

강화학습을 활용한 전이중 및 반이중 IAB 네트워크에서의 자원 할당

김조은, 전영일*, 이문식*, 권태수

서울과학기술대학교, *한국전자통신연구원

worth3910@naver.com, *youngil@etri.re.kr, *moonsiklee@etri.re.kr, tskwon@seoultech.ac.kr

Reinforcement Learning Based Resource Allocation for Full Duplex and Half Duplex IAB networks

Joeun Kim, Youngil Jeon*, Moon-Sik Lee*, Taesoo Kwon

Seoul National Univ. of Science & Technology, *ETRI

요약

IT 관련 산업들이 고도화되면서 요구되는 데이터 속도가 매우 높아졌다. 이를 위해 네트워크 고밀화를 할 때 5G에서는 백홀을 무선으로 연결하여 추가적인 기지국 설치를 쉽게 할 수 있게 하는 통합 액세스 및 백홀(Integrated Access and Backhaul, IAB)을 고려하고 있다. 여기서 IAB 노드는 백홀과 액세스 둘 다 무선으로 연결되어 한 번에 관리하기에 자원 할당 전략이 매우 중요하지만 이것을 푸는 문제는 혼합 정수 비선형 문제로 이는 복잡도가 크다. 따라서 본 논문에서는 IAB 네트워크에서의 자원 할당 문제를 효과적으로 해결하기 위해 강화학습을 활용하고 전이중 통신을 기반으로 한 강화학습을 반이중 통신으로 동작하도록 일반화하는 방안을 제안한다.

I. 서론

IT 산업이 발전함에 따라 한 번에 처리해야 하는 데이터의 양이 많아졌고 5G 표준에서 요구하는 데이터 속도 또한 높아졌다. 이를 만족하기 위해 기지국 고밀화가 한 방법이 될 수 있는데, 수많은 기지국을 광섬유로 각각의 백홀을 연결하여 설치하는 것은 비용이 너무 높다. 이를 해결하기 위해 5G에서는 통합 액세스 및 백홀(Integrated Access and Backhaul, IAB) 네트워크를 지원하고 있다. IAB 네트워크에서는 IAB 노드라 불리는 기지국들이 광섬유 백홀이 연결된 매크로 기지국으로부터 백홀을 무선으로 제공받기 때문에 추가적인 기지국 설치가 유용하여 커버리지를 증가시킬 수 있다. 따라서 IAB 노드는 액세스와 백홀 연결을 한 번에 관리하게 되므로 자원을 얼마나 효율적으로 할당할지에 대한 문제는 네트워크 성능에 큰 영향을 끼치고 이것의 최적인 해를 얻는 것은 중요하다.

[1]~[2]는 IAB 네트워크에서 확률 기하를 기반으로 수학적 및 시스템 레벨 시뮬레이션을 통해 성능을 분석하였다. 이때, 자원 할당 전략이 고정되어 네트워크 상황에 맞추어 동적으로 최적인 자원 할당 전략을 사용하지 않는다는 한계를 가진다. 하지만 자원 할당 문제는 혼합 정수 비선형 문제라서 최적인 해를 얻는 것은 복잡도가 크다. 따라서 [3]~[4]에서는 효율적으로 최적인 해를 얻기 위해 강화 학습을 활용하는 방안을 제안했다. 하지만 [3]에서는 사용자가 움직이지 않는 네트워크 환경으로 제한하였고 [4]에서는 전이중 통신만을 고려하여 성능을 분석하였다. 이에 따라, 본 논문에서는 [4]와 같이 전이중(Full Duplex, FD) IAB 네트워크에서 강화학습을 활용해 자원을 할당하고 이를 반이중(Half Duplex, HD) IAB 네트워크에도 적용할 수 있는 알고리즘을 제안하고자 한다.

II. 시스템 모델 및 강화학습

IAB 노드들에게 무선으로 백홀을 제공하는 매크로 기지국 b_0 는 원점에 위치한다. IAB 노드들과 단말들은 균일 포아송 점 과정(Poisson Point Process, PPP)에 따라 분포하고 단말은 랜덤 워크 모델에 따라 움직인다. 각 단말들은 수신 신호가 가장 높은 기지국과 연결되고 기지국은 이 중

