

基于深度强化学习的超密集网络中 多用户上行功率控制方法

毛 晋^{1,2}, 熊 轲^{1,2}, 位 宁^{3,4}, 张 煜⁵, 张锐晨^{1,2}

(1. 北京交通大学计算机与信息技术学院, 北京 100044)

(2. 交通数据分析与挖掘北京市重点实验室, 北京 100044)

(3. 中兴通讯股份有限公司, 广东 深圳 518057)

(4. 移动网络和移动多媒体技术国家重点实验室, 广东 深圳 518055)

(5. 国网能源研究院有限公司, 北京 102209)

[摘要] 针对超密集网络中由于用户数量多、相互距离近, 通信过程中彼此之间干扰大, 导致频谱利用率不高的问题, 建立了通过优化控制发射功率同时提升系统信息容量和满足服务质量的优化问题。由于该问题非凸且功率控制为离散变量, 将其建模为马尔科夫决策过程。在此基础上, 提出了基于深度强化学习的功率控制算法, 并设计了相应的动作空间、状态空间及奖励函数。仿真结果表明, 所提算法与最大发射功率策略和随机发射功率策略相比, 分别提高了至少 15.9% 的信息容量和至少 10.7% 的用户服务质量满足率。与不考虑用户服务质量满足率提升的算法相比, 所提算法通过适当降低信息容量, 提升了用户服务质量满足率。

[关键词] 超密集网络, 功率控制, 信息容量, 服务质量, 深度强化学习

[中图分类号] TP391 **[文献标志码]** A **[文章编号]** 1672-1292(2022)01-0016-08

Power Control in Ultra Dense Network: A Deep Reinforcement Learning Based Method

Mao Jin^{1,2}, Xiong Ke^{1,2}, Wei Ning^{3,4}, Zhang Yu⁵, Zhang Ruichen^{1,2}

(1. School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China)

(2. Beijing Key Laboratory of Traffic Data Analysis and Mining, Beijing 100044, China)

(3. ZTE Corporation, Shenzhen 518057, China)

(4. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China)

(5. State Grid Energy Research Institute Co., Ltd., Beijing 102209, China)

Abstract: For ultra-dense networks, in view of the problem of low spectrum utilization due to excessive users and large interference, an optimization problem is formulated to increase the system information capacity and satisfy the number of users with the quality of service (QoS) by optimizing the transmission power. Since the problem is non convex and the power control is a discrete variable, it is modeled as a Markov decision policy process. To this end, a power control algorithm based on deep reinforcement learning is proposed, and the corresponding action space, state space and reward function are designed. Simulation results show that compared with the maximum transmit power strategy and random transmit power strategy, the proposed algorithm improves the information capacity by at least 15.9% and the satisfaction of users' QoS by at least 10.7%. Moreover, compared with the algorithm without considering the improvement of user's QoS, the proposed algorithm improves the user's QoS by appropriately reducing the information capacity.

Key words: ultra-dense networks, power control, information capacity, QoS, deep reinforcement learning

随着移动互联网应用的大规模普及, 网络流量呈现爆炸式增长。根据 Triton Market Research 调研预测^[1], 全球网络流量在 2019 至 2028 年期间, 将以 23.05% 复合年增长率增加。另外, 5G 技术的商用和物联

收稿日期: 2021-08-31.

基金项目: 国家自然科学基金项目 (62071033)、国家重点研发计划项目 (2020YFB1806903)、国网能源研究院有限公司研究项目 (526700190002)。

通讯作者: 熊轲, 博士, 教授, 研究方向: 无线网络、物联网、网络信息论。E-mail: kxiong@bjtu.edu.cn

网技术的发展,使得将有超过一千万亿的设备^[2]逐步接入到无线网络中.高流量密度和高密集接入导致无线网络部署密集化,很多场景已发展成为超密集网络.

由于无线频谱资源非常有限,已成为稀缺资源,很难再通过增加频谱带宽提升网络容量.通过功率控制有效协调无线用户间的干扰已成为提升超密集网络频谱效率的重要手段.

当前,网络直播、线上会议等应用逐渐发展成为主流无线业务.与此同时,工业自动化、自动驾驶、虚拟现实(virtual reality, VR)、增强现实(augmented reality, AR)等新型上行数据密集应用也正在普及^[3-4].用户上行数据传输的服务质量(quality-of-service, QoS)变得越来越重要.在超密集网络中,用户间距离相对较近,彼此之间的干扰较大,加之上行带宽在总带宽中占比不高,上行链路的干扰模式与相邻小区用户相关^[5],使得提升网络上行吞吐量和用户服务质量更具挑战.文献[6]针对多蜂窝小区网络,利用频率复用提升了系统的上行吞吐量.文献[7-8]在多用户多小区网络中,通过控制用户发射功率,提升了系统的上行吞吐量.文献[9]联合功率和无线频谱资源分配提升了设备对设备(device-to-device, D2D)通信网络的信息容量.文献[10]通过干扰消除技术提升了多用户多小区网络的上行吞吐量.

