

강화학습 기반 UAV 스케줄링 및 자원 관리 기술 동향

이현수¹, 정소이^{2*}, 김중헌^{1*}
고려대학교, *한림대학교

¹hyunsoo@korea.ac.kr, ²sjung@hallym.ac.kr, ^{1*}joongheon@korea.ac.kr

Trends in Reinforcement Learning-based UAV Task and Charge Scheduling

Hyunsoo Lee¹, Soyi Jung^{2*}, Joongheon Kim^{1*}

¹Korea Univ., ²Hallym Univ.

요약

강화학습은 제약이 지정된 상황에서도 최선의 결과를 낼 수 있다는 점에서 다양한 산업 분야에 활용되고 있다. 특히 UAV의 자원 최적화, 경로 탐색, 충전 및 작업 스케줄링 등 다양한 분야에서 연구되고 있다. 본 논문에서는 강화학습이 적용된 UAV 기술 동향에 대해 살펴본다. 크게 스케줄링 및 자원 관리로 나누어 기술 동향을 분석하고, 향후 연구 방향에 대해 논의한다.

I. 서론

UAV (Unmanned Aerial Vehicle)는 운수업, 농업, 제조업, 통신 사업 등 향후 많은 산업 분야에서 활용될 가능성이 있는 블루칩 기술로 평가받고 있다. 하지만 최근의 배터리 기술 발전에도 불구하고, 최대 수십 분에 불과한 운용 시간과 온보드 프로세서의 성능 제약으로 인해 자원 관리의 제약이 필연적으로 발생한다.

이에 최대한 배터리 사용량을 늘리고, 자원을 효율적으로 활용하는 방안으로 심층강화학습이 널리 연구되고 있다. 심층강화학습은 문제를 마르코프 결정 프로세스(Markov Decision Process, MDP)로 모델링한 뒤에, 적합한 보상 함수를 설계하여 심층 신경망을 통해 학습을 진행하는 방법으로, UAV 뿐만 아니라 자율주행, 스트리밍, 이미지 화질 향상 등 다양한 분야에 활용되고 있다 [1-5]. 이에 본 논문에서는 충전 스케줄링 문제와 자원 관리 문제를 해결하기 위해 강화학습이 적용된 논문들에 대해 살펴본다.

II. 강화학습 기반 UAV 작업 스케줄링

모바일 엣지 컴퓨팅 (Mobile Edge Computing, MEC) 시스템에 UAV가 활용되면, 기존 MEC 서버에 추가 연산 리소스를 제공함으로써 모바일 디바이스의 수명을 연장하고, 계산을 가속화하여 통신 혼잡을 완화할 수 있다. 하지만 제한된 payload와 배터리 용량 때문에, 연산 능력과 비행 시간에 한계를 가진다. 따라서, UAV가 활동하는 도중에 처리하는 작업량과 시스템의 안정도를 최대화하고, 에너지 소비를 최소화하는 컨트롤 모델을 제작하기 위해, [6]에서는 UAV의 경로 설정 및 작업 오프로딩을 최적화할 수 있는 종단 간 심층강화학습 모델을 제안하였다. 작업 전송 시간과 오프로딩에 걸리는 시간, 연산 시간을 합쳐 Total time cost를 계산하고, 각각에 사용되는 에너지의 양을 더하여 Total energy consumption을 계산한 뒤에, 에너지 및 시간 소비는 최소화하면서, 정해진 시간 내에 완료한 작업의 양은 최대화하는 방식으로 문제를 formulate 하였다. 이에 대한 보상 함수는 각 episode가 끝나기 직전까지 수행한 작업량에서 에너지 및 시간의 소비량을 뺀 값으로 계산한 종단 간 연산을 최적화하는 Deep Q-Network (DQN) 모델을 제작하였다. 해당 논문의 시뮬레이션 결과는 기존의 DQN, Q-Learning, Greedy 알고리즘에

비해 더 빠른 시간 내에 많은 작업을 효율적으로 수행하는 것을 보여주었다.

