# Research Progress: End-to-end Performance Optimization for Mixed FSO/RF-aided Vertical Networks: A DRL Approach

**Tinh Nguyen**

Computer Communications Lab.,
The University of Aizu

# Outline

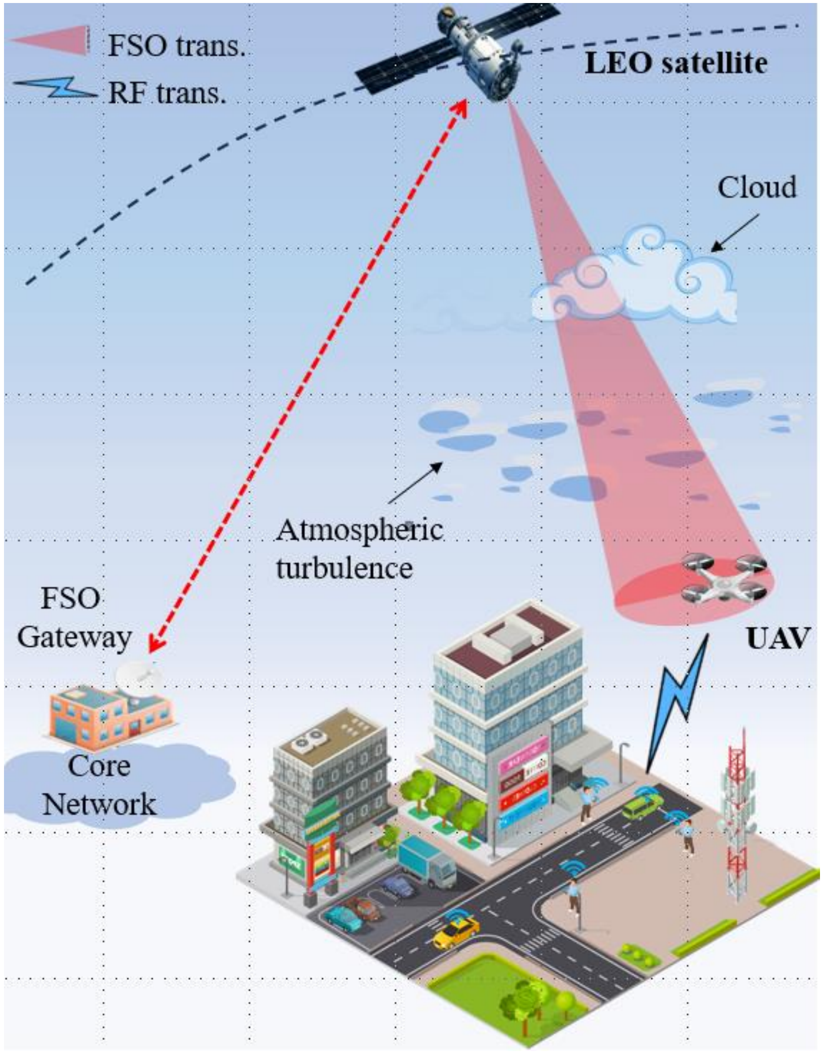
- Introduction
  - Aerospace Integrated Network
  - Critical Issues
  - Motivations
  - Literature Review and Goals of the Study
- System Description
  - System Model
  - Problem Formulation
- Deep Reinforcement Learning for UAV Deployment

# Introduction

# Aerospace Integrated Network

- Recently, ***free-space optics (FSO)*** is envisioned as a promising candidate for backhaul networks thanks to its ***extremely high-speed connection***
- ***Aerospace integrated network***, incorporating satellites, HAP, and UAV, can guarantee ***seamless and ubiquitous access services***, especially in remote, hotspot, or emergency areas.

