

트래픽 예측 및 안전 심층 강화학습 기반 슬립 모드 제어를 통한 이기종 네트워크에서 소형셀 에너지-QoS 트레이드오프 최적화 기법

조아림, 윤재영, 이호원

아주대학교

{jal9160, dbswo0617, howon}@ajou.ac.kr

Small Cell Energy-QoS Trade-Off Optimization Techniques in Heterogeneous Networks with Traffic Prediction and Safe Deep Reinforcement Learning-Based Sleep Mode Control

Arim Cho, Jaeyoung Yun, Howon Lee

Ajou Univ.

요약

본 논문에서는 이기종 네트워크에서 큰 비중을 차지하는 소형셀 기지국의 에너지 소비를 줄이기 위해 안전 심층 강화학습을 기반으로 소형셀의 다단계 슬립 모드 제어 기법을 제안한다. 그러나, 기지국이 장시간 슬립 모드를 유지하면 이에 따른 서비스 품질(Quality of service, QoS) 저하가 발생하기 때문에, 에너지 소모와 QoS 간 트레이드오프를 최적화하는 것을 목표로 한다. 이를 위해, long short-term memory(LSTM) 기반 트래픽 예측을 활용하여 과도한 슬립 모드 지속으로 인한 QoS 저하를 선제적으로 방지하고, 안전 심층 강화학습을 활용하여 허용되지 않은 모드 전환을 제약함으로써 이기종 네트워크의 안정적인 운영을 보장한다.

I. 서론

다양한 모바일 애플리케이션의 급속한 발전으로 무선 네트워크에서 모바일 데이터 트래픽이 폭발적으로 증가하고 있다. 그러나 소형셀 기지국의 밀집 배치로 인해 네트워크 전체의 에너지 소비가 무시할 수 없는 수준으로 커지고 있다 [1],[2]. 소형셀 에너지 절감을 위한 가장 효과적인 접근 방식 중 하나가 슬립 모드이며, 슬립 모드 상태에서도 커버리지와 서비스 품질(Quality of service, QoS)이 유지되도록 인접 활성 기지국의 전송 전력을 소폭 상향 조정하여 사용자를 지원할 수 있다. 이처럼 소형셀을 끄거나 에너지 절약 모드를 적용하면서도 QoS를 보장하는 방식이 에너지 효율을 개선하는 대표적인 방법이다 [3]. 따라서, 본 논문에서는 트래픽 예측을 활용한 안전 심층 강화학습(Safe deep reinforcement learning, SDRL) 기반 슬립 모드 제어를 통해 에너지와 QoS 간의 트레이드오프 최적화 기법을 제안한다.

II. 트래픽 예측 기반 안전 심층 강화학습을 활용한 슬립 모드 제어 기법

본 논문에서는 이기종 네트워크에서 트래픽 예측을 활용한 안전 심층 강화학습 기반 기지국의 다단계 슬립 모드 제어 기법을 제안한다. Long short-term memory(LSTM) 기반 예측된 트래픽 정보를 바탕으로 안전 심층 강화학습을 활용하여 기지국의 슬립 모드 제어를 수행한다. 이를 통해 과도한 슬립 모드 지속으로 인한 QoS 저하를 선제적으로 방지하고, 허용되지 않은 모드 전환을 제약함으로써 안정적인 운영을 보장하고자 하였다. 또한, 에너지 소모를 줄이기 위해 슬립 모드로 전환하면서 발생하는 지연으로 QoS에 영향이 생길 수 있어 전력 소모와 QoS 간의 트레이드오프 관계를 고려하는 것이 필수적이다. 이러한 문제를 DQN으로 풀기 위해 Markov decision process(MDP)로 정의한다.

- Agent: 네트워크에 배치된 모든 소형셀, $SBS(n) \in \{1, \dots, S\}$.
- State: 이전 타임스텝의 전환 시간, 소모된 전력량, 현재 타임스텝에서 선택된 모드 전환에 필요한 남은 타임스텝 수, 현재 타임스텝의 트래픽, $2\tau+1$ 길이의 예측된 트래픽 정보로 구성된다.

$$s_n^t = \{T_n^{t-1}, P_n^{t-1}, NT_n^t, \rho_n^t, \rho_n^{t-\tau:t+\tau}\}.$$

- Action: $a_n^t = \{active, micro\ sleep, light\ sleep, deep\ sleep\}$.
- Reward: 에너지와 QoS 간의 트레이드오프를 고려하기 위해 가중치 α 를 사용하며, 모드 별 전환 시간에 대한 보상(r_{td})과 에너지 절감에 따른 보상(r_{ps})으로 구성된다.

$$r_n^t = \alpha \times r_{ps} + (1 - \alpha) \times r_{td}, \quad (1)$$

$$r_{ps} = \sum_{n=1}^S \frac{P_{active}^{max} - P_{n,x}}{P_{active}^{max}}, r_{td} = \sum_{n=1}^S \frac{T_{deep\ sleep} - T_{n,x}}{T_{deep\ sleep}}, x \in action\ set.$$

III. 시뮬레이션 결과 및 결론

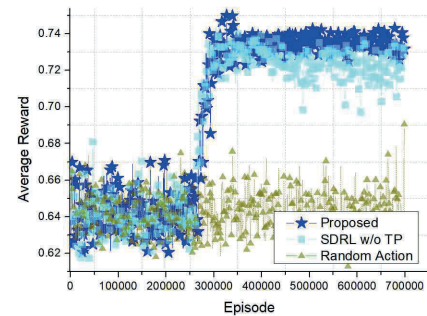


그림 1. Average reward vs. episode.

본 논문에서는 제안방안과 트래픽 예측을 사용하지 않고 안전 심층 강화학습을 수행하는 방안(SDRL w/o TP)과 랜덤한 액션을 수행하는 방안(Random Action)을 비교하였다. 시뮬레이션 결과에서 제안방안이 가장 높은 값의 리워드로 수렴하였으며, 이를 통해 제안방안이 에너지와 QoS 간의 트레이드오프를 고려한 슬립 모드 최적화를 수행함을 확인하였다.

ACKNOWLEDGMENT

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

참고문헌

- [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] S. Wu, R. Yin and C. Wu, "Heterogeneity-Aware Energy Saving and Energy Efficiency Optimization in Dense Small Cell Networks," in *IEEE Access*, vol. 8, pp. 178670–178684, 2020.
- [3] C. Liu, B. Natarajan and H. Xia, "Small Cell Base Station Sleep Strategies for Energy Efficiency," in *IEEE Transactions on Vehicular Technology*, vol. 65, no. 3, pp. 1652–1661, March 2016.