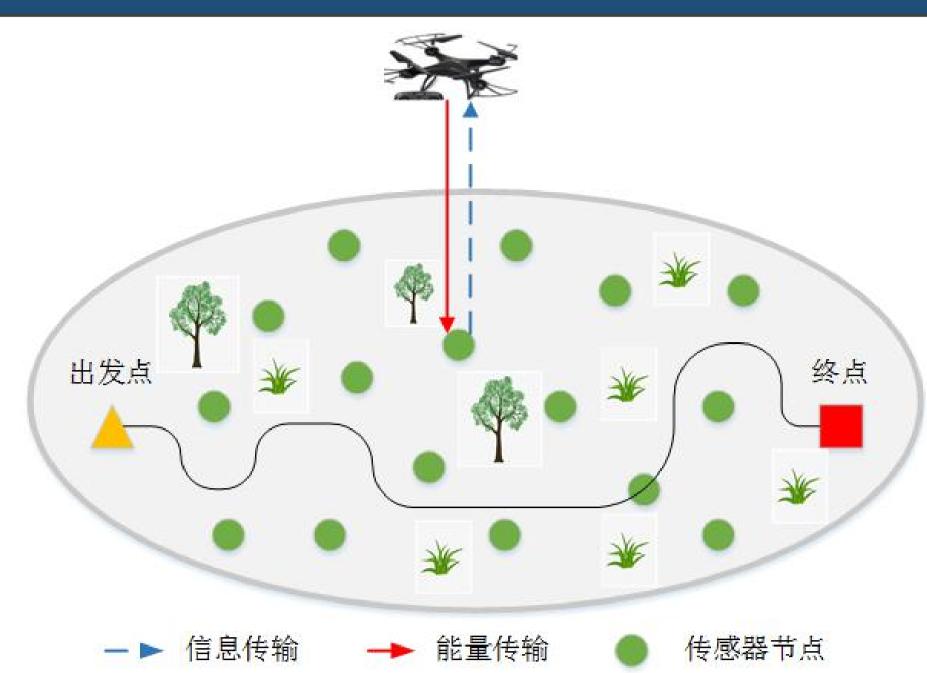
CCML 2021 第十八屆中国机器学习会议The 18th China Conference on Machine Learning

2021年8月6-8日 中国•长沙 AUG 6-8, 2021 CHANGSHA, CHINA

1.研究问题及难点



➤ 研究问题:

考虑到无人机辅助无线供能网络中不同用户对AoI的需 求不同,本章研究了在信息年龄受限下,以无人机能 耗为指标的无人机辅助无线供能网络设计。

▶ 问题难点:

由于所建立优化问题非凸,且问题中的信息年龄AoI、 能量、加速度以及转角变量相互耦合,传统方法难以 有效求解。

2.解决思路及方法

▶ 思路描述:

构建优化问题,将该问题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无人机控制算法, 并设计了相应的状态空间、动作空间及奖励函数。

