

이종 네트워크에서의 에너지 절감 및 QoS 보장 강화학습 에이전트 구현

이광훈, 김태윤, 배찬빈, 홍준규, 구희모, 백상현
*고려대학교

{leegh963, kimtyoun123, bin6050, clickhjk3267, hmku0415, shpack}@korea.ac.kr

Implementation of Reinforcement Learning Agent for Energy Saving and QoS Provisioning using Network Simulator

Lee Kwanghoun, Kim Taeyun, Bae Chanbin, Hong Junkyu, Ku Hwimo, Sangheon Pack
*Korea Univ.

요 약

현대 네트워크는 급증하는 트래픽 수요로 인해 단일 접속망 기반 운용에 효율성이 저하되는 한계에 직면하였으며, 이를 해결하기 위한 이종 네트워크 구조가 주목받고 있다. 그러나 이러한 이종 네트워크에서 기존의 규칙 기반 제어는 동적인 환경 변화에 유연하게 대응하기 어렵다. 본 논문은 NetSimGym 프레임워크 기반의 강화학습 에이전트를 제안하여, 서비스 품질(QoS) 준수와 기지국 에너지 절감을 동시에 달성하는 지능형 트래픽 오프로딩 기법을 연구하였다. 시뮬레이션 결과, 부하 변동에 따른 에이전트의 적응적 판단 능력을 입증하였으며, 이는 향후 지능형 망 운용을 위한 핵심적인 기술적 토대가 될 것이다.

I. 서 론

현대 모바일 네트워크는 5G 고도화와 차세대 6G로의 진화를 거치며, 서로 다른 특성을 가진 다중 접속 기술(Multi-access)이 통합된 이종 네트워크(Heterogeneous network) 구조로 발전하고 있다. 이러한 환경에서는 폭증하는 데이터 트래픽을 효율적으로 분산시키고 사용자에게 일정한 서비스 품질(Quality of Service)을 제공하기 위해, 가용 자원을 실시간으로 파악하여 적절한 망으로 데이터를 전달하는 트래픽 오프로딩(Traffic offloading) 기술의 중요성이 증대되고 있다.

그러나 기존의 규칙 기반(Rule-based) 오프로딩 방식은 사전에 정의된 임계치나 고정된 알고리즘에 의존하기 때문에, 동적으로 급변하는 채널 품질이나 불규칙한 트래픽 패턴에 유연하게 대응하기 어렵다는 한계가 있다. 특히, 사용자의 QoS를 보장하면서도 네트워크 운용 비용과 직결되는 기지국의 전력 소모를 줄이는 문제는 서로 상충하는 관계(Trade-off)를 가지므로, 복잡한 변수들을 동시에 고려할 수 있는 지능형 제어 알고리즘이 필수적으로 요구된다.

본 논문에서는 NetSimGym [1] 기반의 네트워크 시뮬레이션 환경을 구축하고, 다중 접속 환경에서 사용자의 QoS를 준수함과 동시에 기지국의 에너지 효율을 극대화하는 지능형 강화학습 에이전트를 제안한다. 제안하는 에이전트는 네트워크 상태 정보를 실시간으로 수집하여 이종 네트워크 간의 최적 트래픽 배분 정책을 도출하며, 트래픽이 적은 시간대에 기지국을 비활성화 상태로 전환하여 에너지 절감을 추구한다. 본 연구의 성과는 시뮬레이션을 통해 망 부하 변동에 따른 에이전트의 적응적 판단 능력을 입증함으로써, 향후 지능적 네트워크 운용을 위한 기술적 토대를 제공하고자 한다.

II. 본론

1) NetSimGym 개요

NetSimGym은 네트워크 시뮬레이션 도구인 NetSim과 OpenAI의 강화학습 인터페이스인 Gymnasium을 연동하는 래퍼(Wrapper) 프레임워크이다. 본 연구에서는 NetSim 시뮬레이터를 통해 네트워크 환경을 모사하고, 이를 강화학습 에이전트에 실시간 상태 정보를 제공하는 환경으로 활용한다. 에이전트는 NetSim으로부터 전달받은 상태(State) 및 관측(Observation) 데이터를 바탕으로 최적의 행동(Action)을 결정하며, 이는 다시 시뮬레이터에 반영되어 네트워크의 동적 변화를 제어한다.

