



专题：可见光通信

基于 DQN 的 UUV 辅助水下无线光通信轨迹规划系统

胡珈玮¹, 刘晓谦¹, 唐昕柯², 董宇涵^{1,2}

(1. 清华大学深圳国际研究生院, 广东 深圳 518055; 2. 鹏城实验室, 广东 深圳 518055)

摘要: 无人水下航行器 (unmanned underwater vehicle, UUV) 作为重要潜基通信平台可以辅助水下无线光通信 (underwater wireless optical communication, UWOC)。然而, 在实际应用中, 水体波动特性、不同水质环境、多用户接入等给 UUV 辅助 UWOC 系统带来很大挑战, 因而适当的路径规划策略可以应对上述挑战并最大限度地提升系统整体和每一个用户的性能。将深度强化学习 (deep reinforcement learning, DRL) 用于无人载具路径规划中, 提出了一种 UUV 辅助 UWOC 系统的轨迹规划方案。通过 DRL 中深度 Q 网络 (deep Q-network, DQN) 方法让 UUV 自动决策航行方向, 从而提升系统和用户的链路通信容量。同时, 研究了不同水质对容量提升效果的影响。仿真实验表明, DQN 输出策略可以在一定程度上提升系统整体和各个用户的链路通信容量, 并且 UUV 在清澈海水中的容量提升效果优于纯净海水但低于沿岸水。

关键词: 无人水下航行器; 无线光通信; 深度强化学习; 轨迹规划

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2023108

Trajectory planning of UUV-assisted UWOC systems based on DQN

HU Jiawei¹, LIU Xiaoqian¹, TANG Xinke², DONG Yuhan^{1,2}

1. Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China

2. Peng Cheng Laboratory, Shenzhen 518055, China

Abstract: As a key submarine-based communication platform, unmanned underwater vehicle (UUV) can facilitate underwater wireless optical communication (UWOC). However, fluctuating characteristics of water body, different water qualities, multi-user access present challenges to UUV-assisted UWOC systems, which could be alleviated by an appropriate path planning to maximize the system and each user performance. Deep reinforcement learning (DRL) technology was applied in the path planning of autonomous vehicles, a trajectory planning scheme for UUV-assisted UWOC systems was proposed. The UUV automatically decides the navigation direction through deep Q-network (DQN) method, thereby improving the communication capacity of the system and each user. The impact of distinct water qualities on the capacity enhancement was also investigated. Simulation results suggest that the outputted strategy of DQN can improve the link capacity of the system and each user. This capacity improvement in clear seawater is better than that in pure seawater but lower than that in coastal water.

Key words: UUV, optical wireless communication, DRL, trajectory planning

收稿日期: 2023-04-15; 修回日期: 2023-05-10

通信作者: 董宇涵, dongyuhan@sz.tsinghua.edu.cn

0 引言

由于无人水下航行器 (unmanned underwater vehicle, UUV) 具有灵活性、机动性和全自主性等优点, 其辅助的水下通信系统在“空-天-地-海”一体化进程中扮演重要的潜基通信平台角色, 并在应急、侦察、探测等领域具有广阔发展前景^[1]。水下通信系统一般可以通过射频波、声波和光波这 3 种典型的无线载体实现。然而电磁波在水下受到较高的吸收和衰减, 导致水下链路通信范围被限制在几米范围内, 声波虽然可以在一定程度上解决通信距离过短问题, 但低传输速率和较大传输时延的缺陷限制了其在水下通信领域发展^[2]。水下无线光通信 (underwater wireless optical communication, UWOC) 相比于水下射频和水声通信技术, 具有速率高、功耗低、抗截获性强等优点, 从而在近年来得到了较为广泛的研究关注^[3-4]。

UUV 辅助 UWOC 综合了 UUV 和 UWOC 的优点, 成为水下物联网的使能方案。然而在实际应用中, 水体波动特性、不同水质环境、多用户接入等给 UUV 辅助 UWOC 系统带来很大挑战。目前已有部分文献^[4-7]尝试研究 UUV 辅助 UWOC 系统的光场分布、路径规划和能量消耗等因素, 试图从不同途径应对上述挑战。然而相比于无人机 (unmanned aerial vehicle, UAV) 辅助大气无线光通信研究^[8-11], 很少有工作研究 UUV 轨迹规划对提升 UWOC 系统性能的影响。针对这一问题, 本文使用深度强化学习 (deep reinforcement learning, DRL) 中深度 Q 网络 (deep Q-network, DQN) 技术对 UUV 轨迹进行自动规划, 并研究了纯净海水、清澈海水和沿岸水这 3 种水质中对 UUV 辅助 UWOC 系统容量的提升效果。