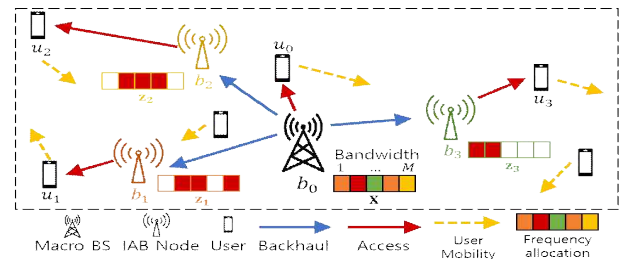


그림 1. IAB 네트워크에서의 자원 할당

하나의 단말에게 서비스한다. IAB 노드의 개수가 N 이라면, IAB 노드는 $n = 1, \dots, N$ 일 때 b_n , 각 기지국에 연결된 단말을 u_n 이라고 정의한다. 각 기지국은 사용 가능한 전체 대역폭을 M 개의 직교하는 부채널로 나누고 이 중 1개 이상을 선택해 단말에게 서비스한다. 매크로 기지국의 자원 할당 벡터는 $i \in T_1 = \{u_0, b_1, \dots, b_N\}$ 일 때 $\mathbf{x}_i = [x_i^1, \dots, x_i^M]^T$ 이고, n 번째 IAB 노드의 자원 할당 벡터는 $j \in T_2 = \{u_1, \dots, u_N\}$ 일 때 $\mathbf{z}_j = [z_j^1, \dots, z_j^M]^T$ 이다. 여기서 $x_i^m, z_j^m \in \{0, 1\}$ 이며 0은 해당 부채널을 할당하지 않음, 1은 할당함을 의미한다. 따라서 자원 할당 행렬은 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_{1+N}]^T$ 와 $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_N]^T$ 로 나타낸다. IAB 노드는 전이중 통신을 하고 자기간섭(Residual Self-interference, RSI) 제거 이득 ξ 를 가진다. 채널 이득 g 는 레일레이 페이딩 이득, 경로 감쇄 이득, 그리고 웨도잉 효과를 포함한다. 또한 송신전력은 P , 해당 자원 할당 벡터의 요소는 a , 다른 기지국으로부터의 간섭은 I , 간섭 잡음은 σ^2 이라고 할 때, SINR은 $\Gamma = \frac{Pga}{I + \sigma^2}$, 데이터 속도 $C(\mathbf{X}, \mathbf{Z})$ 는 $\log_2(1 + \Gamma(\mathbf{X}, \mathbf{Z}))$ 로 나

타내고 단말 속도는 서빙 기지국의 백홀과 액세스의 속도 중 작은 값이다. 이를 기반으로 단말 속도에 대한 효용함수를 최대화하는 주파수 자원 할당 문제는 그림 1에 도식화되었고 식 (1a)~(1c)와 같이 표현할 수 있다. 식 (1b)의 Ω 는 최소 통신서비스 품질(Quality of Service, QoS)을, (1c)는 자원 할당 행렬의 제약 조건을 각각 의미한다. 효용함수 $f(x)$ 는 모든 단말

$$\max_{\mathbf{x}, \mathbf{z}} \sum_{n=0}^N f(C_{u_n}(\mathbf{x}, \mathbf{z})), \quad (1a)$$

$$s.t. C_{u_n} \geq \Omega \quad (1b)$$

$$\sum_{i \in T_1} x_i^m = 1, \sum_{i \in T_1, m=1}^M x_i^m \leq M, \sum_{m=1}^M z_j^m \leq M \quad (1c)$$

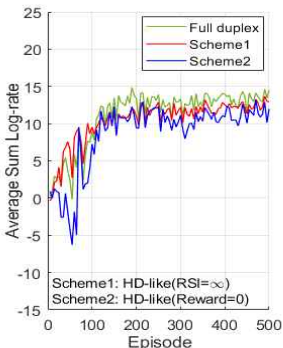
속도가 일정 수준 이상이 되는 동시에 합이 최대가 되도록 $\log(x)$ 의 형태로 설계한다. 위와 같은 문제는 혼합 정수 비선형 문제로 최적인 해를 얻는 것이 매우 복잡하고 어렵다. 따라서 위 문제를 효과적으로 해결하기 위해서 강화학습을 활용하고자 한다.

강화학습은 한 상황의 상태에 따라서 어떤 행동을 선택하는 것이 최대 보상을 얻을지에 대한 정책을 학습하는 것이다. 상태에 따라 선택하는 행동의 경우의 수를 표로 만들어서 각각의 Q-value를 업데이트해나가는 방법을 Q-Learning이라고 한다. 그리고 이러한 표를 대신하여 심층 신경망(Deep Neural Network, DNN)으로 Q-value를 계산하는 것을 DQN(Deep Q-Network)이라고 한다. 더 나아가 DQN에서 Q-value를 과대평가하는 것을 줄이기 위해 두 개의 DNN을 가지고 하나는 행동을 선택하는데 하나는 평가하는데 활용하는 것을 DDQN(Double DQN)이라고 한다. 본 논문에서는 이러한 DDQN을 활용해 자원 할당 문제를 해결한다. DDQN에서 시간 t 에서의 상태 \mathbf{s}_t 는 $\mathbf{s}_{t, u_n} \in \{0, 1\}$ 일 때, $\{\mathbf{s}_{t, u_0}, \dots, \mathbf{s}_{t, u_N}\}$ 로 정의하고 0은 각 단말의 속도가 최소 QoS Ω 를 만족하지 않음을, 1은 만족함을 의미한다. 이에 따라 선택하게 되는 행동 \mathbf{a}_t 는 자원 할당 행렬 $[\mathbf{x}_t, \mathbf{z}_t]^T$ 로 정의한다. 또한 보상은 $0 \leq t \leq T$ 에 대해서 $\sum_{n=0}^N f(C_{u_n}(\mathbf{x}, \mathbf{z}))$ 로 정의되며 (1a)의 최적인 해를 구하도록 설계되었다. 이러한 [4]의 방안에 기반한 DDQN 알고리즘을 반이중 통신으로 동작할 수 있도록 설계한 방안을 다음 장에서 살펴보고자 한다.