2) NetSim 시뮬레이션 환경

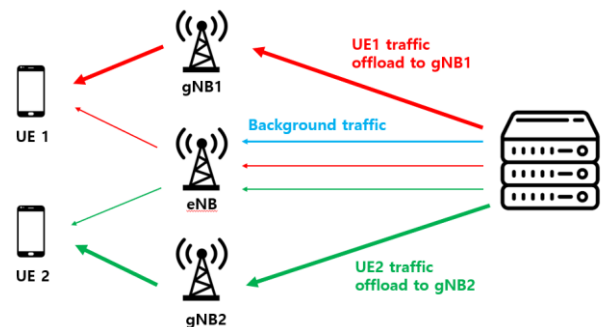


그림 1. NetSim 시뮬레이션 환경

본 연구의 실험을 위해 그림 1과 같은 네트워크 시뮬레이션 환경을 구축하였다.

- 망 구성: 이기종 네트워크 설정을 위해 gNB와 eNB가 동시에 존재하는 5G NSA 환경 설정.

- 트래픽 시나리오: 한 에피소드의 시뮬레이션 시간은 30 초로 설정하였다. UE1 의 트래픽은 gNB1 과 eNB 로, UE2 의 트래픽은 gNB2 와 eNB 로 각각 분산되어 전송된다.
- 부하 변동 모사: 실제 환경의 동적인 부하 변화를 재현하기 위해 3~15 초 및 20~30 초 구간에 eNB 를 점유하는 배경 트래픽 (Background traffic)을 인위적으로 발생시켰다.

A. 상태 공간 (State space)

강화학습 에이전트는 현재 네트워크의 상태를 인지할 수 있도록 eNB2 를 통해 UE2 로 전달되는 패킷의 일부를 샘플링하여 측정된 종단 간 지연 (End to end delay)을 주요 observation (관측) 정보로 NetSim 으로부터 전달받고 해당 정보를 QoS delay 를 기준으로 하여 4 개의 구간 (Bucket)으로 이산화하여 state 를 정의한다.

- State1: 종단 간 지연이 QoS 보다 크게 작은 경우
- State2: 종단 간 지연이 QoS 보다 조금 작은 경우
- State3: 종단 간 지연이 QoS 를 조금 초과한 경우
- State4: 종단 간 지연이 QoS 를 크게 초과한 경우

B. 행동 공간 (Action space)

에이전트는 총 8 가지의 action 을 취할 수 있으며, 이는 gNB1 과 eNB, gNB2 와 eNB 사이의 트래픽 분배 비율 조합으로 구성된다. 각 action 은 특정 기지국의 전력 소모 및 전체 망의 부하량에 직접적인 영향을 미친다. 이 중 action 1,5 는 gNB2 와 eNB 를 통해 나눠서 전달되는 UE2 의 트래픽을 모두 eNB 로 할당하고 gNB2 는 비활성화 하여 에너지 절감 효과를 보는 action 이다. Action 6,7 은 gNB2 와 eNB 를 통해 나눠서 전달되는 UE2 의 트래픽 중 가장 많은 부분을 gNB2 를 통해서 전달하는 action 이다. 이를 통해 백그라운드 트래픽이 많아지는 경우 eNB 에서의 처리량을 줄여 QoS 를 만족할 수 있게 한다.

C. 보상 함수 (Reward)

에이전트의 학습은 에너지 소비량과 QoS 위반 페널티를 결합한 가중치 기반 보상 함수를 통해 이루어진다.

에너지 소비량 향에 대해서 기지국의 전력 사용량을 최소화하기 위해 action 1 또는 5 와 같이 gNB2 기지국을 비활성화하는 행동을 선택할 경우 다른 action 들 보다 더 적은 음의 보상을 부여하여 에너지 절감을 유도한다.

