

doi:10.13756/j.gtxyj.2026.260051.

专题:面向6G的无线光通信

贾林琼,王梓颀,汤昀霏,等.面向用户移动性的VLC辅助的UAV网络能效优化[J].光通信研究,2026(3):13-19.

Jia L Q, Wang Z J, Tang Y F, et al. Energy Efficiency Optimization for VLC-enabled UAV Networks Serving Mobile Users[J]. Study on Optical Communications, 2026(3):13-19.

# 面向用户移动性的VLC辅助的UAV网络能效优化 (特邀)

贾林琼<sup>1</sup>,王梓颀<sup>1</sup>,汤昀霏<sup>1</sup>,桂林卿<sup>2</sup>

(1. 南京理工大学 电子工程与光电技术学院,南京 210094; 2. 南京邮电大学 计算机学院,南京 210003)

**摘要:**【目的】在可见光通信(VLC)辅助的无人机(UAV)网络中,为满足地面用户的动态接入,UAV也需进行动态位置调整以保持与地面用户的视距(LoS)连接。因此UAV的有限机载能量需同时用于通信与飞行推进。【方法】为提升该网络的能效(EE),文章提出了一种针对移动用户的UAV轨迹优化迭代算法。该算法迭代求解UAV最优位置和最优轨迹两个子问题,其中UAV最优位置、用户关联及功率分配可通过软演员-评论家(SAC)算法联合求解;UAV移动至下一个最优位置的轨迹功耗最小化问题可以建模为一个凸问题,并通过优化工具箱求解。【结果】数值仿真结果验证了SAC算法的收敛性,并给出了网络EE最大化的UAV位置及用户与UAV的关联情况;展示了用户动态移动下通过迭代算法得出的最优轨迹。【结论】文章所提迭代算法可以有效降低单位时间内VLC辅助的UAV网络的能耗,以此延长UAV的工作时间。

**关键词:**可见光通信;无人机;能效;轨迹优化;软演员-评论家算法

**中图分类号:**TN929 **文献标志码:**A

## Energy Efficiency Optimization for VLC-enabled UAV Networks Serving Mobile Users

JIA Linqiong<sup>1</sup>, WANG Zijie<sup>1</sup>, TANG Yunfei<sup>1</sup>, GUI Linqing<sup>2</sup>

(1. School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China;

2. School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

**Abstract:** 【Objective】In Visible Light Communication (VLC)-enabled Unmanned Aerial Vehicle (UAV) networks, UAVs adjust their locations to keep Line-of-Sight (LoS) links with moving ground users. As a result, the limited onboard energy of a UAV must cover both communication tasks and flight propulsion. 【Methods】To improve the Energy Efficiency (EE) of these networks, we propose an iterative trajectory optimization algorithm. Instead of solving a coupled global problem, we divide it into two alternating subproblems: optimal positioning and trajectory planning. First, we apply the Soft Actor-Critic (SAC) algorithm to jointly optimize the UAV's target position, user association, and power allocation. Once the target position is set, we model the flight trajectory to minimize propulsion energy as a convex problem, which can be solved based on standard optimization tools. 【Results】Simulation results confirm the convergence of the SAC algorithm in our setup. We show the optimized UAV locations and user association schemes that lead to the highest network EE, along with the actual flight trajectories generated for dynamic user mobility scenarios. 【Conclusion】Overall, our iterative framework successfully lowers the time-averaged energy consumption of the network, which directly extends the flight endurance of the UAVs.

**Key words:** VLC; UAV; EE; trajectory optimization; SAC algorithm

## 0 引言

可见光通信(Visible Light Communication, VLC)辅助的无人机(Unmanned Aerial Vehicle, UAV)网络能够动态建立光通信链路,为通信受限或射频拥塞的场景(如偏远地区、灾后搜救以及人群密集场所)提供灵活的网络覆盖<sup>[1-2]</sup>。

VLC辅助的UAV网络部署面临着UAV位置

优化和动态用户接入管理等关键挑战。针对上述问题,现有研究提出了多种解决方案:文献[3]提出了基于哈里斯鹰优化和神经网络的算法,通过联合优化UAV位置与功率分配,实现系统和速率的最大化;文献[4]利用深度学习预测未来背景光分布,通过联合优化UAV位置和用户关联,实现了UAV发射功率最小化。由于UAV机载能量有限,能效(Energy Efficiency, EE)优化问题备受关注。

收稿日期:2026-02-13; 修回日期:2026-04-02; 纸质出版日期:2026-06-10

基金项目:国家自然科学基金资助项目(62301257)

作者简介:贾林琼(1990-),女,山东滨州人。副教授,博士,主要研究方向为无线光通信。

通信作者:贾林琼,副教授。E-mail: jialinqiong@njjust.edu.cn

© Editorial Office of *Study on Optical Communications*. This is an open access article under the CC BY-NC-ND license.

