

doi:10.13682/j.issn.2095-6533.2025.06.002

基于深度强化学习的矿井5G优化方案

梁沫^{1,2}, 王军选³

(1. 中煤科工集团北京华宇工程有限公司西安分公司, 陕西 西安 710000; 2. 中国煤炭建设协会勘察设计委员会
信息化专业部, 北京 100013; 3. 西安邮电大学通信与信息工程学院, 陕西 西安 710121)

摘要: 针对煤矿井下复杂环境中第五代移动通信技术(5th-Generation Mobile Communication Technology, 5G)网络覆盖性能优化的问题, 提出一种基于深度强化学习的矿井5G优化方案。面向10 km主运巷道场景, 综合考虑巷道截面尺寸、壁面粗糙度、设备遮挡等多重传输损耗因素, 建立融合视距/非视距路径损耗模型与粗糙度衰减因子的信号传播数学模型。将深度Q网络作为价值函数近似器的强化学习智能体, 并通过基站部署与发射功率将在线优化转化为多目标决策问题, 以最小基站数量实现覆盖率最大化。采用动态功率调整机制, 以实时优化基站发射功率, 从而适应局部信号衰减的突发变化。实验结果表明, 该方案可以实现95%以上的覆盖率, 相较于传统静态方案减少了28%基站部署, 显著提升了井下5G网络覆盖性能, 并能够降低部署成本与运行功耗。

关键词: 井下移动网络; 强化学习; 深度Q网络; 网络布局策略; 路径损耗模型

中图分类号: TN929.4; TP18

文献标志码: A

文章编号: 2095-6533(2025)06-0011-10

Optimization scheme for mine 5G based on deep reinforcement learning

LIANG Mo^{1,2}, WANG Junxuan³

(1. China Coal Science Beijing Huayu Engineering Co., Ltd. Xi'an Branch, Xi'an 710000, China;
2. China Coal Construction Association Survey and Design Committee Information Technology Department,
Beijing 100013, China; 3. School of Communications and Information Engineering, Xi'an University of Posts
and Telecommunications, Xi'an 710121)

Abstract: Aiming at the optimization of 5G network coverage performance in the complex environment of coal mine underground, a 5G optimization scheme based on deep reinforcement learning is proposed. For the 10 km main transport roadway scenario, multiple transmission loss factors such as roadway cross-section size, wall roughness, and equipment occlusion are comprehensively considered, and a signal propagation mathematical model integrating the line-of-sight/non-line-of-sight path loss model and the roughness attenuation factor is established. A deep Q-network is adopted as the value-function approximator for the learning agent, transforming the joint online optimization of base-station placement and transmit power into a multi-objective decision problem that maximizes the coverage while minimizing the number of base stations. Adopt a dynamic power-adjustment mechanism, enabling real-time adaptation to abrupt local signal degradations. Experimental results confirm that the scheme achieves a coverage exceeding 95%, while reducing the number of deployed base stations by 28% compared with a conventional static layout, thereby markedly enhancing the underground 5G coverage and lowering the deployment costs and operational power consumption.

收稿日期: 2025-03-26

基金项目: 国家自然科学基金项目(62071377)

引文格式: 梁沫, 王军选. 基于深度强化学习的矿井5G优化方案[J]. 西安邮电大学学报, 2025, 30(6): 11-20.

LIANG M, WANG J X. Optimization scheme for mine 5G based on deep reinforcement learning[J]. Journal of Xi'an University of Posts and Telecommunications, 2025, 30(6): 11-20.

Keywords: mobile network under the mine; reinforcement learning; deep Q-network; network deployment strategy; path loss model

在能源安全战略与“双碳”目标驱动下,煤炭行业正加速向智能化转型,以破解井工煤矿在高危环境下事故频发、人力依赖度高的发展困局^[1-2]。少人化、无人化的智能化生产成为提升安全水平的重要突破口,同时,构建智能矿山需要依赖强大的底层通信技术支撑。因此,第五代移动通信技术(5th-Generation Mobile Communication Technology, 5G)凭借大带宽、低时延、高可靠的特性,成为支撑智能矿山关键场景的核心通信基础技术^[3],引起学术界与产业界的广泛关注。

目前,中国矿业大学(北京)与中国通信标准化协会已分别从能源和通信行业视角,完成矿用5G通信系统标准的制定工作。在系统架构研究领域,研究者们围绕智慧矿山系统架构与矿用5G装备远程控制技术展开探索,提出了提升发射功率阈值方案并构建5G智能综采远控系统。然而,煤矿井下复杂的巷道结构、金属设备密集分布以及严苛的防爆安全要求,导致传统地面5G部署方案难以直接应用^[4]。同时,井下特有的信号衰减模型差异、动态遮挡效应和防爆安全限制,使得现有研究在网络覆盖性能优化方面暴露出显著短板:理论模型多基于简化场景假设,难以精准刻画井下复杂传播环境;部署策略普遍采用静态规划,无法适应工作面迁移等动态变化;功率分配机制缺乏自适应调节能力,难以应对局部信号衰减突变问题^[5]。

以上问题在井下复杂环境中相互制约,阻碍了5G网络的高效部署与稳定覆盖。现有理论模型依赖简化假设,使刻画井下真实传播环境变得困难,导致预测偏差和覆盖盲区与干扰区。静态部署方案无法响应工作面迁移等动态变化,造成网络覆盖失效和关键业务中断。功率分配机制反应迟钝及叠加防爆安全的严格功率限制,难以缓解突发遮挡或地质变化引发的信号突降,无法保障高要求应用的稳定连接。因此,已有研究从多维度对井下5G网络覆盖进行了探索^[6-9],分别在干扰建模、频段优选、功率安全及覆盖影响因素等方面共同推进了矿用5G网络覆盖优化的研究,为后续研究奠定了重要的基础。文献[10]构建了面向5G无线网络规划与优化的人工智能(Artificial Intelligence, AI)方法体系,实现对5G天线方位角的动态、精准与自动化调整,以优化网络覆盖质量与用户体验;文献[11]首创地提出了基于波束空间特征与机器学习融合

