Performance Optimization for Semantic Communications: An Attention-Based Reinforcement Learning Approach

Yining Wang, *Student Member, IEEE*, Mingzhe Chen, *Member, IEEE*, Tao Luo, *Senior Member, IEEE*,
Walid Saad, *Fellow, IEEE*, Dusit Niyato, *Fellow, IEEE*, H. Vincent Poor, *Life Fellow, IEEE*,
and Shuguang Cui, *Fellow, IEEE*

Abstract—In this paper, a semantic communication framework is proposed for textual data transmission. In the studied model, a base station (BS) extracts the semantic information from textual data, and transmits it to each user. The semantic information is modeled by a knowledge graph (KG) that consists of a set of semantic triples. After receiving the semantic information, each user recovers the original text using a graph-to-text generation model. To measure the performance of the considered semantic communication framework, a metric of semantic similarity (MSS) that jointly captures the semantic accuracy and completeness of the recovered text is proposed. Due to wireless resource limitations, the BS may not be able to transmit the entire semantic information to each user and satisfy the transmission delay constraint. Hence, the BS must select an appropriate resource block for each user as well as determine and transmit part of the semantic information to the users. As such, we formulate an optimization problem whose goal is to maximize the total MSS by jointly optimizing the resource allocation policy and determining the partial semantic information to be transmitted. To solve this problem, a proximal-policy-optimization-based reinforcement learning (RL) algorithm integrated with an attention network is proposed. The proposed algorithm can evaluate the importance of each triple in the semantic information using an attention network and then, build a relationship between the importance distribution of the triples in the semantic information and the total MSS. Compared to traditional RL algorithms, the proposed algorithm can dynamically adjust its learning rate thus ensuring convergence to a locally optimal solution. Simulation results show that the proposed framework can reduce by 41.3% data that the BS needs to transmit and improve by two-fold the total MSS compared to a standard communication network without using semantic communication techniques.

Index Terms—Semantic communications, resource allocation, attention networks, policy gradient, reinforcement learning (RL), semantic similarity.

Yining Wang and Tao Luo are with the Beijing Laboratory of Advanced Information Network, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: wyy0206@bupt.edu.cn; tluo@bupt.edu.cn).

Mingzhe Chen is with the Department of Electrical and Computer Engineering and Institute for Data Science and Computing, University of Miami, Coral Gables, FL 33146 USA (e-mail: mingzhe.chen@miami.edu).

Walid Saad is with the Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, Arlington, VA 22203 USA (e-mail: walids@vt.edu).

Dusit Niyato is with the School of Computer Science and Engineering (SCSE), NTU, Singapore (e-mail: dniyato@ntu.edu.sg).

H. Vincent Poor is with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

Shuguang Cui is with the School of Science and Engineering (SSE), Guangdong Provincial Key Laboratory of Future Networks of Intelligence, Future Network of Intelligence Institute (FNii), The Chinese University of Hong Kong, Shenzhen 518172, China, also with the Shenzhen Research Institute of Big Data, Shenzhen 518172, China, and also with the Peng Cheng Laboratory, Shenzhen 518066, China (e-mail: shuguangcui@cuhk.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JSAC.2022.3191112.

Digital Object Identifier 10.1109/JSAC.2022.3191112

I. INTRODUCTION

NEW wireless applications such as extended reality and digital twinning are generating unprecedented amounts of data (at zetta-byte scale) which will strain the capacity of wireless networks [2]. To support these human-centered services and applications, wireless networks must be carefully designed based on content, human-related requirements, human-related knowledge, and experience-based metrics [3]. One of the solutions for these challenges is *semantic communication*, which allows the meaning of the data (behind digital bits) to be extracted and exploited during wireless transmission [3]–[5]. Semantic communication has recently attracted significant interest due to its advantages in terms of providing human-oriented services and improving communication efficiency. However, deploying semantic communications in wireless networks faces several challenges including extraction of data meaning, semantic-oriented resource allocation, semantic encoding and decoding, measurement of semantic communication performance, and security of semantic communications.

A. Related Works

Recently, a number of works such as [4]–[9] have investigated semantic communications over wireless networks.

