

# 다중에이전트 강화학습 기반 LEO 위성에서의 동적 트래픽 대응 빔 호핑 최적화

이재열, 김태윤\*, 문태한\*, 이재영\*, 김재현\*\*

아주대학교 우주전자정보공학과, \*아주대학교 AI융합네트워크학과, \*\*아주대학교 전자공학과

{jaeyel98, \*xodbsxogjs, \*ansxogks3, \*dlwodudd4865, \*\*jtkim}@ajou.ac.kr

## Optimizing Beam-Hopping in LEO Constellation Systems with Random Cell Traffic Using Multi-Agent Reinforcement Learning

Jaeyeol Lee, Tae-Yoon Kim\*, Taehan Moon\*, JaeYeong Lee\*, Jae-Hyun Kim\*\*

Department of Space Survey Information Technology, Ajou University,  
\*Department of Artificial Intelligence Convergence Network, Ajou University,  
\*\*Department of Electrical and Computer Engineering, Ajou University

### 요약

저궤도 위성에 적용되는 빔 호핑 기술은 지상 셀의 트래픽 수요에 따라 자원을 동적으로 할당하는 장점을 제공한다. 하지만, 지상 셀 트래픽의 동적인 변화, 가시 위성의 수, 그리고 고도각 제약 등을 종합적으로 고려한 스케줄링 기법에 대한 연구는 아직 미흡하다. 따라서 본 논문에서는 동적 트래픽 분포, 그리고 가시성 조건을 함께 고려하여 저궤도 위성 자원 할당 효율을 최적화하는 심층 강화학습 기반의 빔 호핑 스케줄링 알고리즘을 제안한다.

### I. 서론

6G 시대에는 통신 서비스의 공간적 한계를 극복하기 위해 위성 기반 비지상 네트워크에 대한 연구가 지속되고 있다. 특히, 동적인 지상 트래픽 수요에 맞춰 자원을 유연하게 할당하는 빔 호핑 기술이 주목받고 있으나, 기존 연구들은 주로 정지궤도 위성에 집중하거나 다중 위성 환경의 현실적인 제약을 고려하지 않는 한계가 있다. 이러한 문제점을 해결하기 위해, 본 논문에서는 다중 빔 저궤도 위성 네트워크를 위한 다중 에이전트 심층 강화학습 기반의 빔 호핑 스케줄링 기법을 제안한다. 수많은 지상 셀의 불규칙한 트래픽 요구를 만족시키기 위해 MCS 인덱스 기반의 보상 함수를 설계하였고, 지상 셀들의 변하는 트래픽 수요를 효과적으로 충족시키도록 빔 호핑 스케줄링을 최적화하고, 동시에 모든 활성 셀을 서비스하는 데 필요한 최적의 위성 수를 결정하는 multi-agent deep Q-network (MADQN) 빔호핑 알고리즘을 제안한다.

### II. 시스템 모델

본 논문에서는  $S$ 개의 저궤도 위성으로 구성된 군집 위성을 고려하면 위성은  $S = \{s | s = 1, 2, \dots, S\}$ 으로 구성된다. 위성은  $G$ 개의 지상 셀에 연결하고  $G = \{g | g = 1, 2, \dots, G\}$ 으로 구성된다. 위성은 시간 슬롯  $0 \leq t \leq T$ 에서 Earth-Centered, Earth-Fixed (ECEF) 좌표는  $l_s(t) = (x_s(t), y_s(t), z_s(t))$ 로 표현할 수 있다. 모든 위성은 고도  $h$ 에서 운용되며 각 지상 셀의 트래픽 수요는  $b_g$ 로 불규칙적으로 생성된다. 고려하는 LEO 위성군 네트워크는 시간 슬롯 방식으로 동작하며, 총 서비스 시간  $T$ 는 서비스 시간이  $\delta_t$ 인 다수의 슬롯으로 나눌 수 있다. 이에 따라 위성의 움직임은  $l_s(t+1) - l_s(t) = V\delta_t$ 로 표현할 수 있으며  $V$ 는 위성의 속도이다. 각 시간 슬롯  $t$  동안, 저궤도 위성은 지상 셀

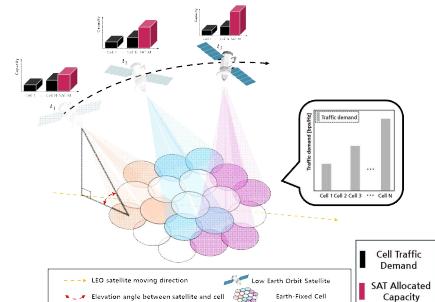


그림 1. 제안하는 빔호핑 알고리즘

들의 불균일한 트래픽 수요를 처리하기 위해 빔 호핑 스케줄링을 수행한다. 시스템 모델은 3rd Generation Partnership Project (3GPP) 문서를 기반으로 하였다[1, 2]. 그림 1은 빔 호핑을 위한 시스템 모델이다.

