

# A2G 네트워크에서 중앙집중식 심층강화학습 기반 UAV-BS 에너지 효율 성능 분석 이승민\*, 반태원<sup>o</sup>, 이호원\*

한경대학교\*, 경상대학교<sup>o</sup>

{julsin1, hwlee}@hknu.ac.kr, twban35@gnu.ac.kr

## Performance Analysis on Energy-Efficiency of UAV-BS Based on Centralized Deep Reinforcement Learning in A2G Networks

Seungmin Lee\*, Tae-Won Ban<sup>o</sup>, Howon Lee\*

Hankyong National University\*, Gyeongsang National University<sup>o</sup>

### 요약

본 논문에서는 unmanned aerial vehicle-base station (UAV-BS)를 이용하는 공중 무선 네트워크 환경에서, 아웃티지 사용자를 고려하면서 송신 에너지 효율을 최대화하기 위한 심층강화학습 기반 UAV-BS 제어 기법의 성능을 비교 분석한다. Round robin (RR)방식으로 UAV-BS를 순차적으로 제어하는 RR-Deep Q-Network (RR-DQN) 알고리즘, 한 타임스텝에 모든 UAV-BS를 동시에 제어하는 simultaneous control-DQN (SC-DQN) 알고리즘, 그리고 분산형 큐러닝 기반 UAV-BS 제어 알고리즘에 대하여 살펴본다. UAV-BS의 이동성과 전송전력 제어를 행동(action)으로 정의하고, 아웃티지 사용자를 고려한 에너지 효율을 보상(reward)으로 정의하여 그 성능을 비교해 본다.

### I. 서론

Unmanned aerial vehicle (UAV) 기반 air-to-ground (A2G) 네트워크는 차세대 네트워크를 주도할 핵심 서비스 중에 하나로 주목 받고 있으며, 이에 대한 연구가 매우 활발히 수행되고 있다 [1]. 특히 UAV의 제한적인 비행시간 때문에, 네트워크 동작 시간 연장을 위해 에너지 효율개선을 목적으로 한 연구들이 진행되고 있다 [2] [3]. 이에 따라서, 본 논문에서는 사용자 아웃티지를 고려하면서 에너지 효율의 성능을 분석하기 위해 심층강화학습 기반의 알고리즘들에 대하여 살펴본다. Round robin (RR)방식의 RR-Deep Q-Network (RR-DQN), 한 타임스텝에 모든 UAV-BS를 동시에 제어하는 simultaneous control-DQN (SC-DQN) 알고리즘, 그리고 분산형 큐러닝 기반 UAV-BS 제어 알고리즘에 대해서 살펴보고 그 성능을 비교 분석해 본다.

### II. 심층강화학습 기반 UAV-BS 제어 알고리즘

본 논문에서 사용하는 RR-DQN과 SC-DQN 알고리즘은 중앙집중식으로 학습이 진행된다. 따라서, 두 알고리즘에서는 UAV-BS  $u \in U$ 를 제어할 중앙 컨트롤러 (central controller)가 존재하며 중앙 컨트롤러가 에이전트로서 역할을 수행한다. 분산 큐러닝에서는 각 UAV가 에이전트로서 역할을 수행한다. 각 알고리즘에 대한 상태 (state)는 다음과 같이 정의한다.

$$S_i = \begin{cases} [o_1^{t-(U-1)}, o_2^{t-(U-2)}, \dots, o_U^t]^T, & \text{if } \text{mod}(t, U) = 1 \\ [o_2^{t-(U-2)}, o_3^{t-(U-3)}, \dots, o_1^{t+1}]^T, & \text{if } \text{mod}(t, U) = 2 \\ \vdots & \vdots \\ [o_U^t, o_1^{t+1}, \dots, o_{U-1}^{t+U}]^T, & \text{if } \text{mod}(t, U) = 0 \end{cases}, \quad (1)$$

$$S_i = \begin{bmatrix} o_1^t & o_2^t & \dots & o_U^t \\ o_2^t & o_3^t & \dots & o_1^t \\ \vdots & \vdots & \dots & \vdots \\ o_U^t & o_1^t & \dots & o_{U-1}^t \end{bmatrix}, \quad (2)$$

