Deep Reinforcement Learning-Based Power Allocation in Multi-Cell Massive MIMO

Youngwoo Oh
Dept. of Computer Engineering
Chosun University
Gwangju, South Korea
ywo@chosun.kr

Wooyeol Choi
Dept. of Computer Engineering
Chosun University
Gwangju, South Korea
wyc@chosun.ac.kr

***Abstract*—In this paper, we consider a massive multiple-input multiple-output (MIMO) system that has a large number of transmit antennas at the base station serving multiple users in a downlink multi-cell system. In the massive MIMO system, the number of radio frequency chains and the total transmit power is increasing due to a large number of deployed antennas. The conventional energy-efficient optimization techniques are based on iterative numerical algorithms requiring high computational complexity. To solve this problem, we present a deep reinforcement learning-based power allocation scheme to improve the sum-rate and reduce the complexity. The simulation results demonstrate that the reinforcement learning-based power allocation methods achieve higher energy efficiency with lower complexity than existing optimization algorithms.**

***Index Terms*—Massive MIMO, deep reinforcement learning, power allocation**

I. Introduction

The massive multiple-input multiple-output (MIMO) has been utilized to enhance spectral efficiency by serving multiple users simultaneously with a large number of transmitting antennas at a base station (BS). Since the total consumption is proportional to the antennas in the BS, the energy consumption in a massive MIMO system has significantly increased. To this end, various optimization techniques for maximizing the sum-rate and minimizing power consumption have been performed to improve energy efficiency in the massive MIMO system. However, most of the optimization problems are non-convex [1], [2] and require a high computational complexity.

In this paper, we consider a deep reinforcement learning (DRL)-based approach to solve these problems. DRL-based techniques have been rapidly developed in wireless communications. These approaches have demonstrated finding the optimal value faster than the conventional methods and are robust against dynamic changes [3], [4]. Therefore, we propose a DRL-based power allocation scheme to improve the sum-rate performance and reduce the complexity in the massive MIMO systems.

This research was supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under grant (No. NRF-2021R1I1A3050535), and in part by 'Technology Commercialization Collaboration Platform' program of the Korea Innovation Cluster (No. 2022-DD-RD-0065).

II. DRL-Based Power Allocation

We consider the downlink power allocation for a multi-cell massive MIMO system. All user equipment (UE) is randomly located in N cells, and a BS is deployed in the center of each cell. The BS is equipped with M transmitting antennas and serves K UEs with a single antenna. The received signal at the k-th UE associated with n-th BS is the sum of the desired signal transmitted from n-th BS, the interference from adjacent BSs, and the noise, and can be expressed as

$$y_{n,k} = p_{n,k} g_{n,k} w_{n,k} x_{n,k} + \sum_{j \neq n}^{N} \sum_{k=1}^{K} p_{j,i} g_{j,i} w_{j,i} x_{j,i} + n_{n,k}, \tag{1}$$

where $p_{n,k}$, $w_{n,k}$, and $x_{n,k}$ is the transmitted power, the zero-forcing precoding vector, and the downlink signal of n-th BS to k-th UE, respectively, and $n_{n,k}$ is additive white Gaussian noise with variance σ^2. Therefore, the signal-to-interference and noise ratio (SINR) of the k-th UE connected to n-th BS can be defined as

$$\gamma_{n,k} = \frac{p_{n,k} \| g_{n,k} w_{n,k} \|^2}{\sum_{j \neq n}^{N} \sum_{k=1}^{K} p_{j,i} \| g_{j,i} w_{j,i} \|^2 + \sigma^2}. \tag{2}$$

The downlink rate for k-th UE is given by $C_{n,k} = \log_2 (1 + \gamma_{n,k})$. The problem of maximizing the sum-rate can be defined as

$$\max_{p_{n,k}} f(p_{n,k}) \tag{3}$$
$$\text{s.t. } 0 < p_{n,k} \leq P_{\text{max}}, \ \forall n, k,$$

where the $f(p_{n,k}) = \sum_{n}^{N} \sum_{k}^{K} C_{n,k}$ and P_{max} is maximum transmission power. The conventional power allocation methods consume high computational time to solve the above problem. Therefore, we apply the DRL algorithms considering the complexity and sum-rate performance. To solve the maximization problem by the reinforcement learning algorithms, we transform it into a Markov decision process (MDP).

The state space consists of the SINR, and the objective function, and can be expressed as

$$\mathrm{S}(t) = \{\gamma_{n,k}(t), f(p_{n,k}(t))\}, \ \forall n, k, \tag{4}$$

The action space $\mathrm{A}(t)$ is the set of the downlink transmission power for all BSs to UEs, which is selected by dividing it into

