九州大学 エネルギーウィーク 2021
Energy Week 2021
Q-PIT Kyushu University Platform of Inter/Transdisciplinary Energy Research

Energy-Efficient Reinforcement Learning-Based UE Pairing in Non-Orthogonal Multiple Access Wireless Communication Systems

Ahmad Gendia[1], Osamu Muta[2]

[1]Graduate School of Information Science and Electrical Engineering, Kyushu University
[2]Center of Japan-Egypt Cooperation in Science and Technology, Kyushu University

1. Introduction

Non-orthogonal multiple access (NOMA) allows multiple user equipment (UE) to simultaneously share the same resource blocks using varying levels of transmit power at the base station (BS) side. Proper selection of candidate users for pairing over the same resource block is critical for an efficient utilization of the available resources. In this article, a reinforcement learning (RL)-based scheme is proposed for user pairing in downlink NOMA systems, where optimized UE selection is accomplished by the continuous interaction between an RL agent and the NOMA environment with the aim of increasing the total system throughput while conserving the energy consumed for NOMA transmissions.

Keywords : Wireless communication networks, NOMA, Reinforcement Learning, Energy efficiency.

NOMA Networks

- NOMA can achieve very high spectrum efficiency operation by aggregating the data of multiple UEs over the same resource block (RB).
- NOMA can utilize energy efficiently through proper UE pairing so as to combat the inherent interference.

