BRNO UNIVERSITY OF TECHNOLOGY · FACULTY OF INFORMATION TECHNOLOGY

Student: Bc. Filip Jahn
Supervisor: Ing. Jiří Novák, Ph.D.

Excel@FIT 2026

Optimization of Multi-Agent Unmanned Aerial Systems Behavior

Abstract

One of the promising safety-critical applications of multi-agent system optimization and control is autonomous aerial firefighting. From a computer science perspective, the core challenge solved in this thesis is the coordination of a **heterogeneous swarm**—combining agile quadcopter scouts with heavy fixed-wing drones. Using **Centralized Training with Decentralized Execution (CTDE)**, these mechanically and temporally distinct agents are trained *together* using **MAPPO** algorithm. They learn to develop a shared communication protocol, enabling them to navigate, share targeting data, and extinguish fires completely without any central command.

1. Simulation Environment & The Heterogeneous Swarm

High-fidelity UAV/UAS simulations with thermal updraft effect

The simulation models dangerous thermal updrafts directly affecting the flight dynamics of the agent.

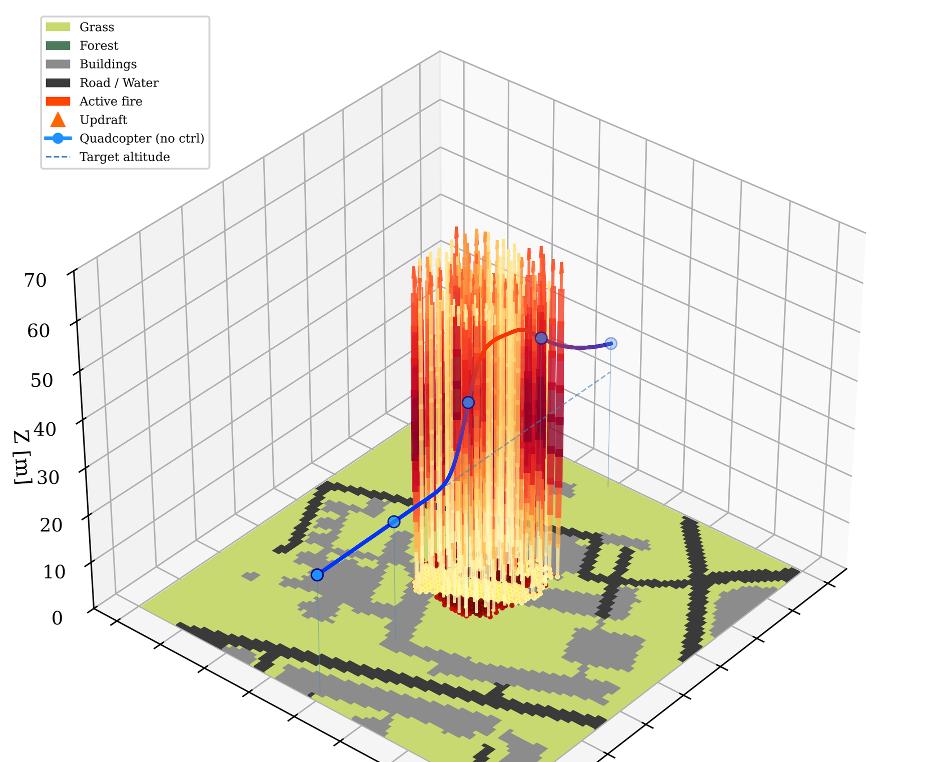


Figure 1: Thermal updrafts

Stochastic fire spread simulation model

Fire spread is dynamically simulated using stacked 2D cellular automata layers representing fuel, moisture, and burn intensity.

