

# Tethered UAV 지원 통합 액세스 백홀 네트워크에서 저복잡도 심층강화학습 기반 합 전송률 최대화 기법

이예린<sup>1</sup>, 유희정<sup>2</sup>, 이호원<sup>1</sup>

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

{qkek0379, hwlee}@hknu.ac.kr, heejungyu@korea.ac.kr

## Low-Complexity Deep Reinforcement Learning-Based Sum-rate Maximization in Tethered UAV-Aided Integrated Access and Backhaul Network

Yerin Lee<sup>1</sup>, Heejung Yu<sup>2</sup>, Howon Lee<sup>1</sup>

Hankyong National Univ.<sup>1</sup> Korea Univ.<sup>2</sup>

### 요약

통합 액세스 백홀(Integrated Access and Backhaul, IAB) 네트워크 기술은 차세대 이동통신 네트워크의 유연하고 비용 절감적인 운용을 가능하게 한다. 그러나 링크 간의 심각한 상호 간섭과 사용자 트래픽 시변성에 의해 네트워크 무선 자원 관리에서 발생하는 많은 문제들로 인하여 지상 및 공중 네트워크 간 통합 자원관리 기술의 적용은 필수적이다. 이에 본 논문에서는 지상으로부터 전력을 끊임없이 공급받는 tethered unmanned aerial vehicle-base station(TUBS)을 IAB 노드로 활용하는 3차원 공간 네트워크를 고려한다. 특히, 본 논문에서는 심층강화학습을 기반으로 IAB 네트워크에서의 합 전송률 최대화를 위해 액세스 및 백홀의 채널할당, 전력 조절, TUBS 위치 제어를 공동으로 최적화한다. 동적으로 변화하는 IAB의 네트워크에서 기존의 최적화 기법들을 활용하여 문제를 해결하는 것은 알고리즘 복잡도 및 네트워크의 가변성 등으로 인해 현실적인 많은 어려움이 존재하기 때문에 어렵다. 이에 본 논문에서는 동적 환경에 대응하며 계산 복잡도를 줄이기 위해 분산형 DDQN 알고리즘을 제안한다. 시뮬레이션을 통해 제안 방안이 비교 방안들에 비해 합 전송률 측면에서 그 성능이 우수함을 보인다.

### I. 서론

모바일 데이터 트래픽의 폭발적인 증가는 6G 네트워크에서 기존보다 더 높은 데이터 전송률을 필요로 하고 있으며, 이에 따라 밀리미터파 대역과 테라헤르츠 대역의 사용이 적극적으로 고려되고 있다. 높은 주파수 대역은 데이터 전송률에 대한 요구 사항을 충족시키지만 높은 직진성, 신호 감쇠, 전파 반사의 고유의 특성으로 좁은 서비스 커버리지를 가진다. 이는 고밀도 기지국 배치를 통해서 일부 해결이 가능하지만 높은 자본 비용 및 운용 비용 문제를 야기한다. 이를 위해, the 3rd generation partnership project(3GPP)는 비용을 절감하면서 네트워크의 유연한 배치가 가능한 통합 액세스 및 백홀 네트워크(Integrated Access and Backhaul, IAB) 아키텍처를 제안하였다 [1]. IAB는 기존 액세스 링크를 위해 사용되던 무선 자원들을 백홀 링크와 동시에 사용하는 아키텍처로 IAB 공여자 기지국과 IAB 노드 기지국으로 구성된다. 이러한 IAB 네트워크는 커버리지 향상에 용이하며 주파수 효율을 높이면서 운용비용을 절감할 수 있다. 그러나 백홀 링크와 액세스 링크 사이의 co-tier interference, 고정형 기지국 배치문제, 사용자 트래픽 시변성으로 인해 발생하는 다양한 네트워크 운용 문제들을 해결하기 위해서 unmanned aerial vehicle(UAV)를 고려한 지상 및 공중 기지국 협력 네트워크의 활용이 매우 중요하다 [2]. 하지만 UBS는 제한된 배터리 용량으로 장시간 비행이 불가능하여 지속가능한 통신 서비스를 제공하는데 한계가 있다. 이에 따라 지상으로부터 연결된 tether를 통해 지속적으로 에너지를 공급받는 tethered-UAV(TUAV)가 제안되었다 [3]. 이에 본 논문에서는 TUBS를 IAB의 노드 기지국으로 활용하고 IAB 네트워크의 합 전송률을 최대화하기 위하여 심층강화학습 기반의 액세스 및 백홀 네트워크 채널 할당, 전송 전력 조절, TUBS 위치 제어 공동 최적화 기법을 제안한다.

