

6G 비지상 네트워크 최적화를 위한 최신 강화학습 기술 동향

김준영*, 장희연*, 박서영*, 장지석*, 정소이

아주대학교 AI 융합네트워크학과*, 아주대학교 전자공학과

{*junzero0615, *timd0801, *syjm0819, *star12191254, sjung}@ajou.ac.kr

Reinforcement Learning for 6G Non-Terrestrial Network Optimization: Recent Advances and Trends

Junyoung Kim*, Huiyeon Jang, Seoyeong Park, Jiseok Jang, Soyi Jung

*Dept. of Artificial Intelligence Convergence Network, Ajou University,
Dept. of Electrical and Computer Engineering, Ajou University

요약

6G 비지상 네트워크(non-terrestrial networks, NTN)는 위성 및 고고도 플랫폼을 통해 전 지구적 통신을 제공하지만, 고속 이동성, 장거리 지연, 동적 토폴로지 등으로 인해 전통적 제어 방식의 한계를 드러낸다. 강화학습(reinforcement learning, RL)은 이러한 복잡하고 동적인 환경에서 최적 정책을 도출하기 위한 인공지능 기술로써 NTN 최적화의 핵심 기술로 주목받고 있다. 이에 본 논문에서는 6G NTN을 위한 RL 기술의 최신 동향을 3rd generation partnership project (3GPP) 표준을 기반으로 NTN의 핵심 특성과 기술적 과제에 따른 RL 기술의 최신 연구 동향을 제시한다.

I. 서론

차세대 6G 이동통신은 지상, 해상, 공중 영역을 아우르는 초연결 시대를 지향하며, 비지상 네트워크(non-terrestrial networks, NTN)는 이를 실현하는 핵심 인프라이다. 그러나, NTN은 빠른 이동성으로 인해 빈번한 핸드오버, 자원할당 스케줄링 최적화 및 연산량 급변화 등 다양한 이슈가 존재한다. 해당 문제점들을 최적화하기 위해 6G에서는 인공지능(artificial intelligence, AI)을 접목하여 연산량 경량화 및 시스템 성능 향상을 목표로 이동통신 시스템에 AI를 접목하여 최적화하고자 한다. 특히, AI 기술 중 강화학습(reinforcement learning, RL)은 NTN의 동적으로 변화하는 토폴로지와 빠른 이동성을 고려한 최적화 기술로써 주목받고 있다. 이에, 본 논문에서는 3rd generation partnership project (3GPP)에서 논의된 주요 6G key issues를 중심으로 AI/RL 관점에서의 최적화를 진행한 최신 강화학습 기술 동향에 대하여 논한다 [1].

II. 3GPP NTN key issue 기반 강화학습 최적화 연구 동향

2.1 3GPP NTN key issue : 핸드오버 최적화 연구 동향

저궤도 위성 (low Earth orbit, LEO)를 활용한 NTN에서는 LEO의 빠른 이동성으로 인해 지상 단말의 빈번한 핸드오버는 필연적이다. 비효율적인 핸드오버는 통신 중단, 높은 지연 시간 및 패킷 손실을 유발하여 사용자의 quality of service (QoS)를 심각하게 저하될 수 있다. 이를 위해, 핸드오버 결정 행동을 최적화하기에 적합한 deep Q-network (DQN) 기법을 통한 다양한 연구가 진행되었다 [2][3]. 더하여, LEO mega-constellation 시나리오에서는 상호 영향을 고려하는 협력적인 결정이 필수적이다. 이에, 멀티 에이전트 강화학습 기법 (multi-agent reinforcement learning, MARL)을 활용하고 조건부 핸드오버 절차를 기반으로 한 대규모 단말 에이전트의 MARL 분산 핸드오버 연구가 진행되었다 [4].

2.2 3GPP NTN key issue : 빔포밍 최적화 연구 동향

빔포밍은 위성 안테나의 방사 패턴을 특정 사용자에게 집중시켜 신호 품질을 향상시키고 사용자 간섭을 줄이는 핵심 기술이다. NTN의 동적 채널

변화 및 다중 사용자 간섭 환경 내에서 최적 빔포밍 벡터를 실시간 결정하기 위해 각 위성 에이전트 간 MARL을 기반하여 협력적 빔포밍 벡터 결정 행동을 통한 최적화 연구가 진행되었다 [6].

