

RESEARCH

Open Access



# A DRL-based resource allocation for IRS-enhanced semantic spectrum sharing networks

Yingzheng Zhang<sup>1,2</sup>, Jufang Li<sup>1</sup>, Guangchen Mu<sup>3</sup> and Xiaoyu Chen<sup>4\*</sup>

\*Correspondence:  
zyz@hait.edu.cn

<sup>1</sup> School of Electronic Information Engineering, Henan Institute of Technology, XinXiang 453003, China

<sup>2</sup> Henan Industrial Equipment Intelligent Engineering Technology Research Center, Luoyang, China

<sup>3</sup> Department of Science, Henan Institute of Technology, XinXiang 453003, China

<sup>4</sup> Engineering Technology Center, Henan Institute of Technology, XinXiang 453003, China

## Abstract

Semantic communication and spectrum sharing are pivotal technologies in addressing the perennial challenge of scarce spectrum resources for the sixth-generation (6G) communication networks. Notably, scant attention has been devoted to investigating semantic resource allocation within spectrum sharing semantic communication networks, thereby constraining the full exploitation of spectrum efficiency. To mitigate interference issues between primary users and secondary users while augmenting legitimate signal strength, the introduction of Intelligent Reflective Surfaces (IRS) emerges as a salient solution. In this study, we delve into the intricacies of resource allocation for IRS-enhanced semantic spectrum sharing networks. Our focal point is the maximization of semantic spectral efficiency (S-SE) for the secondary semantic network while upholding the minimum quality of service standards for the primary semantic network. This entails the joint optimization of parameters such as semantic symbol allocation, subchannel allocation, reflective coefficients of IRS elements, and beamforming adjustment of secondary base station. Recognizing computational intricacies and interdependence of variables in the non-convex optimization problem formulated, we present a judicious approach: a hybrid intelligent resource allocation approach leveraging dueling double-deep Q networks coupled with the twin-delayed deep deterministic policy. Simulation results unequivocally affirm the efficacy of our proposed resource allocation approach, showcasing its superior performance relative to baseline schemes. Our approach markedly enhances the S-SE of the secondary network, thereby establishing its prowess in advancing the frontiers of semantic spectrum sharing (S-SE).

**Keywords:** Semantic communication, Intelligent reflection surface, Spectrum sharing, Semantic spectral efficiency, Deep reinforcement learning

## 1 Introduction

The problem of spectrum scarcity is further exacerbated by the massive smart devices and massive connections that characterize 6 G wireless communication networks [1–3]. However, it is difficult for the existing conventional communication paradigms to further improve the spectral efficiency, as revealed in the Shannon rate limit [4–6]. Recently, the artificial intelligence (AI)-driven semantic communication paradigm has received much

attention due to the great promise it shows in breaking the Shannon limit and improving spectral efficiency [7]. Specifically, semantic communication is a communication paradigm over the semantics, where the parties communicate the intention. In task-oriented semantic communications, the task-related necessary information is transmitted, while the unnecessary information, e.t. the task-unrelated information, is ignored [8]. Semantic encoding networks are able to utilize the powerful knowledge representation capabilities of machine learning to extract and encode key task-relevant information, greatly avoiding information redundancy [9, 10]. Under the semantic communication paradigm, joint source channel coding techniques have been developed and demonstrate better performance than separated coding [11].

Spectrum sharing technology has been widely used due to its high spectrum utilization efficiency [12]. Specifically, the secondary network is able to share the spectrum of the primary network while minimizing the impact on the primary network [13]. However, under the same active channel, the interference between primary and secondary users (SUs) is unavoidable during the sharing period. This may have an impact on the main network. Recently, spectrum sharing networks assisted by smart reflective surfaces have been widely studied [14]. Intelligent Reflective Surfaces (IRS), a future-proof technology, can enhance target signal strength and attenuate interfering signals in a low-energy way [15, 16]. Specifically, a smart reflective surface consists of a planar array of passive transmitting elements, which can be programmed to adjust the reflection factor of the reflective elements to achieve signal-phase adjustment during the signal reflection process [17].

Due to the superior ability of solving large-scale complex problems and real-time performance [18], deep reinforcement learning (DRL) has recently attracted a lot of attention. This trend of interest reflects the growing recognition of the superior effectiveness of DRL in dealing with complex situations where traditional methods may not perform well. It is likely that the citation refers to an in-depth discussion of specific advances or applications within the field, contributing to a more comprehensive understanding of the importance and impact of DRL in contemporary problem solving paradigms. The advantage of DRL over traditional convex optimization methods is its ability to handle high-dimensional, nonlinear, and complex problems, and to adapt to diverse environments and tasks by extracting features and optimization strategies from experience through learning [19]. However, the traditional convex optimization methods usually face problems with low dimensionality and linear structure, which make it difficult to effectively deal with complex real-world scenarios. Deep reinforcement learning is able to learn autonomously and gradually improve its performance through end-to-end learning of neural networks and thus has more advantages when facing variable and uncertain problems in the real world [20]. Hence, it is promising to propose an DRL-based AI-native resource allocation scheme for semantic spectrum sharing networks.

## 1.1 Related works

### 1.1.1 Semantic coding network

A great number of works have paid attention to semantic communication networks. According to the data source modality, the type of semantic coding research can be divided into text [21, 22], image [23, 24], and speech [25]. The authors in [21] exploited

inference rule from the knowledge graph, which aimed to obtain the inexplicable and inflexible of the semantic communication networks. The authors in [24] investigated the unmanned aerial vehicle (UAV) image-sensing-driven semantic communication for a triple-based scene construction. In [25], the authors considered a speech semantic communication network, where speech synthesis at the receiver entails a dedicated process wherein the regeneration of speech signals transpires. This involves inputting the recognized text and speaker information into a neural network module for the purpose of generating the synthesized speech signals. There is little work [26–28] that pays attention to the resource allocation for semantic communication networks. In [26], the authors defined the semantic spectral efficiency (S-SE) for the first time. Then, based on the [26], the authors in [27] further considered the quality of experience of the users. The authors in [28] proposed a novel semantic-bit quantization method and considered an adaptive resource allocation scheme for semantic communications over the physical wireless channels. In [29], the IRS-assisted secure semantic communication network was investigated, while the spectrum sharing technology was not considered for further improvement of spectrum efficiency. Moreover, IRS was used to counter semantic eavesdropping in [29], while IRS has not been used to eliminate inter-user interference and ensure the quality of service of the primary network. As the best knowledge of the authors, there has been little work considering the IRS-assisted semantic communication networks with spectrum sharing.

### **1.1.2 IRS-assisted spectrum sharing networks**

Given that IRS employs passive reflective elements devoid of signaling methods on received signals, the IRS demonstrates a capacity to reshape signals with minimal overhead [30]. Notably, a substantial body of research has concentrated on unraveling the extensive potential inherent in IRS technology [31–36]. The authors in [32] considered the IRS-assisted secure spectrum sharing network, where the IRS can significantly enhance the legal signal and suppress the eavesdropping at the eavesdroppers. The authors in [34] introduced the multiple IRS for wide convergence and the secure performance improvement of the secondary network. Nevertheless, it is noteworthy that only a limited number of studies have delved into the augmentation of semantic spectrum efficiency (S-SE) for task performance in the context of semantic spectrum sharing networks operating under low signal-to-noise ratios. Furthermore, a majority of the previously mentioned investigations [31, 33–36] have predominantly relied on conventional methodologies grounded in convex optimization, a less time-efficient paradigm when confronted with the complexities of large-scale connections. Consequently, there exists a compelling imperative to delve into and develop time-efficient intelligent resource allocation schemes [37].

