

doi: 10.11835/j.issn.1000-582X.2023.04.009

5G 基站自适应天馈系统设计与建模

沈煜航, 王 晟

(电子科技大学 信息与通信工程学院, 成都 611731)

摘要: 为了提供一个各方面更优的全自动天面自适应调整方案, 在降低维护成本的同时实现更优覆盖效果, 从 5G 天面的信号辐射方向调整方法入手, 对 5G 基站自适应天馈系统的智能调节系统设计关键技术进行研究, 提出对基于深度强化学习的基站天面自适应调节策略。基于此设计了 5G 基站自适应天馈系统, 可以使用电信公司 RSRP 信号覆盖地图作为数据源, 获取当前状态的观测值并自动分析数据, 对天面进行自动调整。在虚拟环境下, 对基于强化学习的系统进行了模拟搭建与仿真训练, 结果符合预期。

关键词: 5G 基站; 强化学习; 天馈系统; 自适应调整; 系统设计; 仿真分析

中图分类号: TN929.5

文献标志码: A

文章编号: 1000-582X(2023)04-089-08

Design and modeling of 5G base station adaptive antenna feed system

SHEN Yuhang, WANG Sheng

(School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, P. R. China)

Abstract: To provide a fully automatic antenna adaptive adjustment scheme with advantages of better performance, wider coverage and lower maintenance cost, the key design technologies of intelligent adjustment system of adaptive antenna feed system of 5g-based station are studied from the perspective of signal radiation direction adjustment of antenna panel. An adaptive adjustment strategy for base-station antenna based on deep reinforcement learning is proposed. The adaptive antenna feed system designed with the proposed strategy can use telecom RSRP coverage map as a data source, and obtain the current state of the observed values to automatically analyze data and adjust the antenna panels. In a virtual environment, the system based on reinforcement learning is simulated and trained, and the results are in line with expectations.

Keywords: 5G base-station; reinforcement learning; antenna feed system; adaptive adjustment; systematic design; simulated analysis

收稿日期: 2021-12-25 网络出版日期: 2022-06-09

基金项目: 国家自然科学基金资助项目(62001087, 62072079)。

Supported by National Natural Science Foundation of China(62001087, 62072079).

作者简介: 沈煜航(1999—), 男, 硕士研究生, 主要从事智慧通信网络与智能信息处理方向研究, (E-mail)shenyh327@163.com。

通信作者: 王晟, 男, 教授, 博士生导师, 主要从事网络规划, 下一代互联网与下一代光网络方向研究, (E-mail)wsh_keylab@uestc.edu.cn。

近年来,中国在第五代移动通信技术的研发上抢占先机^[1-2]。移动通信网络中,天馈系统是整个系统中最重要部分,它直接影响用户的移动接入、数据传输质量。电信公司为了提升优化效率,设计了一套可视化的基站信号覆盖地图。但是,用覆盖地图优化的工单系统,大部分工单是在进行天面的调整。现有研究过于纠结如何将移动网络优化经验统合成专家系统,这类研究比较适合应用于基站建设初期的规划布局,在移动网络运营中,难以在复杂多变的空间环境高效地达到最优解。因此,电信公司开始将研究重心转移到结合人工智能技术的5G网络极简运营方向。引入自适应的机器学习^[3-6]进行系统优化,消减不必要的成本开销,提高5G网络的效率,这将是天馈系统维护优化与移动通信系统极简运营的主要发展态势。

为了提供一个各方面都更优的全自动天面自适应调整方案,让天面调整优化策略能够完全脱离人工干涉,并在最大限度降低维护成本的同时实现更优覆盖效果,研究设计一个5G基站自适应天馈系统,它基于深度强化学习技术^[7-13],用电信公司信号覆盖地图作为数据源,能够自动分析数据、获取当前状态的观测值,并根据观测值判断哪些天面需要调整、应该在什么方向上调整多少角度。希望该系统具有泛用性,最终的学习效果应能快速适应空间环境变化,可更快找到最优的调整方案,且调整方案比人工更优。针对5G网络天馈系统的自适应优化策略,以天馈系统信号辐射方向的智能调整为切入点,对5G基站自适应天馈系统的智能调节系统设计关键技术进行了研究,提出基于深度强化学习的基站天面自适应调节策略,进行系统设计分析与建模。