*→ The aerospace integrated network jointly comprised of the UAV access and LEO satellite/HAP-based FSO backhaul is a promising framework for future 6G networks.*



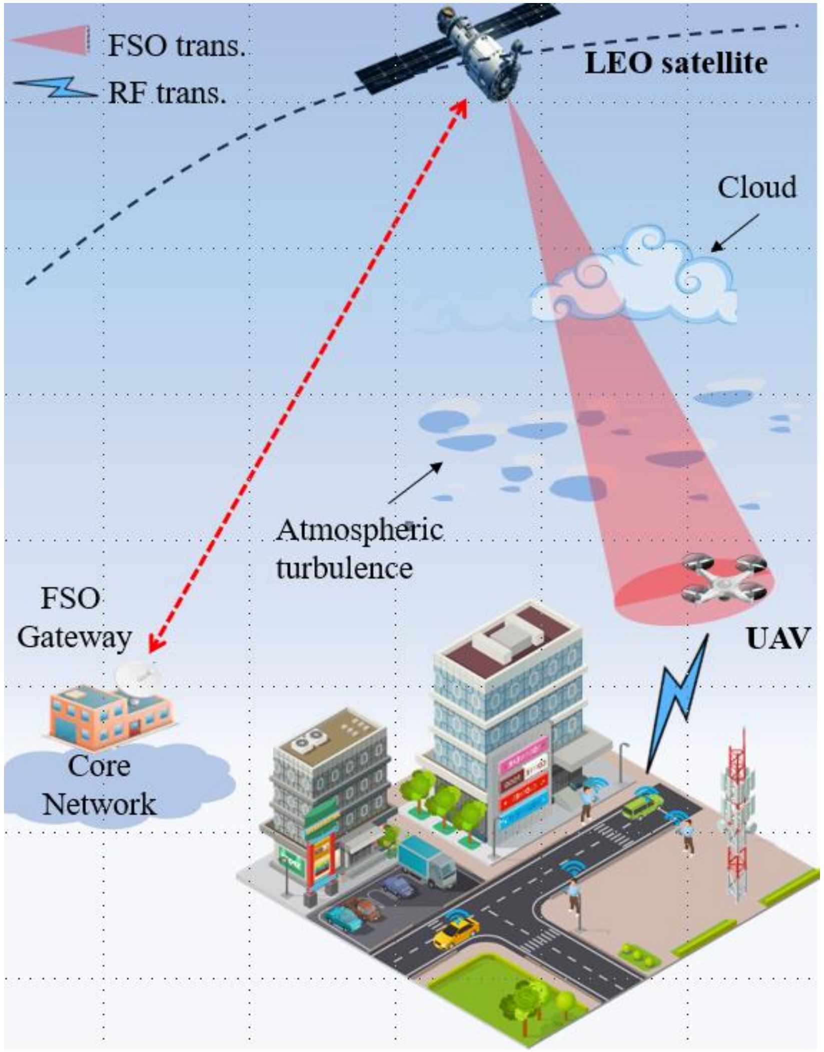
# Critical Issues

- Backhaul link:
  - ***Cloud attenuation:*** the liquid water particles in clouds cause the scattering phenomenon
  - ***Atmospheric turbulence:*** air pockets with different refractive indexes cause the scintillation effect
  - ***Pointing error:*** misalignment between the center of satellite beam footprint and that of the UAV detector

  → *Unstable channel* → *Limited capacity of the backhaul link*

- Access link:
  - ***Dynamic network:*** the GUs move over time

  → *Time-varying network topology*



# Motivations

- Aerospace integrated FSO backhaul links:
  - FSO channel conditions, e.g., atmospheric turbulence, cloud, and pointing error conditions, have a considerable impact on the link capacity
  - Limited backhaul capacity can be a bottleneck, which directly affects the end-to-end performance

  → *It is necessary to investigate the FSO channel conditions on the end-to-end performance*

- Dynamic environment:
  - The movement of ground users, time-varying conditions of backhaul link

  → *Flying platforms, e.g., UAVs, can adapt the spatial positions to maintain stable connections with LoS connectivity*

# Literature Review

- Some major papers that consider UAV placement for end-to-end network

| Ref. | Year | Journal | Title |
|---|---|---|---|
| 1 | 2021 | Trans. Veh. Technol. | Latency Aware 3D Placement and User Association in Drone-Assisted Heterogeneous Networks With FSO-Based Backhaul |
| 2 | 2023 | Trans. Wireless Commun. | Backhaul-Aware Drone Base Station Placement and Resource Management for FSO-Based Drone-Assisted Mobile Networks |
| 3 | 2022 | Trans. Veh. Technol. | Spectral-Efficient Network Design for High-Altitude Platform Station Networks with Mixed RF/FSO Systems |
| 4 | 2023 | Photon. J. | Outage Probability Analysis and Joint Optimization for UAV-aided FSO/RF Systems with Nonlinear Power Amplifiers |
| 5 | 2023 | Internet Things J. | Cooperative UAV Trajectory Design for Disaster Area Emergency Communications: A Multi-Agent PPO Method |

