

무선 데이터 수집과 에너지 전송을 위한 Deep Q-network 기반 UAV 경로 계획

이충녕, 채승호*

한국공학대학교

lc9902130509@tukorea.ac.kr, *shchae@tukorea.ac.kr

Deep Q-network Based UAV Path Planning for Wireless Data Collection and Energy Transfer

Chungnyeong Lee, Seong Ho Chae*

Tech University of Korea

요약

본 논문에서는 제한된 배터리를 가지는 UAV(Unmanned Aerial Vehicle) 및 IoT(Internet of Things) 디바이스에 대해, 무선 에너지 전송과 데이터 수집이 가능한 UAV의 경로 계획 방법을 제안한다. 구체적으로, 모든 IoT 디바이스의 AoI(Age of Information)를 최소화하는 DQN(Deep Q-network) 기반 UAV 경로 계획을 제안한다. 시뮬레이션을 통해 제안된 경로 계획 방법과 무작위 경로 계획 방법을 IoT 디바이스의 AoI와 배터리 잔량을 비교함으로써 제안된 경로 계획 방법이 기존 경로 계획 기법 대비 우수한 성능을 가짐을 보였다.

I. 서론

최근 UAV(Unmanned Aerial Vehicle)를 활용한 다양한 상업 서비스들(물류 배송, 감시 등)이 활발히 개발되고 있다[1]. 특히, UAV를 활용한 IoT(Internet of Things) 디바이스 데이터 수집 어플리케이션이 사람들의 많은 관심을 받고 있으며, 최근 심층 강화학습을 활용한 다양한 UAV 데이터 수집 경로 계획 방법이 제안되었다[2],[3]. 구체적으로, UAV가 비행하며 IoT 디바이스로 에너지를 전송하고 IoT 디바이스로부터 데이터 수집하는 환경에서, DDPG(Deep Deterministic Policy Gradient) 알고리즘을 활용하여 데이터의 AoI(Age of Information)를 최소화하는 UAV 경로 계획법이 제안되었다[2]. 또한, [3]에서는 DQN(Deep Q-network) 알고리즘을 활용하여 IoT 디바이스로부터 데이터 수집 양을 최대화하는 UAV 경로 계획 방법이 제안되었다. 하지만, UAV의 제한된 배터리 용량[1],[4]은 데이터 수집 및 에너지 전송을 위한 UAV 경로 계획에 매우 큰 영향을 미침에도 불구하고, 위의 연구들은 해당 영향을 간과하였다. 따라서, 본 논문에서는 UAV의 제한된 배터리 제약 조건을 고려하여 AoI를 최소화하기 위한 DQN 기반 UAV 경로 계획 방법을 제안하고 기존 경로 계획 방법들과 성능을 비교 분석한다.

II. 시스템 구성

그림 1은 본 논문에서 고려하는 UAV를 활용한 무선 에너지 전송 및 데이터 수집 시스템을 보여준다. 가로 및 세로 길이 d 의 셀 $M \times M$ 개로 구성된 그리드 월드(Grid World)에서 복수 개의 IoT 디바이스와 중심점에 위치한 지상 무선 충전소가 서로 위치 중복 없이 존재함을 가정한다. UAV는 그리드 월드 내를 연속적으로 비행하여 이동하는데, IoT 디바이스 중 하나의 위치로 이동 및 호버링(hovering)을 통해 무선 에너지 전송 및 데이터 수집을 수행하고, 충전소 위치로 이동 및 호버링을 통해 배터리를 무선 충전할 수 있다. UAV는 모든 IoT 디바이스와 중앙 지상 무선 충전소의 위치를 알고 있음을 가정한다.

전체 시간 T 를 지속시간 τ 의 N 개의 시간 스텝(time step)으로 분할하고 시스템 환경은 각 시간 스텝마다 준정적(Quasi-static) 상태라 가정한다. UAV와 IoT 디바이스들이 각각 서로 다른 배터리 용량을 가지며, 항시 일

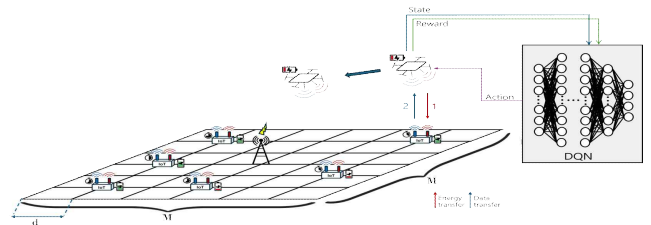


그림 1. UAV를 활용한 무선 에너지 전송 및 데이터 수집 시스템

정 수준 이상의 배터리 상태를 유지하도록 고려한다. UAV 에너지 전송 및 데이터 수집 시스템은 TDD(Time Division Duplex) 방식으로 작동한다. 구체적으로, 각 시간 스텝은 두 부분으로 나뉘어, 첫 번째 부분은 에너지 전송을 수행하고 두 번째 부분은 데이터 수집을 수행한다. UAV의 데이터 수집을 위한 에너지 소모는 UAV의 비행 및 에너지 전송에 비해 무시할 수 있을 정도로 적기 때문에 우리는 UAV의 비행 및 에너지 전송의 에너지 소모만을 고려한다[2].

본 논문에서는 수집된 정보의 최신 정도를 측정하기 위해 성능 메트릭(Metric)으로 AoI를 고려한다. AoI는 UAV가 IoT 디바이스로부터 수집한 정보가 해당 IoT 디바이스에서 생성된 이후 경과된 시간으로 정의된다. 시간 스텝 t 에서 UAV가 IoT 디바이스 i 로부터 성공적으로 데이터를 수집하면, IoT 디바이스 i 의 AoI는 다음과 같이 정의된다.

$$A_i^t = t - C_i(t), \quad (1)$$

여기서 $C_i(t)$ 는 IoT 디바이스 i 에서 데이터가 업데이트된 시간 스텝을 나타낸다. 본 논문의 최종 목표는 UAV와 IoT 디바이스의 배터리 제약을 고려하여 모든 IoT 디바이스의 AoI를 최소화하는 UAV의 경로 계획을 도출하는 것이 목적이다.

III. 제안된 Deep Q-network 기반 UAV 경로 계획 알고리즘

본 장에서는 IoT 디바이스들의 AoI를 최소화하기 위한 DQN 기반 UAV 경로 계획 방법을 제안한다. 이를 위해, 먼저 시스템 모델을 다음과 같은 MDP(Markov Decision Process)으로 정의한다.

1) 에이전트(Agent): UAV가 에이전트로서 환경과 상호작용한다.

- 2) **상태(State)**: UAV는 각 시간 스텝마다 자신의 배터리 잔량 및 위치 정보를 알고, IoT 디바이스 및 중앙 무선 충전소의 위치를 상태 정보로 안다고 가정한다. 이때 UAV는 고정된 IoT 디바이스와 중앙 무선 충전소의 위치 정보만 알고 있고, IoT 디바이스의 배터리 정보는 모른다.
- 3) **행동(Action)**: UAV는 매시간 스텝마다 셀 크기인 d 만큼 동,서,남,북 방향으로 움직이며 일정 크기의 비행 에너지를 소모한다. UAV가 비행 후 해당 위치에서 IoT 디바이스와 연결되면, UAV는 무선 에너지 전송 및 데이터 수집을 수행하고 무선 충전 에너지를 소모한다. 또한 UAV가 비행 후 중앙 무선 충전소에 연결되면, UAV는 자신의 배터리의 일정 비율을 충전한다. 이때 UAV는 모든 시간 스텝 동안 자신과 IoT 디바이스의 배터리 잔량이 일정 이상 유지하도록 행동한다.
- 4) **보상(Reward)**: 우리는 보상을 세 가지 경우로 나눠 정의한다. 첫째, UAV의 위치가 IoT 디바이스 위치와 동일한 경우 임의의 크기의 양의 보상을 받는다. 둘째, UAV의 위치가 중앙 무선 충전소 위치와 동일한 경우에도 임의의 크기의 양의 보상을 받는다. 마지막으로, UAV의 배터리가 방전된 경우는 음의 크기의 페널티를 받는다. 이렇게 설정함으로써 UAV는 에너지 제약을 만족하면서 IoT 디바이스에게 에너지를 전송하고 데이터를 수집하도록 행동을 유도한다.

그림 1은 본 논문에서 제안한 DQN 기반 UAV 경로 계획 방법을 보여준다. 본 논문에서는 페이지 제약으로, 구체적인 DQN 알고리즘 동작에 대한 기술은 생략하며, 참고문헌 [5]를 참조하면 수행 가능하다.

IV. 시뮬레이션 결과

본 장에서는 제안된 경로 계획 방법과 무작위 알고리즘 기반 경로 계획 방법과 AoI 및 배터리 잔량 성능을 비교 분석한다. 무작위 알고리즘은 UAV가 무작위로 움직이며 배터리 잔량이 일정 이하일 때 중앙 무선 충전소로 이동해 충전하는 방식이다. 학습을 위해 사용된 DQN 알고리즘의 하이퍼파라미터는 참고문헌[5]과 동일하게 설정하였다.

그림 2는 한 에피소드 동안 UAV의 비행 경로를 보여준다. 중앙의 '●' 기호는 중앙 무선 충전소, '×' 기호는 IoT 디바이스를 나타낸다. 그림 2를 통해, UAV가 모든 IoT 디바이스를 방문하여 에너지 전송 및 데이터를 수집하고, 무선 충전소 방문을 통해 배터리를 충전한 것을 확인할 수 있다.

그림 3은 UAV 비행 시간 스텝에 따른 각 IoT 디바이스의 AoI 변화를 보여준다. 이를 통해 우리가 제안한 경로 계획 알고리즘의 AoI 성능이 비교 알고리즘보다 뛰어난 것을 확인할 수 있다. 표 1은 알고리즘에 따른 UAV 및 IoT 디바이스의 평균 배터리 잔량을 보여준다. 결과적으로, UAV의 배터리 잔량은 제안된 알고리즘보다 무작위 알고리즘이 더 높고, IoT 디바이스들의 배터리 잔량은 제안된 알고리즘이 더 높은 것을 확인할 수 있다. 이는 그림 3에서 확인할 수 있듯이, 무작위 알고리즘이 IoT 디바이스를 거치는 횟수가 적어서 AoI가 높고, 그에 따라 IoT 디바이스로의 에너지 전송량도 부족하기 때문이다.

	배터리 잔량	
	제안된 알고리즘	무작위 알고리즘
UAV	17%	30%
IoT 디바이스 1	36%	3%
IoT 디바이스 2	16%	4%

표 1. 알고리즘에 따른 UAV 및 IoT 디바이스 평균 배터리 잔량 비교

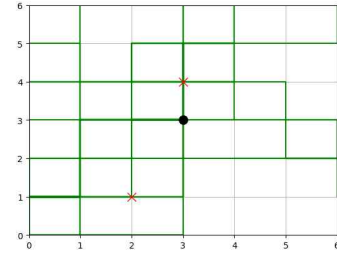


그림 2. UAV 비행 경로 예시

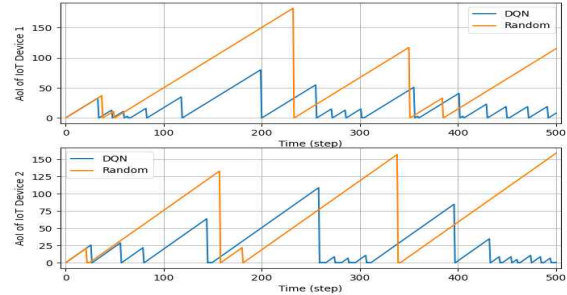


그림 3. UAV 비행 시간 스텝에 따른 IoT 디바이스의 AoI 변화

V. 결론

본 논문에서는 제한된 배터리를 가지는 UAV 및 IoT 디바이스에 대해, 무선 에너지 전송과 데이터 수집이 가능한 UAV의 경로 계획 방법을 제안하였다. 구체적으로, 모든 IoT 디바이스의 AoI를 최소화하기 위한 DQN 기반 UAV 경로 계획을 제안하였다. 시뮬레이션 결과를 통해 제안한 경로 계획 알고리즘이 비교 알고리즘보다 AoI 성능이 우수함을 보였다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 ICT혁신인재4.0사업(IITP-2024-RS-2022-00156326, 50%)과 지역지능화혁신인재양성사업(IITP-2024-2020-0-01741, 50%) 연구임

참고 문헌

- [1] S. Lim, S. H. Chae, and H. Lee, "RE-ORA: Residual Energy-Aware Online Random Access for Improving the Lifetime of Slotted ALOHA-Based Swarming Drone Networks," *IEEE Access*, vol. 9, pp. 45504-45511, Mar. 2021.
- [2] O. S. Oubbati, M. Atiquzzaman, H. Lim, A. Rachedi, and A. Lakas, "Synchronizing UAV teams for timely data collection and energy transfer by deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6682-6697, Jun. 2022.
- [3] C. Lee and S. H. Chae, "DQN Based UAV Path Planning for Wireless Data Harvesting," in Proc. of *KICS Winter Conf.*, Feb. 2023.
- [4] S. Lee, S. Lim, S. H. Chae, B. C. Jung, C. Y. Park and H. Lee, "Optimal frequency reuse and power control in multi-UAV wireless networks: Hierarchical multi-agent reinforcement learning perspective," *IEEE Access*, vol. 10, pp. 39555-39565, Apr. 2022.
- [5] H. van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in Proc. of the *30th AAAI Conf on Artificial Intelligence*, vol. 30, no. 1, pp. 2094-2100, Mar. 2016.