



基于无线传播环境的无蜂窝大规模 MIMO 系统接入点部署优化

姜静¹, 刘永强¹, 严冯洋¹, 陶莎¹, Worakrin Sutthiphan²

(1. 西安邮电大学陕西省信息通信网络及安全重点实验室, 陕西 西安 710121;

2. 泰国国家电信有限公司, 泰国 曼谷 10700)

摘要: 无蜂窝大规模多输入多输出 (MIMO) 系统通过在覆盖区域内部署大量的接入点 (AP), 可以为用户提供均匀、可靠的服务。传统的无蜂窝大规模 MIMO 系统采用随机部署, 未考虑 AP 周围的路径损耗、阴影衰落散射物以及环境遮挡对覆盖质量的影响。为了考虑实际环境下无蜂窝大规模 MIMO 能实现均匀、一致的覆盖, 提出了基于无线传播环境的 AP 部署方案。首先, 通过混合概率路径损耗模型对无线传播环境进行表征, 其次构建了以最大化平均吞吐量为目标的 AP 部署优化问题, 最后将问题转化为马尔可夫博弈过程, 并且基于多智能体深度确定性策略梯度 (MADDPG) 算法得出最优的 AP 部署策略。仿真结果表明, 相比于传统的随机部署和现有 AP 部署策略, 所提方案可明显改善复杂环境下的非均匀覆盖问题, 为用户提供良好一致的均匀覆盖。

关键词: 无蜂窝大规模 MIMO; AP 部署; 混合概率路径损耗模型; MADDPG 算法

中图分类号: TN929.5

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2024022

Environment-aware based access point deployment optimization for cell-free massive MIMO system

JIANG Jing¹, LIU Yongqiang¹, YAN Fengyang¹, TAO Sha¹, WORAKRIN Sutthiphan²

1. Shaanxi Key Laboratory of Information Communication Network and Security, Xi'an
University of Posts and Telecommunication, Xi'an 710121, China

2. National Telecommunication Public Company, Bangkok 10700, Thailand

Abstract: Cell-free massive multiple-input multiple-output (MIMO) systems deploy a large number of access point (AP) across the coverage area which can provide uniform high-rate services to users. However, the quality of coverage would be affected by path loss, shadow fading scatters, and environmental occlusions around the randomly placed AP in conventional cell-free massive MIMO systems that do not consider their impact. Considering the impact of ac-

收稿日期: 2023-08-08; 修回日期: 2024-02-01

基金项目: 国家自然科学基金资助项目 (No.61871321, No.61901370); 陕西省国际科技合作项目重点计划项目 (No.2023-GHZD-37); 陕西省重点产业链项目 (No.2023-ZDLGY-49)

Foundation Items: The National Natural Science Foundation of China (No.61871321, No.61901370), Key Program for International Science and Technology Cooperation Projects of Shaanxi Province (No.2023-GHZD-37), Key Industrial Chain Projects of Shaanxi Province (No.2023-ZDLGY-49)



tual wireless propagation environments, an AP deployment scheme was proposed to acquire uniform and consistent coverage. Firstly, a hybrid probabilistic path loss model was utilized to characterize various wireless propagation environments. Then, the AP deployment optimization problem was solved with the objective of maximizing the average throughput. Finally, the problem was transformed into a Markov game process and solved by the multi-agent deep deterministic policy gradient (MADDPG) algorithm. The simulation results demonstrate that the proposed scheme can provide more uniform coverage in complex environments and serve users with reliable and consistent service compared to random AP deployment and existing AP deployment methods.

Key words: cell-free massive MIMO, AP deployment, hybrid probabilistic path loss model, MADDPG algorithm

0 引言

无蜂窝大规模多输入多输出 (multiple-input multiple-output, MIMO) 系统作为 6G 的关键潜在技术之一, 通过引入“以用户为中心”的思想能够在覆盖区域内为用户提供一致良好的服务^[1-3]。无蜂窝大规模 MIMO 系统的关键思想是在覆盖区域内部署大量的接入点 (access point, AP), 所有 AP 通过回程链路连接到中央处理单元 (central processing unit, CPU), 并且利用相同的时间-频率资源协作服务所有用户^[4-5]。无蜂窝大规模 MIMO 突破了“以小区为中心”的传统蜂窝覆盖思想, 有效消除了蜂窝边界, 避免了频繁的小区间切换, 大幅提升了用户的服务体验, 并且多 AP 协作传输带来的宏分集增益, 可以为用户提供更加稳定可靠的链路质量^[6-8]。

在现有无蜂窝大规模 MIMO 的研究中, 通常假设覆盖区域内的 AP 均匀、随机地部署^[9-10]。然而这种假设过于理想化, 不符合无蜂窝网络的实际部署情况。文献[11]将无蜂窝大规模 MIMO 系统中的 AP 位置建模为一个更符合实际的空间随机性模型, 即泊松点过程 (Poisson point process, PPP), 并借助随机几何工具推导了下行链路覆盖概率和可达速率, 其结果表明, 覆盖率随着 AP 密度的增加越来越高, 并在达到峰值后饱和。文献[12]同样采用 PPP 模型表征 AP 的空间随机性, 利用随机几何工具得到平均频谱效率的下界, 并在提出的实际无蜂窝大规模 MIMO 系统的功率消

耗模型下, 以导频复用因子和 AP 密度的形式获得了单位面积最优能效闭式解。文献[13]假设用户移动性和传播信道条件给定的情况下, 通过将遗传算法 (genetic algorithm, GA) 与射线追踪 (ray tracer, RT) 信道模拟器相结合, 生成了以最大化系统容量为目标的最优 AP 部署方案。文献[14]针对无蜂窝大规模 MIMO 系统, 分别提出了最大化最小吞吐量和最大化总吞吐量算法对 AP 位置进行优化, 其数值结果表明, 与传统的随机部署算法相比, 所提算法显著提升了总吞吐量。

上述工作证明 AP 部署优化可以显著改善系统性能, 然而在无蜂窝大规模 MIMO 系统应用的实际环境中, 由于 AP 的地理分布性质和各种障碍物 (如城市地区的建筑物、树木等) 的影响, AP 分布位置的不同可能会造成与用户之间的信道经历不同的路径损耗, 进而导致覆盖区域内用户的服务质量出现差异化, 部分用户的通信质量严重降低。因此在 AP 部署时应充分考虑无线传播环境造成的大尺度衰落、阴影衰落以及建筑物的影响, 针对环境信息优化 AP 部署, 从而实现真实环境下无蜂窝大规模 MIMO 系统均匀、一致的覆盖。