上述工作通过凸优化或博弈论对相关优化问题进行分析求解,所提算法多采用迭代或启发式搜索进行,对小型网络有效,随着网络规模的扩大,执行复杂度会急剧增加,因此难以适用于大规模网络.此外,由于这些算法均由模型驱动,需收集网络全局的信道状态信息(channel state information, CSI),在实际的动态无线环境中,信令交互开销大,运行效率低.

近年来,深度强化学习(deep reinforcement learning, DRL)因其无模型的特性及自适应学习的能力,受到了工业界和学术界的广泛关注^[11].深度Q网络(deep-q-network, DQN)使用神经网络拟合价值函数,更适合处理高维度数据(连续状态)的情况.目前,DQN已被用于解决无线网络的优化设计问题.文献[12-14]基于DQN算法优化了蜂窝小区的功率分配方案,提升了系统的总吞吐量.文献[15]基于DQN算法联合优化信道选择和功率分配以提升网络的信息容量.然而,上述工作仅关注了网络总吞吐量的提升,忽略了用户对上行QoS的需求.虽然文献[16]针对两层异构网络,研究了基于强化学习的自适应功率控制方法,提升了微小区的信息总容量,并确保了用户的QoS,但其采用了基于用户到基站的距离的奖励函数,用户与基站之间的距离并不容易获取,因此该方法在超密集网络中并不适用.若采用集中式算法对该问题进行求解,存在复杂度高的问题,加之集中式算法需要基站间进行通信,代价较高且不易实现.

为此,本文针对超密集网络场景,以提升上行网络的吞吐量为目标,同时考虑提升满足服务质量的用数量,对用户进行有效的功率控制和调节.由于问题非凸且功率控制为离散变量,故基于深度强化学习设计分布式算法对其进行求解.

1 系统模型

1.1 网络模型

考虑一个上行多小区蜂窝网络模型,如图1所示.该网络由 N 个蜂窝小区组成,其中每个小区中心部署一个基站(base station, BS),一个小区同时服务 K 个用户.同一小区中的用户使用相互正交的频段进行上行通信,因此不存在小区内干扰.不同小区之间共享同一段频谱资源,故小区间干扰不能忽略.为了表达清晰,使用 $n \in N = \{1, 2, 3, \dots, N\}$ 和 $k \in K = \{1, 2, 3, \dots, K\}$ 分别表示基站和用户的索引.

1.2 信道模型

同时考虑大尺度和小尺度衰落.由于实际无线场景中信道随时间不断变化,因此采用时变信道模型,即 t 时隙时小区 n' 中用户 k 到基站 n 之间的信道增益 $g_{n',k \rightarrow n}^{(t)}$ 可表示为

$$g_{n',k \rightarrow n}^{(t)} = |h_{n',k \rightarrow n}^{(t)}|^2 \beta_{n',k \rightarrow n}, \quad (1)$$

式中, $h_{n',k \rightarrow n}^{(t)}$ 、 $\beta_{n',k \rightarrow n}$ 分别为 t 时隙时小区 n' 的用户 k 与基站 n 之间的小尺度瑞利衰落和路径损失.根据

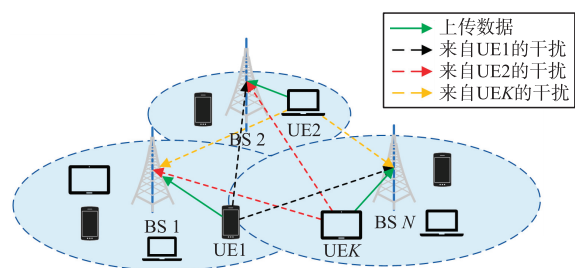


图1 系统模型

Fig. 1 The system model

Jake's mode 小尺度瑞利衰落 $h_{n',k \rightarrow n}^{(t)}$ 在时间上的相关性^[17], 可以被建模为一阶复高斯-马尔可夫过程, 表示为

$$h_{n',k \rightarrow n}^{(t)} = \rho h_{n',k \rightarrow n}^{(t-1)} + \sqrt{1-\rho^2} e_{n',k \rightarrow n}^{(t)}, \quad (2)$$

式中, $e_{n',k \rightarrow n}^{(t)} \sim \mathbb{C} \mathcal{N}(0, 1)$ 为复高斯随机变量; ρ 为两个相邻时隙的信道增益的相关系数, 被定义为

$$\rho = J_0(2\pi f_d T_s), \quad (3)$$

式中, $J_0(\cdot)$ 、 f_d 、 T_s 分别为第一类零阶贝塞尔函数、最大多普勒频率和相邻时刻之间的时间间隔. 不失一般性, 各个小区的第 k 个用户使用同频载波, 在时隙 t 用户 k 到基站 n 之间的上行链路的信号与干扰加噪声比(signal to interference plus noise ratio, SINR)为

$$\text{SINR}_{n,k}^{(t)} = \frac{g_{n,k \rightarrow n}^{(t)} p_{n,k}^{(t)}}{\sum_{n' \in N, n' \neq n} g_{n',k \rightarrow n}^{(t)} p_{n',k}^{(t)} + \sigma^2}, \quad (4)$$