的5G天线方位角反演模型,实现了天线方位角的智能动态调整;而文献[12]在系统分析煤矿智能化发展现状与传统通信局限的基础上,提出了5G技术驱动的全场景应用框架。与此同时,矿山智能化场景研究持续深化,文献[13]基于贝叶斯-多层感知器协同建模构建了5G网络动态优化系统。但是,现有研究仍存在显著局限,一方面,针对矿井环境的动态特性缺乏动态响应机制;另一方面,在基站部署、频谱分配及功率控制等多维资源协同优化层面尚未形成系统性解决方案^[14-15]。这些不足导致当前矿用5G网络难以满足智能矿山对高可靠性、强适应性的严苛需求。

针对上述问题,基于深度强化学习,拟提出一种基于深度强化学习的矿井5G优化方案。通过构建融合视距(Line-of-Sight, LoS)/非视距(Non-Line-of-Sight, NLoS)路径损耗模型与粗糙度衰减因子的信号传播数学模型,将基站部署问题转化为多目标优化问题。同时,利用深度Q网络(Deep Q-Network, DQN)算法设计动态功率调整机制,实现对井下复杂环境的实时感知与自适应调节,以期提升矿井5G网络覆盖性能与资源利用效率。

1 矿下5G基站覆盖范围影响因素

通信信号以电磁波形式在空间传播时,会因波速扩散效应出现能量损失^[16]。在非理想条件下,散射和折射等情况还会进一步导致能量损耗,且具有较大随机性,从而影响通信信号的覆盖范围。与地面、露天煤矿等较为空旷的环境相比,煤矿有空间受限、大型金属设备多等特点^[17],使得井下矿用5G信号的覆盖性能更差。因此,根据煤矿的环境特点和影响因素,对矿用5G的覆盖性能进行差异化分析。

1.1 系统模型

考虑一种煤矿井下5G通信的巷道覆盖系统模型,该模型由线性主运巷道、部署于巷道壁面的多个基站节点,以及随机分布的用户终端组成。在该系统模型中,主运巷道横截面为固定矩形结构,壁面粗糙且散布各类机电设备,易引入多径、阴影遮挡和NLoS效应。为真实反映信道特性,建立融合LoS/NLoS路径损耗模型与粗糙度衰减因子的综合传播模型。在此基础上,将最少基站数下最大化覆盖率的问题转化为强化学习中的状态-动作决策

过程,再利用 DQN 算法结合动态发射功率控制,实现在最小基站数量下最大化系统覆盖率的部署方案。

1.2 工作频段

在现代无线通信领域,频谱资源被视为至关重要的基础,工作频段的选择对通信系统的性能有着决定性的影响。不同的频段具有独特的传播特性和损耗表现,直接关系到信号的覆盖范围、穿透能力以及数据传输的稳定性。在自由空间中,路径损耗 $L^{[18]}$ 是衡量信号在传输过程中能量损失的一个关键指标,其表达式为

$$L=32.4+20\lg f+20\lg D \quad (1)$$

式中: f 为信号频率,单位是 MHz; D 为传输距离,单位是 km。

信号的路径损耗与传输距离和频段大小呈正比关系。具体来说,当传输距离 D 固定时,频率 f 越高,路径损耗 L 也随之增大;同样,当频率 f 固定时,传输距离 D 越远,路径损耗 L 也会相应增加。这意味着,在同等条件下,高频信号的路径损耗大于低频信号。在设计无线通信系统时,合理选择工作频段是优化通信性能的关键因素之一,特别是在需要长距离传输或良好覆盖的应用场景中,如农村地区、大型工业厂区以及地下矿井等,低频段的优势尤为显著。

在无线通信的实际应用场景中,除了自由空间的路径损耗外,还需综合考虑非理想条件下的多种影响因素。这些因素对信号的传播特性有着显著的影响,特别是在不同频段的表现更为明显。一方面,高频信号的绕射能力相对较弱,意味着当信号遇到建筑物、山脉或其他障碍物时,其能够绕过这些障碍物继续传播的能力较差,从而导致在障碍物后方的信号覆盖区域出现阴影效应,信号强度急剧下降,通信质量受到严重影响。另一方面,空气中存在多种介质,如水蒸气、二氧化碳等,这些介质会对电磁波产生吸收作用,进而导致额外的损耗。而高频信号由于其波长较短,与空气中的分子或其他微小颗粒的相互作用更为频繁,因此更容易受到这些空间介质的影响,衰减更为明显。

在矿下环境中,信号的传输特性与其他环境类似,但由于矿下环境相对恶劣,存在更多的障碍物和复杂的地质条件,高频信号的衰减程度往往更大。例如,矿井中的巷道狭窄且纵横交错,煤壁凹凸不平,这些都会对信号的传播造成严重的阻碍和反射,使得信号在传输过程中的能量损失更为严重。此外,矿下环境中还存在着大量的粉尘、水分等,这些介质会进一步吸收和散射电磁波,加剧信号的衰减。在煤矿环境下,工作频段越低,信号的

波长越长,其绕射能力和穿透能力就越强,能够更有效地克服矿下环境中的各种障碍物和介质的影响,从而使得信号衰减越小,矿用 5G 的覆盖能力也就越强。因此,在设计矿用 5G 通信系统时,合理选择较低的工作频段对于提高信号的覆盖范围和通信质量至关重要。

1.3 矿下环境因素

在矿井通信环境中,信号的衰减特性不仅受频率影响,还受巷道几何尺寸、壁面/顶底板材料及表面粗糙度的影响^[19]。这些因素共同决定了信号在矿井环境中的衰减程度,进而影响通信系统的覆盖范围和信号质量。信号强度最低次水平极化衰减系数 α_{Eh} 和垂直极化系数 α_{Ev} ^[20] 的表达式分别为

$$\alpha_{\text{Eh}}=4.343\lambda^2\left(\frac{\epsilon_1}{W^3\sqrt{\epsilon_1-1}}+\frac{1}{H^3\sqrt{\epsilon_2-1}}\right) \quad (2)$$

$$\alpha_{\text{Ev}}=4.343\lambda^2\left(\frac{1}{W^3\sqrt{\epsilon_1-1}}+\frac{\epsilon_2}{H^3\sqrt{\epsilon_2-1}}\right) \quad (3)$$