1 系统模型

UUV 辅助 UWOC 系统模型如图 1 所示。本文假设一个固定高度为 H 的 UUV 在限定空间 V 中,

为水下 3 个位置随机且静止的用户提供通信和照明服务。当系统运行时, UUV 自动决策其在下一时隙的航行方向, 从而达到提升系统整体通信容量的目标。

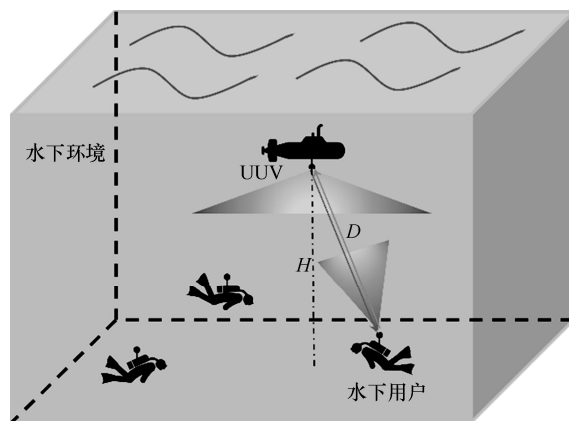


图 1 UUV 辅助 UWOC 系统模型

1.1 信道模型

在 UWOC 中, 光子会受到海水吸收、散射及湍流的影响, 能量的损失和传播方向发生改变。近年来, 研究人员提出了多种方法模拟水下无线光通信的链路损耗特性, 其中朗伯比尔 (Lambert-Beer) 定律是用于评估光损耗最简单的模型。Lambert-Beer 定律将接收功率描述为链路距离及海水消光系数 (也称衰减系数 c) 的函数, 且接收功率呈指数下降趋势。该方法因为形式简单、计算方便而被广泛使用。然而, 该方法假设散射光子完全丢失, 背离实际情况, 具有一定误差。

针对上述问题, Elamassie 等^[12]基于 Lambert-Beer 定律提出了一个适用于不同水质 UWOC 链路的闭合形式路径损耗表达式。该表达式考虑了散射分量的贡献, 与真实水下链路衰减特性更为吻合, 并适用于不同的波束发散角和各种接收器 (探测器) 直径。因此, 本文使用文献[12]中的路径损耗模型, 即光子在水下的路径损耗为:

$$G(d) = 10 \lg \left(\left(\frac{D_R}{\theta_{1/e} d} \right)^2 e^{-cd \left(\frac{D_R}{\theta_{1/e} d} \right)^r} \right) \quad (1)$$



其中, D_r 是接收器孔径; $\theta_{i/c}$ 是发射机的波束发散角; c 为海水消光系数, 在不同水质下具有不同的取值; d 是发射器和接收器之间的链路距离; T 是校正系数, 可以通过对蒙特卡洛仿真数据的曲线拟合确定; $G(d)$ 的单位为 dB。

在确定好路径损耗模型之后, UUV 同水下用户 j 之间的链路通信容量可以表示为^[8]:

$$C_j: \begin{cases} C_j^{\text{DL}} = \frac{1}{2} \lg \left(1 + \frac{e}{2\pi} \left(\frac{\zeta P^{\text{DL}} G(d)}{n_w} \right)^2 \right) \\ C_j^{\text{UL}} = \frac{1}{2} \lg \left(1 + \frac{e}{2\pi} \left(\frac{\zeta P^{\text{UL}} G(d)}{n_w} \right)^2 \right) \end{cases} \quad (2)$$

其中, C_j^{DL} 是下行链路信道容量, 即 UUV 到水下用户 j 之间的链路; C_j^{UL} 是上行链路信道容量, 即水下用户 j 到 UUV 之间的链路; ζ 是调光目标; P^{DL} 是 UUV 的 LED 发射机功率; P^{UL} 是水下用户的 LED 发射机功率; n_w 是环境中的噪声。

