

# Modular DQN을 활용한 다중 빔 저궤도 위성 네트워크의 빔 호핑 스케줄링, 커버리지, 전송 파워 통합 최적화

신예진, 이충녕, 채승호\*

한국공학대학교

{syy4622, lc9902130509, shchae}@tukorea.ac.kr

## Joint Optimization of Beam Hopping, Coverage, and Transmit Power in Multi-beam LEO Satellite Networks via Modular DQN

Yejin Shin, Chungnyeong Lee, Seong Ho Chae\*

Tech University of Korea

### 요약

본 논문은 다중 빔 저궤도 위성 시스템에서 시스템 처리량과 에너지 효율을 최대화하기 위한 modular Deep Q-Network 기반 빔 호핑 스케줄링, 빔 커버리지, 빔 송신 파워 동시 최적화 방안을 제안한다. 시뮬레이션을 통해, 제안된 알고리즘이 기존 연구 대비 시스템 처리량과 에너지 효율 측면에서 모두 우수한 성능을 보였다.

### I. 서론

저궤도(Low Earth Orbit: LEO) 위성은 광범위한 커버리지, 높은 신뢰성 및 낮은 지연 특성으로 인해 6G 통신의 주요 인프라로 많은 관심을 받고 있다[1],[2]. 빔 호핑(Beam hopping)은 시변하는 트래픽 수요에 대응하여 자원 낭비를 최소화하면서 빔 자원을 할당하는 기술로 다중 빔 위성 시스템의 핵심 기술로 주목받고 있다[3],[4]. 최근, 참고 문헌 [5]와 [6]에서 다중 빔 정지궤도(Geostationary Earth Orbit: GEO) 위성 시스템에서 빔 호핑에 관한 연구가 수행되었다. [5]에서는 빔 호핑 스케줄링과 빔 커버리지를 최적화하지만 고정된 송신 파워를 고려하고, [6]에서는 빔 호핑 스케줄링과 송신 파워를 최적화하지만 고정된 커버리지 반경을 고려하였다. 이는 빔의 자원을 효율적으로 활용하지 못하는 한계를 가진다.

따라서, 본 논문에서는 정지 궤도 위성이 제어하는 다중 빔 저궤도 위성 시스템에서 시스템 처리량과 에너지 효율 동시 최대화를 위한 modular DQN(Deep Q-Network) 기반 빔 호핑 스케줄링, 빔 커버리지, 빔 전송 파워를 동시 최적화 방안을 제안한다. 제안된 기법은 동적으로 변하는 시스템 환경에 따라 최적화된 빔 호핑 스케줄링, 빔 커버리지, 빔 송신 파워를 할당함으로써 시스템 처리량과 에너지 효율을 최대화하고, 시뮬레이션 결과를 통해 기존 연구 대비 우수한 성능을 가짐을 보였다.

### II. 시스템 모델

그림 1은 본 논문에서 고려하는 다중 빔 저궤도 위성 시스템을 보여준다. 본 시스템은 다수의 저궤도 위성이 동일한 궤도를 따라 이동하며, 하나의 정지궤도 위성이 SNOF(Satellite Network OpenFlow) 채널[7]을 통해 모든 저궤도 위성의 빔을 제어한다. 전체  $I$ 개의 저궤도 위성이  $K$ 개의 빔을 생성하여 매 시간 슬롯(time slot)마다  $N$ 개의 지상 셀에 서비스를 제공한다. 각 시간 슬롯 동안에는 네트워크 상태가 준정적(Quasi-static) 상태임을 가정한다.

정지궤도 위성은 지상 셀이 요청하는 트래픽을 저장하기 위한 버퍼를 가지고 있으며, 버퍼는 트래픽이 도착한 순서대로 저장하고 처리하는 FIFO(First-In-First-Out)방식으로 동작한다. 각 시간 슬롯  $t$ 에서의 트래픽 도착량은 포아송 분포(poisson distribution)를 따른다고 가정한다. 위성  $i$ 의 각 빔  $k$ 는 해당 위성  $i$ 의 커버리지  $C_i^{i,k,n}$  내에 포함된 지상 셀들 중