式中, $g_{n,k \rightarrow n}^{(t)}$ 为时隙 t 小区 n 中用户 k 到基站 n 的信道增益; $p_{n,k}^{(t)}$ 为小区 n 中用户 k 的发射功率; σ^2 为加性高斯白噪声. 由香农公式可知, 用户 k 到基站 n 的链路的信息容量为

$$C_{n,k}^{(t)} = \log_2(1 + \text{SINR}_{n,k}^{(t)}), \quad (5)$$

因此, 网络上总容量为

$$C^{(t)} = \sum_n \sum_k C_{n,k}^{(t)}. \quad (6)$$

定义用户的速率大于等于阈值速率为满足其 QoS, 则在时隙 t 系统总满足 QoS 的用户数量为

$$I^{(t)} = \sum_n \sum_k \mathbb{I}(C_{n,k}^{(t)} \geq C_{th}), \quad (7)$$

式中, $\mathbb{I}(\cdot)$ 为指示函数. 记功率控制策略向量 $\mathbf{p}^{(t)} = \{p_{n,k}^{(t)}, \forall t, n \in N, k \in K\}$, 网络设计的目标应为提升信息容量, 同时提升满足服务质量的用户的数量, 可将其表示为

$$\max_{\mathbf{p}^{(t)}} (C^{(t)} + \beta I^{(t)}), \quad (8)$$

式中, β 为权重.

由式(5)可知, 提升网络容量的关键是提升用户的 SINR. 式(4)表明可以通过增加用户发射功率、降低干扰用户的功率的方式提升当前用户的 SINR. 因此, 合理地控制用户的发射功率, 能有效提升信息容量, 并提升满足服务质量的用户的数量.

2 优化问题建模

记 t 时隙小区 n 的用户 k 的发射功率为 $p_{n,k}^{(t)}$, 则优化问题可描述为

$$\begin{aligned} & \max_{\mathbf{p}^{(t)}} (C^{(t)} + \beta I^{(t)}) \\ \text{s.t. } & C_1: 0 \leq p_{n,k}^{(t)} \leq p_{\max}, \\ & C_2: n \in N, k \in K, \end{aligned} \quad (9)$$

式中, C_1 表示用户的发射功率约束; p_{\max} 表示用户的最大发射功率. 由于优化目标是非凸的, 故式(9)是一个非凸问题. 当接入用户增多时, 收集整个网络的信道信息并以此优化整个系统, 会造成非常大的信令开销以及计算成本, 使用传统的数学解法难以直接求解. 因此, 本文提出一种基于数据驱动的深度强化学习算法进行求解.

3 基于 DQN 的 QoS 感知功率控制算法

集中式算法复杂度高, 算法需要基站间进行通信, 代价较高且不易实现. 为了降低算法复杂度并减少开销, 设计分布式算法. 将每个基站当作一个智能体, 故式(9)对应的功率控制问题可被构建为多智能体系统. 现有工作基于多智能体深度确定性策略算法(multi-agent deep deterministic policy gradient, MADDPG), 在训练过程中每个智能体需对应一套网络(Actor 网络和 Critic 网络). 当基站数量增加时, 训练过程中存在大量的模型, 使得训练难度增加. 且训练过程中 Critic 网络需要获取所有智能体 Actor 网络的状态和动作, 当智能体数量较多时, Critic 网络的状态空间较大, 不利于算法的收敛. 由于所有基站具有

相同的动作空间、状态空间和奖励函数,同时考虑到多智能体训练涉及到的神经网络参数多、训练不稳定的问题^[18],考虑将 DQN 算法设计为集中式训练分布式执行的模式. 训练过程中只需要一套网络,参数规模小;训练过后将整套网络的参数复制到其他基站分布式执行,无需其余开销. 具体来说,在集中式训练过程中,由一个智能体获取到系统中每个基站的相关信息,即该智能体可获得整个系统中所有基站产生的经验以用于训练. 该智能体学习到系统中所有用户功率分配策略,然后将该智能体的参数进行拷贝,发送到其余基站中. 各个基站使用该网络参数获取与当前用户相关的局部信息即可进行相关决策.

图 2 为基于 DQN 的 QoS 感知功率控制算法框架. 首先,智能体从环境中获得初始状态 s ,智能体通过当前状态决策一个动作 a ,该动作作用于环境后环境变化到达下一状态,同时给智能体反馈一个奖励 r . 智能体通过该奖励 r 判断动作的好坏,学习好的动作,避免坏的动作. 根据该学习原则,经多次迭代后,智能体最终学习到能获得更多正值奖励的策略.

