NOMA-Enhanced Intelligent Semantic Communication Networks using Deep Reinforcement Learning

Thanh Phung Truong, Donghyeon Hur, Junsuk Oh, Donghyun Lee, Dongwook Won, Jeongyeup Paek, and Sungrae Cho

School of Computer Science and Engineering, Chung-Ang University, Seoul 06974, South Korea.
Email: {tptruong, dhhur, jsoh, dhlee, dwwon}@uclab.re.kr, jpaek@cau.ac.kr, srcho@cau.ac.kr.

***Abstract*—This research explores an intelligent semantic communication network where the base station (BS) processes and transmits semantic data to users through a wireless link. To improve communication between the BS and users, we apply the non-orthogonal multiple access (NOMA) technique in this system. Within this system, we consider the delay model for transmitting original data from the BS to users using semantic communication. This model consists of three main phases: *(i)* extracting semantic data from the original data at the BS; *(ii)* transmitting this semantic data from the BS to the users; and *(iii)* reconstructing the original data from the received semantic data at the users. To optimize the overall system delay, we formulate a problem involving optimizing beamforming vectors and computation resource allocation at the BS. To address this problem, we propose a deep reinforcement learning (DRL) framework that utilizes the deep deterministic policy gradient algorithm. In simulations, we analyze the convergence of our proposed framework, which shows stable performance under varying algorithm parameters.**

***Index Terms*—non-orthogonal multiple access, semantic communication networks, deep reinforcement learning**

I. Introduction

The quick promotion of emerging applications, such as the metaverse and digital twin, necessitates that wireless networks be capable of delivering high data transmission rates and ultra-low latency. Nevertheless, the limited resources of wireless communication links may only support transmitting a little data, especially in real-time applications that use speech, images, and text. To this end, semantic communication becomes a viable solution for reducing the transmitted data [1]. Unlike traditional communication systems, semantic communication places its emphasis on the extraction of fundamental semantics through the process of encoding source messages (referred to as semantic encoding). This method filters out any content that is unrelated to the predefined conveyed information, effectively minimizing the volume of data transmitted, all the while retaining the integrity of the original semantics [2]. Accordingly, the application of AI that enables semantic communication can be referred to as a system named intelligent semantic communication [3]–[5].

Besides, the rapid increase in the number of devices may reduce communication performance due to interference. As a practical solution to massive connectivity, non-orthogonal multiple access (NOMA) has been studied and applied in many research, which proves its effectiveness in improving system performance in dense environments [6]–[8]. In particular, the authors in [6] delve into the synergy between NOMA and full-duplex (FD) communication, aiming to enhance the user fairness and spectral efficiency of the system. A NOMA-assisted mobile edge computing (MEC) system is considered in [7]. This work proposes a deep reinforcement learning (DRL) framework to design the offloading policy and resource allocation in the MEC system to minimize the computational overhead. The DRL is also used in [8], where they optimize the offloading strategy and the transmit power to minimize the cost in high altitude platform-aided vehicular networks. With the increase of use of DRL in many research, it has emerged as a powerful tool for addressing numerous challenges in communication systems [9]–[11]. It offers a viable solution for tackling complex problems and scenarios with dynamic observations. In response to the above observations, this research considers the use of NOMA in an intelligent semantic communication system, where we formulate a system delay minimization problem by optimizing the transmit beamforming vectors and computation resource allocation variables. To solve the problem, we propose a DRL framework that applies a well-known algorithm named deep deterministic policy gradient (DDPG) . In summary, the main contributions of our work are list as bellows.

- We consider a novel system that employs NOMA to enhance the intelligent semantic communication system. Herein, we assess the delay model of transmitting the original data from the BS to the users via semantic communication with the NOMA technique in the wireless

This work was supported in part by the National Research Foundation of Korea (NRF), South Korea grant funded by the Korea government (MSIT) (No. 2022R1A4A5034130), in part by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2024-RS-2022-00156353) supervised by the IITP (Institute for Information and Communications Technology Planning and Evaluation), South Korea.

link.

- We formulate a system delay minimization problem by optimizing the transmit beamforming vectors and the computation resource allocation variables. To solve the problem, we propose a DRL framework that applies the DDPG algorithm to train the agent.
- We show the convergence of the proposed algorithm in the simulation, where we assess the training performance in different algorithm parameters.

The rest of this paper is organized as follows. Section II introduces the system model, delay model, and formulates the considered problem. In section III, we propose the DRL framework to solve the problem. Then, we demonstrate the simulation results in Section IV. Finally, we conclude the work in Section V.