QoS 위반 페널티에 대해서는 지연 시간이 QoS 기준을 초과할 경우 강력한 페널티를 부여하되, 종단간 지연이 QoS 를 위반한 상태에서 지연 시간이 이전 단계보다 감소 추세를 보일 경우 양의 보상을 주어 QoS 회복을 학습하게 한다.

3) 학습 결과

강화학습 에이전트의 행동 선택 및 에너지 소비 추이를 분석한 결과, 에이전트가 네트워크 상태에 따라 적응적으로 판단함을 입증하였다. 그림 2 는 시뮬레이션 경과에 따른 에이전트의 action 선택 및 기지국들의 에너지 사용량을 나타낸다. 에이전트는 배경 트래픽이 없는 부하량이 작은 구간 (0~3s, 15~20s)에서는 에너지 절감을 위해 트래픽을 전부 gNB2 로 할당하고 eNB 를 비활성화 하는 action 1

또는 5 를 선택하며 이에 따라 에너지 소비량이 다른 구간보다 더 적은 것을 확인할 수 있다. 그 외의 배경 트래픽이 큰 구간 (3~15s, 20~30s)에서는 eNB 로 향하는 트래픽이 증가함에 따라 처리가 지연되므로 이를 방지하고 QoS 를 만족하기 위해 gNB2 를 활성화하고 트래픽의 많은 부분을 gNB2 에서 처리하는 action 6,7 을 선택하는 것을 확인할 수 있다.

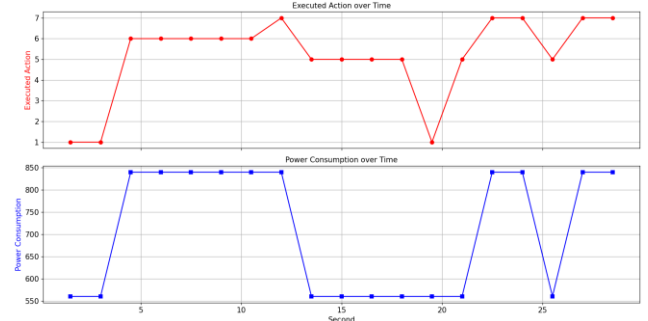


그림 2. RL agent 의 시간대 별 action 선택 (위) 및 전력 사용량 (아래)

III. 결론

본 논문에서는 NetSim 시뮬레이션을 바탕으로 다중 접속망 간의 트래픽을 동적으로 제어하여 QoS 요구사항을 충족하고 에너지 효율을 극대화하는 강화학습 기반 지능제어 에이전트를 개발하였다.

검증 시뮬레이션 결과, 제안된 에이전트는 배경 트래픽 급증으로 인한 지연 시간 상승 시 접속망 간 트래픽 전달 비율을 조정하여 QoS 요구사항을 준수하는 능력을 보여주었다. 동시에, 트래픽 부하가 적은 구간에서는 불필요한 기지국을 비활성화하는 행동을 선택함으로써 전반적인 시스템 에너지 소모를 효과적으로 절감하는 성과를 거두었다.

향후에는 현재 시나리오보다 더 복잡한 상황에서도 적절하게 동작 할 수 있도록 에이전트의 state 파라미터 다양화 및 action 의 고도화 등을 구현할 예정이다.

ACKNOWLEDGMENT

이 논문은 2025 년도 정부(과학기술정보통신부)의 재원으로 한 국연구재단 중견연구 (No. RS-2024-00341965) 및 IITP 네트 워크 연구센터 (NRC) (No. RS-2024-00398948)의 지원을 받아 수행된 연구임

참 고 문 헌

[1] I. Cha, S. Shin, K. Kim, Y. Ahn, T. Kim, and S. Pack, "NetSimGym: The gymnasium for reinforcement learning in networking research," in *Proc. 2025 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, pp. 339-344, Jun. 2025