优化问题构建

 $(\mathbf{P}_1) \min_{\{a,o,c,T\}} E_{\mathbf{u}}^{\mathrm{fly}}$

s.t. $A_k(N) < (N+1)\delta, k \in \{1, 2, ..., K\},$

 $\frac{1}{N} \frac{1}{K} \sum_{n=1}^{N} \sum_{k=1}^{K} A_k(n) \leq A_{\ell},$

 $q(0) = q_0, q(N) = q_F$

 $a(n) \in [a_{\min}, a_{\max}],$

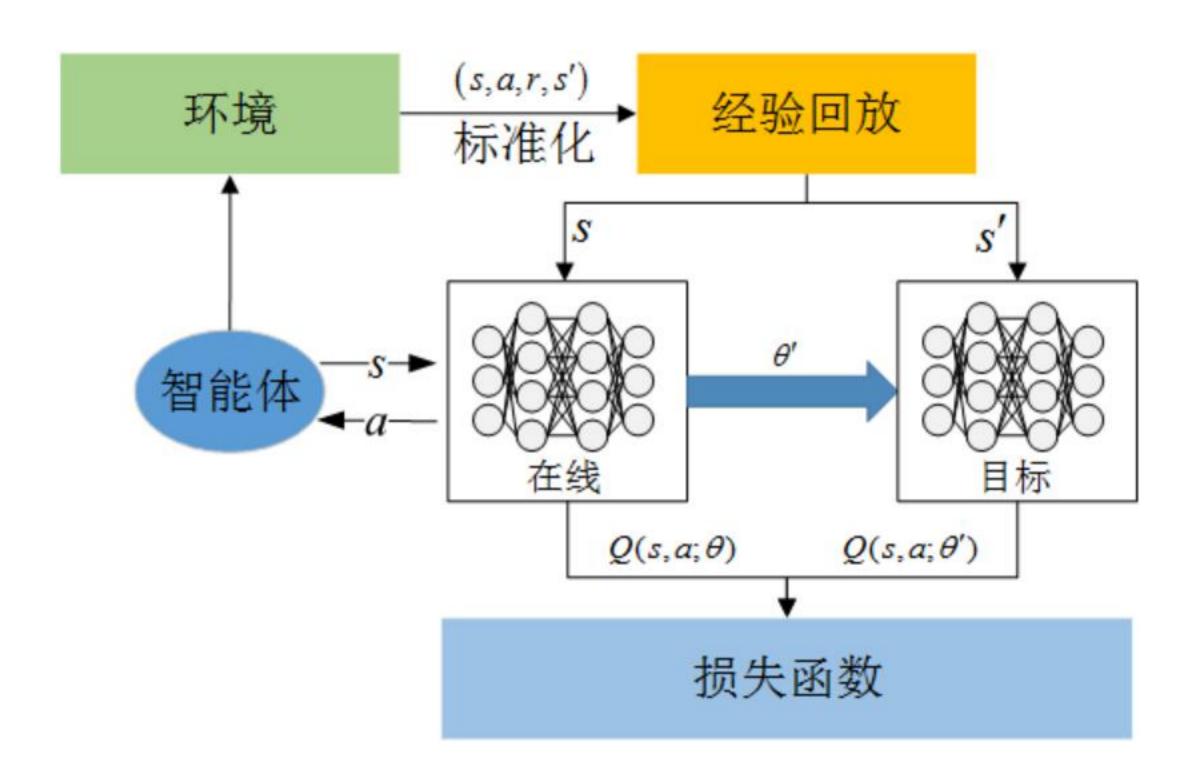
 $o(n) \in [o_{\min}, o_{\max}],$

 $v(n) \in [v_{\min}, v_{\max}],$

构建优化问题,旨在在 信息年龄受限的约束下, 最小化无人机能耗。

算法框架设计

下图为基于DQN的无人机控制算法框架图。首先,将 无人机视为一个智能体,将在空中的飞行运动,与传 感器节点的交互如信息能量传输等视为环境。在每个 训练周期中,智能体需要感知周边环境,即传感器节 点的电量、上传信息的AoI等信息状态,从而根据当前 环境决定下一时刻的动作,执行动作后,智能体便能 获得环境相应的反馈奖励,并继续观察下一时刻状态。



基于DQN的无人机控制算法框架

算法实体设计

状态空间 (State)

$$S_n = \left\{ \boldsymbol{B}, \boldsymbol{A}_{\mathbf{u}}, \boldsymbol{q}(n), v(n), p(n), E(n), n, d(n) \right\}$$

(电量AoI,无人机轨迹,速度,转角,能耗,时间,距离)

动作空间 (Action)

