

doi: 10.14132/j.cnki.1673-5439.2025.05.002

无人机携带STAR-RIS辅助的超密集网络效益优化研究

张国斌^{1,2}, 陈瀚³, 杨芷晴³, 李晓滨³

1. 广州航海学院 低空装备与智能控制学院, 广东 广州 510725
2. 东莞理工学院 电信工程与智能化学院, 广东 东莞 523808
3. 深圳大学 电子与信息工程学院, 广东 深圳 518060

摘要: 针对5G时代超密集网络部署成本与能效的挑战, 提出一种结合无人机以及同时透射和反射信号的可重构智能表面(STAR-RIS)的网络规划设计方法。该方法通过构建包含宏基站、用户终端和携带STAR-RIS模块的无人机超密集网络模型, 利用软演员评论家(Soft Actor-Critic, SAC)算法优化网络配置, 以实现在满足通信速率需求的前提下最大化系统效益。首先构建了速率模型和功率模型, 并基于模型设计了能量和成本效益目标函数。然后通过SAC算法求解优化问题, 得到在满足功率、速率和成本限制条件下的最优网络配置。仿真结果表明, 该方法能够有效提升网络能量和成本效益, 减少系统能耗和部署成本。

关键词: 超密集网络; 无人机; STAR-RIS; SAC算法; 效益优化

中图分类号: TP928

文献标志码: A

文章编号: 1673-5439(2025)05-0014-09

Efficiency optimization in UAV-STAR-RIS assisted UDNs

ZHANG Guobin^{1,2}, CHEN Han³, YANG Zhiqing³, LI Xiaobin³

1. School of Low-Altitude Equipment and Intelligent Control, Guangzhou Maritime University, Guangzhou 510725, China
2. School of Electrical Engineering & Intelligentization, Dongguan University of Technology, Dongguan 523808, China
3. College of Electronics and Information Engineering, Shenzhen University, Shenzhen 518060, China

Abstract: In the 5G era, ultra-dense networks (UDNs) face significant challenges in cost and energy efficiency optimization. This study proposes a network planning and design approach that combines unmanned aerial vehicles (UAVs) and simultaneously transmitting and reflecting reconfigurable intelligent surfaces (STAR-RISs). An ultra-dense network model is constructed, incorporating macro base station, user terminals, and UAV-carried STAR-RIS modules. The soft actor-critic (SAC) algorithm is employed to optimize network parameters with the goal of maximizing system efficiency while meeting communication rate requirement. Rate and power models are developed and an objective function of energy and cost efficiency is formulated. The SAC algorithm is adopted to solve the optimization problem and determine the optimal network configuration under constraints of power, rate, and cost. Simulation results indicate that the proposed approach effectively enhances network energy and cost efficiency, thereby reducing energy consumption and deployment cost of the system.

收稿日期: 2025-04-29; 修回日期: 2025-08-26 本刊网址: <http://nyzr.njupt.edu.cn>

基金项目: 广东省基础与应用基础研究基金粤莞联合基金-地区培育项目(2023A1515140003)资助项目

作者简介: 张国斌, 女, 博士, 副教授, guobinzh@163.com

引用本文: 张国斌, 陈瀚, 杨芷晴, 等. 无人机携带STAR-RIS辅助的超密集网络效益优化研究[J]. 南京邮电大学学报(自然科学版), 2025, 45(5): 14-22.

Citation: ZHANG Guobin, CHEN Han, YANG Zhiqing, et al. Efficiency optimization in UAV-STAR-RIS assisted UDNs [J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2025, 45(5): 14-22.

Keywords: ultra-dense network (UDN); unmanned aerial vehicle (UAV); simultaneously transmitting and reflecting reconfigurable intelligent surface (STAR-RIS); soft actor-critic (SAC) algorithm; efficiency optimization

5G技术以其高速率、低时延和高连接密度的特点,为万物互联提供了技术支撑。5G网络的构建不仅需要宏基站实现大范围覆盖,还需要低功率基站进行热点区域或室内覆盖,以应对数据传输需求^[1]。超密集网络通过增加低功率节点,提高网络容量和覆盖范围,被认为是5G的关键技术之一^[2]。然而,超密集网络的实施面临部署成本和能效挑战,低功率基站的密集部署增加了网络复杂性,可能导致信号干扰,影响通信质量。此外,随着用户数量的增加,网络能耗问题也日益突出。因此,如何在保证通信质量的同时降低能耗和部署成本,成为超密集网络开发中亟待解决的问题。

为了解决这些问题,研究者们提出了可重构智能表面(Reconfigurable Intelligent Surface, RIS)技术^[3]。RIS通过改变入射信号反射角度实现信号的智能调控,从而优化无线通信链路^[4-5]。特别是同时透射和反射信号的STAR-RIS^[6-7],能够同时服务于位于STAR-RIS前面和后面的用户,为无线通信提供了新的解决方案。Zhong等^[8]提出一种STAR-RIS方案,考虑了耦合相移模型,通过定义一个联合无源和有源波束成形优化问题,在耦合相移限制和最小数据速率约束下最小化长期广播的功耗。Wang等^[9]提出一种改进的STAR-RIS,使用复杂的控制器来改变STAR-RIS组件的电磁特性,入射信号可以分为反射信号和透射信号,从而实现360°覆盖,仿真结果表明,STAR-RIS比RIS性能更优。

