

# 심층강화학습 활용한 레이저 충전 기반 다중 UAV 네트워크 성능 향상 기법 연구

전지민<sup>1</sup>, 이승민<sup>1</sup>, 유희정<sup>2</sup>, 이호원<sup>1</sup>

한경대학교 전자전기공학부<sup>1</sup>, 고려대학교 전자및정보공학과<sup>2</sup>

{jimin0516, julsin1, hwlee}@hknu.ac.kr<sup>1</sup>, heejungyu@korea.ac.kr<sup>2</sup>

## Performance Improvement of Laser-Charged Multi-UAV Networks Based on Deep Reinforcement Learning

Jimin Jeon<sup>1</sup>, Seungmin Lee<sup>1</sup>, Heejung Yu<sup>2</sup>, and Howon Lee<sup>1</sup>

Hankyong National University<sup>1</sup>, Korea University<sup>2</sup>

### 요 약

본 논문에서는 레이저 충전 기반 다중 UAV(Unmanned Aerial Vehicles) 네트워크를 고려하며, 지상 사용자들에게 하향링크 커버리지를 제공하기 위해 UAV가 공중 기지국의 역할을 수행한다. 특히, UAV의 배터리 제약 문제를 해결하기 위해 레이저 충전 기반의 다중 UAV 통신 프레임워크 환경에서 UBS(UAV Base-Station)의 아웃티지 수를 고려하며 총 비행시간을 연장하기 위한 DQN(Deep Q-Network) 기반 UBS 위치 제어 기법을 제안한다. 시뮬레이션을 통해 제안방안이 고정된 배치 기반 레이저 충전 UBS 제어 기법, 그리고 랜덤 행동 기반 레이저 충전 UBS 제어 기법보다 더 좋은 성능을 가짐을 확인한다.

### I. 서론

UAV(Unmanned Aerial Vehicles)는 높은 이동성, 배치의 유연성, 그리고 조작성의 용이성과 같은 다양한 장점들로 인해 최근 다양한 분야에서 많은 관심을 받고 있으며 특히, 6G에서는 핵심 기술이자 서비스로 중요하게 고려되고 있다[1]. 하지만, UAV는 크기와 무게 제약으로 인해 제한된 배터리 용량을 가진다는 치명적인 단점을 가지는데, 이를 해결하기 위해 다양한 연구가 이루어지고 있다[2]. 본 논문에서는 이러한 UAV의 배터리 제약 문제를 극복하기 위해 레이저 충전 기반 UBS(UAV Base-Station) 통신 환경에서의 DQN(Deep Q-Network) 기반 UBS 위치 제어 기법을 제안한다. 시뮬레이션을 통해 제안방안이 총 비행시간 관점에서 다른 비교방안들에 비해 좋은 성능을 가짐을 보인다.

### II. 시스템 모델

본 논문에서는 그림 1과 같이  $N$  공중 UBS,  $U$  지상 사용자, 그리고  $L$  지상 레이저 충전소로 구성된 무선 통신 네트워크 환경을 가정한다. 여기에서, 지상의 레이저 충전소가 공중의 UBS에 지속적으로 에너지를 전송하고 UBS가 지상 사용자들에게 하향링크 통신 커버리지를 제공한다. 전체 서비스 시간  $T$ 는 길이가  $t^{slot}$ 인 time slots으로 나누어지며 매 time slot마다 UBS는 다음 위치로 이동한다. 이 때, 지상 레이저 충전소를 통한 충전과 사용자에게 제공되는 통신은 지속적으로 이루어진다고 가정한다. 또한, UBS는  $C$  채널을 가지고 있으며 지상 사용자는 UBS로부터 수신한 신호의 RSSI가  $RSSI^{th}$ 보다 클 때, UBS로부터 서비스를 제공받으며 그 반대의 경우에는 아웃티지 사용자가 된다.

### II.1 A2G(Air-to-Ground) Channel Model

A2G(Air-to-Ground) 환경에서 LoS(Line-of-Sight)를 고려한 확률적 채널 모델은 다음과 같다[3].

$$R_{i,j}[n] = B_{i,j}[n] \log_2 \left( 1 + \frac{P_{i,j}^r \phi_{i,j}^{LoS}}{N_0 B_{i,j}[n]} \right) p^{LoS} + B_{i,j}[n] \log_2 \left( 1 + \frac{P_{i,j}^r \phi_{i,j}^{NLoS}}{N_0 B_{i,j}[n]} \right) p^{NLoS}, \quad (1)$$

위 수식에서,  $R_{i,j}[n]$ 은 time slot  $n$ 에서 UBS  $i$ 와 사용자  $j$ 간의 평균 데이터 전송률이다.  $p_{i,j}$ 는 UBS  $i$ 에 의해 사용자  $j$ 에 할당된 전송 전력,  $N_0$ 는 white Gaussian noise의 전력 스펙트럼 밀도,  $B_{i,j}$ 는 UBS  $i$ 에 의해 사용자  $j$ 에 할당된 대역폭, 그리고  $\phi_{i,j}^{LoS}$ 와  $\phi_{i,j}^{NLoS}$ 는 각각 UBS  $i$ 로부터 사용자  $j$ 로의 LoS와 non-LoS 경로 손실을 의미한다.