### **1.1.3 Intelligent resource allocation approach**

DRL-based resource allocation approaches have been widely used due to its powerful computational ability [38–43]. The authors in [39] proposed a DRL-based intelligent resource allocation scheme, which can rapidly solve the tricky non-convex problem. The authors in [40] proposed a hybrid intelligent resource allocation to address the formulated optimization problem for CR networks. In the work [41], the authors intricately devised

DRL-based schemes with the explicit goal of diminishing output dimensionality, elevating learning efficiency, and formulating a judicious resource allocation policy. Moreover, the DRL-based resource allocation scheme was used for mobile edge computing in railway Internet of Things (RIoT) networks to jointly optimize the subcarrier assignment, offloading ratio, power allocation, and computation resource allocation. Note that the computational complexity comparison conducted in the [34], the DRL-based scheme performed better time efficiency and achieved the close performance compared to the traditional mathematical methods. Further considering the semantic resource allocation in the UAV-assisted semantic communication system, DRL-based resource allocation scheme and intelligent trajectory planning scheme were proposed in [43].

## 1.2 Motivations and contributions

We investigate IRS-assisted semantic spectrum sharing networks in this paper to optimize the S-SE of the secondary network. The noteworthy contributions of this paper are as follows.

- In this paper, we explore for the first time IRS-assisted semantic spectrum sharing communication networks. Specifically, IRS is able to simultaneously enhance the performance of semantic tasks in the secondary network while minimizing the interference of the secondary network to the primary network. Semantic spectrum efficiency is utilized to evaluate the secondary network spectrum efficiency. The S-SE is maximized by jointly optimizing the allocation of subchannel, semantic symbols, IRS reflection array elements, and the beamforming of secondary base station.
- To solve our proposed complex non-convex problem, based on dueling double-deep Q networks (D3QN) twin-delayed deep deterministic policy (TD3), an intelligent resource allocation scheme is introduced for semantic spectrum sharing networks. Specifically, the discrete action space, i.e., semantic symbol number allocation and subchannel allocation, is handled by utilizing D3QN. TD3, on the other hand, can effectively address continuous action spaces, i.e., tuning the transmit beam and optimizing the IRS reflection elements. Such a hybrid algorithm design can fully utilize the powerful Q-value computation capability of D3QN and the powerful exploration capability of TD3 in high-dimensional space.
- Simulation results demonstrate that our proposed IRS-assisted semantic spectrum sharing network can significantly enhance the S-SE of the secondary network while guaranteeing the communication quality of service in the primary network compared to the benchmark scheme lacking IRS and the conventional communication scheme. In addition, we demonstrate that our proposed hybrid intelligent resource allocation scheme can reach convergence in a short period of time, proving its powerful exploration capability.

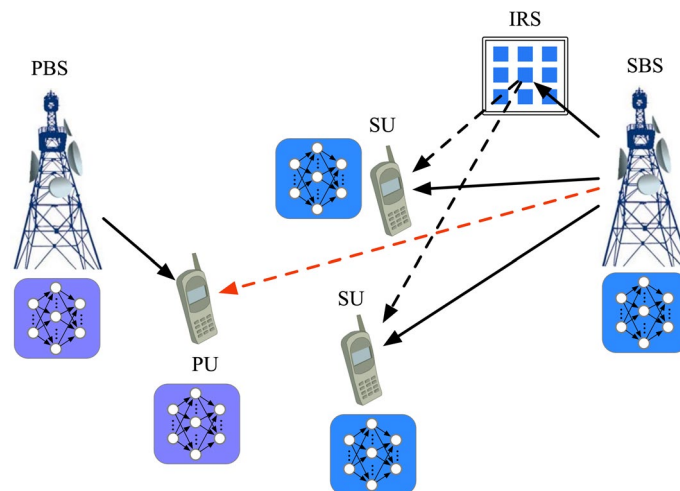
The rest of the paper is organized as follows: Section II presents the IRS-assisted semantic spectrum sharing network model. Section III presents the problem of maximizing the semantic spectral efficiency of the sub-network. Section IV presents the hybrid resource allocation scheme. Section V demonstrates the simulation results. Finally, Section VI gives the conclusion of this paper.

## 2 System model and optimization problem

### 2.1 Semantic coding framework

Here, we consider an IRS-enhanced semantic spectrum sharing network designed for the text recovery task, featuring a transmitter and a receiver and a IRS (Fig. 1). The transmitter incorporates a semantic encoder and channel encoder for extracting features and semantically encoding text sources, while the receiver includes the channel decoder and semantic decoder to decode semantics and recover mission-critical information. Let  $\mathbf{S} = [s_1, \dots, s_i, \dots, s_l]$  represent a sentence, where  $l$  is the sentence length,  $s_i$  is the  $i$ -th word. We employ DeepSCs with multi-level semantic symbol outputs for higher transmission efficiency. Specifically, the different DeepSC networks can produce the semantic information with different length. Let  $\mathcal{O} = \{1, \dots, O\}$  be the set of different DeepSC set.

The encoding process for the input sequence  $\mathbf{s}$  involves the semantic encoder, yielding semantic information  $\mathbf{P}$  through the expression  $\mathbf{P} = E_{\rho_o}(\mathbf{s})$ , where  $E_{\rho_o}(\cdot)$  represents the encoder  $o$  characterized by parameters  $\rho_o$ . This encoded information, denoted as  $\mathbf{X} = C_\alpha(\mathbf{P})$ , is then passed through channel encoder  $C_\alpha(\cdot)$ , defined by parameter  $\alpha$ . At the receiver, the received signal is expressed as  $\mathbf{Y} = \mathbf{h}\mathbf{X} + \mathbf{n}$ , where  $\mathbf{h}$  represents channel coefficients and  $\mathbf{n}$  denotes the background noise. Decoding of the received symbols is carried out by the channel decoder, represented as  $\mathbf{X}' = C_\beta^{-1}(\mathbf{Y})$ . Finally, semantic information  $\mathbf{X}'$  is reconstructed at the receiver through the decoder, denoted as  $\mathbf{m} = S_{\sigma_o}^{-1}(\mathbf{X}')$ , where  $S_{\sigma_o}^{-1}(\cdot)$  signifies the inverse operation of the semantic decoder with parameters  $\sigma_o$ . Let  $\zeta_{k,o}$  denote whether user  $k$  select  $o$ -th DeepSC. If user  $k$  select  $o$ -th DeepSC, then  $\zeta_{k,o} = 1$ ; otherwise,  $\zeta_{k,o} = 0$ .



**Fig. 1** The proposed semantic spectrum sharing communication

The received signal, denoted as  $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ , is succinctly expressed by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \widehat{\mathbf{H}}\mathbf{x} + \mathbf{n}. \tag{1}$$

Here,  $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$  denotes the channel matrix,  $\widehat{\mathbf{H}} \in \mathbb{C}^{N_r \times N_t}$  denotes the IRS reflective coefficients, and  $\mathbf{n} \in \mathbb{C}^{N_r \times 1} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$  is additive white Gaussian noise (AWGN), where  $\sigma^2$  denotes noise variance and  $\mathbf{I}$  denotes identity matrix.

