Deep Reinforcement Learning-based Resource Allocation and Mode Selection for Semantic Communication

Hyeonho Noh*, Sojeong Park†, and Hyun Jong Yang*
*Department of Electrical and Computer Engineering, Seoul National University, Korea
†Department of Electrical Engineering, Pohang University of Science and Technology, Korea

Abstract—In this paper, we aim to solve the joint resource allocation and mode selection problem, in which an agent adaptively allocates communication users to appropriate resource units and toggles between bit and semantic transmission modes while determining the count of transmitted semantic symbols in semantic communication mode. Specifically, in contrast to the common yet unrealistic assumptions of prior research, which posits the possibility of limitless data transmission over infinite periods, our focus shifts towards the realities of unsaturated traffic conditions, where users transmit a finite amount of data within restricted time frames. In order to evaluate the efficiency of data transmission within the semantic domain under unsaturated traffic conditions, we propose a short-term semantic transmission rate (SR), as an evaluation metric of the joint problem. Under these unsaturated traffic scenarios, the challenge emerges from the need to address a combinatorial issue, optimizing resource allocation, transmission mode selection, and symbol lengths simultaneously across the time-frequency axis. This task is compounded by the high degree of complexity and a significant number of unknown variables, making it a formidable challenge for conventional optimization techniques to solve effectively. In response, we propose a deep reinforcement learning-based method that in each time step allocates users to each resource units, determines the communication transmission mode, and selects data size according to communication environment and users' packet states. Extensive experiments demonstrate superior performance over conventional schemes in terms of semantic transmission performance.

Index Terms—Semantic communication, Resource allocation, Deep reinforcement learning, Semantic rate, Mode selection

I. INTRODUCTION

In beyond 5G and 6G, wireless communication demands serving much more user equipments (UEs) with larger amounts of data, resulting in the challenge of a shortage in the frequency spectrum [1], [2]. However, traditional wireless communication has been primarily focused on the transmission and reception of data without comprehending its actual content [3], [4]. As a result, the amount of data that can be transmitted is strictly limited by the frequency spectrum in use.

To address the frequency spectrum shortage problem in conventional communication, task-oriented semantic communication, which can surpass the Shannon capacity in terms of performing specific tasks, has been proposed and is actively under research [3], [5]–[10]. Semantic communication extracts, compresses, and transmits features relevant to the intended task from data, rather than transmitting the raw data itself. Thus, semantic communication employs lossy data compression, but it excels in the realm of task performance efficiency [11].

In the field of text transmission, semantic communication models like DeepSC [11] have demonstrated excellent performance. However, they maintain a fixed transmission symbol size regardless of channel state information (CSI), analogous to keeping the coding rate and modulation fixed in conventional communication. To take into account the benefits of channel diversity, a resource allocation (RA) model that combines channel assignment and transmission volume control of semantic symbols was proposed [12]. Specifically, they defined the spectral efficiency in the realm of semantic communication when transmitting infinite sentences over very long transmission times [12]–[14]. However, this assumption does not align with real-world scenarios, where user traffic tends to be unsaturated, meaning that transmission time and packet lengths are bounded by strict limitations [15].

This paper goes beyond by addressing the joint RA and mode selection (MS) problem in unsaturated traffic scenarios, where UEs participate in uplink communication while holding data of different sizes and numbers. The main contributions are as follows:

- Building on the definition of semantic spectral efficiency in a long-term perspective, we propose a short-term semantic transmission rate (SR) to evaluate the data transmission rate in unsaturated traffic conditions. The SR reflects more realistic communication scenarios, where the the frame length is strictly limited the length of data varies.
- Under the definition of SR, the performance superiority between bit communication and semantic communication changes depending on various signal-to-noise ratios (SNRs) and data sizes. Therefore, we propose a joint RA and MS problem that dynamically allocates UEs into resource units (RUs) in the frequency domain, adaptively selects transmission mode between bit and semantic communication, and determines the number of transmitted semantic symbols for semantic communication.
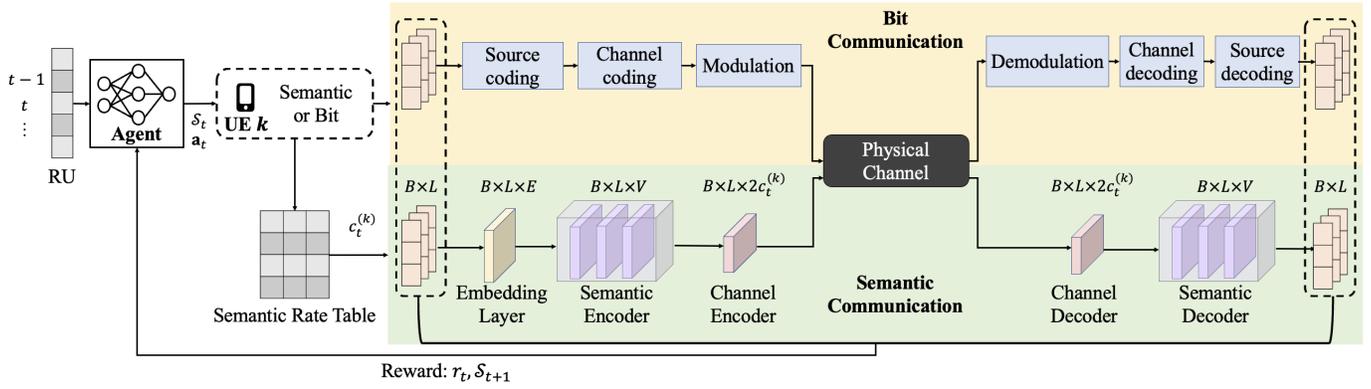- To solve the proposed RA and MS optimization problem

Fig. 1. The proposed deep reinforcement learning-based RA and MS protocol

while considering both UEs' SNR and data size, which is an intractable problem due to its combinatorial aspect [16], we propose an algorithm based on deep reinforcement learning (DRL), which has proven to be a powerful tool for solving complex resource management problems in recent year [5], [17], [18].

As a case study, we evaluate the proposed DRL-based RA and MS algorithm in the field of text transmission. Our results demonstrate that the proposed DRL-based RA and MS algorithm can achieve superior performance in terms of sentence similarity [11], [12], [19], [20] over various conventional schemes such as DeepSC and bit communication.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. Scenario

We consider a scenario in which a base station (BS) communicates with K UEs. Given the CSI and sentences to transmit of the UEs, the BS allocates each UE to N RUs while also selecting the optimal transmission mode, which could be either conventional bit or semantic communication. Additionally, if the BS decides to serve UE with semantic communication, it needs to determine the number of transmitted semantic symbols. The primary objective of the RA and MS process is to maximize task-specific performance metrics within the predefined packet length for all UEs. The RA and MS process is shown in Fig. 1.

B. Wireless Communication Model

We define $a_{n,k}$ as a binary RU assignment variable such that $a_{n,k} = 1$ if the k-th UE is allocated on n-th RU, and $a_{n,k} = 0$ otherwise. Then, we can represent the constraints on the RA as follows:

$$\sum_{n \in \mathcal{N}} a_{n,k} \leq 1, \quad \forall k \in \mathcal{K} \tag{1a}$$

$$\sum_{k \in \mathcal{K}} a_{n,k} \leq 1, \quad \forall n \in \mathcal{N} \tag{1b}$$

where $\mathcal{N} = \{0, 1, \ldots, N-1\}$ and $\mathcal{K} = \{0, 1, \ldots, K-1\}$. Constraint (1a) indicates the unique user assignment along all

RUs. Constraint (1b) imposes the restriction by which each user can only occupy at most one channel.

Let $h_{n,k} \in \mathbb{C}$ denote the uplink communication channel between the BS and the k-th UE on the n-th RU. Then, the SNR for the k-th UE on the RU n is given by $\Gamma_{n,k} = P_{n,k} |h_{n,k}|^2 / \sigma^2$. where $P_{n,k}$ is the transmit power of the k-th UE on the RU n, and σ^2 is the noise variance.

C. Text Transmission Performance

Many researchers rely on the specific yet well-developed large language model, known as bi-directional encoder representations from transformers (BERT) [21], to measure how accurate the semantic information is transmitted in text transmission field [11], [12], [19], [20]. In this paper, we adopt the calculate sentence similarity [12], which is defined by

$$F(\mathbf{s}, \hat{\mathbf{s}}) = \frac{\mathbf{B}(\mathbf{s})\mathbf{B}(\hat{\mathbf{s}})^{\mathsf{T}}}{\|\mathbf{B}(\mathbf{s})\| \|\mathbf{B}(\hat{\mathbf{s}})\|}, \tag{2}$$

where $\mathbf{B}(\mathbf{s})$ represents the output embedding vector using the BERT model for a sentence $\mathbf{s}$. We leverage a pre-trained BERT model to compute the sentence similarity. Note that from the similarity definition in (2), we have $0 \leq F(\mathbf{s}, \hat{\mathbf{s}}) \leq 1$, with 1 indicating the highest similarity and 0 indicating no relationship between two sentences.