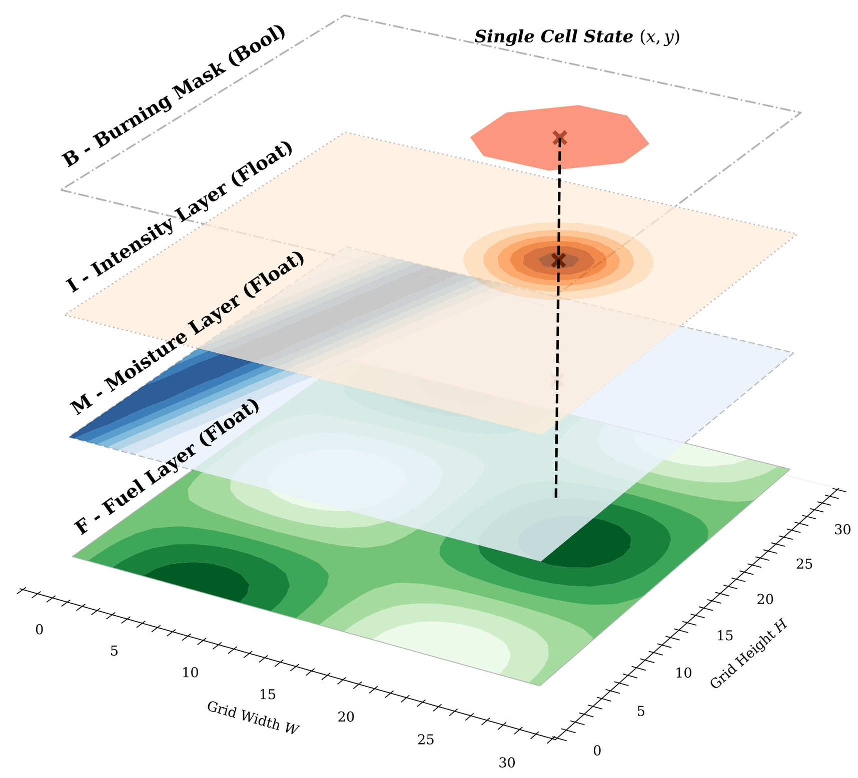


Figure 2: Environment memory representation

Agent Roles in Firefighting

- **SCOUT:** Quadcopter. Acts as a persistent visual sensor, broadcasting the fire location and intensity.
- **COMMANDER:** Fixed-wing. Carries payload, blind to fire, relies purely on decoded messages from Scouts.

Real-World to Simulation

Parsing OpenStreetMap GIS data into a 2D semantic grid. This enables drones to operate over realistic terrain topologies.

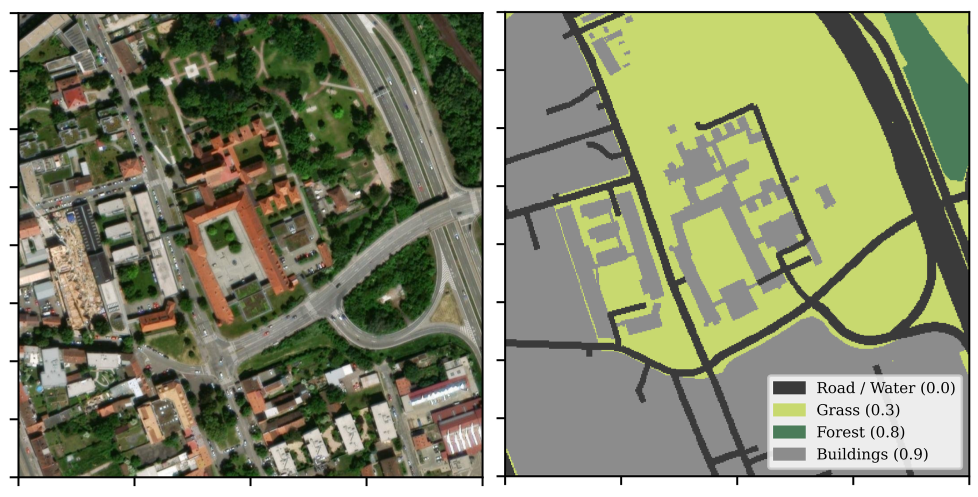


Figure 3: OSM to internal representation

2. CTDE Methodology & Network Architecture

MAPPO & CTDE

The algorithm utilizes a centralized critic that observes the privileged global state during training. However, the actors (drones) make execution decisions based solely on local observations and swarm communication.

Scout (Standard MAPPO):

$$\mathcal{L}_{Scout} = \underbrace{\mathcal{L}_{CLIP}}_{\text{Policy}} - \underbrace{c_1\mathcal{S}}_{\text{Entropy}}$$

Commander (Auxiliary Enhanced):

$$\mathcal{L}_{Cmdr} = \mathcal{L}_{CLIP} - c_1\mathcal{S} + \underbrace{c_2\mathcal{L}_{AUX}}_{\text{Fire Pos. Prediction}}$$

Curriculum Training Phases:

1. **Assisted Navigation** (explicit compass)
2. **Altitude Envelope** (vertical limits)
3. **Mission Cycle** (Extinguish → Refill)
4. **Pure Autonomy** (no training assists)