### 2.2 IRS-enhanced semantic spectrum sharing communication

We introduce an IRS, which is equipped with  $E$  reflective elements, strategically positions itself to improve the transmission efficiency of the secondary network. Let  $\Theta = \text{diag}(\varpi_1 e^{j\phi_1}, \varpi_2 e^{j\phi_2}, \dots, \varpi_n e^{j\phi_e}) \in \mathbb{C}^{E \times E}$  denote the IRS phase coefficients, where  $\varpi_e \in [0, 1]$  signifies IRS amplitude coefficients, while  $\phi_e \in [0, 2\pi]$  symbolizes IRS phase coefficients. Let  $\mathcal{D} = \{1, \dots, D\}$  be primary users (PUs), and  $\mathcal{K} = \{1, \dots, K\}$  be SUs. Let  $\mathbf{h}_{p,d} \in \mathbb{C}^{M \times 1}$ ,  $\mathbf{G}_{p,r} \in \mathbb{C}^{E \times M}$  and  $\mathbf{g}_d \in \mathbb{C}^{E \times 1}$  be the channel coefficients from base station (BS) to PU  $d$ , from the primary base station (PBS) to IRS and from IRS to PU  $d$ , respectively. Similarly, let  $\mathbf{h}_{s,k} \in \mathbb{C}^{M \times 1}$ ,  $\mathbf{G}_{s,r} \in \mathbb{C}^{E \times M}$ , and  $\mathbf{g}_{r,k} \in \mathbb{C}^{E \times 1}$  be the channel coefficients from BS to SU  $k$ , from SBS to IRS, and from IRS to SU  $k$ , respectively. Let  $\mathcal{C} = \{1, \dots, C\}$  be subchannel set. The channel allocation of  $k$ -th SU is expressed as  $\rho_k = \{\rho_{k,1}, \dots, \rho_{k,c}, \dots, \rho_{k,C}\}$ , where  $\rho_{k,c} \in \{0, 1\}$  represents whether  $k$ -th user uses  $c$ -th subchannel. If  $k$ -th SU occupies  $c$ -th subchannel, then  $\rho_{k,c} = 1$ ; otherwise,  $\rho_{k,c} = 0$ .

Let  $\mathbf{f}_d^p$  be the transmit beamforming of the  $d$ -th PU and  $\mathbf{f}_k^s$  be the transmit beamforming of the  $k$ -th SU. The transmission rate from PBS to  $d$ -th PU is denoted as

$$R_d = \frac{B}{C} \log_2 \left[ 1 + \frac{\left| \left( \mathbf{h}_{p,d}^T + \mathbf{g}_{r,d}^T \Theta \mathbf{G}_{s,r} \right) \mathbf{f}_d^p \right|^2}{\Gamma_{p,d} + \sigma_d^2} \right], \tag{2}$$

where  $B$  denotes the bandwidth. The interference  $\Gamma_{p,d}$  can be given by

$$\Gamma_{p,d} = \sum_{k=1}^K \delta_{kd} \left| \left( \mathbf{h}_{s,d}^T + \mathbf{g}_{r,d}^T \Theta \mathbf{G}_{p,r} \right) \mathbf{f}_k^s \right|^2, \tag{3}$$

where  $\mathbf{h}_{s,d}$  denotes the subchannel from the PBS to the  $d$ -th user. The  $\delta_{kd}$  denotes whether  $d$ -th PU and  $k$ -th SU share the same channel. If the PU  $d$  and the SU  $k$  share the same subchannel, then  $\delta_{kd} = 1$ ; otherwise,  $\delta_{kd} = 0$ . The transmission rate from BS to  $k$ -th user is expressed by

$$R_k = \frac{B}{C} \log_2 \left[ 1 + \frac{\left| \left( \mathbf{h}_{s,k}^T + \mathbf{g}_{r,k}^T \Theta \mathbf{G}_{s,r} \right) \mathbf{f}_k^s \right|^2}{\Gamma_{s,k} + \sigma_k^2} \right]. \tag{4}$$

The interference  $\Gamma_k^s$  can be given by

$$\Gamma_{s,k} = \sum_{d=1}^D \delta_{dk} \left| \left( \mathbf{h}_{p,k}^T + \mathbf{g}_{r,k}^T \Theta \mathbf{G}_{p,r} \right) \mathbf{f}_d^p \right|^2, \tag{5}$$

where  $\mathbf{h}_{p,k}$  is the subchannel from PBS to  $k$ -th user.

### 2.3 Semantic similarity metrics

Similar to the semantic similarity metric proposed in [22], the Bidirectional Encoder Representations from Transformers (BERT) model is used. BERT is a cutting-edge natural language processing model characterized by its bidirectional attention mechanism, enabling it to capture contextual information in a given sequence. This innovative architecture, based on the transformer model, facilitates a deeper understanding of word relationships and semantic nuances within sentences, leading to superior performance in a wide array of language understanding tasks such as sentiment analysis, named entity recognition, and question answering. BERT's pre-training strategy involves training on large corpora, empowering the model to grasp intricate linguistic patterns and foster transfer learning for downstream applications, making it a pivotal milestone in the field of NLP. The BERT-based semantic similarity can be obtained by

$$\xi = \frac{\mathbf{B}(S)\mathbf{B}(\hat{S})^T}{\|\mathbf{B}(S)\|\|\mathbf{B}(\hat{S})\|}, \quad (6)$$

where  $\mathbf{B}(\cdot)$  is the pretrained BERT model [44] and  $\xi \in [0, 1]$ .

### 2.4 Semantic Spectral Efficiency

Following the principles outlined in [26], the semantic unit (sut) is introduced for semantic information representation. The unit of the rate of semantic transmission (S-R) is sut/s. and the unit of the S-SE is suts/s/Hz. We introduce a text dataset  $\mathcal{D} = \{\mathbf{s}_d\}$ , where  $\mathbf{s}_d$  is the  $d$ -th sentence. We considered a prior known probability of each sentence  $d$  used, expressed by  $p(\mathbf{s}_d)$ . Let  $\mathbf{P} = \sum_{d=1}^D \mathbf{P}_d p(\mathbf{s}_d)$  be the semantic information. Let  $L = \sum_{d=1}^D L_d p(\mathbf{s}_d)$  be the number of words per sentence. Considering a long-term semantic communication processing,  $\mathbf{P}$  and  $L$  are fixed, which are randomly given in this paper. Let  $v_k$  be the semantic symbols used. Hence, the semantic symbols used for sentence representation denoted as  $\Gamma_k = v_k L$ . Consistent with [26], the S-R is equivalent to channel bandwidth, where the S-R of  $k$ -th user over  $c$ -th subchannel is denoted by  $\Omega_{k,c} = \frac{W\mathbf{P}}{\Gamma_k} \xi_{k,c}$ , where  $\xi_{k,c}$  represents the semantic similarity difference. The  $\xi_{k,c}$  is decided by symbol allocation, channel allocation, transmit beamforming of SBS, and the coefficients of the IRS reflective elements. Therefore, the achievable S-SE of  $k$ -th SU over  $c$ -th subchannel is expressed by

$$\Psi_{k,c} = \frac{\Omega_{k,c}}{W} = \frac{\mathbf{P}}{\Gamma_k} \xi_{k,c}. \quad (7)$$

