

다중궤도 LEO 위성의 링크 버짓 및 가시성을 고려한 DRL 기반 전송률 최대화 기법

문찬준, 이예린, 이호원
아주대학교

{mcjmcyj0412, yerin1205, howon}@ajou.ac.kr

DRL-Based Throughput Maximization Considering Link Budget and Visibility in Multi-Orbit LEO Satellite Networks

Chanjun Moon, Yerin Lee, Howon Lee
Ajou University

요약

본 논문에서는 다중궤도 low earth orbit (LEO) 위성 네트워크에서 링크 예산과 가시 시간을 고려한 deep reinforcement learning (DRL) 기반 전송률 최대화 기법을 제안한다. 제안 기법은 위성 간 협력 학습을 통해 제한된 가시 시간 내에서도 효율적인 정책 학습이 가능하다. 시뮬레이션을 통해 제안 기법이 기존 벤치마크 기법 대비 빠른 수렴 속도와 높은 누적 보상 성능을 가지는 것을 보인다.

I. 서론

차세대 무선 통신 시스템은 전 지구적 커버리지와 시공간적 제약이 없는 서비스 제공을 목표로 한다. 이를 위해 low earth orbit (LEO) 위성 네트워크가 지상 네트워크와 상호 보완적으로 활용되고 있다 [1]. 그러나 LEO 위성은 빠른 이동성과 가시성 변화로 인해 링크 예산 관리와 동적 자원 할당이 어렵다는 한계가 있다 [2-3]. 본 논문에서는 다중궤도 LEO 위성 네트워크에서 링크 버짓과 가시 시간을 고려한 협력 deep reinforcement learning (DRL) 기반 전송률 최대화 기법을 제안한다. 제안 기법은 실제 시스템 제약을 반영한 최적화 문제를 해결하기 위해, 위성 간 학습 정보 공유를 통한 협력 학습 프레임워크를 활용한다.

II. 시스템 모델 및 제안기법

본 논문에서는 지구를 반지름 R 의 구로 가정하며 고도 h 에서 궤도 $j \in J$ 에 LEO 위성 $s \in S_j$ 가 존재한다. 각 궤도에는 N_s 개의 위성이 공존한다. 위성과 지상 사용자 링크 l 의 signal-to-interference-plus-noise ratio (SINR, $\Gamma_{b,u}^J$)은 다음과 같이 계산한다.

$$\Gamma_{s,u}^J = \frac{P_{s,u}^{RX,l}}{\sum_{s' \neq s} I_{b',u}^J + \sigma^2}. \quad (1)$$

여기서 $P_{s,u}^{RX,l}$ 는 위성 s 로부터 사용자 u 로의 l 링크에서의 수신전력, $I_{b',u}^J$ 는 l 링크에서 빔 b' 로부터의 간섭, σ^2 는 잡음전력을 나타낸다. 제안 기법은 순차적 다중 에이전트 의사결정 구조를 적용하여 지상 게이트웨이가 먼저 서비스를 제공할 위성을 선택하고 이후 선택된 위성에서 자원 할당을 수행한다. 지상 사용자들의 sum-rate을 최대화하기 위한 최적 위성 선택 및 자원할당 문제를 위한 Markov Decision Process (MDP)는 다음과 같이 정의한다.

● Agent Type-1: Gateway

$$S_G(t) = [E_i(t), T_{vis}(t), X(t), Y(t), Z(t), t], A_G(t) = [S | S \in T_{vis}(t)]. \quad (2)$$

게이트웨이의 상태 $S_G(t)$ 는 위성과 지상 사용자 링크 평균, 가시 위성들의 고각, 남은 가시 시간, 위성의 위치로 구성되며 행동 $A_G(t)$ 은 최적 위성 선택을 포함한다. 게이트웨이는 sum-rate 최대화를 위한 서비스 위성 선택 최적화를 수행한다.

