IAB 네트워크에서 분산형 심층강화학습 기반 자원 할당과 테더드 UAV 위치 제어 공동 최적화

이예린, 이호원
아주대학교
yerin1205@ajou.ac.kr, howon@ajou.ac.kr

Joint Optimization of Resource Allocation and Tethered UAV Positioning in IAB Networks Based on Distributed Deep Reinforcement Learning

Yerin Lee, Howon Lee
Ajou University

요 약

본 논문에서는 tethered unmanned aerial vehicle base station(TUBS) 지원 integrated access and backhaul(IAB) 네트워크에서 합 전송률 최대화를 위한 분산형 심층강화학습 기반 자원 할당 및 위치 제어 공동 최적화 기법을 제안한다. TUBS는 테더를 통한 지속적인 전력 공급으로 배터리 제약 없이 3차원 위치 조정이 가능하며, IAB 기술로 채널 대역폭을 통합적으로 사용할 수 있다. 제안 방안은 IAB를 구성하는 각 기지국을 독립적인 에이전트로 설계하고 분산형 DDQN 학습과 공동 보상 함수를 활용하여 협력적 자원 최적화를 달성한다. 시뮬레이션 결과, 제안 방안은 기존 방안 대비 우수한 성능을 달성하며 위치 제어와 분산 학습의 효율성을 입증한다.

Ⅰ. 서 론

차세대 무선 네트워크 환경에서는 시공간적으로 급변하는 사용자 필요 트래픽 요구를 유연하게 수용하기 위한 지능형 공중 기지국 운용 기술이 중요한 과제로 부상하고 있다[1]-[3]. 이러한 요구에 대응하기 위해 지상으로부터 지속적인 전력 공급이 가능한 tethered unmanned aerial vehicle base station(TUBS)이 공중 기지국으로서의 대안으로 주목받고 있다. TUBS는 배터리 제약에서 자유로워 운용 안정성을 확보할 수 있다[2],[3]. 본 연구에서는 TUBS를 공중 기지국으로 활용한 integrated access and backhaul(IAB) 아키텍처 기반의 네트워크 구조를 제안한다. 이를 통해 백홀과 액세스 링크를 단일 주파수 자원에서 효율적으로 운영할 수 있다. 본 연구에서는 테더 길이로 인한 TUBS의 제한된 이동 범위 내에서도 합전송률 최대화를 위해 주파수 자원 할당과 3차원 위치 제어를 동시에 최적화하는 기법을 제안한다. 특히, 분산 에이전트 간 협력 학습을 위해 통합된 공유 보상을 설계하여 네트워크 전체 성능을 향상시키고 사용자 위치 변화에 효과적으로 대응할 수 있는 방안을 제안한다.

Ⅱ. 본론

제안 방안은 MBS와 TUBS를 독립적인 에이전트로 설계하고 Markov Decision Process(MDP)로 정형화하여 각 에이전트가 분산적으로 학습을 수행하되 공동 보상을 통해 협력적으로 최적화를 달성하는 구조이다. 에이전트 MBS i의 상태(s_i^M)는 채널 할당 정보와 전송 전력으로 정의하며, 행동(a_i^M)으로는 채널 할당 변경과 전력 조절을 수행한다. 에이전트 TUBS j의 상태(s_j^T)는 채널 할당 정보, 전송 전력, 구면 좌표 기반의 TUBS이 3차원 위치 정보로 정의하며, 행동(a_j^T)은 채널 할당 변경, 전력 조절과 함께 TUBS의 3차원 위치 제어를 포함한다. 각 에이전트 간 협력 학습을 위한 공유 보상 함수(r_S)는 다음과 같이 정의한다.

$$r_S = \sum(i \in M_M) R_i^{M-SR} + \sum(j \in M_T) R_j^{T-SR} \quad (1)$$

여기서 R_i^{M-SR}과 R_j^{T-SR}은 각각 MBS와 TUBS의 합 전송률을 의미하며 M_M과 M_T는 각각 MBS와 TUBS의 집합을 나타낸다. 공유 보상 함수를 통해 모든 에이전트가 네트워크 전체 성능 향상에 기여하도록 유도한다.