1 相关工作

在5G场景下,基站天面的角度可以进行数字程控^[14],且5G基站能够进行高效率的站间通信。移动网络接入侧部署有高算力的边缘计算节点^[15],这让基于人工智能技术的天面自调整具有现实可行性。可以将5G自适应天馈系统分为了2个阶段:1)云端训练阶段。系统被部署在云端的计算密集型服务器上,根据所辖区域进行虚拟环境建模,并在虚拟环境中执行所有智能体策略的预训练。2)边缘部署阶段。云端服务器把完成预训练的神经网络参数迁移到对应的边缘站中,基于该初始参数,智能体在真实业务场景中进行更安全的策略探索。

电信公司提供了实时的参考信号接收功率(reference signal receiving power, RSRP)信号覆盖地图。5G场景下,服务区内的用户被接入移动通信网络后,移动终端将采集信道特征(信号强度、比特率、时延、丢包率等)反馈给接入站。接入站通过移动通信基站定位技术算出用户坐标,并按地理区域划分用户集,将采集到的用户数据分组汇集,得到实时的信号覆盖情况分布数据。通过上述手段采集用户信号质量RSRP值,以 Δt 为更新间隔,在地图上对应的位置上以不同的颜色呈现当前时刻各位置的信号覆盖情况。

如图1所示的电信公司RSRP信号覆盖地图中,将区域划分成了以边长为 Δx m的众多方格,每个方格上

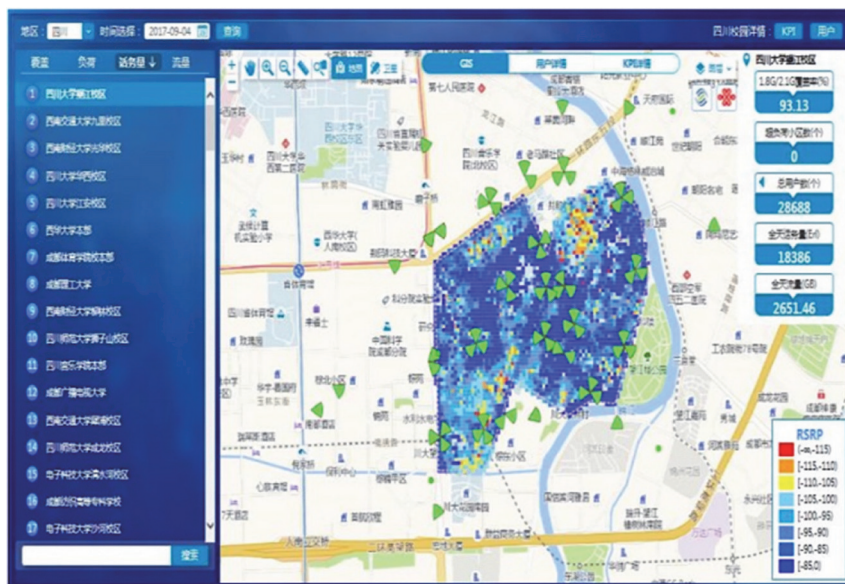


图1 电信公司RSRP信号覆盖地图

Fig. 1 RSRP signal coverage map of Telecom company

用不同的颜色表示出该 $\Delta x^2 \text{ m}^2$ 范围内的平均信号覆盖质量。因此,该马尔科夫决策问题^[7]的优化目标也可以表示成小区内不同信号覆盖等级的格子数量加权后进行平均,即

$$\max 4n_{\text{RSRP}_1} + 3n_{\text{RSRP}_2} + 2n_{\text{RSRP}_3} + n_{\text{RSRP}_4} - 2n_{\text{RSRP}_5} - 4n_{\text{RSRP}_6} - 8n_{\text{RSRP}_7}, \quad (1)$$

其中: n_{RSRP_k} 代表 RSRP 信号覆盖等级为 k 的方格数量, k 是从好到坏分成 7 个等级的 RSRP 取值范围,其目的是对覆盖较好的情况给予一个正向激励,对覆盖较差的情况给予负向激励。同时,环境观测值也可以抽象成小区内不同颜色的格子分布,可以当作一张灰度图进行处理。因此,整个过程也可以被描述为:根据当前时刻小区的灰度图特征,预测如何对小区内所属天面的角度进行调整,从而让整个小区的覆盖情况趋向于最好。