● Agent Type-2: LEO satellite

$$S_s(t) = [(P_1 \dots P_B), X(t), Y(t), Z(t)], A_G(t) = [\pm \Delta P_b, \Delta P_b = 0]. \quad (3)$$

선택된 위성의 상태 $S_s(t)$ 는 전송전력과 각 위성의 위치로 구성되며, 행동 $A_s(t)$ 은 전송 전력 증대/감소/유지를 포함한다. 선택된 위성은 서비스 영역 내 간섭 최소화 및 서비스 최대화를 달성할 수 있는 전송 전력 최적화를 수행한다. 분산형 다중 에이전트 환경에서 개별 보상을 사용할 경우 전역 최적해 도달의 어려움이 발생한다. 이를 해결하기 위해 본 논문에서

는 sum-rate 최대화를 위해 처리량과 남은 가시 시간의 가중합으로 구성된 공유 보상을 다음과 같이 설계한다.

$$r^{shared}(t) = \alpha R_{sumrate} + (1 - \alpha) R_{vis}. \quad (4)$$

LEO 위성의 빠른 이동성으로 특정 지역에 대한 가시 시간이 매우 짧아 충분한 학습 데이터를 확보하기 어렵다는 문제가 있다. 이를 해결하고자 본 논문에서는 동일 지역을 서비스하는 위성들이 학습 과정에서 얻은 replay buffer와 weight 값들을 상호 공유함으로써 학습 효율성을 높이는 위성 간 협력 학습 기반의 DRL 기법에 기반하여 최적화를 수행한다.

III. 시뮬레이션 결과 및 결론

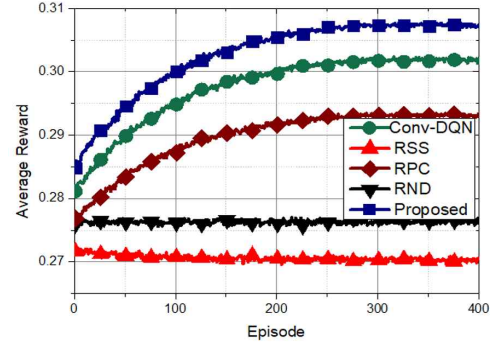


그림 1. Average reward vs. episode

그림 1은 제안한 DRL 기반 위성 선택 및 전력 제어 기법과 비교 기법들의 에피소드에 따른 누적 보상 추이를 보여준다. 궤도 당 20 개의 위성이 배치된 4 개의 다중 궤도로 구성된 위성 군집 환경을 가정하였으며, 위성은 고도 600km에서 운용된다. 벤치마크 기법으로는 'Conventional DQN (Conv-DQN)', 'Random Power Control (RPC)', 'Random Satellite Selection (RSS)', 'All-Random (RND)'를 고려한다. 제안 기법은 모든 벤치마크 대비 빠른 수렴과 높은 성능을 보이며 특히, Conv-DQN 대비 약 1.63%의 성능 향상을 통해 협력 학습으로 위성들이 짧은 가시 시간 내에서도 효율적 정책 학습이 가능함을 입증한다. 이는 제안 기법이 다중궤도 LEO 위성망에서 링크 예산과 가시성을 고려한 전송률 최적화에 효과적임을 시사한다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원 (No. RS-2025-02303435)과 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원(No. RS-2024-00396992)과 2025년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(RS-2025-00563401)을 받아 수행된 연구임.

참고 문헌

- [1] H. Lee et al., "Towards 6G Hyper-Connectivity: Vision, Challenges, and Key Enabling Technologies," IEEE/KICS Journal of Communications and Networks, vol. 25, no. 3, pp. 344-354, Jun. 2023.
- [2] J. -H. Lee et al., "Handover Protocol Learning for LEO Satellite Networks: Access Delay and Collision Minimization," in IEEE Transactions on Wireless Communications, vol. 23, no. 7, pp. 7624-7637, July 2024.
- [3] Y. Cho et al., "Multi-Agent Deep Reinforcement Learning for Interference-Aware Channel Allocation in Non-Terrestrial Networks," in IEEE Communication Letters, vol. 27, no. 3, pp. 936-940, March 2023.