

# Slimmable Neural Network 강화학습 기반 지구 저궤도 위성 라우팅

김규선, 백한결, 박수현, 김종헌

고려대학교

kingdom0545@korea.ac.kr, 67back@korea.ac.kr, soohyun828@korea.ac.kr, joongheon@korea.ac.kr

## Slimmable Neural Networks Reinforcement Learning-based Robust Low Earth Orbit (LEO) Satellite Routing

Gyu Seon Kim, Hankyul Baek, Soohyun Park, Joongheon Kim

Korea Univ

### 요약

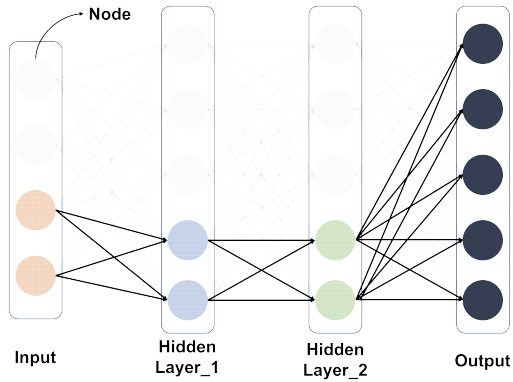
빠르게 발전하는 위성 통신 분야에서 6G와 같은 차세대 통신 시스템의 출현은 글로벌 인터넷을 제공하는 위성의 잠재력을 부각시켰다. 여러 위성 중 특히 지구 저궤도 (Low Earth Orbit, LEO) 위성은 상대적으로 낮은 고도로 인해 다른 위성시스템에 비해 통신 프레임워크 구성에 뚜렷한 이점이 존재한다. 그러나 LEO의 경우, 다수의 위성의 빠른 공전속도로 인해 토폴로지가 동적으로 빠르게 변화하고 이는 여러 통신 성능을 저하시킨다. 따라서 이를 극복하고 효율적인 글로벌 인터넷 서비스를 제공하기 위해서는 향상된 통신 성능을 위한 고급 라우팅 기술이 필수적이다. 이에, 본 논문에서는 위성 네트워크의 역동적이고 복잡한 토폴로지에 기계학습 방법 중 하나인 강화학습을 적용하여 강건한 LEO 라우팅 알고리즘을 제안한다. 여러 원인으로 인해 link quality의 불확실성이 존재하는 실제 우주 환경을 묘사하여 랜덤하게 장애 LEO를 배치시키고 이로 인해 발생하는 신경망의 상태 차원 변화 문제를 해결하기 위해 기존의 fully connected layer가 아닌 slimmable neural network를 활용한다. 실제 우주 환경을 묘사한 실험을 통해 해당 알고리즘의 우수성을 평가함과 동시에 급격히 성장하는 위성 통신 분야에 인공지능 (Artificial Intelligence, AI) 기술을 적용하기 위한 새로운 길을 열어준다.

### I. 서론

글로벌 6G 통신 시스템의 발전은 전세계 인터넷 접근성을 향상시키는 획기적인 기술로 떠오르고 있는 위성 통신 네트워크의 발전을 전제로 한다. 특히, 저궤도(Low Earth Orbit, LEO) 위성은 중궤도(Medium Earth Orbit, MEO), 정지궤도(Geostationary Orbit, GEO) 위성에 비해 상대적으로 지구와 더 가까운 위치에서 운용되기 때문에 지연 시간이 줄어들어 통신의 효율성이 높아진다. 이러한 기술적 이점을 활용한 SpaceX의 Starlink 프로젝트는 LEO를 활용한 글로벌 인터넷 서비스가 더 이상 공상 과학에서만 등장하는 소재가 아닌, 현실에 적용 가능한 기술임을 보여주었다 [1]. SpaceX는 Falcon 9과 같은 재사용 로켓의 안정적인 발사 기술을 활용하여 약 2,000개의 LEO 위성을 배치했으며 심지어 우크라이나와 같은 전쟁 지역에서 상용화되고, 40ms 대기 시간 및 50Mbps 다운로드 등 인상적인 통신 성능을 자랑한다. 그러나 짧은 공전주기로 인해 네트워크 토폴로지가 동적으로 빠르게 변화하는 LEO 통신시스템 특성상, 고속의 통신성능을 유지하기 위해서 효율적인 라우팅 기술이 필수적이다. 위성 라우팅은 inter-satellite 또는 satellite-ground station (GS) 간 통신 경로를 설정하는 것을 의미한다. 한편, 실제 우주 환경에서는 우주선, 태양풍 등 위성의 성능 및 궤도요소 그리고 link quality에 영향을 줄 수 있는 다양한 변수들이 존재하여, 원활하게 글로벌 인터넷을 공급하지 못하는 장애 위성이 발생할 수 있다. 본 논문에서는 이러한 현실적인 상황을 묘사하기 위해 무작위로 장애 위성을 배치시킨다. 이때 신경망 상태 입력 차원이 장애 위성의 개수에 따라 바뀌는 문제점이 발생한다. 기존의 완전 연결 계층 (fully connected layer)으로는 학습 중에 바뀌는 신경망 상태 입력 차원에 대응할 수 없다. 이러한 문제점을 해결하기 위해 본 논문에서는 slimmable neural network를 활용하여 장애 위성의 개수가 학습 중에 무작위로 바뀌어도 불구하고 학습이 가능한 강화학습 기반 LEO 라우팅 알고리즘을 제안한다.