$$S_u = (x_u, y_u, h_u, p_{tu}). \quad (3)$$

수식 (1), (2) 그리고 (3)은 각각 RR-DQN, SC-DQN 그리고 분산 큐러닝에 대한 상태이며,  $o_u = (x_u, y_u, h_u, p_{tu})$ 는 UAV-BS  $u$ 의 3차원 공간에서의 좌표와 전송전력 상태를 나타낸다. 세 알고리즘에 대한 행동과 보상은 다음과 같이 정의된다.

$$A_i = \begin{cases} a_1, & \text{if } \text{mod}(t, U) = 1 \\ a_2, & \text{if } \text{mod}(t, U) = 2 \\ \vdots & \vdots \\ a_U, & \text{if } \text{mod}(t, U) = 0 \end{cases}, \quad (4)$$

$$A_i = [a_1, a_2, \dots, a_U], \quad (5)$$

$$A_u \in \{\pm \Delta x_u, \pm \Delta y_u, \pm \Delta h_u, \pm \Delta p_{tu}\} \quad (6)$$

$$R = \left( \sum_{u=1}^U \chi_u \right) \times e^{-\frac{O_G}{\psi}}. \quad (7)$$

수식 (4), (5) 그리고 (6)은 각각 RR-DQN, SC-DQN 그리고 분산 큐러닝의 행동이다. 또한,  $a_u \in \{\pm \Delta x, \pm \Delta y, \pm \Delta h, \pm \Delta p_t, 0\}$ 이다. 세 알고리즘의 보상은 수식 (7)로 동일하며, 수식 (7)에서  $\chi_u$ 와  $O_G$ 는 각각 UAV-BS  $u$ 의 에너지 효율과 아웃티지 사용자 수를 나타내며,  $\psi$ 는 패널티 인자를 나타낸다.

### III. 시뮬레이션 결과 및 분석

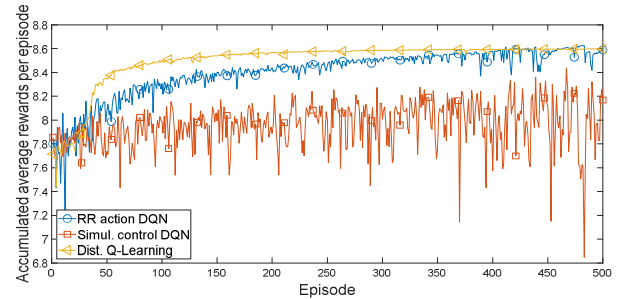


그림 1. 에피소드 진행에 따른 누적 평균 보상 시뮬레이션 결과

그림 1은 2-에이전트 시나리오이며, 네트워크 범위는  $R_{area} = 200 \times 200$ 로 설정되었고, 사용자는  $R_{area}$  내에 무작위로 분포하였다. 또한,  $\psi = 8$ ,  $\{|x_u|, |y_u|\} = 21, |y_u| = 7, |p_{tu}| = 5$ 이고 고도는  $120m \sim 150m$ 로 고려하며, 전송전력은  $0.1W \sim 0.5W$ 로 설정하였다. 네트워크 환경이 복잡하지 않은 관계로 분산 큐러닝이 RR-DQN에 비해 빠른 수렴성을 보이고, SC-DQN은 한 에이전트의 다음 상태가 다른 에이전트 행동에 의해 상당한 영향을 받기 때문에 발산하는 것을 확인할 수 있다.

### ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1F1A1063606).

### 참고문헌

- [1] H. Yu, H. Lee, and H. Jeon, "What is 5G? emerging 5G mobile services and network requirements," MDPI Sustainability, vol. 9, pp. 1-22, Oct. 2017.
- [2] 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, 2021.
- [3] S. Lee and H. Lee, "Reinforcement learning based optimal altitude control for throughput improvement in air-to-ground communications," KICS Winter Conference, pp. 870-870, Feb. 2020.