近年来有研究将STAR-RIS部署在无人机上,动态调整网络覆盖区域,提高信号传输质量,尤其是在直射链路受阻的区域。Shi等^[10]提出一种无人机配备STAR-RIS辅助的多用户无线通信系统,设计了一种迭代方法同时优化能量分裂和模式切换协议的变量,比固定的STAR-RIS更灵活并提高了系统总速率。Zhang等^[11]联合优化STAR-RIS的反射、透射系数和无人机轨迹,在功率、轨迹耦合约束下最大化多用户和速率。Wang等^[12]分析了无人机携带STAR-RIS网络的传输中断概率和可达遍历速率,推导出速率近似上界的闭式表达式。Xiao等^[13]将STAR-RIS搭载于无人机,利用时隙调度和联合波束成形显著提升了任务卸载与信息传输的加权性能。

深度强化学习(Deep Reinforcement Learning,

DRL)方法在处理大规模变量问题上性能优越,逐渐被用于STAR-RIS辅助的无人机网络的资源管控。Yang等^[14]提出一种基于优化的分级DRL算法,用于STAR-RIS辅助的无人机数据收集系统,通过优化无人机轨迹、设备调度和STAR-RIS波束成形,显著降低了无人机能耗。Zhao等^[15]设计无人机轨迹、主动波束成形和STAR-RIS被动传输/反射波束成形联合解决方案,以实现最大化系统和速率,并通过分布式鲁棒强化学习算法提升了系统鲁棒性。Aung等^[16]采用基于近端策略优化的DRL方法,联合优化无人机轨迹、STAR-RIS幅度、相移系数及功率分配,在系统时延、无人机飞行路径、速率限制下最小化无人机和物联网设备总能耗。

在各种网络场景的研究中,STAR-RIS在提高网络性能方面发挥了重要作用,但STAR-RIS网络的资源分配问题尚未得到充分探索。目前关于空中STAR-RIS辅助通信的场景研究仍然比较局限,大部分研究只针对应用单个STAR-RIS的通信网络终端设备的性能进行研究,然而,单个STAR-RIS的应用无法实现网络所需性能,由于超密集网络场景中节点众多,网络对大容量的需求激增,因此研究多个STAR-RIS网络模型对超密集网络发展具有重要意义。此外,现有工作主要对STAR-RIS网络的系统速率和能量消耗进行优化,未考虑网络综合部署及运维成本,缺乏超密集节点场景下对系统相移耦合、功率消耗与部署成本的联合设计。随着节点规模增大,传统求解方法应用于高维混合整数非凸问题面临计算困难,因此研究高效强适应性算法求解多STAR-RIS网络资源管理问题具有关键作用。

本研究旨在超密集网络背景下,构建网络模型、定义速率与功率模型,针对STAR-RIS辅助的网络规划配置问题,在满足最小通信速率需求的前提下,考虑最大功率、STAR-RIS数目限制以及成本约束,优化用户功率分配、STAR-RIS元件的反射透射角度和部署数量,最大化系统效益,实现提升网络能量效率和降低部署成本的目标。本文采用DRL方法对效益目标函数进行求解。在训练中构建了适应环境的状态空间、动作空间和奖励函数,以表示系统状态、可行动作与性能指标,并采用SAC算法对状态空间进行探索,选择最优动作以获得最大

收益。与现有技术相比,本研究在网络参数配置的准确性与系统效率方面均表现出显著优势,为超密集网络的智能优化配置提供了新的解决方案,也为未来移动通信网络的发展提供了理论支持和技术参考。

1 系统模型

在超密集网络中引入多个STAR-RIS,对宏基站的信号进行反射和透射。一方面,由于STAR-RIS提供了两种信号传输路径,可以更有效地利用空间资源,覆盖更广泛的区域。同时,由于STAR-RIS具备同时控制反射和透射信号的能力,可以更灵活地控制信号的传播方向和强度,实现更好的波束成形和干扰管理。另一方面,由于无人机具有高度的灵活性和机动性,可以实现快速按需部署,与传统的固定基站相比,无人机的成本相对较低,具有更强的环境适应性。因此,本研究将无人机与STAR-RIS技术相结合,提高系统通信性能的同时降低部署成本。考虑下行链路,系统模型如图1所示。

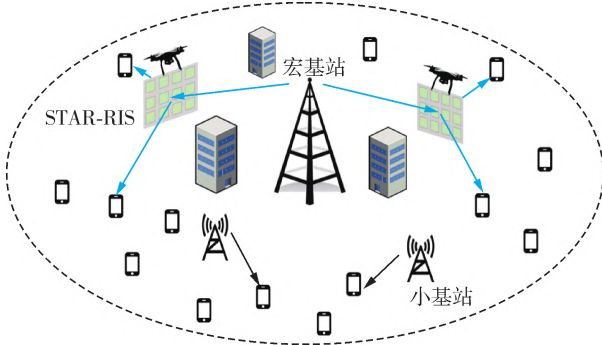


图1 系统网络模型

无人机携带STAR-RIS在固定高度飞行,用户终端随机地分布在网络覆盖区域,数目、位置已知且固定。假设系统通信为多输入单输出传输,基站配备多个天线,每个用户终端都配备单个天线。系统关键符号及其含义如表1所示。小基站集合为 $\{S_1, S_2, \dots, S_n, \dots, S_N\}$, STAR-RIS集合为 $\{R_1, R_2, \dots, R_i, \dots, R_I\}$ 。用户分为两类,被小基站服务的用户定义为小基站用户,被STAR-RIS服务的用户定义为STAR-RIS用户。定义二进制变量 $\alpha_{x,y}$,当 $\alpha_{x,y} = 1$ 表示在坐标 (x, y) 处配置一个STAR-RIS;否则表示该位置没有配置。同理定义二进制变量 $\beta_{x,y}$,当 $\beta_{x,y} = 1$ 表示在坐标 (x, y) 处配置一个小基站;否则表示该位置没有配置小基站。假定总共有 I_{\max} 个STAR-RIS和 N_{\max} 个小基站可用于配置。为了

