Decentralized Deep Reinforcement Learning for Cooperative Multi-Agent Flight Trajectory Planning in Adverse Weather

Extended Abstract

Bizhao Pang
Air Traffic Management Research Institute, Nanyang Technological University
Singapore
bizhao001@e.ntu.edu.sg

Xinting Hu
Air Traffic Management Research Institute, Nanyang Technological University
Singapore
xinting.hu@ntu.edu.sg

Mingcheng Zhang
Air Traffic Management Research Institute, Nanyang Technological University
Singapore
m200057@e.ntu.edu.sg

Sameer Alam
Air Traffic Management Research Institute, Nanyang Technological University
Singapore
sameeralam@ntu.edu.sg

Guglielmo Lulli
Dept of Informatics, Systems and Communication, University of Milano-Bicocca
Milan, Italy
guglielmo.lulli@unimib.it

ABSTRACT

Adverse weather, especially thunderstorms, disrupts air traffic operations and requires real-time trajectory adjustments to ensure aircraft safety. Existing methods often rely on centralized or single-agent approaches, lacking the coordination and robustness needed for scalable solutions. This paper presents a decentralized multi-agent method for cooperative trajectory planning, where each aircraft operates as an autonomous agent. The problem is modeled as a Decentralized Markov Decision Process (DEC-MDP) and solved with a proposed Independent Deep Deterministic Policy Gradient (IDDPG) algorithm. Experimental results show that the proposed method outperforms the state-of-the-art baselines in maintaining safe separation and optimizing rerouting efficiency under dynamically evolving thunderstorm cells.

KEYWORDS

Multi-Agent Systems; Cooperative Path Planning; Deep Reinforcement Learning; Air Traffic Management; Thunderstorm Weather

ACM Reference Format:
Bizhao Pang, Xinting Hu, Mingcheng Zhang, Sameer Alam, and Guglielmo Lulli. 2025. Decentralized Deep Reinforcement Learning for Cooperative Multi-Agent Flight Trajectory Planning in Adverse Weather: Extended Abstract. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 3 pages.

1 INTRODUCTION

Thunderstorms pose significant challenges to air traffic management (ATM), leading to loss of separation between aircraft and increased delays [6]. Effective trajectory planning for multiple aircraft in dynamic thunderstorms is essential to mitigate these risks and improve operational efficiency [2, 10]. The primary challenge lies in adjusting trajectories cooperatively in real-time, particularly in areas with limited air traffic control coverage, such as oceanic airspace [17], where timely and effective coordination is crucial.

Traditional trajectory planning under thunderstorms has utilized geometry [8], optimization [24], and heuristics methods [19]. Early works [7] emphasized iterative approach, while studies applied dynamic programming [16] and constrained optimization [14] for aircraft rerouting. Methods addressing uncertainties, such as [13, 18, 22], lacked real-time applicability or coordination in thunderstorm conditions. Optimal control approaches [9, 20, 25] improved robustness but still faced scalability challenges in complex and evolving environments. Heuristic methods [2, 11] showed effectiveness in trajectory planning but struggled with dependency and scalability, limiting large-scale and real-time applications.

Recent advancements in reinforcement learning (RL) have contributed to trajectory planning [27]. Proximal Policy Optimization (PPO) [4, 5] has been used for conflict resolution in air traffic. Deep Deterministic Policy Gradient (DDPG) [[23] has optimized aircraft vectoring under uncertainties. Models [1, 28] incorporating physics-based knowledge provide explainable solutions, while studies [12, 21] validated multi-agent RL models in simulations, emphasizing the need for alignment with air traffic control procedures. Despite these advancements, existing RL approaches still struggle to ensure coordination and scalability, especially when adapting to dynamic and unpredictable weather, such as thunderstorms.

This paper proposes a cooperative and scalable multi-agent RL method for trajectory planning problem under dynamic thunderstorms. The problem is modeled as a decentralized Markov Decision Process (DEC-MDP) and solved using the proposed Independent Deep Deterministic Policy Gradient (IDDPG) algorithm. The experimental results show that our method outperforms the baseline algorithms in maintaining safe separation and improving scalability in diverse evaluation scenarios.