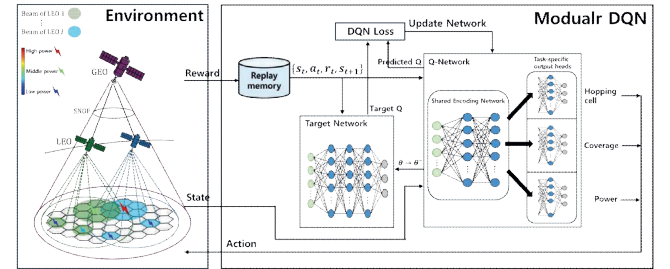


그림 1. 시스템 모델

하나를 중심 셀로 선택하며, 이를 이진 변수  $x_t^{i,k,n}$ 로 정의한다. 시간 슬롯  $t$ 에서 위성  $i$ 의 빔  $k$ 의 커버리지는 이산 반경  $r_t^{i,k} \in \{R_{\min}, R_{\min} + \Delta R, R_{\min} + 2\Delta R, \dots, R_{\max}\}$ 를 갖는 원형 영역으로 모델링한다.  $x_t^{i,k,n} \in \{0,1\}$ 는 셀  $n$ 이 위성  $i$ 의 빔  $k$ 에 의해 조명되었는지를 나타내는 변수로, 중심 셀과 빔의 반경이 결정되면 자동으로 결정된다.  $V_t^{i,k,n} = \{(i,k) \in I \times K | x_t^{i,k,n} = 1\}$ 는 시간 슬롯  $t$ 에 셀  $n$ 을 서비스하는 위성의 모든 빔들의 집합을 의미한다. 이때, 시간 슬롯  $t$ 에서 위성  $i$ 의 빔  $k$ 로부터 셀  $n$ 의 채널 용량은 다음과 같이 정의된다.

$$C_t^{i,k,n} = x_t^{i,k,n} \frac{B_{i,k}}{L_{i,k}} \log_2 \left( 1 + \frac{h_t^{i,k,n} P_t^{i,k}}{\frac{B_{i,k}}{L_{i,k}} N_0 + \varphi_t^{i,k,n}} \right), \quad (1)$$

$$\varphi_t^{i,k,n} = \sum_{(i',k') \in V_t^{i,k,n}, k' \neq k} h_t^{i',k',n} P_t^{i',k'} + \sum_{(i',k') \in V_t^{i,k,n}, i' \neq i} h_t^{i',k',n} P_t^{i',k'}. \quad (2)$$

여기서,  $B_{i,k}$ 는 각 위성  $i$ 의 빔  $k$ 의 대역폭을 나타내고,  $L_{i,k}$ 는 위성  $i$ 의 빔  $k$ 가 커버하는 셀의 수를 나타내고,  $N_0$ 는 백색 가우시안 잡음을 나타낸다.  $h_t^{i,k,n}$ 은 위성  $i$ 의 빔  $k$ 에서 셀  $n$ 으로의 채널 이득을 나타내고,  $P_t^{i,k}$ 는 위성  $i$ 의 빔  $k$ 의 송신 파워를 나타낸다.  $\varphi_t^{i,k,n}$ 는 셀  $n$ 이 받는 간섭을 나타내며, 첫 번째 항은 빔 간 간섭을, 두 번째 항은 위성 간 간섭을 의미한다. 따라서, 실제 전송되는 트래픽 양은 시간 슬롯  $t$ 에서 해당 셀의 버퍼에 저장된 데이터 수요와 채널 용량 중 더 작은 값으로 결정된다.

본 논문의 궁극적인 목표는 다중 빔 저궤도 위성 시스템에서 시스템 처리량 및 에너지 효율을 최대화하는 것이다. 이를 위해 빔 호핑 스케줄링

$w_t^{i,k,n}$ , 빔 커버리지  $r_t^{i,k}$ 와 빔 송신 파워  $P_t^{i,k}$ 를 동시 최적화하는 문제  $\mathcal{P}$ 를 다음과 같이 정의한다.