例如,文献[5]运用改进的粒子群算法优化 UAV 运动轨迹,有效权衡了用户速率与通信功耗,提升了网络 EE 与覆盖;文献[6]提出了一种二维部署策略,用以求解满足通信质量约束的最大化 EE 问题。

但现有的研究多局限于地面用户位置固定的理想假设,而对于随机移动的地面用户,UAV 需动态调整位置以保持连接,其有限的机载能量在通信传输与飞行推进之间面临严峻的分配权衡。因此本文提出了一种面向用户动态移动的 UAV 轨迹优化迭代算法,该算法将位置优化与轨迹规划分解为两个子问题进行迭代求解。仿真结果验证了本文所提算法的收敛性与有效性。

## 1 VLC 辅助的 UAV 网络模型

图 1 所示为  $U$  架旋翼 UAV 和  $K$  个地面用户组成的下行 VLC 网络,其中 UAV 的集合可表示为  $U = \{1, \dots, u, \dots, U\}$ ,第  $u$  架 UAV 的三维位置定义为  $q_u = (q_{u,x}, q_{u,y}, q_{u,z})$ 。假设每架 UAV 的悬停高度是可调节的,调节范围为  $h_L \leq q_{u,z} \leq h_H$ ,其中  $h_L$  和  $h_H$  分别为悬停最小和最大允许高度,UAV 的位置集合为矩阵  $Q = [q_1, q_2, \dots, q_U] \in \mathbb{R}^{3 \times U}$ ,其中  $\mathbb{R}$  为实数空间。地面用户使用光电二极管作为接收器,地面用户集合为  $K = \{1, \dots, k, \dots, K\}$ 。第  $k$  个用户的位置为  $w_k = (w_{k,x}, w_{k,y}, w_{k,z})$ ,所有用户的位置矩阵为  $W = [w_1, \dots, w_K] \in \mathbb{R}^{3 \times K}$ 。第  $k$  个用户与第  $u$  架 UAV 之间的距离为  $d_{k,u} = \|w_k - q_u\|_2$ ,距离矩阵为  $D = [d_{k,u}] \in \mathbb{R}^{K \times U}$ ,其中  $\|\cdot\|_2$  为 L2 范数。

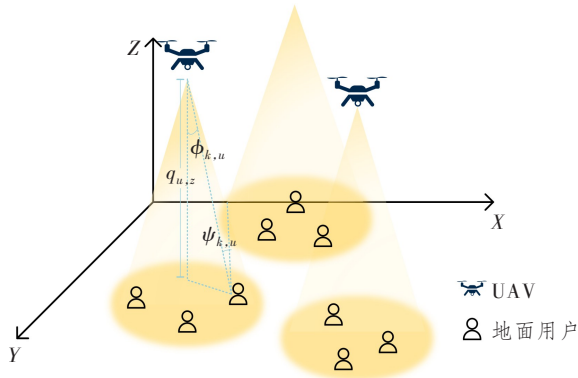


图 1 VLC 辅助的 UAV 网络

Figure 1 VLC-enabled UAV network

根据朗伯辐射模型,第  $k$  个用户和第  $u$  架 UAV 之间的信道增益  $h_{k,u}$  可表示为<sup>[7]</sup>

$$h_{k,u} = \begin{cases} \frac{(m+1)A}{2\pi d_{k,u}^2} g(\psi_{k,u}) \cos^m(\phi_{k,u}) \cos(\psi_{k,u}) & 0 < \psi_{k,u} < \Psi_C \\ 0 & \psi_{k,u} \geq \Psi_C \end{cases}, \quad (1)$$

式中: $m$  为朗伯辐射阶数; $A$  为光电二极管探测面积; $g(\psi_{k,u})$  为聚光器增益; $\psi_{k,u}$  为辐射角; $\phi_{k,u}$  为入射角; $\Psi_C$  为接收半视场角,若  $\psi_{k,u} \leq \Psi_C$ ,  $g(\psi_{k,u}) = n^2 / \sin^2(\Psi_C)$ ,  $n$  为折射系数。

VLC 中信号光强随驱动电流  $I$  在其工作范围内  $I \in [I_l, I_h]$  线性变化,  $I_l$  和  $I_h$  分别为该范围的最小值和最大值。  $I_{sw} = I_h - I_l$  为发光二极管(Light Emitting Diode, LED)的最大电流摆幅,  $i_{sw}^u$  为第  $u$  架 UAV 通信所需的电流摆幅,电流摆幅向量为  $i_{sw} = (i_{sw}^1, \dots, i_{sw}^U)$ 。

采用  $M$  进制脉冲幅度调制( $M$ -ary Pulse Amplitude Modulation,  $M$ -PAM),第  $u$  架 UAV 的信号星座集合可表示为  $S = \{I_l, \dots, I_l + m\Delta i^u, \dots, I_l + (M-1)\Delta i^u\}$ ,式中  $M$  为调制阶数; $\Delta i^u = i_{sw}^u / (M-1)$  为相邻星座点之间的幅度间隔。