D. Definition of Semantic Rate

With the definition of sentence similarity, SR is proposed in [12] for measuring the semantic information transmission rate using BERT model. However, unlike the conventional approach, which calculates the average value of SR over infinite frame length when sending a large amount of data, in real communication environments, each user transmits limited data of different sizes. Furthermore, all users must transmit data within a predetermined frame length to synchronize the uplink transmission. To address these practical issues, we newly define the SR in this paper.

Let $\mathcal{D}_k = \{\mathbf{s}_{j,k} = [w_{j,k,0}, w_{j,k,1}, \ldots, w_{j,k,L_{j,k}-1}]\}_{j=0}^{D_k-1}$ denote the text dataset for the k-th UE with size D_k, where $\mathbf{s}_{j,k}$ is the j-th sentence with length $L_{j,k}$ and $w_{j,k,l}$ is the l-th word of the j-th sentence of the k-th UE. In addition, one can define

2

the amount of semantic information of $\mathbf{s}_{j,k}$ as $I_{j,k}$ (suts). Each sentence is transmitted via either bit communication or semantic communication, as shown in Fig. 1. We denote $m_{n,k}$ as the binary transmission mode variable of the k-th UE on the n-th RU such that $m_{n,k} = 0$ represents bit communication while $m_{n,k} = 1$ means semantic communication.

In bit communication, the transmitter protects information from impairments such as noise or distortion by performing rate adaptation through source coding and channel coding based on the current SNR $\Gamma_{n,k}$. In the case of semantic communication, successful transmission of semantic information is guaranteed by extracting semantic information and compressing the sentence length to $c_{n,k}$ according to the SNR $\Gamma_{n,k}$ through semantic encoding and channel encoding. The encoded symbol stream then can be represented by

$$\mathbf{x} = \begin{cases} C_{\mathrm{bc}}(\mathbf{s}; \Gamma_{n,k}, m_{n,k}), & \text{if } m_{n,k} = 0, \\ C_{\mathrm{sc}}(\mathbf{s}; \Gamma_{n,k}, c_{n,k}, m_{n,k}, \boldsymbol{\beta}), & \text{if } m_{n,k} = 1. \end{cases} \quad (3)$$

where C_{sc} includes channel encoding, semantic encoding, while C_{bc} includes channel encoding, source encoding, and modulation, $\boldsymbol{\beta}$ is the parameter set of semantic and channel encoder networks. If $\mathbf{x}$ is sent, the signal received at the receiver will be $\mathbf{y} = h\mathbf{x} + \mathbf{z}$, where $\mathbf{z}$ is the additive white Gaussian noise (AWGN) that follows $\mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$. With the received signal, the decoded sentence can be represented as

$$\hat{\mathbf{s}} = \begin{cases} C_{\mathrm{bc}}^{-1}(\mathbf{y}; \Gamma_{n,k}, m_{n,k}), & \text{if } m_{n,k} = 0, \\ C_{\mathrm{sc}}^{-1}(\mathbf{y}; \Gamma_{n,k}, c_{n,k}, m_{n,k}, \boldsymbol{\beta}), & \text{if } m_{n,k} = 1, \end{cases} \quad (4)$$

where inverse operation for C means the reverse process of C. Finally, the SR (suts/s) on n-th RU for k-th UE is defined by

$$\phi(\mathcal{D}_k; \Gamma_{n,k}, c_{n,k}, m_{n,k}) = \frac{\sum_{j=0}^{D_k-1} WI_{j,k} \cdot F(\mathbf{s}_{j,k}, \hat{\mathbf{s}}_{j,k})}{L_{\mathrm{frame}}}, \quad (5)$$

where W is the bandwidth and L_{frame} is the frame length.