2 基于深度强化学习的自适应天馈系统的实现

2.1 MADDPG 算法

在处理天馈系统这种复杂的合作-竞争环境时,不妨引入多智能体深度确定性策略梯度(MADDPG)算法的 Actor-Critic 框架设计^[7-13]。MADDPG 很容易解决天馈系统优化过程中的非稳定环境问题。在 MADDPG 中,每个智能体都有自己的执行者 (Actor)网络与评价者 (Critic)网络。其中,Actor 网络中的算法和原始 DDPG 中相应算法差别不大。用于训练过程的 Critic 网络算法有较大改进。在 Critic 网络中,为了获取全局信息,用系统中所有智能体的观测值与行为作为参数传入网络。

除了小区内的多智能体问题,相邻小区之间也是合作-竞争共存的关系。在 MADDPG 中,Critic 网络负责集中式训练,需要区域内的全局信息;Actor 网络负责分布式执行,只需要智能体自己的环境观测数据。因此,Actor 与 Critic 网络可以传入不同的观测数据,不妨让 Critic 网络使用小区内的环境状态信息,让 Actor 网络不再局限于智能体所属的小区,而是使用智能体潜在覆盖范围内的状态信息。一个天面的潜在覆盖区域包含它在调整过程中所有可能覆盖的范围,以潜在覆盖区域作为 Actor 网络的输入,就可以让智能体执行决策时考虑到跨区域的信息。

在此,给出基于 MADDPG 自适应优化系统的流程总览,如图 2 所示。由于强化学习总是伴随着大量的探索,为了避免造成服务区信号质量下降,将系统流程划分为预训练与业务场景部署 2 个阶段。在预训练阶段,根据真实业务场景的地理空间环境,为每个小区建立了一个虚拟环境。在一个计算密集型的高算力平台上,基于该虚拟环境进行虚拟智能体的训练。此时,虚拟智能体可以将行为信息简单地同步,因此 MADDPG

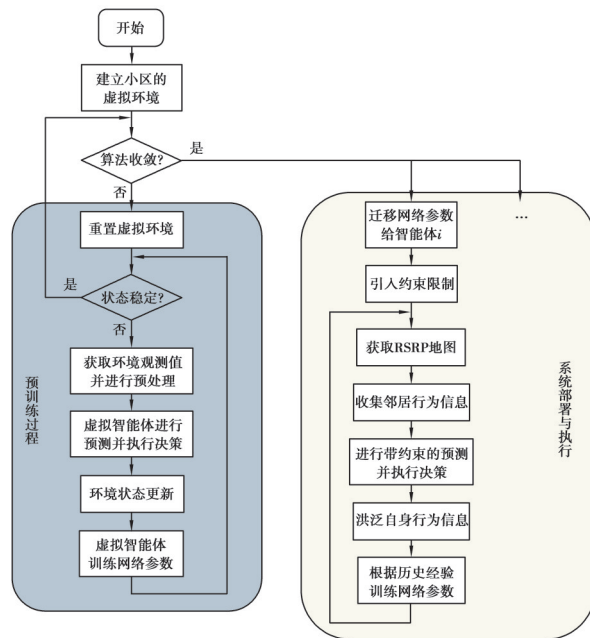


图 2 基于 MADDPG 的 5G 基站自适应天馈系统部署流程总览

Fig. 2 An overview of the deployment process of 5G base station adaptive antenna feed system based on MADDPG

的训练过程可以像单智能体算法一样,直接用一个共享的经验回放池更新所有智能体的 Actor 与 Critic 网络。整个过程可以是单线程枚举的。每次迭代中,首先枚举所有智能体,让它们获取自己的环境观测值并执行确定性行为预测。执行所有智能体的行为后,对环境的更新就进入下一个状态,并将本次交互信息放入全局的经验回放池中。之后再次枚举每个智能体,为它们随机取出一批经验数据进行训练,然后进入下一次迭代。待环境状态稳定后,判断算法是否收敛,若未收敛,则重置虚拟环境继续进行预训练。

2.2 环境归一化算法

在多智能体强化学习中,Distributed-Q^[16-17]、MA-DQN 以及 A3C 等算法要求在不同智能体中,将环境状态的维度与特征分布统一起来,这就是状态统一性问题。在研究的场景中,状态统一性问题源自潜在覆盖范围的形态和大小,取决于不同天面间的参数差异,尤其是铁塔站高度、天面水平辐射范围、垂直辐射范围、主瓣参数等,让天面的潜在覆盖区域呈现为不同弧长与半径的扇形,甚至某些部分可以为环形扇面。