### II. 1. 저궤도 위성 통신 시스템

네트워크 기술의 지속적인 발전에도 불구하고, 자연재해나 전쟁으로 인해 여전히 많은 지역에서 인터넷이 단절되는 등, 원활한 글로벌 인터넷 서비스를 제공하는 것은 여전히 어려운 과제로 남아 있다. 36,000km 궤도를 도는 GEO 위성에 의존했던 기존의 위성 통신 시스템은 장거리 신호 지연으로 인해 통신 속도가 크게 저하되는 문제를 겪고 있다. 이러한 문제점을 해결하기 위해, LEO 위성 네트워크는 (i) 위성 고도를 줄여 대기 시간을 줄이고 (ii) 위성 수를 늘려 커버리지를 확장함으로써 본질적으로 지구에서 약 500km 떨어진 곳에 수많은 위성을 배치하여 통신 시스템을 구축하는 방식으로 해결책을 제시한다. LEO 위성 시스템에서 라우팅은 데이터 패킷을 소스에서 목적지로 효율적으로 전송하는 프로세스를 의미한다. 그러나 위성은 미리 설계된 궤도요소(Orbital Element, e.g. inclination angle, mean anomaly)에 따라 지구에 대한 상대적 위치가 계속해서 바뀌기 때문에 지상 네트워크 (Terrestrial Network, TN)에서의 라우팅과 다른 접근 방식이 필요하다. 앞서 언급된 것처럼, LEO 위성의 빠른 궤도 속도(약 7.5km/s)로 인한 잦은 핸드오버는 사용자의 편의성에 악영향을 미친다. 즉, 지속적으로 변화하는 동적 네트워크 토폴로지가 생성되어 대기 시간 증가 및 처리량 감소 등이 일어나 통신 성능에 부정적인 영향을 끼친다. 이러한 LEO 위성 통신 시스템만의 고유한 특성 때문에, 강화학습은 라우팅 문제를 해결하기 위한 유망한 해결책으로 떠오르고 있다. 강화학습은 환경(Environment)에 놓여진 에이전트(Agent)가 환경과 계속해서 상호작용을 하며 학습을 진행하기 때문에, 동적이고 불확실성이 존재하는 환경에 보다 쉽게 대응을 할 수 있다. 앞서 언급된 것처럼, LEO 위성 시스템에서 라우팅은 매우 동적이며 불확실성이 존재하는 환경이다. 강화학습은 이러한 환경에 쉽게 적응하며 최적에 가까운 정책을 학습할 수 있기 때문에 LEO 라우팅에 적합하다.



[그림 1] Slimmable Neural Network 구조

## II. 2. 강화학습과 Deep Q-Learning

강화학습은 에이전트가 환경과의 상호 작용을 통해 학습하여 누적 보상을 최대화하는 기계 학습 방법론 중 하나이다 [2]. 환경을 탐색하는 에이전트는 다양한 상태에서 행동의 가치를 학습하고 이 가치를 극대화하기 위해 지속적으로 행동을 조정하며 학습해 나간다. 에이전트와 환경 사이의 반복적인 상호작용이 강화학습의 핵심이며, LEO 라우팅 같은 순차적 의사 결정 문제에 강화학습은 특히 적합하다. 강화학습 알고리즘은 가치를 학습하느냐, 정책을 학습하느냐에 따라서 크게 ‘value based method’와 ‘policy gradient based method’로 나뉜다. 본 논문에서는 value based method의 가장 대표적인 방법이자, Q-value를 신경망을 통해 학습하는 deep Q learning (DQN)을 통해 위성 라우팅 알고리즘을 설계한다. DQN은 Q-Learning과 유사한 원리로 작동하며 Bellman 최적 방정식을 학습 대상(target)으로 간주하고 학습하기 때문에, 학습에 사용되는 손실 함수는 식 (1)로 정의된다.