式中: λ 为电磁波波长; ϵ_1 和 ϵ_2 分别为巷道壁和顶底板的相对介电常数; W 为巷道宽度; H 为巷道高度。

为便于后续计算,定义平均极化衰减系数为

$$\alpha_{\text{p}}=\frac{\alpha_{\text{Eh}}+\alpha_{\text{Ev}}}{2} \quad (4)$$

式中: α_{p} 的物理意义是在巷道中任意极化状态下信号沿传播方向的线性衰减率。

电磁波在狭窄巷道传输时会反复反射,巷道壁粗糙度会影响信号入射角和发射角,进而影响信号能量损耗^[21]。同时,不同粗糙度的巷道壁对信号极化特性改变程度不同。因此,类比光线传播特性研究电磁波反射,得出多次折射后的综合损耗因子 l 和衰减损耗 L_r ^[22] 的表达式分别为

$$l=\exp\left(-\pi^2\lambda\left(\frac{h_1^2}{W^4}+\frac{h_2^2}{H^4}\right)\right) \quad (5)$$

$$L_r=-10\lg l=4.343\pi^2\lambda\left(\frac{h_1^2}{W^4}+\frac{h_2^2}{H^4}\right) \quad (6)$$

式中: h_1 和 h_2 分别为巷道壁和顶底板的粗糙度均方根。由于 L_r 与距离无关,可视为一次性散射损耗并在总路径损耗中统一叠加。

考虑到收发两端的位置及传播路径上的遮挡物,信号传播方式可分为完全视距传输、近视距传输和非视距传输 3 种。完全视距传输是收发两端无遮挡物,电磁波直线传播,信号能量损失小,覆盖距离最大;近视距传输是存在部分遮挡物,信号通过衍射和透射到达接收端,但有能量损耗,井下环境复杂,暂不考虑此方式;非视距传输是存在完全遮挡物,信号通过反射和折射到达接收端,但能量损耗较小。

第三代合作伙伴计划(3rd Generation Partnership Project, 3GPP)未针对煤矿特殊场景建模,但可借鉴其定义的室内工厂场景。设备密集时,视距传输路径损耗(Loss of Line of Sight, LLoS)和非视距传输路径损耗(Loss of Non Line of Sight LNLLoS)的表达式^[23]分别为

$$L_{\text{LoS}} = 31.84 + 21.50 \lg d_{3\text{D}} + 19.00 \lg f_c \quad (7)$$

$$L_{\text{NLLoS}} = \max(L_{\text{IF}}, L_{\text{LoS}}) \quad (8)$$

式中: $d_{3\text{D}}$ 为收发设备的3D距离; f_c 为中心频率;室内工厂场景下 L_{IF} 的计算公式为

$$L_{\text{IF}} = 33.63 + 21.9 \lg d_{3\text{D}} + 20 \lg f_c \quad (9)$$

同等条件下,LNLLoS大于LLoS。煤矿井下遮挡物多,非视距传输概率大,信号衰减严重,覆盖距离减小。

通过将传播损耗 $L_{\text{LoS/NLLoS}}$ 、粗糙度散射损耗系数 A_m 和沿程极化衰减系数 $\alpha_p/d_{3\text{D}}$ 叠加,可得距离 $d_{3\text{D}}$ 下的总路径损耗为

$$L_{\text{tot}}(d_{3\text{D}}) = \begin{cases} L_{\text{LoS}}(d_{3\text{D}}) + A_m + \alpha_p d_{3\text{D}} \\ L_{\text{NLLoS}}(d_{3\text{D}}) + A_m + \alpha_p d_{3\text{D}} \end{cases} \quad (10)$$

2 基于强化学习的矿井5G覆盖设计

随着矿业数字化与智能化转型,矿井5G通信系统成为关键支撑。为保障其稳定与高效,提出基于强化学习的矿井5G覆盖优化设计,聚焦DQN算法,通过智能体与环境交互获取最优策略,以实现动态调整5G基站位置与参数,从而提升覆盖效果。考虑到井下设备算力有限且基站部署、功率调节动作本质离散,轻量级策略DQN算法能在离线仿真中高效训练并周期下发权重,相较于需同策略更新的策略梯度算法,其计算负担更低、部署更便捷,从而成为最优选型。

2.1 DQN算法原理

DQN将经典Q-Learning的增量式价值更新机制与深度神经网络近似器 $Q_\theta(s, a)$ 结合,在高维状态和离散动作场景中仍能有效逼近最优动作价值^[24]。核心迭代的表达式为

$$Q_\theta(s_{t+1}, a_{t+1}) \leftarrow Q_\theta(s_t, a_t) +$$

$$\partial [R_{t+1} + \gamma \max_a Q_\theta(s_{t+1}, a) - Q_\theta(s_t, a_t)] \quad (11)$$

式中: θ 为在线Q网络的参数向量; ∂ 为学习率; γ 为折扣因子; s_t 为 t 时刻的状态; a_t 为 t 时刻选择的动作; R_{t+1} 为即时奖励。

为克服Q-Learning在函数逼近时的非稳定性,DQN引入经验回放与目标网络两项关键机制。经验回放持续缓存智能体与环境交互得到的四元组 (s, a, R, s') ,训练阶段随机抽取小批量样本以打破

时间相关性,并提升样本利用率,目标网络 \hat{Q}_θ^- 则以滞后参数向量 θ^- 计算目标项 $R_t + \gamma \max_a Q_\theta^-(s', a)$,并周期性地从主网络复制参数,从而缓解估计震荡并稳定收敛。

2.2 系统建模

将坑道总长度 M 按1m间隔离散为 M 个采样点,记距坑道起点 i 处的位置为 $x_i = i (i=0, 1, 2, \dots, M-1)$ 。假设第 j 个基站发射功率为 P_j ,天线增益已包含在 P_j 内,综合链路裕量为 I_t ,则位置 x_i 处(与第 j 个基站的距离为 $d_{j,i}$)的接收功率的表达式为

$$P_{\text{rx}}(d_{j,i}) = P_j - L_{\text{tot}}(d_{j,i}) - I_t \quad (12)$$

式中: $P_{\text{rx}} \geq P_{\text{th}}$ 时判定为已覆盖,阈值 P_{th} 可由业务服务质量(Quality of Service, QoS)要求得到。

$c[i]$ 表示距坑道起点 i 位置处的覆盖状态,当 $c[i]=1$ 表示覆盖,当 $c[i]=0$ 表示没有覆盖,具体可表示为

$$c[i] = \begin{cases} 1, P_{\text{rx}}(x_i) \geq P_{\text{th}} \\ 0, P_{\text{rx}}(x_i) < P_{\text{th}} \end{cases} \quad (13)$$