简单起见,用户连接距离最近的STAR-RIS或者小基站以保证较好的通信质量,小基站用户只接收小基站的信号,STAR-RIS用户接收来自宏基站直射、反射和透射的信号。

表1 关键符号介绍

符号	含义
\mathcal{N}	小基站集合
\mathcal{I}	STAR-RIS集合
\mathcal{L}	小基站用户集合
\mathcal{K}	STAR-RIS用户集合
$h_{l,n}$	小基站 n 到用户 l 的信道增益
$h_{k,i}^M$	宏基站到第 i 个STAR-RIS服务的用户 k 的信道增益
h_i^L	宏基站到第 i 个STAR-RIS的信道增益
$h_{k,i}^L$	第 i 个STAR-RIS到用户 k 的信道增益
$\theta_{k,i}^R$	第 i 个STAR-RIS到用户 k 的反射角度
$\theta_{k,i}^T$	第 i 个STAR-RIS到用户 k 的透射角度
C	总成本
B	信道带宽
$P_{l,n}^S$	小基站 n 到用户 l 的发射功率
$P_{k,i}^M$	宏基站到第 i 个STAR-RIS服务的用户 k 的功率

假定第 n 个小基站记为 $S_n, \forall n \in \mathcal{N} = \{1, 2, \dots, N\}$,服务的用户 l 为 $U_{l,n}^S, \forall l \in \mathcal{L} = \{1, 2, \dots, L\}$,从小基站到用户的信道被认为是直射路径,则小基站 n 到用户 l 的信道增益可表示为 $h_{l,n}$ 。考虑通信信道由路径损耗、频率选择性衰落和阴影衰落所组成的块衰落,其中路径衰落为 h_p ,与 $(d/d_0)^\eta$ 成正比, η 为路径衰落指数, d 为小基站 n 到用户 l 的距离, d_0 为参考距离;频率选择性衰落 h_f 服从 σ_f 的瑞利分布;阴影衰落 h_s 为 $10^{-\xi/10}$ 的形式, ξ 为服从 $(0, \sigma_s^2)$ 正态分布的随机变量。

假定第 i 个STAR-RIS记为 $R_i, \forall i \in \mathcal{I} = \{1, 2, \dots, I\}$,服务的用户 k 为 $U_{k,i}^R, \forall k \in \mathcal{K} = \{1, 2, \dots, K\}$,宏基站到用户 k 的信道增益为 $h_{k,i}^M$,宏基站到第 i 个STAR-RIS的信道增益为 h_i^L ,第 i 个STAR-RIS到用户 k 的信道增益为 $h_{k,i}^L$ 。 R_i 服务的用户 $U_{k,i}^R$ 可以直接接收宏基站的直射信号,同时也能接收宏基站通过STAR-RIS反射和透射的信号。宏基站到第 i 个STAR-RIS和第 i 个STAR-RIS到用户 k 的路径损耗分别可表示为^[17]

$$h_i^L = 20\log(d_{i,M}) + 20\log(f) + 20\log\left(\frac{4\pi}{c}\right) + \eta \quad (1)$$

$$h_{k,i}^L = 20\log(d_{k,i}) + 20\log(f) + 20\log\left(\frac{4\pi}{c}\right) + \eta \quad (2)$$

其中, $d_{i,M}$ 和 $d_{k,i}$ 分别表示宏基站到第 i 个STAR-RIS和第 i 个STAR-RIS到用户 k 的距离, η 表示建筑物所

引起的阴影和散射引起的附加损耗。考虑传输过程中的相位变化,并将复信道增益表示为

$$\mathbf{h}_i^l = \mathbf{h}_i^l e^{j\theta_{i,M}} \quad (3)$$

$$\mathbf{h}_{k,i}^l = \mathbf{h}_{k,i}^l e^{j\theta_{k,i}} \quad (4)$$

其中, $\theta_{i,M}$ 和 $\theta_{k,i}$ 分别为宏基站到第 i 个 STAR-RIS 和第 i 个 STAR-RIS 到用户 k 的随机相位。为了描述 STAR-RIS 的定向行为,定义反射角 $\theta_{k,i}^R$ 和透射角 $\theta_{k,i}^T$, 分别表示用户 k 在第 i 个 STAR-RIS 的入射信号与反射信号之间的夹角以及入射信号与透射信号之间的夹角。通过控制反射角和透射角,可以将信号更精准地导向前向和后向目标用户,增强接收信号强度,改善目标区域覆盖或避开干扰源,提高信号质量和通信速率。STAR-RIS 元件反射角和透射角对应的对角相移分别满足

$$\Theta_{k,i}^R = e^{j\theta_{k,i}^R} \quad (5)$$

$$\Theta_{k,i}^T = e^{j\theta_{k,i}^T} \quad (6)$$

定义二进制连接变量 $\omega_{x,y}^R[k]$ 和 $\omega_{x,y}^S[l]$, 当值为 1 时表示用户 k 或者用户 l 连接坐标为 (x, y) 处的 STAR-RIS 或小基站, 值为 0 时表示没有建立连接。 S_n 服务的用户 $U_{l,n}^S$ 在接收 S_n 发送的信号的信噪比为

$$\text{SNR}_{l,n}^S = \frac{\omega_{x,y}^S[l] P_{l,n}^S h_{l,n}^S}{\sigma^2} \quad (7)$$