## 3 Problem formulation

The S-SE of the secondary network is maximized by semantic symbol allocation, subchannel allocation, beamforming of SBS, and the coefficients of the IRS reflective elements are jointly optimized. Hence, the problem can be formulated as follows

$$\mathbf{P} : \sum_{k=1}^K \sum_{c=1}^C \rho_{k,c} \max_{\rho, \zeta, \Theta, \mathbf{F}} \Psi_{k,c} \quad (8a)$$



$$\text{s.t. } \rho_{k,c} \in \{0, 1\}, \forall c \in \mathcal{C}, \forall k \in \mathcal{K}, \quad (8b)$$

$$\phi_e \in \{0, 2\pi\}, \forall e \in E, \quad (8c)$$

$$|\Phi_e| = 1, \forall e \in E, \quad (8d)$$

$$\zeta_{k,o} \in \{0, 1\}, \forall k \in \mathcal{K}, \forall o \in \mathcal{O}, \quad (8e)$$

$$\sum_{k=1}^K \sum_{c=1}^C \|\mathbf{f}_{k,c}\|^2 \leq TP_s^{\text{th}}, \quad (8f)$$

$$\xi_d \geq \xi_d^{\text{th}}, \forall d \in \mathcal{D}, \quad (8g)$$

where  $\xi_d^{\text{th}}$  is the minimum task requirements of the PUs, and  $TP_s^{\text{th}}$  represents the upper bound of SBS transmit power. Constraint (8b) limits the subchannel assignment, where each subchannel can only be occupied by one SU. Constraints (8d) and (8e) limit the IRS reflective elements. Constraint (8f) presents the upper bound of SBS transmit power. Constraint (8g) aims to maintain the requirements of QoS of the PUs.

#### 4 Proposed DRL-based resource allocation scheme for semantic spectrum sharing networks

We present our design of a DRL-based resource allocation method for semantic spectrum sharing network. We propose a hybrid intelligent resource allocation method based on D3QN-TD3, which is able to deal with discrete action spaces, including subchannel assignment and semantic symbol assignment, and continuous action spaces, including IRS reflective elements and the transmit beamforming of the SBS, efficiently, as detailed below.

##### 4.1 MDP formulation

The Markov Decision Process (MDP) serves as cornerstones of reinforcement learning theory, providing a structured framework for modeling decision-making problems. In the context of this paper, the optimization problem defined in (8) is initially transformed into an MDP problem, laying the groundwork for employing reinforcement learning algorithms to achieve optimal performance. Within this MDP framework, the environment is conceptualized as the IRS-enhanced semantic spectrum sharing networks, with intelligent agents residing in the control unit of the SBS. The definition of the state space, action space, reward function, and transition probability becomes pivotal in formulating the RL problem specific to our IRS-assisted semantic spectrum sharing communication network. This MDP-based approach enables the reinforcement learning algorithm to navigate the dynamic environment, making informed decisions to enhance the overall performance and effectiveness of the system.

We designate  $\mathcal{S}$  to denote the state space. At time step  $t$ , a state, represented as  $s_t \in \mathcal{S}$ , encapsulates a comprehensive set of information, including subchannel information, selected actions, S-SE ( $\Psi_t$ ), and obtained rewards. To be more explicit, the state  $s_t$  at



time step  $t$  is characterized by the S-SE denoted as  $\Psi_t$ . This holistic representation ensures that the state captures the relevant aspects of the system's history, facilitating the reinforcement learning algorithm's ability to make informed decisions based on the accumulated knowledge of actions, subchannel dynamics, S-SE, and rewards.

$$s_t = \left\{ \mathbf{H}_t, \Psi_t, a_{t-1}, r_t \right\}. \quad (9)$$

The action space of the considered semantic spectrum sharing network is denoted as  $\mathcal{A}$ . Specifically, at time step  $t$ , the semantic symbol allocation, subchannel allocation, the beamforming of SBS, and IRS reflective coefficients are, respectively, represented as  $\rho_t$ ,  $\zeta_t$ ,  $\mathbf{F}_t$  and  $\Theta_t$ . This multi-faceted action space reflects the diverse choices and configurations influencing the IRS-assisted semantic spectrum sharing network's behavior and performance at each time step. The action is expressed by

$$a_t = \left\{ \rho_t, \zeta_t, \mathbf{F}_t, \Theta_t \right\}. \quad (10)$$

The agent's action selection strategy is significantly influenced by the pivotal role played by the reward function. This function establishes the goal of maximizing S-SE and assesses the agent's performance following each iteration step. As a result, the reward function design holds paramount importance in determining the maximized S-SE. The loss function is designed as

$$r = \sum_{k=1}^K \sum_{c=1}^C \rho_{k,c} \Psi_{k,c} + \sum_{d=1}^D (\Psi_d - \Psi_d^{\text{th}}). \quad (11)$$

The policy serves as a crucial component in reinforcement learning, representing the probability with which the agent chooses a specific action  $a$  through current state  $s$ . It encapsulates the agent's strategy and decision-making process as it interacts with its surroundings. The overarching goal of the intelligent agent lies in gaining insights into an optimal resource allocation strategy. In pursuit of this objective, the long-term reward is intricately linked to the action selection process guided by the policy. This reward is a measure of the agent's success in achieving its goals and is essential for shaping the learning process. The formulation of an effective policy becomes pivotal, as it significantly influences the agent's ability to navigate and make informed decisions in the dynamic environment, ultimately contributing to the attainment of superior long-term performance, represented by

$$R = \sum_t \gamma^t [r_{t+1} | s_t, a_t]. \quad (12)$$

In this context, the symbol  $\gamma \in [0, 1)$  represents a discount factor, a pivotal parameter governing the impact of past decision-making steps in the reinforcement learning process. The value of  $\gamma$  operates within the range  $[0, 1)$ , where a larger  $\gamma$  indicates a further consideration of rewards. Conversely, a small  $\gamma$  suggests that the agent places greater emphasis on more recent decisions, tailoring its strategy to prioritize the most immediate and relevant information for optimal decision-making within the dynamic environment. The judicious selection of the discount factor is integral to shaping the agent's

temporal perspective and influencing its adaptability to different scenarios, thus playing a crucial role in achieving effective and context-aware learning.

Let  $\pi$  represent the policy. Therefore, the formulation of the optimal policy can be represented by