本文针对无蜂窝大规模 MIMO 系统提出了基于无线传播环境的 AP 部署方案。首先, 通过视距 (line of sight, LOS) 链路概率模型对 LOS 和非视距 (non line of sight, NLOS) 路径进行加权, 将路径损耗建模为混合概率路径损耗模型对真实的无线传播环境进行表征, 然后将 AP 部署问题

建模为以最大化平均吞吐量为目标的 0-1 决策问题，最后将原问题转化为马尔可夫博弈过程，通过提出的基于多智能体深度确定性策略梯度（multi-agent deep deterministic policy gradient, MADDPG）的 AP 部署优化算法求解出最优 AP 部署策略。仿真结果表明，与传统的随机部署和现有 AP 部署算法相比，本文提出的 AP 部署策略可以实现均匀一致覆盖并提升系统性能。

1 系统模型

去蜂窝大规模 MIMO 系统模型示意图如图 1 所示，考虑具有 M 个 AP 的无蜂窝大规模 MIMO 系统，其中每个 AP 配备有 N_T 个天线，在覆盖区域内同时同频服务于 K ($K \ll M$) 个单天线用户，所有的 AP 通过回程链路与 CPU 连接。假设整个系统工作在时分双工（time-division duplex, TDD）模式，第 m 个 AP 和第 k 个用户之间的信道表示如下。

$$\mathbf{h}_{mk} = \sqrt{\beta_{mk}} \boldsymbol{\alpha}_{mk} \quad (1)$$

其中， $\mathbf{h}_{mk} \in \mathbf{C}^{N_T \times 1}$ ； β_{mk} 表示第 m 个 AP 和第 k 个用户之间的大尺度衰落，表征路径损耗和阴影衰落。此外， $\boldsymbol{\alpha}_{mk} \in \mathbf{C}^{N_T \times 1}$ 为小尺度衰落向量，且假设 $\boldsymbol{\alpha}_{mk} \sim \mathcal{CN}(0, \mathbf{I}_{N_T})$ 。

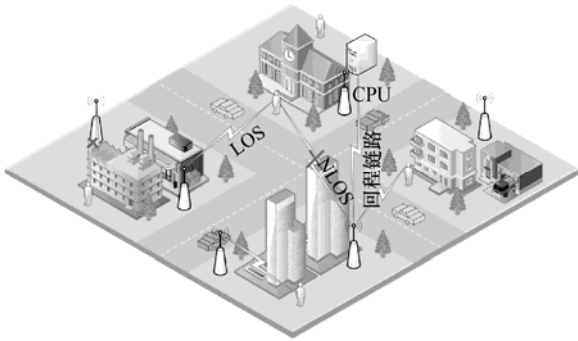


图 1 去蜂窝大规模 MIMO 系统模型示意图

1.1 上行链路导频训练

设 $\boldsymbol{\phi}_k \in \mathbf{C}^{\tau \times 1}$ 为第 k 个用户发送的长度为 τ 的上行导频信号，不同用户之间的导频序列相互正交，且满足 $\|\boldsymbol{\phi}_k\|^2 = 1$ 。则第 m 个 AP 接收到的第 k 个用

户发送的导频信号 $\mathbf{y}_m \in \mathbf{C}^{N_T \times \tau}$ 为：

$$\mathbf{y}_m = \sum_{k=1}^K \sqrt{\rho_t} \mathbf{h}_{mk} \boldsymbol{\phi}_k^H + \mathbf{w}_m \quad (2)$$

其中， $\mathbf{w}_m \in \mathbf{C}^{N_T \times \tau}$ 为加性高斯白噪声矩阵，其中矩阵元素为服从独立同分布（independent identically distributed, i.i.d.）随机变量。基于接收到的导频信号 $\mathbf{y}_m$ ，第 m 个 AP 采用最小均方误差^[15]（minimum mean squared error, MMSE）对信道 $\mathbf{h}_{mk}$ 进行估计可以得到 $\hat{\mathbf{h}}_{mk}$ ：

$$\hat{\mathbf{h}}_{mk} = c_{mk} (\mathbf{y}_m \boldsymbol{\phi}_k)^H \quad (3)$$

其中， $c_{mk} = \frac{\sqrt{\tau \rho_t} \beta_{mk}}{\tau \rho_t \sum_{i=1}^K \beta_{mi} |\boldsymbol{\phi}_i^H \boldsymbol{\phi}_k|^2 - 1}$ 为标准化参数，

信道估计误差向量可以表示如下。

$$\tilde{\mathbf{h}}_{mk} = \mathbf{h}_{mk} - \hat{\mathbf{h}}_{mk} \quad (4)$$

其中， $\hat{\mathbf{h}}_{mk} \sim \mathcal{CN}(0, \mathcal{G}_{mk})$ ， $\tilde{\mathbf{h}}_{mk} \sim \mathcal{CN}(0, \beta_{mk} - \mathcal{G}_{mk})$ ，

且估计信道的方差 $\mathcal{G}_{mk} = \frac{\tau \rho_t \beta_{mk}^2}{\tau \rho_t \sum_{i=1}^K \beta_{mi} |\boldsymbol{\phi}_i^H \boldsymbol{\phi}_k|^2 - 1}$ 。

1.2 下行链路数据传输与系统性能分析

在下行数据传输阶段，AP 使用信道估计来正确设计预编码器。通过全导频迫零（full-pilot zero-force, fpZF）预编码器^[16]对发射的数据信号进行预编码，可以有效地消除用户间的干扰。第 m 个 AP 向所有用户发送的数据信号为：

$$\mathbf{x}_m = \sum_{k=1}^K \sqrt{\rho_t} \boldsymbol{\Psi}_{mk} q_k \quad (5)$$

其中， q_k 为第 k 个用户发送的数据信号，满足 $E\{|q_k|^2\} = 1$ 。此外， $\boldsymbol{\Psi}_{mk} \in \mathbf{C}^{N_T \times 1}$ 为 fpZF 预编码向量，可以表示为：

$$\boldsymbol{\Psi}_{mk} = \frac{\bar{\mathbf{H}}_m (\bar{\mathbf{H}}_m^H \bar{\mathbf{H}}_m)^{-1} \mathbf{e}_k}{\sqrt{E\left\{\left\|\bar{\mathbf{H}}_m (\bar{\mathbf{H}}_m^H \bar{\mathbf{H}}_m)^{-1} \mathbf{e}_k\right\|^2\right\}}} \quad (6)$$

