Drone Routing Problems Challenge

Demonstration Track

Shiyao Ding
Kyoto University
Kyoto, Japan
ding@i.kyoto-u.ac.jp

Fumito Uwano
Okayama University
Okayama, Japan
uwano@okayama-u.ac.jp

Hideki Aoyama
Panasonic Holdings Corporation
Osaka, Japan
aoyama.hideki@jp.panasonic.com

Donghui Lin
Okayama University
Okayama, Japan
lindh@okayama-u.ac.jp

ABSTRACT

Drone routing problems (DRP) focus on optimizing routes for a fleet of drones to ensure efficient, timely, and cost-effective delivery of goods. The problem is formulated as a multi-agent path finding (MAPF) problem, aiming to compute collision-free, optimal paths that minimize the total travel time of all agents. While traditional MAPF research is typically conducted on 4-connected grid graphs, DRP extends this to a non-grid environment suited for drone delivery. This competition incorporates non-grid maps based on real-world environments and a gym-standard environment for benchmarking various pathfinding algorithms.

KEYWORDS

Multi-Agent Path Finding; Drone Delivery

ACM Reference Format:
Shiyao Ding, Fumito Uwano, Hideki Aoyama, and Donghui Lin. 2026. Drone Routing Problems Challenge: Demonstration Track. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. https://doi.org/10.65109/HXAJ9978

Figure 1: Aerial delivery robots [6].

1 INTRODUCTION

The drone routing problems (DRP) are a critical challenge in logistics and autonomous systems, focusing on optimizing collision-free routes for multiple drones to enhance efficiency. This competition aims to minimize total fleet travel time while considering constraints such as payload capacity, battery life, and dynamic routing requirements [2, 5].

Unlike most traditional grid-based maps, the competition environment features a non-grid space where drones navigate along predefined edges, necessitating advanced path-planning algorithms [3, 4, 8, 9]. Participants can employ reinforcement learning, graph theory, swarm intelligence, and optimization techniques to develop sustainable routing solutions that contribute to reducing logistics-related carbon emissions. **The DRP Challenge website is available at [1] , and a demo video can be accessed at [2]**.

2 DRONE ROUTING PROBLEMS (DRP)

2.1 Rules and Guidelines

In this competition, we have designed an environment that simulates real-world drone delivery conditions, where participants navigate the environment to complete assigned tasks.

The environment consists of a non-grid space represented as a topological map with nodes and edges, where edges have assigned distances derived from real-world maps, as shown in Fig. 2. Drones move across the map based on predefined policies and are constrained to be located only on nodes and edges. Each agent has a unique departure point and destination node. An agent can detect the positions of other drones only when they appear on an adjacent node. Additionally, once a drone enters an edge, it must maintain its direction until it reaches the next node.

Drones continue moving until they either collide with another drone or all agents successfully reach their respective destinations. Given the dynamic nature of this environment, participants must apply interdisciplinary problem-solving approaches.

[1]https://drp-challenge.com/
[2]https://www.dropbox.com/scl/fi/01is8ref0934h4bnsl4qn/DRPDemo.mp4?rlkey=sgogcr92bhmfz2t3fb4712u9f&st=ma3b226q&dl=0

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/HXAJ9978

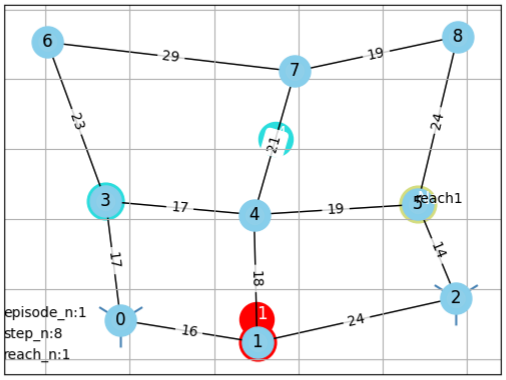

Figure 2: The development environment of the DRP Challenge.

