

# 다중 에이전트 강화학습 기반 셀룰러 네트워크에서의 상향링크 전력 제어 기법

박준영, 유정엽, 고영채  
고려대학교

megaduke7@korea.ac.kr, youj0305@korea.ac.kr, koyc@korea.ac.kr

## Multi-Agent Reinforcement Learning-based Uplink Power Control Technique in Cellular Networks

Joon-Young Park, Jeong-Yeop You, Young-Chai Ko  
Korea Univ.

### 요약

본 논문은 셀룰러 네트워크의 상향링크 간섭 문제를 완화하고 셀 용량을 개선하기 위한 다중 에이전트 강화학습 기반의 전력 제어 기법을 제안한다. 사용자 단말 간의 협력적 학습을 통해 power racing 문제를 효과적으로 해결하며, 시뮬레이션을 통해 제안 기법이 스펙트럼 효율을 향상시키고 통신 단절 확률을 감소시키는 것을 검증하였다. 이에 더해, 다중 에이전트 강화학습에서 협력적인 학습을 위한 구조체가 필요하며, 이가 매우 효율적임을 증명한다.

### I. 서론

최근 5G NR 시스템의 확산으로 인해 다중 셀 (multi-cell), 다중 사용자 (multi-UE) 환경에서의 통신 트래픽이 급격히 증가하고 있다. 이와 더불어 높은 사용자 (UE, User Equipment) 밀도로 인해 셀 내 혼잡도가 심화되면서, 특히 상향링크 (UL, Uplink) 간섭 문제로 인한 성능 저하가 심각한 수준에 이르고 있다. 전통적인 상향링크 전력 제어 (ULPC, Uplink Power Control) 방식은 수학적 접근 및 이를 기반으로 한 최적화 기법을 통해 해당 문제를 해결하려고 하였으나, 다수의 UE를 고려하는 동시에 더욱 네트워크 환경이 복잡해진 현재의 이동 통신 네트워크에서는 이를 적용하기가 불가능하다. ULPC를 적용하기 위해 상정할 수 있는 환경으로, 다중 셀 환경에서, 특정 UE가 자신의 신호 대비 잡음비 (SINR, Signal-to-Interference-plus-Noise Ratio)를 개선하기 위해 전력을 증가 시킴에 따라, 다른 셀의 UE에게 추가적인 간섭을 초래하고, 이에 따라 다른 셀의 UE 또한 기존의 SINR을 회복하기 위해 자신의 전력을 높이게 된다. 이로 인해, 각 UE의 전력은 계속 증가하는 동시에 상호 간섭이 증가 하기에 이는 곧 UL 셀 용량의 저하로 이어지며, 또한 하향링크(DL, Downlink)에도 악영향을 주어 전반적인 네트워크의 성능을 크게 하락시키는 악순환, 즉 Power racing 현상이 발생한다. 이러한 현상은 특히 셀 경계 지역 (Cell-edge)의 사용자들에게 심각한 서비스 품질 저하와, 통신 망에서 제외되는 아웃지 (Outage) 발생 위험을 가중시킨다. 이를 해결 하기 위해, 최근에는 복잡한 환경에서도 높은 성능을 제공할 수 있는 인공지능 (AI, Artificial Intelligence) 및 머신러닝 (ML, Machine Learning)을 기반으로 한 접근법이 대안으로 부상하고 있다 [1]-[2].

본 연구는 강화 학습 (Reinforcement Learning, RL) 기반의 접근 방식을 통해 power racing 환경에서의 ULPC 문제를 해결하고자 한다. RL에서는 에이전트 (agent)가 주어진 환경과의 상호작용을 통해 최적의 행동의 집합인 정책 (policy)을 학습하는 방식으로, 본 환경과 같은 통신 시스템의 동적이고 복잡한 환경에 적합하다. 본 연구에서는 RL에 신경망을 결합 및 다수의 에이전트를 활용하는 다중 에이전트 심층 강화 학습 (MADRL,

Multi-agent Deep Reinforcement Learning)을 활용하여, 중앙 훈련 장치를 통해 각 UE의 상태 정보를 수집 후, 독립적으로 최적의 전력 제어 결정을 실행하는 QMIX 기반 프레임워크를 제안하며, 동시에 시스템 전체 성능을 극대화하기 위한 협력적 접근 방식을 제안한다.