其中, $P_{l,n}^S$ 表示小基站 S_n 分配给用户 l 的功率。因此小基站用户 l 的速率计算表达式为

$$R_l^S = B \sum_{n=1}^N \log(1 + \text{SNR}_{l,n}^S) \quad (8)$$

其中 B 表示带宽。

STAR-RIS 用户 $U_{k,i}^R$ 接收宏基站反射信号或透射信号的信噪比分别为

$$\text{SNR}_{k,i}^R = \frac{\omega_{x,y}^R[k] P_{k,i}^M (h_{k,i}^M + |\mathbf{h}_{k,i}^{l_2} \Theta_{k,i}^R \mathbf{h}_i^{l_1}|)}{\sigma^2} \quad (9)$$

$$\text{SNR}_{k,i}^T = \frac{\omega_{x,y}^R[k] P_{k,i}^M (h_{k,i}^M + |\mathbf{h}_{k,i}^{l_2} \Theta_{k,i}^T \mathbf{h}_i^{l_1}|)}{\sigma^2} \quad (10)$$

其中, $P_{k,i}^M$ 表示宏基站分配给第 i 个 STAR-RIS 服务的用户 k 的功率。定义 $\mu_{k,i}^R$ 和 $\mu_{k,i}^T$ 表示用户 k 的信号指示变量, $\mu_{k,i}^R = 1$ 表示用户 k 接收第 i 个 STAR-RIS 的反射信号, $\mu_{k,i}^T = 1$ 表示用户 k 接收折射信号, $\mu_{k,i}^R = 0$ 和 $\mu_{k,i}^T = 0$ 表示未接收。STAR-RIS 用户 k 的速率计算表示为

$$R_k^M = B \sum_{i=1}^I \log(1 + \mu_{k,i}^R \text{SNR}_{k,i}^R + \mu_{k,i}^T \text{SNR}_{k,i}^T) \quad (11)$$

系统的速率之和定义为所有用户的速率累加, 表示为

$$R = \sum_{l=1}^L R_l^S + \sum_{k=1}^K R_k^M \quad (12)$$

功率之和定义为所有用户的功率累加, 表示为

$$P = \sum_{l=1}^L \sum_{n=1}^N P_{l,n}^S + \sum_{k=1}^K \sum_{i=1}^I P_{k,i}^M \quad (13)$$

系统的效益定义为速率之和与功率之和、成本的比值, 表达式为

$$\eta = \frac{R}{CP} = \frac{R}{(C_1 I + C_2 N) P} = \frac{R}{\left(C_1 \sum_{x=1}^X \sum_{y=1}^Y \alpha_{x,y} + C_2 \sum_{x=1}^X \sum_{y=1}^Y \beta_{x,y} \right) P} \quad (14)$$

其中, C 表示总成本, C_1 和 C_2 分别表示一个 STAR-RIS 和小基站的硬件成本, X 和 Y 表示位置范围的最大值。

2 优化问题

本文研究目标是在满足通信速率需求、功率、STAR-RIS 数目、小基站数目等限制下, 优化用户的功率分配、STAR-RIS 元件的角度调节和配置数目, 从而最大化整个系统的效益。优化问题定义为

$$\begin{aligned} \max_{P, \theta_{k,i}^R, \theta_{k,i}^T, \alpha_{x,y}, \beta_{x,y}} \quad & \eta = \frac{R}{CP} \\ \text{s.t.} \quad & C_1: 0 \leq P \leq P_{\max} \\ & C_2: 0 \leq \theta_{k,i}^R \leq 2\pi, \forall k \in \mathcal{K}, i \in \mathcal{I} \\ & C_3: 0 \leq \theta_{k,i}^T \leq 2\pi, \forall k \in \mathcal{K}, i \in \mathcal{I} \\ & C_4: R_l^S \geq R_{\min}, \forall l \in \mathcal{L} \\ & C_5: R_k^M \geq R_{\min}, \forall k \in \mathcal{K} \\ & C_6: C \leq C_{\max} \\ & C_7: I \leq I_{\max} \\ & C_8: N \leq N_{\max} \end{aligned} \quad (15)$$

其中, 系统效益定义为能量效率和网络系统总成本的比值。硬件成本通过市场上的供求关系和小基站的规模来确定, 为了降低复杂度, 系统的总成本由网络中小基站成本、运维电力消耗和 STAR-RIS 的硬件成本构成, 不考虑后期人工维护成本等其他成本。 C_1 表示所有用户功率总和不超过最大功率, C_2 和 C_3 表示 STAR-RIS 元件的角度范围, C_4 和 C_5 表示系统的速率达到最小通信需求, C_6 表示成本限制, C_7 和 C_8 分别表示 STAR-RIS 和小基站的数目限制。

3 基于深度强化学习的求解方法

优化问题式(15)包含多个优化变量, 且变量之间存在复杂的耦合关系, 优化目标为速率与成本的非线性比值函数, 约束条件多且具备非凸性。因此, 整个问题呈现出高维、强耦合、非线性且非凸的

特征,使用传统的凸优化方法难以准确求解或容易陷入局部最优解。

与之相比,深度强化学习方法具备处理高维连续动作空间、非线性目标函数以及不确定环境下策略自适应调整的能力。SAC算法通过引入最大熵思想提升策略的探索性,并采用“软策略”来平衡探索与利用,在多目标约束和动态环境中展现出较强的鲁棒性和收敛能力。因此,本文引入基于深度强化学习的SAC算法,适应高维状态空间下网络资源动态优化的需求,从而实现系统效益的最大化。

3.1 SAC算法

系统在基站端集中执行SAC算法来实现网络配置策略。由于引入了交叉熵,SAC算法对环境的探索和利用将会达到一个平衡,因而更适合复杂、随机的环境,同时也能避免DQN算法带来的 Q 值过高估计问题。SAC算法首先建立3个网络来拟合状态值函数 $V_\psi(s_t)$ 、 Q 函数 $Q_\theta(s_t, a_t)$ 和策略函数 $\pi_\phi(a_t|s_t)$,对应的网络分别称为 V 评判者网络、 Q 评判者网络和行动者网络,对应的网络参数分别为 ψ 、 θ 和 ϕ ,如图2所示。

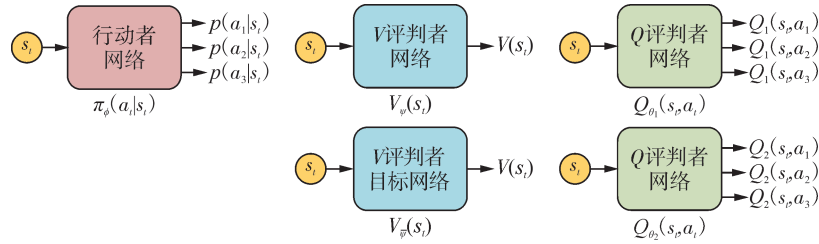


图2 SAC网络示意图

对 Q 值的过度估计会存在于行动者—评判者的方法中,因此SAC算法增加一个 Q 评判者网络。每次训练时,只会使用两个 Q 网络中值较小的一个网络,从而缓解值过高估计的问题。在优化 V 网络时,SAC算法沿用DQN中固定目标值的方法,增加一个目标 V 网络,通过每隔一段时间将 $V_\psi(s_t)$ 网络的参数拷贝到目标网络,以使学习过程更加稳定。

图3表示 V 评判者网络的更新过程,当前状态分别通过两个 Q 评判者网络后得到对应的状态动作函数值,取两者输出最小值与当前状态通过行动者网络得到的概率计算熵,结果与当前状态通过 V 评判者网络得到的 $V(s_t)$ 计算损失函数,通过均方误差最小化损失函数来优化 V 评判者网络的参数。

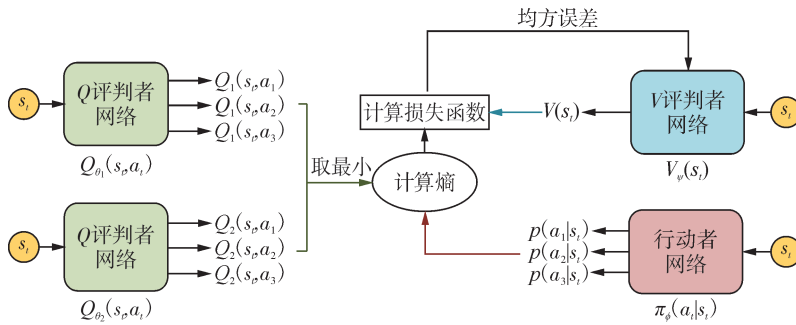


图3 V评判者网络更新流程

图4表示 Q 评判者网络的更新过程,SAC算法使用两个 Q 评判者网络,每次在两个网络中选择最小的 Q 值,两个 Q 网络同步更新。通过 V 评判者网络得到的 $V(s_t)$ 与通过 Q 评判者网络得到的 $Q(s_t, a_t)$ 计算损失函数,使用均方误差最小化该损失函数来优化 Q 评判者网络中的参数 θ 。

图5为行动者网络的更新流程,将行动者网络得到的状态下每个动作的概率与通过 Q 评判者网络得到的状态动作 Q 值进行计算损失函数操作,同样使用均方误差对行动者网络参数进行优化。

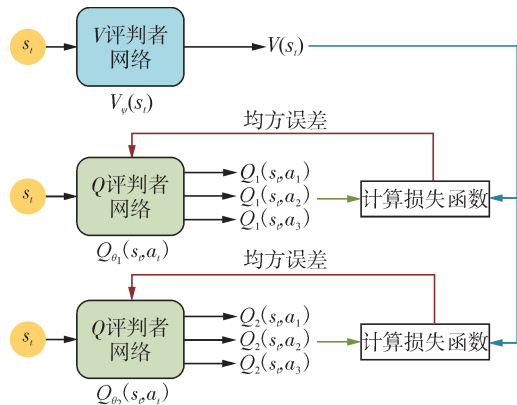


图4 Q评判者网络更新流程

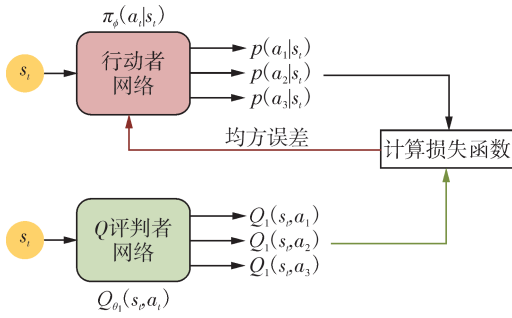


图5 行动者网络更新流程

3.2 算法求解过程

选取宏基站作为智能体实现SAC算法,该智能体包含所有用户的信息,如所有用户的状态、动作、信道状态信息等。用户分别将自己的信道状态等信息上传到宏基站,宏基站每次训练完成后,将各个用户的信息分发给用户,用户接收到后更新自己的状态,再次上传到基站,如此重复执行训练,在此过程中,宏基站承担绝大部分计算工作。

强化学习的关键元素可以表示为一个元组 $(\mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P})$,其中 \mathcal{S} 表示状态空间, \mathcal{A} 表示行动空间, \mathcal{R} 表示奖励函数, \mathcal{P} 表示指从状态 s 到状态 s' 的状态转移概率,分别设置如下:

(1) 状态空间。将小基站和STAR-RIS配置作为智能体的状态集合,状态空间表示为 $\mathcal{S} = \{s_t = (P, I, N, \vartheta) | \vartheta = (\theta_{k,i}^R, \theta_{k,i}^T)\}$ 。

(2) 动作空间。定义智能体动作空间为宏基站和小基站发送功率、STAR-RIS角度和是否放置小基站或STAR-RIS,空间集合表示为 $\mathcal{A} = \{a_t = (P_{l,n}^U, a_k^M, \alpha_{x,y}, \beta_{x,y}) | a_k^M = (P_{k,i}^M, \theta_{k,i}^R, \theta_{k,i}^T)\}$ 。

(3) 奖励函数。使用系统整体的效益作为奖励,表示为 $\mathcal{R} = \{r = r_t | r_t = \eta\}$ 。

本文采用集中训练式SAC算法,流程如算法1所示。SAC算法的复杂度主要由行动者网络和评判者网络结构决定,使用 G_a 和 G_c 分别表示行动者网络和评判者网络的层数,相应地,使用 n_g^a 和 n_g^c 表示二者在第 g 层中的神经元数量,假设在达到收敛前训练了 S 个时隙以及 Y 次迭代,算法复杂度可以表示为 $\mathcal{O}(SY(\sum_{g=0}^{G_a-1} n_g^a n_{g+1}^a + \sum_{g=0}^{G_c-1} n_g^c n_{g+1}^c))$ 。

算法1 SAC算法流程

1. 初始化状态值网络参数 ψ 、状态值目标网络 $\bar{\psi}$ 、Q网络 θ 和策略网络 ϕ ;
2. For episode = 1 to M
3. For $t = 1$ to T
4. 通过策略采样得到动作 $a_t \sim \pi_\phi(a_t | s_t)$;

5. 执行动作 a_t ,得到环境反馈的 r_t ;
6. 从环境中采样得到状态 $s_{t+1} \sim p(s_{t+1} | s_t, a_t)$;
7. 将 s_t, a_t, r_t, s_{t+1} 组成元组 (s_t, a_t, r_t, s_{t+1}) 并存入经验回放池 D 中;
8. 从经验回放池 D 中随机抽选数组 (s_t, a_t, r_t, s_{t+1}) 元组;
9. End for
10. For step = 1 to step_{\max}
11. 更新状态值函数 $V_\psi(s_t): \psi \leftarrow \psi - \lambda_V \nabla_\psi J_V(\psi)$;
12. 更新Q值函数 $Q_\theta(s_t, a_t): \theta_i \leftarrow \theta_i - \lambda_Q \nabla_{\theta_i} J_Q(\theta_i), \forall i \in \{1, 2\}$;
13. 更新策略函数 $\pi_\phi(a_t | s_t): \phi \leftarrow \phi - \lambda_\pi \nabla_\phi J_\pi(\phi)$;
14. 更新状态值函数的目标网络: $\bar{\psi} \leftarrow \tau\psi + (1 - \tau)\bar{\psi}$;
15. End for
16. End for

4 仿真实验

本研究使用计算机仿真软件进行数值模拟,评估提出方案的性能,实验基于python 3.7版本、Tensorflow 1.15框架进行。实验计算机的硬件配置为英特尔的i5-11300 3.1 GHz CPU以及16 GB运行内存。

4.1 实验参数设置

构建一个如图1所示的超密集网络,网络覆盖范围为半径1 000 m的圆形区域,无人机飞行高度为200 m。宏基站位于网络中心,小基站随机分布在宏小区内,无人机携带STAR-RIS模块在区域内飞行。用户终端和建筑物随机分布在网络覆盖区域内且位置固定不变。通过仿真训练中的变量 $\alpha_{x,y}$ 和 $\beta_{x,y}$ 来决定STAR-RIS是否替换小基站。假设宏基站到STAR-RIS链路和STAR-RIS到用户链路的散射路径数为3。网络的带宽为100 MHz。STAR-RIS的分辨率设置为2,吸收系数设定为 0.0033 m^{-1} 。

本方案使用SAC强化学习策略,用神经网络去模拟值函数,共生成5个神经网络。在神经网络的设计中,本研究统一采用了3层结构,其中,输入层接收输入数据,隐藏层处理数据并提取特征,输出层则生成最终的预测结果。本研究设置了256个隐藏层神经元,并将其连接到输出层。这样可以使得模型能够捕获更多的特征信息和非线性关系,提高预测精度。本研究还使用了常见的Relu函数作为输出层的激活函数,其具有快速收敛和有效降低过拟合的特点。

为了优化神经网络的参数,采用梯度下降优化器反向传播的方法。首先计算网络预测值与真实值之间的误差,然后通过反向传播算法更新各层之间的权重和偏置,从而不断优化模型的预测能力。SAC算法的记忆回放池容量设置为1 000,每次训练

从中随机抽取 256 批次进行训练,每经过 200 回合更新一次神经网络参数。温度系数 α 设为 0.02,折扣因子 γ 设置为 0.9。定义一个单位成本,假定 $C_2 = \lambda C_1$,其中 $\lambda > 0$ 表示 STAR-RIS 成本的权重。综合不同类型 STAR-RIS 元件成本、基站硬件成本、运维成本等^[18-19],为了简单分析起见,取 $\lambda = 3$ 。仿真实验中具体参数的设置如表 2 所示。

表 2 仿真实验参数设置

参数名称	参数值
记忆池大小	100
训练批次大小	256
神经网络更新频率	200
折扣因子 γ	0.9
学习率 δ	0.01
温度系数 α	0.02
激活函数	Relu
优化器	梯度下降优化器
带宽/Hz	10^9
中心频率/THz	0.34
STAR-RIS 相移分辨率/bit	2
分子吸收系数/ m^{-1}	0.003 3
BS 到 STAR-RIS 路径数目	3
STAR-RIS 到用户路径数目	3
高斯白噪声/dBm	10^{-10}