第  $u$  架 UAV 在照明与通信模式下消耗的电功率  $P_{T,u}$  可表示为

$$P_{T,u} = \frac{1}{M} \sum_{m=0}^{M-1} (I_l + m \cdot \Delta i^u)^2 = I_l^2 + I_l i_{sw}^u + \left[ \frac{1}{3} + \frac{1}{6(M-1)} \right] (i_{sw}^u)^2, \quad (2)$$

式中: $P_{T,u}$  为每架 UAV 用于通信的电功率向量  $p = (P_{T,1}, \dots, P_{T,U})$  的分量。UAV 发射光信号时,其照明强度由平均驱动电流  $I_b = I_l + i_{sw}^u / 2$  决定,该电流基于  $M$ -PAM 中星座点等概率出现的假设推导得出。为确保充足的照明,  $I_b$  必须保持恒定并满足约束条件  $I_b \geq \beta_u I_{sw}$ ,式中  $\beta_u$  为第  $u$  架 UAV 的照明系数。通信消耗的电功率  $P_{C,u} = P_{T,u} - I_b^2 = \left[ \frac{1}{12} + \frac{1}{6(M-1)} \right] (i_{sw}^u)^2$ ,式中,  $I_b^2$  为 UAV 在照明模式下消耗的电功率。

第  $u$  架 UAV 服务的第  $k$  个用户的信干噪比  $\gamma_{k,u}$  可表示为

$$\gamma_{k,u} = \frac{\eta h_{k,u} P_{C,u}}{\eta \sum_{l \in U, l \neq u} h_{k,l} P_{C,l} + \sigma^2}, \quad (3)$$

式中: $\eta$  为光电探测器的响应度; $\sigma^2$  为加性高斯白噪

声的方差,其他 UAV 信号被视为干扰。

VLC 信道容量可通过容量下界公式近似代替<sup>[8]</sup>,用户  $k$  的可达信道传输速率  $r_{k,u}$  可表示为

$$r_{k,u} = \frac{1}{2} B_b \log_2 \left( 1 + \frac{e}{2\pi} \gamma_{k,u} \right), \quad (4)$$

式中:  $e$  为自然对数;  $B_b$  为系统带宽。

一架 UAV 服务的多个用户采用时分多址接入<sup>[9]</sup>。时隙  $T$  内,设第  $u$  架 UAV 分配给第  $k$  个用户的服务时间为  $t_{k,u}$ ,  $\alpha_{k,u} = t_{k,u}/T$  为 UAV 服务用户的时间占比,  $\alpha_{k,u} = 0$  表示第  $k$  个用户未被第  $u$  架 UAV 服务;  $\alpha_{k,u} = 1$ , 则表示当前时隙第  $u$  架 UAV 仅服务第  $k$  个用户。

用户的随机移动建模为简化的随机游走模型。当用户移动到新位置  $\hat{\mathbf{w}}_k = (\hat{w}_{k,x}, \hat{w}_{k,y}, \hat{w}_{k,z})$ , UAV 需要移动至新的最佳位置  $\hat{\mathbf{Q}} = [\hat{q}_1, \hat{q}_2, \dots, \hat{q}_U]^T \in \mathbb{R}^{3 \times U}$ ,  $(\cdot)^T$  为矩阵转置。将 UAV 飞行调度矩阵定义为  $\mathbf{S} = [s_{i,j}] \in \mathbb{R}^{U \times U}$ ,  $s_{i,j} \in \{0, 1\}$ ,  $s_{i,j} = 1$  表示将原位置  $i$  的 UAV 调度至新位置  $j$ 。

当 UAV 以速度  $v$  飞行时,其在轨迹上的推进功率消耗  $P_p(v)$  可表示为<sup>[10]</sup>

$$P_p(v) = P_0 \left( 1 + \frac{3v^2}{U_{\text{tip}}^2} \right) + P_{\text{rot}} \left( \sqrt{1 + \frac{v^4}{4v_0^4}} - \frac{v^2}{2v_0^2} \right)^{1/2} + P_v \quad (5)$$

式中:  $P_0$  为悬停时的叶型阻力功率;  $U_{\text{tip}}$  为桨叶叶尖速度;  $P_{\text{rot}}$  为克服旋翼诱导阻力所需的功率;  $v_0$  为悬停时旋翼的平均诱导速度;  $P_v$  为克服除旋翼外气动部件产生阻力所需的功率。