总体覆盖率 C 的计算公式为

$$C = \frac{\sum_{i=0}^{M-1} c[i]}{M} \times 100\% \quad (14)$$

最终的优化目标是实现矿道内5G基站总体覆盖率最大化,同时尽量减少基站的部署数量和基站功率调整次数。此时,优化问题可以表示为

$$\max C - \lambda \cdot N \quad (15)$$

$$\text{s. t. } C_1: P_j \leq P_{\text{max}}, j=1, \dots, N,$$

$$C_2: 0 \leq y_j \leq M, j=1, \dots, N,$$

$$C_3: N \leq N_{\text{max}}.$$

式中: N 表示部署基站的数量;第 j 个基站的位置为 y_j ; $\lambda > 0$ 用于在覆盖性能与部署成本之间进行权衡; C_1 为基站最大发射功率 P_{max} 的约束,用于限制基站的发射功率不超过最大值;约束 C_2 限制基站的位置不得超过矿道的总长度; C_3 为基站最大部署数量 N_{max} 的约束。

所优化的问题具有非凸特性,使得寻找最优解变得困难,而强化学习可通过智能体与环境的交互解决这类问题^[25-26]。该优化目标可等价映射为智能体的即时奖励函数,以确保与强化学习框架一致。智能体在每步观察状态向量 s_t (含当前覆盖率、基站数量占比等信息),从离散动作集 \mathbf{A} 中选择动作 a_t (如是否在候选位置部署基站及调整其发射功率),与环境交互得到 t 时刻奖励 R_t 并转移至新状态 s_{t+1} ,其中 s_{t+1} 为在执行 a_t 后由环境反馈得到的下一时间步的状态。这一过程可以用马尔可夫决策过程描述,通过马尔可夫决策过程和神经网络训练,深度强化学习能够找到

最优的决策策略,从而最大化系统的总体覆盖率。强化学习框架示意图如图 1 所示。

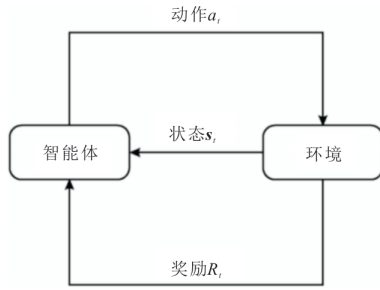


图 1 强化学习框架示意图

将负责制定基站部署和功率调整策略的决策核心视为一个智能体,其本质是一个 DQN 驱动的强化学习模型。为确保功率动态调整过程与矿井防爆安全标准相衔接,在设计中引入了触发阈值与安全冗余机制。一是,功率调整的触发条件基于接收功率与覆盖率双重判据,当局部接收功率低于 -90 dBm 或系统覆盖率在一个采样周期内下降超过 5% 时,触发功率上调动作;当接收功率高于 -80 dBm 且覆盖率稳定超过 95% 时,触发功率下调动作。二是,设定基站发射功率硬约束不超过 40 dBm,并在强化学习动作空间中引入安全冗余区间,即智能体可调节的功率范围限定为 $30 \sim 38$ dBm,以保证任何情况下功率不会突破防爆极限。通过上述协同机制,动态功率调整既能适应井下突发信道衰减,又能满足矿井防爆安全的严格要求。环境设定为模拟煤矿井下物理空间和通信特性的虚拟系统,负责提供状态反馈和奖励信号,智能体与环境产生交互。定义以下五元组 (S, A, P, R, γ) 。

1) 状态空间 S 。将五维归一化在 t 时刻的状态向量定义为

$$\mathbf{s}_t = (s_1, s_2, s_3, s_4, s_5) = \left(\frac{x_i}{M}, C, \frac{N}{N_{\max}}, 1 - \frac{N}{N_{\max}}, \min_i \frac{|x_i - y_j|}{M} \right) \quad (16)$$

式中: $s_1 = x_i/M$ 为当前位置归一化; $s_2 = C$ 为当前总体覆盖率; $s_3 = N/N_{\max}$ 为当前基站数量占比; $s_4 = 1 - s_3$ 为剩余部署预算; $s_5 = \min(x_i - y_j | M)$ 为最近基站归一化距离。

2) 动作空间 A 。动作空间 A 可定义为

$$\mathbf{A} = \{a_0, a_1(x_i), a_2(j), a_3(j)\} \quad (17)$$

式中: a_0 为保持现状,不进行基站部署与功率调整; $a_1(x_i)$ 为在位置 x_i 新部署一个发射功率初始设定为 P_0 的新基站; $a_2(j)$ 为第 j 个基站降低功率 $P_j \rightarrow \max(P_j - \Delta P, P_{\min})$; $a_3(j)$ 为第 j 个基站提高功率 $P_j \leftarrow \min(P_j + \Delta P, P_{\max})$ 。为了让 Q 网络结构在训练前就确认维度,离散动作集合 A 的规模固定为

$$|\mathbf{A}| = 1 + M + 2N_{\max} \quad (18)$$

同时,在训练或推理时,如果某些动作暂时“无效”,例如,对应网格已部署基站,或指定基站当前已达功率上限,就对这些动作打掩码,强制 Q 值不参与 Argmax 选择。

3) 状态转移概率 P 。执行动作后,环境依据式(10)计算 $L_{\text{tot}}(d_{j,i})$,随后通过式(12)得到各位置的接收功率 P_{rx} ,并按式(13)更新 $c[i]$,随后返回下一个状态 \mathbf{s}_{t+1} 。

4) 奖励函数 R 。为兼顾覆盖提升与部署/能耗成本两类目标,总奖励可定义为

$$R_{\text{tot}} = r_{\text{cov},t} - J_{\text{dep},t} - J_{\text{pow},t} - J_{\text{scl},t} \quad (19)$$

式中: $r_{\text{cov},t} = \alpha_{\text{cov}}(C_t - C_{t-1})$ 为 t 时刻覆盖增益的增量式奖励; α_{cov} 为覆盖增益缩放系数; $J_{\text{dep},t}$ 为 t 时刻部署数量惩罚,当执行基站部署动作时扣除固定成本; $J_{\text{pow},t}$ 为功率调整惩罚,当执行调整功率动作时给予惩罚; $J_{\text{scl},t}$ 为规模惩罚,随已部署基站占比增加而线性增长。