MADDPG 允许每个智能体关注毫不相关的局部环境状态,使用完全独立的回报函数,并不要求对环境状态进行统一化处理。不过,为了加速整个算法的收敛速度,让虚拟环境下训练好的神经网络能够快速部署到业务场景中,依然需要对每个智能体的观测值进行归一化处理。

2.3 归一化采样

在天面 i 获取状态 s_i 的局部观测值 o_i 时,可能会因为基站参数、地理环境等的影响,得到一个与其他天面维度不同的张量。同时,基于地面的 RSRP 覆盖情况获得 o_i ,其分布情况并不能很好地对应天面的角度朝向。所以,需要从天面的角度出发对环境状态进行采样,以方位角为横轴、俯仰角为纵轴,从而保证状态分布的一致性。如图 3 所示,对于每一个天面,先找出它的整个可覆盖角度的区间范围,并对其垂直角度区间范围从上到下按照 1:2:3:4 的比例划分成 4 组,每一组均匀采样 5 个角度,得到 20 个纵轴刻度;对其水平角度区间,均匀地采样 20 个角度值,作为 20 个横轴刻度。这样,就获得了一个 20×20 的采样矩阵。然后按照采样的俯仰角和方位角来计算每个样本信号辐射在地面上的坐标,采集该点的 RSRP 值作为样本值。这样,就能保证所有观测值与智能体行为之间的映射关系具有普遍联系。

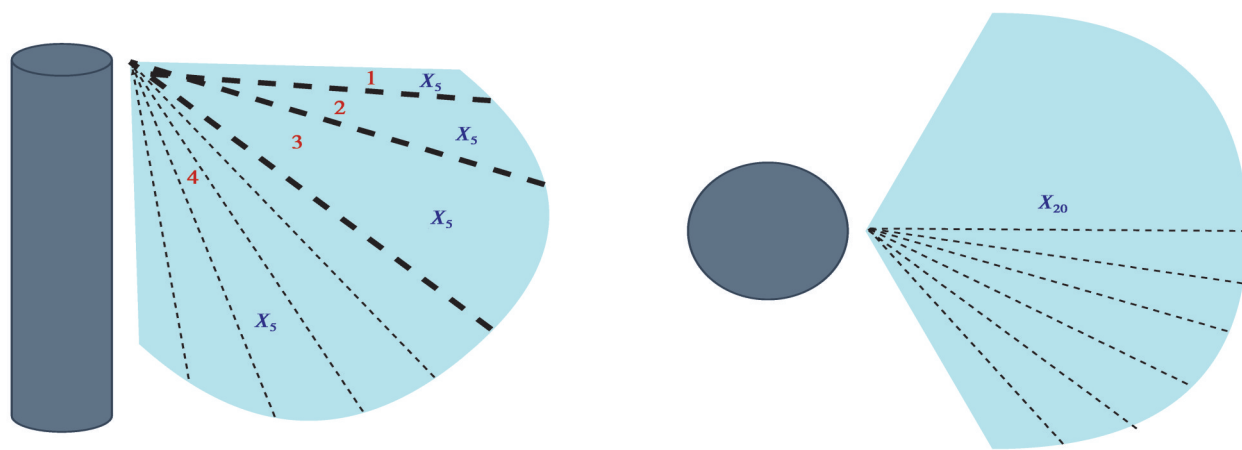


图3 归一化采样示意图

Fig. 3 Schematic diagram of normalized sampling

该采样过程有 2 个基本前提。首先,单个运营商的基站都是同一规格,各天面的可调范围基本一致。其主要影响因素是基站高度与预制倾角,因此可以通过分割角度的方式进行采样。另一方面,应在最远覆盖范围内进行采样,也就是说可调节俯仰角的上限应低于天面在可容忍衰减区间内覆盖到地面的角度。如果超过该角度,则天面的信号将在严重衰减后射到地面,此时可认为天面对地面设备无覆盖。

2.4 线性探测补点

在归一化采样时,根据天面不同角度到地面的投影来采集样本点,很容易导致部分样本点采到空值。在预训练过程中,无论是在虚拟环境还是真实业务场景下,RSRP 覆盖地图都可能存在大量无法获得采样的空