其中， $\bar{\mathbf{H}}_m \in \mathbf{C}^{N_T \times \tau} = c_{mk}^{-1} \hat{\mathbf{h}}_{mk} \mathbf{e}_k$ 为全秩矩阵， $\mathbf{e}_k$ 为 $\mathbf{I}_\tau$ 的第 k 列元素。第 k 个用户处的接收信号可以建模为：



$$\begin{aligned}
y_k = & \sum_{m=1}^M \sqrt{\rho_t} \mathbf{h}_{mk}^H \boldsymbol{\Psi}_{mk} q_k + \sum_{m=1}^M \sum_{i \neq k}^K \sqrt{\rho_t} \mathbf{h}_{mk}^H \boldsymbol{\Psi}_{mi} q_i + w_k = \\
& \sum_{m=1}^M \sqrt{\rho_t} E \left\{ \mathbf{h}_{mk}^H \boldsymbol{\Psi}_{mk} \right\} q_k + \\
& \sum_{m=1}^M \left(\sqrt{\rho_t} \mathbf{h}_{mk}^H \boldsymbol{\Psi}_{mk} - \sqrt{\rho_t} E \left\{ \mathbf{h}_{mk}^H \boldsymbol{\Psi}_{mk} \right\} \right) q_k + \\
& \sum_{m=1}^M \sum_{i \neq k}^K \sqrt{\rho_t} \mathbf{h}_{mk}^H \boldsymbol{\Psi}_{mi} q_i + w_k
\end{aligned} \quad (7)$$

$$\text{SINR}_k = \frac{\rho_t \sum_{m=1}^M \left| E \left\{ \mathbf{h}_{mk}^H \boldsymbol{\Psi}_{mk} \right\} \right|^2}{\rho_t \sum_{m=1}^M \left(E \left\{ \left| \mathbf{h}_{mk}^H \boldsymbol{\Psi}_{mk} \right|^2 \right\} - \left| E \left\{ \mathbf{h}_{mk}^H \boldsymbol{\Psi}_{mk} \right\} \right|^2 \right) + \rho_t \sum_{m=1}^M \sum_{i \neq k}^K E \left\{ \left| \mathbf{h}_{mk}^H \boldsymbol{\Psi}_{mi} \right|^2 \right\} + 1} \quad (9)$$

为了进一步计算 SINR_k ，首先求出 $\hat{\mathbf{h}}_{mk} \boldsymbol{\Psi}_{mk}$ 的简便形式。采用 fpZF 预编码可以有效抑制所有用户之间的干扰：

$$\hat{\mathbf{h}}_{mk} \boldsymbol{\Psi}_{mi} = \frac{c_{mk}}{c_{mi}} \mathbf{e}_k^H \mathbf{e}_i \sqrt{g_{mk} (N_T - \tau)} = \begin{cases} 0, & i \neq k \\ \sqrt{g_{mk} (N_T - \tau)}, & i = k \end{cases} \quad (10)$$

将式 (10) 代入式 (9)，计算期望值，可以得到用户 k 的下行可达速率下界的封闭形式，如式 (11) 所示。

$$\begin{aligned}
R_k = & \text{lb}(1 + \text{SINR}_k) \\
= & \text{lb} \left(1 + \frac{(N_T - \tau) \left(\sum_{m=1}^M \sqrt{\rho_t g_{mk}} \right)^2}{\sum_{m=1}^M \sum_{i=1}^K \rho_t (\beta_{mi} - g_{mi}) + 1} \right) \quad (11)
\end{aligned}$$

2 基于无线传播环境的 AP 部署策略

2.1 概率路径损耗模型

在无蜂窝大规模 MIMO 系统的 AP 部署中，其周围的散射物和环境遮挡的差异，会导致 AP 至用户的无线信道经历不同的路径损耗和阴影衰落。阴影衰落是由于用户移动时环境发生变化产

生的影响，在对数分布上呈缓慢变化趋势，而路径损耗依据环境差异会有很大差别，是影响系统性能的主要原因，因此在进行 AP 部署时主要考虑路径损耗的影响。将式 (1) 中的大尺度衰落系数表示如下。

$$R_k = \text{lb}(1 + \text{SINR}_k) \quad (8)$$

其中， SINR_k 为第 k 个用户的信干噪比，可以写为：

生的影响，在对数分布上呈缓慢变化趋势，而路径损耗依据环境差异会有很大差别，是影响系统性能的主要原因，因此在进行 AP 部署时主要考虑路径损耗的影响。将式 (1) 中的大尺度衰落系数表示如下。

$$\beta_{mk} = \text{PL}_{mk} \cdot 10^{\frac{\sigma_{sh} \ell_{mk}}{10}} \quad (12)$$

其中， PL_{mk} 表示第 m 个 AP 和第 k 个用户之间的路径损耗， $10^{\frac{\sigma_{sh} \ell_{mk}}{10}}$ 表示标准偏差为 σ_{sh} 的阴影衰落，且 $\ell_{mk} \sim \mathcal{CN}(0, 1)$ 。

在实际传播环境中，AP 和用户之间的信道可能是 LOS 或 NLOS，主要取决于二者之间是否被障碍物遮挡。假设在一片特定的区域内，AP 和用户之间的 LOS 链路概率取决于两者的高度以及区域内的地理分布参数，通过国际电信联盟 (International Telecommunications Union, ITU) 阻塞模型可以给出 LOS 链路概率的完整模型^[17]。LOS 链路概率如下。

$$P_{\text{LOS}} = (1 - \omega) \sqrt{\varepsilon \eta d_{mk}} \quad (13)$$

其中， $\omega \triangleq \sqrt{\frac{\pi}{2}} \frac{\varpi}{\varepsilon_m \varepsilon'_k} \left[\text{erf} \left(\frac{\varepsilon_m}{\varpi \sqrt{2}} \right) - \text{erf} \left(\frac{\varepsilon'_k}{\varpi \sqrt{2}} \right) \right]$,

$\text{erf}(z) \triangleq \frac{1}{\sqrt{\pi}} \int_{-z}^z e^{-t^2} dt$, ϖ 为障碍物的平均高度, ξ 为区域内障碍物的面积占比, η 为单位面积的平均阻塞数, d_{mk} 为第 m 个 AP 和第 k 个用户之间的水平距离, ε_m 和 ε'_k 分别表示 AP 和 UE 的平均高度。

传统的远距离传播路径损耗模型在给定的测量范围外容易产生路径损耗估计偏差, 将 LOS 和 NLOS 模型结合在一起以改进路径损耗估计, 分别使用 3GPP 协议中定义的城市微小区

(urban micro-cellular, UMi) 场景下的 LOS 和 NLOS 模型将 LOS 与 NLOS 传播进行结合, 并使用式 (13) 表示的 LOS 链路概率来加权 LOS 和 NLOS 路径损耗。混合概率路径损耗模型表示为以下形式:

$$\text{PL}_{\text{prob}}[\text{dB}] = P_{\text{LOS}} \times \text{PL}_{\text{LOS}} + (1 - P_{\text{LOS}}) \times \text{PL}_{\text{NLOS}} \quad (14)$$

其中, PL_{LOS} 为 LOS 路径下的路径损耗, PL_{NLOS} 为 NLOS 路径下的路径损耗, 二者采用 3GPP 协议中对 UMi 的定义, 其经验模型如下。

$$\text{PL}_{\text{LOS}} = \begin{cases} 32.4 + 21 \lg d_{3D} + 20 \lg f_c, & 10 \text{ m} \leq d_{2D} \leq d'_{\text{BP}} \\ 32.4 + 40 \lg d_{3D} + 20 \lg f_c - 9.5 \lg(d'_{\text{BP}})^2 + (\varepsilon_m - \varepsilon'_k)^2, & d'_{\text{BP}} \leq d_{2D} \leq 5 \text{ km} \end{cases} \quad (15)$$

$$\begin{aligned} \text{PL}_{\text{NLOS}} &= \max(\text{PL}_{\text{LOS}}, \text{PL}'_{\text{NLOS}}), \quad 10 \text{ m} \leq d_{2D} \leq 5 \text{ km} \\ \text{PL}'_{\text{NLOS}} &= 13.54 + 39.08 \lg d_{3D} + 20 \lg f_c - 0.6(\varepsilon'_k - 1.5) \end{aligned} \quad (16)$$

其中, f_c 为载波频率, d_{3D} 为用户到 AP 的直线距离, d'_{BP} 为该模型下设置的断点距离。

式 (14) 表明, AP 与用户之间的 LOS 链路概率越大, AP 与用户之间的路径损耗越小, 反之则越大, 因此在部署 AP 时应尽量将 AP 部署在与用户之间 LOS 链路概率更大的位置, 从而提升与用户之间的通信性能。

2.2 问题描述

在满足每个用户发射功率约束要求的情况下, 为了给覆盖区域内的用户提供均匀一致的覆盖, 本文提出了以最大化平均吞吐量为标准优化 AP 部署的位置。为了便于问题的求解, 将连续的二维空间划分为 U 个大小相同的覆盖单元, 每个单元的大小由覆盖范围和 AP 个数决定, 并假设 AP 只能放置在预设的覆盖单元内, $\mathcal{N} = \{1, 2, \dots, U\}$ 表示所有覆盖单元的集合。将 AP 部署决策向量表示为 $\mathbf{L} = \{l_1, l_2, \dots, l_U\}^T$, 其中 l_u 为二元决策向量, 表示给定 AP 部署位置的部署情况, 如式 (17) 所示。

$$l_u = \begin{cases} 1, & \text{覆盖单元 } u \text{ 内部署 AP} \\ 0, & \text{覆盖单元 } u \text{ 内未部署 AP} \end{cases} \quad (17)$$

因此将连续空间栅格化后, 式 (11) 可以改写为与 AP 位置相关的向量乘积的形式:

$$R_k(\mathbf{L}) = \text{lb} \left(1 + \frac{(N_T - \tau) \rho_t \left(\sum_{u=1}^U l_u \sqrt{g_{uk}} \right)^2}{\rho_t \sum_{u=1}^U \sum_{i=1}^K l_u (\beta_{ui} - g_{ui}) + 1} \right) \quad (18)$$

以最大化平均吞吐量为目标的 AP 部署问题可以表示为:

$$\begin{aligned} \max_{\mathbf{L}} \quad & \frac{1}{K} \sum_{k=1}^K \text{lb} \left(1 + \frac{(N_T - \tau) \rho_t \left(\sum_{u=1}^U l_u \sqrt{g_{uk}} \right)^2}{\rho_t \sum_{u=1}^U \sum_{i=1}^K l_u (\beta_{ui} - g_{ui}) + 1} \right) \\ \text{s.t.} \quad & \text{C1: } R_k \geq R_{\min} \\ & \text{C2: } l_u \in \{0, 1\}^{\mathcal{N}}, \quad u = 1, \dots, U \end{aligned} \quad (19)$$

其中, C1 表示单用户速率需要满足最低用户速率阈值要求, C2 表示 AP 部署约束, 即 AP 部署限制在给定的区域内。

2.3 问题求解

由于目标函数为非凸的, 且存在非凸约束 C2, 式 (19) 属于非凸的 0-1 整数规划问题, 采用传



统优化算法难以求解,因此将其转化为马尔可夫博弈过程,进一步利用 MADDPG 算法对其进行求解。

2.3.1 MADDPG 算法

在传统的单智能体深度强化学习过程中,智能体通过与环境进行交互来改进其策略,进而从环境中得到最优策略,但是在面对更加复杂多变的环境时,单个智能体很难在学习过程中找到最优策略。为了保证学习过程的收敛性,提出了 MADDPG 算法^[18]作为深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法的拓展。

DDPG 算法采用行动者-评论家(actor-critic, AC)网络架构,行动者网络用于表示确定性策略梯度,评论家网络用于近似 Q 值函数,并且二者都包含两个网络结构相同的子网络:在线网络和目标网络。给定当前状态 s^t 、动作 a^t 和确定性策略 μ ,行动者网络和行动者目标网络分别表示为 $\mu(s^t|\nu)$ 和 $\mu'(s^t|\nu')$,评论家网络和目标网络分别表示为 $Q(s^t, a^t|\theta)$ 和 $Q'(s^t, a^t|\theta')$,其中 ν 、 ν' 、 θ 、 θ' 为这 4 个神经网络对应的参数。