### II.2 UAV Power Consumption and Charging Model

본 논문에서는 UBS와 레이저 충전소 사이의 free space optical 전파를 고려한다. 따라서 UBS  $i$ 가 레이저 충전소  $k$ 로부터 수신한 전력  $P_{k,i}^r$ 는 다음과 같이 표현된다[4].

$$P_{k,i}^r = P_{k,i}^t G_k^t G_i^r \tau_k^t \tau_i^r F. \quad (2)$$

위 수식에서,  $P_{k,i}^r$ 는 레이저 충전소의 전송전력,  $G_k^t$ 와  $G_i^r$ 은 충전소 전송전력이득과 UBS의 수신전력이득을 의미하며  $\tau_k^t, \tau_i^e, \tau_i^r$ 는 각각 레이저 충전소와 UBS의 optical 효율, 그리고 환경적 감소 파라미터를 의미한다.  $F$ 는 free space 경로 손실이다. 수식(2)에서 계산된 수신전력은 UBS에게 안정적인 전압을 제공한다고 가정한다. 따라서 time slot  $n+1$ 에서 UBS  $i$ 의 남은 비행시간  $\Gamma_i[n+1]$ 은 다음과 같이 계산된다. UBS  $i$ 의 남은 비행시간은 레이저 충전을 통해 증가한 비행시간과 이동으로 인해

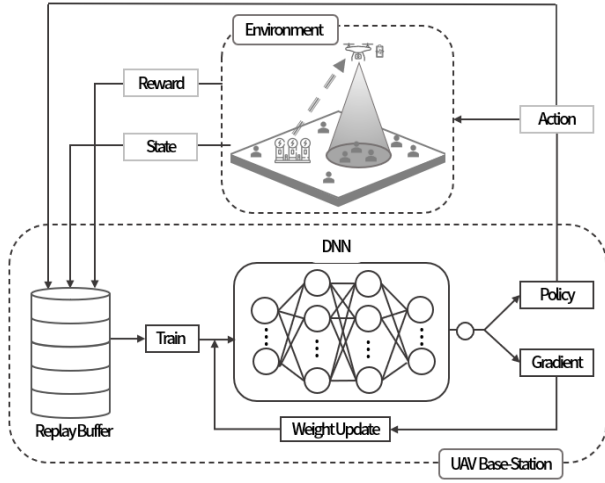


그림 1. 레이저 충전 UBS의 최적 위치 제어를 위한 DQN Framework  
감소한 비행시간 모두를 고려하여 다음과 같이 계산될 수 있다[3].

$$R_i[n+1] = R_i[n] - \Gamma_i^{used}[n+1] + \Gamma_i^{charged}[n+1], \quad (3)$$

$$R_i[n+1] = R_i[n] - \left( t^{slot} + \frac{E_i[n+1]}{I_i^d V_i} \right) + \left( \frac{\eta_i^m \eta_i^e \eta_i^c t^{slot}}{I_i^d V_i} P_{k,i}^r[n+1] \right). \quad (4)$$

수식(3)에서  $E_i[n+1]$ 은 UBS  $i$ 의 운동 에너지이며  $I_i^d$ 와  $V_i$ 는 각각 UAV 작동 전류와 작동 전압을 의미한다. 또한,  $\eta_i^m$ ,  $\eta_i^e$ ,  $\eta_i^c$ 는 각각 본 논문에서 고려하는 UAV의 motor 효율, 배터리 방전 효율, 그리고 충전 회로 converting 효율을 의미한다.

### III. DQN을 활용한 레이저 충전 기반 UBS 위치 제어 기법

본 논문에서 고려하는 문제를 MDP(Markov Decision Process)로 정의하며  $N$ 개의 UBS가 에이전트로서 역할을 수행한다. 제안방안의 DQN 프레임워크는 그림 1과 같다.