### II. 본론

본 연구에서는 중심각이 5도인 3개의 셀이 인접하여 배치되고, 각 셀에 1개의 기지국 (BS, Base Station) 및 5명의 UE가 배치되어 있는 상향링크 환경을 상정한다. 해당 UE들은 path-loss의 정도에 따라 강전계/중전계/약전계 UE로 구분한다. 각각의 UE들을 하나의 에이전트로 설정하여, 상향링크에서 다음과 같은 보상 함수를 극대화하는 MADRL 체계를 구축한다. 해당 보상 함수는 sum PF metric을 최대화하는 동시에 누적된 outage 지표에 따라, 시스템의 전반적인 outage probability를 낮추는 것을 목표로 한다.

$$\sum_{i=1}^3 \sum_{k=1}^5 \log \left( \frac{T_{i,k}^{tot}(t) - T_{i,k}^{out}(t)}{T_{i,k}^{tot}(t)} * \log_2(1 + SINR_{i,k}(t)) \right)$$

각 UE는 현재 time slot에서의 채널 이득, 직전 time slot에서의 측정된 SINR 및 outage 여부, 그리고 전력 수준을 자신의 상태 정보로 사용한다. 이를 기반으로, 각 UE는 직전 전력 수준 대비  $\{-2, -1, 0, 1, 2\}$  (dBm)을 선택하는 5차원의 행동 공간을 가진다. 이 때, power racing 환경을 고려 하였을 때, 단순한 독립적 학습 기반 RL 진행 시 UE들은 타 UE의 행동에 의해 간섭량 증가-자신의 전력 증가라는 악순환이 발생하므로, 이를 해결 할 수 있도록 협력적인 알고리즘인 QMIX를 적용한다 [3]. QMIX는 이산적 행동 공간에서, [그림 1]과 같이 중앙 훈련 장치인 mixing network를 설정, 각 에이전트의 행동을 통한 상태-가치 함수 (Q function)의 값을 수 합하여, 이에 대한 가중치를 신경망을 통해 조절하여 시스템의 총합 total Q-value를 최대화하는데, 이를 단조 증가시켜 지속적으로 시스템의 성능을 극대화하는 방향을 설정할 수 있다 [4]-[5]. Mixing network에는, 기존

UE들의 개별적인 (local) 상태 정보 외에도, 각 셀의 평균 outage probability, 각 셀의 스펙트럼 효율 (SE, Spectral Efficiency) 및 평균적인 간섭량 등이 포함된다.

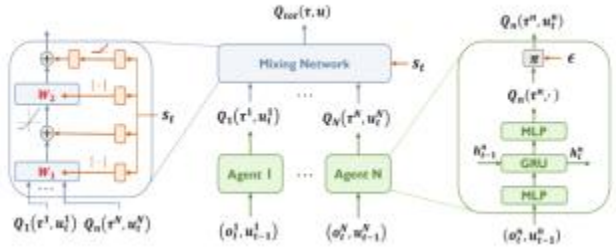


그림 1. QXMI 기반 MADRL-ULPC 프레임워크

본 연구에서, 협력적인 MADRL의 효능을 확인하기 위해서, QMIX와 중앙 훈련 장치를 사용하지 않는 독립적 Q-러닝 프레임워크 (MA-DQN) 알고리즘을 대조군으로 사용한다. MA-DQN의 경우, UE 간 협력이 없으므로, 그림 2와 같이 자신의 정보를 기반으로 outage probability를 낮추기 위해 이기적인 전력 증가만을 고려하며, 이는 power racing과 동일하다.

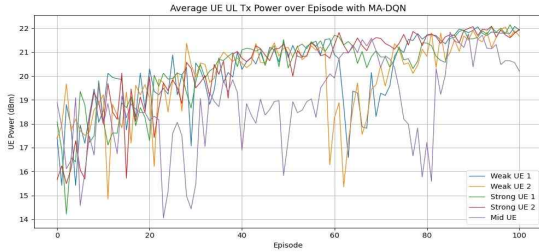


그림 2. MA-DQN 기반 ULPC 경향성 확인

반면, QMIX-MADRL의 경우, 그림 3과 같이 강전계 UE는 하한선에 가까운 2dBm에 근접한 전력을 사용하고, 중, 약전계 UE는 상한선에 가까운 23dBm에 근접한 전력을 사용한다. 이는, 주어진 보상 함수를 고려한 시스템 성능 극대화의 방향성을 고려했을 때, 채널 상태가 좋은 강전계 UE의 경우 전력을 최소화하여 채널 상태가 불량한 중, 약전계 UE에 대한 간섭을 최소화하여, 이들의 outage 가능성을 낮추면서도, 자신들은 outage 없이 충분히 통신이 가능한, 전반적으로 최선의 방향을 찾음을 의미한다...