针对随机移动用户场景<sup>[11]</sup>下 VLC 辅助的 UAV 网络,本文将轨迹优化分解为位置优化(P1)和轨迹规划(P2)的两个子问题。

子问题 P1: 为了在有限的机载能量下最大化系统的通信收益,本文将 EE 作为子问题 P1 的优化目标,其表示为系统总吞吐量  $R_{\text{tot}}$  与总能耗  $P_{\text{tot}}$  的比值  $R_{\text{tot}}/P_{\text{tot}}$ , 式中  $R_{\text{tot}} = \sum_{u=1}^U \sum_{k=1}^K \alpha_{k,u} r_{k,u}$  为系统可达和速率;  $P_{\text{tot}} = \sum_{u=1}^U P_{\text{T},u}$  为照明与通信模式下消耗的功率总和。优化子问题 P1 可表示为

$$(P1) \max_{i_{\text{min}}, \alpha_{k,u}, \mathbf{Q}} \frac{R_{\text{tot}}}{P_{\text{tot}}} = \frac{\sum_{u=1}^U \sum_{k=1}^K \alpha_{k,u} r_{k,u}}{\sum_{u=1}^U P_{\text{T},u}},$$

$$\begin{aligned} \text{s.t.} \quad & \sum_{u=1}^U \alpha_{k,u} \leq 1, \forall k \in K, \\ & \sum_{k=1}^K \alpha_{k,u} \leq 1, \forall u \in U, \\ & \alpha_{k,u} \geq 0, \forall k \in K, \forall u \in U, \quad (6) \\ & 2(\beta_u I_{\text{sw}} - I_l) \leq i_{\text{sw}}^u \leq I_{\text{sw}}, \\ & h_L \leq q_{u,z} \leq h_H, \\ & \|\mathbf{q}_i - \mathbf{q}_j\| \geq d_{\text{min}}, \end{aligned}$$

式中:  $2(\beta_u I_{\text{sw}} - I_l)$  为由照明约束  $I_b \geq \beta_u I_{\text{sw}}$  决定的照明最小阈值,功率分配通过优化 UAV 的电流摆幅  $I_{\text{sw}}$  确定;  $d_{\text{min}}$  为避免碰撞而设定的任意两架 UAV 之间的最小距离;  $i, j$  分别为不同 UAV 的索引。

子问题 P2: 当地面用户随机移动时, UAV 需要移动至新的最佳位置  $\hat{\mathbf{Q}}$ 。通过优化调度矩阵  $\mathbf{S}$  来最小化飞行中的推进能耗,假设 UAV 以速度  $v_m$  匀速飞行,  $P_p/v_m$  为单位距离消耗的能量,  $\|\hat{\mathbf{Q}}\mathbf{S}^T - \mathbf{Q}\|_F$  为 UAV 的空间距离,最小化轨迹能耗问题可表示为

$$(P2) \min_{\mathbf{S}} \frac{P_p}{v_m} \|\hat{\mathbf{Q}}\mathbf{S}^T - \mathbf{Q}\|_F, \quad (7)$$

$$\text{s.t.} \quad s_{i,j} \in \{0, 1\}, \mathbf{1}_U^T \cdot \mathbf{S} = \mathbf{1}_U^T, \mathbf{S} \cdot \mathbf{1}_U = \mathbf{1}_U,$$

式中,  $\mathbf{1}_U$  为长度为  $U$  的全 1 列向量;  $\|\cdot\|_F$  为 Frobenius 范数, Frobenius 范数可以作为对整体 UAV 移动代价的度量,同时, Frobenius 范数的严格凸性确保子问题 P2 存在唯一的全局最优解。

## 2 轨迹规划算法

本文提出了一种 UAV 的动态轨迹优化方法,如图 2 所示。给定 UAV 位置  $\mathbf{Q}$  和用户位置矩阵  $\mathbf{W}$ , 在用户随机移动过程中,每个时间间隔先更新用户位置  $\hat{\mathbf{W}}$ , 然后通过软演员-评论家(Soft Actor-Critic, SAC)算法求解子问题 P1 得到 UAV 的最优位置  $\hat{\mathbf{Q}}$ , 再通过子问题 P2 得到调度矩阵  $\mathbf{S}$ , 驱动 UAV 飞行到新位置。如此循环,在用户移动场景下持续保证高效接入。

位置优化问题 P1 属于非凸的分式规划问题,传统的凸优化方法难以对其直接求解。本文将等效建模为马尔可夫决策过程,并采用图 3 所述的 SAC 算法进行求解。马尔可夫决策过程由元组  $(\mathcal{S}, \mathcal{A}, R, \zeta, \pi)$  定义,其分别为状态空间、动作空间、奖励函数、折扣因子和策略。其中,状态空间  $\mathcal{S} = (\mathbf{W}, h_{k,u})$ ; 动作空间  $\mathcal{A} = (\alpha_{k,u}, \mathbf{Q}, i_{\text{sw}})$ ; 奖励函数