Note that the sentence similarity heavily depends on the design of C_{sc} and C_{bc}. In bit communication, the design of C_{bc} is standardized according to SNR Γ. Then, it must be satisfied that $\sum_{j=0}^{D_k-1} \hat{L}_{j,k} \leq L_{\mathrm{frame}}$ where $\hat{L}_{j,k}$ is the length of $C_{\mathrm{bc}}(\mathbf{s}_{j,k}; \Gamma_{n,k})$. In semantic communication, the optimal channel coding dimension with respect to SNR has not been thoroughly surveyed. Thus, we define the channel coding dimension of semantic communication for the n-th RU for the k-th UE as $c_{n,k}$. Then, semantic communication transmits each word by packing it with a size of $c_{n,k}$. We determine this value to regulate the number of transmitted semantic symbols. Similar to the approach in bit communication, it is essential to satisfy the condition $\sum_{j=0}^{D_k-1} c_{n,k}L_{j,k} \leq L_{\mathrm{frame}}$ for the k-th UE on the n-th RU.

E. Problem Formulation

From (1) and (5), we formulate the joint RA and MS optimization problem that maximizes sum of SR (S-SR) as
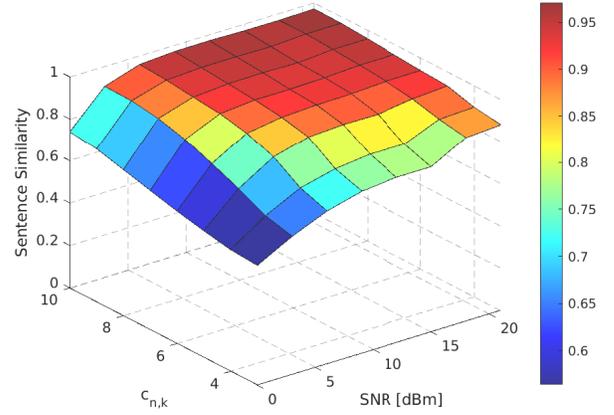

Fig. 2. Semantic rate table according to SNR and data size $c_{n,k}$.

follows:

$$\max_{\mathbf{a},\mathbf{c},\mathbf{m}} \quad \Phi = \sum_{n \in \mathcal{N}} \sum_{k \in \mathcal{K}} a_{n,k}\phi(\mathcal{D}_k; \Gamma_{n,k}, c_{n,k}, m_{n,k}), \quad (6a)$$

$$\text{s.t.} \quad (1a), (1b) \quad (6b)$$

$$\sum_{j \in \mathcal{D}_k} c_{n,k}L_{j,k} \leq L_{\mathrm{frame}}, \forall n \in \mathcal{N}, \forall k \in \mathcal{K}, \quad (6c)$$

$$\sum_{j \in \mathcal{D}_k} \hat{L}_{j,k} \leq L_{\mathrm{frame}}, \forall n \in \mathcal{N}, \forall k \in \mathcal{K}, \quad (6d)$$

$$c_{n,k} \in \mathbb{N}, \forall n \in \mathcal{N}, \forall k \in \mathcal{K}, \quad (6e)$$

$$a_{n,k}, m_{n,k} \in \{0, 1\}, \forall n \in \mathcal{N}, \forall k \in \mathcal{K}, \quad (6f)$$

where $\mathbf{a}$, $\mathbf{c}$, and $\mathbf{m}$ are the set of all variable $a_{n,k}$, $c_{n,k}$, and $m_{n,k}$ for $n \in \mathcal{N}$ and $k \in \mathcal{K}$, respectively. Clearly, due to its nonconcave aspect, it is intractable to solve the RA and MS optimization problem [16].

III. PROPOSED DRL-BASED RA OPTIMIZATION

A. Proposed DRL structure

We propose a DRL structure consisting of an agent, which performs RA and MS, based on the SNR and the data size. If the allocated UE decides to utilize semantic communication, the dimension of channel encoder and decoder $c_{n,k}$, i.e., the number of symbols for each word is selected to maximize the S-SR Φ in (6). We obtain the solution by precomputing the Φ for all possible $c_{n,k}$ and organizing the results into an SR table, as shown in Fig. 2. In the case where the agent chooses bit communication for data transmission, the sentence is conveyed using the conventional bit communication protocol.

B. Definitions of Parameters in DRL

Here, we define the result of RA and MS, whether it's bit communication or semantic communication, as an action. The BS selects actions corresponding to each RU index at each time step based on the current state. Therefore, one can set $t \in \mathcal{N}$. Then, the state space, action space, and reward functions of the agent are defined below.