#### ■ State

에이전트  $i$ 의 상태는 time slot  $n$ 에서 UBS  $i$ 의 3차원 좌표  $x_i(n), y_i(n), h_i(n)$ 와 사용자에게 할당되지 않은 UAV의 채널 개수  $o_i(n)$ 으로 구성된다.

$$S_i(n) = [x_i(n), y_i(n), h_i(n), o_i(n)]. \quad (5)$$

#### ■ Action

에이전트  $i$ 는 매 time slot마다 3차원 이동으로 구성된 7가지 행동 중 하나의 행동을 취할 수 있다.

$$A_i(n) \in \{+\Delta x, +\Delta y, +\Delta h, -\Delta x, -\Delta y, -\Delta h, 0\}. \quad (6)$$

#### ■ Reward

본 논문에서는 아웃티지 발생을 최소화하며 UBS의 비행시간을 동시에 최대화하는 것을 목표로 한다. 따라서 제안방안의 보상에는 네트워크 성능을 향상시키기 위해 UBS  $i$ 와 모든 지상 사용자 간의 ADR(Average data rate)을 포함하며, 비행시간을 최대화하기 위해 다음 time slot  $n+1$ 에서 충전되는 UBS  $i$ 의 배터리를 의미하는 늘어난 비행시간  $\Gamma_i^{charged}[n+1]$ 을 보상에 포함한다.

$$R_i(n) = \begin{cases} \Gamma_i^{charged}[n+1], & \text{if } o_i[n+1] > 0. \\ -\frac{1}{ADR_i[n+1]}, & \text{Otherwise.} \end{cases} \quad (7)$$

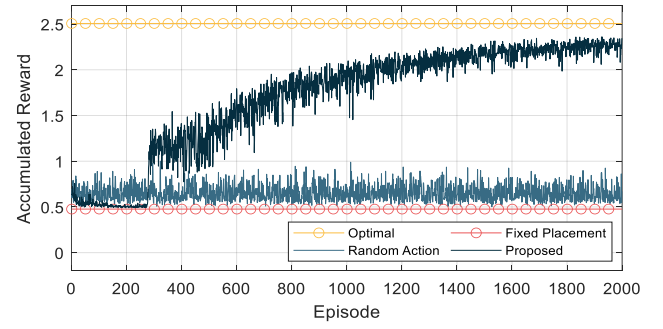


그림 2. 에피소드에 따른 제안방안과 기존방안의 누적 보상

위 수식에서, 아웃티지가 발생한다면 ADR을 보상으로 가지며 보상은 음수의 값을 가진다. 아웃티지가 발생하지 않는다면  $\Gamma_i^{charged}[n+1]$ 을 보상으로 가지며 보상은 양수의 값을 가진다. 이를 통해 DQN 모델이 아웃티지 발생을 최소화하며 충전을 통한 비행시간의 증가를 최대화하는 것을 학습하도록 유도한다.

### IV. 시뮬레이션 결과 분석 및 고찰

시뮬레이션은  $N = 1$ ,  $U = 20$ ,  $L = 1$ ,  $C = 10$ ,  $RSSI_{th} = 5[W]$ ,  $P_t = 200[W]$ ,  $T = 30[min]$ ,  $t^{slot} = 1[min]$ 인 환경에서 수행되었으며, 지상 사용자들은  $200[m] \times 200[m]$  크기의 네트워크 환경에서 랜덤하게 분포된다. 또한 학습은 3000번 에피소드로 진행되며 한 에피소드는 UBS의 배터리가 0이 될 때 종료된다. 그림 2를 통해 제안방안의 학습 진행에 따른 누적 보상이 고정된 배치 기반 레이저 충전 UBS 제어 기법, 그리고 랜덤 행동 기반 레이저 충전 UBS 제어 기법과 같은 비교방안을 대비 높은 값에 수렴함을 확인할 수 있다.

### ACKNOWLEDGMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No. 2021-0-00794, 3차원 공간 이동통신 기술 개발, 50%)과 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2022R1A2C1010602, 50%).

### 참 고 문 헌

- [1] S. Lim, H. Yu and H. Lee, "Optimal Tethered-UAV Deployment in A2G Communication Networks: Multi-Agent Q-Learning Approach," in IEEE Internet of Things Journal, vol. 9, no. 19, pp. 18539-18549, Oct. 2022.
- [2] S. Lee, H. Yu and H. Lee, "Multiagent Q-Learning-Based Multi-UAV Wireless Networks for Maximizing Energy Efficiency: Deployment and Power Control Strategy Design," in IEEE Internet of Things Journal, vol. 9, no. 9, pp. 6434-6442, May 2022.
- [3] W. Liu, L. Zhang and N. Ansari, "Laser Charging Enabled UBS Placement for Downlink Communications," in IEEE Transactions on Network Science and Engineering, vol. 8, no. 4, pp. 3009-3018, Oct. 2021.
- [4] A. Majumdar, "Free-space laser communication performance in the atmospheric channel," J. Opt. Commun. Rep. vol. 2, pp. 345-396, Nov. 2005.