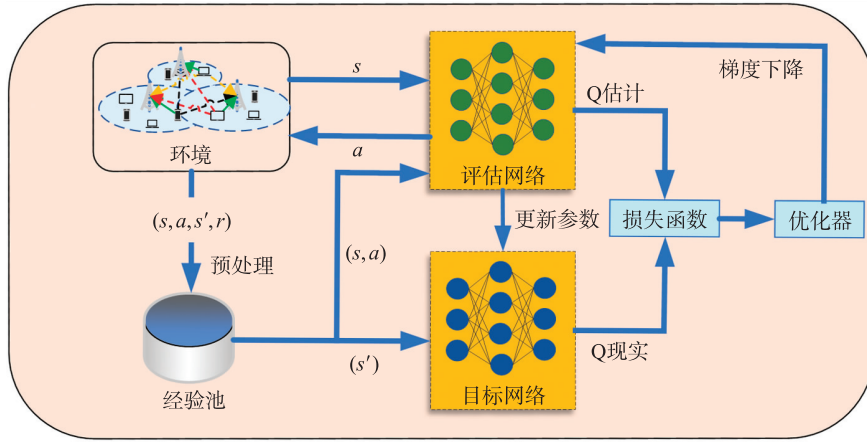


图 2 基于 DQN 的 QoS 感知功率分配算法框架

Fig. 2 The framework of proposed DQN-based QoS aware power control algorithm

用 Q^* 表示最优动作价值函数,则用神经网络逼近该函数可表示为

$$Q(s, a | \theta_i) \approx Q^*(s, a), \quad (10)$$

式中, θ_i 表示在迭代轮数 i 神经网络的参数. $Q(s, a | \theta_i)$ 的更新方式为

$$Q(s, a) \leftarrow (1 - \alpha) Q(s, a | \theta_i) + \alpha (r + \gamma \max_{a'} Q(s', a' | \theta_{i-1})), \quad (11)$$

式中, s' 表示下一个状态. 上述估计 Q 值的过程中,若仅使用一个网络估计当前状态采取动作的价值更新自身,会使得当前 Q 值和目标 Q 值相关性过大,训练过程变得不稳定. 因此构建一个和估计网络结构完全相同但参数不相同的目标网络用以计算目标 Q 值,每过一段时间从估计网络中将参数复制过来(延时更新),这样可使得在一段时间内目标 Q 值不变,一定程度上可使训练更加稳定. 可通过最小化估计网络的损失函数序列 $L_i(\theta_i)$ 来训练神经网络,定义为

$$L_i(\theta_i) = (y_i - Q(s, a | \theta_i))^2, \quad (12)$$

式中, y_i 为迭代轮数 i 的目标,表达式为

$$y_i = r + \gamma \max_{a'} Q(s', a' | \theta_{i-1}), \quad (13)$$

式中, a' 为下一状态 s' 对应 Q 值最大的动作; θ_{i-1} 为前一轮固定下来的神经网络参数,即目标网络. 根据权重对损失函数进行微分,其梯度为

$$\nabla_{\theta_i} L_i(\theta_i) = (r + \gamma \min_{a'} Q(s', a' | \theta_{i-1}) - Q(s, a | \theta_i)) \times \nabla_{\theta_i} Q(s, a | \theta_i). \quad (14)$$

由于生成经验之间可能存在一定相关性,加入经验回放机制,通过随机采样方式对网络进行更新去除样本之间的相关性. 将算法构建为马尔可夫决策过程,构建三元组 (s, a, r) , 其中, s 为状态空间, a 为动作空间, r 为奖励函数,具体定义如下:

(1) 动作空间 (Action)

智能体的动作是用户的发射功率,但 DQN 的动作空间必须是有限的,定义智能体的动作个数为 m ,将发射功率指数离散化分为 m 个等级. 动作集合 \mathbf{P} 为

$$\mathbf{P} = \left\{ 0, P_{\min}, P_{\min} \left(\frac{P_{\max}}{P_{\min}} \right)^{\frac{1}{m-2}}, \dots, P_{\max} \right\}, \quad (15)$$

式中, P_{\max} 、 P_{\min} 分别表示最大和最小发射功率.

(2) 状态空间(State)

状态空间包含两部分:该用户与对该用户有干扰的用户的信道增益信息和上一时隙用户的速率以及使用的发射功率. 因此小区 n 中用户 k 的状态 $S_{n,k}^{(t)}$ 表示为

$$S_{n,k}^{(t)} = \{ \Theta_{n,k}^{(t)}, P_{n,k}^{(t-1)}, C_{n,k}^{(t-1)} \}. \quad (16)$$

标准化后的信道增益定义为 $\Theta_{n,k}^{(t)}$, 计算式为

$$\Theta_{n,k}^{(t)} = \log_2 \left(\frac{\mathbf{g}_{n',k}^{(t)}}{\mathbf{g}_{n,k}^{(t)}} \otimes \mathbf{1}_k \right), \quad (17)$$

式中, 运算符 \otimes 、 $\mathbf{1}_k$ 分别表示 Kronecker 积和 k 个全为 1 的向量. 同时由于在实际环境中存在大尺度衰落, 使用对数可以有效减少信道振幅数量级不相同带来的影响. $P_{n,k}^{(t-1)}$ 、 $C_{n,k}^{(t-1)}$ 分别表示与 $\Theta_{n,k}^{(t)}$ 对应的上一时隙使用的发射功率和达到的速率.