II. Problem Statement

A. System Model

We consider an intelligent semantic communication system, where an M-antenna BS serves K single-antenna users. Herein, the communication between users and BS is enhanced by using the NOMA technique. At BS, the original data with large size is extracted to small-size semantic data to transmit to the users via wireless communication. This extraction reduces the amount of transmitted data, bringing benefits in resource-saving and network constraints [3], and can be obtained by applying semantic model using deep neural networks (DNNs) [12]. To reconstruct the original data from the received semantic data, the BS shares its semantic model with all users, referred to as knowledge data. To enhance the communication efficiency, we employ the NOMA technique to the communication between BS and users. Accordingly, the users utilizes successive interference cancellation technique to decode their own signal.

B. Delay Model in Semantic Communication

We build the delay model for the considered semantic communication according to the work [13], which can be divided into three phases: *(i)* extracting semantic data from the original data at the BS, *(ii)* transmitting this semantic data from the BS to the users, and *(iii)* reconstructing the original data from the received semantic data at the users. Before transmitting the data, the BS extracts the original data into semantic data, where it allocates the computation resource f_k to extract data for user k. Accordingly, this semantic data is transmitted to the user via wireless communication enhanced by NOMA. Also, the BS transmits the knowledge data to users for constructing the original data. At the user side, the original data is reconstructed from the received semantic data and knowledge data. As a result, the total delay in transmitting the original data from the BS to the $k-$th user in a semantic communication network, t_k, is calculated as the sum of the delay in three phases.

C. Problem Formulation

In this study, we aim to minimize the total delay for all users while satisfying the delay constraint for them. To do so, we formulate a delay minimization problem by optimizing the beamforming vectors and the computation resource allocation at the BS.

III. Proposed DRL Framework

In this section, we propose a DRL framework that applies the well-known DRL algorithm named deep deterministic policy gradient (DDPG) to solve the problem. First, we model the problem as an RL-based problem, where the agent is implemented at the BS, and the environment is the whole system. At each time step, the agent decides the action to interact with the environment based on the observed state and gets back the reward. Herein, the state space, action space, and reward function at time slot t is defined as

- *State space:* The state space indicates the environment observation affecting the reward. In this work, the state space, therefore, includes the channel coefficient vectors, the required computation cycles, and the size in bits of semantic data of all users.
- *Action space:* It includes all optimization variables the agent has to decide at each time slot, which are the beamforming vectors and the computation resource allocation at the BS.
- *Reward function:* With the aim of maximizing the reward in DRL training algorithms, we define the reward as a negative value of the total delay so that maximizing the reward leads to minimizing the total delay. Besides, for users violating the delay constraint, we penalize them with a penalty value.

To train the agent, we utilize the DDPG algorithm, which combined from actor and critic networks, each has a main and a target networks [14]. The agent observes the state and decide action at each time slot using the main actor network. The decided action is measured by using the main critic network. Besides, the target networks are used to enhance the training performance.

The proposed algorithm includes interaction and training processes. At each time step of the interaction process, the agent determines action by the main actor networks. Accordingly, the obtained variables are interacted to the environment. Then, a sample including state, action, reward, and next state is stored to the buffer for training. At each step of the training process, a random batch of samples is selected from the buffer. Then, the DDPG algorithm is the adapted to train the neural network. After training, the optimal actor network is obtain to interact with the environment.

IV. Numerical results

This section evaluates the performance of the proposed framework via simulation. We simulate a scenario where a BS serves ten users from 10 to 200 meters away from it. The channel coefficient vectors are modeled as [6]. The semantic data size is set in the range of 200 to 600 (bits), the amount of

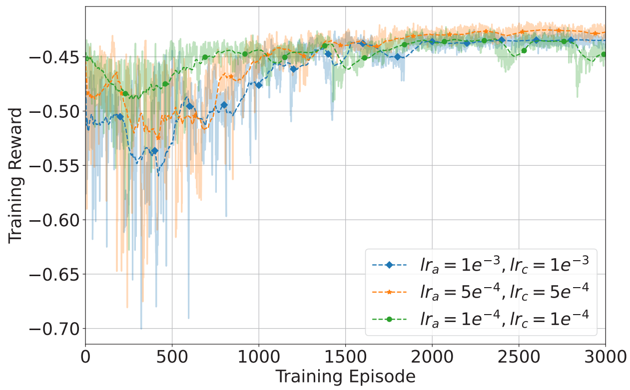

Fig. 1: Model convergence according to learning rates.

computation cycle required for extracting and reconstructing the data is set from 800 to 1000 (cycles).