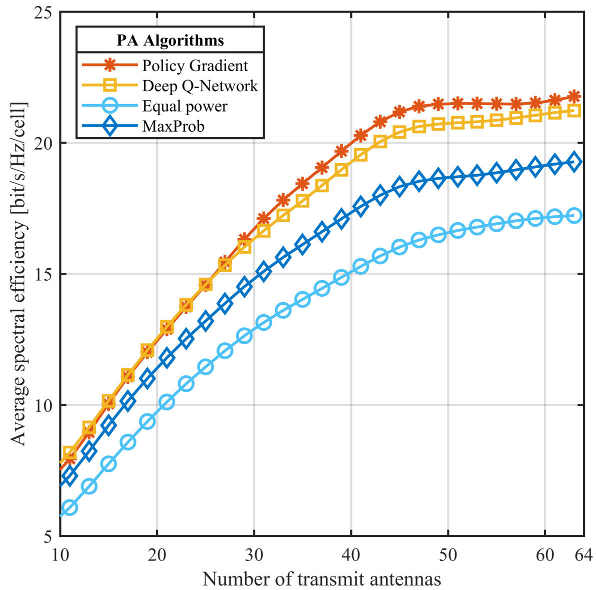

Fig. 1. A performance evaluation of different methods at various BS antennas.

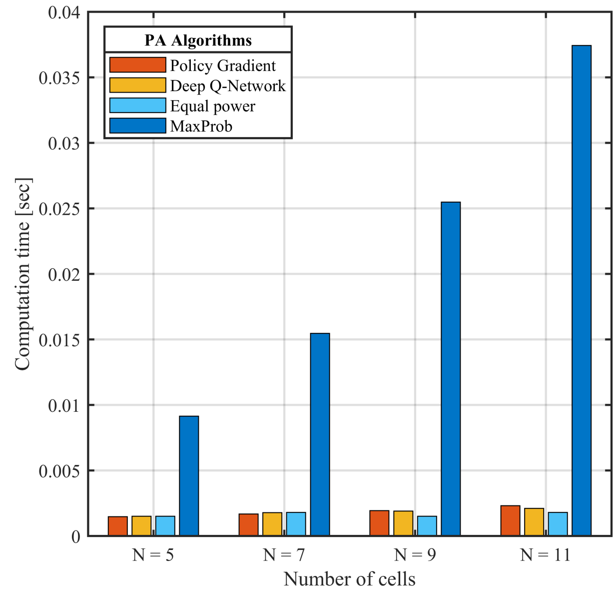

Fig. 2. A comparison of the computational time for each method at different number of cells.

specific steps between $P_{\min}$ and $P_{\max}$ to reduce the complexity. The reward by the state and action spaces is the same as the objective function to maximize the sum-rate, and can be defined as

$$\mathrm{R}(t) = f\left(p_{n,k}(t)\right) \tag{5}$$

Therefore, the DRL agent trains to obtain as much reward as possible, considering the maximize the sum-rate.

III. Simulation Results

This section evaluates the numerical simulation performance for the applied DRL-based power allocation in the multi-cell massive MIMO system. For the massive MIMO system, we consider the number of BS antennas installed to be 64, and the total number of users is 10. In addition, the cell is a radius of 500 m, and the maximum transmitted power of BS is 32 dBm.

We conducted the simulations using both value-based and policy-based DRL algorithms. The deep Q-network (DQN) consists of three fully connected layers of 128, 64, and 64 neurons with a Relu activation function. In addition, we rely on the Adam optimizer and the mean squared error (MSE) loss function. We evaluate the performance of the applied DRL algorithms based on the average spectral efficiency and computational time. The maximum product SINR and equal power algorithms are suggested as benchmarks to compare the DRL algorithms, where we denote the maximum product SINR algorithm as MaxProb.

Fig. 1 presents the performance of the average spectral efficiency of each algorithm at a different number of BS antennas. The results show that DQN and policy gradient achieve higher average spectral efficiency than the suggested benchmark algorithms. The reason is that the DRL agents acquire environmental information, such as states and transmission power, and update it after each episode to determine the optimal action. On the other hand, the MaxProb method has a lower performance due to the number of repetitions, the interference of adjacent cells, and the effect of large-scale fading.

Fig. 2 shows the computational time of each algorithm as the number of cells increases. The result demonstrates that the computational time of MaxProb increases exponentially, whereas DRL-based and equal power methods are stable, even though the number of cells increases.

IV. Conclusion

This paper presented the DRL algorithms to maximize the sum-rate in a multi-cell massive MIMO system. Through the extensive simulations, we confirmed that the DRL methods can achieve a higher performance of the sum-rate with lower complexity than the existing optimization techniques. Our future work is to perform a joint antenna selection and power allocation method to maximize the spectral efficiency of the massive MIMO systems.

References

[1] L. You, Y. Huang, D. Zhang, Z. Chang, W. Wang, and X. Gao, "Energy efficiency optimization for multi-cell massive mimo: Centralized and distributed power allocation algorithms," *IEEE Transactions on Communications*, vol. 69, no. 8, pp. 5228–5242, 2021.

[2] J. Choi, N. Lee, S.-N. Hong, and G. Caire, "Joint user selection, power allocation, and precoding design with imperfect csit for multi-cell mu-mimo downlink systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 162–176, 2020.

[3] Y. S. Nasir and D. Guo, "Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 10, pp. 2239–2250, 2019.

[4] L. Luo, J. Zhang, S. Chen, X. Zhang, B. Ai, and D. W. K. Ng, "Downlink power control for cell-free massive mimo with deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 6772–6777, 2022.