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

2 PROBLEM AND METHODS

We model the problem as a Decentralized Markov Decision Process (DEC-MDP) [3], where each aircraft operates as an independent agent making decisions to optimize a joint reward while ensuring safe separations. For n agents, the DEC-MDP is defined by the joint state space $S = S_1 \times S_2 \times \cdots \times S_n$, joint action space $A = A_1 \times A_2 \times \cdots \times A_n$, joint reward function $R(s_1, \ldots, s_n, a_1, \ldots, a_n)$, and state transition probabilities $P = (P_1, P_2, \ldots, P_n)$. The state space for aircraft agent i at t includes its position $p_i(t)$, velocity $v_i(t)$, thunderstorm cell information $O^W(t)$, and the remaining distance to the exit waypoint $d_t^{i\_\text{Exit}}$. The action space consists of the heading change $a_t^i = \Delta h_i(t)$, a continuous variable within $[-30, 30]$ degrees. The reward function includes separation assurance, dynamic thunderstorm avoidance, heading change minimization, and distance minimization, all of which assign negative rewards for undesirable actions. The only positive reward is goal reaching. The total system reward is the weighted sum of these individual components, with the weights fine-tuned to reflect task priorities.

To solve the DEC-MDP, we propose an Independent Deep Deterministic Policy Gradient (IDDPG) method, leveraging a shared actor-critic architecture within a Centralized Training and Decentralized Execution framework [15]. During training, a centralized critic network $Q_i(s_t, a_t^i \mid \theta^{Q_i})$ evaluates actions for all agents using global state information s_t. This allows the critic to compute value functions that reflect the joint impact of actions, facilitating coordination across agents. Each agent maintains a decentralized actor network $\mu_i(s_t^i \mid \theta^{\mu_i})$, mapping its local observations s_t^i to actions a_t^i. In execution, each agent independently chooses actions based on its local state, ensuring scalability and real-time applicability.

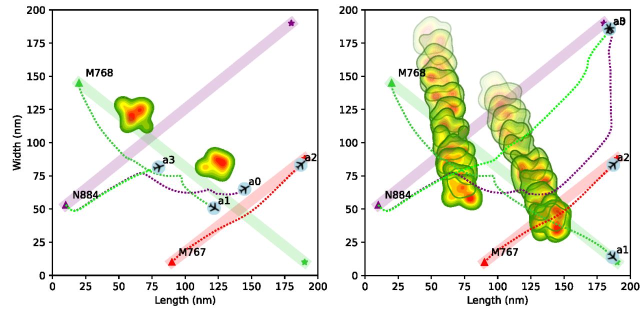
3 EXPERIMENTAL RESULTS

The multi-agent decentralized IDDPG method is adapted for aircraft rerouting under dynamic thunderstorms. Training and evaluations were conducted in a self-built simulator. Multiple evaluations were performed to assess the effectiveness and scalability of the proposed method. Performance metrics include safety (loss of separation (LOS) rate) and efficiency (goal reach rate and flight distance ratio).

The effectiveness of the proposed method is demonstrated in Fig. 1, where all aircraft trajectories successfully avoid dynamic thunderstorm cells while maintaining safe separation over time. Scalability results in Table 1 reveal that as the number of aircraft increases from 4 to 8, the method's performance remains robust, with Aircraft LOS rates consistently below 1% even in high-density scenarios. Baseline comparisons in Table 2 demonstrate that the proposed IDDPG method outperforms the state-of-the-art Fast Marching Tree (FMT) [11] and DDPG [23] methods with significant improvements, achieving a 98% goal reach rate compared to 87% for FMT and 90% for DDPG, while maintaining zero Aircraft LOS in contrast to 12% for FMT. These results show that the IDDPG enables superior coordination and efficiency in dynamic thunderstorm conditions.

4 CONCLUSIONS

This paper presents a decentralized multi-agent RL method for trajectory planning under dynamic thunderstorms. Simulations show a significant improvement in maintaining safe separation, with a reduction in conflict rates by up to 12% compared to baseline methods.



(a) t = 1/2 time step (b) t = final time step

Figure 1: All aircraft successfully avoids dynamic thunderstorms (evolving contours) and maintains safe separation. Entry and exit waypoints are marked by solid triangles and stars, and dashed lines represent rerouted trajectories.

Table 1: Scalability analysis in 100 test scenarios.

Performance metrics	Number of aircraft				
	4	**5**	**6**	**7**	**8**
Aircraft LOS rate	0	0	0	1%	0
Thunderstorm LOS rate	0	2%	2%	3%	5%
Goal reach rate	100%	98%	98%	96%	95%
Distance ratio (Mean ± S.D.)	(1.08 ±0.06)	(1.17 ±0.15)	(1.17 ±0.16)	(1.16 ±0.17)	(1.17 ±0.17)

Table 2: Comparisons of state-of-the-art methods in 100 diverse test scenarios with six aircraft.

Methods	**Aircraft LOS rate**	**Thunderstorm LOS rate**	**Goal reach rate**
FMT	12%	1%	87%
DDPG	2%	8%	90%
IDDPG (Ours)	**0**	**2%**	**98%**

The proposed IDDPG enables robust and scalable coordination between multiple aircraft trajectories during evolving thunderstorms, providing a promising method for real-time air traffic management in complex airspace affected by thunderstorms and other dynamic disruptions, such as space launch failures [26]. In future work, the method could be expanded to incorporate uncertainties in storm evolution and trajectory prediction.

ACKNOWLEDGMENTS

This research is supported by the Italian Ministry of Foreign Affairs and International Cooperation (MAECI) and the Agency for Science, Technology and Research (A*STAR), Singapore, under the First Executive Programme of Scientific and Technological Cooperation between Italy and Singapore for the years 2023–2025. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Italian Ministry of Foreign Affairs and International Cooperation or the Agency for Science, Technology and Research (A*STAR), Singapore.

REFERENCES

[1] Shaull Almagor and Morteza Lahijanian. 2020. Explainable Multi Agent Path Finding. In *19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020)*. Auckland, New Zealand, 34–42.

[2] Eduardo Andrés, Daniel González-Arribas, Manuel Soler, Maryam Kamgarpour, and Manuel Sanjurjo-Rivo. 2021. Informed scenario-based RRT* for aircraft trajectory planning under ensemble forecasting of thunderstorms. *Transportation Research Part C: Emerging Technologies* 129 (August 2021), 103232.

[3] Daniel S. Bernstein, Robert Givan, Neil Immerman, and Shlomo Zilberstein. 2002. The Complexity of Decentralized Control of Markov Decision Processes. *Mathematics of Operations Research* 27, 4 (2002), 819–840.

[4] Marc Brittain and Peng Wei. 2019. Autonomous Air Traffic Controller: A Deep Multi-Agent Reinforcement Learning Approach. In *International Conference on Machine Learning (ICML), Reinforcement Learning for Real Life (RL4RealLife) Workshop*.

[5] Marc Brittain, Xuxi Yang, and Peng Wei. 2020. A Deep Multi-Agent Reinforcement Learning Approach to Autonomous Separation Assurance. In *arXiv*.

[6] Performance Review Commission. 2024. *Performance Review Report (PRR) 2023: An Assessment of Air Traffic Management in Europe*. Performance Review Report. European Organisation for the Safety of Air Navigation (EUROCONTROL). Available: www.eurocontrol.int/air-navigation-services-performance-review.

[7] Heinz Erzberger. 2005. Automated Conflict Resolution for Air Traffic Control. In *25th International Congress of the Aeronautical Sciences (ICAS 2006)*. 1–27.

[8] Heinz Erzberger, Todd A. Lauderdale, and Yung-Cheng Chu. 2012. Automated conflict resolution, arrival management, and weather avoidance for air traffic management. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering* 226, 8 (August 2012), 930–949.

[9] Daniel González-Arribas, Manuel Soler, and Manuel Sanjurjo-Rivo. 2018. Robust aircraft trajectory planning under wind uncertainty using optimal control. *Journal of Guidance, Control, and Dynamics* 41, 3 (2018), 673–688.

[10] Daniel González-Arribas, Manuel Soler, Manuel Sanjurjo-Rivo, Maryam Kamgarpour, and Juan Simarro. 2019. Robust aircraft trajectory planning under uncertain convective environments with optimal control and rapidly developing thunderstorms. *Aerospace Science and Technology* 89 (June 2019), 445–459.

[11] Andréas Guitart, Daniel Delahaye, Félix Mora Camino, and Eric Feron. 2023. Collaborative Generation of Local Conflict Free Trajectories With Weather Hazards Avoidance. *IEEE Transactions on Intelligent Transportation Systems* 24, 11 (November 2023), 12831–12842.

[12] Yash Guleria, Duc-Thinh Pham, Sameer Alam, Phu N. Tran, and Nicolas Durand. 2024. Towards conformal automation in air traffic control: Learning conflict resolution strategies through behavior cloning. *Advanced Engineering Informatics* 59 (2024), 102273.

[13] Daniel Hentzen, Maryam Kamgarpour, Manuel Soler, and Daniel González-Arribas. 2018. On maximizing safety in stochastic aircraft trajectory planning with uncertain thunderstorm development. *Aerospace Science and Technology* 79 (August 2018), 543–553.

[14] Maryam Kamgarpour, Vera Dadok, and Claire Tomlin. 2010. Trajectory Generation for Aircraft Subject to Dynamic Weather Uncertainty. In *49th IEEE Conference on Decision and Control (CDC)*. IEEE, 2063–2068.

[15] Ryan Lowe, Yi I. Wu, Aviv Tamar, Jean Harb, Pieter Abbeel, and Igor Mordatch. 2017. Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments. *Advances in Neural Information Processing Systems (NeurIPS)* 30 (2017).

[16] Hok Kwan Ng, Shon Grabbe, and Avijit Mukherjee. 2009. Design and evaluation of a dynamic programming flight routing algorithm using the convective weather avoidance model. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*. 5862.

[17] Civil Aviation Authority of Singapore. 2024. Aeronautical Information Publication. Available: https://www.caas.gov.sg/docs/default-source/docs—ats/aip-singapore—16-may-2024.pdf.

[18] Bizhao Pang, Kin Huat Low, and Vu N. Duong. 2024. Chance-constrained UAM traffic flow optimization with fast disruption recovery under uncertain waypoint occupancy time. *Transportation Research Part C: Emerging Technologies* 161 (April 2024), 104547.

[19] Bizhao Pang, Kin Huat Low, and Chen Lv. 2022. Adaptive conflict resolution for multi-UAV 4D routes optimization using stochastic fractal search algorithm. *Transportation Research Part C: Emerging Technologies* 139 (2022), 103666.

[20] Jessica Pannequin, Alexandre Bayen, Ian Mitchell, Hoam Chung, and Shankar Sastry. 2007. Multiple aircraft deconflicted path planning with weather avoidance constraints. In *AIAA Guidance, Navigation, and Control Conference*. American Institute of Aeronautics and Astronautics Inc., 2467–2488.

[21] George Papadopoulos, Alevizos Bastas, George A. Vouros, Ian Crook, Natalia Andrienko, Gennady Andrienko, and Jose Manuel Cordero. 2024. Deep reinforcement learning in service of air traffic controllers to resolve tactical conflicts. *Expert Systems with Applications* 236 (February 2024).

[22] Duc-Thinh Pham, Yash Guleria, Sameer Alam, and Vu Duong. 2023. Probabilistic Conflict Detection using Heteroscedastic Gaussian Process and Bayesian Optimization. *IEEE Access* (2023).

[23] Duc-Thinh Pham, Phu N. Tran, Sameer Alam, Vu Duong, and Daniel Delahaye. 2022. Deep reinforcement learning based path stretch vector resolution in dense traffic with uncertainties. *Transportation Research Part C: Emerging Technologies* 135 (February 2022).

[24] Gauthier Picard. 2022. Trajectory Coordination based on Distributed Constraint Optimization Techniques in Unmanned Air Traffic Management. In *21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022)*. 1065–1073.

[25] Dinesh B. Seenivasan, Alberto Olivares, and Ernesto Staffetti. 2020. Multi-aircraft optimal 4D online trajectory planning in the presence of a multi-cell storm in development. *Transportation Research Part C: Emerging Technologies* 110 (January 2020), 123–142.

[26] Zhengyi Wang, Imen Dhief, Sameer Alam, Sven Kaltenhäuser, Tobias Rabus, and Henk Blom. 2023. Integrated Air and Space Traffic Management: An Agent-Based Simulation for Analysis of Space-Launch Impact on Air Traffic. In *13th SESAR Innovation Days 2023*. Sevilla, Spain. 2023-11-30.

[27] Yaoxin Wu, Mingfeng Fan, Zhiguang Cao, Ruobin Gao, Yaqing Hou, and Guillaume Sartoretti. 2024. Collaborative Deep Reinforcement Learning for Solving Multi-Objective Vehicle Routing Problems. In *23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024)*. 1956–1965.

[28] Peng Zhao and Yongming Liu. 2022. Physics Informed Deep Reinforcement Learning for Aircraft Conflict Resolution. *IEEE Transactions on Intelligent Transportation Systems* 23, 7 (July 2022), 8288–8301.