### III. 최적화 문제 정의

본 논문에서는  $t$ 에서 위성  $s$ 에 의해 서비스되는 지상 셀의 수를 최대화하는 것을 목표로 한다. 이는 다중 빔 LEO 위성의 제한된 용량과 불균일한 트래픽 요구사항을 고려하면서, 지상 셀의 불규칙한 트래픽 수요  $D_{s,g}$ 를 만족시킨다. 최적화 수식은 아래와 같다.

$$P1 : \underset{x, y, z, I}{\text{maximize}} \sum_{s=1}^S \sum_{g=1}^G \sum_{t=1}^T S^{s,g}(t), \quad (1)$$

$$c1 : I_{s,g}(t) \in \{0, 1\}, \forall_s, \forall_g, \forall_t, \quad (2)$$

$$c2 : \sum_{s=Sg}^S C_{s,g}(t) = 1, \quad (3)$$

$$c3 : \theta_{s,g}^e \geq \theta_{\min}, \forall_s, \forall_g, \forall_t, \quad (4)$$

$$c4 : D_{s,g} \geq D_g^{th}, \forall_s, \forall_g, \forall_t, \quad (5)$$

표 1. 시뮬레이션 파라미터

Parameters	Value	Parameters	Value
위성 고도	530 km	EIRP 밀도	34 dB
주파수	2 GHz	최소 고도각	25°
총 시스템 트래픽	263 MHz	위성 속도	7.5 km/s
지상 셀 수	40	지상 셀 반경	25 km
위성 안테나 이득	30 dB	에피소드	10,000
안테나 구경 반지름	2	최대 트래픽	14 MHz
수신 안테나 이득	0 dB	최소 트래픽	3 MHz
잡음 지수	9 dB	평균 트래픽	9 MHz

수식 P1에서 목적 함수는 시간  $t$ 에 다중 범 저궤도 위성에 의해 서비스되는 지상 셀의 수  $S^{s,g}(t)$ 를 최대화하는 것이다. 제약 조건 C1은  $I^{s,g}(t)$ 가 불리언 변수임을 나타낸다. 제약 조건 C2는 특정 시간  $t$ 에 하나의 지상 셀은 위성군 내의 단 하나의 위성에 의해서만 서비스를 받는 것을 보장한다. 제약 조건 C3은 LEO 위성과 지상 셀 사이의 고도각이 최소 고도각  $\theta_{\min}$  이상이어야 함을 보장하며, 제약 조건 C4는 위성  $s$ 와 지상 셀  $g$  간의 트래픽이 MCS 테이블에 명시된 불규칙 트래픽 수요를 초과해야 함을 보장한다.

State  $s_s(t)$ 는 아래와 같이 구성한다.

$$s_s(t) = (l_s(t), \{b_1(t), \dots, b_g(t)\}, \{I_1(t), \dots, I_g(t)\}, B_s(t)), \quad (6)$$

$l_s(t)$ 는 시간  $t$ 에서  $s$  번째 저궤도 위성의 ECEF 좌표를 나타내고  $\{b_1(t), \dots, b_g(t)\}$ 는 시간  $t$ 에서 각 지상 셀의 트래픽 수요를 나타내고,  $\{I_1(t), \dots, I_g(t)\}$ 는 시간  $t$ 에서 각 지상 셀의 불리언 변수를 나타내고  $B_s(t)$ 는 위성의 잔여 용량을 나타낸다.

빔 호핑스케줄링은 매 sub-step에서 각 에이전트의 ECEF 좌표를 기반으로 최적화되며, action space은 두 개의 파라미터로 구성된다. 여러 위성이 동시에 동일한 셀을 서비스하려는 충돌을 방지하기 위해, 사전에 정의된 순차적 스케줄링 전략이 사용된다. 이 전략에 따라,  $S$ 개의 위성 각각은 할당된 스케줄링 우선순위에 맞춰 고정된 순서로 자신의 서비스 셀을 선택한다. Action은 아래와 같이 구성한다.

$$a_s(t) = (\pm V_x \delta_t, \pm V_y \delta_t, \pm V_z \delta_t, \bigcup_{g=1}^G I_g(t)), \quad (7)$$

$$\pm V_x \delta_t, \pm V_y \delta_t, \pm V_z \delta_t \text{는 } x, y, z \text{ 방향 이동을 나타내며 } \bigcup_{g=1}^G I_g(t)$$

은 시간  $t$ 에서 에이전트가 선택하는 셀의 집합을 나타낸다.