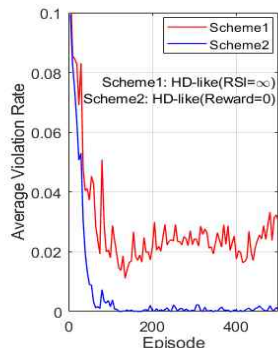
III. 성능 분석

표 1. IAB 네트워크 환경변수

변수	값
P_D, P_S : 송신전력	43 dBm, 33 dBm
f_c, W : 반송주파수, 대역폭	2.4 GHz, 20 MHz
σ^2 : 잡음 전력	$-174 \text{ dBm/Hz} + 10\log_{10} W$
N, M : IAB 노드 개수, 부채널 개수	4, 10
K_D, K_S : 경로감쇄이득	-34 dB, -37 dB
α_D, α_S : 경로감쇄지수	4, 3
ξ : 자기간섭제거 이득	-130 dB
ν, θ : 사용자 움직임 속도, 각도	$\nu \sim U(0, 2) \text{ m/s}, \theta \sim U(0, 2\pi)$



(a) 로그 속도 합의 평균



(b) 평균 위반율

그림 2. IAB 네트워크에서 자원 할당을 위한 DDQN 성능 그래프

이번 장에서는 이전 장에서 설계한 DDQN을 표 1의 환경에서 구현하여 이를 반이중 통신처럼 동작하도록 하는 2가지 방안을 제안하고 성능 분석하고자 한다. 이때 표 1의 아래첨자 D, S 는 각각 매크로 기지국과 IAB 노드의 값을 의미한다. 그림 2의 (a)를 보면 전이중 통신보다 반이중 통신으로 일반화한 것들의 로그 속도 합이 낮다. 이는 반이중 통신이 활용할 수 있는 자원이 더 제한적이기에 발생하는 결과라고 볼 수 있다. 반이중 통신과 유사한 동작으로 제한하는 방안으로는 2가지 방안을 고려했는데, 방법 1은 자기간섭제거 능력을 무한대로 보내 전이중 통신이 일어나면 해당 채널의 데이터 속도는 0에 가까운 값이 나와 보상을 적게 만드는 것이다. 그리고 방법 2는 IAB 노드가 동시에 송수신하면 즉, 전이중 통신이 일어나면 보상을 무조건 0으로 설정하는 것이다. 그림 2의 (b)에서는 두 방법의 위반율을 비교하고 있는데 위반율은 IAB 노드들의 전체 부채널 개수에 대한 전이중 통신이 일어난 채널의 비율을 나타낸다. 이러한 위반율은 방법 1에서 조금 더 높게 나타난다. 즉, 방법 2가 좀 더 반이중 통신에 가깝다는 것을 의미한다. 이는 보상을 얼마나 보정했느냐에 따라서 차이가 발생했다고 할 수 있다. 또한 학습이 완료된 DDQN으로 테스트한 결과 위반율은 방법 1이 0.0023, 방법 2가 0으로 두 방법 모두 거의 반이중 통신으로 동작하고 있음을 알 수 있다. 따라서 이는 반이중 통신 성능의 참조 성능이 될 수 있으므로 반이중 통신을 설계할 때 성능에 대한 가이드 라인을 제공할 수 있다.

IV. 결론

본 논문에서는 전이중 통신을 가정하는 IAB 네트워크에서 자원 할당 문제를 효과적으로 해결하기 위해 DDQN을 활용하고 이를 반이중 통신처럼 동작하도록 일반화하여 성능을 분석했고, 이것이 반이중 통신 알고리즘 성능에 대한 참조가 될 수 있음을 확인하였다. 향후에는 본 논문의 결과를 활용하여 반이중 통신 알고리즘을 설계하고 좀 더 확장하여 밀리미터파에서의 자원 할당 문제를 위한 강화학습의 성능 분석을 하고자 한다.

ACKNOWLEDGMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No. 2019-0-01360, 동적 기능분할을 지원하는 개방형 기지국 분산 유닛(DU) 기술 개발)

참고 문헌

- [1] C. Madapatha, B. Makki, C. Fang, O. Teyeb, E. Dahlman, M. S. Alouini, and T. Svensson, "On Integrated Access and Backhaul Networks: Current Status and Potentials," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1374-1389, Sept. 2020.
- [2] C. Saha and H. S. Dhillon, "Millimeter Wave Integrated Access and Backhaul in 5G: Performance Analysis and Design Insights," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 12, pp. 2669-2684, Dec. 2019.
- [3] Q. Cheng, Z. Wei and J. Yuan, "Deep Reinforcement Learning-based Spectrum Allocation and Power Management for IAB Networks," *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1-6, Montreal, QC, Canada, Jul. 2021.
- [4] W. Lei, Y. Ye and M. Xiao, "Deep Reinforcement Learning-Based Spectrum Allocation in Integrated Access and Backhaul Networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 970-979, Sept. 2020.