### II. 채널모델

본 논문에서는 MBS, TUBS의 두 가지 유형의 기지국이 존재하는 멀티 셀 IAB 네트워크를 고려한다. 여기서 MBS는 IAB의 공여자 기지국이며 TUBS는 릴레이 역할을 수행하는 IAB 노드 기지국으로 고려된다.  $\mathbf{M}_M$ ,  $\mathbf{M}_T$ ,  $\mathbf{U}_{tot}$  ( $= (\bigcup_i \mathbf{U}_i^M) \cup (\bigcup_j \mathbf{U}_j^T)$ )는 각각 MBSs, TUBSs, 총 지상

사용자의 집합을 나타낸다. 여기서  $\mathbf{U}_l^M = \{u_l^M | l = 1, 2, \dots, |\mathbf{M}_M|\}$ 는 MBS  $l$ 에서 서비스를 받는 지상 사용자 집합,  $\mathbf{U}_l^T = \{u_l^T | l = 1, 2, \dots, |\mathbf{M}_T|\}$ 는 TUBS  $l$ 로부터 서비스를 받는 지상 사용자 집합을 나타낸다.  $\mathbf{M}_{tot} = \mathbf{M}_M \cup \mathbf{M}_T$ 는 모든 기지국을 포함하는 집합을 나타낸다. IAB 네트워크는 2개의 계층으로 구성된다. 1계층 링크는 MBS와 관련 지상 사용자 간의 액세스 링크 및 MBS와 TUBS 간의 백홀 링크가 포함된다. 2계층 링크는 TUBS와 관련 지상 사용자 간의 액세스 링크가 포함된다.

네트워크의 총 대역폭은  $K$ 개의 직교 하위채널로 나뉘며 모든 하위채널의 집합을  $\mathbf{K}_{tot} = \{1, 2, \dots, K\}$ 로 나타낸다. 액세스 링크와 백홀 링크는 하위채널은 함께 공유한다고 가정한다. MBS에서 1계층  $l$ 번째 수신자(MBS에 속하는 사용자 집합과 TUBSs가 포함된다.)에게 할당되는 채널 할당 벡터를  $\mathbf{c}_l = [c_l^1, \dots, c_l^k, \dots, c_l^K]$ ,  $c_l^k \in \{0, 1\}$ ,  $k \in \mathbf{K}$  같이 나타낸다. TUBS에서 2계층의  $l$ 번째 수신자(TUBS에게 서비스를 받는 사용자)에게 할당되는 채널할당 벡터를  $\bar{\mathbf{c}}_j = [\bar{c}_j^1, \dots, \bar{c}_j^k, \dots, \bar{c}_j^K]$ ,  $\bar{c}_j^k \in \{0, 1\}$ ,  $k \in \mathbf{K}$ 로 나타낸다.

G2G 채널 모델에서  $k$ 번째 하위채널을 사용하는 송신자  $i$ 와 수신자  $j$  사이의 채널 이득은 다음과 같이 계산한다.

$$\mathbf{G}_{i,j}^k = \mathbf{L}_{i,j} \mathbf{h}_{i,j}^k, \quad (1)$$

여기서,  $\mathbf{L}_{i,j}$ 는 경로 손실 및 음영 효과를 포함하는 거리 기반의 대규모 페이딩 계수,  $\mathbf{h}_{i,j}^k$ 는 레일리 페이딩을 겪는 소규모 페이딩을 나타낸다. A2G의 전파 특성을 반영하기 위해 고도각을 고려한 경로 손실 모델을 이용한다. 송신기  $i$ 와 수신기  $j$ 사이의 line-of-sight(LoS) 확률 및 non-LoS(NLoS) 확률은 다음과 같이 계산한다 [4].

$$P_{i,j}^{LoS} = a \left( \text{actan} \left( \frac{z_i}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}} \right) - 15 \right)^b \quad (2)$$

$$P_{i,j}^{NLoS} = 1 - P_{i,j}^{LoS} \quad (3)$$

여기서,  $a$ 와  $b$ 는 도심 시나리오에 따른 환경 파라미터이며 위 수식을 이용한 채널 링크 상태에 따른 평균 경로 손실은 다음과 같이 계산한다.

$$\eta_{i,j}^{avg} = P_{i,j}^{LoS} \times \eta_{i,j}^{LoS} + P_{i,j}^{NLoS} \times \eta_{i,j}^{NLoS} \quad (4)$$