3 仿真结果及分析

3.1 仿真环境设置

构建一个基于强化学习的矿道 5G 信号覆盖优化仿真环境,详细设定矿道的物理特性、通信参数以及与强化学习相关的配置,具体参数如表 1 所示。表 1 中的参数涵盖了矿道的基本物理尺寸、信号传播特性、基站部署限制以及功率调整规则等。这些参数能够准确反映矿道环境下的通信状况,在仿真中模拟出接近真实的通信场景,从而为优化矿道 5G 信号覆盖策略提供有力支持。

表 1 参数说明和默认值设置

参数	数值
巷道长度 M/m	10 000
巷道宽度 W/m	2.5
巷道高度 H/m	2.5
粗糙度均方根 h_1, h_2	0.15
工作频段 f/GHz	0.9
初始发射功率 P_0/dBm	30
接收功率阈值 P_{th}/dBm	-90
视距传输概率 P_{LoS}	0.3
空间分辨率 $\Delta x/\text{m}$	100
最大基站数量 N_{\max}	20
功率调整大小 $\Delta P/\text{dB}$	5 dB
基站最大发射功率 P_{\max}/dB	40 dB
防爆安全功率上限 $P_{\text{safe}}/\text{dBm}$	40
功率调整冗余区间 $\Delta P_{\text{buf}}/\text{dBm}$	≤ 38
功率调整触发阈值(上调) P_{up}/dBm	< -90
功率调整触发阈值(下调) $P_{\text{down}}/\text{dBm}$	> -80

3.2 结果分析

为评估奖励函数中权重配置对智能体学习过程及最终性能的影响,对系统进行敏感性测试。测试主要对比3组权重组合(覆盖奖励,部署惩罚,功率调整惩罚,规模惩罚):(1,10,2,5)、(1,10,5,5)和(1,5,2,5),并从学习收敛性与覆盖性能两个方面进行分析。3组不同权重的训练奖励值的收敛性如图2所示。

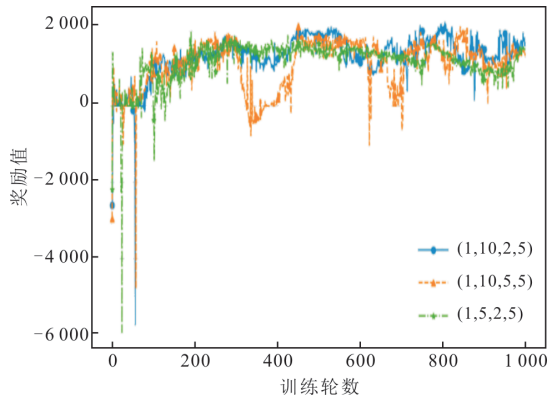


图2 3组不同权重的训练奖励值的收敛性

由图2可以看出,3组权重均能使智能体成功学习,奖励值随训练轮次增加逐渐从波动走向稳定。权重组合(1,10,2,5)与权重组合(1,5,2,5)的收敛轨迹高度重合,均在约600轮后达到稳定,在最终奖励表现上,(1,10,2,5)突破1800,而(1,5,2,5)约为1600,存在明显差距。权重组合(1,10,5,5)虽能收敛,但其稳定后的奖励值显著较低,约为1500,表明降低部署惩罚权重反而导致智能体获得的整体收益下降。

为了验证所提方案的性能,对其3组不同权重下可实现的覆盖率进行分析,具体如图3所示。

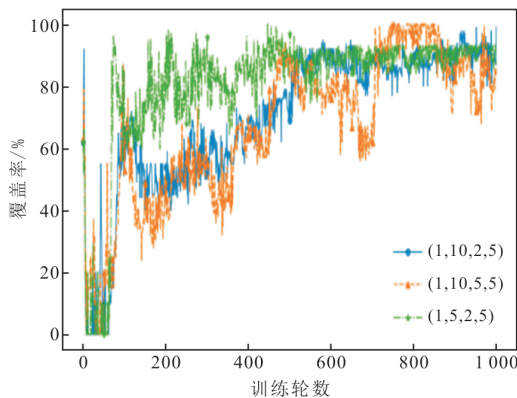


图3 3组不同权重下的覆盖率

由图3可知,权重组合(1,10,2,5)表现出最优且较为稳定的性能,其覆盖率随训练快速提升并最终稳定在98%的高水平。这表明该配置能有效

引导智能体在严格控制部署成本的同时,通过优化决策实现极高的网络覆盖。权重组合(1,10,5,5)取得了相近的覆盖率,但稳定性略逊于权重组合(1,10,2,5)。权重组合(1,5,2,5)的最终覆盖率可达93%,这主要是由于较低的部署惩罚使智能体倾向于部署更多基站,从而提升了覆盖范围,无法实现覆盖与成本之间的有效平衡。综合分析表明,部署惩罚系数是影响方案性能的核心参数。选定的权重组合(1,10,2,5)被证明为最优配置,其能使智能体学习到在严格成本约束下实现高性能的卓越方案,最终在保障高达98%覆盖率的同时,确保学习过程的高效与稳定。该敏感性测试充分验证了本文权重选择的合理性与科学性。

为确保仿真模型的可靠性并评估不同空间尺度对优化结果的影响,系统性地对比了所提方案在20 m、50 m及100 m这3种分辨率下的训练过程。通过重点考察奖励值的收敛特性与稳定性,为分辨率这一关键参数的设置提供实证依据,从而在计算复杂度与仿真精度之间寻求最佳平衡。具体的训练奖励值的收敛性分别如图4~图6所示。