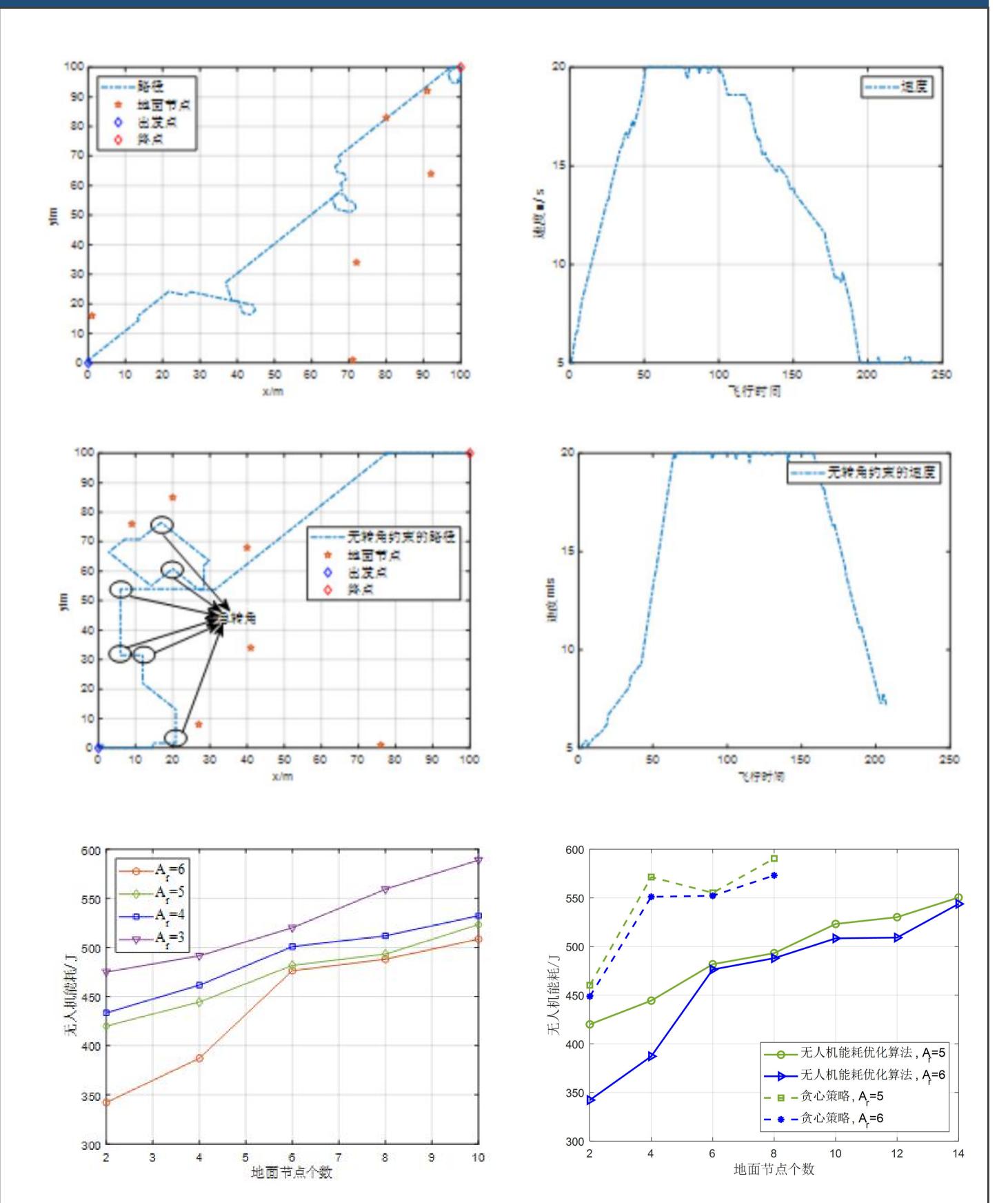
 $a_n = \{a(n), o(n), c(n)\} \ \langle$

加速度离散化 转角离散化 传输策略

奖励函数 (Reward)

$$r_{n} = \begin{cases} k_{1} - \frac{E_{u}^{fly}}{k_{2}}, & d(n) = 0 \text{ and } \frac{1}{N} \frac{1}{K} \sum_{k=1}^{K} A_{k} (N) \leq A_{r} \\ -\frac{d(n)}{k_{3}} - \frac{E(n)}{k_{2}}, \text{ others} \end{cases}$$

3. 研究成果及总结



- > 与传统的贪心策略相比,基于DQN的无人机能耗优化 算法相比传统的贪心策略能够降低约8%-30%的无人 机能耗。在仿真场景中,当传感器个数超过8个时,传 统的贪心策略很难求解,而提出的DQN算法依然能够 解决。
- > 无人机的能量消耗随着AoI的减小或传感器数量的增加 而增加。无人机在飞行过程中无人机会尽可能飞至传 感器节点附近进行信息收集和能量传输,且飞行速度 先迅速增加后下降。

投稿人 姓名: 刘玲珊

单位:北京交通大学