(3) 奖励函数(Reward)

奖励函数包含系统总信息容量和当前用户的信息容量两部分. 分段式奖励函数为

$$r_{n,k}^{(t)} = \begin{cases} k_c C_k^{(t)} + k_{u1} \frac{(C_{n,k}^{(t)} - \varphi_u)}{\varphi_u}, & C_{n,k}^{(t)} > \varphi_u; \\ k_c C_k^{(t)} - k_{u2} \frac{(\varphi_u - C_{n,k}^{(t)})}{\varphi_u}, & C_{n,k}^{(t)} \leq \varphi_u. \end{cases} \quad (18)$$

式中, $C_k^{(t)}$ 为与用户 k 同频率所有用户的速率; $C_{n,k}^{(t)}$ 为小区 n 用户 k 的信息容量; $k_c, k_{u1} \in (0, 1)$, $k_{u2} \in (1, \infty)$ 为超参数. 当信息总容量增加时, 奖励也应当增加. 同时, 若当前用户的速率满足阈值速率时, 则奖励应增加得较少, 反之当前用户的速率不能满足速率阈值时应给予一个较大的惩罚.

基于 DQN 的 QoS 感知功率分配算法伪代码如表 1 所示, 首先初始化超参数和网络参数; 在每个训练周期初始化智能体的初始状态; 在 0 到 t 的每个时隙内, 智能体使用 ε -greedy 策略选择动作, 防止陷入局部最优; 智能体获得奖励, 并进入下一状态, 将刚才的经验存入记忆库中; 从记忆库中随机抽取批量经验, 计算损失后进行梯度下降.

表 1 基于 DQN 的 QoS 感知功率控制算法

Table 1 DQN-based QoS aware power control algorithm

算法 1 基于 DQN 的 QoS 感知功率控制算法	
输入: 智能体的状态	10. 获得下一状态 $s_{n,k}^{(t+1)}$
输出: 智能体功率选择策略	11. 将 $(\varphi(s_{n,k}^{(t)}), a_{n,k}^{(t)}, r_{n,k}^{(t)}, \varphi(s_{n,k}^{(t+1)}))$ 存入经验库
1. 初始化超参数, 网络参数 θ , 预处理操作 $\varphi(\cdot)$	12. 更新 $s_{n,k}^{(t)} = s_{n,k}^{(t+1)}$
2. While 迭代次数 < 最大迭代次数 Do	13. 从经验库中随机抽取批量经验 $(\varphi(s_{n,k}^{(t)}), a_{n,k}^{(t)}, r_{n,k}^{(t)}, \varphi(s_{n,k}^{(t+1)}))$
3. 初始化智能体状态 $s_{n,k}^{(0)} = s_{\text{init}}$	14. 根据式(12)计算损失
4. For 0 到 t Do	15. 根据式(14)梯度下降
5. If $\text{rand}(0, 1) < \varepsilon$ Then	16. 每隔 J 轮更新目标网络的参数 $\theta' = \theta$
6. 在动作空间中随机选择动作 $a_{n,k}^{(t)}$	17. End For
7. Else	18. End While
8. $a_{n,k}^{(t)} = \arg \max_a Q(\varphi(s_{n,k}^{(t)}), a \theta)$	
9. 执行动作 $a_{n,k}^{(t)}$, 获得一个奖励 $r_{n,k}^{(t)}$	

4 实验与结果

本节从算法的收敛性、有效性(与最大发射功率策略和随机发射功率策略相比提升的信息容量和用户服务质量满足率)对所提算法进行仿真验证. 在仿真中, 大尺度路径衰落建模为

$$PL = 10^{-(120.9 + 37.6 \log_{10}(d/1000))/20}. \quad (19)$$

用户上行的信息容量阈值的默认值 $\varphi_u = 0.6$ (bps/Hz). 智能体使用 4 层前向传播网络,其中第一层为输入层,中间两层为隐藏层,第四层为输出层. 实验中,智能体的学习率 $lr = 10^{-3}$,经验重放池的大小为 20 000,每次训练的批次数量为 128. 初始的 $\varepsilon = 0.1$,随着训练轮数增加, ε 衰减至 0.001. 仿真参数如表 2 所示.

4.1 算法收敛性

图 3(a)-(e)展示了小区个数分别为 4、9、16、25、36 场景下的累计奖励变化趋势,其中横坐标表示训练的轮数,纵坐标表示累计奖励. 从图中可以看出,随着小区数量的增加,算法收敛所需要的训练轮数也在增加,这是因为小区数量增加导致相同频率上的用户数量增加,环境的情况也就更复杂,智能体需要更多的训练轮数使得算法收敛. 此外,智能体的奖励在训练初期抖动较大,这是因为 ε -greedy 使得在训练初期探索概率比较大,之后环境变化会引起小幅抖动,使得每次获得的奖励不同.