III. 강화학습 기반 UAV 충전 스케줄링

다수의 UAV가 태양광 발전 시스템이 구축된 다수의 충전 타워를 사용하는 경우, 비용 효율적으로 에너지 저장 시스템 (Energy Storage System, ESS)을 이용하기 위해 [7]에서는 ESS 간 정보를 주고받는 Co-operative 충전 구조를 설계하였다. UAV는 충전 타워와 충전량, 충전 시간을 결정하면서, 거리가 가까운 충전 타워를 선택하여 불필요한 비행시간을 최소화하고, UAV의 배터리 잔량을 최대한 소비한 상태에서 충전하며, 한 번에 충전하는 에너지 양은 최대한 많도록 설계하였다. 충전 타워 측면에서의 에너지 공유 시에는 한번에 결정하는 것이 아닌, 강화학습을 적용하여 동적으로 변화하는 상황에 맞게 원하는 목적을 달성하기 위한 연속적인 의사 결정 프로세스를 수행한다. 이때 추구하는 목적은 에너지가 부족할 때 전력 회사로부터 구매하는 에너지를 최소화하기 위해 충전 타워간의 에너지 공유에 대한 결정을 수행하는 것으로, 이는 시간이 지나면서 시평균 값으로 정리되어 최적 값을 도출할 수 있다. 제안하는 알고리즘은 블록 최적화 스케줄링 및 다중에이전트 강화학습 방식을 통해 잉여자원을 최소화하면서 구매 에너지를 최소화하고, 에너지 공유를 최대화할 수 있다.

IV. 결론

본 논문에서는 강화학습 기반의 UAV 스케줄링 기술 동향에 대해 논의하였다. 기존에 풀기 어려웠던 복잡한 문제들을 강화학습을 통해 처리할 수 있다는 것을 확인하였고, 추후 다중에이전트 강화학습을 통해 여러 에이전트의 협업으로 성능이 향상되는 연구가 수행될 수 있다.

ACKNOWLEDGMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2022-0-00907, (2세부) AI Bots 협업 플랫폼 및 자기조직 인공지능 기술개발). 본 논문의 교신저자는 정소이, 김중헌임.

참 고 문 헌

- [1] D. Kwon, J. Kim, D. A. Mohaisen and W. Lee, "Self-adaptive power control with deep reinforcement learning for millimeter-wave Internet-of-vehicles video caching," *Journal of Communications and Networks*, vol. 22, no. 4, pp. 326-337, Aug. 2020
- [2] W. J. Yun, S. Park, J. Kim, M. Shin, S. Jung, D. Mohaisen, and J-H. Kim, "Cooperative Multi-Agent Deep Reinforcement Learning for Reliable Surveillance via Autonomous Multi-UAV Control", *IEEE Transactions on Industrial Informatics (Early Access)*, January 2022
- [3] H. Baek et al., "Joint Superposition Coding and Training for Federated Learning over Multi-Width Neural Networks," *arXiv preprint arXiv:1612.04340*, December 2021.
- [4] M. Shin and J. Kim, "Randomized Adversarial Imitation Learning for Autonomous Driving," in *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*, Macau, China, August 2019.
- [5] M. Shin, J. Kim and M. Levorato, "Auction-Based Charging Scheduling With Deep Learning Framework for Multi-Drone Networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4235-4248, May 2019
- [6] L. Zhang, Z. Zhang, L. Min, C. Tang, H. Zhang; Y. Wang, and P. Cai, "Task Offloading and Trajectory Control for UAV-Assisted Mobile Edge Computing Using Deep Reinforcement Learning," *IEEE Access*, vol. 9, pp. 53708-53719, April 2021
- [7] S. Jung, W. J. Yun, M. Shin, J. Kim and J. -H. Kim, "Orchestrated Scheduling and Multi-Agent Deep Reinforcement Learning for Cloud-Assisted Multi-UAV Charging Systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 5362-5377, June 2021