1.2 问题表述

UUV 的目标是通过合适的轨迹规划方法最大化上行和下行链路的信道容量之和, 从而提升系统性能。然而在实际应用中, UUV 不仅需要根据信道容量的变化改变自身的位置信息, 同时还必须满足一些限制条件, 例如, UUV 不能移动超出限定空间 V , 其上下行链路的最小信道容量不能低于通信阈值。综合上述考虑, 本文提出的轨迹规划问题在数学上可以表述为:

$$\begin{aligned} & \max_{x_t, y_t} \sum_{j \in J} (C_j^{\text{DL}} + C_j^{\text{UL}}) \\ & \text{s.t. } (x_j, y_j, z_j) \in V \\ & (x_t, y_t, H) \in V \\ & C_j^{\text{DL}} \geq \delta^{\text{DL}}, C_j^{\text{UL}} \geq \delta^{\text{UL}} \end{aligned} \quad (3)$$

其中, (x_j, y_j, z_j) 是第 j 个水下用户的三维坐标; (x_t, y_t, H) 是 UUV 在第 t 个时隙的三维坐标; δ^{DL} 和 δ^{UL} 分别表示下行链路和上行链路的通信阈值, 即 UUV 必须保证所有用户的上下行链路容量都大于对应的阈值, 并在限定的空间中对整体通信容量之和进行提升。

2 基于 DRL 的 UUV 轨迹规划

由于 UUV 需要同时为多个水下用户提供服务并满足若干限制条件, 这使得式 (3) 成为一个非凸问题, 从而导致传统优化算法无法获得最优路径甚至无法求解。而 DQN 通过与环境不断交互的方式获取 UUV 航行动作对应的奖励值 r 影响其航迹规划策略, 从而实现 UUV 在无须访问所有可用环境数据的情况下学习得到最佳航行策略的目标。

DQN 从传统基于 Q 表进行决策的 Q 学习 (Q-learning) 技术改进而来。它使用 Q 网络和目标 Q 网络这两个神经网络 (neural network, NN) 对原有的 Q 表进行拟合, 使得 NN 的输出结果与 Q 表中的值接近, 从而解决状态空间过大、过复杂导致的 Q 表维护困难和单 NN 拟合稳定性较差的问题。为了保证系统具备一定的探索性, UUV 以概率 P 选择 NN 输出的最大 Q 值所对应的动作选项 a , 以 $1-P$ 的概率随机选择下一步动作。之后系统与环境进行交互获取奖励值并通过不断降低损失函数的方式更新 NN 的参数, 从而逐渐学习到一个航迹规划策略。DQN 算法结构如图 2 所示。

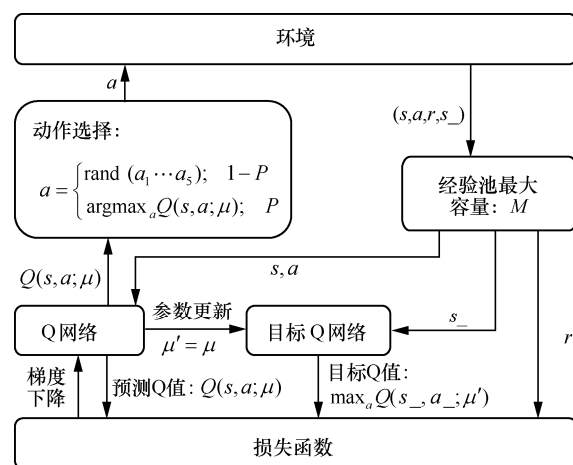


图2 DQN 算法结构

图 2 中两个 NN 均由 1 个输入层、3 个全连接层和 1 个输出层构成, 其中全连接层的每一层都包含 256 个节点, 输入层和输出层分别有 3 个和 5 个节点, 它们各自代表 UUV 的三维坐标和动作选项。

由于 UUV 处于固定高度, 故其有 5 个动作选择, 分别为前、后、左、右和静止。当 UUV 在状态 s 处选择动作 a 之后, 其在环境中移动到了下一状态 s_+ 并获得了一个奖励值 r , 随后系统将 (s, a, r, s_+) 打包作为一个经验组存入最大容量为 M 的经验池中, 这些经验组将被用于后续规划策略的学习。

