

# 심층강화학습 기반 에너지 효율적 전송 전력 할당 기법

문지훈, 이안호, 심병효  
서울대학교

{jymoon, ahlee, bshim}@islab.snu.ac.kr

## Deep Reinforcement Learning-based Energy-Efficient Transmission Power Allocation

Jihoon Moon, Anho Lee, and Byonghyo Shim  
Seoul National University

### 요약

본 논문에서는 최근 늘어나는 데이터 요구량을 만족하기 위해 도입한 고밀도 네트워크 환경에서의 에너지 효율적 전송 전력 할당 기법에 대해 논의한다. 높은 계산량과 복잡도를 줄이기 위해 딥러닝 기법 중 학습 데이터가 따로 필요하지 않은 심층강화학습을 네트워크 환경에 활용해 고밀도 네트워크의 높은 에너지 소모를 효과적으로 줄일 수 있다. 제안하는 기법은 Actor-Critic Network를 활용해 네트워크의 에너지 소모량을 최소화하면서도 각 사용자의 요구 조건을 만족시킬 수 있음을 시뮬레이션을 통해 보여준다.

### I. 서론

본 논문에서는 최근 늘어나는 데이터 요구량을 달성하기 위해 도입한 고밀도 네트워크 환경에서 전송 전력 할당 방식에 대해 논의한다[1]. 하지만 고밀도 네트워크에서는 전력 할당 문제를 해결하는 과정이 높은 계산량과 복잡도로 인해 최적화가 어렵다. 심층강화학습은 학습용 데이터가 별도로 필요하지 않으면서도 계산량과 복잡도를 낮춰 효과적으로 네트워크 문제들을 해결할 수 있어 최근 몇 년 동안 많은 무선통신 환경에 적용되어 왔다[2]. 따라서 본 연구에서는 고밀도 네트워크에서 심층강화학습을 활용해 전송 전력 할당 문제를 효율적으로 해결하고자 한다.

### II. 본론

고밀도 네트워크 기지국  $M$ 개와  $K$ 명의 사용자가 있는 환경을 가정한다. 모든 기지국들은 중앙 처리 장치에 연결되어 있어 각 채널 정보를 전달하고 전력 할당 방식을 전달받는다. 기지국 번호를  $m \in \{1, 2, \dots, M\}$ 으로, 사용자 번호를  $k \in \{1, 2, \dots, K\}$ 으로 정한다. 일반적인 셀룰러 네트워크와 다르게 기지국 여러 개가 동시에 협력적으로 각 사용자에게 데이터를 전송할 수 있다고 가정한다. 이때, 각 사용자 별 데이터 수신 속도는 다음과 같다.

$$R_k = \log_2 \left( 1 + \frac{\sum_{m=1}^M \rho_{m,k} h_{m,k}}{\sum_{m=1}^M \sum_{j \neq k} \rho_{m,j} h_{m,k} + \sigma_n^2} \right).$$

여기서  $h_{m,k}$ 는 기지국  $m$ 과 사용자  $k$  사이의 채널이고,  $\rho_{m,k}$ 는 기지국  $m$ 이 사용자  $k$ 에게 전송하는 전력이다. 기지국 당 최대 전력은  $P_{\max}$ 로 제한되어 있다. 기지국  $m$ 과 연결된 사용자의 집합을  $U_m$ 이라 하면,

$$\sum_{k \in U_m} \rho_{m,k} \leq P_{\max},$$

이다. 따라서, 풀어야 하는 에너지 소모량 최소화 문제는 다음과 같다.

$$P1: \min_{\{\rho_{m,k}\}} \sum_{m=1}^M \sum_{k=1}^K P_{m,k}$$

$$\text{s. t.} \quad \sum_{k \in U_m} \rho_{m,k} \leq P_{\max}, \\ R_k \leq P_{\max} \quad k = 1, 2, \dots, K,$$

심층강화학습 위에 제시된 전력 할당 문제를 해결하기 위해 Actor-Critic 강화학습 기법을 적용한다[3]. Actor-Critic은 특정 상태(state,  $s$ )와 행동(action,  $a$ ), 보상(reward,  $r$ )를 토대로 심층 신경망으로 구성된 actor network와 critic network를 학습시키는 방식이다. 상태를 채널 상태 정보(CSI)와 이전 전력 할당 상황, 행동을 현재의 전력 할당, 보상을 그 결과로 나온 전력 소모량으로 하면, 어떤 전력 수준을 할당했을 때 에너지 소모량이 최소화 되면서도 각 사용자의 데이터 요구량을 만족시

킬 수 있는지 학습할 수 있다. 위 값들을 메모리에 저장해 진행하면서 훈련 단계에서 일정 크기만큼 불러와 두 개의 심층 신경망을 훈련시킨다. 이때 최적화해야 하는 값(loss)는 actor와 critic에 대해 각각

$$L_{\text{actor}} = -Q(s_t, a_t),$$

$$L_{\text{critic}} = \left( r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \right)^2,$$

이다.

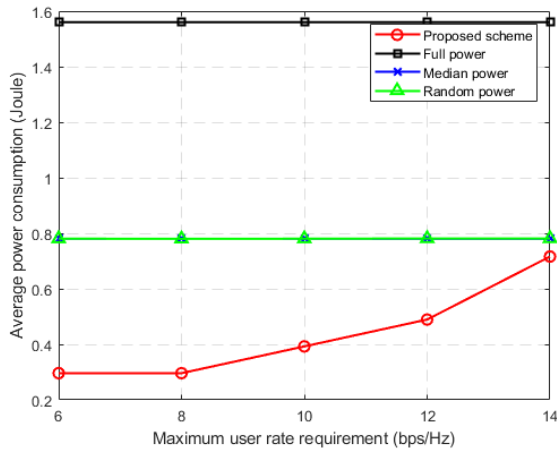


그림 1. 사용자의 데이터 전송률 요구 조건 변화에 따른 평균 전력 소모량 그래프

시뮬레이션 결과는 그림 1과 같다. 제안하는 심층강화학습 기반 기법이 여러 가지 사용자의 데이터 전송률 요구 조건에 맞춰 다른 전력 할당 방식에 비해 평균적으로 낮은 전력을 소모하는 것을 볼 수 있다.

### III. 결론

고밀도 네트워크 환경에서 심층강화학습을 활용하여 전송 전력 할당을 최적화할 수 있는 것을 보았다. 한번 심층신경망이 훈련되고 나면 적용 시 계산량이 일반 최적화 알고리즘에 비해 적기 때문에 시간에 계산 효율적으로 최적화할 수 있다. 이후 사용자가 많이 움직이는 상황이나 더 많은 사용자가 존재하는 환경에서 적용 해보는 것이 미래 과제이다.

### ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (2020R1A2C2102198)

### 참 고 문 헌

- [1] O. Naparstek and K. Cohen, "Deep multi-user reinforcement learning for dynamic spectrum access in multichannel wireless networks," in IEEE Transactions on

Wireless Communications, vol. 18, no. 1, pp. 310-323, 2018.

- [2] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-dense networks: A survey," IEEE Communications Surveys and Tutorials, vol. 18, no. 4, pp. 2522-2545, 2016.

- [3] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in International conference on machine learning, 2014, pp. 387-395.