*Current studies:*

- *mainly focus on the access network and usually* ***ignore the impacts of FSO channel conditions***
- *mainly address the* ***UAV placement problem with static networks****, i.e., when there are changes in the environment, those algorithms need to start from the beginning*

# Goals of the Study

- GOAL: find the **optimal placement of the UAV** to **optimize the end-to-end throughput performance** under the constraints of (1) the unstable channel of the backhaul links and (2) the dynamic network of the access links

1. Backhaul link:

   - We thoroughly consider the FSO impairment factors, including atmospheric turbulence, pointing error, and cloud attenuation

2. Access link:

   - We implement Deep Reinforcement Learning (DRL) as an effective approach to handle the real-time UAV placement with dynamic network

# System Description

# System Model (1)

**Application:** to provide the internet connection to **rural/remote areas** or **temporary events** where **ground base stations are unavailable**

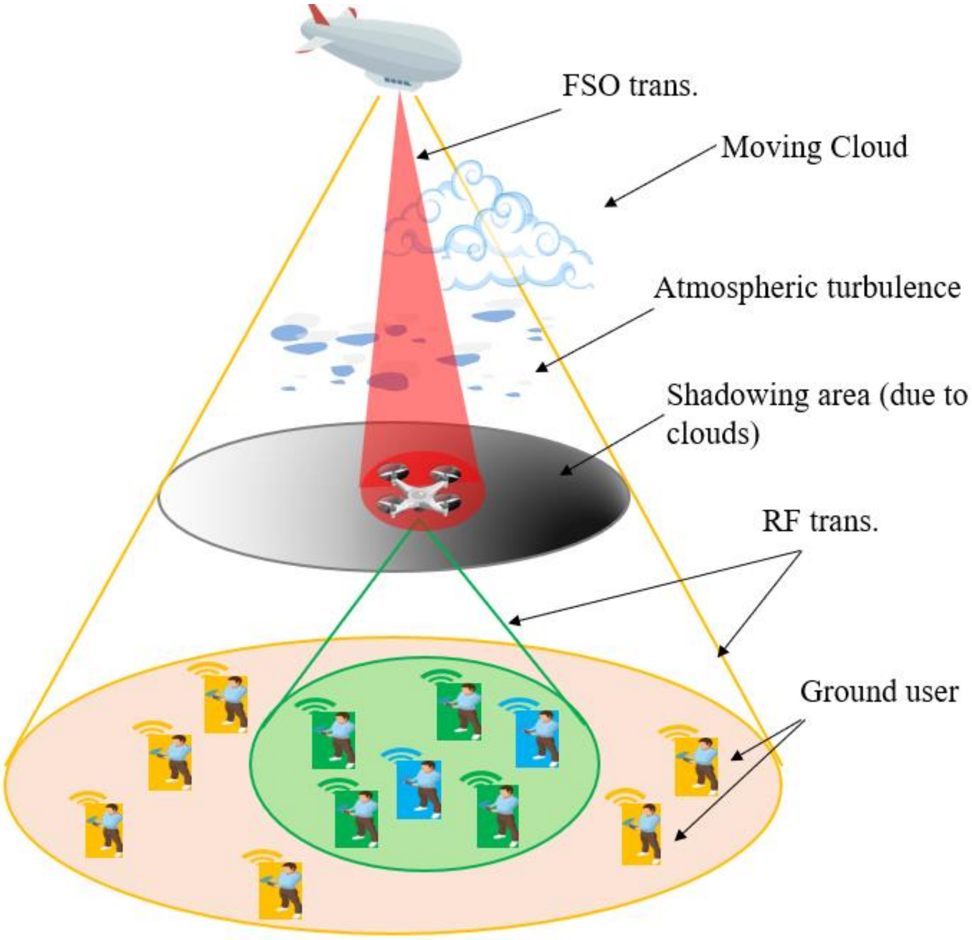
The end-to-end network scenario includes 2 main transmission links:

1. **Backhaul link:** from HAP to UAV

→ *FSO transmission is used for high-speed connection*

2. **Access link:** from UAV to GU and HAP to GU

→ *RF transmission is used*



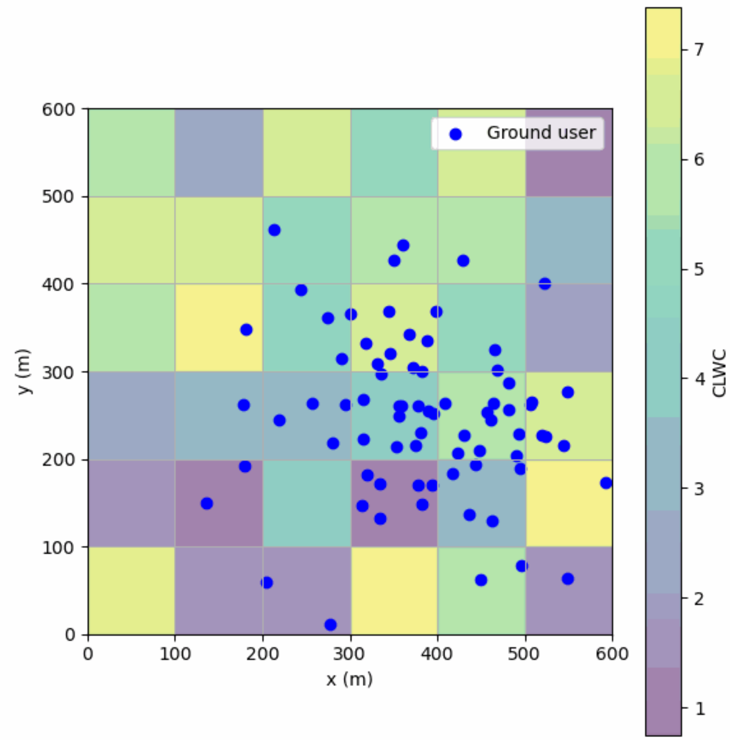
# System Model (2)

We consider an area of $600 \times 600\ m^2$

- Users
  - The users are normally distributed with the standard deviation of 100m. The mean is randomly chosen in the whole area
  - A part of users follow the Gauss-Markov mobility model, while others stay unchanged

- Cloud model
  - We consider moving cloud with heterogeneous clwc*
  - The shadowing area is modeled as a 6x6 grid that demonstrates different clwc
  - The cloud moves to the west with a velocity of 4 m/s

** : clwc (cloud liquid water content) - a measure of the total liquid water contained in a cloud in a vertical column of atmosphere (the less, the better)*