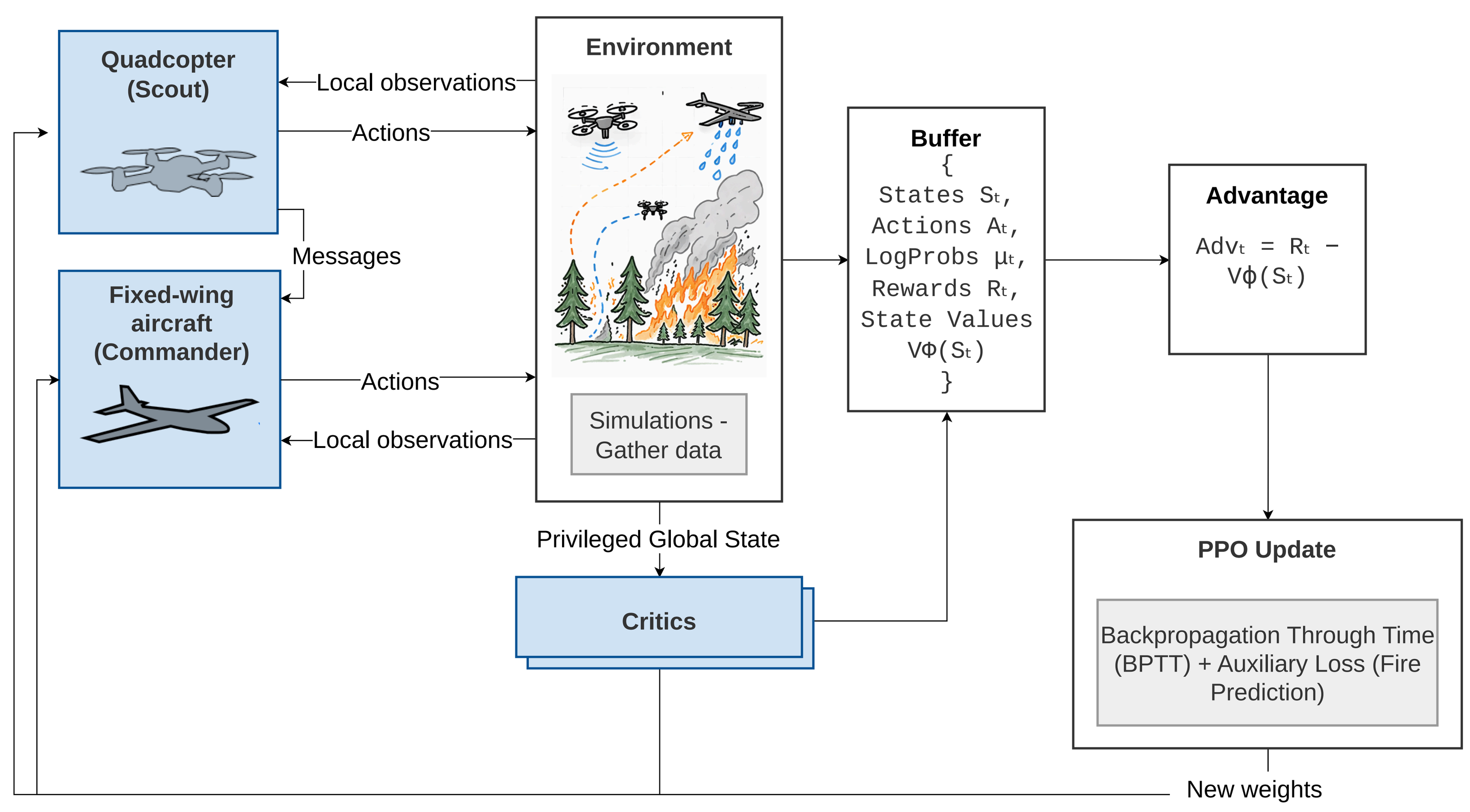
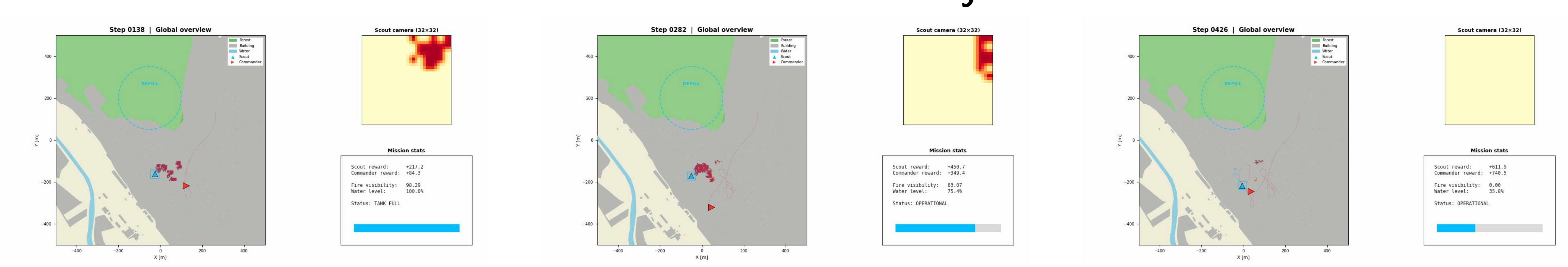


Figure 4: NN training pipeline

3. Training Results & Action Demo

Autonomous Mission Cycle



A) Heading closer to fire B) Approach to fire C) Drop water on the fire

Figure 5: Sequence of Commander navigates, drops water, and retreats.

Scan for gif results: