AI-Driven Multi-Dimensional Resource Scheduling for UAV-Enabled Multi-Access Edge Computing

Team Members: Xiaotian Jia, Weiyue Zhang, Shengran Ding, and Yutian Cui

Supervisor: Dr. Yuanjian Li

Xi'an Jiaotong-Liverpool University · SURF Code: SURF-2026-0240

Abstract

In a **cooperative multi-UAV mobile edge computing (MEC) system**, UAVs serve computation-constrained ground users, process part of their tasks onboard, and relay the remainder to terrestrial edge servers in infrastructure-sparse, delay-sensitive areas. We propose **DSPAC-Attn, an attention-enhanced multi-agent deep reinforcement learning algorithm** that jointly optimizes UAV trajectories and user-specific task-offloading ratios under latency, energy, coverage, and flight constraints. With centralized training and decentralized execution, independent actors and **shared twin critics equipped with Transformer attention** learn coordinated policies. We compare DSPAC-Attn with MADDPG, MATD3, and DSPAC in terms of system cost and training stability.

Missdrop/UAV-MEC-MADRL

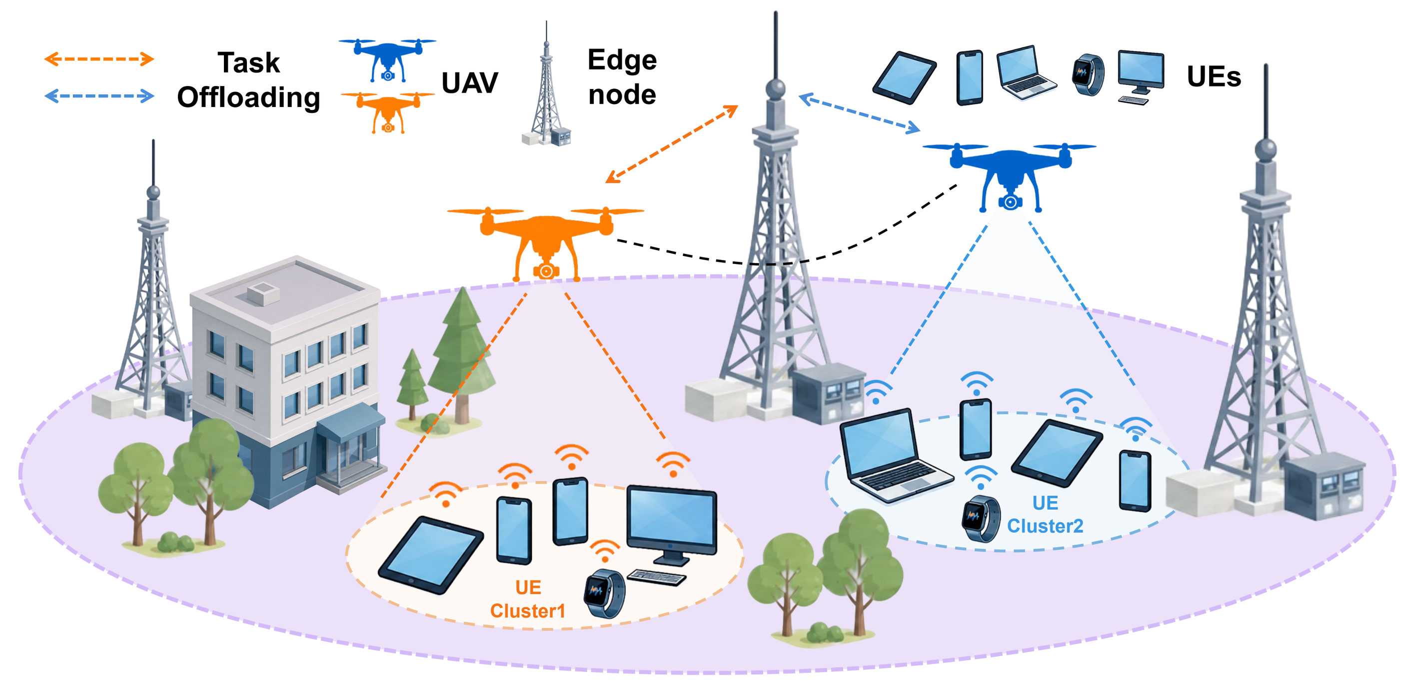

Figure 1. Cooperative UAV-assisted MEC system.

Literature review

⟹ Early UAV-assisted MEC research jointly optimized **computation scheduling, bandwidth allocation, and UAV trajectory**, demonstrating energy and latency benefits [1].

⟹ Cooperative multi-UAV studies added **task scheduling, partial offloading, and resource allocation** to improve processing delay and load balance [2].

⟹ MADRL enables joint **trajectory, task-allocation, and power control**. TD3 handles continuous actions, while DSPAC stabilizes energy-efficient coordination [3, 4].

⟹ Attention-based centralized critics selectively aggregate inter-agent information for scalable coordination [5], motivating our **Transformer-attention shared critic**.

Methodology

Multi-dimensional scheduling: DSPAC-Attn jointly controls UAV mobility, onboard-versus-edge computation, and communication service quality.

State and action: Each actor observes normalized UAV/UE/edge geometry, UE CPU-cycle and data-size demands, and connection indicators. Its continuous action contains a bounded 2-D movement vector and one ratio per UE, splitting tasks between onboard computation and the nearest edge server.

Reward design: All agents share the reward $r_t = -\eta U_t$ and therefore optimize the same system objective:

$$U_t = \underbrace{E_t + T_t + \frac{1}{\mathrm{Th}_{b,t}}}_{\text{operational cost}} + \lambda_u N_t^{\mathrm{unc}} + \lambda_c N_t^{\mathrm{unc}} \mathbb{1}[N_t^{\mathrm{unc}} > N_c] + \lambda_b B_t.$$

Here, E_t, T_t, and $\mathrm{Th}_{b,t}$ are total energy, delay, and minimum served-UE throughput. They aggregate UE upload, UAV computation, UAV-edge relay, and edge computation. Additional terms penalize unserved UEs, uncovered demand above N_c, and boundary violations.

Attention-enhanced critics: Replayed joint transitions update shared twin critics. Each state-action pair becomes an agent token; a Transformer encoder captures inter-UAV dependencies, and learned softmax pooling produces the global representation used to evaluate joint actions.

Stable updates: Clipped target-action noise and the smaller twin-target value reduce optimistic bias. Critics minimize temporal-difference loss; actors update less frequently, and Polyak averaging tracks the target networks.

Decentralized execution: After training, each UAV acts from its local observation without online critics or global state aggregation.

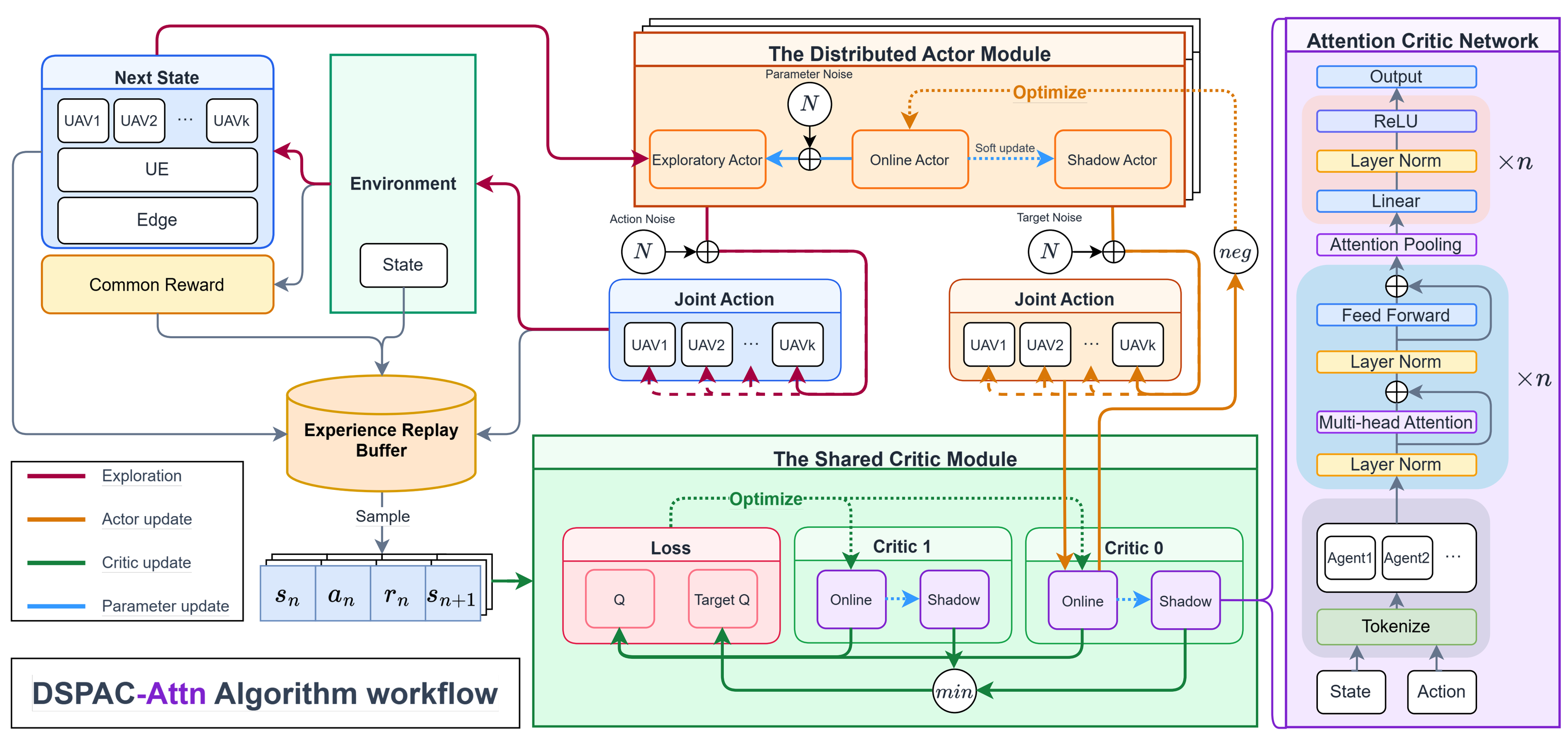

Figure 2. Workflow of the proposed DSPAC-Attn algorithm

Results

Evaluation protocol: DSPAC-Attn is compared with MADDPG, MATD3, and DSPAC over 1,000 episodes. Learning curves, five 200-episode stages, and the final-100 distribution evaluate the multi-dimensional cost of energy, delay, bottleneck throughput, coverage, and flight feasibility.

Main finding: DSPAC-Attn reaches the low-cost region substantially earlier than all baselines and retains the strongest final-stage distribution, demonstrating persistent rather than isolated improvement.