2.2 Definition of the DRP

We define the Drone Routing Problem (DRP) in a general non-grid environment, following [1]. A team of drones is denoted by the index set $N = \{1, \ldots, i, \ldots, |N|\}$, operating in a two-dimensional space modeled as a graph $G = \langle V, E \rangle$.

The vertex set $V = \{v_1, \ldots, v_{|V|}\}$ consists of navigable locations, where each node v_k is associated with a planar coordinate $l_k = (l_k^x, l_k^y)$. The edge set $E = \{(v_k, v_l) \mid \text{a link exists between } v_k \text{ and } v_l\}$ represents navigable connections between nodes. For each node v_k, its adjacent nodes are defined as: $v_k^{nei} = \{v_l \mid (v_k, v_l) \in E\}$.

At the beginning of each episode, every drone i is assigned a start node $s^i \in V$ and a destination $g^i \in V$. During execution, drones may occupy not only nodes but also intermediate positions along edges, represented by continuous coordinates $l_k = (l^x, l^y)$.

The DRP is subject to the following feasibility constraints: 1) No two drones i and j are allowed to be located at the same position at the same time, i.e., $l^i \neq l^j$; 2) Two drones are prohibited from simultaneously traversing the same edge in opposite directions.

For a finite planning horizon of T time steps, the trajectory of drone i is represented as $path^i = (l^i[0], l^i[1], \ldots, l^i[T])$, with $l^i[0] = s^i$. If the drone reaches its target at step t, then $l^i[t] = g^i$, and the drone remains at the goal for all subsequent steps, i.e., $l^i[t'] = g^i$ for every $t' > t$.

The movement cost of a trajectory is measured by the cumulative Euclidean distance: $cost(path^i) = \sum_{t=0}^{T-1} \|l^i[t+1] - l^i[t]\|_2$.

The DRP aims to determine a set of trajectories for all drones that minimizes the overall travel cost while satisfying the above constraints. This objective is formulated as:

$$\begin{gathered} \min \sum_{\text{epi}} \sum_{i} cost(path^i_{\text{epi}}) \\ \text{subject to } \forall i \in N, \quad l^i[T] = g^i_{\text{epi}}. \end{gathered} \tag{1}$$

The terminal condition $l^i[T] = g^i_{\text{epi}}$ guarantees that each drone remains at its designated goal at the end of the episode. The notation g^i_{epi} reflects that goal assignments may differ between episodes. The overall objective is therefore to minimize the total movement cost aggregated over all episodes.

Table 1: Top-three methods in each DRP Challenge edition (lower cost is better)

Edition	Rank	Approach	Cost	Method
1st (AAMAS 2024)	1st	RL	11,902	QMIX [7]
	2nd	Planning	12,014	Improved heuristic function
	3rd	Planning	12,220	Priority-based search
2nd (AAMAS 2025)	1st	Planning	10,903	Improved CBS [10]
	2nd	Planning	10,991	Improved Dijkstra
	3rd	Planning	11,355.8	Shortest path + genetic algorithm

2.3 Evaluation Criteria

We evaluate all methods on three benchmark maps with different scales and topological characteristics: `map_3x3`, `map_aoba01`, and `map_shibuya`. For each map, we define multiple configurations with varying numbers of drones and distinct start–goal assignments. We refer to each unique combination of map, number of drones, and start–goal configuration as a problem instance. In total, 30 problem instances are used for evaluation, as specified in the predefined configuration file (`problem/problems.py`), which is fixed and cannot be modified by participants.

Cost for a single problem. For each problem p, we execute the policy over 10 independent iterations and compute the average cost: $cost_p = \frac{1}{10} \sum_{i=1}^{10} \sum_{j \in drones} cost_{ij}$, where i indexes the iteration and $drones$ denotes the set of drones in problem p. The per-drone cost $cost_{ij}$ is defined as:

$$cost_{ij} = \begin{cases} step_{ij}, & \text{if drone } j \text{ reaches its goal in iteration } i, \\ max\_steps, & \text{if a collision occurs.} \end{cases} \tag{2}$$

Here, $step_{ij}$ represents the number of time steps taken by drone j to reach its destination in iteration i, max_steps is the maximum steps in one episode. Collisions or failures to reach the goal are heavily penalized by assigning a fixed cost of 100.