In particular, in [4], the authors proposed a deep learning approach to optimize the mutual information between the original and decoded signals. The authors in [6] designed a distributed semantic communication system for capacity-limited networks. In [4] and [6], the authors used a deep neural network based autoencoder to encode data and considered the output of the autoencoder as the semantic information of the data. However, the output of the autoencoder being a vector of decimals is incomprehensible for humans and does not have any physical meaning. In [7] and [8], the authors developed a semantic communication framework for real-time control systems and considered the control signals as the semantic information of the data. However, using control signals as semantic information is not applicable for text or image data transmission since control signals cannot represent the content related to a text or image. The authors in [9] used a hierarchical structure to represent the relationships among the objects in the real world so as to understand the meaning underlying the data to be transmitted. In [5], the authors provided an overview on the use of semantic detection and knowledge modeling techniques for semantic information extraction. However, neither [5] nor [9] considered the design of semantic metrics that can measure the performance of semantic communications.

In [10]–[13], the authors proposed a number of metrics for evaluating the performance of semantic communication systems. The work in [10] derived the semantic channel capacity based on the mutual information between the transmitted symbols and the desired meaning. In [11], the authors proposed a semantic communication system whose goal is to minimize the average age of incorrect information (AoII) that captures the accuracy and freshness of the real-time data. The authors in [12] used the perceptual evaluation of speech distortion as well as speech signal metrics to evaluate the performance of a semantic-based speech transmission framework. The work in [13] defined the semantic distance between two words based on the distribution of words in a given corpus so as to measure the semantic similarity of two sentences. While interesting, the metrics developed in [10]–[13] only capture whether the information in the received data is correct without guaranteeing that the received data contains all the information in the original data to be transmitted. Meanwhile, these prior works [10]–[13] assumed that all the semantic information extracted from the original data can be transmitted over the network and did not consider the semantic communication over wireless resource-constrained networks. In practice, due to wireless resource limitations (e.g., bandwidth), the BS may not be able to transmit the entire semantic information to all users [14]. A number of existing works such as in [15]–[19] studied the problem of resource allocation in wireless networks. However, the solutions in these existing works [15]–[19] cannot be applied in semantic communications since they do not consider the effects of semantic information extraction and transmission on resource allocation.

B. Contributions

The main contribution of this work is a novel semantic communication framework that jointly considers semantic information extraction and transmission, data recovery, as well as performance evaluation. The key contributions are summarized as follows:

- We consider a semantic communication network in which a base station (BS) uses semantic communication techniques to extract the meaning of the text data and transmits it to its associated users. The meaning of the text data is defined as the semantic information and modeled by a knowledge graph (KG) that consists of a set of semantic triples in the form of "*entity-relation-entity*". Based on the received semantic information, each user uses a graph-to-text generation model to recover the original text.
- We propose a mathematical metric, called the metric of semantic similarity (MSS), to capture the semantic communication performance. The proposed MSS can measure whether the information in the recovered text is correct and whether the recovered text contains all the information in the original text.
- To satisfy the delay constraint, the BS must optimize the resource allocation for each user and transmit a part of the semantic information to each user. This problem is formulated as an optimization problem whose goal is to maximize the total MSS by optimizing the resource allocation policy and determining the semantic information that needs to be transmitted.
- To solve this problem, we propose an attention proximal policy optimization (APPO) algorithm that can evaluate the importance of each triple in the semantic information. Then, the proposed algorithm can analyze the relationship between the importance distribution of the triples in the semantic information and the total MSS, thus finding the effective policies for resource allocation and semantic information transmission.
- We analyze the convergence of the proposed APPO algorithm. The analytical result shows that the APPO algorithm is guaranteed to converge to a locally optimal solution of the studied total MSS maximization problem.

Simulation results show that, compared to a standard communication network that does not consider semantic communications, the proposed APPO algorithm can reduce the number of words that the BS needs to transmit by up to 41.3% while achieving a 2-fold improvement in the total MSS. To our knowledge, *this is the first work that introduces a mathematical model for semantic communication enabled wireless networks and optimizes resource allocation and semantic information transmission to improve the performance of semantic-driven wireless networks.*

The rest of this paper is organized as follows. The proposed semantic communication framework and the problem formulation are described in Section II. The use of APPO algorithm for resource allocation and partial semantic information determination is introduced in Section III. In Section IV, the numerical results are presented and discussed. Finally, conclusions are drawn in Section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a wireless network that consists of a BS using semantic communication techniques to transmit the meaning