State Space: The state includes the CSI and dataset to

transmit of the UEs, which is defined as $\tilde{\mathbf{s}}_{n,k} = \{\Gamma_{n,k}, \mathcal{D}_k\}$. Additionally, the initial state for all RUs and all UEs is defined as $\mathcal{S}_0 = \bigcup_{n \in \mathcal{N}} \bigcup_{k \in \mathcal{K}} \tilde{\mathbf{s}}_{n,k}$. When the k-th UE is selected as an action during the DRL procedure, we set the $\Gamma_{n,k} = -1$ for all n to mark it as an unavailable option.

Action Space: The action is defined by $\mathbf{a}_t \in \mathcal{A}$, which represents the result of RA and MS on the t-th RU. Thus, we can represent the action as $\mathbf{a}_t = \{(k, m_{t,k}) | a_{t,k} = 1, \forall k \in \mathcal{K}\}$.

Reward Function: We define the reward function of the agents as $r_t = \sum_{k \in \mathcal{K}} a_{t,k} \phi_{t,k}$.

C. DRL Training Process

Initialization: We introduce the Deep Q-network (DQN) [22] as the learning framework of the agent. Thus, we utilize a parameter θ that defines an action-value function $Q(\mathcal{S}, \mathbf{a}; \theta)$ for the agent. In addition, we initialize replay memories $\mathcal{E}$ for the agent to capacity E.

Experience collection: At each time step t, the agent iteratively collects experience by selecting the actions. Each action is drawn in an epsilon-greedy fashion with linear decay, i.e., $\epsilon(e) = \max\{1 - e/Z, 0.01\}$, where Z is the decaying rate constant, and e is the episode step. The agent first selects a random action $\mathbf{a}_t$ with probability $\epsilon(e)$ or selects $\mathbf{a}_t = \arg\max_{\mathbf{a}} Q(\mathcal{S}_t, \mathbf{a}; \theta)$, otherwise. The agent stores transition at each time-step $(\mathcal{S}_t, \mathbf{a}_t, r_t, \mathcal{S}_{t+1})$ in $\mathcal{E}$.

Updating model parameters: With the stored experiences in the replay memories, the agent updates learning parameters, θ. In the case of θ, the agent samples random mini-batch of B transitions $(\mathcal{S}_j, \mathbf{a}_j, r_j, \mathcal{S}_{j+1})$ from $\mathcal{E}$. We set $y_j = r_j$ if $\mathcal{S}_{j+1}$ is a terminal state or $y_j = r_j + \gamma \max_{\mathbf{a}} Q(\mathcal{S}_{j+1}, \mathbf{a}; \theta)$, otherwise. Then, we get the training loss $J(\theta) = \sum_j (y_j - Q(\mathcal{S}_j, \mathbf{a}_j; \theta))^2 / B$. The agent performs a gradient descent step on $J(\theta)$ and updates θ.

IV. SIMULATION RESULTS

To evaluate the performance of the proposed DRL-based RA and MS algorithm under scenario where both semantic and bit communication are available, we have conducted simulations with the proposed DRL algorithm and baseline methods.

A. Experimental Setup

We adopt the datasets named European parliament proceedings parallel Corpus [23]. It includes around 2.0 million sentences and 53 million words. We sample 200,000 sentence from the datasets and divides them into a training dataset and a test dataset. In addition, we collect the sentence with the length of 4 to 30.

We examine baselines in RA methods and communication types. In RA methods, we investigate two methods: random and max-SNR [24], [25]. The random method chooses UEs regardless of SNR and data size while the max-SNR prioritizes UEs based sorely on SNR. In terms of communication types, semantic communication-based and bit communication-based systems are considered. In the semantic communication-based system, we refer to it as "DeepSC" [11] when the channel

	Random + BC	Random + SC	Max-SNR + BC	Max-SNR + SC
S-SR	1,776	2,464	2,169	2,498

	DRL + BC	DRL +SC	Proposed	
S-SR	2,374	3,091	3,113	

coding dimension is fixed at eight and "Semantic" when the channel coding dimension is optimized according to SNR. In the bit communication-based system, we adopt Huffman coding as a source coding and low-density parity check (LDPC) as a channel coding. We follow the 5G standard in terms of coding rate and modulation and [26] to get modulation and coding scheme index according to SNR.

We set the bandwidth $W = 180$ kHz and the frame length $L_{\text{frame}} = 1024$. We assume that the amounts of semantic information of all sentence are equivalent, i.e., $I_{j,k} = 1$, for all (j, k). In all experiments, the number of users is set to 5, and the number of resource blocks is fixed at 5 3.

