

# THz 대역 무선 네트워크에서 강화학습 기반 빔포밍 기술 동향 분석

박준영, 나웅수\*

국립공주대학교

jjunyoung99@gmail.com, \*wsna@kongju.ac.kr

## Trend Analysis of Reinforcement Learning-Based Beamforming Techniques in THz Wireless Networks

Park Jun Young, Na Woong Su\*

Kongju National Univ.

### 요약

본 논문은 초고속·초저지연 전송이 요구되는 VR 및 홀로그램 환경에서, THz 대역 무선 네트워크의 빔포밍 최적화를 위해 강화학습 기반의 연구들을 분석하였음. MAB 기반의 트레이닝 최적화와 FDRL 기반의 분산 학습 기법은 각각 멀티 홉 및 셀룰러 환경에서 스펙트럼 효율과 간섭 완화 성능을 효과적으로 향상시킴을 확인하였음. 이를 통해 강화학습은 THz 환경에서 MAC 프로토콜 수준의 네트워크 성능 향상에 유의미한 기여가 가능함을 보여줌.

### I. 서론

VR 및 홀로그램과 같은 초대용량 콘텐츠의 무선 전송을 위해서는 높은 처리량과 낮은 지연시간을 가진 5G 및 6G 통신이 필수적으로 요구되어지며[1] mmWave와 THz 대역 네트워크에서 빔 관리는 매우 중요한 문제임. 본 논문에서는 THz 대역의 무선 네트워크에서 빔 포밍에 강화학습을 응용 및 사용함으로써 MAC 프로토콜 수준에서 네트워크의 성능을 향상시킨 연구들을 소개하고자 함.

### II. 본론

[2]는 고전적인 강화학습 기법인 Multi-Armed Bandit (MAB)을 활용하여, 고정된 멀티 홉(다중 홉) THz 무선 통신 환경에서의 빔 트레이닝 최적화 문제를 해결하는 새로운 기법을 제안하였음. 제안된 방법은 기존의 계층적 빔 트레이닝 방식에서 모든 홉에 대해 고정된 수의 트레이닝 레벨을 사용하는 구조의 한계점, 즉 과도한 트레이닝 시간 오버헤드를 해결하기 위해, 각 홉의 무선 링크 상태에 따라 동적으로 트레이닝 레벨 수를 선택하는 방식임. 이를 위해 저자들은 전체 멀티 홉 링크의 스펙트럼 효율을 극대화하는 최적화 문제를 수립하고, 채널 상태 정보에 대한 사전 지식 없이도 학습을 수행할 수 있도록, MAB 기반의 강화학습 알고리즘을 설계하였음. 본 알고리즘은 학습과정에서 각 홉별로 최적의 트레이닝 레벨 수를 선택하는 문제를 ‘팔 선택(arm selection)’ 문제로 모델링하여 해결하며, 시간 블록 단위로 반복 학습을 통해 최적의 정책에 수렴함. 시뮬레이션 결과에 따르면, 제안된 방법은 기존의 고정 레벨 방식 및 무작위 레벨 방식과 비교하여, 최대 75%까지의 스펙트럼 효율 향상을 달성하였으며, 평균적인 오차(regret)가 빠르게 수렴함을 통해 학습의 안정성과 효율성도 입증하였음. 또한, 전송 SNR이 낮은 환경에서도 오탐지율의 증가 없이 우수한 성능을 유지하여, 제안된 기법의 실용성을 입증하였음.

[3]는 심층 강화학습(DRL)과 연합 학습을 결합한 연합 심층 결정론적 정책 그라디언트(FDRL) 기법을 제안하였음. 제안된 FDRL은 THz 대역의 고정형 셀룰러 네트워크 환경에서, 각 기지국이 제한된 채널 상태 정보를

바탕으로 독립적으로 DDPG 기반 강화학습을 수행하고, 이를 에지 서버와 공유하여 최적의 빔 포밍 정책을 학습하는 방식으로 구성됨. THz 대역은 mmWave보다 훨씬 넓은 대역폭(0.1~10 THz)을 제공하지만, 심각한 전파 감쇠, 차폐, 분자 흡수 등의 문제로 인해 고정밀 빔포밍이 필수적임. 하지만 현실적인 네트워크 환경에서는 CSI를 완벽하게 수집하는 것이 어렵고, 기지국 간 간섭이 발생하기 때문에 기존의 중앙 집중형 학습 방식이나 완전한 채널 정보 기반 최적화 방식으로는 성능 한계가 존재함. 이 문제를 해결하기 위해 [2]에서는 각 기지국이 로컬에서 DDPG 기반 빔포밍 정책을 학습하고, 그 결과로 얻은 DNN 파라미터(모델 가중치)를 에지 서버에 주기적으로 공유하는 연합 학습 구조를 도입하였음. 에지 서버는 각 기지국에서 전송된 학습 모델을 통합하여 간섭 정보를 포함하는 전역 표현을 학습하고, 다시 각 기지국에 배포함으로써 전체 네트워크의 빔포밍 성능을 점진적으로 향상시킴. 또한, 부분 모델 업데이트를 통해 통신 오버헤드를 줄이면서도 전역 성능을 유지하는 구조를 제안하였음. 시뮬레이션 결과에 따르면, 제안된 FDRL은 기존의 DDPG 단독 기법이나 ZF(Zero Forcing), MMSE(Minimum Mean Square Error) 등의 전통적인 빔포밍 기법보다 우수한 스펙트럼 효율을 달성하였으며, 특히 기지국 수가 많고 안테나 수가 클수록 높은 처리량을 유지하면서 간섭 완화 성능도 탁월함을 보였음. 또한, 전체 모델을 업로드하지 않고도 10% 수준의 부분 가중치만 공유하였을 때도 거의 동일한 성능을 확보함으로써, 에지 기반 분산 학습의 실용성과 확장 가능성을 입증하였음.

### III. 결론

본 논문에서는 초고속·초저지연 통신이 요구되는 VR 및 홀로그램 콘텐츠 전송 환경에서, THz 대역 무선 네트워크의 빔포밍 최적화를 위한 강화학습 기반 연구들을 살펴보았음. 두 연구 모두 THz 통신 환경에서 강화학습이 빔 관리 및 MAC 계층 성능 향상에 실질적인 기여를 할 수 있음을 보여주며, 향후 차세대 무선 네트워크 설계에 있어 유망한 방향성을 제시하였음.

## ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구 센터육성지원사업(IITP-2025-RS-2022-00156353) 및 2025년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (RS-2019-NR040074)

## 참 고 문 헌

- [1] T. S. Rappaport et al., "Wireless Communications and Applications Above 100 GHz: Opportunities and Challenges for 6G and Beyond," in *IEEE Access*, vol. 7, pp. 78729–78757, 2019.
- [2] A. Ahmadi and O. Semiari, "Reinforcement Learning for Optimized Beam Training in Multi-Hop Terahertz Communications," *ICC 2021 – IEEE International Conference on Communications, Montreal, QC, Canada, 2021*.
- [3] P. -C. Hsu, L. -H. Shen, C. -H. Liu and K. -T. Feng, "Federated Deep Reinforcement Learning for THz-Beam Search with Limited CSI," *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*, London, United Kingdom, 2022.