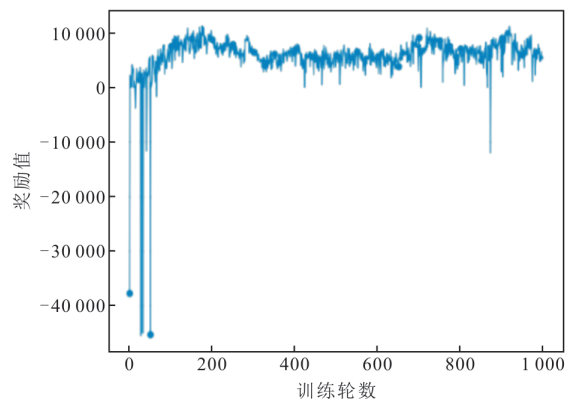


图4 分辨率为20 m的训练奖励值的收敛性

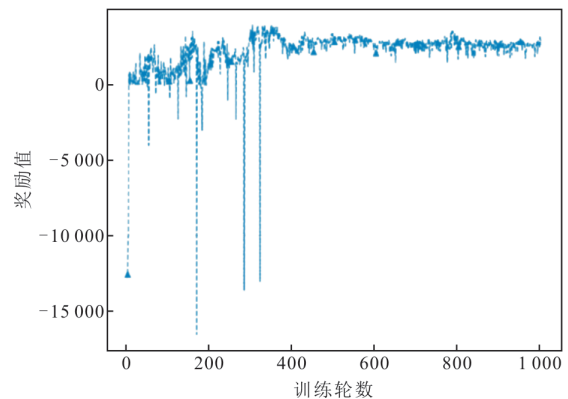


图5 分辨率为50 m的训练奖励值的收敛性

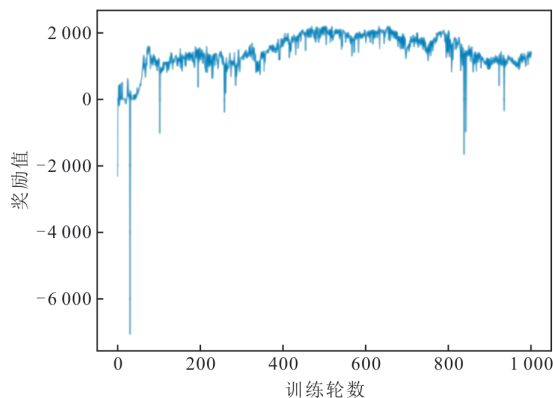


图 6 分辨率为 100 m 的训练奖励值的收敛性

分辨率对智能体的训练动态具有决定性影响。在图 4 中,20 m 分辨率下的奖励曲线表现出极不稳定的特征:训练前期出现多次大幅负向尖峰(最低接近 $-40\ 000$),中后期虽有所缓解,但始终在 0 值线上下持续震荡,未能达到稳定状态。虽然其最终奖励值可达 5 000 左右的高位,但这种极差的稳定性表明超精细的分辨率导致状态空间过于复杂,智能体难以学习到稳健的方案,导致训练难度巨大、收敛过程非常缓慢且不稳定。在图 5 中,50 m 分辨率下的奖励曲线呈现出另一种问题模式:前 400 个训练轮次波动剧烈,频繁出现大幅惩罚(降至 $-15\ 000$ 附近),后期整体稳定在 0 附近的极低水平。这表明在当前奖励函数设计下,该分辨率导致智能体行为的经济成本持续高于其带来的覆盖收益,智能体学习到的是一个相对保守的次优方案。相比之下,在图 6 中,100 m 分辨率下的奖励曲线展现了最佳的综合性能。奖励值曲线在经历初期(约前 200 轮)的必要探索后迅速上升并趋于稳定,后期奖励值稳定在 2 000 左右的较高水平,且波动幅度很小。这表明该分辨率下智能体成功地从探索阶段过渡到了利用阶段,得到一个既高性能又稳定可靠的方案。

基于以上实验结果,虽然 20 m 分辨率能提供最精细的环境信息,但其训练过程的不稳定性使其难以实用;50 m 分辨率则显示出奖励函数失衡的问题;而 100 m 分辨率在保证状态空间信息量的同时,有效控制了问题复杂度,使训练既高效又可靠,是精度与效率的最佳平衡点。因此,本文采用的 100 m 分辨率设置能够满足矿井无线覆盖宏观规划研究的精度需求,同时保证了研究的可行性和结论的可靠性。

为了验证本文系统的有效性,以及所提方案的性能,对所提方案的系统覆盖率性能、基站数量与覆盖率的相关性进行分析,具体仿真结果分别如图 7 和图 8 所示。

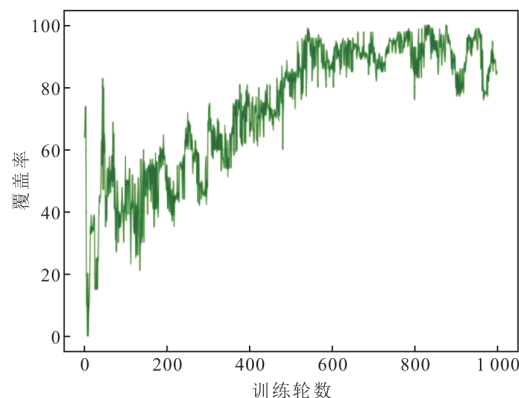


图 7 所提方案的系统覆盖率

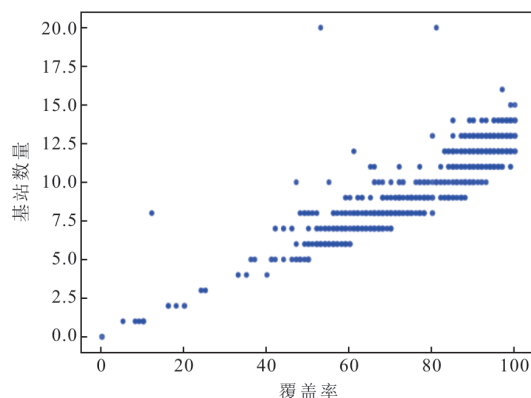


图 8 基站数量与覆盖率的相关性

通过训练一个强化学习智能体,并观察其训练过程的收敛性和性能表现(见图 6)可以看出,智能体的奖励趋于稳定,说明方案已经基本收敛,达到了较优的效果。由图 7 可以看出,随着训练轮次的增加,系统覆盖率呈现明显的 3 阶段提升特征:在初始阶段(0~200 轮次),覆盖率从 0% 快速攀升至 40%,表明所提方案具备良好的初始探索能力;进入中期阶段(200~600 轮次)后,覆盖率以稳定斜率持续增长至 80%,表明所提方案正在有效优化覆盖方案;最终在后期阶段(600~1 000 轮次)逐渐收敛至接近 100% 的覆盖率,验证了所提方案的有效性和稳定性。值得注意的是,600 轮次后性能提升趋于平缓,因此可在此阶段采用动态调整学习率或提前终止训练的方案,以提高训练效率。整体而言,该方案展现出优秀的覆盖性能和学习效率,为实际应用中的区域覆盖任务提供了可靠的技术方案。由图 8 可以看出,基站数量与网络覆盖性能存在