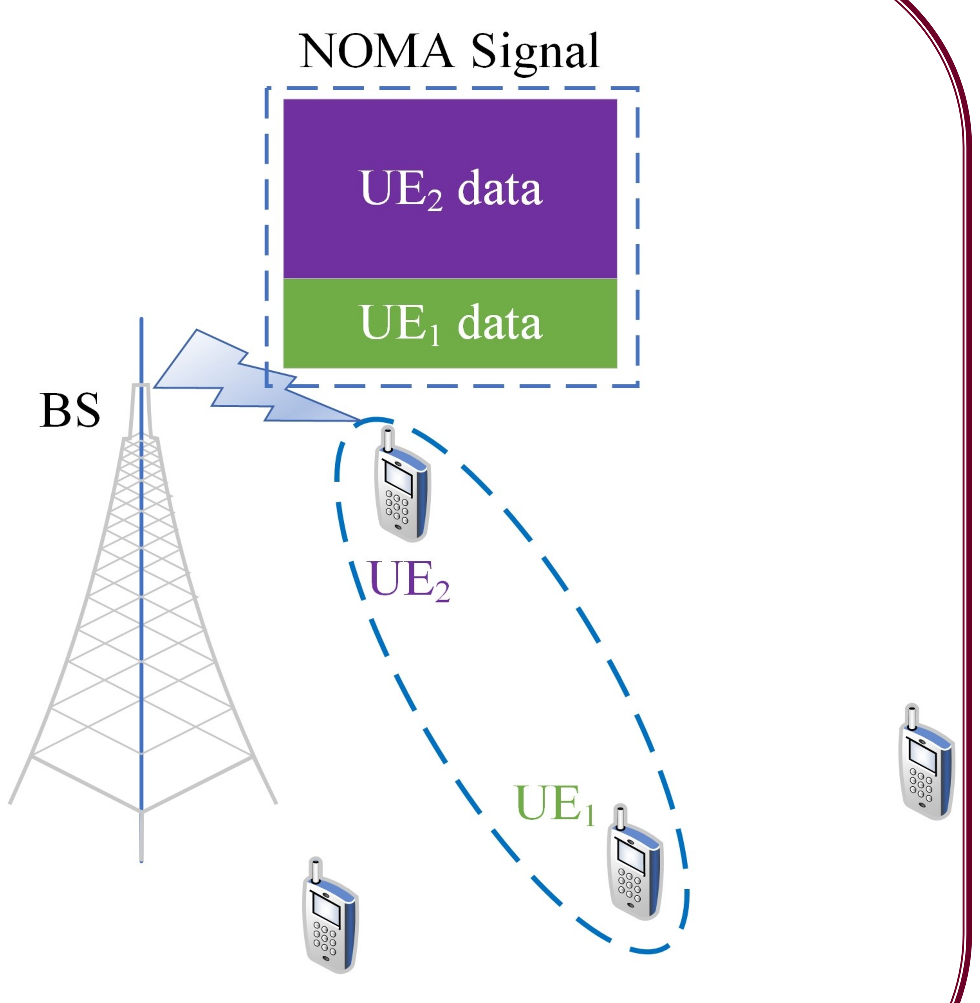


Fig. 1: Downlink NOMA system transmission

Agent-Environment Dynamic Interaction

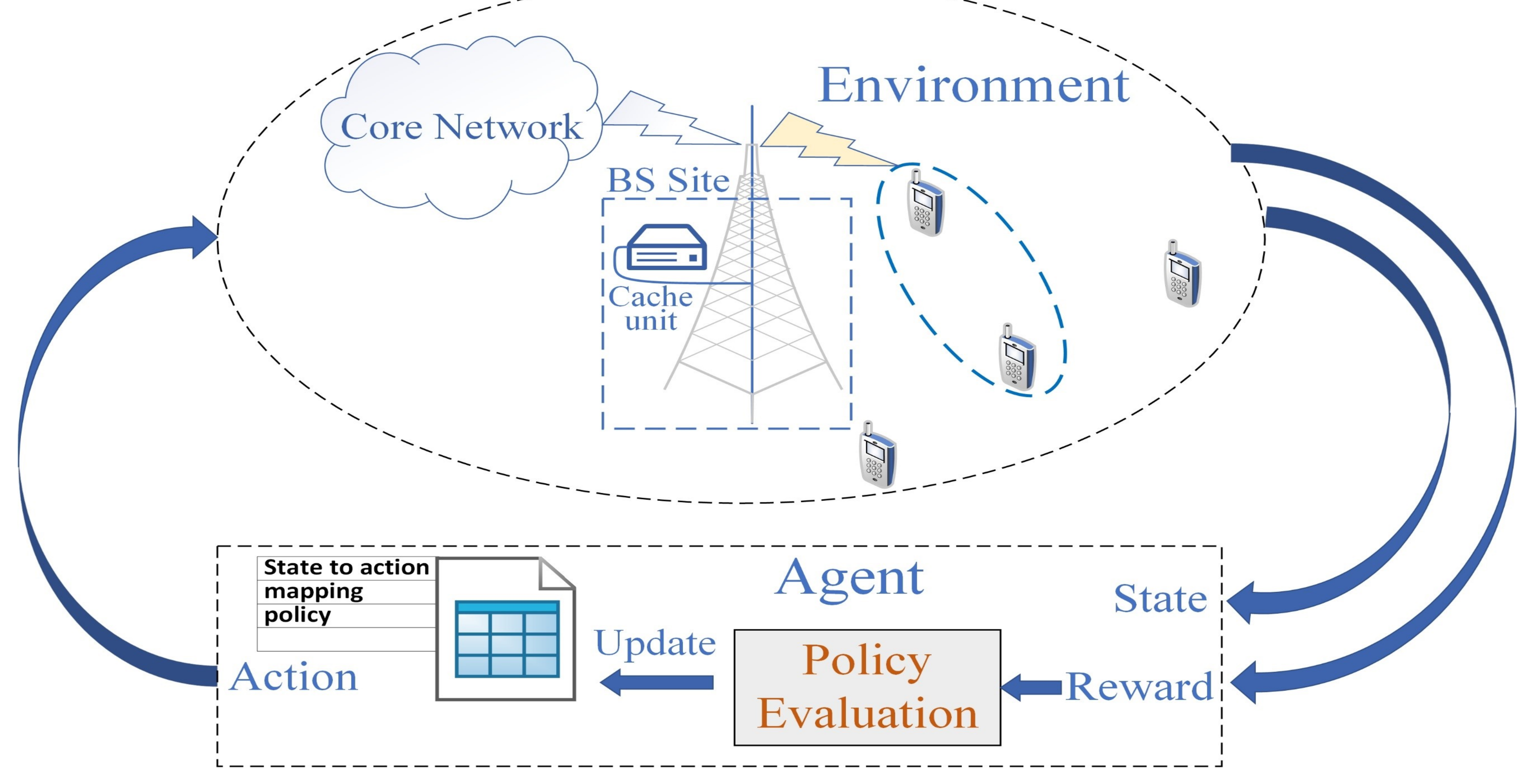


Fig. 2: Top-level view of the proposed scheme

- **The agent takes an action based on its internal policy, receives a reward and makes a transition to a new state.** This dynamic interaction enables the agent to adapt to the effects of the various factors within the environment that impact the sum-rate and energy consumption.

System Model

- Consider a downlink NOMA system where N candidate UEs are requesting popular contents (e.g. popular internet data) for download:

1. The BS has to first fetch the requested information via capacity-limited backhaul link to the core network.

2. For M-user NOMA transmission over a certain RB, $M < N$, the BS then compiles and forwards a NOMA message containing the content of the M UEs with maximum potential contribution to the system throughput:

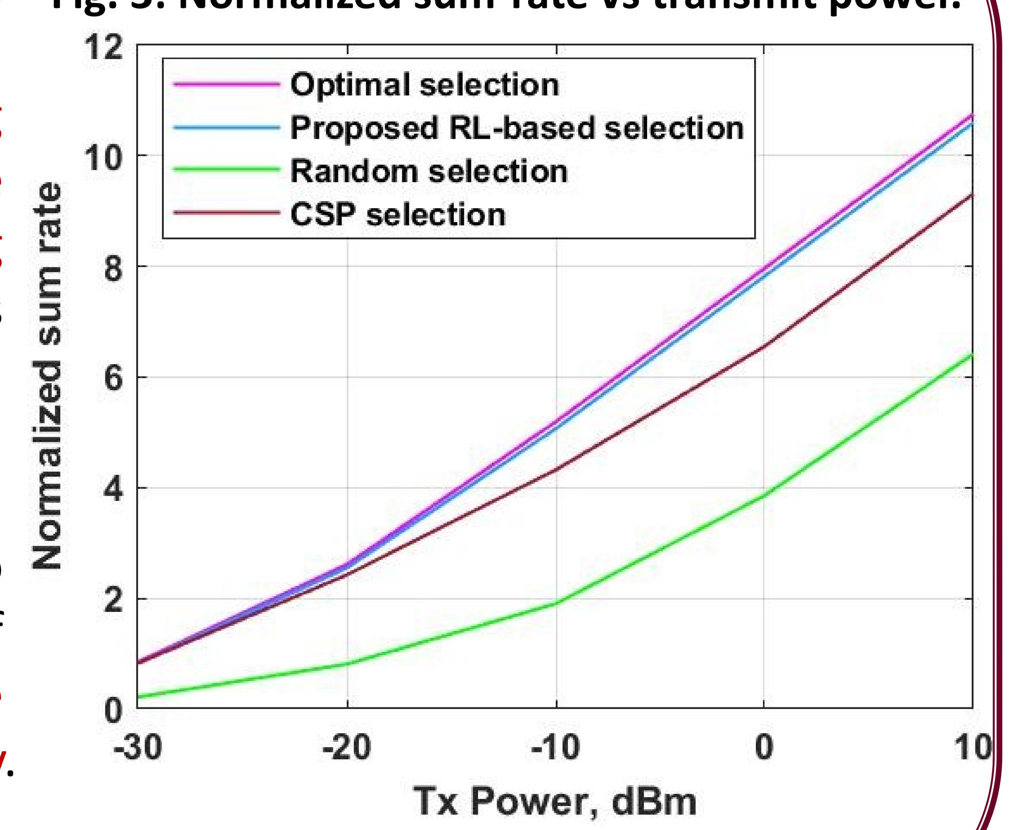
$$R = \sum_{m=1}^{M} \min\left(C, B\, log_2\left(1 + \frac{\alpha_m |h_m|^2}{\sum_1^{m-1} \alpha_i |h_m|^2 + \sigma^2}\right)\right)$$

where C is the backhaul capacity allocated for an active UE, B is the transmission bandwidth, and σ^2 is the noise power.

Simulation Results

- The proposed scheme is compared to the optimal and random pairing approaches as well as the channel-sorting pairing (CSP) scheme where UEs with high channel-gain separation are paired [2].
- The RL agent learns to optimize the selection of paired UEs to utilize available energy efficiently.

Fig. 3: Normalized sum-rate vs transmit power.



[2] H. Zhang, D. Zhang, W. Meng, and C. Li, "User pairing algorithm with sic in non-orthogonal multiple access system," in IEEE ICC, pp. 1–6.

Reinforcement Learning-Based UE Pairing

- The proposed scheme deploys a caching unit at the BS side to store high-demand contents to alleviate the data throttling caused by the backhaul link [1].
- An RL-based agent interacts with the NOMA environment to learn the pairing decisions resulting in higher sum rates and optimized energy consumption for the activated UEs:

- ✓ Initially, the agent starts with a random policy that maps the state of the environment to a random pairing action.
- ✓ the action is executed and a corresponding reward is observed.
- ✓ The agent then updates its policy based on the cumulative rewards observed over many interactions.

[1] F. R. Yu and Y. He, "Deep reinforcement learning for interference alignment wireless networks," in Deep Reinforcement Learning for Wireless Networks. Springer, 2019, pp. 21–44.

Summary

- **We have proposed an RL-based scheme for user-pairing in NOMA systems [3-4]** with a limited-capacity backhaul connection to achieve near-optimal sum-rate levels while **conserving the energy consumed for UE data transmission** as indicated in the table below.

Scheme	Transmit power savings (dB)
Optimal	15.5
Proposed	15
CSP	10.75

$$\text{Power saving} = 10\log_{10}(\gamma_r/\gamma_s)\ \text{dB}$$

γ_r is the random pairing transmit power threshold and $\gamma_s \in \{\gamma_{Optimal}, \gamma_{Proposed}, \gamma_{CSP}\}$ is the transmit power of the respective scheme.

- **The proposed scheme achieves 6 bps/Hz at 15 dB transmit power saving.**

[3] A. H. Gendia, Osamu Muta, "Unified Reinforcement Learning-Based User Pairing in NOMA Systems," in IEICE general conference, March 2021.
[4] __, "User Pairing in NOMA Systems via Reinforcement Learning Selection," in RCS, March 2021.