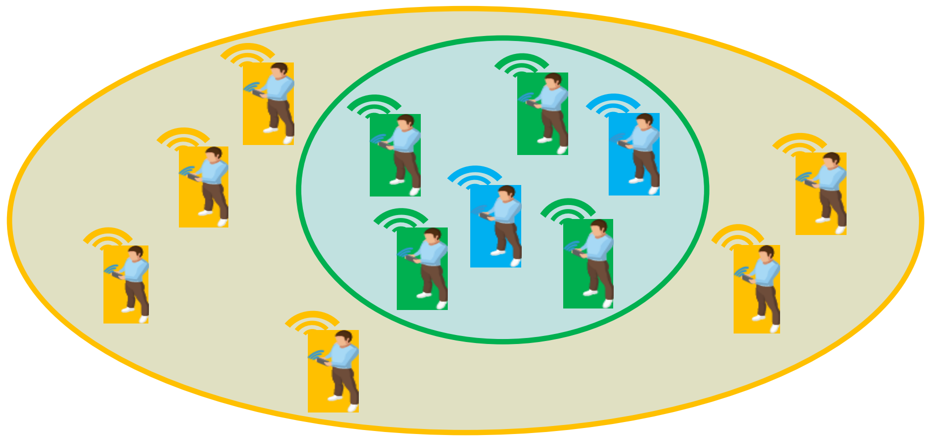
# Problem Formulation

 Users supported by UAV → **Group $U_1$**

 Users in the coverage of UAV but supported by HAP → **Group $U_2$**

 Other users supported by HAP → **Group $U_3$**

UAV coverage    HAP coverage



- The optimization problem ($\mathcal{P}$) is formulated as

$$\mathscr{P}: \quad \max_{(x_t, y_t)} \sum_{t=0}^{T-1} \sum_{u \in U} R_u(t)$$

Find optimal position of UAV to maximize total rate of all GUs

$$\text{s.t.} \quad C_1: \quad \sum_{u_1 \in U_1} R_{u_1}(t) \leq R_{\text{FSO}}(t)$$

Total rate of GUs supported by UAV must not exceed the backhaul capacity

$$C_2: R_{\text{FSO}}(t) = f\left(M_c(t)\right)$$

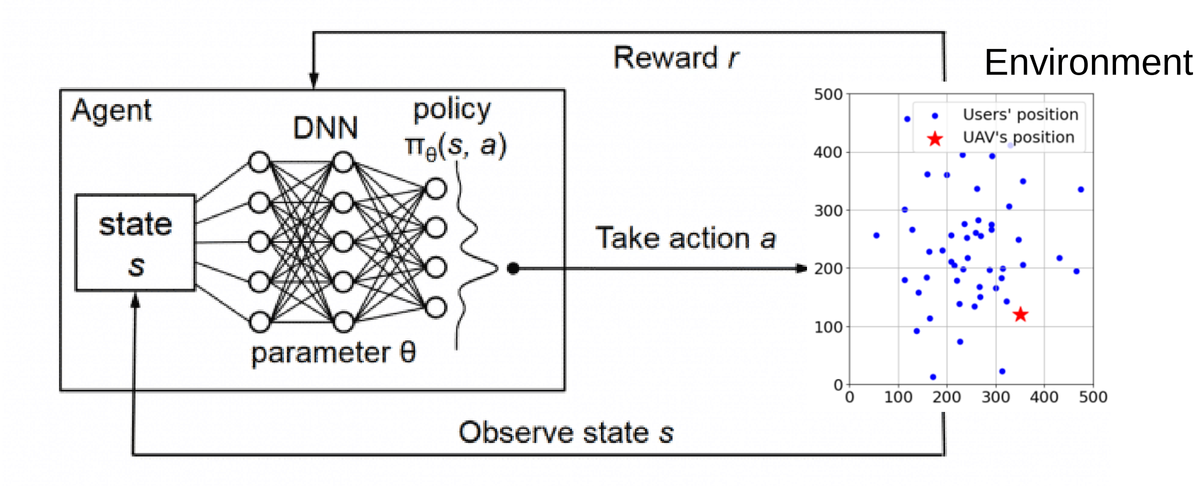
FSO capacity depends on the current clwc of clouds