白值区域,归一化采样将不可避免地取得部分空值。于是,希望找到一种方法来填补这些空白样本。借鉴邻插值的思想,对无法采样的点使用二维线性探测的方式进行修补。

图 4 中给出了算法的具体流程。线性探测补点的方式保证了归一化采样的有效性。进一步地,可以在业务场景下对该补点方式进行优化。在业务场景下采样到空值,一般是由于在该区域,用户稀疏,于是不是每次 RSRP 地图更新期间内都有用户存在。所以,可以在一定程度上继承该区域的历史 RSRP 值,以增强线性探测补点的可靠性。具体做法是,接入站保留每个小区域的最近的一次非空 RSRP 平均值,在本轮更新信号覆盖地图时,如果在该区域无 RSRP 反馈信息,则先用线性探测补点方法预测出该区域的 RSRP_{predict},再找到历史数据 RSRP_{last},并记录本轮更新的更迭周期数 τ 。在本轮更新中,对该区域估计的填补值取为: $RSRP_{now} = (1 - \varepsilon/\tau) RSRP_{predict} + \varepsilon/\tau RSRP_{last}$, 其中, ε 是历史数据的权重因子, $\varepsilon < 1$ 。

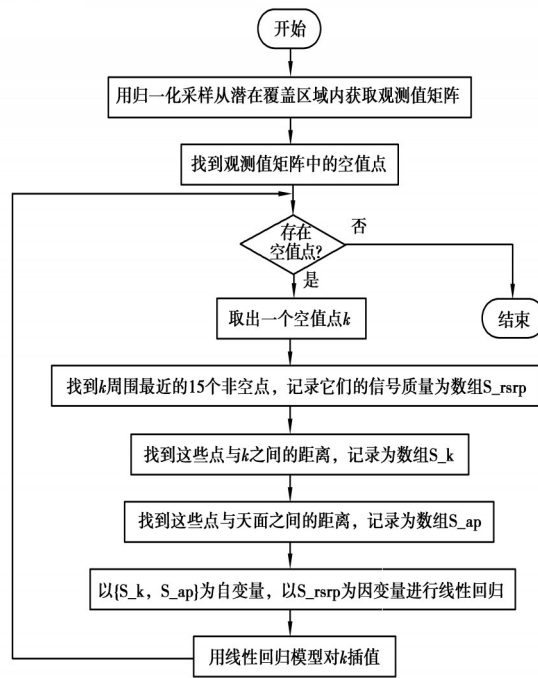


图 4 线性探测补点的代码流程

Fig. 4 Code flow of linear detection complement

2.5 业务场景部署

前面详细描述了整个系统的设计与算法训练过程,在此将给出系统部署到真实业务场景下的流程,以及基于带约束的马尔可夫决策问题(constraint markov decision process, CMDP)的站间通信过程。

为了保证系统的鲁棒性,提出先在虚拟环境预训练,再将网络参数迁移到业务场景的系统设计。在部署到业务场景之前,首先需要根据真实的小区环境,建立对应的虚拟环境模型。该虚拟环境建模需要考虑真实的基站布局、信道衰减以及空间环境影响,并提供合理的干扰仿真机制。基于真实环境,为该小区搭建专属的 MADDPG 网络,每一个基站的 3 个天面各拥有一张 Actor-Critic 网络,每个基站将在自己的移动边缘计算节点上同时部署 3 个天面的强化学习算法。强化学习的训练过程是先探索再学习。为了避免在部署前期出现过于激进的探索,可以借鉴迁移学习的思想,让算法先在已建好模的虚拟环境中训练,然后取收敛到趋近平稳状态的神经网络梯度值作为初始值,部署到真实环境中,让预训练后的网络指导现实中的天面调整,同时也在真实环境下执行下一步训练过程。

3 虚拟环境建模与训练

研究将天馈系统按地理位置与行政关系划分成多个小区,将单一小区内部的信号覆盖优化问题建模成多智能体的马尔可夫决策问题^[7]。在该问题背景下,智能体之间既有合作又有竞争,导致该博弈环境具有非

稳定性。使用多智能体深度确定性策略梯度算法,通过集中式训练、分布式执行的方式,在每个智能体中,将全局的行为决策作为环境的一部分。这种新环境具有马尔可夫假设^[7]的稳定性。为了避免智能体在真实业务场景下进行激进探索,提出先预训练再迁移的算法部署流程,并为原问题引入约束条件进行CMDP建模。在系统部署伊始,为对应的业务小区进行虚拟环境建模,并在该虚拟环境下进行MADDPG算法的预训练。