Overall evaluation metric. The final evaluation score is computed as the sum of costs over all problem instances: $Final\ Cost = \sum_{p \in problems} cost_p$, where $problems$ denotes the full set of 30 problem instances. The objective of the DRP Challenge is to **minimize** this overall cost.

2.4 Competition Results

Table 1 summarizes the top-three methods and their costs in the past two competitions. The 1st Challenge attracted 8 teams, while the 2nd grew to 17 teams, indicating improved solution quality.

Approaches fall into two categories: planning-based and reinforcement learning (RL). Planning methods such as CBS [10] and priority-based search dominate. Although QMIX [7] (RL) won the first edition, all top-three methods in the second used planning. This is because instances vary in map size and robot count: RL often requires retraining, whereas planning generalizes across graphs and agent numbers without additional training.

ACKNOWLEDGMENTS

This work was supported by KAKENHI Grants JP23K11230, JP24K03001, and JP25K03227.

REFERENCES

[1] Shiyao Ding, Hideki Aoyama, and Donghui Lin. 2023. MARL4DRP: Benchmarking Cooperative Multi-agent Reinforcement Learning Algorithms for Drone Routing Problems. In *Pacific Rim International Conference on Artificial Intelligence*. Springer, 459–465.

[2] Kevin Dorling, Jordan Heinrichs, Geoffrey G Messier, and Sebastian Magierowski. 2016. Vehicle routing problems for drone delivery. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 47, 1 (2016), 70–85.

[3] Ariel Felner, Roni Stern, Solomon Eyal Shimony, Eli Boyarski, Meir Goldenberg, Guni Sharon, Nathan Sturtevant, Glenn Wagner, and Pavel Surynek. 2017. Search-based optimal solvers for the multi-agent pathfinding problem: Summary and challenges. In *Tenth Annual Symposium on Combinatorial Search*.

[4] Jiaoyang Li, Andrew Tinka, Scott Kiesel, Joseph W Durham, TK Satish Kumar, and Sven Koenig. 2020. Lifelong Multi-Agent Path Finding in Large-Scale Warehouses.. In *AAMAS*. 1898–1900.

[5] Amirhossein Moadab, Fatemeh Farajzadeh, and Omid Fatahi Valilai. 2022. Drone routing problem model for last-mile delivery using the public transportation capacity as moving charging stations. *Scientific Reports* 12, 1 (2022), 6361.

[6] Panasonic Newsroom. 2023. https://news.panasonic.com/jp/press/jn231106-1.

[7] Tabish Rashid, Mikayel Samvelyan, Christian Schroeder, Gregory Farquhar, Jakob Foerster, and Shimon Whiteson. 2018. Qmix: Monotonic value function factorisation for deep multi-agent reinforcement learning. In *International Conference on Machine Learning*. PMLR, 4295–4304.

[8] Oren Salzman and Roni Stern. 2020. Research challenges and opportunities in multi-agent path finding and multi-agent pickup and delivery problems. In *Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems*. 1711–1715.

[9] Guillaume Sartoretti, Justin Kerr, Yunfei Shi, Glenn Wagner, TK Satish Kumar, Sven Koenig, and Howie Choset. 2019. Primal: Pathfinding via reinforcement and imitation multi-agent learning. *IEEE Robotics and Automation Letters* 4, 3 (2019), 2378–2385.

[10] Guni Sharon, Roni Stern, Ariel Felner, and Nathan R. Sturtevant. 2015. Conflict-based search for optimal multi-agent pathfinding. *Artificial Intelligence* 219 (2015), 40–66. https://doi.org/10.1016/j.artint.2014.11.006