4.2 性能评估

图 6 显示了在不同学习率下 SAC 算法的收敛过程。图 6(a) 为学习率在 0.000 1 下的收敛过程,可以发现训练过程在迭代 500 次左右开始逐渐收敛。图 6(b) 中学习率为 0.001,训练过程在第 300 次左右开始收敛。图 6(c) 中学习率为 0.01,训练过程在第 100~200 次左右逐渐收敛,可以发现训练的收敛速度取决于学习率的大小,学习率越大,利用的误差函数越多,参数更新速度越快,但学习率对训练的结果影响较小。同时可以看出,用户数目越多,系统效益越大。

图 7 分析了最小速率需求对系统效益的影响。系统效益随着最小速率需求先增加后减小,当最小速率需求较大时,训练中会给该用户分配更高的功率,以使其达到速率限制,但这样会造成单个用户能量效率的降低,从而造成整体系统效益的降低。同时,由图 7 可见,当最大功率限制越大时,系统整体的效益越小。

改变 STAR-RIS 数目限制分析系统效益的变化,如图 8 所示。最大 STAR-RIS 数目限制 $I_{\max} = 0$ 表

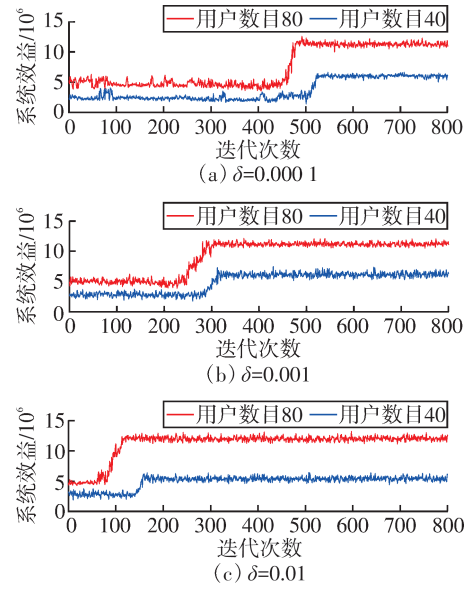


图 6 不同学习率下 SAC 算法的收敛图

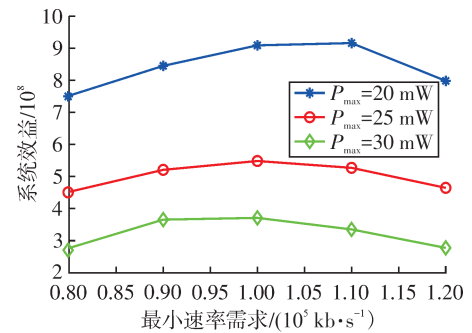


图 7 不同功率限制下系统效益对比

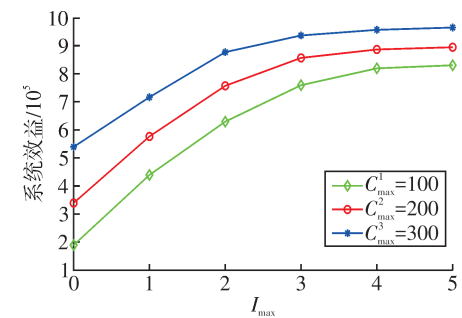


图 8 不同成本限制下系统效益对比

示没有 STAR-RIS 辅助通信。可以看出,随着 STAR-RIS 数目限制的增多,网络环境中可以被取代的小基站也会增多,因此使用 STAR-RIS 辅助的网络效益会增大。但由于 STAR-RIS 只能反射和透射信号,并不能产生信号,因此曲线呈现逐渐平缓趋势。同时图 8 显示,在同一 STAR-RIS 数目限制下,总成本越多可部署的小基站和 STAR-RIS 也越多,系统的效益越大。当小基站和 STAR-RIS 部署的足够多,再增加小基站和 STAR-RIS 会造成成本增加,

导致系统效益增量减少。

对比不同算法之间的性能差异,如图9所示。从图中可以看出,随着用户增多,无论使用何种算法系统效益依旧是递增的趋势,但本文使用的SAC算法相较于DQN算法更加稳定高效,这是由于引入了熵奖励项和策略梯度,通过采用随机性策略探索更大的动作空间,而不是贪婪地选择最优的动作。同时对比使用Chai等^[20]提到的Dinkelbach算法求解优化问题,强化学习算法求解性能更优,因为Dinkelbach算法在处理多约束的优化问题时需要做非凸转换,这个过程中将丢失一部分约束信息。

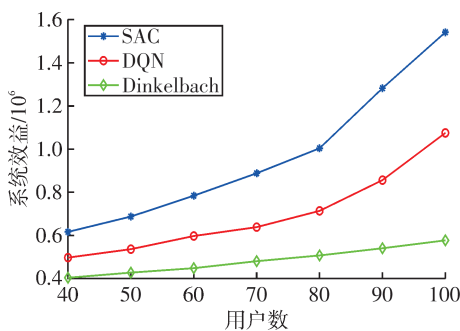


图9 不同算法系统效益对比

5 结束语

针对超密集网络中的网络配置优化问题,提出了一种基于SAC算法的解决方案。通过构建包含宏基站、小基站、用户终端和携带STAR-RIS模块的无人机的超密集网络模型,定义了速率和功率模型,并构建了效益目标函数。利用SAC算法在满足功率、速率和成本约束的条件下求解优化问题,得到最优网络参数配置。仿真结果表明,该方法能够有效提升网络效益,降低部署成本,同时提高通信的可靠性。在5G及未来移动通信网络建设中,该方法可为超密集网络的优化配置提供新的思路和技术支持。然而,当前研究在复杂网络环境下的算法稳定性和收敛速度仍有待提高。未来研究可从以下方向展开:一是探索更高效的优化算法,以适应更大规模和更复杂的网络环境;二是研究多目标优化问题,综合考虑网络性能、成本和用户体验等多方面因素;三是开展实际网络环境中的测试和验证,以进一步验证算法的有效性和可行性。

参考文献:

[1] 尤肖虎,潘志文,高西奇,等. 5G移动通信发展趋势与若干关键技术[J]. 中国科学(信息科学), 2014, 44(5):

551-563.