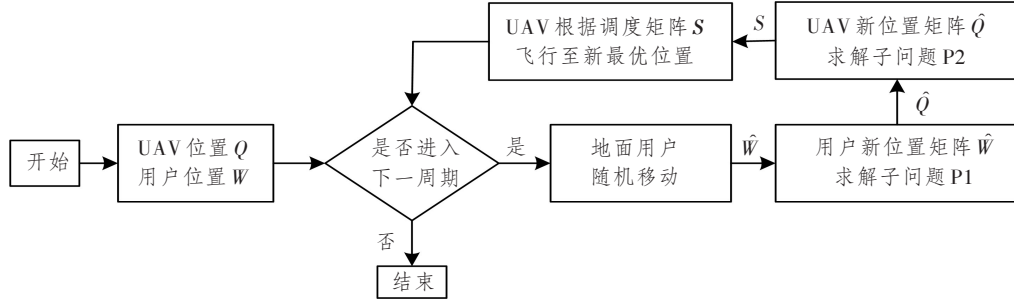
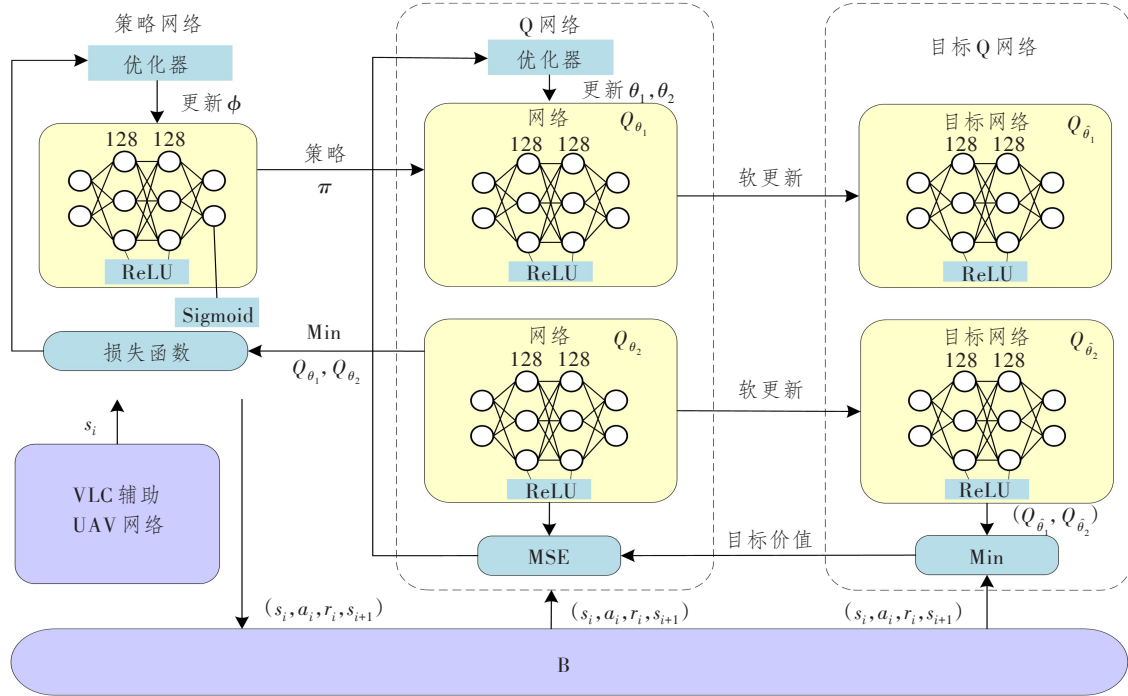


图2 VLC辅助的UAV网络的动态轨迹优化算法

Figure 2 Dynamic trajectory optimization algorithm for VLC-assisted UAV networks



注:MSE为均方误差;Min为最小值操作。

图3 求解UAV最优位置的SAC算法图(问题P1)

Figure 3 SAC algorithm diagram for solving UAVs' optimal position in P1

$$R = \left( \sum_{u=1}^U \sum_{k=1}^K \alpha_{k,u} r_{k,u} / \sum_{n=1}^U P_{T,u} \right) \circ$$

SAC架构由一个参数为 $\phi$ 的策略网络、两个参数为 $\theta_1$ 和 $\theta_2$ 的动作价值网络(简称Q网络)以及两个参数为 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 的目标Q网络组成。SAC旨在最大化期望累积奖励的同时,最大化策略在 $\rho_\pi$ 上的期望熵,其目标函数 $J(\pi)$ 可表示为

$$J(\pi) = \sum_{i=0}^{\infty} \mathbb{E}_{(s_i, a_i) \sim \rho_\pi} \left[ R(s_i, a_i) + \xi \mathcal{H}(\pi(\cdot | s_i)) \right], \quad (8)$$