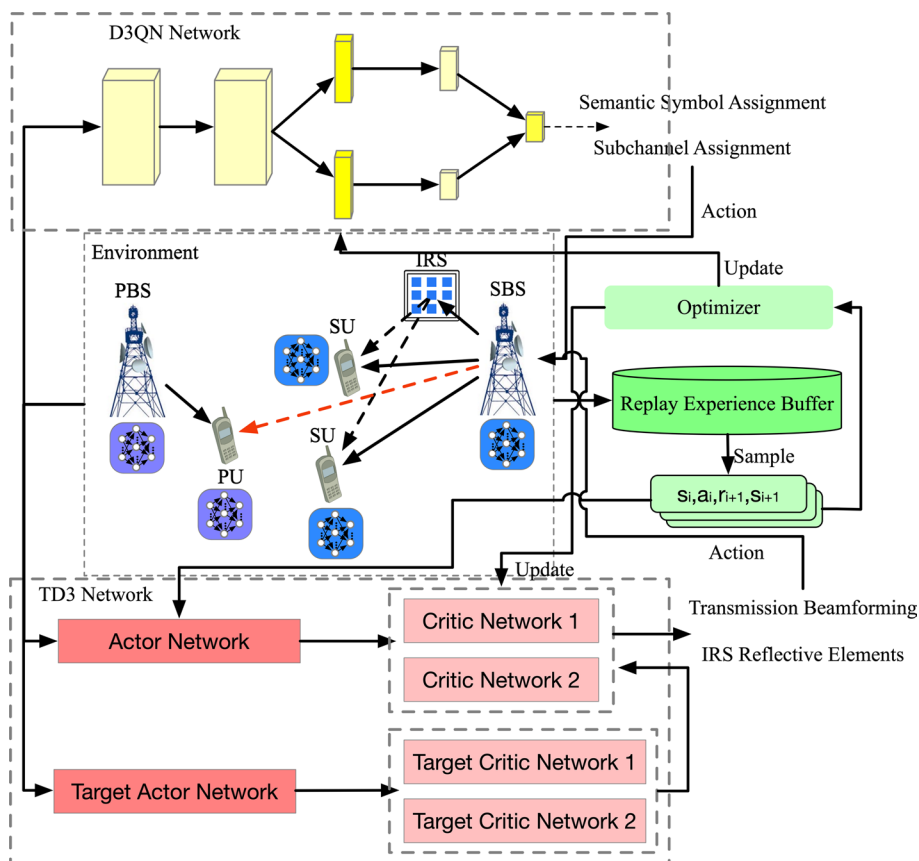
$$\pi^*(s, a) = \arg \max_{\pi} \mathbb{E}_{\pi} [r(s, a)]. \tag{13}$$

### 4.2 Intelligent resource allocation approach based on D3QN-TD3

We consider a DRL-based resource allocation approach using D3QN-TD3 for semantic spectrum sharing communication networks. Our proposed scheme can be seen in Fig. 2.

#### 4.2.1 D3QN algorithm for semantic symbol allocation and subchannel assignment

The incorporation of an additional advantage function into the DQN framework enhances the precision of Q-value estimations in Dueling DQN, requiring fewer discrete action data and thereby improving sample efficiency. Conversely, the concern of Q-value overestimation is tackled by double DQN through the prediction of Q-values using two sets of Q networks.



**Fig. 2** Proposed DRL-based resource allocation scheme for the IRS-enhanced semantic spectrum sharing communication

D3QN, a synergistic integration of dueling DQN and double DQN, leverages the strengths of both algorithms. In this context, D3QN is employed for the DeepSC allocation and channel allocation. The Q-value can be calculated by

$$Q(s, a; v, \eta, \tau) = \left( A(s, a; v, \tau) + V(s; v, \eta) - \frac{1}{|\mathcal{A}|} \left( \sum_{a'} A(s, a'; v, \tau) \right) \right), \quad (14)$$

where  $v$ ,  $\tau$  and  $\eta$  are the parameters of hidden layers, action network, and value network, respectively.

Sample tuples  $(s_t, a_t, r_{t+1}, s_{t+1})$  with the size of  $N$  are randomly selected from the replay buffer. The target Q value in D3QN is expressed by

$$Q = r + \gamma Q' \left( s_{t+1}, \arg \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}; v); v' \right), \quad (15)$$

where  $v'$  represents the parameters of hidden layers in target networks. The loss function is expressed by

$$\mathcal{L}(v) = \frac{1}{N} \sum_{t=1}^T (Q - Q(s_t, a_t; v))^2. \quad (16)$$

#### 4.2.2 TD3 algorithm for IRS reflective elements and transmit beamforming

The TD3 algorithm stands out in reinforcement learning by effectively addressing instability and sample inefficiency. Through the utilization of twin critics, TD3 enhances stability by minimizing overestimation bias in value function estimates, while the introduction of delayed policy updates prevents premature convergence to suboptimal policies. Additionally, the incorporation of target policy smoothing regularization promotes exploration and prevents the algorithm from becoming overly deterministic. These innovations collectively position TD3 as a robust solution, showcasing superior performance and potential applicability in diverse real-world scenarios, hence to address the continuous actions including IRS reflective elements and transmit beamforming of the SBS.

Let  $\vartheta$  and  $\vartheta^-$ , respectively, be parameters of the actor network and target actor network. Let  $\epsilon_1$  and  $\epsilon_1^-$ , respectively, be parameters of the critic networks, while  $\epsilon_2$  and  $\epsilon_2^-$  represent the parameters of the target critic networks. The target Q value in TD3 can be obtained by

$$Q = r_{t+1} + \gamma \min_{i=1,2} Q_{i, a \sim \pi}(s_{t+1}; \vartheta_i^-). \quad (17)$$

The weights  $\{\epsilon_i\}$  of the critic networks are updated by

$$\mathcal{L}(\epsilon_i) = \frac{1}{N} \sum_i [Q - Q_i(s_t, a_t; \epsilon_i)]^2. \quad (18)$$

#### 4.3 Training processing for semantic communication networks

To encapsulate our proposed algorithm succinctly, we utilize the current system state as a decisive input for the next action, encompassing semantic symbol allocation, channel

allocation, transmit beamforming, and IRS reflective coefficients. This decision-making process is facilitated through the integration of D3QN and TD3 networks. Primarily, we introduce a semantic similarity estimation algorithm based on DeepSC [22] to gauge semantic similarity and subsequently compute S-SE. This involves a domain transfer of the network, tailoring it to the specifics of our considered semantic task domain. The experiential knowledge accumulated through each training step finds residence in replay buffer  $\mathcal{E}$ . At predefined intervals  $I$ , the parameters of the evaluation network are synchronized with those of the target network. For clarity, the salient steps of our proposed approach, employing D3QN-TD3, are encapsulated in Algorithm 1.

**Algorithm 1** D3QN-TD3-Based Hybrid Intelligent Resource Allocation Scheme for Semantic Spectrum Sharing Networks

---

```

1: Initialize the semantic communication system with the parameters  $\vartheta, \epsilon_i, v, \eta, \tau$ ;
2: Initialize action  $a_0$ ;
3: Obtain the semantic similarity table based on [22];
4: for episode  $t = 1 \rightarrow T$  do
5:   Obtain the initial environment state with  $a_0$ ;
6:   for step  $i = 1 \rightarrow L$  do
7:     Obtain the discrete action decision based on D3QN  $a_t^1 = \{\rho_t, \zeta_t\}$ ;
8:     Obtain the continuous action decision based on TD3  $a_t^2 = \{\Psi_t, \mathbf{F}_t\}$ ;
9:     Execute  $(a_t^1, a_t^2)$  in the environment and obtain the next state  $s_{t+1}$ ;
10:    Compute  $r_t$  based on the semantic similarity table;
11:     $(s_t, a_t, r_t, s_{t+1}) \cap \mathcal{E} \rightarrow \mathcal{E}$ 
12:    Sample  $\{\{s_n, a_n, r_n, s_{n+1}\}\}_{n=1, \dots, B}$  from the memory buffer;
13:    Computer the loss functions;
14:    Generating the gradient  $\nabla_{\vartheta} Q, \nabla_{\epsilon_i} Q, \nabla_v Q, \nabla_{\eta} Q, \nabla_{\tau} Q$ ;
15:    Every  $I$  step, update  $\vartheta, \epsilon_i, v, \eta, \tau$  and copy the target parameters;
16:   end for
17: end for

```

---

## 5 Numerical analysis

In the specified scenario, there are three PUs, denoted by  $D=3$ , positioned at coordinates  $(200, 120, 0)$ ,  $(170, 160, 0)$ , and  $(165, 170, 0)$ , with a PBS located at the origin  $(150, 150, 30)$  equipped with six antennas. Simultaneously, there are three SUs, denoted by  $K=3$ , located at  $(40, 30, 0)$ ,  $(20, 50, 0)$ , and  $(0, 30, 20)$ , with a six-antenna SBS at  $(0, 0, 50)$ . The system configuration includes  $C=3$  subchannels, the minimum SBS transmit power  $TP_s^{\text{th}} = 20$  dBm, and the PBS transmit power of  $TP_p = 30$  dBm. The total bandwidth is set to  $B = 120$  kHz. A IRS is strategically located at  $(60, 60, 30)$  with  $E = 36$  IRS reflective elements. The variance of background noise is set to  $\sigma^2 = 0.01$ .