MADDPG 算法通常以集中训练和分布式执行的方式工作,在训练过程中使用全局信息,而在测试时并不使用该信息。假设 M 个智能体的连续策略集和相应策略的参数集分别为 $\mu = \{\mu_1, \mu_2, \dots, \mu_M\}$ 和 $\nu = \{\nu_1, \nu_2, \dots, \nu_M\}$ 。智能体以期望奖励最大化为目标函数,根据局部观测结果进行独立决策,即 $a_i^n = \mu_{\nu_i}(o_i^n)$ 。通过目标函数相对于 ν_i 的梯度对智能体 i 的连续策略 μ_i 进行优化:

$$\nabla_{\nu_i} \mathcal{J}(\mu_i) \approx \frac{1}{N} \sum_n \nabla_{\nu_i} \mu_i(o_i^n) \nabla_{a_i} Q_i^{\theta_i}(s^n, a_1^n, \dots, a_M^n) \big|_{a_i^n = \mu_i(o_i^n)} \quad (20)$$

其中, $Q_i^{\theta_i}(s^n, a_1^n, \dots, a_M^n)$ 为集中式动作值函数, $s^t = (o_1^n, o_2^n, \dots, o_M^n)$ 包含所有 M 个智能体的观测值,并且 $a^n = (a_1^n, a_2^n, \dots, a_M^n)$ 、 r^n 和 s^{n+1} 分别代表

所有智能体的动作、全局奖励和下一时刻状态。 T 为批处理样本的大小。通过经验回放机制,将状态转移四元组 (s^n, a^n, r^n, s^{n+1}) 存放在经验回放池中,在训练时从经验回放池中随机抽取样本用于网络参数的更新,该机制可以消除训练样本之间的时间相关性。通过该方式,评论家网络可以获得所有智能体的联合状态和联合动作。

然后,第 i 个智能体的集中式动作值函数可以根据以下损失函数进行更新。

$$\mathcal{L}(\theta_i) = \frac{1}{N} \sum_n (y^n - Q_i^{\theta_i}(s^n, a_1^n, \dots, a_M^n))^2 \quad (21)$$

其中,目标 Q 值函数 y^n 基于马尔可夫决策过程(Markov decision process, MDP)可以计算为:

$$y^n = r^n + \gamma Q_i^{\theta_i'}(s^{n+1}, a_1^{n+1}, \dots, a_M^{n+1}) \big|_{a_i^{n+1} = \mu_i(o_i^{n+1})} \quad (22)$$

其中, γ 为折扣因子,用来衡量当前奖励相对于未来奖励的重要性。

目标行动者网络参数和目标评论家网络参数采用软更新的方式,软更新可以使得目标网络的参数缓慢更新,进而使得利用目标网络计算的目标值更加平稳,保证了评论家网络学习过程中的稳定性。用 δ 表示目标网络更新过程中的学习率,则目标网络的参数更新过程如下。

$$\begin{aligned} \theta_i' &\leftarrow \delta \theta_i + (1 - \delta) \theta_i' \\ \nu_i' &\leftarrow \delta \nu_i + (1 - \delta) \nu_i' \end{aligned} \quad (23)$$

2.3.2 基于 MADDPG 算法的 AP 部署优化策略

基于上述 MADDPG 算法网络结构及参数更新过程,本文提出了基于 MADDPG 算法的 AP 部署优化策略,并通过定义以下模型指定多智能体的随机博弈策略。

(1) 状态:在时隙 t 时,第 i 个智能体从网络环境中观察其本地信息,即 AP 的位置 $A_p(t)$ 、用户的位置 $U_p(t)$,以及 AP 的部署范围 $\mathcal{N}$ 。因此,可以将智能体 i 的状态定义为 $\mathbf{o}_i^t = \{A_p(t), U_p(t), \mathcal{N}\}$,则所有智能体在时隙 t 的

联合状态为 $\mathbf{s}^t = \{\mathbf{o}_1^t, \mathbf{o}_2^t, \dots, \mathbf{o}_M^t\}$ 。

(2) 动作：基于系统的当前状态和环境，在时隙 t 时智能体 i 的动作空间可以写为 $\mathbf{a}_i^t = l_u(t)$ ，表示当前的 AP 部署策略，因此 M 个智能体在时隙 t 的联合动作为 $\mathbf{a}^t = \{a_1^t, a_2^t, \dots, a_M^t\}$ 。AP 需要根据系统当前的状态和环境选择要执行的动作。

(3) 奖励：在 MADDPG 算法中，多个智能体相互合作以最大化系统效用。本文针对 AP 的部署位置进行优化，以达到最大化系统平均吞吐量的目的，因此可以根据与式 (19) 中相同的效用函数来定义所有智能体的奖励，奖励函数构建如下。

$$r^t = r(\mathbf{s}^t, \mathbf{a}^t) = \frac{1}{K} \sum_{k=1}^K R_k(t) - c^t \quad (24)$$

其中， c^t 表示不满足约束条件时的惩罚项，如果 AP 未满足约束条件，则累积奖励得到负回报作为惩罚，通过获取累积奖励得到最终的优化目标。

在对上述的状态、动作、奖励进行定义之后，

给出了基于 MADDPG 算法的 AP 部署优化策略，如算法 1 所示，基于 MADDPG 算法的 AP 部署优化策略框架如图 2 所示。第 i 个智能体从环境中观察到当前状态 $\mathbf{o}_i^t$ ，并将其输入行动者网络中，得到 AP 部署的初始策略。接着每个智能体分别执行当前动作，得到奖励值 r^t ，并更新状态环境为 $\mathbf{s}^{t+1}$ 。然后智能体将四元组 $(\mathbf{s}^t, \mathbf{a}^t, r^t, \mathbf{s}^{t+1})$ 存储在经验回放池中，重复该过程直到经验回放池满之后，从经验回放池中随机抽取 F 个样本用以训练行动者网络和集中式评论家网络。为了提高训练速度和搜索效率，采用奥恩斯坦-乌伦贝格 (Ornstein-Uhlenbeck, OU) 过程^[19]加入与时间相关的随机噪声 Φ 。

