

레이저 충전 기반 UAV 네트워크에서 전송-충전 트레이드오프를 고려한 분산 심층강화학습 기반 위치 및 전송전력 최적화 기법

전지민, 이호원

아주대학교

jimin0516@ajou.ac.kr, howon@ajou.ac.kr

Distributed Deep Reinforcement Learning-Based Positioning and Transmission Power Optimization Considering Communication-Charging Trade-Off in Laser Charging-Based UAV Networks

Jimin Jeon, Howon Lee

Ajou University

요약

무인 항공기(Unmanned Aerial Vehicle, UAV)는 제한된 온보드 배터리 용량으로 인해 운영 수명에 심각한 제약을 받고 있다. 이러한 문제를 해결하기 위해, 다양한 충전 기술들이 연구되어왔으며, 그 중 레이저 충전 기술은 고출력의 에너지를 UAV에 무선으로 전송함으로써 지속적인 운영을 가능하게 한다. 본 연구는 UAV가 지상 사용자에게 다운로드 서비스를 제공하기 위해 기지국 역할을 수행하는 레이저 충전 기반 다중 UBS(UAV-based Base Station) 네트워크를 고려한다. 이러한 네트워크 환경에서, 에너지 효율성과 통신 처리량을 동시에 향상시키기 위해, 본 논문에서는 DQN(Deep Q-Network) 기반 UBS 위치 및 전송 전력 제어 기법을 제안한다. 제안된 기법은 에너지 충전 및 소비, 통신 품질 등 상태 정보를 바탕으로 통신 성능과 에너지 효율 간의 트레이드오프를 고려한다. 시뮬레이션을 통해 제안방안이 기존 기법 대비 통신 성능과 에너지 효율 모두에서 우수한 성능을 가지는 것을 보인다.

I. 서론

무인항공기(Unmanned Aerial Vehicle, UAV)는 저비용, 유연한 배치, 운용의 용이성 등의 장점으로 다양한 분야에 활용되고 있다. 특히 6세대 이동통신에서는 UAV가 지상 사용자에게 서비스를 제공하는 UAV 기반 기지국(UAV-based Base Station, UBS)의 활용을 위한 연구도 활발히 진행되고 있다 [1]. UBS는 3차원 공간에서 위치를 유연하게 조정할 수 있어 LoS(Line-of-Sight) 링크 형성이 용이하고, 고품질 통신이 가능하다. 그러나 제한된 배터리 용량으로 운용 시간이 짧다는 한계가 있으며, 이를 보완하기 위한 다양한 에너지 공급 방식이 제안되었다. 그 중 레이저 충전 기술은 수십 km 거리에서도 안정적인 에너지 전송이 가능해 장기 운용에 적합한 대안으로 주목받고 있다 [2]. UBS 네트워크의 효율적 운용을 위해서는 위치 및 전송 전력의 정교한 제어가 필수적이며, 특히 통신 품질과 에너지 효율을 동시에 고려해야 한다. 본 논문에서는 이러한 두 목표를 균형 있게 달성하기 위해, 레이저 충전 기반 UBS 네트워크를 대상으로 심층 강화학습 기반 제어 기법을 제안한다.

II. 시스템 모델

본 논문에서는 그림 1과 같이 N 개의 공중 UBS, U 개의 지상 사용자(Ground User, GU), L 개의 지상 레이저 충전 스테이션(Laser Charging Station, LCS)으로 구성된 무선 통신 네트워크 환경을 고려한다. 지상의 LCS는 레이저 빔을 통해 UBS에 에너지를 지속적으로 전달하고, UBS는 GU에 다운로드 통신 서비스를 제공한다. 또한, UBS가 LCS를 통해 충전을 하면서 동시에 지상 노드와 통신할 수 있다고 가정한다. UBS는 각 타임 슬롯마다 에이전트로서 행동을 수행한다.

II.1 A2G 채널 모델

A2G(Air-to-Ground) 환경에서 LoS를 고려한 확률적 채널 모델은 다음과 같다 [3].

$$D_i[n] = \sum_{j=1}^U b_{i,j} \log_2 \left(1 + \frac{P_{i,j}^t - \Phi_{i,j}^{avg}}{N_0 b_{i,j} + \Gamma_j[n]} \right), \quad (1)$$

여기서, $b_{i,j}$ 는 UBS i 가 GU j 에게 할당된 주파수 대역폭이다, $P_{i,j}^t$ 는 UBS i 의 통신 전송전력이며, $\Phi_{i,j}^{avg}$ 는 LoS 및 NLoS 링크를 고려한 평균 경로 손실이다. N_0 는 화이트 가우시안 잡음의 전력 스펙트럼 밀도이며,

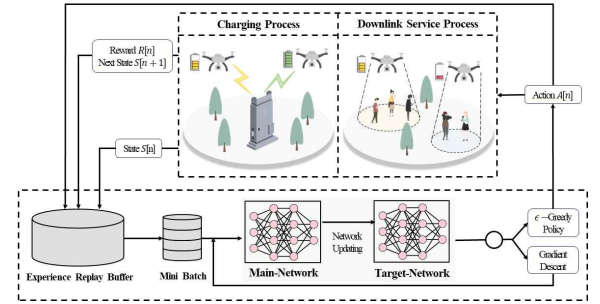


그림 1. 레이저 충전 기반 UAV 네트워크에서 전송-충전 트레이드 오프를 고려한 DQN 프레임워크
Fig. 1. DQN framework considering communication-charging trade-off in laser charging-based UAV networks.

$\Gamma_j[n]$ 은 타임 슬롯 n 에서의 동일 채널 간섭이다.

II.2 레이저 충전 모델

LCS k 로부터 UBS i 에서 충전된 에너지는 자유공간 광통신(Free-space Optical Communication, FSO)를 기반으로 다음과 같이 표현된다 [4].

$$E_i^{charged}[n] = \zeta_i^m \zeta_i^e \zeta_i^c P_{k,i}^r[n], \quad (2)$$

위 수식에서 ζ_i^m , ζ_i^e , ζ_i^c 는 각각 모터 효율, 배터리 방전 효율, 충전 효율을 의미한다.

II.3 UAV 전력 소비 모델

본 연구에서는 통신 에너지, 호버링 에너지, 이동 에너지를 포함하는 UAV 전력 소비 모델을 고려한다 [5], [6], [7].

$$E_i^{used}[n] = E_i^{comm}[n] + E_i^{hover}[n] + E_i^{mov}[n], \quad (3)$$

위 수식은 다음과 같이 표현될 수 있다.

$$E_i^{used}[n] = (u_i[n] \cdot P_{k,i}^t + P_i^{ca}) \cdot t_{comm} + \left(\frac{MG^{3/2}}{\sqrt{2} \rho \pi \beta^2} \right) \cdot t_{hover} + \left(P_h \frac{\sqrt{\Delta x^2 + \Delta y^2}}{v_h} + P_v \frac{\Delta h}{v_v} \right) \cdot t_{mov}. \quad (4)$$

여기서, t_{comm} , t_{hover} , t_{mov} 은 각각 통신, 호버링, 이동에 소요되는 시간을 의미하며, $u_i[n]$ 은 타임슬롯 n 에서 UAV i 와 연결된 GU의

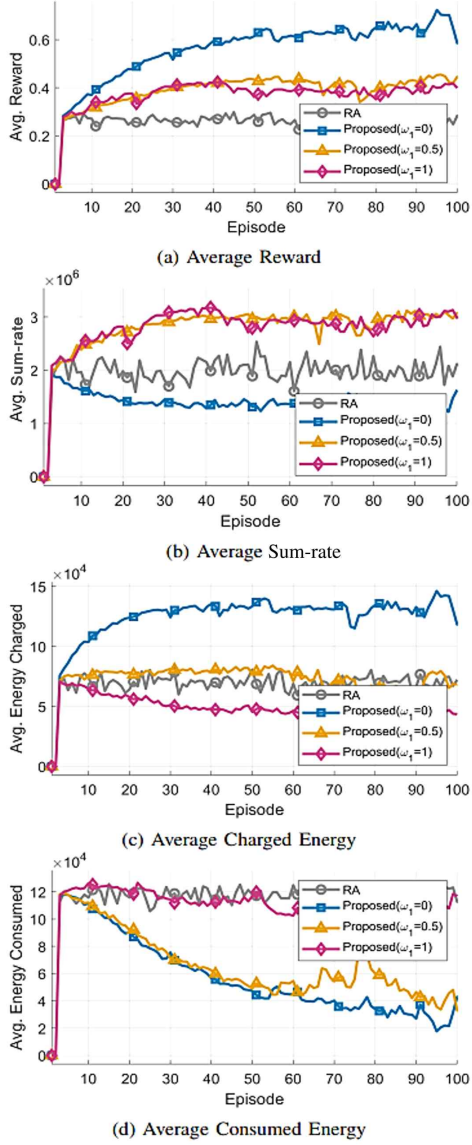


그림 2. 2 에이전트 시나리오에서 에피소드 진행에 따른 평균 보상, 통신품질, 충전 및 소모 에너지 성능 비교
Fig. 2. Performance comparison of average reward, communication quality, charging energy and energy consumption over episodes in 2-agent scenarios.

수이다. P_i^{cu} 는 온보드 회로 전력 소비를 의미한다. M 은 UAV의 로터 수, G 는 뉴턴 단위의 추력, ρ 는 공기의 유체 밀도, β 는 로터 디스크 반경이다. Δx 와 Δy 는 수평 이동 거리, Δh 는 수직 이동 거리를 나타낸다. P_h 와 P_v 는 각각 수평 및 수직 방향의 전력 소비를 나타내며, v_h 와 v_v 는 각 방향의 속도이다.