3.1 虚拟环境建模

这里,虚拟环境基于电信公司RSRP信号覆盖地图。虚拟环境包含一个区域,区域内有 M 个基站,每个基站上有3个天线,每个天线有其垂直覆盖角度、水平覆盖角度、俯仰角调整范围、方位角调整范围等参数,每个基站有其高度、位置等参数。本系统以区域为单位进行训练,区域内的每一个天线都是单独的智能体,因此智能体的数量有 $3 \times M$ 个。参考RSRP信号覆盖地图,以10 m为每个格子的边长,整个区域拥有 $xSize \times ySize$ 个格子,每个格子的颜色代表该100 m²内的平均RSRP值,从-115~-65均匀分成了7个级别。其中,红色、橙色、黄色代表-95以下的信号覆盖水平,信号较好的部分细分成了4种由深到浅的蓝色。

虚拟环境参考OpenAI Gym的格式设计接口,执行的行为是调整方位角 $[-10, 10]^\circ$ 与调整俯仰角 $[-10, 10]^\circ$,要达成的状态空间是重组成一维数据后的RSRP覆盖地图。可以通过“reset”接口来重置环境,通过“step”接口来传入行为(需要指定作为行动者的天线)并获得执行操作后的环境状态。

为了模拟真实的城市环境,使用瑞利衰减模型^[17]来计算天线辐射信号的传输增益。根据5G的空分特性^[15],相邻基站间只有1/8的几率会发生冲突。图5给出了用Tkinter软件绘制的虚拟环境可视化窗口界面。设置了300×300 m²的正方形虚拟小区,每个方格区域的边长 $\Delta x = 10$ m,虚拟RSRP信号覆盖地图的尺寸为 $xSize = ySize = 30$ m。在Tkinter的画板上,用黑色圆点标识出了小区内的所有基站。

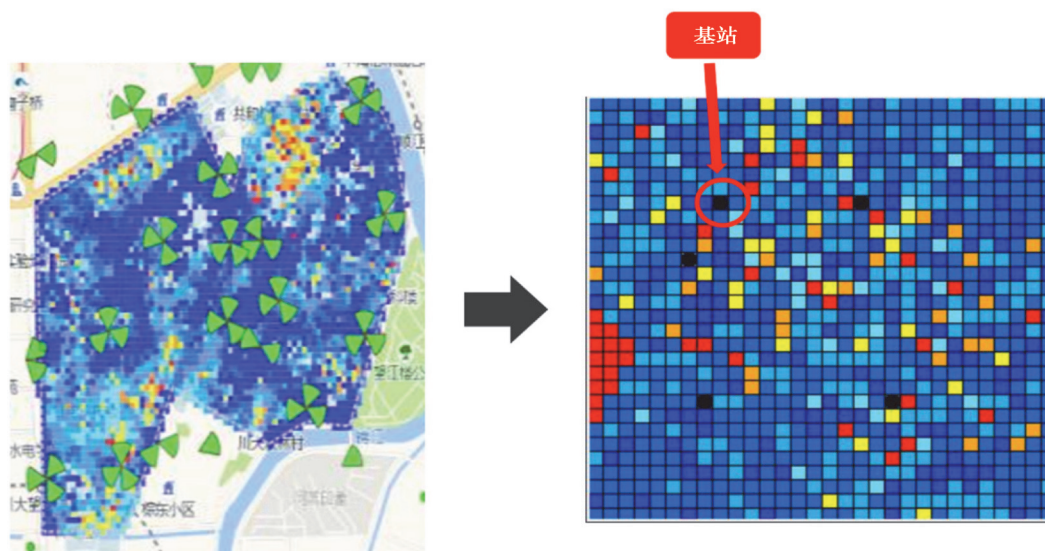


图5 虚拟环境搭建与基于Tkinter的可视化界面绘制

Fig. 5 Virtual environment construction and visual interface drawing based on Tkinter

3.2 实验训练实施及其结果

实验基于python 3.9下的pytorch 1.8环境,在macOS Big Sur 11.2(处理器Intel Core i9 9880H,内存16 GB 2667 MHz DDR4,显卡AMD Radeon Pro 5500M 8 GB)平台完成开发与可执行性调试,并在Windows 10 20H2(处理器Intel Xeon Gold 6133,内存256 GB DDR4 ECC,显卡GTX 1070ti 8 GB×2)平台下使用cuda_11.1进行模型训练。