式中: $\rho_\pi$ 为基于当前策略 $\pi$ ,状态动作对 $(s_i, a_i)$ 的概率分布; $\mathbb{E}$ 为期望运算; $\mathcal{H}(\pi(\cdot | s_i))$ 为策略的熵; $\xi$ 为温度参数,用于决定熵与奖励之间的权衡。每个Q网络均配有一个对应的目标Q网络,用于计算的目标Q值 $Q_{\text{tar}}$ ,可表示为

$$Q_{\text{tar}} = R + \zeta \left[ \min_{j=1,2} Q_{\hat{\theta}_j}(s_{i+1}, a_{i+1}) - \xi \ln \pi_\phi(a_{i+1} | s_{i+1}) \right], \quad (9)$$

式中, $\pi_\phi$ 为参数为 $\phi$ 的策略网络。对两个目标Q网络取最小值能够有效缓解过估计偏差。Q网络根据下式进行更新

$$\mathcal{L}(\theta_j) = \mathbb{E}_{(s,a) \in B} \left[ \left( Q_{\theta_j}(s, a) - Q_{\text{tar}} \right)^2 \right], \quad j=1, 2, \quad (10)$$

式中, $\mathcal{L}(\theta_j)$ 为第 $j$ 个Q网络的损失函数; $B$ 为经验回放池; $Q_{\theta_j}$ 为参数为 $\theta_j$ 的Q网络。由于服从高斯分布的动作采样过程是不可微的,为解决此问题,采用了重参数化技巧对动作 $\tilde{a}_i$ 进行采样,并依据如下损失函数更新策略网络:

$$J_\pi(\phi) = \frac{1}{N} \sum_{i=1}^N \left[ \xi \ln \pi_\phi(\tilde{a}_i | s_i) - \min_{j=1,2} Q_{\theta_j}(s_i, \tilde{a}_i) \right]. \quad (11)$$

求解 UAV 最优位置(问题 P1)的 SAC 算法如图 3 所示。首先对策略网络  $\pi_\phi$  和 Q 网络及  $Q_{\theta_1}, Q_{\theta_2}$  其对应的目标网络  $Q_{\hat{\theta}_1}, Q_{\hat{\theta}_2}$  进行参数初始化,设定学习率  $l_a, l_c$ ; 软更新因子  $\tau$ ; 以及温度参数  $\xi$ , 经验回放池 B。在每一个训练回合中,初始化 UAV 的环境状态,在随后的每个时间步内,智能体根据当前策略网络  $\pi_\phi$  采样执行动作  $a_i$ 。环境反馈奖励  $r_i$  并转移至下一状态  $s_{i+1}$ 。将样本数据  $(s_i, a_i, r_i, s_{i+1})$  存入 B 中。在每次参数训练迭代中,从 B 中随机抽取  $N$  步数据,通过以下步骤训练更新网络参数:①目标值计算:根据式(9)计算目标动作价值;②参数优化:基于式(10)和式(11)的损失函数,利用梯度下降法更新 Q 网络参数  $\theta$  和策略网络参数  $\phi$ ;③温度调节:通过计算对数策略概率与目标熵  $\mathcal{H}_0$  的差值,自动调整温度参数  $\xi = \xi - \lambda_\xi \nabla_\xi \{\xi [\ln \pi_\phi(a_i|s_i) + \mathcal{H}_0]\}$ , 以平衡探索与利用,其中  $\nabla_\xi$  为温度参数的梯度,  $\lambda_\xi$  为温度参数的更新系数;④软更新:采用软更新机制对目标 Q 网络参数  $\hat{\theta}$  进行平滑处理,软更新机制为  $\hat{\theta}_j =$

$$\tau\theta_j + (1-\tau)\hat{\theta}_j \quad (j=1, 2)。$$

上述网络训练完成以后得到一个最优策略,该策略网络的输入为用户位置,信道矩阵,输出为最优的 UAV 位置  $\hat{Q}$ , 用户接入关联  $\hat{a}_{k,u}$  及每个 UAV 的功率分配  $\hat{i}_{sw}$ 。

根据最优 UAV 位置  $\hat{Q}$  求解轨迹规划子问题 P2, 只需将整数约束松弛为  $0 \leq s_{i,j} \leq 1$ , 轨迹规划子问题 P2 可转化为凸优化问题, 因为其目标函数是凸函数且约束条件为线性的, 凸优化问题可直接利用 python 中的凸优化工具包 CVXPY 高效求解。

### 3 实验结果分析

VLC 辅助的 UAV 网络部署在一个  $200\text{ m} \times 200\text{ m} \times 5\text{ m}$  的室外区域内, 包含 4 架 UAV 和 40 个地面用户。用户随机分布在高度为  $0.5\text{ m}$  的水平面上。半数用户沿随机方向  $(0, 2\pi]$  并以随机速度  $[0, 5]\text{ m/s}$  移动。4 架 UAV 的初始位置均匀分布在高度为  $2.5\text{ m}$  的水平面上。其余系统参数如表 1 所示。