III. 분산 심층강화학습 기반 위치 및 전송전력 최적화 기법

본 연구에서는 DQN 기반의 위치 및 전송 전력 제어 기법을 제안한다. 각 UBS는 독립된 에이전트로 작동하며, 환경 변화에 따라 실시간으로 상태를 인식하고 최적의 행동을 학습한다.

- 상태(State): 에이전트 i 의 상태는 타임 슬롯 n 에서 UBS i 의 3차원 좌표 및 통신 품질, 충전량, 그리고 소모량으로 정의된다.

$$s_i[n] = [x_i[n], y_i[n], h_i[n], D_i[n], E_i^{charged}[n], E_i^{used}[n]]. \quad (5)$$

- 행동(Action): 에이전트 i 는 각 타임 슬롯에서 3차원 이동과 전송 전력을 제어하는 9개의 행동 중 하나를 취할 수 있다. 행동 집합은 다음과 같

이 정의된다.

$$a_i[n] \in \{\pm \Delta x, \pm \Delta y, \pm \Delta h, \pm \Delta P_{i,j}^T, 0\}. \quad (6)$$

- 보상(Reward): 본 연구에서는 통신 성능과 에너지 효율을 함께 고려하기 위해, 두 요소를 결합한 형태의 보상 함수를 설계하였다.

$$r_i[n] = \omega_1 D_i[n] + \omega_2 (E_i^{charged}[n] - E_i^{used}[n]), \quad (7)$$

여기서 ω_1 은 통신 품질에 대한 가중치 계수이고 ω_2 는 충전과 소비 사이의 에너지 차이에 대한 가중치 계수이다. 두 계수는 $\omega_1 + \omega_2 = 1$ 의 관계로 설정된다. ω_1 이 증가할수록 에이전트는 통신 품질 향상에 더 많은 가중치를 두게 되고, ω_2 가 증가할수록 에너지 효율성에 더 많은 가중치를 두게 된다.

IV. 시뮬레이션 결과 및 결론

그림 2는 2 에이전트 시나리오에서 에피소드 진행에 따른 제안방안 및 벤치마크의 성능을 나타낸다. 그림 2(a)에서 모든 제안 방안이 매 슬롯마다 무작위 행동을 하는 RA(Random Action)보다 높은 평균 보상을 달성함을 보인다. 그림 2(b)의 평균 통신 품질에서는 $\omega_1 = 1$ 인 경우에 가장 우수한 통신 성능을 보이는 반면, 그림 4(c)와 (d)의 충전 에너지 및 소모 에너지 측면에서는 $\omega_1 = 0$ 인 경우에 가장 효율적인 에너지 관리를 보인다. $\omega_1 = 0.5$ 인 경우는 통신 품질과 에너지 효율 사이의 균형을 유지하며, 실제 네트워크 운영에서 효과적인 절충안을 제공할 수 있음을 시사한다. 시뮬레이션 결과를 통해 기존 벤치마크 기법에 비해 제안 기법이 평균 보상, 데이터 품질, 에너지 활용도 측면에서 우수한 성능을 달성함을 확인하였다.

ACKNOWLEDGMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No.2022-0-00704, 초고속 이동체 지원을 위한 3D-NET핵심 기술 개발), 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(RS-2024-00396992, 저궤도 위성통신 핵심 기술 기반 큐브 위성 개발), 그리고 한국연구재단의 지원(RS-2025-00563401, 3차원 공간에서 에너지 효율적 멀티 레벨 AI-RAN 구현을 위한 AI-for/and-RAN 핵심 원천기술 연구)을 받아 수행된 연구임.

참고 문헌

- [1] H. Lee, B. Lee, H. Yang, J. Kim, S. Kim, W. Shin, B. Shim, and H. V. Poor, "Towards 6g hyper-connectivity: Vision, challenges, and key enabling technologies," *Journal of Communications and Networks*, vol. 25, no. 3, pp. 344-354, 2023.
- [2] A. K. Majumdar, "Free-space laser communication performance in the atmospheric channel," *Journal of Optical and Fiber Communications Reports*, vol. 2, no. 4, pp. 345-396, 2005.
- [3] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal lap altitude for maximum coverage," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 569-572, 2014.
- [4] A. K. Majumdar, "Free-space laser communication performance in the atmospheric channel," *Journal of Optical and Fiber Communications Reports*, vol. 2, p. 345-396, 2005.
- [5] H. Qi, Z. Hu, H. Huang, X. Wen, and Z. Lu, "Energy efficient 3-D UAV control for persistent communication service and fairness: A deep reinforcement learning approach," *IEEE Access*, vol. 8, pp. 53 172-53 184, 2020.
- [6] K. Dorling, J. Heinrichs, G. G. Messier, and S. Magierowski, "Vehicle routing problems for drone delivery," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 1, pp. 70-85, 2017.
- [7] J. Lu, S. Wan, X. Chen, Z. Chen, P. Fan, and K. B. Letaief, "Beyond empirical models: Pattern formation driven placement of uav base stations," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 3641-3655, 2018.