当经验池存满时, 系统开始从中随机采样, 将采样得到的 r 和 s_+ 发送给目标 Q 网络从而计算目标 Q 值; 将 s 和 a 发送给 Q 网络计算预测 Q 值。随后系统计算二者之间的均方误差并通过梯度下降的方法更新 Q 网络中的参数 μ 。与此同时, 每隔 N 步将 Q 网络的参数同步到目标 Q 网络的参数 μ' 当中。均方误差和梯度下降方案分别如式 (4) 和式 (5) 所示。

$$L(\sigma) = E \left[((r_s^a + \gamma \max_{a_-} Q(s_-, a_-; \mu')) - Q(s, a; \mu))^2 \right] \quad (4)$$

$$\nabla_{\mu} L(\sigma) = E \left[(r_s^a + \gamma \max_{a_-} Q(s_-, a_-; \mu') - Q(s, a; \mu)) \nabla_{\mu} Q(s, a; \mu) \right] \quad (5)$$

其中, r_s^a 表示 UUV 在状态 s 处选择动作 a 后获得的奖励值, γ 表示奖励折扣。其中 $\gamma < 1$ 并且 γ 的存在让 UUV 在关注当前奖励 r 的同时评估长期奖励的影响, 从而学习到一个能让长期回报最大化的轨迹规划策略。

3 仿真结果和分析

本文对所设计的方案进行了仿真实验。考虑 UUV 辅助 UWOC 系统, UUV 搭载 532 nm 光源, 在纯净海水、清澈海水及沿岸水 3 种不同水质中向 3 个水下固定位置用户提供通信服务。采用如式 (1) 所示的路径损耗模型, 其中的校正系数 T 由蒙特卡洛仿真计算得到。在轨迹规划求解与性能仿真评估方面, 实验所用的设备为一台具有 4 个 NVIDIA RTX 2080 Ti 的 GPU 服务器, 并使用 PyTorch 1.10.2 和 Python 3.6 执行轨迹规划程序。仿真参数汇总见表 1。

不同水质情况下所提 UUV 轨迹规划策略对系统通信容量的提升效果如图 3 所示, 展示了在 3 种

不同水质情况下 DQN 学习到的轨迹规划策略对系统总通信容量的影响。可以看出, 3 条曲线都呈先上升后趋于平稳的趋势, 这说明该轨迹规划系统实现了提升整体通信容量的目标。与 UUV 初始位置对应的系统通信容量相比, UUV 执行轨迹规划策略后, 系统通信容量在纯净海水中提升了 19.41%, 在清澈海水中提升了 25.63%, 在沿岸水中提升了 32.75%。

表 1 仿真参数汇总

参数名称	参数值
接收器孔径/cm	5
接收机视场角	180°
水下用户发射机功率/W	1
UUV 高度/m	4
清澈海水消光系数/m ⁻¹	0.15
纯净海水校正系数	0
沿岸水校正系数	0.13
发射器波束发散角	6°
UUV 发射机功率/W	10
调光目标	0.8
纯净海水消光系数/m ⁻¹	0.056
沿岸水消光系数/m ⁻¹	0.305
清澈海水校正系数	0.05
限定空间/(m × m × m)	12 × 12 × 6

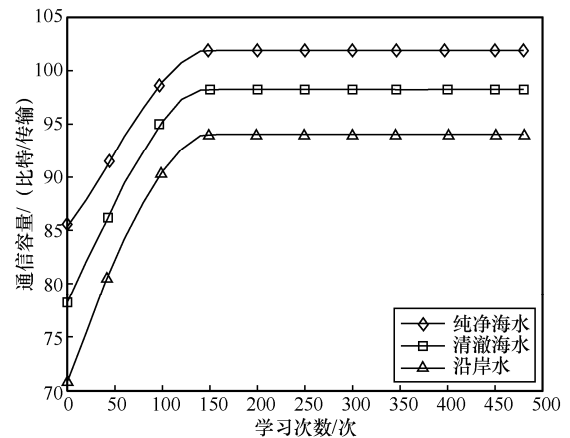


图 3 不同水质情况下所提 UUV 轨迹规划策略对系统通信容量的提升效果

不同水质情况下所提 UUV 轨迹规划策略对不同水下用户通信容量的提升效果如图 4 所示, 展示了不同水质情况下 DQN 学习到的轨迹规划策略对不同用户通信容量的影响。可以看出, 无论是哪一个水下用户在何种水质中, UUV 轨迹规划策略最