表 1 系统参数

Table 1 System parameters

参数	数值	参数	数值
最大电流摆幅 $I_{sw}/\text{A}$	1	光电探测器响应度 $\eta/\text{A/W}$	0.4
最小驱动电流 $I/\text{A}$	0.1	光电探测器接收面积 $A/\text{cm}^2$	1
UAV 最小高度 $h_l/\text{m}$	2.5	单边功率谱密度 $\Lambda^2/\text{Hz}$	$7.02 \times 10^{-23}$
UAV 最大高度 $h_H/\text{m}$	5	PAM 调制阶数 $M$	8
朗伯辐射阶数 $m$	1	推进功率 $P_p/\text{W}$	31.06
调制带宽 $\text{MHz}$	20	光学聚光器增益 $g(\psi_{k,u})$	2.25
UAV 飞行速度 $v/\text{m/s}$	5	经验回放池大小	$1 \times 10^6$
软更新因子 $\tau$	0.05	最小训练批次	512
时隙长度 $\text{s}$	1	UAV 最小距离约束 $d_{\min}$	5

在 SAC 算法中,所有网络均为包含两个隐藏层的全连接多层感知机。每层包含 128 个神经元, 并采用 ReLU 激活函数。策略网络利用 Sigmoid 激活函数来生成动作, 而 Q 网络则包含一个线性输出层用于 Q 值估计。策略网络和 Q 网络均使用 Adam 优化器进行训练。

由图 4 可知,不同用户数量情况下 SAC 算法均能快速收敛,这得益于 SAC 算法引入了熵正则化增加了系统的探索性,使得策略更加随机,降低了策略陷入局部最优解的可能性。并且在上述网络中,随着用户数量的增加,网络最大 EE 有所下降,这是由于有限资源的可分配量减少,且 UAV 的能耗增加。

图 5 所示为 40 个用户场景下,由 SAC 算法求

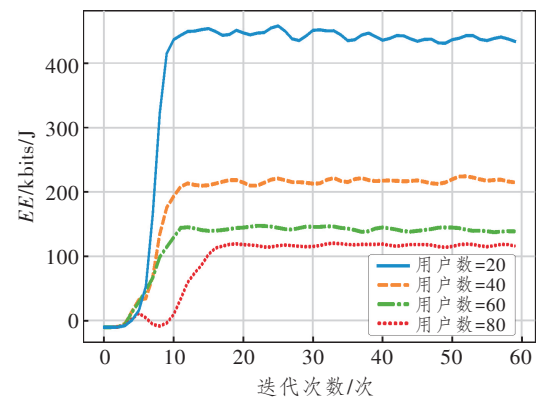


图 4 不同用户数的 SAC 算法收敛情况

Figure 4 Convergence of the SAC algorithm with different numbers of users

解的 UAV 与用户关联的情况。每架 UAV 位于其关联用户群的平均位置的上方。值得注意的是,不同 UAV 的悬停高度存在显著差异,主要受相邻 UAV 间干扰和负载均衡需求的影响。

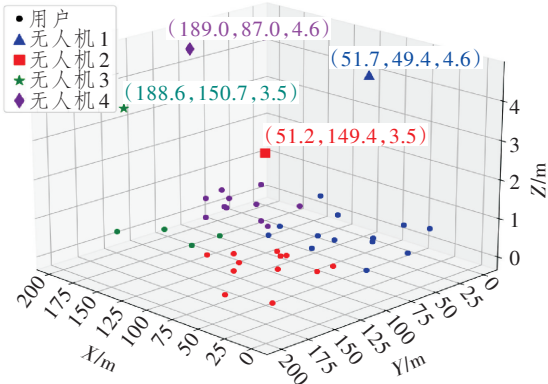


图5 用户与 UAV 的关联关系图

Figure 5 The diagram of User-UAV association

图6所示为不同调制阶数下 SAC 算法与深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)算法、粒子群算法以及 UAV 位置固定基准下网络的 EE 表现<sup>[12]</sup>。结果表明,在相同调制阶数下,SAC 算法的 EE 表现始终优于 DDPG 算法和粒子群优化算法。而随着调制阶数的增加,3 种优化算法之间的 EE 差距逐渐缩小。另外,SAC 算法相比于 DDPG 算法对超参数的敏感度更低,从而在实际训练中表现出更稳定的性能。

