

# 공정성 향상을 위한 심층 강화학습 기반 전이중 UAV 기지국 위치 최적화 알고리즘

김태운, 김재현\*

아주대학교 AI융합네트워크학과, \*아주대학교 전자공학과

{xodbsxogjs, \*jkim}@ajou.ac.kr

## Deep Reinforcement Learning based Full-Duplex UAV Base Station Location Optimization to Improve Fairness

Tae-Yoon Kim, Jae-Hyun Kim\*

Department of Artificial Intelligence Convergence Network, Ajou University,

\*Department of Electronics and Computer Engineering, Ajou University

### 요약

본 논문에서는 공정성 있는 통신을 위한 심층 강화학습 기반 전이중 UAV 위치 최적화를 수행한다. 랜덤한 위치에 UAV를 배치시키고 전이중 쌍을 결정하는 SINR의 임계값을 변경해가며 시뮬레이션을 진행하였다. 시뮬레이션 결과 임계값에 따라 최적 UAV의 위치가 달라지는 것을 확인하였다. 또한, 제안하는 알고리즘과 UAV를 중앙에 배치한 경우를 비교해본 결과, 제안하는 알고리즘이 모든 임계값의 경우에서 더 좋은 성능을 보이는 것을 확인하였다.

### I. 서론

전이중 통신이 가능한 unmanned aerial vehicle (UAV) 기지국은 기존 지상국의 위치적 한계점을 극복하고 효율적인 주파수 사용이 가능하다는 점에 있어 많은 연구가 진행되고 있다[1][2]. 전이중 통신을 위해서는 두 개의 노드가 전이중 쌍을 맺어야 하는데, UAV의 위치에 따라 신호의 수신 세기가 달라지기 때문에 UAV의 위치는 전이중 쌍을 결정하는 데 큰 영향을 준다. 또한, 전이중 쌍에 따라 노드가 채널을 사용하는 빈도수가 결정되기 때문에 노드들의 공정한 통신을 위해서 UAV 위치는 전이중 통신에 중요한 요소이다.

따라서 본 논문에서는 공정성 있게 전이중 쌍을 이뤄 통신할 수 있도록 심층 강화학습 기반 전이중 UAV 기지국 위치 최적화 연구를 수행한다. UAV 기지국은 모든 노드의 전이중 쌍을 관리하는 테이블을 만든다. 제안하는 알고리즘은 심층 강화학습을 기반으로 전이중 쌍 테이블에 최대한 많은 노드가 기록될 수 있도록 UAV의 위치를 최적화한다.

### II. 시스템 모델

시스템 모델은 기지국 역할을 하는 UAV 한 대와 N개의 노드로 구성된다. 초기 UAV는 랜덤한 위치에 배치하고, 노드들의 위치는 UAV 통신 환경 안에서 랜덤하게 배치한다. 노드들은 UAV와만 통신하고, UAV만 전이중 통신이 가능한 상황을 가정한다. 본 논문에서 고려하는 네트워크는 CSMA/CA 기반으로, 만일 노드가 채널을 차지하면 UAV는 전이중 통신을 하고, UAV가 채널을 차지하면 반이중 통신을 하도록 한다.

### III. 전이중 쌍 테이블 구성

UAV는 특정 위치에서 특정 노드가 채널을 차지할 때 구성될 수 있는 전이중 쌍을 테이블로 저장한다. 전이중 쌍은 signal to noise ratio (SINR)을 통해 계산한다. 특정 노드가 채널을 차지할 때 uplink (UL) 신호의 SINR과 다른 모든 노드와의 downlink (DL) 신호의 SINR을 UAV가 계산해서 UL와 DL SINR 값이 모두 특정 SINR 한계값을 넘으면 전이중 쌍이 이루어질 수 있다고 판단한다.

UAV의 위치에 따라서 SINR 값이 달라지기 때문에 전이중 쌍도 바뀌게 된다. 만일 전이중 쌍이 특정 노드들로만 구성되면 통신의 공정성을 보장할 수 없다. 따라서, 공정성 있는 통신을 위해 UAV가 최대한 다양하고, 많은 노드가 포함될 수 있는 위치에 배치되어야 한다. 따라서 본 논문에서는 공정성 있는 통신을 위한 심층 강화학습 기반 전이중 UAV 기지국 위치 최적화 알고리즘을 제안한다.

