合波束成形设计. 通信学报, 2025, 46: 101–114]

26 New W K, Wong K K, Xu H, et al. An information-theoretic characterization of MIMO-FAS: optimization, diversity-multiplexing tradeoff and q-outage capacity. IEEE Trans Wireless Commun, 2024, 23: 5541–5556

27 Vega-Sánchez J D, López-Ramírez A E, Urquiza-Aguiar L, et al. Novel expressions for the outage probability and diversity gains in fluid antenna system. IEEE Wireless Commun Lett, 2024, 13: 372–376

28 Chen Y, Chen M, Xu H, et al. Joint beamforming and antenna design for near-field fluid antenna system. IEEE Wireless Commun Lett, 2025, 14: 415–419

29 Lu Y, Dai L. Near-field channel estimation in mixed LoS/NLoS environments for extremely large-scale MIMO systems. IEEE Trans Commun, 2023, 71: 3694–3707

Fluid antenna assisted federated unlearning system

Yixuan CHEN[1], Zhaohui YANG[1*], Hao XU[2], Cheng ZENG[3], Yuqi CHEN[1], Kai-Kit WONG[4,5] & Zhaoyang ZHANG[1]

1. *College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China*
2. *National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China*
3. *School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China*
4. *Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, UK*
5. *Department of Electronic Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do 17104, Republic of Korea*

* Corresponding author. E-mail: yang_zhaohui@zju.edu.cn

Abstract Federated unlearning (FU) enables users to revoke their data contributions from federated learning models, but its multiple model-calibration rounds introduce considerable computation and communication overhead. To reduce the FU completion time in wireless edge networks, this paper investigates a fluid antenna (FA)-assisted wireless FU system. Under a spatially correlated Rician channel, a mixed discrete-continuous problem is formulated to jointly optimize the active FA port sets, CPU frequencies, uplink transmit powers, transmission delays, and bandwidth allocation. A joint optimization algorithm based on dynamic programming and nested bisection search is proposed. Dynamic programming determines the optimal port sets subject to the minimum spacing constraint, while nested bisection search obtains the communication resource allocation. Simulation results show that the proposed scheme effectively reduces the FU completion time compared with the fixed-antenna, equal-bandwidth, and random-port-selection schemes. At a maximum transmit power of 10 dBm, it achieves an approximately 31.4% delay reduction over the fixed-antenna scheme. Fashion-MNIST experiments further evaluate the post-unlearning model utility and unlearning effectiveness.

Keywords federated unlearning, fluid antenna, port selection, resource allocation, latency minimization