여기서,  $R_{i,j}$ 는 송신기  $i$ 와 수신기  $j$ 사이의 거리,  $\eta_{i,j}^{LoS} = 20\log_{10}R_{i,j} + \xi^{LoS}$ ,  $\eta_{i,j}^{NLoS} = 20\log_{10}R_{i,j} + \xi^{NLoS}$ 로 계산되며  $\xi^{LoS}$ 와  $\xi^{NLoS}$ 는 각각 초과 경로 손실을 의미한다. 식(4)로부터 송신기  $i$ 에서 수신기  $j$ 로의 수신 신호 세기는 다음과 같이 계산한다.

$$P_{i,j}^r = P_i^{TX} - \eta_{i,j}^{avg} \quad (5)$$

여기서,  $P_i^{TX}$ 는 송신기  $i$ 의 전송 전력을 나타낸다. 채널 모델을 사용하여  $k$ 번째 채널을 사용하여 송신기  $i$ 가 수신기  $j$ 사이의 간섭은 동일한 채널에서의 다운 링크 전송으로 고려된다. 1계층에서의 송신자 MBS에서 관련 사용자 사이의 신호대잡음비(SINR)는 다음과 같이 계산된다.

$$\gamma_{\mathbf{M}_M, u_l^M}^k = \frac{P_{\mathbf{M}_M}^{TX} G_{\mathbf{M}_M, u_l^M}^k c_l^k}{\sum_{i \in \mathbf{M}_T} (P_{ij}^r c_j^k) + \sigma^2} \quad (6)$$

여기서,  $\sigma^2$ 는 노이즈를 나타낸다. 송신자 MBS에서 수신자 TUBS사이의 SINR은 다음과 같이 계산된다.

$$\gamma_{\mathbf{M}_M, \mathbf{M}_T}^k = \frac{P_{\mathbf{M}_M}^{TX} G_{\mathbf{M}_M, \mathbf{M}_T}^k c_l^k}{\sum_{i \in \mathbf{M}_T} (P_{ij}^r c_j^k) + \sigma^2} \quad (7)$$

2계층의 송신자 TUBS에서 관련 사용자 사이의 SINR은 다음과 같이 계산된다.

$$\gamma_{\mathbf{M}_T, u_l^T}^k = \frac{P_{\mathbf{M}_T}^r c_l^k}{\sum_{i \in \mathbf{M}_T} (P_{ij}^r c_j^k) + \sum_{i \in \mathbf{M}_M} (P_{ij}^{TX} c_j^k) + \sigma^2} \quad (8)$$

이에 따른 합 전송률은 다음과 같이 계산된다.

$$\Gamma_{i,j}^k = \frac{\mathbf{K}_{tot}}{K} \times \log_2(1 + \gamma_{ij}) \quad (9)$$

### III. 분산 DDQN 기반 최적 TUBS 제어 기법

강화학습은 환경에 대한 사전정보 없이도 스스로 학습하여 정답을 찾아내는 기법으로 적응성이 뛰어나며 동적인 환경에서 사용하기에 적합하다. 딥러닝 기반의 DDQN은 과대평가로 인하여 주어진 상태에서 최적이지 않은 동작을 선택하는 큐러닝과 DQN의 문제를 해결하기 위한 기법으로 제안되었다. DDQN은 각기 다른 매개변수를 갖는 온라인 네트워크와 타깃 네트워크의 구조로 이루어져 있으며, 과대평가 문제를 예방하기 위해 행동을 선택할 때는 온라인 네트워크의 매개 변수를 사용하고 행동을 평가할 때는 타깃 네트워크의 파라미터를 사용한다. 본 시나리오에서는 MBS와 TUBS가 에이전트가 되며 상태, 행동, 보상은 다음과 같이 정의된다.

#### ● 에이전트 MBS 모델링

■ 상태( $S_t$ ): MBS  $i$ 의 상태는 전송 전력( $P_{\mathbf{M}_M}^{TX}$ ), 채널 할당 정보를 포함하여 다음과 같이 정의된다.

$$S_t^i = \{P_{\mathbf{M}_M}^{TX}, c_l^1, \dots, c_l^k, \dots, c_l^K\}, \forall P_{\mathbf{M}_M}^{TX} \in \{P_{\mathbf{M}_M}^{\min}, \dots, P_{\mathbf{M}_M}^{\max}\} \quad (10)$$

■ 행동( $A_t$ ): MBS  $i$ 의 행동은 전송 전력 제어와 채널 할당을 포함하여 다음과 같이 정의된다.

$$A_t^i = \{\pm \Delta P_{\mathbf{M}_M}^{TX}, c_l^1, \dots, c_l^k, \dots, c_l^K\} \quad (11)$$

여기서,  $\pm \Delta P_{\mathbf{M}_M}^{TX}$ 는 MBS 전송 전력의 증가/감소/유지를 의미한다.