显著的正相关性。随着基站数量的增加,系统覆盖率和平均覆盖率均呈现明显提升趋势。但是,同一覆盖率水平下对应的基站数量并不唯一,呈现出明显的离散分布。例如,在80%覆盖率附近,基站数量可能在8~12之间波动。这种一对多的关系说明网络覆盖性能不仅取决于基站的数量,还与其空间部署位置、干扰情况等因素密切相关。在低基站密度区域,如覆盖率在50%以下时,每新增一个基站带来的覆盖性能改善尤为显著。当基站数量达到中等规模时,系统可实现80%以上的覆盖率,虽然总体覆盖率仍在缓慢上升,但其增益呈现出更强的波动性,边际收益显著递减,系统进入性能饱和阶段。此时,即使增加额外的基站,其对覆盖率的提升也非常有限,反而可能因干扰增强等因素带来系统效率下降的风险。因此,仿真结果不仅验证了基站数量对提升覆盖性能的正面作用,也揭示了基站部署优化的重要性。在实际网络规划中,仅依靠增加基站数量并不能线性提升覆盖性能,应结合基站选址优化、功率控制策略、用户分布特征等因素,实现覆盖性能与部署成本之间的最优平衡。

为深入验证所提方案的性能,对DQN算法在矿井5G优化场景的适用性进行分析。DQN算法与近端策略优化(Proximal Policy Optimization, PPO)算法在训练收敛性和训练奖励方面的性能对比结果如图9所示。

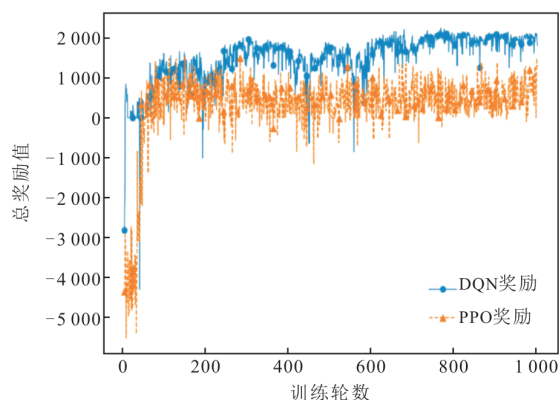


图9 PPO算法与DQN算法的性能对比结果

由图9可以看出,PPO算法的奖励曲线在达到正值后呈现出持续且显著的上下波动(震荡范围较大),表明其方案在训练过程中存在较高的不稳定性。相比之下,采用DQN算法的所提方案,虽然初始学习速度相对较慢,但其奖励曲线在收敛后(约

200回合后)波动幅度显著小于PPO算法,表现出方案具有更优的稳定性。结合矿井边缘设备低算力约束及对通信服务稳定性的严苛要求,DQN算法在推理阶段的计算开销极低,且其方案的确定性更有利于保障网络性能的平稳可靠。尽管PPO算法在奖励峰值上可能更高,但DQN算法在满足核心优化目标(覆盖率>95%)的前提下,有效减少了基站部署数量(28%),并控制了功率调整频率,其低开销、高稳定性的特性使其更贴合矿井实际部署需求。因此,综合考虑边缘算力限制、方案稳定性要求及核心优化目标的达成度,DQN算法被确定为更适合本场景的算法选择。

4 结语

针对传统井下5G部署中理论传播模型与实际偏差大、静态部署难适配动态场景、功率分配无自适应性且难满足防爆要求的问题,提出了一种基于深度强化学习的动态基站部署方案。首先,构建融合LoS/NLoS路径损耗、巷道壁/顶底板粗糙度衰减及沿程极化衰减的综合信号传播模型,精准刻画井下电磁波传播特性。其次,将基站部署与发射功率协同优化转化为“最小基站数量-最大覆盖率”多目标决策问题,采用DQN构建强化学习智能体,通过经验回放与目标网络机制解决函数逼近不稳定性。然后,设计含防爆约束的动态功率调整机制,确保功率调节可响应局部信号衰减突变,且符合矿井防爆标准。最后,仿真结果表明,所提方案在保证95%以上覆盖率的同时,展现出“高损耗区域强化覆盖”的智能特性,有效地提升了井下5G网络覆盖性能,减少了基站部署规模,能够降低网络部署成本与设备运行功耗,且具备适配井下复杂场景的智能特性,能够为智能矿山5G通信系统工程化应用提供理论依据与技术参考。

参考文献

- [1] 孙继平. 煤矿智能化与矿用5G[J]. 工矿自动化, 2020, 46(8):1-7.
SUN J P. Coal mine intelligence and mine-used 5G[J]. Industry and Mine Automation, 2020, 46(8):1-7. (in Chinese)
- [2] 王国法. 煤矿智能化最新技术进展与问题探讨[J]. 煤

