

다중 에이전트 강화학습을 이용한 UAV-자원 할당 최적화에 대한 기술 동향

조현희, 김중헌

성균관대학교, 고려대학교

gt7746d@g.skku.edu, joongheon@korea.ac.kr

Survey on Reinforcement Learning Technologies for Resource Allocation in Unmanned Aerial Vehicles (UAVs)

Hyunhee Cho and Joongheon Kim

Sungkyunkwan University, Korea University

요약

예상하기 어려운 다양한 환경(Environment) 변화나 UAV(Unmanned Aerial Vehicle) 개체의 에너지 또는 컴퓨팅 자원의 한계를 극복하려는 목적으로 강화학습(Reinforcement Learning, RL) 알고리즘을 UAV 시스템에 도입한 연구는 활발하게 진행되고 있다. UAV 군집 주행과 같이 다수의 개체가 존재하는 환경에서 다중 에이전트 강화학습을 기반으로 영상 처리 및 정보 보안을 위해 데이터 할당을 최적화하는 목적의 UAV 시스템에 도입되는 연구를 살펴보았다. 본 논문은 UAV 기반 시스템에 도입한 다중 에이전트 강화학습을 각 에이전트가 할당받는 데이터의 양질 및 서비스 최적화 목적과 데이터 손실과 같은 비용(Cost) 최소화 목적으로 분류하여 최신 연구 동향을 분석하였다.

I. 서론

UAV는 높은 기동성과 넓은 적용 범위, 사용자와 Line-of-Sight (LOS) 링크를 맺을 수 있는 장치로서 주목받고 있다[1]. 현재 크기, 비행 반경 및 고도, 이용 목적에 따라 다양한 종류의 UAV가 [2]와 같이 분류되고 있어 실제 연구 및 산업 분야에 활용되고 있다. UAV는 경찰이나 전투와 같은 군사적 용도와 통신 중계, 무인 배송, 조난 구조 시스템과 같이 민간용으로도 사용되는데 그 중 드론 촬영이나 데이터 처리 분야에 관심이 높아지면서 UAV의 데이터 품질을 고려한 자원 할당 시스템이 주목받고 있다.

본 논문에서는 최근 UAV의 자원 할당 최적화 목적으로 적용되는 강화학습(Reinforcement Learning, RL) 기술 동향을 분석한다. UAV 시스템에서 고려해야 할 자원을 1) 각 에이전트가 소비하는 시간을 적절히 분배하는 시간 할당(Time allocation), 2) 수집하는 데이터의 양과 그 데이터의 품질을 최적화하는 데이터 할당(Data Allocation), 그리고 3) 개별 에이전트가 차지할 수 있는 공간의 양을 지정하는 공간 할당(Space Allocation)으로 크게 분류한다. 이 중 데이터 할당은 UAV들이 개별적인 데이터 수집을 할 때, 데이터의 품질 및 사용자가 받는 서비스의 최적화와 통신량으로 발생하는 잡음 문제 최소화 등으로 세분화할 수 있다.

II. 본론

인공지능 기술을 통한 학습보다는 UAV 성능 자체의 개선에 대하여 진행되었던 기존 연구에서와 달리, 최근의 UAV 기반 연구들은 환경(Environment) 변화에 대한 빠른 대응과 배터리 용량 등의 한계를 극복하려는 목적[3]으로 강화학습 알고리즘을 도입하여 UAV 시스템의 개선을 연구하는 경향을 보인다. 본 장에서는 UAV 환경에서 자주 응용되는 다중 에이전트 강화학습에 대한 개념을 서술하고, UAV 기반 시스템에 도입한 다중 에이전트 강화학습이 양질의 데이터 수집 및 서비스 최적화 그리고 통신 과정에서 발생할 수 있는 데이터 손실 최소화를 목적으로 세운 최신

연구 동향을 논한다.

2-1 다중 에이전트 강화학습

특정 UAV 시스템에 도입되는 강화학습 알고리즘은 사용하는 UAV의 개수에 따라 달라지기도 한다. UAV가 군집 주행을 하는 경우, 일반적으로 MARL(Multi Agent Reinforcement Learning) 알고리즘이 사용되고 있다. MARL 알고리즘[4]은 단일 에이전트 기반 강화학습만으로 해결될 수 없는 트래픽 제어 또는 작업 분배 시스템 등에 사용되어 강화학습의 열풍과 더불어 재조명받고 있다. 그림 1에서 볼 수 있듯이, MARL은 단일 에이전트 강화학습과 마찬가지로 MDP(Markov Decision Process) 기반의 시스템 최적화 문제를 해결하는데, 상태(State)와 각 에이전트가 받는 보상(Result)은 모든 에이전트의 공동 행동(Joint action)에 영향을 받는다[5].