表 2 仿真参数

Table 2 The simulation parameter

参数	含义	取值
f_d	最大多普勒频率	10 Hz
T_s	时间间隙	20 ms
σ_0^2	噪声功率	-114 dBm
p_{\min}	发射最小功率	5 dBm
p_{\max}	发射最大功率	23 dBm
d	小区半径	1 km
SINR_{up}	信噪比的上界	30 dB

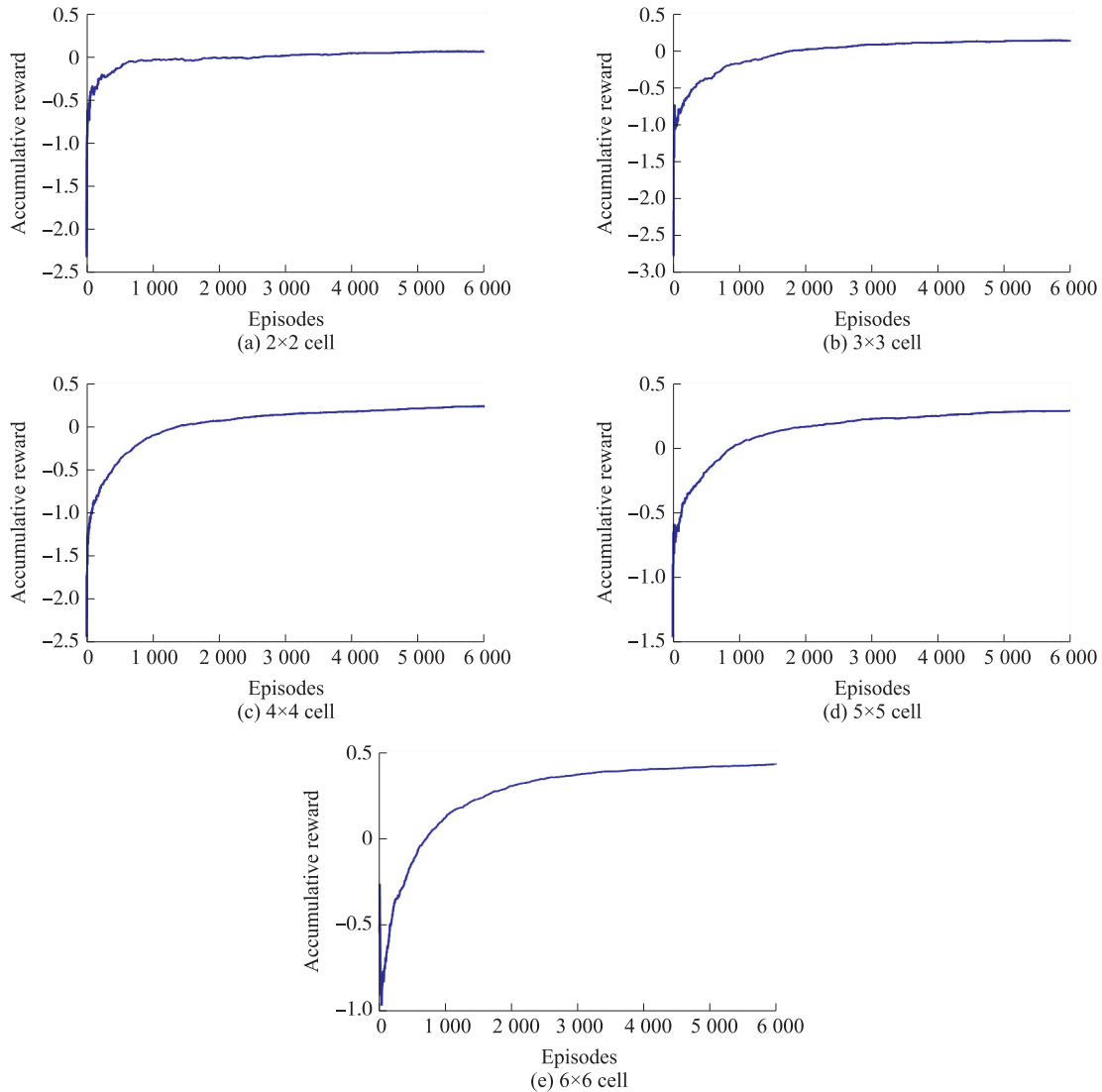


图 3 算法在不同小区数量下的收敛情况

Fig. 3 Algorithm convergence behavior in different number of cells

4.2 算法比较

图 4 和图 5 给出了不同算法随着小区数量增加对系统信息容量和 QoS 满足率的影响情况. 本文算法在可达信息容量和 QoS 满足率两方面均优于传统算法,所提算法与最大发射功率策略和随机发射功率策略相比分别提高了至少 15.9% 的信息容量和至少 10.7% 的用户服务质量满足率. 从图 4 和图 5 可以看出,

随着小区数量的增加,信息容量和 QoS 满足率都在减少. 因为随着小区数量的增加,处于同一载波的用户数量也在增加,彼此之间的干扰随之增加,从而影响了信息容量. 在该环境中,动作的选择对于下一状态的影响很大,而 Dueling DQN 更适合于动作与状态解耦的环境,因此在本文描述环境下 Dueling DQN 会产生过多的无用参数从而影响了效果.