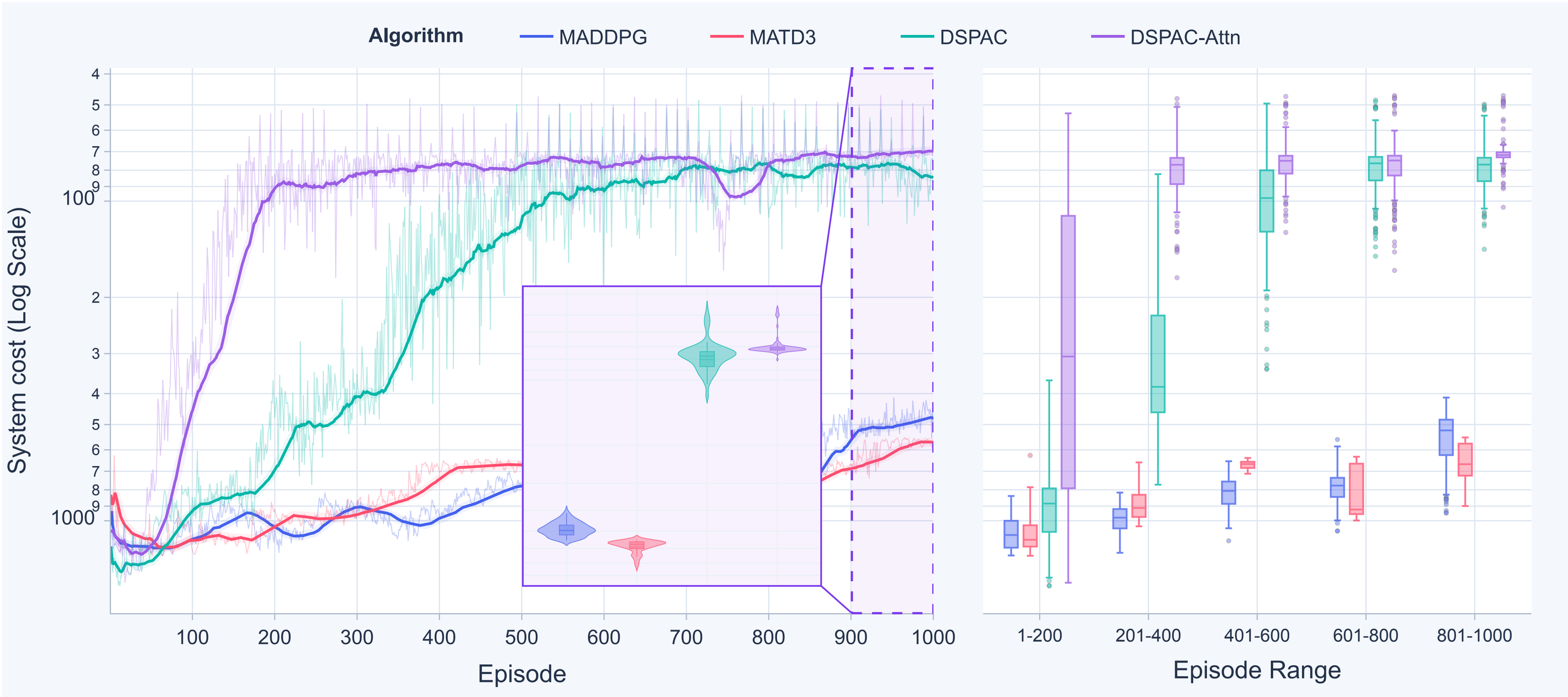

Figure 3. Training stats over 1,000 episodes

Effect of attention: DSPAC improves only after longer training, whereas DSPAC-Attn reduces cost rapidly and stabilizes earlier. Their separation supports Transformer attention for extracting useful inter-UAV interactions within the shared critic.