#### ● 에이전트 TUBS 모델링

■ 상태( $S_t$ ): TUBS  $i$ 의 상태는 전송 전력, 좌표표제로 표현되는 TUBS의 3차원 위치, 채널 할당 정보를 포함하여 다음과 같이 정의된다.

$$S_t^i = \{P_{\mathbf{M}_T}^{TX}, r_t, \theta_t, \phi_t, c_l^1, \dots, c_l^k, \dots, c_l^K\}, \forall P_{\mathbf{M}_T}^{TX} \in \{P_{\mathbf{M}_T}^{\min}, \dots, P_{\mathbf{M}_T}^{\max}\}, \forall r_t \in \{r^{\min}, \dots, r^{\max}\}, \forall \theta_t \in \{\theta^{\min}, \dots, \theta^{\max}\}, \forall \phi_t \in \{\phi^{\min}, \dots, \phi^{\max}\} \quad (12)$$

■ 행동( $A_t$ ): TUBS  $i$ 의 전송 전력 제어와 이동성 제어, 채널 할당을 포함하여 다음과 같이 정의된다.

$$A_t^i = \{\pm \Delta P_{\mathbf{M}_T}^{TX}, \pm \Delta r, \pm \Delta \theta, \pm \Delta \phi, 0, \bar{c}_l^1, \dots, \bar{c}_l^k, \dots, \bar{c}_l^K\} \quad (13)$$

여기서,  $\pm \Delta r, \pm \Delta \theta, \pm \Delta \phi, 0$ 는 TUBS의 반경, 고도각, 방위각의 증가/감소/유지를 의미한다.

■ 보상( $R_t$ ): 식(9)의 합 전송률을 보상으로 고려하며 네트워크 전체의 합 전송률 최대화를 위해 에이전트는 다음과 같은 동일한 수식을 가진다.

$$R_t = \sum \Gamma_{i,j}^k \quad (14)$$

### IV. 시뮬레이션 결과 분석 및 결론

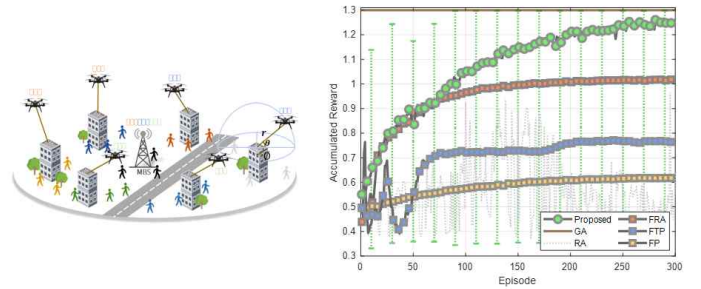


Fig. 1. (left) Network, (right) Accumulated Reward vs. Episode

시뮬레이션은 단일 MBS와 다중 TUBS의 환경에서 진행하였다. 제안 방안의 성능 검증을 위해 ‘fixed-resource allocation(FRA)’, ‘fixed-transmit power control (FTP)’, ‘fixed-positioning(FP)’, ‘random-action(RA)’, ‘greedy-action(GA)’과 비교한다. 시뮬레이션 결과를 통해 ‘Proposed’가 ‘FRA’, ‘FTP’, ‘FP’ 대비 각각 23.49%, 63.14%, 103.04%의 성능 이득을 가지는 것을 확인할 수 있다. 또한 에피소드가 진행됨에 따라 ‘Proposed’가 ‘GA’에 근접하게 수렴해 가는 것을 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%).

### 참고 문헌

- [1] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, “Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond,” *IEEE Access*, vol. 7, pp. 78729 - 78757, May. 2019.
- [2] S. Lee and H. Lee, “UAV-BS Energy Efficiency Maximization Based on Multi-Agent Distributed Q-Learning,” in *Proc. KICS Winter Conf.*, Pyeongchang, Korea, Feb. 2022.
- [3] Z. Xie, X. Song, J. Cao and W. Qiu, “Providing Aerial MEC Service in Areas Without Infrastructure: A Tethered-UAV-Based Energy-Efficient Task Scheduling Framework,” in *IEEE Internet of Things Journal*, vol. 9, no. 24, pp. 25223-25236, Dec. 2022.
- [4] A. Al-Hourani, S. Kandeepan, and S. Lardner, “Optimal LAP Altitude for Maximum Coverage,” *IEEE Wireless Communications Letters*, vol. 3, pp. 569-572, Dec. 2014.