$$\mathcal{P}: \max_{\{w_t^{i,k,n}\}, \{r_t^{i,k}\}, \{P_t^{i,k}\}} \sum_{i \in I, k \in K} \frac{1}{K} \left( \alpha \sum_{i \in I, k \in K} d_t^{i,k} + (1-\alpha) \sum_{i \in I, k \in K} \frac{d_t^{i,k}}{P_t^{i,k}} \right), \quad (3)$$

$$\text{s.t.} \quad \sum_{n \in \mathcal{C}_i} w_t^{i,k,n} = 1, \quad \forall i \in I, \forall k \in K, \forall t \in T, \quad (4)$$

$$\sum_{k=1}^K P_t^{i,k} \leq P_i, \quad \forall i \in I, \forall k \in K, \forall t \in T, \quad (5)$$

$$P_{\min} \leq P_t^{i,k} \leq P_{\max}, \quad \forall i \in I, \forall k \in K, \forall t \in T, \quad (6)$$

$$R_{\min} \leq r_t^{i,k} \leq R_{\max}, \quad \forall i \in I, \forall k \in K, \forall t \in T, \quad (7)$$

$$d_t^n \leq \delta_t^n, \quad \forall n \in N, \forall t \in T. \quad (8)$$

여기서,  $\alpha \in [0,1]$ 는 가중치 계수이고,  $d_t^{i,k}$ 는 위성  $i$ 의 빔  $k$ 가 전송한 총 트래픽 양으로, 시스템 처리량을 의미하며, 이를 송신 파워  $P_t^{i,k}$ 로 나눈 항은 에너지 효율을 나타낸다. 제약조건 (5)와 (6)은 빔의 송신 파워를 제한하며, (7)은 빔 커버리지 반경에 대한 제한을 나타내고, 마지막으로 (8)은 시간 슬롯  $t$ 에서 셀  $n$ 으로 전송되는 총 트래픽 양  $d_t^n$ 이 해당 시간의 버퍼에 저장된 초기 트래픽 수요량  $\delta_t^n$ 을 초과할 수 없는 제약조건을 나타낸다.

### III. 제안된 Modular DQN 알고리즘

본 장에서는 다중 빔 저궤도 위성 시스템에서, 시스템 처리량 및 에너지 효율을 최대화하기 위해 공유 인코더와 세부 작업별 출력 헤드(task-specific output heads)로 구성된 modular DQN 알고리즘을 제안한다. 제안된 modular DQN은 DQN 구조를 기반으로 하되, 에이전트가 통합된 네트워크를 통해 서브 액션들을 공동으로 학습하고, 각 서브 액션에 특화된 출력 헤드를 유지함으로써 보다 효율적인 의사결정을 가능하게 한다. 따라서 제안된 알고리즘은 빔 호핑 스케줄링, 빔 커버리지, 빔 송신 파워를 동시 최적화할 수 있도록 설계되었다. 이를 위해, 본 최적화 문제를 해결하는 데 사용되는 마르코프 결정 과정(Markov Decision Process, MDP)을 다음과 같이 정의한다.

- 1) **에이전트(Agent)**: 정지 궤도 위성을 에이전트로 간주하며, 각 저궤도 위성의 행동을 결정한다.
- 2) **상태(State)**: 에이전트는 환경과 상호작용하여 상태를 인식한다. 시간 슬롯  $t$ 에서 상태는 저궤도 위성들의 좌표, 각 셀의 버퍼에 저장된 트래픽 양, 시간 슬롯  $t$ 에서 모든 위성의 커버리지 정보로 구성된다.
- 3) **행동(action)**: 관측된 상태를 바탕으로, 에이전트는 빔 호핑 스케줄링, 빔 커버리지 반경, 빔 송신 파워를 동시에 결정한다.
- 4) **보상(reward)**: 시스템 처리량과 에너지 효율을 합산한 값으로 정의된다.

### IV. 시뮬레이션 결과

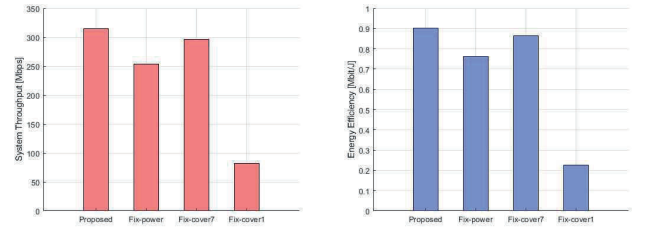
본 장에서는 제안된 알고리즘의 성능을 시뮬레이션을 통해 평가하고, 이를 기존 연구 기법들과 비교한다. 기존 연구 기법들은 본 논문에서 제안하는 시스템 환경에 맞추어 재구성하여 학습을 진행하였다. 시뮬레이션 환경은 98개의 지상 셀과 3개의 빔을 가지는 3개의 저궤도 위성이 동일한 궤도면을 따라 이동하는 설정을 기반으로 하며, 총 25시간 스텝동안 각 저궤도 위성의 빔이 전체 셀을 서비스하도록 학습하였다.