- 炭科学技术,2022,50(1):1-27.
WANG G F. New technological progress of coal mine intelligence and its problems [J]. Coal Science and Technology,2022,50(1):1-27. (in Chinese)
- [3] 何勇,徐元涛. 5G 通信技术在智能化煤矿的应用与研究[J]. 能源与节能,2023(3):167-169.
HE Y,XU Y T. Application and research of 5G communication technology in intelligent coal mines[J]. Energy and Energy Conservation,2023(3):167-169. (in Chinese)
- [4] LEINONEN M E, HOVINEN V, VUOHTONIEMI R, et al. Radio channel characterization measurements in an underground mine for 5G at sub-6 GHz and millimeter wave frequencies[J]. International Journal of Microwave and Wireless Technologies,2025,17(5):850-861.
- [5] 李啸威. 人工智能在 5G 无线网络优化中的设计与实现[J]. 数字技术与应用,2024,42(9):60-62.
LI X W. Design and implementation of artificial intelligence in optimization of 5G wireless network[J]. Digital Technology & Application,2024,42(9):60-62. (in Chinese)
- [6] 尹东升. 5G-700 MHz 技术在三元矿井下应用[J]. 现代矿业,2023,39(7):252-255.
YIN D S. Application of 5G-700 MHz technology in Sanyuan Mine[J]. Modern Mining,2023,39(7):252-255. (in Chinese)
- [7] 孙继平,彭铭,刘斌. 矿井无线传输测试分析与矿用 5G 优选工作频段研究[J]. 工矿自动化,2024,50(10):1-11.
SUN J P,PENG M,LIU B. Analysis of wireless transmission tests in mines and preferred working frequency bands for mining 5G[J]. Journal of Mine Automation,2024,50(10):1-11. (in Chinese)
- [8] 孙继平,彭铭. 矿井无线电波防爆安全发射功率研究[J]. 工矿自动化,2024,50(3):1-5.
SUN J P,PENG M. Research on the safe transmission power of mine radio wave explosion prevention[J]. Journal of Mine Automation,2024,50(3):1-5. (in Chinese)
- [9] 索智文,王亚坤. 煤矿井下重点场所 5G 覆盖性能研究及验证[J]. 工矿自动化,2025,51(2):34-40.
SUO Z W,WANG Y K. Research and verification of 5G coverage performance in key areas of underground coal mines[J]. Journal of Mine Automation,2025,51(2):34-40. (in Chinese)
- [10] 张锬,赵伟. 基于人工智能的 5G 无线网规划和优化分析[J]. 中国宽带,2024(10):61-63.
ZHANG K,ZHAO W. Planning and optimization analysis of 5G wireless network based on artificial intelligence[J]. China Broad Band,2024(10):61-63. (in Chinese)
- [11] 胡贵宾,赵川斌. 基于机器学习和波束配置的 5G 天线方位角智能优化算法及实践[J]. 广东通信技术,2024,44(4):37-40.
HU G B,ZHAO C B. Intelligent azimuth optimization algorithm and practice of 5G antenna based on machine learning and beam configuration[J]. Guangdong Communication Technology,2024,44(4):37-40. (in Chinese)
- [12] 冯卫国. 5G 技术在煤矿智能化中的应用[J]. 能源与节能,2025(2):238-240.
FENG W G. Application of 5G technology in coal mine intelligence[J]. Energy and Energy Conservation,2025(2):238-240. (in Chinese)
- [13] YAN M,FENG G,ZHOU J H, et al. Intelligent resource scheduling for 5G radio access network slicing [J]. IEEE Transactions on Vehicular Technology,2019,68(8):7691-7703.
- [14] AL-TAHMEESSCHI A, TALVITIE J, LÓPEZ-BENÍTEZ M, et al. Multi-objective deep reinforcement learning for 5G base station placement to support localisation for future sustainable traffic [C]//2024 Joint European Conference on Networks and Communications & 6G Summit. Antwerp;IEEE,2024:493-498.
- [15] SORANZO E, GUARDIANI C, WU W. Reinforcement learning for the face support pressure of tunnel boring machines [J]. Geosciences,2023,13(3):82.
- [16] 樊昌信,曹丽娜. 通信原理[M]. 7 版. 北京:国防工业出版社,2015.
FAN C X, CAO L N. Principles of communications [M]. 7th ed. Beijing:National Defense Industry Press,2015. (in Chinese)
- [17] 张侯,陈隆,陈汉章. 基于无线网络通信的煤矿机电设备运行自动化监测系统设计[J]. 电子设计工程,2025,33(17):154-157.
ZHANG H, CHEN L, CHEN H Z. Design of an automated monitoring system for coal mine electromechanical equipment operation based on wireless network communication [J]. Electronic Design Engineering,2025,33(17):154-157. (in Chinese)
- [18] QIAN J Y, WU Y T, SALEEM A, et al. Path loss

- model for 3.5 GHz and 5.6 GHz bands in cascaded tunnel environments[J]. *Sensors*, 2022, 22(12):4524.
- [19] YOU Y S, JING L J, YANG Q H, et al. Analysis and modeling of radio propagation fading characteristics in tunnel construction[J]. *Heliyon*, 2024, 10(4):e26231.
- [20] LI B Y, DING T, WU Y H, et al. Modeling and optimization of wireless signal transmission characteristics of mine roadway based on 3D ray-tracing method[J]. *Applied Sciences*, 2024, 14(4):1534.
- [21] 姚善化, 吴先良, 张量. 矿井巷道壁粗糙度对电磁波传播损耗的影响[J]. *合肥工业大学学报(自然科学版)*, 2010, 33(11):1725-1727.
- YAO S H, WU X L, ZHANG L. Influence of rough wall on electromagnetic waves propagation attenuation in mine tunnels [J]. *Journal of Hefei University of Technology (Natural Science)*, 2010, 33 (11): 1725-1727. (in Chinese)
- [22] SHAHID S, ZAHRA H, QAISAR S B, et al. Radio link model for node deployment in underground mine sensor networks [J]. *Applied Sciences*, 2023, 13 (15):8987.
- [23] 3GPP. Study on channel model for frequencies from 0.5 to 100 GHz: TR 38.901 [S/OL]. [2025-03-01]. https://www.3gpp.org/ftp/Specs/archive/38_series/38.901/38901-i00.zip.
- [24] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning [J]. *Nature*, 2015, 518(7540):529-533.
- [25] 江雨霏, 朱其新. 深度强化学习下的管道气动软体机器人控制[J]. *西安工程大学学报*, 2025, 39(2):65-74.
- JIANG Y F, ZHU Q X. Pipe pneumatic soft robot control based on deep reinforcement learning[J]. *Journal of Xi'an Polytechnic University*, 2025, 39(2):65-74. (in Chinese)
- [26] MISMAR F B, EVANS B L, ALKHATEEB A. Deep reinforcement learning for 5G networks: Joint beamforming, power control, and interference coordination [J]. *IEEE Transactions on Communications*, 2020, 68 (3):1581-1592.

[作者简介]



梁沫(1986—),女,陕西宝鸡人,硕士,中煤科工集团北京华宇工程有限公司高级工程师,主要研究方向为煤矿智能化、物联网技术等。E-mail:290840496@qq.com



王军选(1970—),男,陕西西安人,博士,西安邮电大学教授、博士生导师,主要研究方向为宽带无线通信、物联网技术等。E-mail:wangjx@xupt.edu.cn

[责任编辑:刘武英]