This section is dedicated to evaluating and contrasting the performance of proposed approaches with benchmark schemes. Channel characteristics are meticulously modeled, incorporating Rician fading for channels from BSs to IRS and from IRS to users. Concurrently, Rayleigh fading is assumed for channels from the BSs to the users and from the BSs to the IRS. The path fading is quantified as  $PL = (PL_0 - 10\tau \log_{10}(d/D_0))$  dB, with parameters set to  $PL_0 = 30$  dB and  $D_0 = 1$  m. The loss exponents  $\tau_{bu}$ ,  $\tau_{br}$ , and  $\tau_{ru}$  governing channels from BSs to Users, BSs to IRS, and IRS to users, respectively, are established as 3.6, 2.0, and 2.1. The convert method in [26] that equates the SE of

conventional communication to S-SE of semantic communications is introduced. This conversion is symbolized by  $\Phi'_{n,m} = R_{n,m} \frac{I}{\mu L}$ , where  $R_{n,m} = C_{n,m}/W$  represents the SE. The threshold  $\Psi_{th}$  and  $\xi_d^{th}$  are, respectively, set to 0.2 suts/s/Hz and 0.9. Within the D3QN configuration, two Q networks and two target Q networks are utilized, each incorporating three hidden layers with 256 neurons per layer. The learning rate of D3QN is explicitly set to 0.003. In the TD3 setup, there exist one actor network, two critic networks, two target networks, and one target actor network. Each of these networks is configured with three hidden layers, comprising 512 neurons. The learning rate of TD3 is set to 0.003. The buffer size is  $U = 20,000$ .

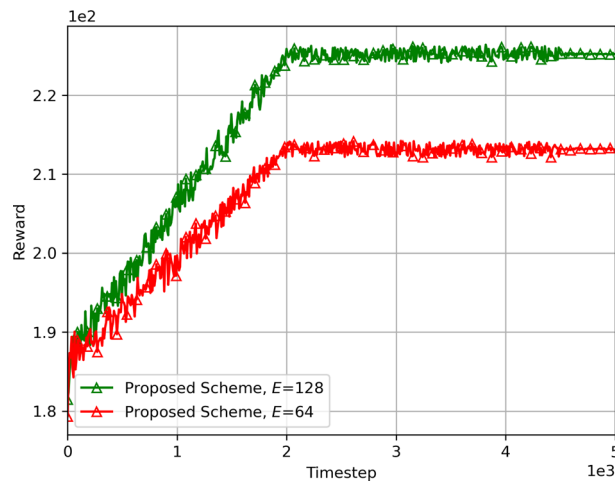
To comprehensively validate and compare the effectiveness of our proposed approach, we introduce several benchmark schemes as outlined below. The random IRS scheme can assess the impact of optimizing IRS reflective coefficients. In this setup, IRS reflection coefficients are assigned randomly, allowing us to evaluate the performance gain achieved through the optimization process. The random scheme is designed to highlight the significance of our proposed intelligent approach; this scheme incorporates random generation of IRS reflection coefficients using the random algorithm. This comparison aims to showcase the added value brought by the intelligent optimization process. The 5 G Communication standard, following the approach outlined in [26], is introduced as a benchmark. This comparison serves to assess the performance of our proposed approach in the context of evolving communication standards, emphasizing its adaptability and superiority in contemporary communication scenarios.

Figure 3 provides a comprehensive assessment of the convergence behavior of our intelligent resource allocation scheme for semantic spectrum sharing communication networks based on the integration of D3QN and TD3 across episodes. Two distinct scenarios are considered in this evaluation: one with  $E = 64$  IRS reflective elements and another with  $E = 128$  IRS reflective elements. The maximum transmit power of the SBS is set to 20 dBm. Notably, the proposed resource allocation approach consistently demonstrates a progressive improvement in rewards, reaching a fast convergence and showcasing substantial performance enhancements. When  $E = 128$  and  $E = 64$ , the algorithm can reach the convergence around  $2 \times 10^2$  due to the fact that our proposed framework is high-efficient in addressing the optimization problem. This observation serves as robust validation for the efficacy of our intelligent resource allocation strategy. It is noteworthy that achieving convergence in scenarios with  $E = 128$  presents a more challenging task compared to the  $E = 64$  scenario, given the heightened complexity associated with a larger number of IRS elements. It is evident that our proposed approach adeptly solves the challenges linked to a high-dimensional adjustment of IRS reflective elements, effectively navigating the exploration of optimal solutions. This effectiveness is attributed to the powerful exploration ability inherent in our proposed scheme, where the Q value of the actions can be evaluated accurately.

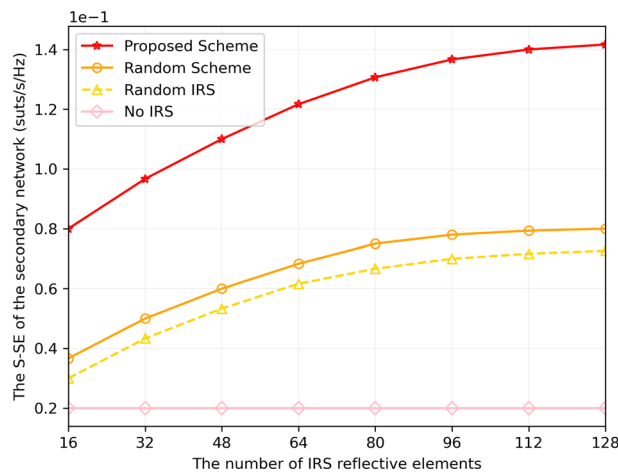
In Fig. 4, we present a comprehensive comparison of the achievable S-SE in the secondary network using proposed resource allocation approach against several benchmark schemes across varying numbers of IRS reflective elements. The results clearly demonstrate a significant improvement of S-SE as the number of IRS reflective elements increases. This improvement can be directly attributed to the growing number of IRS reflective elements, which significantly improves the beamforming accuracy and signal

gain. Consequently, it is easy to make a pronounced enhancement in S-SE, showcasing a significant advantage over schemes that lack IRS integration. Furthermore, our intricately crafted intelligent approach demonstrates the ability to effectively leverage the advantages provided by a substantial number of IRS elements, consistently yielding exceptional performance. In contrast, the fixed IRS scheme displays suboptimal performance, as their incapacity to dynamically adjust IRS array elements impedes their effectiveness in adapting to proposed semantic spectrum sharing network. Figure 4 vividly illustrates the advantageous impact of leveraging IRS to augment S-SE.