#### IV. 실험 결과

제안하는 기법의 성능을 평가하기 위해, 표 1과 같이 3GPP 표준을 기반으로 시뮬레이션 파라미터를 설정하였다. 전송 용량은 SINR에 따라 결정되는 MCS 테이블의 변조 방식을 통해 계산된다 [3]. 위성 궤도는 Starlink의 FCC 문서를 참조하여 Walker-Delta 위성군으로 구성하였으며, 대한민국 상공을 통과하는 궤도를 가정하였다 [4]. 각 에피소드에서는 40개의 지상 셀 중 20개의 셀이 무작위로 선택되어 트래픽을 생성하며, 실험에서는 최소 고도각 요구사항을 만족하는 위성을 최대 4개로 고려하였다.

그림 2는 위성 4개일 때 제안하는 MADQN 범 호핑 알고리즘과 비교 기

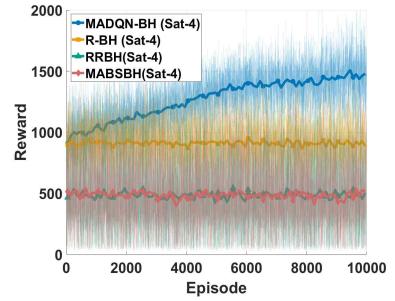


그림 2 위성 4개일 때 reward

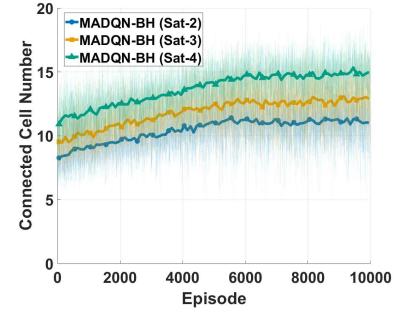


그림 3. 위성 수에 따른 연결된 위성 수

법들의 reward 성능을 비교한다. MADQN-BH는 가장 높은 보상을 달성했으며, 더 많은 지상 셀을 효과적으로 연결했다. 반면, random 범 호핑 (R-BH)은 약 30% 낮은 보상을 보였고, round-robin 범 호핑 (RR-BH)과 Multi armed bandit 범 호핑 (MAB-BH)은 동적 트래픽 환경에서의 무작위 선택 및 제한된 적응성으로 인해 각각 약 69%와 68%의 더 낮은 성능 감소를 나타냈다.

그림 3은 위성 수에 따른 MADQN 범 호핑의 연결 셀 수를 보여준다. 위성 수가 2개에서 4개로 증가했을 때, 연결된 셀의 수는 약 11개에서 15개로 증가함을 확인했다. 비록 위성 수의 증가는 일반적으로 더 높은 간섭을 유발하지만, MADQN 범 호핑은 이 문제를 완화하기 위해 효과적으로 최적의 셀을 선택했다.

#### V. 결론

본 논문에서는 동적인 지상 셀 트래픽 수요 환경에서 다중 LEO 위성 스케줄링을 최적화하기 위해, DQN 기반 범 호핑 (BH) 스케줄링 알고리즘을 제안한다. 제안하는 알고리즘은 간섭을 효과적으로 최소화하고 서비스되는 셀의 수를 최대화한다. 실험 결과, 위성 수가 증가할수록 트래픽 요구 사항을 만족하는 서비스 셀의 수가 향상됨을 확인하였다. 향후 연구에서는 학습된 신경망 가중치를 다른 지역에 적용하여 본 접근법을 확장할 계획이다.

#### 참 고 문 헌

- [1] TR 38.821 V16.4.0, “3rd generation partnership project; study on new radio NR; to support non-terrestrial networks (ntn) (Release 16),” March 2023.
- [2] 3GPP TR 38.811 v15.4.0, “Study on New Radio (NR) to support non-terrestrial networks (Release 15),” Sep. 2019.
- [3] 3GPP TR 38.214 V18.4.0, “3rd generation partnership project; physical layer procedures for data (Release 18),” September 2024.
- [4] Space Exploration Holdings, LLC, “SAT-LOA-20200526-00055,” 2022, pending approval