### IV. 심층 강화학습 프레임워크

UAV 전이중 쌍 테이블의 노드 수를 최대화하기 위해서는 UAV의 최적 위치를 결정해야 한다. 이러한 문제는 markov decision process (MDP)로 공식화될 수 있다. MDP는 상태 ( $S$ ), 행동 ( $A$ ), 보상 ( $R$ ), 전이 확률 ( $P$ ), 감가율 ( $\gamma$ ) 5개의 요소로 구성된다. 시간  $t$ 에서의  $S$ 는 다음과 같이 구성된다.

$$S_t = (x_m(t), y_m(t), PT_m(t)), \quad (1)$$

$x_m(t)$ 와  $y_m(t)$ 는 시간  $t$ 에서 UAV의 x, y 좌표를 의미하고  $PT_m(t)$ 은 전이중 쌍에 존재하는 노드의 개수를 의미한다.

시간  $t$ 에서의  $A$ 는 다음과 같이 구성된다.

$$A_t = (x_m(t) + \alpha, x_m(t) - \alpha, y_m(t) + \alpha, y_m(t) - \alpha, (x_m(t), y_m(t))), \quad (1)$$

각각 순서대로 UAV를 오른쪽, 왼쪽, 위, 아래, 제자리로 움직이게 하는 불연속적인 행동을 의미한다. UAV의 이동속도를 고려하여  $\alpha$ 는 20으로 정의하였다.

시간  $t$ 에서의  $R$ 는 다음과 같이 구성된다.

$$R_t = PT_m(t) + F_m(t) - \xi, \quad (1)$$

$F_m(t)$ 는 시간  $t$ 에서 UAV 전이중 쌍 테이블에서 서로 겹치지 않게 선택될 수 있는 노드 수를 의미하고  $\xi$ 는 만일 UAV가 특정 영역을 벗어날 경우의 처벌 함수를 의미한다.

**Algorithm 1** 제안하는 DQN 기반 전이중 UAV 기지국 위치 최적화

```

1: 초기값: 행동-가치 함수  $Q$ 의 랜덤 가중치  $\theta$ , 목표 행동-가치 함수  $Q$ 의 가중치  $\theta^- = \theta$ , 리플레이 버퍼  $B$ 의 용량  $C$ , minibatch 크기  $M$ , 총 에피소드 수  $E$ ,
2: 입력값: 상태  $S$ , 행동  $A$ , 감가율  $\gamma$ 
3: 출력값: UAV 최적 위치  $(x, y)$ 
4: for  $Episode = 1$  to  $E$  do
5:   UAV 위치, 노드 위치 초기화
6:   for  $timestep : t = 1$  to  $T$  do
7:     Decaying  $\epsilon$ -greedy 방법을 통한 행동  $a_t$  선택
8:      $= \begin{cases} \operatorname{argmax}_a Q(s_t, a; \theta) & , \text{ 확률 } 1 - \epsilon \\ \text{임의의 행동} & , \text{ 확률 } \epsilon \end{cases}$ 
9:     샘플  $(s_t, a_t, r_t, s_{t+1})$ 를  $B$ 에 저장
10:    if  $B > C$  then
11:       $B$ 로 부터 랜덤한 minibatch  $(s_j, a_j, r_j, s_{j+1})$ 
12:      샘플 선택
13:      for  $j = 1$  in  $M$  do
14:         $y_j = r_j + \gamma \max_{a'} \hat{Q}(s_{j+1}, a'; \theta^-)$ 
15:      end for
16:      가중치  $\theta$ 의 관점에서 경사 하강 단계  $(y_j - Q(s_j, a_j; \theta))^2$  수행
17:      매 스텝마다  $Q$ 의 파라미터를  $\hat{Q}$ 로 복사
18:    end if
19:  end for
20: end for
21: end for

```

$P$ 는 선택한  $A$ 가 그대로 수행될 수 있도록 설정하였으며  $\gamma$ 는 0에서 1 사이의 값으로 미래에 받은 보상을 현재 시점에서 고려할 때 감가하는 비율을 의미한다.

본 논문에서는 알고리즘 1과 같이 deep Q-Network (DQN) 모델 [2]을 기반으로 한 심층 강화학습 알고리즘을 사용하여 앞서 정의한 전이중 UAV 기지국 위치 최적화 문제를 해결하였다. 기존 알고리즘에서 사용한 convolutional neural network 기반의 학습 방식을 3개의 은닉층과 rectified linear unit 활성화 함수로 구성된 deep neural network로 수정하였다.

#### IV. 성능 평가

##### 4.1 시뮬레이션 환경

본 논문에서는 PYTHON을 활용하여 시뮬레이션을 수행하였다. Learning rate는 0.0001, 에피소드는 500, replay buffer의 크기  $B$ 는 60,000, buffer capacity  $C$ 는 10,000, mini batch 크기는 32, time step은 총 200초,  $\xi$ 는 -10으로 설정하였다. UAV는 1,000 m  $\times$  1,000 m 맵 안에서 고도 200 m로 랜덤하게 배치하고, 노드는 동일한 맵 안에서 40개를 랜덤하게 배치하였다. 성능 분석을 위해 SINR 임계값을 10, 13, 18, 26 dB로 변경하면서 시뮬레이션을 진행하였다.