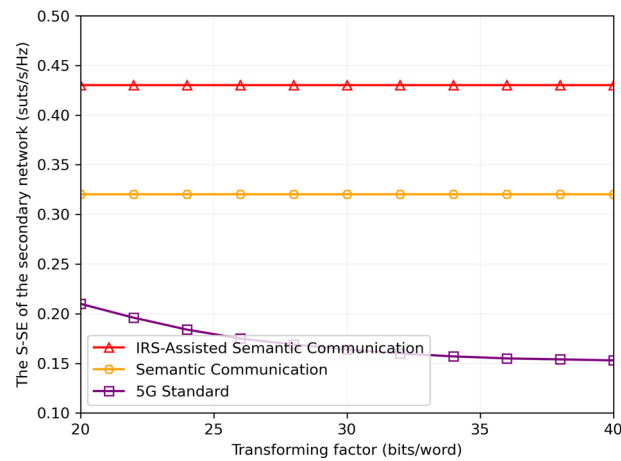
In Fig. 5, we depict the achievable S-SE of the secondary network under varying transforming factors, drawing a comparison with the conventional communication scheme. Notably, semantic communication demonstrates a consistent and stable S-SE performance versus transforming factors, in stark contrast to the declining trend observed in the S-SE of conventional communication. This is because of the fundamental nature of semantic



**Fig. 3** The convergence performance of our proposed approach for semantic spectrum sharing semantic communication networks



**Fig. 4** The S-SE of the secondary network versus different number of IRS reflective elements



**Fig. 5** The S-SE versus different transforming factors

communication, which strives to extract and transmit the important information. Importantly, the IRS-assisted semantic communication exhibits remarkable S-SE performance when compared to conventional communication standards. This is due to the inherent capability of IRS-enhanced communication to further enhance resource utilization efficiency, leading to significant gains in S-SE. The effectiveness of semantic communication schemes assisted by IRS is underscored, especially in scenarios with limited transforming factors.

## 6 Conclusion

We addressed the critical resource allocation challenges where IRS is employed for semantic spectrum sharing. The objective is to ensure QoS for primary network while simultaneously maximizing gains in the secondary network. To achieve this, we jointly optimize semantic symbol assignment, subchannel allocation, transmit beamforming of the SBS, and IRS reflective elements to maximize the S-SE of the secondary network. In order to enhance computational efficiency and intelligence in resource allocation, we introduce an intelligent hybrid method based on D3QN-TD3 to solve the non-convex optimization problem. Specifically, the D3QN component is responsible for determining semantic symbol and subchannel allocation, while the TD3 component focuses on optimizing the transmit beamforming of the SBS and IRS reflective elements. Simulation results validate the effectiveness of our DRL-based resource allocation approach, demonstrating better S-SE performance compared to benchmark schemes.

### Acknowledgements

We appreciate the editors and reviewers who processed and reviewed our manuscript to provide the detailed professional comments on the technical contributions, logical structure, and content presentation of this paper.

### Author contributions

Leibing Yan conducted conceptualization, writing, original draft preparation, and software; Yiqing Cai conducted software, methodology, validation, and figures; Hui Wei conducted formal analysis and editing; all authors reviewed the manuscript.

### Funding

This work was supported by the Key Scientific and Technological Project of Henan Province (Grant Number 222102210066), Key Scientific and Technological Project of Henan Province (Grant Number 222102210212), Henan Provincial Department of Science and Technology 2023 Key research and development and promotion special project (Grant Number 232102210037), Science and Technology Research Project of Henan Province (Grant Number 242102240040).



**Availability of data and materials**

Data sharing is not applicable to this study.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

All authors consent to the publication of this manuscript.

**Competing interests**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received: 22 January 2024 Accepted: 20 May 2024

Published online: 04 June 2024

**References**

1. W. Saad, M. Bennis, M. Chen, A vision of 6g wireless systems: applications, trends, technologies, and open research problems. *IEEE Netw.* **34**(3), 134–142 (2020)
2. C. Chaccour, W. Saad, M. Debbah, Z. Han, H.V. Poor, Less data, more knowledge: building next generation semantic communication networks. *arXiv preprint arXiv:2211.14343* (2022)
3. Y. Wu, S. Tang, L. Zhang, Resilient machine learning based semantic-aware MEC networks for sustainable next-G consumer electronics. *IEEE Transactions on Consumer Electronics* **PP**(99), 1–10 (2023)
4. Z. Yang, M. Chen, Z. Zhang, C. Huang, Energy efficient semantic communication over wireless networks with rate splitting. *IEEE J. Sel. Areas Commun.* **41**(5), 1484–1495 (2023)
5. J. Xia, L. Fan, W. Xu, X. Lei, X. Chen, G.K. Karagiannidis, A. Nallanathan, Secure cache-aided multi-relay networks in the presence of multiple eavesdroppers. *IEEE Trans. Commun.* **67**(11), 7672–7685 (2019)
6. E. Erdemir, T.-Y. Tung, P.L. Dragotti, D. Gündüz, Generative joint source-channel coding for semantic image transmission. *IEEE J. Sel. Areas Commun.* **41**(8), 2645–2657 (2023)
7. W. Yang, H. Du, Z.Q. Liew, W.Y.B. Lim, Z. Xiong, D. Niyato, X. Chi, X. Shen, C. Miao, Semantic communications for future internet: fundamentals, applications, and challenges. *IEEE Commun. Surv. Tutor.* **25**(1), 213–250 (2023)
8. M.K. Farshbafan, W. Saad, M. Debbah, Curriculum learning for goal-oriented semantic communications with a common language. *IEEE Trans. Commun.* **71**(3), 1430–1446 (2023)
9. K. Lu, Q. Zhou, R. Li, Z. Zhao, X. Chen, J. Wu, H. Zhang, Rethinking modern communication from semantic coding to semantic communication. *IEEE Wirel. Commun.* **30**(1), 158–164 (2023)
10. S. Tang, Q. Yang, L. Fan, Contrastive learning based semantic communications. *IEEE Transact. Commun.* **99**, 1–12 (2024)
11. W. Yang, Z.Q. Liew, W.Y.B. Lim, Z. Xiong, D. Niyato, X. Chi, X. Cao, K.B. Letaief, Semantic communication meets edge intelligence. *IEEE Wirel. Commun.* **29**(5), 28–35 (2022)
12. V. Kumar, M.F. Flanagan, R. Zhang, L.-N. Tran, Achievable rate maximization for underlay spectrum sharing mimo system with intelligent reflecting surface. *IEEE Wirel. Commun. Lett.* **11**(8), 1758–1762 (2022)
13. D. Peng, H. Ye, Y. Li, Z. Wang, Integrating terrestrial and satellite multibeam systems toward 6g: techniques and challenges for interference mitigation. *IEEE Wirel. Commun.* **29**(1), 24–31 (2022)
14. L. Liang, H. Ye, G.Y. Li, Spectrum sharing in vehicular networks based on multi-agent reinforcement learning. *IEEE J. Sel. Areas Commun.* **37**(10), 2282–2292 (2019)
15. M. Asif Haider, Y.D. Zhang, E. Aboutanios, Isac system assisted by ris with sparse active elements. *EURASIP J. Adv. Signal Process.* **2023**(1), 1–22 (2023)
16. G. Tian, R. Song, Cooperative beamforming for a double-irs-assisted wireless communication system. *EURASIP J. Adv. Signal Process.* **2021**, 1–10 (2021)
17. X. Guan, Q. Wu, R. Zhang, Joint power control and passive beamforming in IRS-assisted spectrum sharing. *IEEE Commun. Lett.* **24**(7), 1553–1557 (2020)
18. J. Chen, L. Guo, J. Jia, J. Shang, X. Wang, Resource allocation for irls assisted sgf noma transmission: a madrl approach. *IEEE J. Sel. Areas Commun.* **40**(4), 1302–1316 (2022)
19. S. Gong, L. Cui, B. Gu, B. Lyu, D.T. Hoang, D. Niyato, Hierarchical deep reinforcement learning for age-of-information minimization in irls-aided and wireless-powered wireless networks. *IEEE Trans. Wirel. Commun.* **22**(11), 8114–8127 (2023)
20. T.V. Nguyen, T.P. Truong, T.M.T. Nguyen, W. Noh, S. Cho, Achievable rate analysis of two-hop interference channel with coordinated irls relay. *IEEE Trans. Wirel. Commun.* **21**(9), 7055–7071 (2022)
21. F. Zhou, Y. Li, M. Xu, L. Yuan, Q. Wu, R.Q. Hu, N. Al-Dhahir, Cognitive semantic communication systems driven by knowledge graph: principle, implementation, and performance evaluation. *arXiv preprint arXiv:2303.08546* (2023)
22. H. Xie, Z. Qin, G.Y. Li, B.-H. Juang, Deep learning enabled semantic communication systems. *IEEE Trans. Signal Process.* **69**, 2663–2675 (2021)
23. H. Yoo, T. Jung, L. Dai, S. Kim, C.-B. Chae, *Demo: real-time semantic communications with a vision transformer*. In: 2022 IEEE International Conference on Communications Workshops (ICC Workshops), pp. 1–2 (2022)
24. J. Kang, H. Du, Z. Li, Z. Xiong, S. Ma, D. Niyato, Y. Li, Personalized saliency in task-oriented semantic communications: image transmission and performance analysis. *IEEE J. Sel. Areas Commun.* **41**(1), 186–201 (2023)