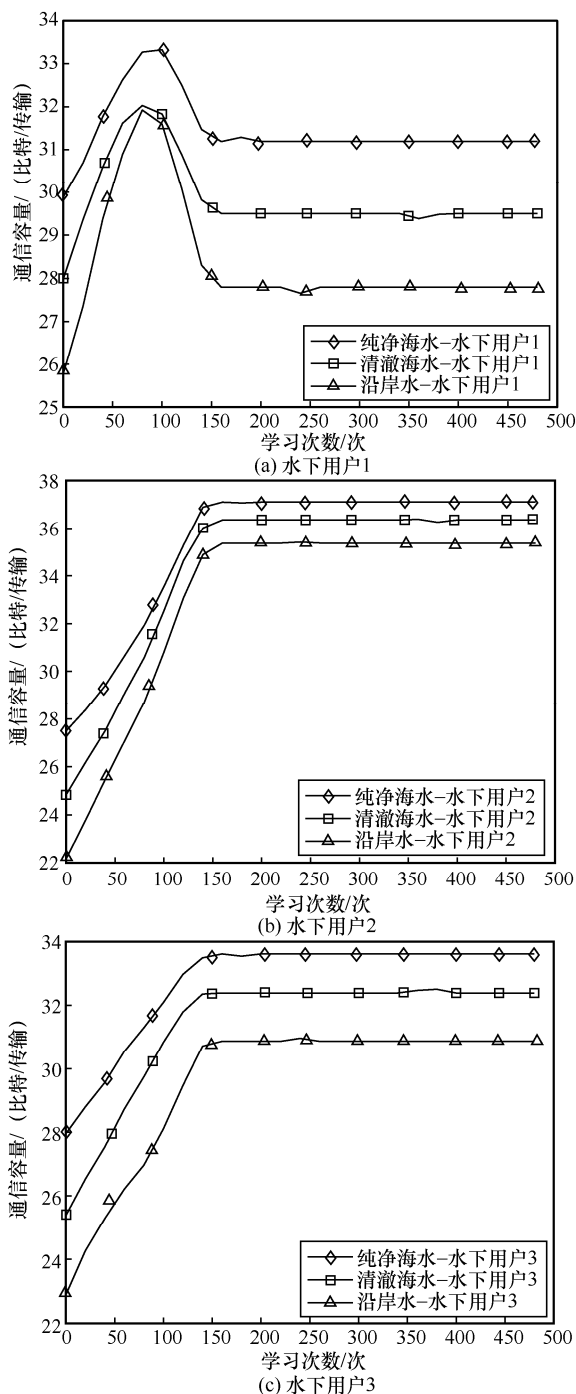


图4 不同水质情况下所提UUV轨迹规划策略对不同水下用户通信容量的提升效果

终都对链路通信容量有不同程度的提高。其中,水下用户1通信容量在纯净海水、清澈海水和沿岸水中分别提升了4.37%、5.42%和7.82%;水下用户2的提升效果分别为34.95%、46.39%和59.53%;水下用户3的提升效果分别为20.22%、27.61%和

34.90%。这从另一个角度验证了图4中UUV轨迹规划策略提升系统容量的结论。

水质对链路容量提升效果的影响如图5所示,将上述所有通信容量提升效果数据进行汇总并按不同的用户归类绘制成柱状图,从而研究水质对链路容量提升效果的影响。通过观察柱状图的高度可以得出结论:UUV轨迹规划策略对清澈海水环境的用户通信容量提升效果优于纯净海水但低于沿岸水。

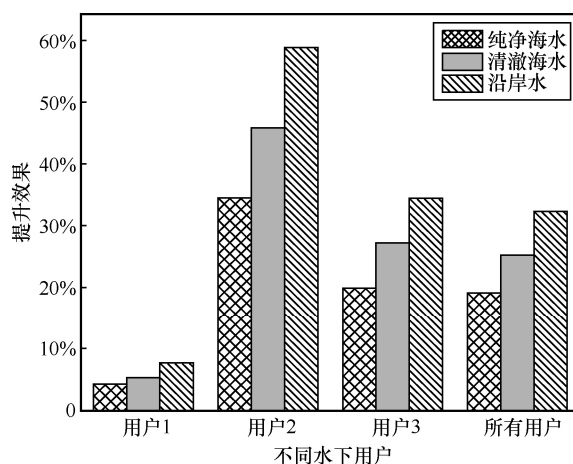


图5 水质对链路容量提升效果的影响

4 结束语

本文研究了轨迹规划策略对UUV辅助的水下无线光通信链路容量的影响,并通过DQN方法实现了UUV的自动航行方向决策。在此基础上,本文还研究了3种不同水质对容量提升效果的影响。仿真结果表明,DQN的输出策略可以提升系统整体和每一个用户的链路通信容量,而UUV在不同水质环境中能达到的容量提升效果并不相同,其在清澈海水水质的提升效果优于纯净海水水质但低于沿岸水水质。

