Seminar: Deep Reinforcement Learning for UAV Networks

Le Viet Hung

Computer Communications Lab, The University of Aizu

hungle.dut@gmail.com

May 22, 2025

Contents

Application of UAVs

- UAVs offer unprecedented flexibility in providing services in remote and inaccessible areas, where traditional infrastructure is unavailable or inadequate.

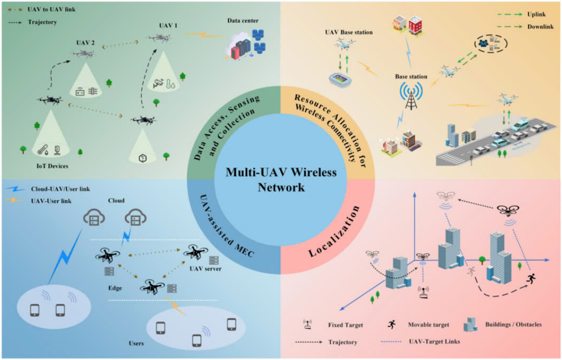


Figure: UAV applications

Application of UAV Networks

Data Access, Sensing, and Collection:

- In IoT applications, transmitting data to nearby Base Stations (BSs) can be challenging for wireless nodes → UAVs gather data from IoT devices and then transmit this data to the Data Center

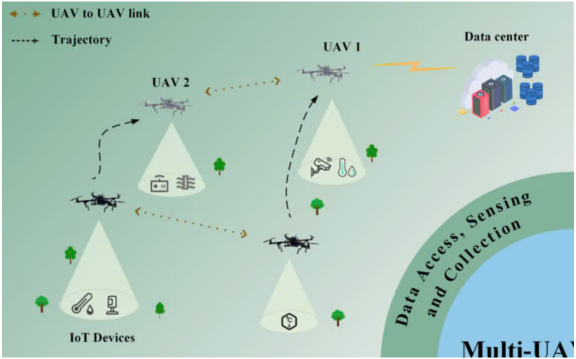


Figure: Data Access, Sensing, and Collection

Application of UAV Networks

Localization:

- UAVs play a crucial role in scenarios like target tracking, and positioning

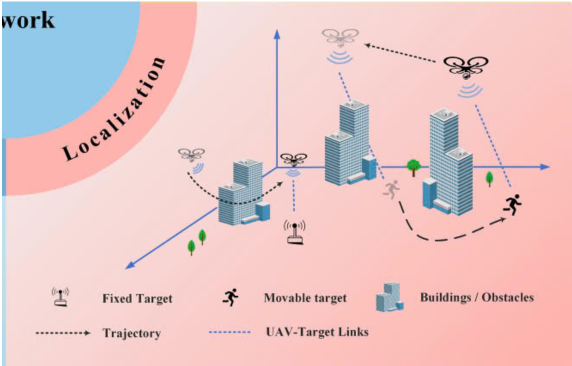


Figure: Localization

Application of UAV Networks

(Non-Terrestrial Networks)NTN-based Communication:

- UAVs are leveraged to assist as aerial base stations in providing network connectivity to ground users.

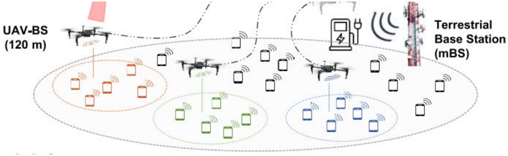


Figure: NTN-based Communication

Introduction to Reinforcement Learning

Traditional methods, such as linear programming and convex optimization are often designed for static environments, while Reinforcement Learning works well in dynamic environments.

Reinforcement Learning:

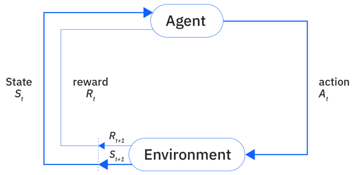


Figure: Reinforcement Learning

- 3 key components of RL: Action, State, Reward.
- Agent interact with Environment: Agent at the State takes an Action and then Environment return Reward and the next State

How to apply Reinforcement Learning for UAV Networks

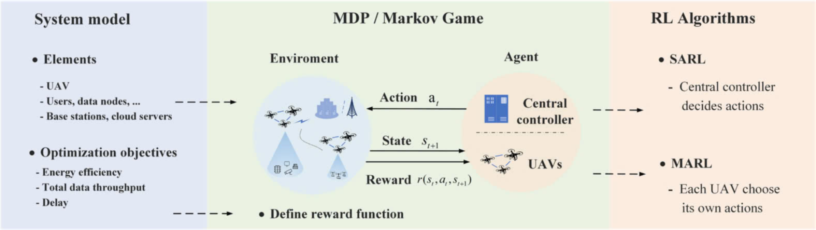


Figure: How to apply DRL to UAV Networks

1. Step 1: Create UAV Environment
2. Step 2: Define State, Action, Reward regarding Elements, Optimization objectives.
3. Step 3: Choose the Deep Reinforcement Learning algorithm for UAV Environment

Challenge of Applying DRL for UAV Networks

High Dimensionality and Computational Complexity

- High-dimensional state and action spaces → Requires signifcant computational resources and extended training times.