25. Z. Weng, Z. Qin, X. Tao, C. Pan, G. Liu, G.Y. Li, Deep learning enabled semantic communications with speech recognition and synthesis. *IEEE Trans. Wirel. Commun.*, 1–1 (2023)
26. L. Yan, Z. Qin, R. Zhang, Y. Li, G.Y. Li, Resource allocation for text semantic communications. *IEEE Wirel. Commun. Lett.* **11**(7), 1394–1398 (2022)
27. L. Yan, Z. Qin, R. Zhang, Y. Li, G. Ye Li, *QoE-aware resource allocation for semantic communication networks*. In: IEEE Glob. Commun. Conf. (GLOBECOM), pp. 3272–3277 (2022)
28. L. Wang, W. Wu, F. Zhou, Z. Yang, Z. Qin, Adaptive resource allocation for semantic communication networks. arXiv preprint [arXiv:2312.01081](https://arxiv.org/abs/2312.01081) (2023)
29. B. Hu, J. Ma, Z. Sun, J. Liu, R. Li, L. Wang, Drl-based intelligent resource allocation for physical layer semantic communication with irs. *Phys. Commun.*, 102270 (2023)
30. P. Jiang, C.-K. Wen, S. Jin, G.Y. Li, RIS-enhanced semantic communications adaptive to user requirements. *IEEE Trans. Commun.* (2024). to be published
31. Y. Wu, F. Zhou, W. Wu, Q. Wu, R.Q. Hu, K.-K. Wong, Multi-objective optimization for spectrum and energy efficiency tradeoff in irs-assisted crns with noma. *IEEE Trans. Wireless Commun.* **21**(8), 6627–6642 (2022)
32. L. Wang, F. Yang, Y. Chen, S. Lai, W. Wu, Intelligent resource allocation for transmission security on IRS-assisted spectrum sharing systems with OFDM. *Phys. Commun.* **58**, 102013 (2023)
33. Y. Wu, F. Zhou, Q. Wu, Y. Huang, R.Q. Hu, Resource allocation for IRS-assisted sensing-enhanced wideband CR networks. In: Proc. IEEE Int. Conf. Commun. Work. (ICC Workshops), pp. 1–6 (2021)
34. L. Wang, W. Wu, F. Zhou, Q. Wu, O.A. Dobre, T.Q.S. Quek, Hybrid hierarchical DRL enabled resource allocation for secure transmission in multi-IRS-assisted sensing-enhanced spectrum sharing networks. *IEEE Transact. Wirel. Commun.* (2023). to be published
35. V. Kumar, M.F. Flanagan, R. Zhang, L.-N. Tran, Achievable rate maximization for underlay spectrum sharing mimo system with intelligent reflecting surface. *IEEE Wirel. Commun. Lett.* **11**(8), 1758–1762 (2022)
36. H. Sadiq, A.K. Hassan, Z.H. Abbas, G. Abbas, System throughput maximization in irs-assisted phase cooperative noma networks. *Phys. Commun.* **58**, 102007 (2023)
37. Y. Guo, R. Zhao, S. Lai, L. Fan, X. Lei, G.K. Karagiannidis, Distributed machine learning for multiuser mobile edge computing systems. *IEEE J. Sel. Top. Signal Process.* **16**(3), 460–473 (2022)
38. Y. Wu, C. Cai, X. Bi, J. Xia, C. Gao, Y. Tang, S. Lai, Intelligent resource allocation scheme for cloud-edge-end framework aided multi-source data stream. *EURASIP J. Adv. Signal Process.* **2023**(1), 56 (2023)
39. W. Wu, F. Yang, F. Zhou, Q. Wu, R.Q. Hu, Intelligent resource allocation for IRS-enhanced OFDM communication systems: a hybrid deep reinforcement learning approach. *IEEE Trans. Wireless Commun.* **22**(6), 4028–4042 (2023)
40. L. Wang, W. Wu, F. Zhou, *Intelligent resource allocation for irs-assisted sensing-enhanced secure communication crns*. In: 2023 International Conference on Ubiquitous Communication (Ucom), pp. 344–349 (2023)
41. X. Wang, Y. Zhang, R. Shen, Y. Xu, F.-C. Zheng, DRL-based energy-efficient resource allocation frameworks for uplink NOMA systems. *IEEE Internet Things J.* **7**(8), 7279–7294 (2020)
42. B. Hazarika, K. Singh, S. Biswas, C.-P. Li, DRL-based resource allocation for computation offloading in IoV networks. *IEEE Trans. Industr. Inf.* **18**(11), 8027–8038 (2022)
43. L. Wang, W. Wu, F. Tian, H. Hu, *Intelligent resource allocation for uav-enabled spectrum sharing semantic communication networks*. In: 2023 IEEE 23rd International Conference on Communication Technology (ICCT), pp. 1359–1363 (2023). IEEE
44. M.E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, L. Zettlemoyer, *Deep contextualized word representations*. In: Proc. North Amer. Chapter Assoc. Comput. Linguistics: Hum. Lang. Tech., pp. 2227–2237 (2018)

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.