B. Result Analysis

We first conduct a comparative analysis between the conventional and proposed schemes in a scenario involving randomly varying data sizes ranging from 1 to 10 and SNR levels distributed uniformly between 3 dB and 15 dB, which is presented in Table I. From the result, we conclude that the proposed DRL-based method achieves the highest S-SR over all conventional methods.

In the following, we assess the S-SR of the bit communication only, semantic communication only, and proposed schemes with the DRL method across different number of sentences, as shown in Fig 3, to ascertain the influence of MS. When UE sends a relatively small number of sentences, it can achieve higher S-SR with bit communication because it can reliably send within the frame length. However, when sending a large number of sentences, compressing sentences into semantic information and transmitting them proves to be much more effective. Thus, the proposed method that allows users to flexibly choose between two modes of bit and semantic communication based on the data size achieves the highest S-SR compared to the other two communication techniques.

Fig. 4 shows the S-SR of the proposed and conventional methods along with different SNRs. In a low SNR environment, the S-SR of bit communication deteriorates due to the failure of complete restoration of data. In contrast, semantic communication provides a significantly better S-SR in low SNR conditions; however, it shows a slightly lower S-SR compared to bit communication when the SNR exceeds or equals 9 dB. While semantic communication experiences some loss in S-SR performance due to lossy compression, bit communication achieves better performance in high SNR
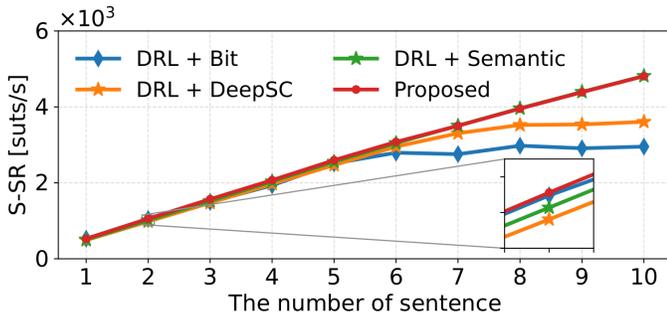
Fig. 3. The S-SR comparison of the proposed and conventional methods with respect to the number of sentences. AWGN channel with a uniform distribution of SNR from 3 dB to 15 dB is considered.
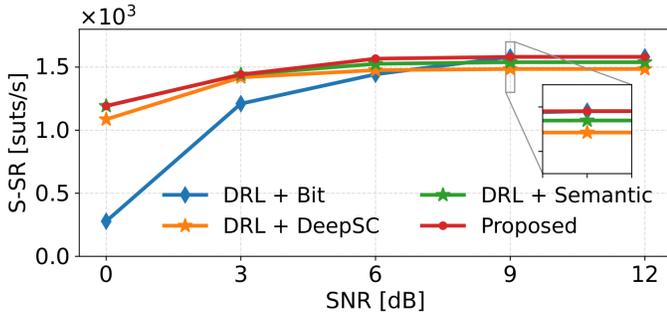

Fig. 4. The S-SR comparison of the proposed and conventional methods with respect to SNR. The number of sentences all UE poses is two.

environments due to its precise data reconstruction. However, the proposed method outperforms all baseline methods across the entire SNR range by adaptively selecting the optimal transmission mode.

V. CONCLUSION

We proposed a DRL-based algorithm for optimizing joint RA and MS, effectively allocating UEs to RUs and determining the optimal transmission mode between semantic and bit-based communication. Our approach dynamically adjusts the number of transmitted semantic symbols, addressing the complexity of unsaturated traffic conditions. Experiments show superior performance over traditional schemes like DeepSC and bit communication, particularly in terms of sentence similarity. Future work will focus on refining the definition and quantification of semantic information in sentence data and expanding the framework to more complex network scenarios. This will enhance the system's adaptability and efficiency, paving the way for more intelligent semantic communication solutions in evolving wireless networks.

VI. ACKNOWLEDGEMENT

REFERENCES

[1] Deniz Gündž et al., "Beyond transmitting bits: Context, semantics, and task-oriented communications," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 5–41, 2023.

[2] Yalin E. Sagduyu, Sennur Ulukus, and Aylin Yener, "Task-oriented communications for nextG: End-to-end deep learning and ai security aspects," *IEEE Wireless Commun.*, vol. 30, no. 3, pp. 52–60, 2023.