算法 1 基于 MADDPG 算法的 AP 部署优化

初始化： 行动者网络 $\mu(\mathbf{s}^t|\nu)$ 和评论家网络

$Q(\mathbf{s}^t, \mathbf{a}^t|\theta)$ 的参数 ν 和 θ ，折扣因子 γ ，经验回放池 $\mathbf{B}$ ，最小批处理大小 F ，随机噪声 Φ ，最大回合数 EP，最大训练步长 T

for episode = 1 : EP do

 每个 AP 根据当前系统的状态环境接收初

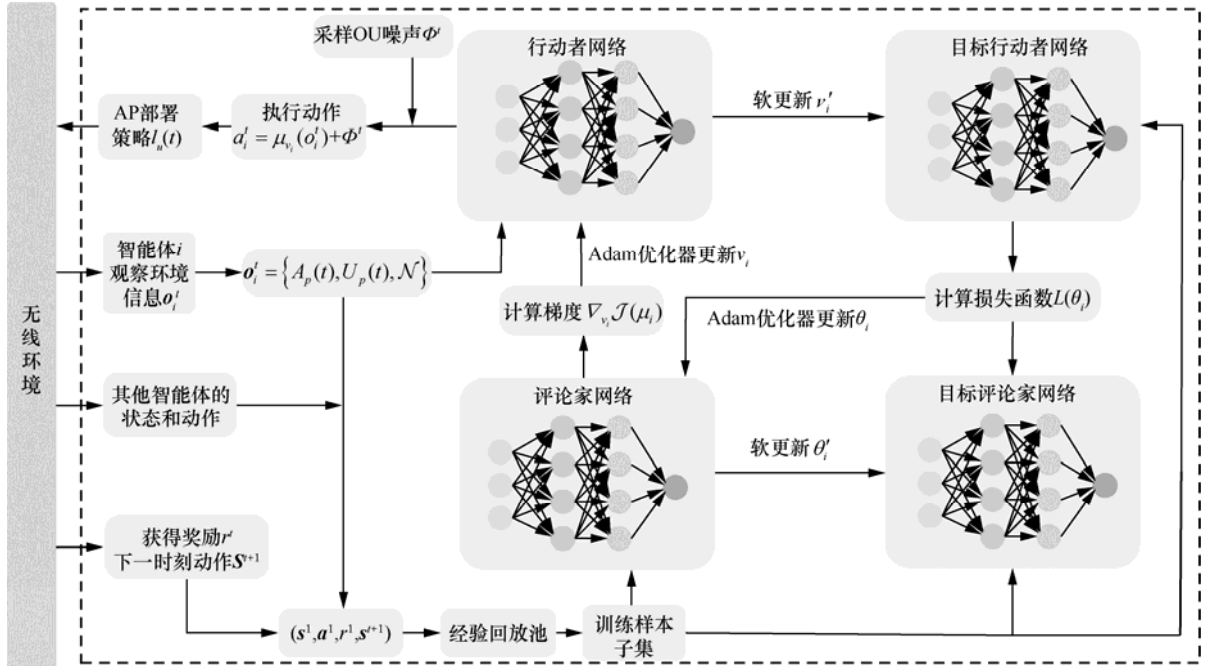


图 2 基于 MADDPG 算法的 AP 部署优化策略框架



始状态 $\mathbf{o}_i^t$ 以及全局状态 $\mathbf{s}^t$;

for $t = 1 : T$ do

对于每个智能体, 根据噪声和当前策略执行动作 $a_i^t = \mu_{v_i}(\mathbf{o}_i^t) + \Phi^t$;

执行动作 $\mathbf{a}^t = \{a_1^t, a_2^t, \dots, a_M^t\}$, 得到全局奖励 r^t , 下一时刻状态 $\mathbf{s}^{t+1}$;

将 $(\mathbf{s}^t, \mathbf{a}^t, r^t, \mathbf{s}^{t+1})$ 存储到经验回放池 $\mathbf{B}$ 中;

令 $\mathbf{s}^t \leftarrow \mathbf{s}^{t+1}$;

for $i = 1 : M$ do

从 $\mathbf{B}$ 中随机抽取 F 个样本作为行动者和评论家网络的训练数据;

根据式 (22) 得到 y^n ;

分别通过式 (20) 和式 (21) 更新行动者网络和评论家网络;

end for

通过式 (23) 更新目标网络参数;

end for

end for

3 仿真结果与分析

3.1 仿真参数

本文考虑一个大小为 $1 \text{ km} \times 1 \text{ km}$ 的区域, 其中有 M 个 AP 和 K 个用户。将该区域划分为大小相同的覆盖单元, 并假设 AP 只能放置在预先设置好的覆盖单元内。为了模拟更真实的用户分布情况, 本文采用高斯混合模型生成用户位置, 如式 (25) 所示。

$$f(\mathbf{n}) = \sum_{i=1}^3 n_i \mathcal{N}(\mathbf{n} | \mu_i, \sigma_i I) \quad (25)$$

其中, 平均位置 $\mu_1 = [0.45, 0.55]^T \text{ km}$ 、 $\mu_2 = [0.85, 0.55]^T \text{ km}$ 、 $\mu_3 = [0.55, 0.25]^T \text{ km}$, 权重 $n_i = \frac{1}{3}$, 方差 $\sigma_i = 200$, $i = 1, 2, 3$ 。

在本节的仿真验证部分, 将所提算法与其他算法的能量效率进行了对比。在无蜂窝大规模

MIMO 系统中, 能量效率被定义为系统总吞吐量与总功耗之比:

$$EE = \frac{B \sum_{k=1}^K R_k}{\rho_t N_0 \sum_{m=1}^M \frac{1}{\zeta_m} \left(\sum_{k=1}^K \mathcal{G}_{mk} \right) + \sum_{m=1}^M (P_{tc,m} + P_{0,m}) + B \left(\sum_{m=1}^M P_{bt,m} \right) \sum_{k=1}^K R_k} \quad (26)$$

其中, B 为系统带宽, N_0 为噪声功率, $0 < \zeta_m < 1$ 为功率放大器效率, $P_{tc,m}$ 是在第 m 个 AP 的电路运行功耗, $P_{0,m}$ 是固定回程功耗, $P_{bt,m}$ 是与第 m 个 AP 到 CPU 处的回程功耗, 部分仿真参数设置见表 1。

表 1 部分仿真参数设置

参数名称	变量名	参数值
AP 天线高度/m	ε_m	10
UE 天线高度/m	ε'_k	1.5
障碍物平均高度/m	ϖ	20
区域内障碍物的面积占比	ξ	0.5
单位面积障碍物的平均数量/ km^2	η	300
阴影衰落标准偏差/dB	σ_{sh}	8
电路运行功耗/W	$P_{tc,m}$	0.2
固定回程功耗/W	$P_{0,m}$	0.825
AP 到 CPU 回程功耗/(W·Gbit $^{-1}$ ·s)	$P_{bt,m}$	0.25
最大回合数	EP	200
最大训练步长	T	1 000
折扣因子	γ	0.95
经验回放池大小	B	5 000