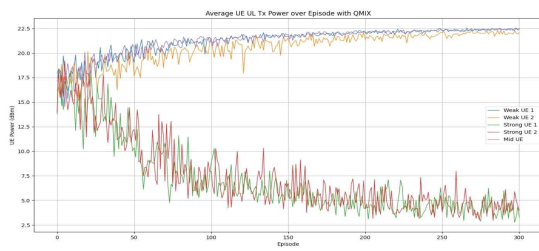


그림 3. QMIX 기반 ULPC 경향성 확인

위 2개의 RL 기반 알고리즘에 더해, RL을 적용하지 않은 ULPC 프레임워크를 추가하여 시스템의 총 SE에 대한 결과는 그림 4와 같이 나타난다. RL을 적용한 두 알고리즘 모두에서 SE의 증가가 나타났는데, 특히 QMIX에서는 강전계 UE의 전력이 저하되면서 강전계 UE 차원에서 제공하는 SE는 줄어들지만, 중, 약전계 UE에 대한 간섭이 저하됨에 따라 outage 가능성이 줄어들고, 이로 인해 중, 약전계 UE가 제공하는 SE의 총량이 기존보다 더욱 증가하여 RL을 적용하지 않았을 때에 대비 오히려 시스템의 총 SE는 증가하는 결과가 나타났다. 이는, UE간 협력을 통한 ULPC 전략을 통해, 시스템의 전반적인 성능이 크게 향상될 수 있음을 입증하는 결과이다.

증하는 결과이다.

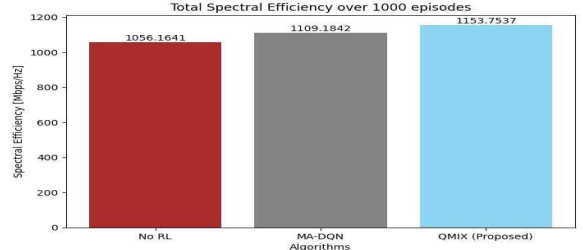


그림 4. 알고리즘별 시스템 총 SE 성능 비교

### III. 결론

본 연구에서는 MADRL 기반 QMIX 알고리즘을 활용하여 power racing 현상에서의 ULPC 문제를 효과적으로 해결하였다. 제안된 방식은 협력적 접근을 통해 다중 셀 환경에서 발생하는 power racing 현상을 억제하고, 중전계 및 약전계 UE의 성능을 보호하는 효과를 보였다. 본 연구는 RL을 기반으로 협력적 MADRL-ULPC 기술이 현실적인 무선 통신 환경에서 높은 성능과 효율성을 제공할 수 있음을 입증하였다. 해당 연구는 더욱 복잡한 환경에서의 해결을 제공할 수 있으며, 추후 간섭량을 집중적으로 제어하는 그래프 신경망 (GNN, Graph Neural Network) 또한 효율적인 AI 기반 ULPC 알고리즘으로 발전할 수 있을 것으로 기대한다].

### ACKNOWLEDGMENT

이 논문은 2024년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (PM2024ST0019).

### 참 고 문 헌

- [1] X. Li, J. Fang, W. Cheng, H. Duan, Z. Chen and H. Li, "Intelligent power control for spectrum sharing in cognitive radios: A. deep reinforcement learning approach," IEEE Access, vol. 6, pp. 25 463 - 25 473, Oct. 2018.
- [2] Y. S. Nasir and D. Guo, "Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks," IEEE J. Sel. Areas Commun., vol. 37, no. 10, pp. 2239 - 2250, Oct. 2019.
- [3] Rashid, T., Samvelyan, M., de Witt, C. S., Farquhar, G., Foerster, J., & Whiteson, S., "QMIX: Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning," in Proceedings of the 35th International Conference on Machine Learning (ICML), 2018.
- [4] R. Jia et al., "Multi-Agent Deep Reinforcement Learning for Uplink Power Control in Multi-Cell Systems," 2022 IEEE International Conference on Communications Workshops (ICC Workshops), Seoul, Korea, Republic of, 2022, pp. 324-330.
- [5] Y. Yang, T. Lv, Y. Cui and P. Huang, "MADRL Based Uplink Joint Resource Block Allocation and Power Control in Multi-Cell Systems," 2023 IEEE Wireless Communications and Networking Conference (WCNC), Glasgow, United Kingdom, 2023, pp. 1-6