2.3 3GPP NTN key issue : 자원할당 최적화 연구 동향

NTN의 동적 특성으로 인한 끊임없이 변화하는 채널 환경 속에서 각 사용자에게 최적의 자원을 실시간으로 분배하는 것은 매우 복잡한 최적화 문제이다. NTN에서의 자원할당 최적화 기법으로는 이산적/연속적 행동으로 다양하게 최적화 문제를 해결할 수 있는데, 먼저 가치 기반 DQN 접근 방식이 존재한다 [7]. 해당 연구에서는 멀티빔 위성 시스템의 동적 채널 할당 문제에서 DQN을 통해 장기적인 스펙트럼 자원 활용률을 최대화하는 채널 할당 정책을 학습한다. 더하여, DQN을 통해 사용자 그룹화, 빔 위치 및 커버리지 반경, 그리고 서버 채널 및 전력 할당을 공동으로 최적화하는 기법도 제안되었다 [8]. 또한, 연속적인 정책을 점진적으로 안정화하는 proximal policy optimization (PPO) 기법을 통해 6G 위성 통신 시스템에서 rate splitting multiple access (RSMA) 기술을 활용하여 전력 및 자원 할당 최적화 연구도 진행되었다 [9].

2.4 3GPP NTN key issue : 라우팅 최적화 연구 동향

NTN에서의 위성 네트워크의 동적 토폴로지와 빠른 이동성은 위성 및 변동하는 링크 품질로 인해 위성 간 링크가 자주 변경되어 기존 TN에서 활용된 라우팅 방법으로는 위성 네트워크에 적용하기가 어렵다. 이에, 저지연 및 글로벌 커버리지를 제공하기 위해 가장 널리 활용되는 기술은 DQN을 활용한 경로 최적화 기술이다 [10]. 추가적으로, 실제 트래픽 패턴과 유사한 가상 트래픽 데이터를 생성하여 에이전트를 학습하는 generative adversarial network (GAN) 기법을 활용한 라우팅 기법 연구도 진행되었다 [11]. 또한, 다양한 네트워크 환경에서 학습하는 방법에 대하여 사전 학습을 진행하여 몇 개의 샘플로 트래픽 분포에 따른 최적 라우팅 정책을 도출해내는 meta RL 연구도 존재한다 [12].

III. 최신 강화학습 기법을 통한 NTN 최적화 연구 동향

3.1 양자 강화학습 기반 NTN 최적화 연구 동향

양자 얽힘 현상과 양자 중첩 현상의 상관관계를 통한 양자 신경망 (quantum neural network, QNN)은 양자 역학적 속성을 활용하여 적은 파라미터로 복잡한 상태 공간 최적화가 가능한 강화학습 기법이다. 이에, mega-constellation LEO 위성 라우팅 시나리오에서 대규모 위성 에이전트의 행동 공간 최적화를 위해 양자 강화학습 기반 경량화 라우팅 알고리즘 기법 연구가 진행되었다 [13]. 더하여, 지상-비지상-우주 다계층 간 네트워킹 라우팅 최적화 기법 연구도 QNN 기반 강화학습 연구가 진행되었다 [14].

3.2 계층적 강화학습 기반 NTN 최적화 연구 동향

차세대 6G 네트워크는 지상-공중-우주 통합 시나리오 운용으로 전 지구적 커버리지를 구축하고 저지연 통신을 지향한다. 특히, 해당 시나리오는 계층적인 구조를 지니고 있으므로, 자원할당, 라우팅, 스케줄링 문제 등 계층적 강화학습을 통해 최적화하기에 적합하다. 이에, 우주-지상 통합 시나리오에서 계층적인 라우팅을 통해 대규모 internet of things (IoT) 단말의 자원을 할당하기 위한 계층적 강화학습 연구가 진행되었다 [15]. 더하여, traffic steering을 통한 상위 계층에서는 steering 결정 정책, 하위 계층에서는 자원 할당 스케줄링을 진행하는 계층적인 구조를 통해 최적 자원 할당을 진행하였다 [16].

3.3 대규모 언어모델 강화학습 기반 NTN 최적화 연구 동향

최근 AI 연구는 대규모 언어모델을 활용한 생성형 AI를 기반으로 다양한 연구가 진행되고 있다. 특히 NTN의 동적인 특성에서 실시간 통신 성능을 만족하기 위해서는 빠른 계산과 추론, 사용자 특성을 반영한 최신 AI 기술 접목이 요구된다. 실제 6G 운용 시나리오에서 large language model (LLM)을 활용하여 사용자 의도를 파악 및 제한된 컴퓨팅 리소스 최적화를 위한 파라미터 최적화, 의도 검증, 네트워크 어플리케이션 선택 진행에 관련된 기법이 제안되었다 [17]. 향후 6G 네트워크의 동적인 특성을 제어하기 위한 다양한 LLM 관련 연구가 진행될 것으로 예상된다.