$$C_3: \quad (x_{\min}, y_{\min}) \leq (x_t, y_t) \leq (x_{\max}, y_{\max})$$

Position constraints of the UAV

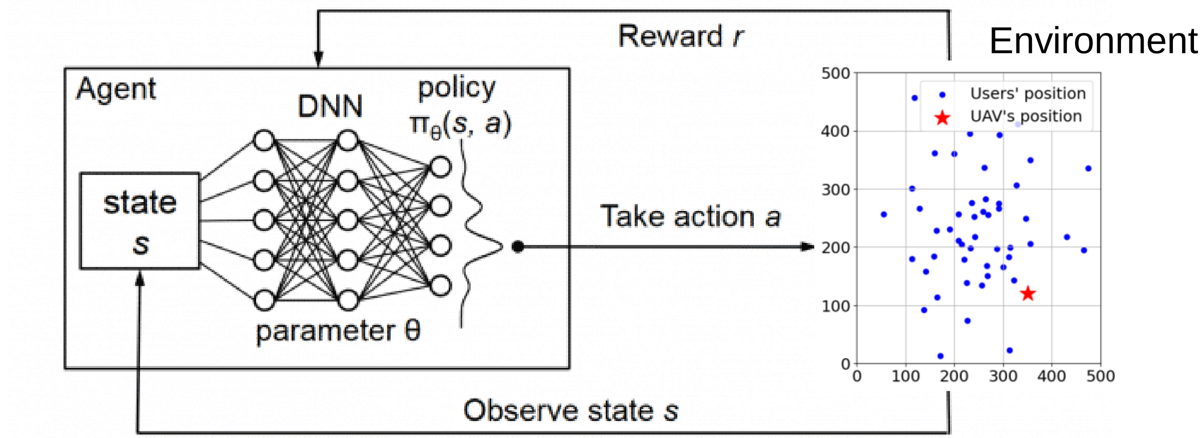
# DRL for UAV Deployment

# DRL Algorithm




- DRL Algorithm:
  - DRL: a process in which an agent learns to make decisions through trial and error
  - The problem is often modeled mathematically as a Markov decision process (MDP), where the agent interacts with the environment based on a particular policy
  - At time step $t$, the agent *chooses an action* $a_t$ from the action space. It then *receives a reward/punishment* $r_t$ from the environment and then *updates its current state* $s_t$
  - The algorithm aims to ***maximize the cumulative received rewards***

# Markov Decision Process (MDP)




- Considered state, action, and reward:
  - **State $s_t$:** the coordinates of UAV, users, and clwc grid
  - **Action $a_t$:** move ***(1)-north, (2)-west, (3)-south, (4)-east, (5)-remain stationary***
  - **Reward $r_t$:** $r_t = \alpha\sqrt{R_t \times R_t^{FSO}}$
    - $\alpha$: normalization factor
    - $R_t = \sum_{\{u1\in U_1\}} R_{u_1}(t)$: total users' rate provided by the UAV at time step $t$
    - $R_t^{FSO}$: backhaul capacity at time step $t$

# System Parameters

- RL Framework: Deep Q-Learning

| Parameter | Value |
|---|---|
| RF transmit power | 20 dBm |
| FSO transmit power | 17 dBm |
| RF bandwidth | 200 MHz |
| FSO bandwidth | 200 MHz |
| Number of GUs | 75 |
| UAV's covered radius | 100 m |
| Learning rate | 0.001 |
| Discount factor | 0.99 |
| Decaying epsilon-greedy parameters $(\epsilon_{\max}, \epsilon_{\min}, \Delta\epsilon)$ | 0.99, 0.01, 0.01 |
| Number of episodes | 2000 |
| Number of iterations/timestep per episode | 900 |

# Episode Reward

- As we can see, the episode reward gradually increases and converges after about 300 episodes
- This means that the agent forms a better movement policy over time