3.2 仿真结果

本文将所提出的 AP 部署方案与随机 AP 部署方案、基于 Lloyd 的 AP 部署方案和文献[14]中的基于最大化和速率的 AP 部署方案进行对比。

随机部署与所提 AP 部署方案路径损耗对比如图 3 所示, 其中深色方形区域表示障碍物, 且不考虑障碍物内部的信号传播。图 3 (a) 显

示了 AP 随机部署时的路径损耗分布, 该方案导致覆盖区域内不同位置的路径损耗出现明显差异, 其原因是被建筑物遮挡的位置为 NLOS 传播, 其路径损耗要高于 LOS 传播。图 3 (b) 为所提 AP 部署方案的路径损耗分布, 由于考虑环境中障碍物的影响, 所提方案的路径损耗分布未出现明显差异, 可以均匀覆盖区域内的用户。

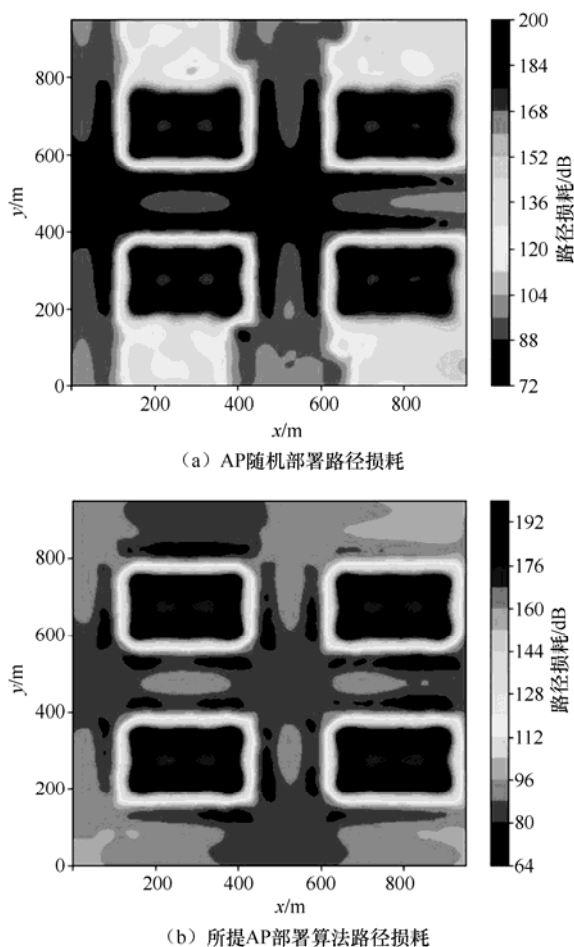


图3 随机部署与所提 AP 部署方案路径损耗对比

单用户可达速率累积分布 ($M=200, K=20$) 如图 4 所示。由图 4 可以看出, 所提算法的单用户可达速率远高于随机部署, 并且优于 Lloyd 算法和最大化和速率算法, 说明所提算法在满足均匀一致覆盖要求的同时可以为用户提供更高的服务质量。

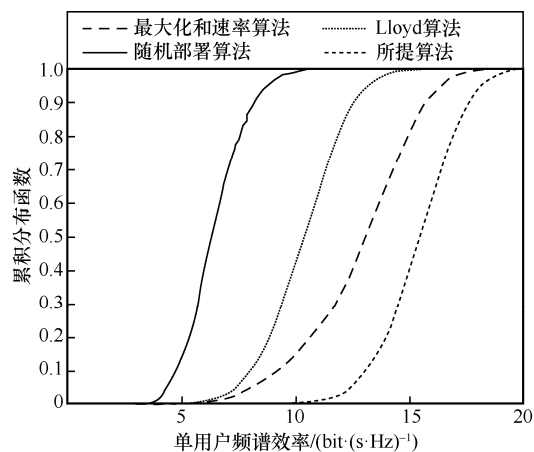


图4 单用户可达速率累积分布 ($M=200, K=20$)

总频谱效率随用户数量变化曲线 ($M=100$) 如图 5 所示, 可以观察到, 总频谱效率随着用户数量的增加呈上升趋势。此外, 所提出的算法比随机部署高约 30%, 相较于最大化和速率算法也有 10%左右的性能提升, 因此所提算法在总体性能上有更好的表现。

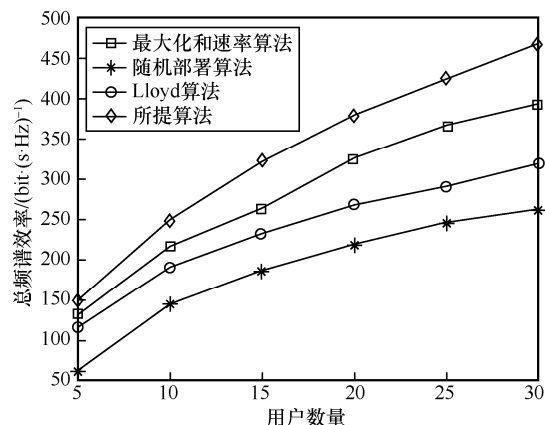
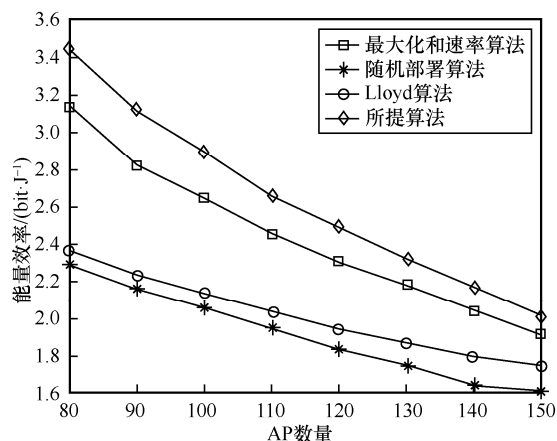


图5 总频谱效率随用户数量变化曲线 ($M=100$)