IV. 결론

본 논문에는 3GPP에서 정의된 주요 NTN key issue에 따른 RL 기반 최적화 연구와 최신 AI/RL이 적용된 NTN 최적화 연구 동향에 대하여 논하였다. 향후 6G use case scenario에 적합한 최적화 연구로 확장한다.

참 고 문 헌

[1] 3GPP TR 38.811 v16.1.0, "Study on new radio (NR) to support non-terrestrial networks (Release 15)," 3rd Generation Partnership Project (3GPP), Technical Report 38.811, September 2020.

[2] N. Badini, M. Jaber, M. Marchese, and F. Patrone, "User-centric satellite handover for multiple traffic profiles using deep Q-learning," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 60, no. 6, pp. 8591-8604, December 2024.

[3] J. Yang, Z. Xiao, H. Cui, J. Zhao, G. Jiang, and Z. Han, "DQN-ALrM based intelligent handover method for satellite-ground integrated network," *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 4, pp. 977-990, August 2023.

[4] H. Liu, Y. Wang, P. Li, and J. Cheng, "A multi-agent deep

reinforcement learning-based handover scheme for mega-constellation under dynamic propagation conditions," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 13579-13596, October 2024.

[6] J. Yu, C. Hua, L. Liu, and P. Gu, "Joint beamforming optimization for user-centric multi-satellite system," in *Proc. IEEE International Conference on Communications (ICC)*, Denver, CO, USA, June 2024.

[7] S. Liu, X. Hu, and W. Wang, "Deep reinforcement learning based dynamic channel allocation algorithm in multibeam satellite systems," *IEEE Access*, vol. 6, pp. 15733-15742, February 2018.

[8] S. Zhang, R. Chai, C. Liang, and Q. Chen, "Dynamic resource allocation for multibeam satellite communication systems," *IEEE Internet of Things journal*, vol. 11, no. 22, pp. 36907-36921, November 2024.

[9] J. Huang, Y. Yang, L. Yin, D. He, and Q. Yan, "Deep reinforcement learning-based power allocation for rate-splitting multiple access in 6G LEO satellite communication systems," *IEEE Wireless Communications Letters*, vol. 11, no. 10, pp. 2185-2189, October 2022.

[10] P. Zuo, C. Wang, Z. Yao, S. Hou, and H. Jiang, "An intelligent routing algorithm for LEO satellites based on deep reinforcement learning," in *Proc. 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, September 2021.

[11] Q. Guo, F. Tang, and N. Kato, "Routing for space-air-ground integrated network with GAN-powered deep reinforcement learning," *IEEE Transactions on Cognitive Communications and Networking*, vol. 11, no. 2, pp. 914-922, April 2025.

[12] Y. Hu, X. Wang, and W. Saad, "Distributed and distribution robust meta reinforcement learning for data pre-storage and routing in cube satellite networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 17, no. 1, pp. 129-141, January 2023.

[13] G. S. Kim, S. Lee, I.-S. Cho, S. Park, and J. Kim, "Quantum reinforcement learning for lightweight LEO satellite routing," *IEEE Internet of Things Journal*, vol. 12, no. 14, pp. 28986-29004, July 2025.

[14] G. S. Kim, Y. Cho, S. Park, S. Jung, and J. Kim, "Quantum reinforcement learning for joint cube satellites and high-altitude long-endurance aerial vehicles in SAGIN," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 61, no. 4, pp. 9490-9510, August 2025.

[15] J. Zhu, Y. Shi, Y. Zhou, C. Jiang, and L. Kuang, "Hierarchical learning and computing Over space-ground integrated networks," *IEEE Transactions on Mobile Computing*, vol. 24, no. 10, October 2025.

[16] M. A. Habib, H. Zhou, P. Rivera, M. Elsayed, M. Bavand, R. Gaigalas, Y. Ozcan, and M. Kantarci, "Hierarchical reinforcement learning based traffic steering in multi-RAT 5G deployments," in *Proc. 2023 IEEE International Conference on Communications (ICC)*, Rome, Italy, May-June 2023.

[17] M. A. Habib, and et.al., "Harnessing the power of LLMs, informers and decision transformers for intent-driven RAN management in 6G," *IEEE Transactions on Network Science and Engineering*, Early Access.