그림 2는 제안된 기법의 성능을 (a)시스템 처리량 및 (b)에너지 효율 측면에서, 고정 파워를 사용하는 기법(Fix-power) [5] 및 고정 커버리지를 사용하는 기법(Fix-cover1, Fix-cover7) [6]과 비교한 결과를 나타낸다. Fix-cover1 기법은 항상 하나의 셀을 커버하므로 시스템 처리량과 에

너지 효율 측면에서 모두 가장 낮은 성능을 보인다. Fix-power 기법은 지상 셀의 트래픽 수요에 관계없이 항상 동일한 송신 파워가 할당되므로 제안된 기법보다 낮은 성능을 확인할 수 있다. 한편, Fix-cover7 기법은 항상 7개의 셀을 서비스하기 때문에 다른 기법들에 비해 많은 간섭과 비효율적인 파워 소모가 발생할 수 있다. 따라서 빔 커버리지 반경과 빔 송신 파워를 동적으로 조절하는 제안된 기법이 기존 연구들에 비해 높은 에너지 효율과 시스템 처리량을 보여준다.

### V. 결론

본 논문에서는 정지궤도 위성이 제어하는 다중 빔 저궤도 위성 시스템에서 시스템 처리량 및 에너지 효율을 최대화하기 위해 modular DQN 기반의 알고리즘을 제안하였다. 제안된 알고리즘은 동적으로 변하는 시스템 환경에서 빔 호핑 스케줄링, 빔 커버리지 반경, 빔 송신 파워를 동시 최적화함으로써 효율적인 빔 자원 할당을 수행한다. 시뮬레이션 결과, 제안된 기법이 기존 연구들과 비교하여 시스템 처리량과 에너지 효율 모두에서 우수한 성능을 보이는 것을 확인하였다.



(a) 시스템 처리량 비교

(b) 에너지 효율 비교

그림 2. 성능 비교

### ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-학·석사연계ICT핵심인재양성의 지원(IITP-2025-RS-2022-00156326, 50%)과 2022년도 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임 (KRIT-CT-22-047, 우주계측 지능통신망 특화연구실)

### 참고 문헌

- [1] S. H. Chae, H. Lim, H. Lee, and B. C. Jung, "Performance analysis of dense low earth orbit satellite communication networks with stochastic geometry," *J. Commun. and Networks*, vol. 25, no. 2, pp. 208-221, Apr. 2023.
- [2] C. Lee, I. Bang, T. Kim, H. Lee, B. C. Jung and S. H. Chae, "Multi-agent deep reinforcement learning based handover strategy for LEO satellite networks," *IEEE Commun. Lett.*, vol. 29, no. 5, pp. 1117-1121, May 2025.
- [3] G. Kim, S. Lee, H. Lim, B. C. Jung, and S. H. Chae, "Coverage probability analysis of LEO satellite communication systems with directional beamforming," in *Proc. of the 14th Int. Conf. on Ubiquitous and Future Netw. (ICUFN)*, Paris, France, Jul. 2023.
- [4] M. Meng, B. Hu, S. Chen and S. Kang, "Joint beamforming and dynamic beam hopping based on MAPPO for LEO satellite communication system," *IEEE Wireless Commun. Lett.*, vol. 14, no. 5, pp. 1461-1465, May 2025.
- [5] G. Xu, F. Tan, Y. Ran, Y. Zhao and J. Luo, "Joint beam-hopping scheduling and coverage control in multibeam satellite systems," *IEEE Wireless Commun. Lett.*, vol. 12, no. 2, pp. 267-271, Feb. 2023.
- [6] Y. Ran *et al.*, "Towards beam hopping and power allocation in multi-beam satellite systems with parameterized reinforcement learning," *IEEE Trans. Vehi. Tech.*, vol. 73, no. 9, pp. 14050-14055, Sep. 2024.
- [7] S. Zhang, A. Liu, C. Han, X. Ding and X. Liang, "A network-flows based satellite handover strategy for LEO satellite networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 12, pp. 2669-2673, Dec. 2021.