[3] Wanting Yang et al., "Semantic communications for future internet: Fundamentals, applications, and challenges," *IEEE Commun. Surv. Tut.*, vol. 25, no. 1, pp. 213–250, 2023.

[4] Christina Chaccour, Walid Saad, Mérouane Debbah, Zhu Han, and H. Vincent Poor, "Less data, more knowledge: Building next generation semantic communication networks," *IEEE Commun. Surveys Tuts.*, pp. 1–1, 2024.

[5] Haijun Zhang et al., "DRL-driven dynamic resource allocation for task-oriented semantic communication," *IEEE Trans. Commun.*, vol. 71, no. 7, pp. 3992–4004, 2023.

[6] Hongwei Zhang et al., "Deep learning-enabled semantic communication systems with task-unaware transmitter and dynamic data," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 170–185, 2023.

[7] Ke Yang et al., "WITT: A wireless image transmission transformer for semantic communications," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2023, pp. 1–5.

[8] Huiqiang Xie, Zhijin Qin, and Geoffrey Ye Li, "Semantic communication with memory," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2658–2669, 2023.

[9] Guangming Shi et al., "From semantic communication to semantic-aware networking: model, architecture, and open problems," *IEEE Commun. Magazine*, vol. 59, no. 8, pp. 44–50, 2021.

[10] Xuewen Luo, Hsiao-Hwa Chen, and Qing Guo, "Semantic communications: Overview, open issues, and future research directions," *IEEE Wireless Commun.*, vol. 29, no. 1, pp. 210–219, 2022.

[11] Huiqiang Xie, Zhijin Qin, Geoffrey Ye Li, and Biing-Hwang Juang, "Deep learning enabled semantic communication systems," *IEEE Trans. Signal Process.*, vol. 69, pp. 2663–2675, 2021.

[12] Lei Yan, Zhijin Qin, Rui Zhang, Yongzhao Li, and Geoffrey Ye Li, "Resource allocation for text semantic communications," *IEEE Wireless Commun. Lett.*, vol. 11, no. 7, pp. 1394–1398, 2022.

[13] Xidong Mu et al., "Heterogeneous semantic and bit communications: A semi-noma scheme," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 155–169, 2023.

[14] Xidong Mu and Yuanwei Liu, "Exploiting semantic communication for non-orthogonal multiple access," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 8, pp. 2563–2576, 2023.

[15] Hyeonho Noh, Harim Lee, and Hyun Jong Yang, "Joint optimization on uplink OFDMA and MU-MIMO for IEEE 802.11ax: Deep hierarchical reinforcement learning approach," *IEEE Commun. Lett.*, pp. 1–5, 2024.

[16] Nan Zhao et al., "Deep reinforcement learning for user association and resource allocation in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5141–5152, 2019.

[17] Haijun Zhang et al., "Power control based on deep reinforcement learning for spectrum sharing," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 4209–4219, 2020.

[18] Shaoyang Wang et al., "Joint resource management for MC-NOMA: A deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5672–5688, 2021.

[19] Zi Qin Liew et al., "Economics of semantic communication system in wireless powered internet of things," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2022, pp. 8637–8641.

[20] Tianxiao Han et al., "Semantic-preserved communication system for highly efficient speech transmission," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 245–259, 2023.

[21] Matthew E. Peters et al., "Deep contextualized word representations," in *Proc. North Amer. Chapter Assoc. Comput. Linguistics: Hum. Lang. Tech.*, New Orleans, Louisiana, June 2018, pp. 2227–2237.

[22] Volodymyr Mnih et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

5

[23] Philipp Koehn, "Europarl: A parallel corpus for statistical machine translation," in *MT summit*, 2005, pp. 79–86.

[24] Shengli Liu et al., "Joint user association and resource allocation for wireless hierarchical federated learning with IID and non-IID data," *IEEE Trans. Wireless Commun.*, vol. 21, no. 10, pp. 7852–7866, 2022.

[25] Amin Abdel Khalek, Constantine Caramanis, and Robert W. Heath, "Delay-constrained video transmission: Quality-driven resource allocation and scheduling," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 1, pp. 60–75, 2015.

[26] Eunmi Chu, Janghyuk Yoon, and Bang Chul Jung, "A novel link-to-system mapping technique based on machine learning for 5G/IoT wireless networks," *Sensors*, vol. 19, no. 5, pp. 1196, 2019.