

# Enhancing UAV Communication Networks through DRL-Based Trajectory Optimization and Power Allocation

Quang Tuan Do, 김동완\*, 조성래

Chung-Ang Univ., \*Dong-A Univ.

[dqtuan@uclab.re.kr](mailto:dqtuan@uclab.re.kr), [dongwankim@dau.ac.kr](mailto:dongwankim@dau.ac.kr), [srcho@cau.ac.kr](mailto:srcho@cau.ac.kr)

## Abstract

Unmanned Aerial Vehicles (UAVs) offer flexible and high-performance communication services, but their limited energy resources necessitate optimized operational strategies. This paper presents a novel approach using Deep Reinforcement Learning (DRL) to optimize UAV trajectory design and power allocation. Specifically, we employ the Deep Deterministic Policy Gradient (DDPG) algorithm to maximize service time and downlink throughput in a UAV-based communication network. Simulation results demonstrate that our DRL-based method significantly improves energy efficiency and data transmission rates compared to traditional optimization techniques.

## I. Introduction

The rapid growth of mobile data demands and advancements in wireless communication technologies have positioned UAVs as a promising solution for providing flexible as well as high-performance communication services. UAVs can act as aerial base stations, enhancing coverage and ensuring reliable communication channels [1]. However, the inherent limitation of UAVs' energy resources poses significant challenges for prolonged operations. Optimizing UAV trajectory design and power allocation is crucial to maximizing their service time and downlink throughput. Traditional optimization methods often fall short in dynamic and complex environments. In contrast, Deep Reinforcement Learning (DRL) offers a powerful framework for addressing these challenges. This paper proposes a DRL-based approach using the Deep Deterministic Policy Gradient (DDPG) algorithm to optimize UAV operations in a communication network.

## II. System Model

We consider a rectangular area of size  $a \times b$  m<sup>2</sup> where  $K$  ground users (GUs) are randomly distributed. A single UAV serves as an aerial base station, providing communication coverage for the GUs. The UAV operates at a fixed altitude  $H$  and has a constant flight speed  $V$  m/s. Two charging stations are positioned at the start and end points of the UAV's trajectory, ensuring that the UAV begins and ends its mission fully charged.

The wireless channel between the UAV and GUs is modeled as a line-of-sight (LoS) channel [2], with the channel power gain given by:

$$h_k(u(t)) = \frac{\beta_0}{(x(t) - x_k)^2 + (y(t) - y_k)^2 + H^2} \quad (1)$$

where  $\beta_0$  is the channel power gain at the reference distance per meter,  $(x(t), y(t))$  is UAV's coordinate, and  $(x_k, y_k)$  is  $k$ -th GU's coordinates.

The power received by the  $k$ -th GU at time  $t$  is:

$$Q_k(u(t)) = \frac{\beta_0 P_k}{(x(t) - x_k(t))^2 + (y(t) - y_k)^2 + H^2} \quad (2)$$

where  $P_k$  is the transmission power allocated to the  $k$ -th GU.

The Signal-to-Noise Ratio (SNR) for the  $k$ -th GU is:

$$SNR_k = \frac{P_k h_k}{\sigma^2} \quad (3)$$

where  $\sigma^2$  is the noise power.

The energy consumption of the UAV comprises communication and propulsion energy. Given that propulsion energy is significantly higher [3], we focus on this aspect only. The propulsion power is modeled as [4]:

$$P(V) = P_0 \left( 1 + \frac{3V^2}{U_{tip}^2} \right) + \frac{P_i v_0}{V} + \frac{1}{2} d_0 \rho s A V^3 \quad (4)$$

where  $P_0, P_i$  represent blade profile and induced power in hovering state,  $U_{tip}$  is tip velocity,  $v_0$  is induced velocity,  $d_0$  is fuselage drag ratio,  $\rho$  is air density,  $s$  is rotor solidity, and  $A$  is rotor disc area.

## III. Problem Formulation

The objective is to maximize the total downlink throughput while ensuring efficient energy usage. The problem can be formulated as:

$$\max \sum_{k=1}^K \log_2(1 + SNR_k) \quad (5)$$

subject to:

1.  $E(t) \geq 0$
2.  $P_k \leq P_{max}$
3.  $V \leq V_{max}, x(t) \in [0, a], y(t) \in [0, b]$

## IV. DRL-Based Optimization Framework

The state, action and reward is defined as follows:

- **State:** The state  $s_t$  includes UAV's coordinate  $(x(t), y(t), H)$ , GUs' positions, and the remaining energy  $E(t)$ .
- **Action:** The action  $a_t$  involves the power allocation  $P_k$  to each GU, and the UAV's movement direction.
- **Reward:** The reward  $r_t$  is designed to maximize the throughput while minimizing energy consumption, defined as:

$$r_t = \sum_{k=1}^K \log_2(1 + SNR_k) + \lambda E(t) \quad (6)$$

where  $\lambda$  is the weight parameter balancing throughput and energy efficiency.

The DDPG algorithm is employed to solve the continuous control problem. The actor network selects actions, while the critic network evaluates them. The learning process involves:

1. **Initialization:** Initialize networks and replay buffer.
2. **Policy Update:** Update actor network and select optimal actions.
3. **Value Update:** Update critic network to evaluate action values.

The detailed training algorithm is shown in **Algorithm 1**.

---

**Algorithm 1:** DDPG-based Trajectory Design and Power Allocation for Energy-Efficient UAV Communication

---

1. **Initialize** experience replay buffer  $B$ , actor and critic networks
2. **For each** episode do:
3.     **Initialize** GUs location, initial and destination location of the UAV, and initial energy.
4.     **For each** time slot do:
5.         Select optimal action  $a_t$
6.         Observe reward  $r_t$  and next state  $s_{t+1}$
7.         Store  $(s_t, a_t, r_t, s_{t+1})$  tuple in buffer
8.     **End for**
9.     Randomly sample  $M$  mini-batches from  $B$
10.     Update actor and critic networks using back-propagation
11.     Soft update target networks using main networks' parameters.
12. **End for**

## V. Conclusion

This paper presents a novel approach for optimizing UAV communication networks using deep reinforcement learning. By leveraging the DDPG algorithm, the proposed method achieves superior performance in terms of throughput and energy efficiency. Future work will explore multi-UAV coordination and 3D trajectory optimization to further enhance network capabilities.

## ACKNOWLEDGMENT

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information

Technology Research Center) support programs (IITP-2024-RS-2022-00156353 and IITP-2024-RS-2023-00258639) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

## REFERENCES

- [1] Y. R. Z. a. T. J. L. Zeng, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Communications Magazine* 54.5, pp. 36-42, 2016.
- [2] Y. R. Z. a. T. J. L. Zeng, "Throughput maximization for UAV-enabled mobile relaying systems," *IEEE Transactions on Communications*, 2016.
- [3] Y. a. R. Z. Zeng, "Energy-efficient uav communication with trajectory optimization," *IEEE Transactions on Wireless Communications*, 2017.
- [4] Y. J. X. a. R. Z. Zeng, "Energy Minimization for Wireless Communication With Rotary-Wing UAV," *IEEE Transactions on Wireless Communications*, 2019.