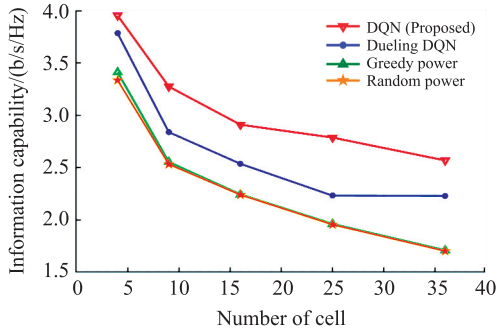


图 4 不同算法下信息容量的变化关系

Fig. 4 Information capability of different algorithms

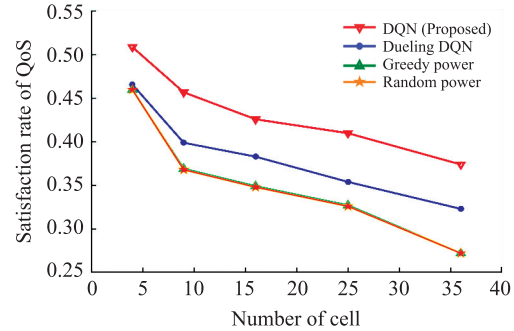


图 5 不同算法下 QoS 满足率的变化关系

Fig. 5 Satisfaction rate of QoS of different algorithms

4.3 用户服务质量满足率目标对信息容量和服务质量满足率的影响

图 6 和图 7 对最大化系统总信息容量的过程中是否提升满足 QoS 的用户数量进行了比较,给出了系统信息容量及 QoS 满足率随着小区数量的增加而变化的趋势. 如图所示,与仅提升网络容量的算法相比,本文算法会适当降低信息容量以提升 QoS 满足率,从而保证用户间的公平性,避免资源过度倾斜于信道好的用户.

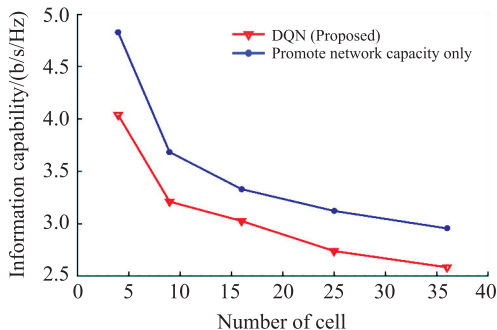


图 6 信息容量与提升满足 QoS 的用户数量的变化关系

Fig. 6 Information capability with considering QoS objective

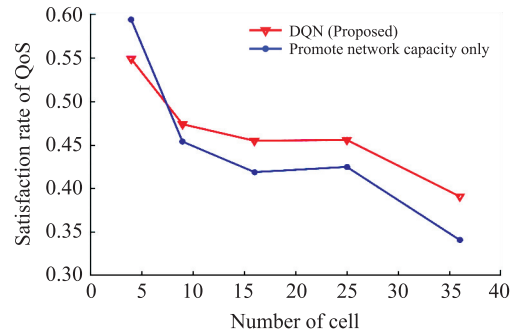


图 7 QoS 满足率与提升满足 QoS 的用户数量的变化关系

Fig. 7 Satisfaction rate of QoS with considering QoS objective

4.4 用户分布对信息容量和服务质量满足率的影响

图 8 和图 9 给出了不同用户分布半径下系统信息容量和 QoS 满足率的变化趋势. 横轴表示用户分布半径距离基站的最小距离,即增加表示用户更接近小区的边缘. 图 8 和图 9 都呈现出递减的趋势,这是因为随着用户距离边缘越来越远,用户间的相互干扰也越来越小,无论是信息容量还是用户 QoS 满足率都在降低.