参考文献:

- [1] 王亭亭, 张南南, 岳才谦, 等. 基于水声通信的AUV组网与协同导航[J]. 水下无人系统学报, 2021, 29(4): 400-406.
WANG T T, ZHANG N N, YUE C Q, et al. AUV networking

- and cooperative navigation based on underwater acoustic communication[J]. Journal of Unmanned Undersea Systems, 2021, 29(4): 400-406.
- [2] HUANG N, GONG C, FU C F, et al. Preliminary investigation of air-to-water visible light communication link under strong ambient light[C]//Proceedings of 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall). Piscataway: IEEE Press, 2021: 1-5.
- [3] 郭银景, 徐锋, 屈衍玺, 等. 水下可见光通信关键技术综述[J]. 光通信研究, 2020(2): 1-6, 19.
- GUO Y J, XU F, QU Y X, et al. A survey of key technologies of underwater visible light communication[J]. Study on Optical Communications, 2020(2): 1-6, 19.
- [4] 吕斌斌, 林酩涑, 万鑫. 水下可见光通信在水下作战体系中的应用设想[J]. 水下无人系统学报, 2022, 30(6): 787-793.
- LYU B B, LIN M L, WAN X. Application of underwater visible light communication in underwater warfare system[J]. Journal of Unmanned Undersea Systems, 2022, 30(6): 787-793.
- [5] ANGUITA D, BRIZZOLARA D, PARODI G, et al. Optical wireless underwater communication for AUV: preliminary simulation and experimental results[C]//Proceedings of OCEANS 2011 IEEE - Spain. Piscataway: IEEE Press, 2011: 1-5.
- [6] MAHMOODI K A, UYSAL M. AUV trajectory optimization for an optical underwater sensor network in the presence of ocean currents[C]//Proceedings of 2021 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom). Piscataway: IEEE Press, 2021: 1-6.
- [7] MAHMOODI K A, UYSAL M. Energy aware trajectory optimization of solar powered AUVs for optical underwater sensor networks[J]. IEEE Transactions on Communications, 2022, 70(12): 8258-8269.
- [8] WANG Y N, CHEN M Z, YANG Z H, et al. Deep learning for optimal deployment of UAVs with visible light communications[J]. IEEE Transactions on Wireless Communications, 2020, 19(11): 7049-7063.
- [9] ZHU Z Y, YANG Y, GUO C L, et al. Power efficient deployment of VLC-enabled UAVs[C]//Proceedings of 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications. Piscataway: IEEE Press, 2020: 1-6.
- [10] CANG Y H, CHEN M, ZHAO J W, et al. Joint deployment and resource management for VLC-enabled RISs-assisted UAV networks[J]. IEEE Transactions on Wireless Communications, 2023, 22(2): 746-760.

- [11] YANG Y, CHEN M Z, GUO C L, et al. Power efficient visible light communication with unmanned aerial vehicles[J]. IEEE Communications Letters, 2019, 23(7): 1272-1275.
- [12] ELAMASSIE M, MIRAMIRKHANI F, UYSAL M. Performance characterization of underwater visible light communication[J]. IEEE Transactions on Communications, 2019, 67(1): 543-552.

[作者简介]



胡珈玮(1999—), 男, 清华大学深圳国际研究生院硕士生, 主要研究方向为深度强化学习和无人机辅助通信。



刘晓谦(1999—), 男, 清华大学深圳国际研究生院硕士生, 主要研究方向为水下无线光通信和下一代短距无线通信技术。



唐昕柯(1989—), 男, 鹏城实验室副研究员, 主要研究方向为无线光通信、量子密钥分发、光纤通信等。



董宇涵(1979—), 男, 清华大学深圳国际研究生院副教授、鹏城实验室副研究员, 主要研究方向为无线通信与网络、无线光通信、机器学习与优化。