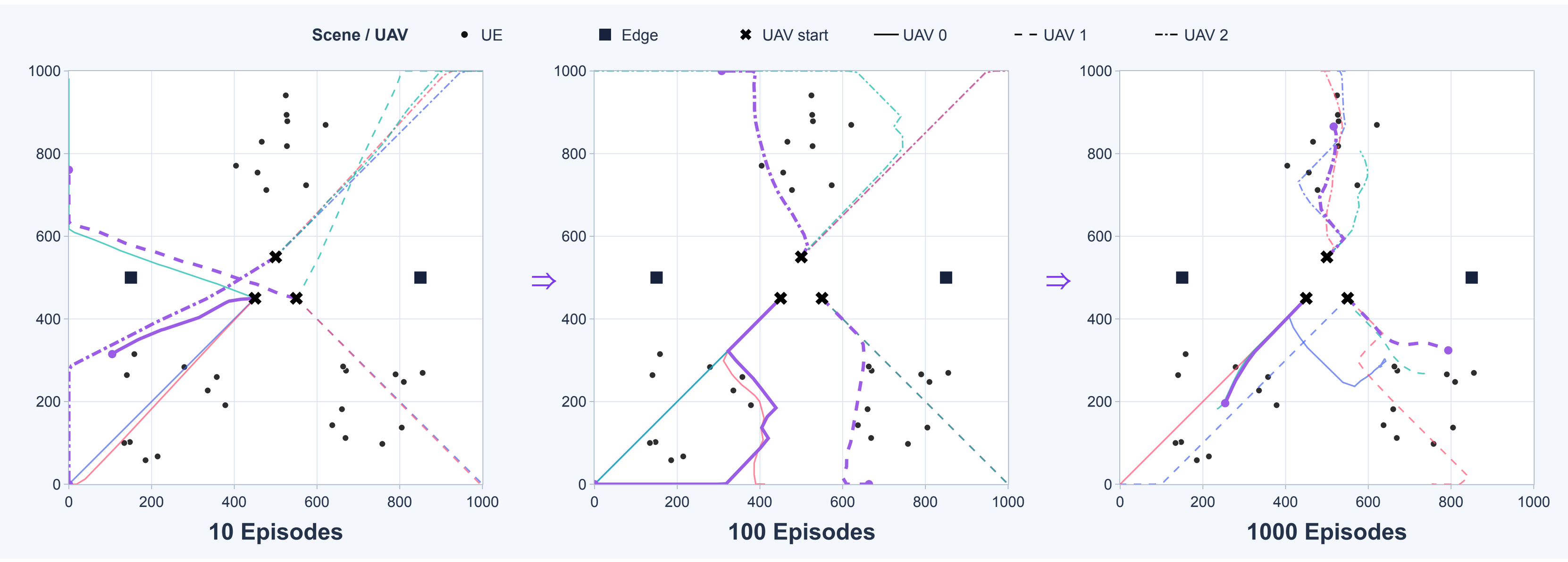

Figure 4. Evolution of DSPAC-Attn UAV trajectories during training

Trajectory evolution: Long, intersecting paths at 10 episodes separate toward different service regions by 100 episodes. At 1,000 episodes, localized trajectories cover UE clusters with less redundant movement, connecting spatial coordination to lower system cost.

Conclusion

Proposed solution: DSPAC-Attn unifies three scheduling dimensions: UAV mobility and coverage, onboard-versus-edge computation, and communication service quality. Decentralized actors jointly output movement and user-specific offloading ratios, while shared twin critics use Transformer attention to learn which inter-UAV interactions matter for the global objective.

Outcome: Compared with MADDPG, MATD3, and DSPAC, the proposed algorithm learns the low-cost region earlier, maintains a stronger final distribution, and develops localized, differentiated UAV trajectories. These results show that attention-enhanced centralized training improves coordinated multi-dimensional scheduling while preserving local-observation-only execution.

References

[1] X. Hu, K.-K. Wong, K. Yang, and Z. Zheng, "UAV-Assisted Relaying and Edge Computing: Scheduling and Trajectory Optimization," *IEEE Trans. Wireless Commun.*, vol. 18, no. 10, pp. 4738–4752, Oct. 2019.
[2] H. Guo, Y. Wang, J. Liu, and C. Liu, "Multi-UAV Cooperative Task Offloading and Resource Allocation in 5G Advanced and Beyond," *IEEE Trans. Wireless Commun.*, vol. 23, no. 1, pp. 347–359, Jan. 2024.
[3] N. Zhao, Z. Ye, Y. Pei, Y.-C. Liang, and D. Niyato, "Multi-Agent Deep Reinforcement Learning for Task Offloading in UAV-Assisted Mobile Edge Computing," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 6949–6960, Sep. 2022.
[4] Y. Li, A. S. Madhukumar, T. Z. H. Ernest, G. Zheng, W. Saad, and A. Hamid Aghvami, "Energy-Efficient UAV-Driven Multi-Access Edge Computing: A Distributed Many-Agent Perspective," *IEEE Trans. Commun.*, vol. 73, no. 9, pp. 8405–8420, Sep. 2025.
[5] S. Iqbal and F. Sha, "Actor-Attention-Critic for Multi-Agent Reinforcement Learning," *in Proc. 36th Int. Conf. Mach. Learn. (ICML)*, vol. 97. PMLR, Jun. 2019, pp. 2961–2970.

SURF

Summer Undergraduate Research Fellowship