상기 정의된 MDP 하에서 최적 정책 학습을 위해 분산형 DDQN을 도입한다. DDQN은 과대평가 문제를 해결하기 위해 온라인 네트워크와 타깃 네트워크를 분리하여 운용한다. 온라인 네트워크는 행동 선택에, 타깃 네트워크는 선택된 행동의 평가에 사용되어 학습 안정성이 향상된다. 분산형 접근법은 중앙집중형 방식 대비 계산 복잡도가 에이전트 수에 따라 선형적으로만 증가하여 네트워크 규모 확장에 효과적으로 대응이 가능하다.

Ⅲ. 시뮬레이션 결과 및 결론

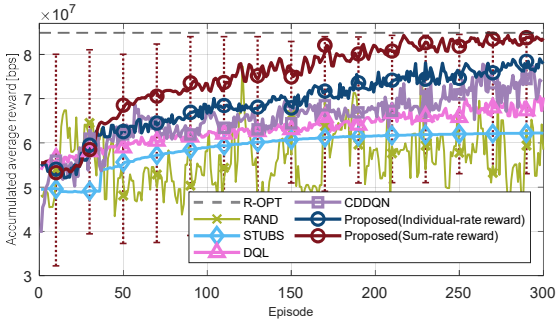

그림 1. 에피소드에 따른 제안 방안 및 비교 방안의 누적 보상

그림 1은 고밀도 도심 환경에서의 3-에이전트 시나리오를 기준으로, 제안 방안과 비교 방안의 누적 성능을 나타낸다. 제안 방안은 약 250 에피소드 이내에 Reward-Optimal(R-OPT)에 근접한 수준까지 도달하며 안정적인 성능을 유지한다. 반면, 중앙집중형 접근법인 CDDQN은 높은 계산 오버헤드로 인해 분산형 구조 대비 수렴 속도가 현저히 저하된다. 또한, Q-러닝 방식인 DQL은 상태 공간 이산화 및 확장성의 제약으로 인해 낮은 수준에서 수렴하고, 위치 제어 기능이 없는 STUBS는 제안 방안 대비 34.58%의 가장 큰 성능 저하를 보여주며, 공중 기지국의 위치 제어가 IAB 네트워크 성능의 핵심 요소임을 시사한다. 제안한 분산형 DDQN 기반 학습 프레임워크는 TUBS 기반 IAB 네트워크에서 자원 할당과 위치 제어를 통합적으로 수행할 수 있는 현실적인 해법으로 기능하며, 동적인 네트워크 환경에서도 효과적으로 대응 가능한 확장성 높은 지능형 제어 구조로 활용 가능하다.

ACKNOWLEDGMENT

이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원(No. RS-2024-00396992, 저궤도 위성통신 핵심 기술 기반 큐브위성 개발)과 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원(No. 2022-0-00704, 초고속 이동체 지원을 위한 3D-NET 핵심 기술 개발)과 2025년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(RS-2025-00563401, 3차원 공간에서 에너지 효율적 멀티레벨 AI-RAN 구현을 위한 AI-for/and-RAN 핵심 원천기술 연구)을 받아 수행된 연구임.

참 고 문 헌

[1] H. Lee et al., "Towards 6G hyper-connectivity: Vision, challenges, and key enabling technologies," in Journal of Communications and Networks, vol. 25, no. 3, pp. 344-354, June 2023.
[2] B. E. Y. Belmekki and M. -S. Alouini, "Unleashing the Potential of Networked Tethered Flying Platforms: Prospects, Challenges, and Applications," in IEEE Open Journal of Vehicular Technology, vol. 3, pp. 278-320, 2022.
[3] M. Polese et al., "Integrated Access and Backhaul in 5G mmWave Networks: Potential and Challenges," in IEEE Communications Magazine, vol. 58, no. 3, pp. 62-68, March 2020.