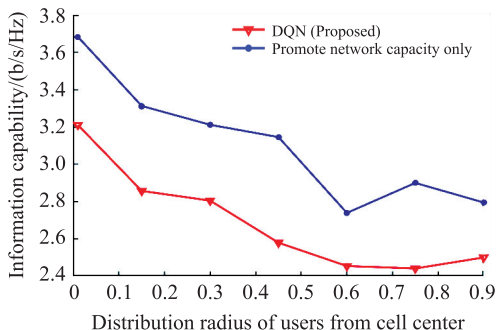


图 8 信息容量与用户分布半径的变化关系

Fig. 8 Information capability with different user distribution radius

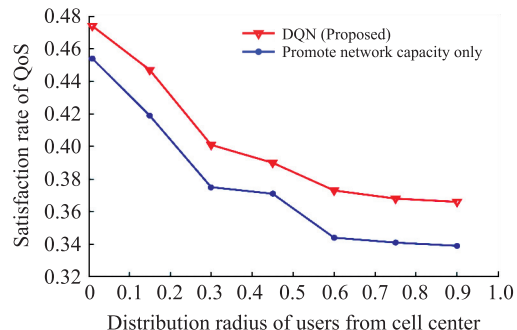


图 9 QoS 满足率与用户分布半径的变化关系

Fig. 9 Satisfaction rate of QoS with different user distribution radius

5 结论

本文研究了超密集网络中服务质量感知的功率控制问题. 为了解决该非凸问题,设计了数据驱动的无模型 DQN 的功率控制算法,通过对小区用户的上行功率进行调节来协调处于相同频率的小区间用户之间的干扰,在提升系统信息容量的同时提升服务质量满足率. 仿真结果证明了算法的收敛性,通过与最大发射功率策略和随机发射功率策略对比,证明了算法的有效性. 实验结果分析表明,提出的算法与单纯提升网络容量相比,会降低信息容量以提升服务质量满足率. 小区数量越多,小区用户距离小区边缘越近,系统的信息容量和服务质量满足率越小.

[参考文献] (References)

- [1] 新天域互联. 全球网络流量分析市场将以 23.05% CAGR 成长 [Z/OL]. [2020-06-23] https://www.sohu.com/a/403637502_100161396.
- [2] OLWAL T O, DJOUANI K, KURIEN A M. A survey of resource management toward 5G radio access networks[J]. IEEE Communications Surveys and Tutorials, 2016, 18(3): 1656–1686.
- [3] NAVARRO-ORTIZ J, ROMERO-DIAZ P, SENDRA S, et al. A survey on 5G usage scenarios and traffic models[J]. IEEE Communications Surveys and Tutorials, 2020, 22(2): 905–929.
- [4] KHURPADE J M, RAO D, SANGHAVI P D. A survey on IOT and 5G network[C]//Proceedings of the 2018 International Conference on Smart City and Emerging Technology (ICSCET). Mumbai, India: IEEE, 2018: 1–3.
- [5] SHEN K M, YU W. A coordinated uplink scheduling and power control algorithm for multicell networks[C]//Proceedings of the 2015 49th Asilomar Conference on Signals, Systems and Computers. Pacific Grove, USA: IEEE, 2015: 1305–1309.
- [6] ELWEKEIL M, ALGHONIEMY M, MUTA O. Dynamic autonomous frequency reuse for uplink cellular networks[C]//Proceedings of the 2018 IEEE International Conference on Consumer Electronics (ICCE). Las Vegas, USA: IEEE, 2018: 1–5.
- [7] VISALI M, SAKURU K L V S. Power control based resource allocation in LTE uplinks[C]//Proceedings of the 2015 International Conference on Communications and Signal Processing. Melmaruvathur, India: IEEE, 2015: 0579–0582.
- [8] SHEN K M, YU W. Fractional programming for communication systems—Part I: power control and beamforming[J]. IEEE Transactions on Signal Processing, 2018, 66(10): 2616–2630.
- [9] NINGOMBAM D D, SHIN S. Radio resource allocation and power control scheme to mitigate interference in device-to-device communications underlying LTE-A uplink cellular networks[C]//Proceedings of the 2017 International Conference on Information and Communication Technology Convergence (ICTC). Jeju, Korea: IEEE, 2017: 961–963.
- [10] ZEINEDDINE K, HONIG M L, NAGARAJ S. Uplink power allocation for distributed interference cancellation with channel estimation error[J]. IEEE Transactions on Wireless Communications, 2016, 15(10): 6785–6796.
- [11] 王云, 韩伟. 一种基于划分和集成思想的多智能体强化学习[J]. 南京师范大学学报(工程技术版), 2008, 8(4): 59–62.
- [12] MENG F, CHEN P, WU L N, et al. Power allocation in multi-user cellular networks: deep reinforcement learning approaches[J]. IEEE Transactions on Wireless Communications, 2020, 19(10): 6255–6267.
- [13] GHADIMI E, CALABRESE F D, PETERS G, et al. A Reinforcement learning approach to power control and rate adaptation in cellular networks[C]//Proceedings of the 2017 IEEE International Conference on Communications. Paris, France: IEEE, 2016: 1–7.
- [14] NASIR Y S, GUO D N. Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks[J]. IEEE Journal on Selected Areas in Communications, 2019, 37(10): 2239–2250.
- [15] TAN J J, ZHANG L, LIANG Y C. Deep reinforcement learning for channel selection and power control in D2D networks[C]//Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM). Waikoloa, USA: IEEE, 2020: 1–6.
- [16] ZHANG R C, XIONG K, GUO W, et al. Q-learning-based adaptive power control in wireless RF energy harvesting heterogeneous networks[J]. IEEE Systems Journal, 2020, 15(2): 1861–1872.
- [17] DENT P, BOTTOMLEY G E. Jakes fading model revisited[J]. Electronics Letters, 1993, 29(13): 1162–1163.
- [18] O'SHEA T, HOYDIS J. An introduction to deep learning for the physical layer[J]. IEEE Transactions on Cognitive Communications and Networking, 2017, 3(4): 563–575.

[责任编辑: 严海琳]