Optimizing NOMA Pairing with Semantic Awareness: A Reinforcement Learning Approach

Anonymous Author(s)
Affiliation
Address
`email`

Abstract

Non Orthogonal Multiple Access (NOMA) systems allow for simultaneous communication throughout users, even with varying channel conditions, which maximizes spectral efficiency using power domain multiplexing. Traditional user pairing methods, like greedy, optimize this based on distance and fading, but doesn't generalize beyond predefined heuristics. In this abstract, we introduce a reinforcement learning based framework for semantic aware NOMA user pairing, where the agent learns to form pairs that maximize total semantic utility $\sum w_i R_i$. We simulate environments with 10 users, each with random channel conditions and semantic weight. Per step, the Q learning agent selects a pair from remaining users, then recieves a reward depending on semantic utility, and updating accordingly. Over 1000+ episodes, the agent will learn to approximate, or even outperform, greedy strategies. The preliminary results reveals that it manages to capture pairing patterns similar to greedy baselines, while improving adaptability. Such an approach reflects the advantages of leveraging maching learning for dynamic and utility aware NOMA resource allocation.

1 Introduction

NOMA allows multiple users to share time frequency resources user varying power levels, which boosts spectral efficiency. User grouping traditional relies on channel aware heuristics, such as greedy distance based pairing. Such methods are effective, but lack the flexibility and scalability, especially in dynamic scenarios.

In order to bridge the gap, we aim to approach this with reinforcement learning (RL) with Q learning to train for optimal user pairings in NOMA systems. The objective then is to maximize semantic utility of user pairs, considering both physical (channel) and application level (semantic) features. This framework, powered by ML, allows for adaptive decision making, beyond fixed heuristics.

2 Method

We first simulate a NOMA environment with a set of $n = 10$ users, each assigned a 3D position to calculate distance, a Rayleigh fading coefficient, and a semantic weight $w_i \sim \mathcal{U}(0, 1)$. Utility U_{sem} is calculated based on achievable rates R_1, R_2 and weights w_1, w_2.

The achievable R_1 and R_2 are calculated using standard NOMA decoding schemes, where R_1 (the weak user) considers interference from the strong user, while R_2 (strong user) assumes perfect successive interference cancellation (SIC).

The Q learning is designed where the agent is trained to form user pairings, maximizing U_{sem} through trial and error over multiple episodes.