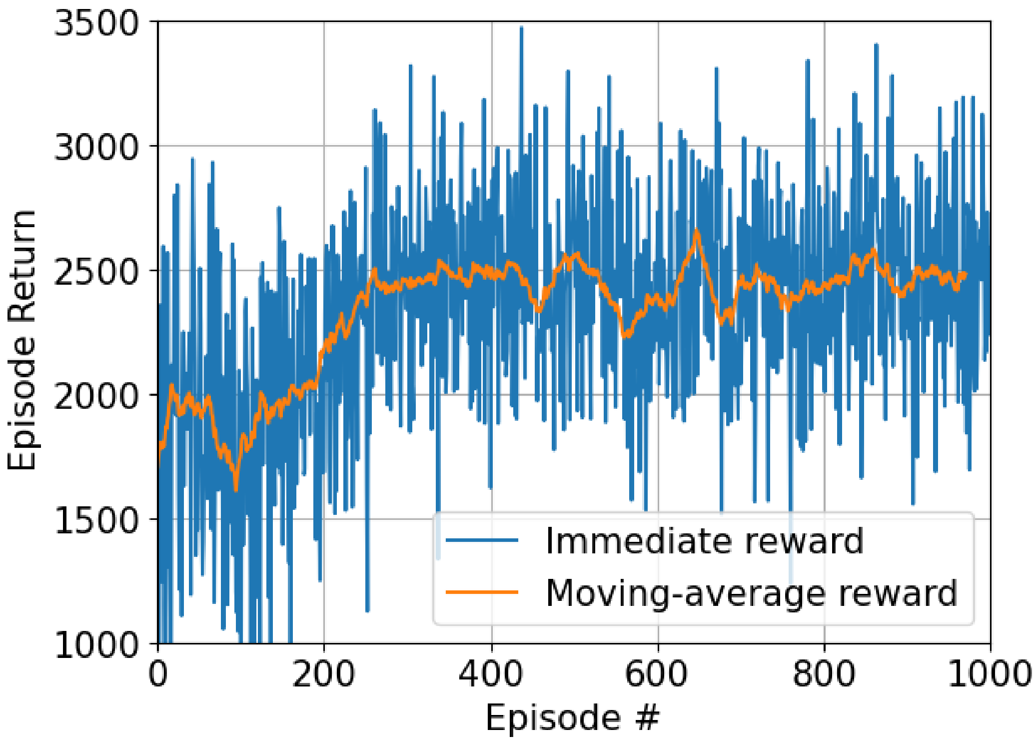



Fig. Episode reward vs. number of episode

# Path of Movement

- The figure demonstrates the movement of the trained agent in a test environment
  - The test environment was set up randomly
  - The agent was not trained during the test

→ *We can see that the UAV tends to move to the area that gathers many users and has low clwc*