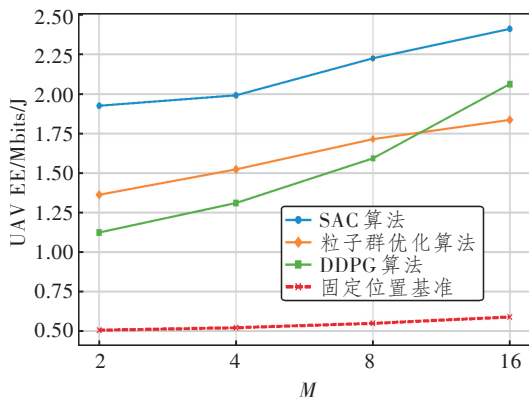
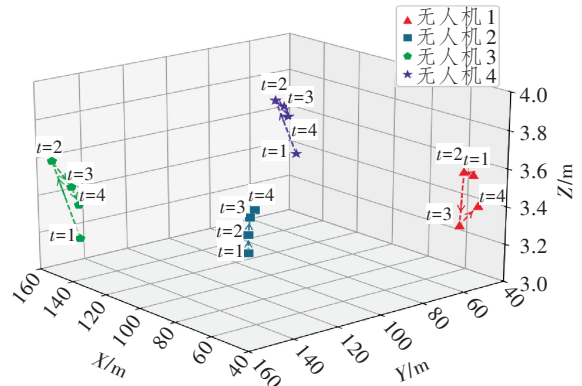


图6 不同调制阶数下 SAC 算法与其他算法 EE 对比图

Figure 6 EE comparison between the SAC algorithm under different modulation orders and other algorithms

图7所示为半数用户处于动态随机移动下的 UAV 三维轨迹优化结果。观察轨迹可知,UAV 主要通过局部微小位置的调整来适应用户的移动,而非进行大范围的移动,此策略有效地避免了长距离飞行,从而节省了推进能耗。



注:t为时间步。

图7 20个用户随机移动后 UAV 飞行轨迹图  
Figure 7 UAV flight trajectories with 20 randomly moving users

## 4 结束语

针对地面用户随机移动及 UAV 机载能量有限的挑战,本文提出了一种最大化系统 EE 的轨迹优化迭代算法。该算法通过联合优化用户关联、UAV 位置和功率分配以提升网络 EE,进一步以最小化推进功耗为目标求解 UAV 最优飞行轨迹,在降低推进功耗的同时实现了网络负载均衡,有效延长了 UAV 的续航时间。由于 UAV 机载算力限制和环境实时需求,我们将进一步探索“离线集中训练、在线分布式执行”的部署方案,降低算法在线推理延迟,以推进该研究在实际 UAV 中的落地应用。

### 参考文献:

- [1] Pham Q V, Huynh-The T, Alazab M, et al. Sum-Rate Maximization for UAV-Assisted Visible Light Communications Using NOMA: Swarm Intelligence Meets Machine Learning[J]. IEEE Internet of Things Journal, 2020, 7(10): 10375–10387.
- [2] 闫红强,江明.跨水面VLC链路对准与信号检测技术研究[J].光通信研究,2025(4):250001.  
Yan H Q, Jiang M. Research on Link Alignment and Signal Detection Technologies for Cross-Water Visible Light Communication[J]. Study on Optical Communications, 2025(4): 250001.
- [3] Wang Y, Chen M, Yang Z, et al. Deep Learning for Optimal Deployment of UAVs with Visible Light Communications[J]. IEEE Transactions on Wireless Communications, 2020, 19(11): 7049–7063.
- [4] Ibraiwish H, Eltokhey M W, Alouini M S. Energy Efficient Deployment of VLC-Enabled UAV Using Particle Swarm Optimization[J]. IEEE Open Journal of the Communications Society, 2024, 5: 553–565.

- [5] Das A, Panda K G, Kumar M L N, et al. 3-D Placement Strategy for VLC Enabled UAV Network with Guaranteed QoS[C]//2022 IEEE 96th Vehicular Technology Conference (VTC2022 - Fall). London, UK: IEEE, 2022: 10012922.
- [6] Anwar D N, Peer M, Lata K, et al. 3-D Deployment of VLC Enabled UAV Networks with Energy and User Mobility Awareness[J]. IEEE Transactions on Green Communications and Networking, 2022, 6(4): 1972–1989.
- [7] Wang J B, Hu Q S, Wang J, et al. Tight Bounds on Channel Capacity for Dimmable Visible Light Communications[J]. Journal of Lightwave Technology, 2013, 31(23): 3771–3779.
- [8] Jiang R, Wang Q, Haas H, et al. Joint User Association and Power Allocation for Cell-Free Visible Light Communication Networks[J]. IEEE Journal on Selected Areas in Communications, 2018, 36(1): 136–148.
- [9] Zeng Y, Xu J, Zhang R. Energy Minimization for Wireless Communication with Rotary-Wing UAV[J]. IEEE Transactions on Wireless Communications, 2019, 18(4): 2329–2345.
- [10] Camp T, Boleng J, Davies V. A Survey of Mobility Models for Ad Hoc Network Research [J]. Wireless Communications and Mobile Computing, 2002, 2(5): 483–502.
- [11] Qi M, Zheng H. Temporal-Constrained DDPG-based Path Planning for UAV in Dynamic Environment [J]. IEEE Sensors Journal, 2026, 26(3): 4301–4312.