YOU Xiaohu, PAN Zhiwen, GAO Xiqi, et al. The 5G mobile communication: the development trends and its emerging key techniques[J]. Scientia Sinica (Informationis), 2014, 44(5): 551-563. (in Chinese)

[2] TENG Y L, LIU M T, YU F R, et al. Resource allocation for ultra-dense networks: a survey, some research issues and challenges[J]. IEEE Communications Surveys & Tutorials, 2019, 21(3): 2134-2168.

[3] 齐峰,岳殿武,孙玉. 面向6G的智能反射面无线通信综述[J]. 移动通信, 2022, 46(4): 65-73.

QI Feng, YUE Dianwu, SUN Yu. A survey of intelligent reflecting surface wireless communications toward 6G[J]. Mobile Communications, 2022, 46(4): 65-73. (in Chinese)

[4] 庞海舰,陈健锋,张广驰,等. 智能反射面辅助的无线信息与能量传输研究综述[J]. 中兴通讯技术, 2022, 28(3): 27-35.

PANG Haijian, CHEN Jianfeng, ZHANG Guangchi, et al. Survey on intelligent reflecting surface-assisted wireless information transmission and power transfer[J]. ZTE Technology Journal, 2022, 28(3): 27-35. (in Chinese)

[5] 王亚彬,梁彦,李飞. IRS辅助的无人机系统安全通信能效优化方案[J]. 南京邮电大学学报(自然科学版), 2025, 45(1): 68-78.

WANG Yabin, LIANG Yan, LI Fei. Secure communication energy efficiency optimization for IRS-assisted UAV systems[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2025, 45(1): 68-78. (in Chinese)

[6] AHMED M, WAHID A, LAIQUE S S, et al. A survey on STAR-RIS: use cases, recent advances, and future research challenges[J]. IEEE Internet of Things Journal, 2023, 10(16): 14689-14711.

[7] 付羿博,王鸿. 相移耦合的STAR-RIS辅助NOMA通信系统低功耗传输方案研究[J]. 南京邮电大学学报(自然科学版), 2024, 44(3): 26-32.

FU Yibo, WANG Hong. A low-power transmission scheme for NOMA communication system assisted by phase-coupled STAR-RIS[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2024, 44(3): 26-32. (in Chinese)

[8] ZHONG R K, LIU Y W, MU X D, et al. Hybrid reinforcement learning for STAR-RISs: a coupled phase-shift model based beamformer[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(9): 2556-2569.

[9] WANG Y R, GUAN P X, YU H K, et al. Transmit power

- optimization of simultaneous transmission and reflection RIS assisted full-duplex communications [J]. *IEEE Access*, 2022, 10: 61192–61200.
- [10] SHI B H, WANG Y, LI D Q, et al. STAR-RIS-UAV-aided coordinated multipoint cellular system for multi-user networks[J]. *Drones*, 2023, 7(6): 403.
- [11] ZHANG Q, ZHAO Y, LI H, et al. Joint optimization of STAR-RIS assisted UAV communication systems [J]. *IEEE Wireless Communications Letters*, 2022, 11(11): 2390–2394.
- [12] WANG Y X, LIU Y X, NI Y Y, et al. Performance analysis for STAR-RIS-assisted-UAV-empowered networks under nakagami-m fading [C]//*IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs)*. 2024: 1–6.
- [13] XIAO H, HU X Y, WANG W J, et al. STAR-RIS and UAV combination in MEC networks: simultaneous task offloading and communications [J]. *IEEE Transactions on Communications*, 2025, 73(8): 6169–6184.
- [14] YANG B W, YU Y, HAO X, et al. OH-DRL: an AoI-guaranteed energy-efficient approach for UAV-assisted IoT data collection[J]. *IEEE Transactions on Wireless Communications*, 2025, 24(6): 5009–5022.
- [15] ZHAO J J, ZHU Y B, MU X D, et al. Simultaneously transmitting and reflecting reconfigurable intelligent surface (STAR-RIS) assisted UAV communications[J]. *IEEE Journal on Selected Areas in Communications*, 2022, 40(10): 3041–3056.
- [16] AUNG P S, NGUYEN L X, TUN Y K, et al. Aerial STAR-RIS empowered MEC: a DRL approach for energy minimization[J]. *IEEE Wireless Communications Letters*, 2024, 13(5): 1409–1413.
- [17] ZHANG G B, HAN Z, XIN H. Power control for NOMA HAP-UAV integrated aerial access networks[C]//*IEEE Global Communications Conference (GLOBECOM)*. 2021: 1–6.
- [18] 朱俏楠. 基于智能反射面系统的资源分配技术研究 [D]. 成都: 电子科技大学, 2022.
- ZHU Qiaonan. Research on resource allocation technologies for intelligent reflecting surface aided wireless systems [D]. Chengdu: University of Electronic Science and Technology of China, 2022. (in Chinese)
- [19] NIKOLIKJ V, JANEVSKI T. State-of-the-art comparative cost modeling of heavily-loaded wireless heterogeneous networks [C]//*4th International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems (VITAE)*. 2014: 1–5.
- [20] CHAI J L, FENG L, ZHOU F Q, et al. Energy-efficient resource allocation based on hypergraph 3D matching for D2D-assisted mMTC networks[C]//*IEEE Global Communications Conference (GLOBECOM)*. 2019: 1–7.

(责任编辑:李小溪)