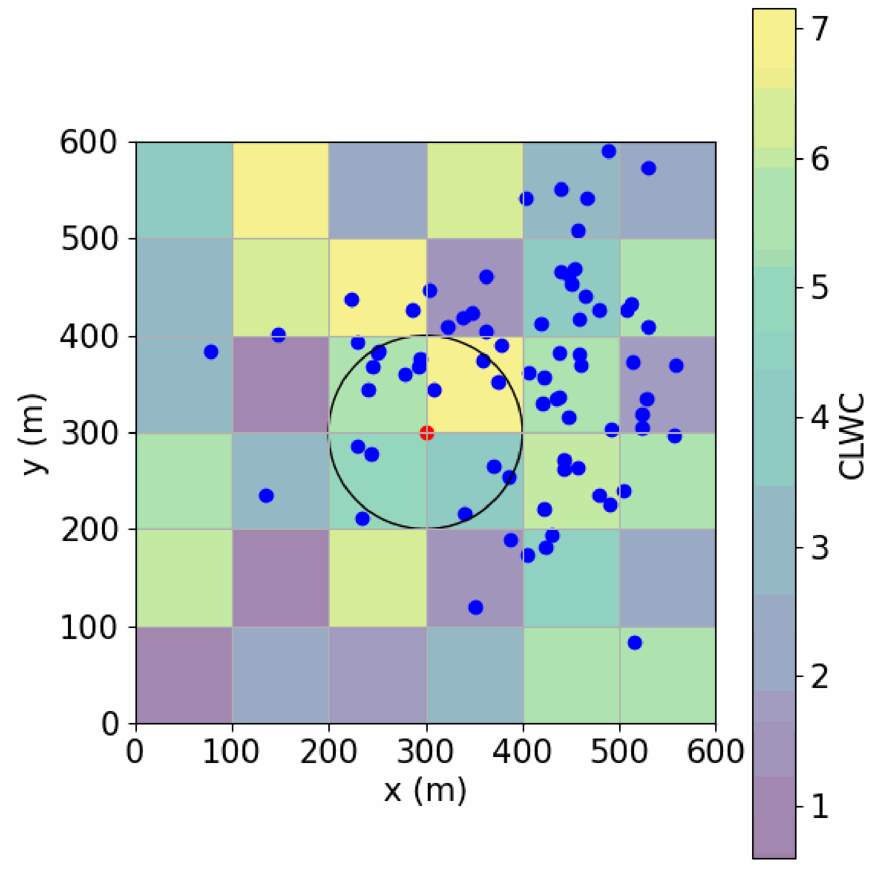



Fig. Behavior of the trained agent in the test environment

# Total Users' Rate

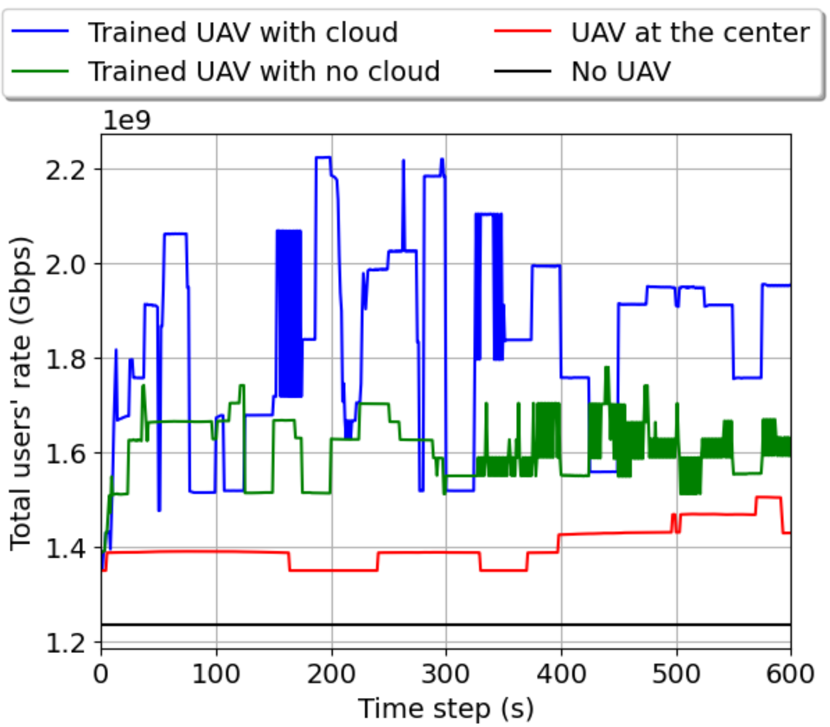



Fig. Total users' rate over time
(1 episode)

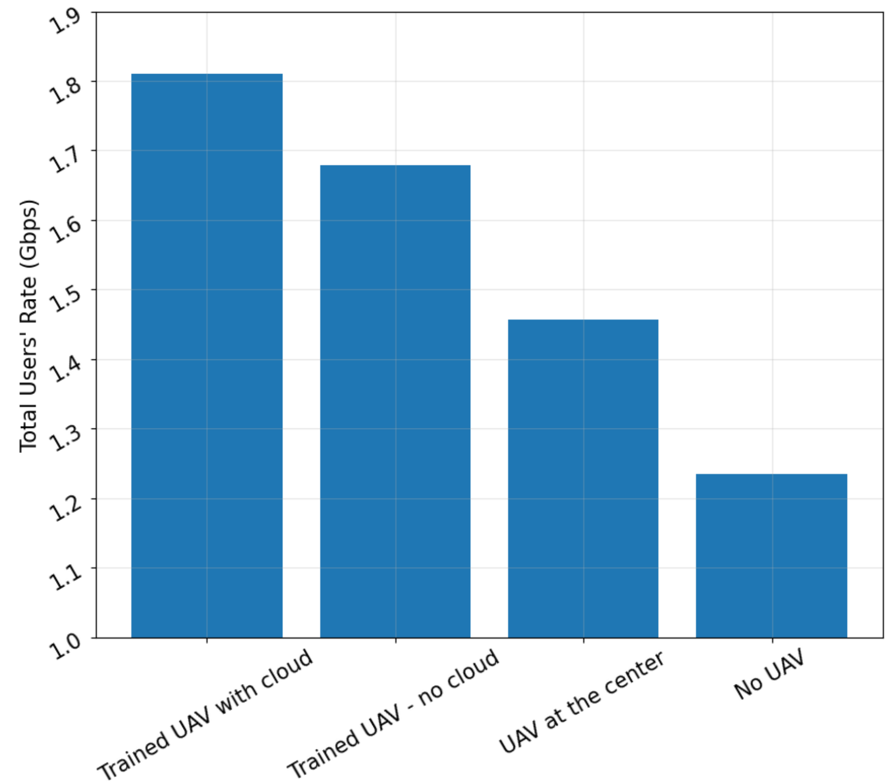



Fig. Average total user's rate over 50 test episodes

- As can be seen, our trained model provides the highest performance in terms of total users' rate over time
- This means that the agent can find good positions that have low clwc and there are many users around

Thank you for your listening!