2-2 데이터 양질 및 서비스 최적화

UAV는 무선 통신(Wireless Communication)을 통해 데이터를 전송 또는 수집할 수 있다. 무선 통신 센서로부터 데이터를 수집하고 중앙 기지국에 데이터를 전송하는 ADCP(Aerial Data Collection Problem) 문제[6]가 제안되었다. 해당 문제를 해결하기 위해 DQN(Deep Q-Network) 알고리

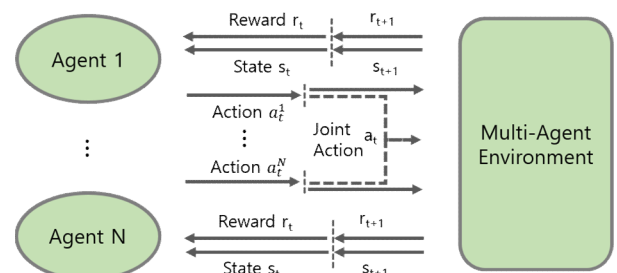


그림 1 다중 에이전트 강화학습 구조

즘이 UAV의 서비스를 받는 사용자들의 QoS(Quality of Service)와 데이터 수집 정도를 파악하는 것을 목적으로 사용되었다.

첫 번째 연구에서는 UAV를 통해 수집된 데이터가 사용자의 개별 휴대용 장치로 전송될 때, 전송되는 데이터의 QoS를 휴대용 장치의 이동통신망(Cellular Network)에서 보장하기 위하여 다중 에이전트 DQN 알고리즘을 고려한 모델[7]이 제시되었다. 또 다른 연구[8]에서는 기존 DQN 알고리즘을 뛰어넘는, PER(Prioritised Experience Replay)와 고정된 Q 타겟(Q-targets)을 기반으로 DDDQN(Double Dueling Deep Q-Network) 모델을 설계하여 오버피팅(Over-fitting) 문제의 해결 및 성능 향상을 이끌어내었다. 해당 연구는 빅데이터 기반의 사례 연구를 이용하여 양적인 QoE(Quality-of-Experience)와 질적인 QoS(Quality-of-Service)를 모두 고려하였다. 마지막으로, MARL 알고리즘 기반으로 분산형 사물인터넷(IoT) 장치로부터 다중 UAV가 수집하는 데이터양을 최대화하는 연구가 있다. 해당 연구에서는 시스템에 존재하는 UAV의 방향을 조정함으로써 수집되는 데이터양을 최적화하는 목적을 달성하였다. 제안된 네트워크 구조[9]는 DRL(Deep Reinforcement Learning)을 통한 Dec-POMDP(Decentralized-Partially Observable Markov Decision Process) 기법으로 변환되어 데이터 수집 작업을 UAV 간 분배하여 최적화하였다. 이처럼 여러 제약 조건을 동시에 고려하면서도 최적화 목적을 수행하는 UAV 시스템이 연구되는 추세이다.

2-3 데이터 손실 최소화

UAV 기반 환경에서 발생하는 데이터 손실은 네트워크 연결 방해(Interference), 통신량의 급증, 또는 직접적인 위험 등을 받을 때 발생한다. 데이터 손실은 개인 정보 유출, 저장되지 않은 정보의 삭제 그리고 촬영 결과물 퀄리티 저하의 이유로 문제가 되고 있다. 문제를 해결하기 위하여 다중 에이전트들끼리의 정보 교환 없이 사용자와 UAV가 상호작용하는 상황에서 자원 할당에 따른 보상을 최대화하려는 목적의 연구가 진행되었다. 첫 번째 연구에서는 모든 에이전트가 독립적으로 의사결정 알고리즘을 실행하지만, Q-러닝 기반의 공통적인 다중 강화학습 구조[10]를 공유하는 에이전트 독립 방법이 제안되었다. 두 번째 연구[11]에서는 사물인터넷 노드에서의 버퍼 오버플로우(Overflow)와 채널 손실로 인한 데이터 전송 오류를 줄이기 위하여 비행경로와 데이터 수집 스케줄링을 POMDP(Partially Observable Markov Decision Process)로 공식화하여 UAV가 수집하는 데이터양을 최대화하였다. 이 연구를 통해 대규모 환경 감지 기술 발전을 위한 사물인터넷 네트워크에서의 데이터를 수집 과정에서 데이터 손실을 줄이는 것이 가능해졌다. 마지막으로 DQN 알고리즘을 개선한 DDPG(Deep Deterministic Policy Gradient) 알고리즘을 기반으로 DeFRAS(Deep Reinforcement Learning based Flight Resource Allocation Scheme)을 제안하여 UAV의 방향과 순간 속도를 조절하여 데이터 손실을 최소화한 연구[12]도 있다.

III. 결론

강화학습 알고리즘을 도입하면서 UAV 시스템은 더 높아진 기동성과 넓어진 활용 범위를 가지게 되었고, 자체적인 물리적 한계를 극복해 나가며 급속도로 발전해오고 있다. 특히, 최근 정보 보안 및 위험 방지 분야에 주목도가 높아지면서 자원 할당 중 데이터 할당 최적화 연구의 발전이 전망되고 있다. 본 논문에서는 자원 할당 최적화 목적의 UAV 기반 시스템에 도입한 다중 에이전트 강화학습을 데이터의 양질 및 서비스 최적화 목적과 데이터 손실의 최소화 목적으로 분류하여 최신 연구 동향을 살펴본다. 다양한 데이터 종류와 환경 특성에 따라 상황별 고려 요소를 추가하

여 확장하는 것이 앞으로의 연구 과제이다.

ACKNOWLEDGMENT

This research is funded by Ministry of Health and Welfare (HI19C0842).

참 고 문 헌

- [1] B. Li, Z. Fei, and Y. Zhang, "UAV Communications for 5G and Beyond: Recent Advances and Future Trends," *IEEE Internet Things Journal*, vol. 6, no. 2, pp. 2241 - 2263, Apr. 2019.
- [2] A. Fotouhi et al., "Survey on UAV Cellular Communications: Practical Aspects, Standardization Advancements, Regulation, and Security Challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3417-3442, 2019.
- [3] A. Carrio, et al., "A Review of Deep Learning Methods and Applications for Unmanned Aerial Vehicles," *Journal of Sensors*, vol. 2017, no. 3296874, Aug. 2017.
- [4] L. Canese, et al., "Multi-Agent Reinforcement Learning: A Review of Challenges and Applications," *Applied Science*, 2021.
- [5] K. Zhang, Z. Yang, and T. Basar, "Multi-Agent Reinforcement Learning: A Selective Overview of Theories and Algorithms," *arXiv preprint arXiv:1911.10635*, 2021.
- [6] C. Caillouet, F. Giroire, and T. Razafindralambo, "Efficient data collection and tracking with flying drones," *Ad Hoc Networks*, vol. 89, pp. 35-46, 2019.
- [7] F. Wu, et al., "Cellular UAV-to-Device Communications: Trajectory Design and Mode Selection by Multi-Agent Deep Reinforcement Learning," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4175-4189, 2020.
- [8] W. Guo, "Partially Explainable Big Data Driven Deep Reinforcement Learning for Green 5G UAV," *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, pp. 1-7, 2020.
- [9] H. Bayerlein, et al., "Multi-UAV Path Planning for Wireless Data Harvesting With Deep Reinforcement Learning," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1171-1187, 2021.
- [10] J. Cui, Y. Liu and A. Nallanathan, "Multi-Agent Reinforcement Learning-Based Resource Allocation for UAV Networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 729-743, Feb. 2020.
- [11] K. Li, et al., "Joint Flight Cruise Control and Data Collection in UAV-Aided Internet of Things: An Onboard Deep Reinforcement Learning Approach," *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 9787-9799, 2021.
- [12] K. Li, W. Ni and F. Dressler, "LSTM-characterized Deep Reinforcement Learning for Continuous Flight Control and Resource Allocation in UAV-assisted Sensor Network," *to appear in IEEE Internet of Things Journal*, 2021.