Multi-agent Coordination

- The number of agents (UAVs) increases → The growing action spaces and interactions → Complexity of learning policies increases between agents

Modeling Realistic Environments

- UAV communication networks operate in highly dynamic environments (mobility, ...) → DRL methods must adapt to dynamic conditions

Reward Function Design and Convergence

- Poorly designed rewards can lead to suboptimal policies or even divergence during training. In UAV networks, multiple objectives are involved → Balancing conflicting rewards poses a challenge

Overview

Problem Statement:

- **UAVs** are deployed to **support mBS** in providing data services to **users**
- To access the internet, the data rate that the user perceives need to be higher than a **minimum data rate threshold**

Objective:

- **UAVs and mBS** maximize the number of satisfied users that can access the internet

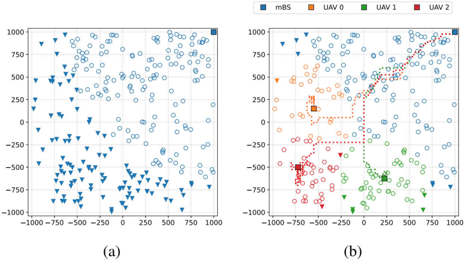


Figure: Comparison between With UAVs' support and Without UAVs' support

Limitation and Recent works

Limitation:

- In fact, UAVs can not move endlessly → They need to be **recharged**
- If 3 UAVs are charged at the same time, there aren't UAVs supporting users for a certain period of time → UAVs need to be **scheduled**

Recent works:

Paper	Journal/Conference	Scheduling UAV Networks	Limitation
[1]	IEEE Globecom 2020	No	Not consider the Scheduling
[2]	IEEE Transactions on CE 2025	No	Not consider the Scheduling
[3]	IEEE WCNC 2024	Using DRL for Scheduling	Consider Scheduling but not consider UAV Trajectory

→ Almost recent works don't consider **both Scheduling and UAV Trajectory**

→ My proposal will consider both of them.

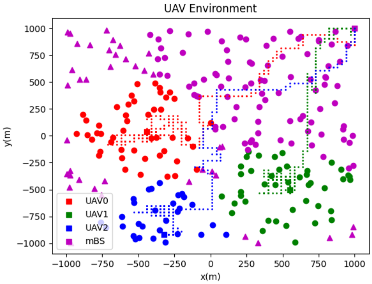
My proposal

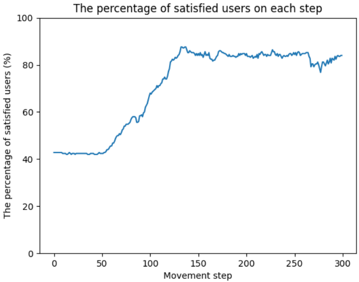
Proposal:

- Create and Train a **Controller** (with Deep Reinforcement Learning) to decide the time to quit/join UAV
- Besides "Quit/Join" Term, we create a new term "**In Service**"

Objective of Controller: 2 UAVs in service; 1 UAV charged (backup) → Schedule the UAV Recharging

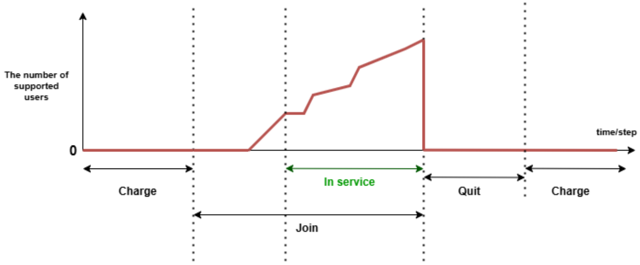
My proposal


UAV Environment
y(m)
x(m)
1000
750
500
250
0
−250
−500
−750
−1000
−1000
−750
−500
−250
0
250
500
750
1000
UAV0
UAV1
UAV2
mBS


The percentage of satisfied users on each step
The percentage of satisfied users (%)
Movement step
100
80
60
40
20
0
0
50
100
150
200
250
300

My proposal

Definition of Charge,Join,Quit, In Service:



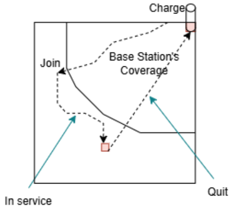


Figure: Charge, Join, Quit, In Service (1 UAV)

My proposal

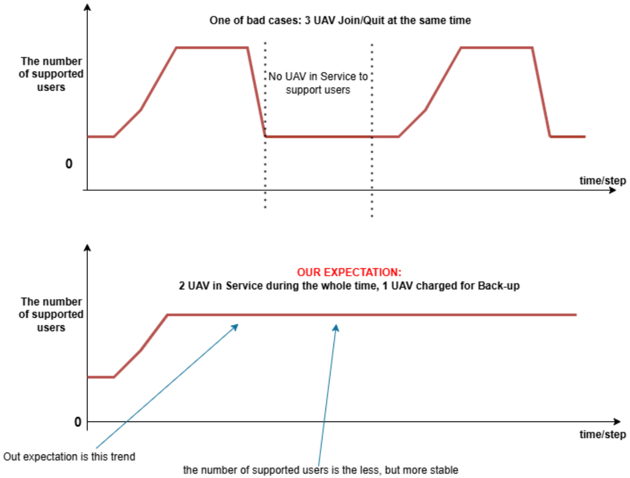


Figure: Objective of Controller

References

[1] Zhang, Ran, et al. "Learning to be proactive: Self-regulation of UAV based networks with UAV and user dynamics." IEEE Transactions on Wireless Communications 20.7 (2021): 4406-4419.
[2] Chen, Zijing, Yijun Guo, and Jianjun Hao. "Dynamic Trajectory Design for Multi-UAV Assisted Long-Term Data Collection Tasks." IEEE Transactions on Consumer Electronics (2025).
[3] Osrhir, Youssef, Btissam El Khamlichi, and Amal El Fallah-Seghrouchni. "PPOSWC: Deep Reinforcement Learning Recharging Scheduling for Effective Service in Multi-UAV Aided Networks." 2024 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2024.

Thank you