能量效率随 AP 数量变化曲线 ($K=30$) 如图 6 所示, 显示了用户数量 $K=30$ 的条件下能量效率与 AP 数量之间的关系。从图 6 中可以看出, 随着 AP 数量的增加, 所有算法的能量效率都呈下降趋势。虽然所提算法相较于其他算法获得了相对较高的能量效率, 但是当无蜂窝大规模 MIMO 网络以高能效为目标时, 更需要关注 AP 的数量要适当地适应于服务的用户数量, 进而平衡用户需求与系统能效之间的关系。

图6 能量效率随 AP 数量变化曲线 ($K=30$)

4 结束语

本文针对无蜂窝大规模 MIMO 系统提出了一种基于无线传播环境的 AP 部署策略,通过构建的混合路径损耗模型反映了实际环境中障碍物对路径损耗的影响,并以最大化平均吞吐量为目标,将问题转化为马尔可夫博弈过程,并通过基于 MADDGP 算法的 AP 部署优化策略得到了最优 AP 部署方案。仿真结果表明,所提 AP 部署方案相较于其他方案能够实现区域内更加均匀一致的覆盖,并在总体性能上也有明显提升。该项工作能很好地解决在实际场景中部署 AP 时造成的非均匀覆盖问题,为用户提供更加良好的服务质量。

参考文献:

- [1] CHEN S F, ZHANG J Y, JIN Y, et al. Wireless powered IoE for 6G: Massive access meets scalable cell-free massive MIMO[J]. China Communications, 2020, 17(12): 92-109.
- [2] NGO H Q, ASHIKHMIN A, YANG H, et al. Cell-free massive MIMO versus small cells[J]. IEEE Transactions on Wireless Communications, 2017, 16(3): 1834-1850.
- [3] NAYEBI E, ASHIKHMIN A, MARZETTA T L, et al. Cell-Free Massive MIMO systems[C]//Proceedings of the 49th Asilomar Conference on Signals, Systems and Computers. Piscataway: IEEE Press, 2015: 695-699.
- [4] BJÖRNSON E, SANGUINETTI L. Making cell-free massive MIMO competitive with MMSE processing and centralized im-

plementation[J]. IEEE Transactions on Wireless Communications, 2020, 19(1): 77-90.

- [5] ZHENG J K, ZHANG J Y, ZHANG L M, et al. Efficient receiver design for uplink cell-free massive MIMO with hardware impairments[J]. IEEE Transactions on Vehicular Technology, 2020, 69(4): 4537-4541.
- [6] 章嘉懿. 去蜂窝大规模 MIMO 系统研究进展与发展趋势[J]. 重庆邮电大学学报(自然科学版), 2019, 31(3): 285-292.
ZHANG J Y. Overview of cell-free massive MIMO system[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2019, 31(3): 285-292.
- [7] REN Z H, DOUFEXI A, BEACH M A. Scalable pilot assignment for user-centric cell-free massive MIMO networks[C]//Proceedings of IEEE International Conference on Communications. Piscataway: IEEE Press, 2022: 2555-2560.
- [8] CHOI T, KANNO I, ITO M, et al. A realistic path loss model for cell-free massive MIMO in urban environments[C]//Proceedings of IEEE Global Communications Conference. Piscataway: IEEE Press, 2022: 2468-2473.
- [9] LIU H, ZHANG J Y, JIN S, et al. Graph coloring based pilot assignment for cell-free massive MIMO systems[J]. IEEE Transactions on Vehicular Technology, 2020, 69(8): 9180-9184.
- [10] WANG Z, ZHANG J Y, AI B, et al. Uplink performance of cell-free massive MIMO with multi-antenna users over jointly-correlated Rayleigh fading channels[J]. IEEE Transactions on Wireless Communications, 2022, 21(9): 7391-7406.
- [11] PAPAFAEIROPOULOS A, KOURTESSIS P, DI RENZO M, et al. Performance analysis of cell-free massive MIMO systems: a stochastic geometry approach[J]. IEEE Transactions on Vehicular Technology, 2020, 69(4): 3523-3537.
- [12] PAPAFAEIROPOULOS A, NGO H Q, KOURTESSIS P, et al. Optimal energy efficiency in cell-free massive MIMO systems: a stochastic geometry approach[C]//Proceedings of the 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications. Piscataway: IEEE Press, 2020: 1-7.
- [13] SHEN K, DE PESSEMIER T, MARTENS L, et al. Genetic algorithm combined with ray tracer for optimizing cell-free mMIMO topology in a confined environment[C]//Proceedings of the 15th European Conference on Antennas and Propagation (EuCAP). Piscataway: IEEE Press, 2021: 1-5.
- [14] NAYEBI E, RAO B D. Access point location design in cell-free massive MIMO systems[C]//Proceedings of the 52nd Asilomar Conference on Signals, Systems, and Computers. Piscataway:

IEEE Press, 2018: 985-989.

- [15] JIANG J, YAN F Y, YE Y H, et al. Traffic demand-oriented cell-free massive MIMO network[J]. IEEE Wireless Communications Letters, 2023, 12(11): 1861-1865.
- [16] ZHANG J Y, ZHANG J, BJÖRNSON E, et al. Local partial zero-forcing combining for cell-free massive MIMO systems[J]. IEEE Transactions on Communications, 2021, 69(12): 8459-8473.
- [17] MUKHERJEE S, CHOPRA R. Performance analysis of cell-free massive MIMO systems in LoS/NLoS channels[J]. IEEE Transactions on Vehicular Technology, 2022, 71(6): 6410-6423.
- [18] LOWE R, WU Y, TAMAR A, et al. Multi-agent actor-critic for mixed cooperative-competitive environments[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 6379-6390.
- [19] CHERIDITO P, KAWAGUCHI H, MAEJIMA M. Fractal Ornstein-Uhlenbeck processes[J]. Electronic Journal of Probability, 2003, 8(3): 1-14.

[作者简介]



姜静（1974—），女，博士，西安邮电大学科研处处长、教授、博士生导师，主要研究方向为人工智能在无线通信中的应用、大规模 MIMO 技术等。



刘永强（1998—），男，西安邮电大学通信与信息工程学院硕士生，主要研究方向为大规模 MIMO 技术。



严冯洋（1999—），男，西安邮电大学通信与信息工程学院硕士生，主要研究方向为大规模 MIMO 技术。



陶莎（1999—），女，西安邮电大学通信与信息工程学院硕士生，主要研究方向为大规模 MIMO 技术。

Worakrin Sutthiphan（1980—），男，泰国国家电信有限公司移动网络规划经理，主要研究方向为无线网络规划、网络资源配置优化等。