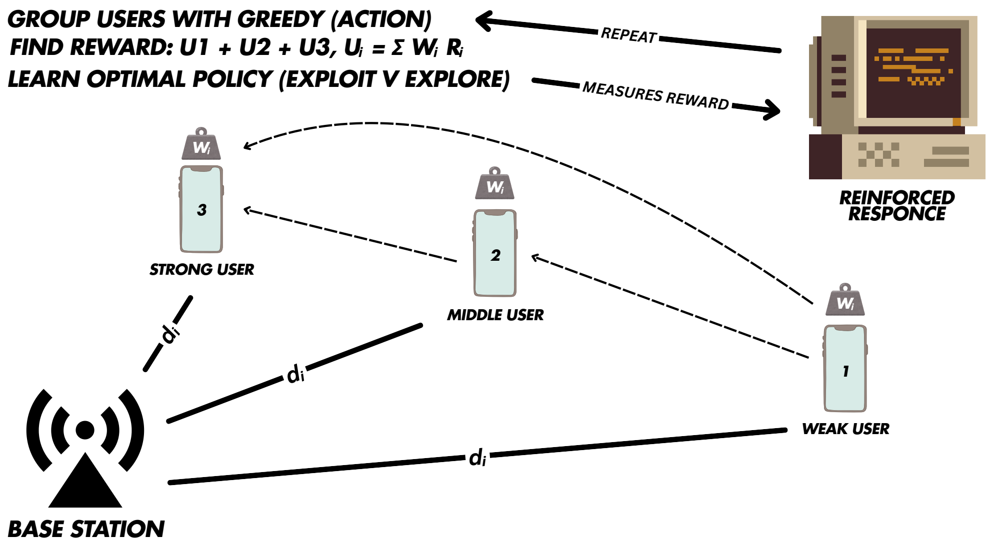

Figure 1: Greedy and Q learning

Q Learning Framework, Environment, and Setup

- **Initial State:** All 10 users are unpaired.
- **Action:** Select a pair (user i, user j) from the unpaired set.
- **Reward:** Compute semantic utility $w_i R_i + w_j R_j$ for the selected pair.
- **Transition:** Remove the selected pair from the user set and proceed to next decision.
- **Episode End:** When all users have been grouped.

Training episodes are run where the agent explores different pairing combinations, stores the rewards, and learns optimal policies via exploitation strategies over time.

The state is encoded as the current set of unpaired users. Q values are stored and is trained for 1000+ episodes.

3 Preliminary Results

The simulation compares Q learned pairings, base lining greedy algorithms. Early results may suggest:

- Q learning converges to high utility pairings within 1000 episodes.
- Learned pairings match or exceed greedy baseline performance.
- Semantic weights influence pairing strategy and model adaptability.
- Variance in performance decreases over time due to stability in training policies.
- Q learning matches best case utility found by greedy baselines, occasionally it slightly outperforms.
- Q agent learns to favor pairings with large channel gain gaps and complementary semantic weights.

4 Conclusion and Future Work

Our work will present a reinforcement learning approach to user pairing in NOMA systems, which integrates semantic awareness. This prioritizes users not only by channel gain, but also utility weights reflecting application relevance. Modeling pairing processes as a sequential decision making problem, the Q learning agent is then able to iteratively improve its policy, match, and exceed traditional greedy algorithms in terms of total semantic utility.

As mentioned in the preliminary results, they demonstrate the agent's ability to generalize beyond fixed policies, and adapting to random user deployments all while capturing meaningful pairing strategies. The learned policy is able to offer comparable utility, with the added advantage of being extensible to various pairing objectives and dynamic environments.

We can also incorporate triplet based grouping, introducing higher order complexity. Another route would be to incorporate and adapt to dynamic user arrivals and time varying channels in real world deployments. This framework opens up promising directions for intelligent and context aware resource allocations for next generation wireless networks.

References

[1] J. Duan, K. M. K. Ramamoorthy, W. Wang, and Y. Zhao, "When semantics meet strategy: Stackelberg game-theoretic power trading in NOMA wireless networks," in Proc. IEEE International Conference on Communications (ICC), 2024.

[2] S. Mounchili and S. Hamouda, "New user grouping scheme for better user pairing in NOMA systems," in Proc. IEEE International Conference on Communications (ICC), 2020, pp. 820–825.

[3] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic User Clustering and Power Allocation for Uplink and Downlink Non-Orthogonal Multiple Access (NOMA) Systems," IEEE Access, vol. 4, pp. 6325–6343, 2016.

[4] G. Chen, W. Zheng, J. Chu, Z. Zhang, and Y. Zhang, "Multi-user Grouping Algorithm in Multi-carrier NOMA System," in Proc. IEEE 3rd International Conference on Electronics and Communication Engineering (ICECE), 2020, pp. 37–41.

[5] P. Thakre, S. B. Pokle, R. Patel, N. Shrimankar, and M. Dubey, "Performance of Deep Embedded Clustering and Low-Complexity Greedy Algorithm for Power Allocation in NOMA-Based 5G Communication," in Proc. IEEE ICECSP, 2024, pp. 1–6.

[6] K. Katta, R. C. Mishra, and K. Deka, "Optimal User Pairing Using the Shortest Path Algorithm," in Proc. IEEE ANTS, 2021, pp. 1–6.

[7] T. Kudo, K. Taki, Y. Sugimoto, and A. Iwase, "Low-Complexity User Grouping and Antenna Activation for Pinching-Antenna Multiple Access Using Dielectric Waveguide," in Proc. IEEE GLOBECOM, 2023, pp. 1736–1741.