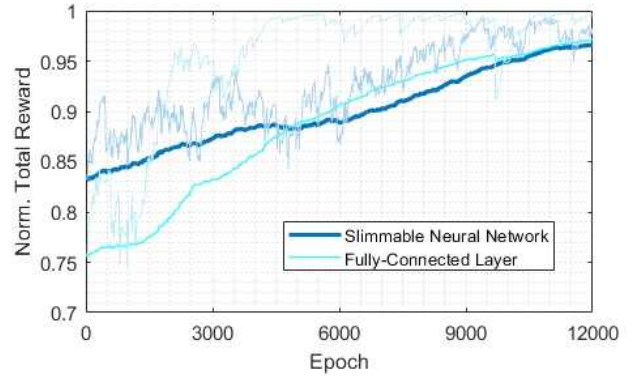
$$L(\theta)' = E[(R + \gamma \max_{a'} Q_{\theta}(s', a') - Q_{\theta}(s, a))^2] \quad (1)$$

식 (1)에서  $L(\theta)'$ ,  $R$ ,  $\gamma$ ,  $Q_{\theta}(s', a')$ , 그리고  $Q_{\theta}(s, a)$ 는 각각 손실 함수, 보상 함수, 감쇄비, 타겟 네트워크에 의해 추정된 Q값, 그리고 에이전트가 얻은 Q값이다. 여기서  $\theta$ 는 신경망 파라미터이다. (1)에서 정의된 것과 같이 에이전트는 보상함수를 극대화하는 방향으로, 손실함수를 최소화하는 방향으로 학습을 진행하기 위해 손실함수에 대한 신경망 파라미터  $\theta$ 의 영향력을 파악해야 한다. 이를 위해 각 신경망 파라미터에 대해 손실 함수를 편미분 해야 하는데, 이에 대한 DQN에서의 신경망 파라미터 업데이트 식은 식 (2)와 같다.

$$\theta' \leftarrow \theta + \alpha (R + \gamma \max_a Q_{\theta}(s_{t+1}, a_{t+1}) - Q_{\theta}(s_t, a_t)) \times \nabla_{\theta} Q_{\theta}(s_t, a_t) \quad (2)$$

## II. 3. Slimmable Neural Networks

[그림 1]에 표현된, slimmable neural network는 단일 신경망 모델이 다양한 폭(width)의 구성으로 동작할 수 있도록 설계된 네트워크를 의미한다. 즉, 일정비용에 따라 input nodes와 hidden layer nodes를 줄여 유연하게 학습이 가능한 신경망 구조이다. 이러한 네트워크는 LEO 시스템과 같이 하드웨어 자원이 제한된 환경에서 특히 유용하며, 여러 가지 컴퓨팅 요구 사항에 따라 신경망을 동적으로 제어할 수 있다. 이를 통해 두꺼운 신경망을 사용할 때는 더 높은 성능을 보여주고, 얇은 신경망을 사용할 때는 더 적은 계산 비용을 필요로 한다. Slimmable neural network의 가장 큰 장점은 ‘유연성’이다. LEO 위성 라우팅과 같이 장애 위성이 무작위로 배치되는 환경에서는 신경망의 상태 입력 (input layer) 차원이 장애 위성의



[그림 2] 훈련에서의 보상값 추이

개수에 따라 유동적으로 변한다. 기존의 fully connected layer는 이러한 유동적 신경망 환경에 대응하지 못 하지만, slimmable neural network는 하나의 네트워크가 다양한 크기로 조정 가능하기 때문에 다양한 환경에서 유연하게 배치될 수 있다. 이는 학습 시 필요한 파라미터와 같은 자원 사용량을 최적화하면서도 성능을 유지할 수 있어 위성 시스템과 같이 자원이 한정적인 환경에 효과적으로 적용될 수 있다.

## II. 4. 성능평가

[그림 2]는 훈련에서 제안하는 알고리즘(slimmable neural network)과 대조군(fully connected layer)의 보상값 추이를 보여준다. 강화학습에서는 보상값을 극대화하는 방향으로 학습이 이루어지기 때문에 보상함수의 추이를 보고 해당 모델의 학습 여부를 판단할 수 있다. Slimmable neural network가 더 적은 수의 파라미터와 자원을 사용했음에도 불구하고 종료 조건에서의 보상값이 fully connected layer와 유사함을 확인할 수 있다. 이는 장애 위성이 무작위로 발생하는 LEO 위성 라우팅 환경에서 slimmable neural network가 효율적으로 적용될 수 있음을 의미한다.

## III. 결론

본 논문에서는 LEO 시스템의 고유한 특성으로 인해 발생하는 라우팅 문제를 강화학습을 통해 해결하였다. 특히, 장애 위성의 발생에 따라 발생하는 신경망 상태 입력 차원 변화에 대응하기 위해 slimmable neural network를 활용하여 효율적인 라우팅 알고리즘을 제안하였다. Slimmable neural network 강화학습 기반 라우팅 알고리즘은 파라미터 사용을 최소화하면서도 fully connected layer와 비슷한 성능을 보여주었다.

## ACKNOWLEDGMENT

This work was supported by Institute of Information communications Technology Planning Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2021-0-00467, Intelligent 6G Wireless Access System)

## 참고 문헌

- [1] M. Neinavaie, J. Khalife and Z. M. Kassas, "Acquisition, Doppler Tracking, and Positioning With Starlink LEO Satellites: First Results," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 3, pp. 2606-2610, June 2022.
- [2] G. -P. Antonio and C. Maria-Dolores, "Multi-Agent Deep Reinforcement Learning to Manage Connected Autonomous Vehicles at Tomorrow's Intersections," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 7, pp. 7033-7043, July 2022.