##### 4.2 시뮬레이션 결과

그림 1은 SINR 임계값에 따른 랜덤한 UAV의 시작 지점으로부터 제안하는 알고리즘을 기반으로 학습을 통해 도출한 최적 지점까지의 UAV 위치 변화를 나타낸다. 10, 13, 18 dB의 경우 최적 지점이 각각 (37, 203), (-2, 163), (37, 163)으로 비슷한 지역에 도출되었지만, 26 dB의 경우는 (-302, 483)으로 차이가 큰 지점이 최적 지점으로 도출된 것을 확인할 수 있었다. 이는 다른 조건들에 비해 26 dB가 전이중 쌍 조건을 만족하는 노드가 거의 없기 때문이라고 판단되었다. 그림 2는 제안하는 알고리즘과 UAV를 맵의 중앙지점 (원점)에 두었을 때, 만족하는 전이중 쌍 노드 개수를 학습이 진행됨에 따라 나타내는 그래프이다. 두 경우 모두 SINR의 임계값이 증가함에 따라 전이중 쌍 노드 개수가 감소하는 경향을 보였다. 하지만, 제안하는 알고리즘은 학습이 진행되면서 모든 경우에서 더 많은 전이중 쌍을 가지게 되고 SINR 임계값이 증가함에 따라 두 경우에서 전이중 쌍 노드 수의 차이 값은 48, 102, 62, 26개로 수렴하는 것을 확인하였다.

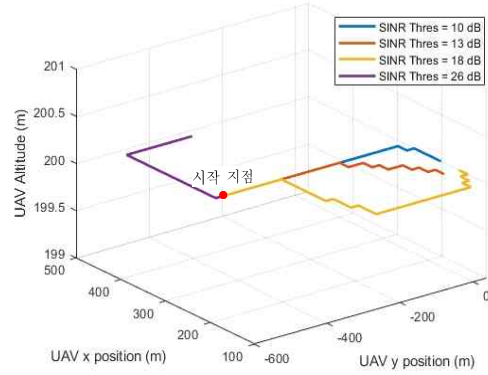


그림 1. 제안하는 알고리즘 기반 UAV 위치 변화

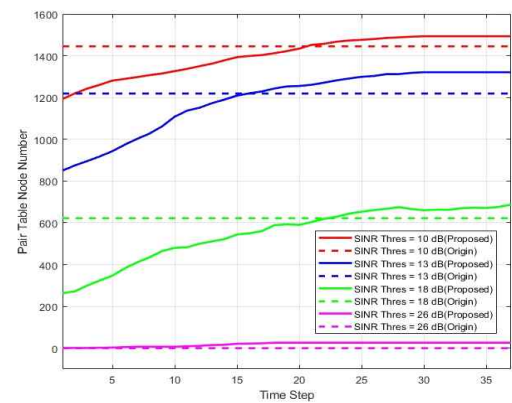


그림 2. 학습 진행에 따른 전이중 쌍 노드 수 비교

#### V. 결론

본 논문에서는 공정성 있는 통신을 위한 심층 강화학습 기반 전이중 UAV 위치 최적화 알고리즘을 제안하였다. 시뮬레이션 결과 제안하는 알고리즘이 SINR 임계값에 따라 최적 위치를 도출하는 것을 확인할 수 있었고, UAV를 맵의 중앙에 위치시킬 때보다 더 좋은 성능을 보이는 것을 확인할 수 있었다.

#### ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2021R1A4A1030775).

#### 참고 문헌

- [1] C. Dai, K. Zhu and E. Hossain, "Multi-agent deep reinforcement learning for joint decoupled user association and trajectory design in full-duplex multi-UAV Networks," IEEE Transactions on Mobile Computing, pp. 1-15, Jul. 2022, doi: 10.1109/TMC.2022.3188473.
- [2] K. Tian, B. Duo, S. Li, Y. Zuo and X. Yuan, "Hybrid uplink and downlink transmissions for full-duplex UAV communication With RIS," IEEE Wireless Communications Letters, vol. 11, no. 4, pp. 866-870, Apr. 2022, doi: 10.1109/LWC.2022.3149096.
- [3] V. Mnih et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, pp. 529-533, Feb. 2015, doi: https://doi.org/10.1038/nature14236