在多智能体强化学习环境中,神经网络的尺寸相对较小,主要的时间开销来自于环境更新与数据预处理过程。尤其是CMDP求解中的策略采样算法,将在一次更新中花费大量时间,多次进行环境状态转移与线性探测补点操作。上述操作都是基于CPU的。选择40核80线程的双路Intel Xeon Gold 6133以提供较好的CPU计算能力。在显卡方面,考虑到这并非核心需求,因此选用了2张GTX 1070ti以提供16 GB的显存

容量。为了直观表现算法的有效性,可以通过可视化的方式呈现充分训练后算法在环境中的预测表现。如果取5 000次经历后的算法数据,即 50×10^4 次更新后的神经网络参数,对一个初始状态的环境进行10次更新,每次更新情况如图6所示。

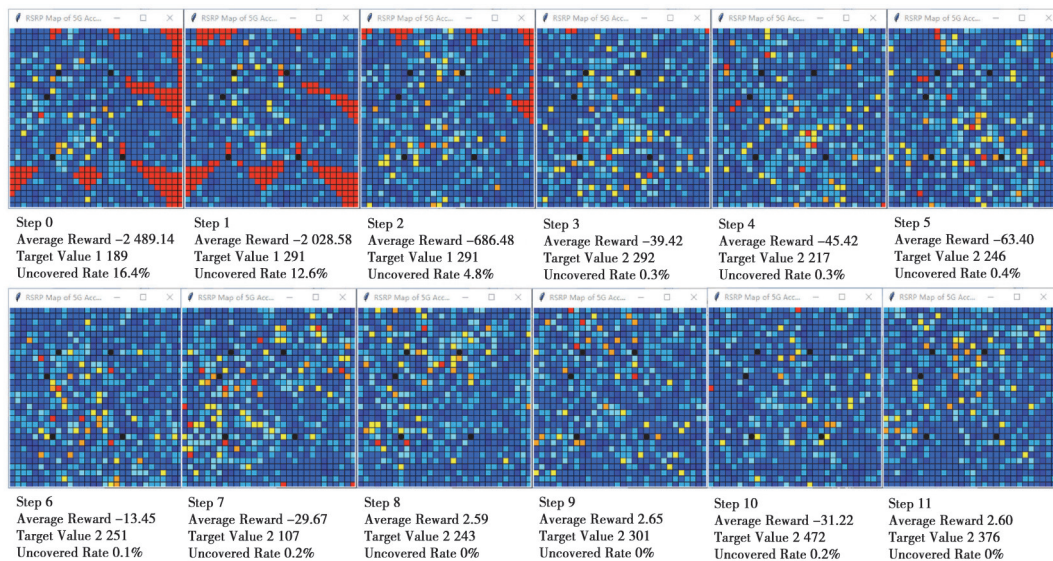


图6 充分训练的算法在10交互中环境的更新情况

Fig. 6 Environment update of the fully trained algorithm in ten interactions

图6给出了每一步更新时虚拟小区内各智能体的平均回报值、整个小区的目标函数值,以及为覆盖率。在该测试中,将约束收缩到了 $u=0$ (u 表示约束上限值),并允许智能体执行 $n_{\text{uncovered}}^{\text{now}} \leq n_{\text{uncovered}}^{\text{last}} + 1$ 的行为。分析该图发现,该系统经过足够的预训练后能快速将信号覆盖率调整至约束条件下,并在10次更新内实现令人满意的覆盖效果。

4 结 语

结合电信公司提供的基站信号覆盖地图中收集到的实时覆盖数据,研究了5G基站自适应天馈系统设计相关问题。通过改进MADDPG算法和修改Critic网络显著以降低整个算法的空间复杂度;为自适应天馈系统设计了一套基于人工智能算法的调度方案,从而实现真实业务场景下的CMDP策略采样。在系统设计中,为每个天面划分潜在的覆盖区域,将该区域的RSRP分布作为每一个状态下的观测值,研究了天馈系统在各观测值下的信号辐射方向调整策略。针对5G环境,在多智能体强化学习方面设计了一套环境观测值的归一化采样方案,让预训练出的算法模型可以迁移到真实业务场景下;提出线性探测补点的方法,从而避免观测值中出现空值点,保证预测的可靠性。所用智能调节算法能够有效避免智能体的激进探索,在大幅度提高算法收敛性能的同时,降低了算法的空间复杂性,可为整个系统在业务场景下的部署提供理论支撑。