We analyze the convergence of the proposed DRL algorithm by training the algorithm in different learning rates. The actor learning rate (lr_a) and critic learning rate (lr_c) is selected from three values: $lr_a = lr_c = \{1e^{-3}, 5e^{-4}, 1e^{-4}\}$, while the discount factor, batch size, and buffer size are corresponding selected as 0.99, 16, and 100000. The neural networks have two hidden layers, where the actor networks have 256 and 128 nodes, and the critic networks have 512 and 256 nodes in their hidden layers. As shown in Fig. 1, the model converges after about 2000 episodes, where the case $lr_a = lr_c = 5e^{-4}$ gives the best results. In particular, it is approximate 3 % better than the remaining cases. Therefore, we use the trained model with the learning rate $lr_a = lr_c = 5e^{-4}$ to evaluate the proposed framework performance.

V. Conclusion

This study considered an intelligent semantic communication network, where the BS extracts the semantic data and transmits it to its users via the wireless communication link. Herein, the NOMA technique was applied to the system to enhance the communication between the BS and users. In this system, we investigated the delay model of transmitting the original data from the BS to the users via semantic communication. Accordingly, we formulated a problem of minimizing the total system delay by optimizing the beamforming vectors and the computation resource allocation at the BS. To solve the problem, we proposed a DRL framework that applies a well-known DDPG algorithm to train the agent. In the simulation, we evaluated the convergence of the proposed framework in different algorithm parameters.

References

[1] B. Güler, A. Yener, and A. Swami, "The semantic communication game," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 787–802, 2018.

[2] E. Calvanese Strinati and S. Barbarossa, "6G networks: Beyond shannon towards semantic and goal-oriented communications," *Computer Networks*, vol. 190, p. 107930, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128621000773

[3] L. Xia, Y. Sun, X. Li, G. Feng, and M. A. Imran, "Wireless resource management in intelligent semantic communication networks," in *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2022, pp. 1–6.

[4] H. Xie and Z. Qin, "A lite distributed semantic communication system for internet of things," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, pp. 142–153, 2021.

[5] H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2663–2675, 2021.

[6] H. V. Nguyen, V.-D. Nguyen, O. A. Dobre, D. N. Nguyen, E. Dutkiewicz, and O.-S. Shin, "Joint power control and user association for NOMA-based full-duplex systems," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 8037–8055, 2019.

[7] V. D. Tuong, T. P. Truong, T.-V. Nguyen, W. Noh, and S. Cho, "Partial computation offloading in NOMA-assisted mobile-edge computing systems using deep reinforcement learning," *IEEE Internet of Things Journal*, vol. 8, no. 17, pp. 13 196–13 208, 2021.

[8] T.-H. Nguyen, T. P. Truong, N.-N. Dao, W. Na, H. Park, and L. Park, "Deep reinforcement learning-based partial task offloading in high altitude platform-aided vehicular networks," in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, 2022, pp. 1341–1346.

[9] T.-H. Nguyen, H. Park, K. Seol, S. So, and L. Park, "Applications of deep learning and deep reinforcement learning in 6G networks," in *2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN)*, 2023, pp. 427–432.

[10] V. D. Tuong, W. Noh, and S. Cho, "Sparse CNN and deep reinforcement learning-based D2D scheduling in UAV-assisted industrial IoT networks," *IEEE Transactions on Industrial Informatics*, pp. 1–11, 2023.

[11] T. P. Truong, T.-H. Nguyen, A.-T. Tran, S. V.-T. Tran, V. D. Tuong, L. V. Nguyen, W. Na, L. Park, and S. Cho, "Deep reinforcement learning-based sum-rate maximization for uplink multi-user SIMO-RSMA systems," in *Intelligence of Things: Technologies and Applications*, N.-N. Dao, T. N. Thinh, and N. T. Nguyen, Eds. Cham: Springer Nature Switzerland, 2023, pp. 36–45.

[12] W. Zhang, K. Bai, S. Zeadally, H. Zhang, H. Shao, H. Ma, and V. C. M. Leung, "DeepMA: End-to-end deep multiple access for wireless image transmission in semantic communication," *IEEE Transactions on Cognitive Communications and Networking*, pp. 1–1, 2023.

[13] Z. Yang, M. Chen, Z. Zhang, and C. Huang, "Energy efficient semantic communication over wireless networks with rate splitting," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 5, pp. 1484–1495, 2023.

[14] T. P. Truong, V. D. Tuong, N.-N. Dao, and S. Cho, "FlyReflect: Joint flying IRS trajectory and phase shift design using deep reinforcement learning," *IEEE Internet of Things Journal*, vol. 10, no. 5, pp. 4605–4620, 2023.