参考文献

- [1] 周俊,权笑,马建辉. 5G无线优化面临的挑战及应对策略[J]. 电信科学, 2020, 36(1): 58-65.
Zhou J, Quan X, Ma J H. Challenge and strategy of 5G radio optimization[J]. Telecommunications Science, 2020, 36(1): 58-65. (in Chinese)
- [2] 赵国锋,陈婧,韩远兵,等. 5G移动通信网络关键技术综述[J]. 重庆邮电大学学报(自然科学版), 2015, 27(4): 441-452.
Zhao G F, Chen J, Han Y B, et al. Prospective network techniques for 5G mobile communication: a survey[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2015, 27(4): 441-452. (in Chinese)
- [3] 杜威,丁世飞. 多智能体强化学习综述[J]. 计算机科学, 2019, 46(8): 1-8.
Du W, Ding S F. Overview on multi-agent reinforcement learning[J]. Computer Science, 2019, 46(8): 1-8. (in Chinese)
- [4] 殷昌盛,杨若鹏,朱巍,等. 多智能体分层强化学习综述[J]. 智能系统学报, 2020, 15(4): 646-655.

- Yin C S, Yang R P, Zhu W, et al. A survey on multi-agent hierarchical reinforcement learning [J]. CAAI Transactions on Intelligent Systems, 2020, 15(4): 646-655. (in Chinese)
- [5] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning [J]. Nature, 2015, 518 (7540): 529-533.
- [6] Goodfellow I, Bengio Y, Courville A, et al. Deep learning [M]. US:MIT press Cambridge, 2016.
- [7] Littman M L. Markov games as a framework for multi-agent reinforcement learning [J]. Machine Learning Proceedings, 1994: 157-163.
- [8] Foerster J N, Assael Y M, de Freitas N, et al. Learning to communicate with deep multi-agent reinforcement learning [EB/OL]. 2016: arXiv: 1605.06676. <https://arxiv.org/abs/1605.06676>.
- [9] Hong Z W, Su S Y, Shann T Y, et al. A deep policy inference Q-network for multi-agent systems [EB/OL]. 2017: arXiv: 1712.07893. <https://arxiv.org/abs/1712.07893>.
- [10] Hessel M, Modayil J, Van Hasselt H, et al. Rainbow: combining improvements in deep reinforcement learning [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1): 3215-3222.
- [11] Foerster J, Nardelli N, Farquhar G, et al. Stabilising experience replay for deep multi-agent reinforcement learning [C]// Proceedings of the 34th International Conference on Machine Learning - Volume 70. August 6 - 11, 2017, Sydney, NSW, Australia. New York: ACM, 2017: 1146 - 1155.
- [12] Babaeizadeh M, Frosio I, Tyree S, et al. GA3C: GPU-based A3C for deep reinforcement learning [EB/OL]. 2016: arXiv: 1611.06256. <https://arxiv.org/abs/1611.06256>.
- [13] Abbas N, Zhang Y, Taherkordi A, et al. Mobile edge computing: a survey [J]. IEEE Internet of Things Journal, 2018, 5(1): 450-465.
- [14] 高松涛, 程日涛, 邓安达. 5G天馈系统下倾角设置原则研究 [C]//5G网络创新研讨会(2020)论文集. 北京:移动通信, 2020: 272-275.
Gao S T, Cheng R T, Deng A D. Research on setting principle of dip angle of 5G antenna feed system [C]//5G Network Innovation Seminar (2020). Beijing, China: Mobile Communications, 2020: 272-275. (in Chinese)
- [15] Gupta A, Jha R K. A survey of 5G network: architecture and emerging technologies [J]. IEEE Access, 2015, 3: 1206-1232.
- [16] Galindo-Serrano A, Giupponi L. Distributed Q-learning for aggregated interference control in cognitive radio networks [J]. IEEE Transactions on Vehicular Technology, 2010, 59(4): 1823-1834.
- [17] Sklar B. Rayleigh fading channels in mobile digital communication systems Part II: Mitigation [J]. IEEE Communications Magazine, 1997, 35(7): 102-109.

(编辑 侯 湘)