TABLE I

LIST OF NOTATION

Notation	Description
U	Number of users
Q	Number of downlink orthogonal RBs
L_i	Original text needed to transmit to user i
$\boldsymbol{\alpha}_i$	RB allocation vector for user i
$w_{i,n}$	Token n in text L_i
$c_i(\boldsymbol{\alpha}_i)$	Downlink channel capacity
N_i	Number of tokens in text L_i
W	Bandwidth of each RB
$e_{i,j}$	Entity j in text L_i
P	Transmit power of the BS
E_i	Number of entities in text L_i
I_q	Interference of RB q
$r_{i,jk}$	Relation between entity $e_{i,j}$ and entity $e_{i,k}$
N_0	Noise power spectral density
R_i	Number of relations in text L_i
T	Transmission delay threshold
$\mathcal{G}_i$	Semantic information of text L_i
ϕ_i	Channel gain between the BS and user i
ε_i^g	Semantic triple g in $\mathcal{G}_i$
$L_i'(\boldsymbol{\alpha}_i, \mathcal{G}_i')$	Text recovered by user i
$Z(\mathcal{G}_i)$	Number of tokens in $\mathcal{G}_i$
M_i	Number of tokens in text $L_i'(\boldsymbol{\alpha}_i, \mathcal{G}_i')$
$\mathcal{G}_i'$	Partial semantic information transmitted to user i
O	Number of bits used to represent a token
$A_i(\boldsymbol{\alpha}_i, \mathcal{G}_i')$	Semantic accuracy
$E_i(\boldsymbol{\alpha}_i, \mathcal{G}_i')$	Metric of semantic similarity
$R_i(\boldsymbol{\alpha}_i, \mathcal{G}_i')$	Semantic completeness
ξ_i	Penalty for short text
φ	Parameter adjusting $A_i(\boldsymbol{\alpha}_i, \mathcal{G}_i')$ and $R_i(\boldsymbol{\alpha}_i, \mathcal{G}_i')$

Fig. 1.   Illustration of the proposed semantic communication framework.

$L_i = \{[\text{little}], [\text{girls}], [\text{are}], [\text{playing}], [.]\}$, where $w_{i,1} = [\text{little}]$, $w_{i,2} = [\text{girls}]$, $w_{i,3} = [\text{are}]$, $w_{i,4} = [\text{playing}]$, and $w_{i,5} = [.]$.

In our model, the semantic information extracted from a text data is modeled by a KG [23]. Hence, the semantic information consists of a set of nodes and a set of edges, as shown in Fig. 2. In particular, each node in the semantic information is an *entity* that refers to an object or a concept in the real world. Hereinafter, we define entity j that consists of a subsequence of text L_i as $e_{i,j}$. For example, in Fig. 2, "*baseform of word*" is an entity that consists of four tokens in the example text. An information extraction system such as the scientific information extractor in [24] can be used to recognize the set $\mathcal{E}_i$ of E_i entities in text L_i.

Edges are the *relations* between each pair of entities. Given a pair of recognized entities $(e_{i,j}, e_{i,k}), j \neq k$, the BS must find the relation $r_{i,jk} \in \mathcal{R}_i$ between them, where $\mathcal{R}_i$ is the set of R_i relations involved in text L_i. For example, in Fig. 2, the relation between entity "*baseform of word*" and "*stochastic lexicon model*" can be formulated as "*part of*". Note that the relations (i.e., the edges of the semantic information) are directional and, hence, we have $r_{i,jk} \neq r_{i,kj}$. We assume that there is a predefined set $\mathcal{R}$ containing all relations in the texts and each relation is a two-token sequence such as "*part of*" and "*evaluate for*", as done in [24]. Hence, given a pair of entities $e_{i,j}$ and $e_{i,k}$ in the original text L_i, the relation $r_{i,jk}$ between $e_{i,j}$ and $e_{i,k}$ can be obtained by classification algorithms such as convolutional neural networks [25]. Here, the input of the relation classification algorithm is the sentence containing the pair of entities in the original text and the output is the predefined two-token relation that can summarize the description of the entities in the input sentence. We assume that the information extraction system can recognize all the entities in the original text. Hence, the extracted triples in the form of "*entity-relation-entity*" can represent *the meaning* of the original text.

Based on the recognized entities and the extracted relations, the semantic information of text L_i can be modeled as

$$\mathcal{G}_i = \{\varepsilon_i^1, \ldots, \varepsilon_i^g, \ldots, \varepsilon_i^{G_i}\}, \qquad (2)$$

where $\varepsilon_i^g = (e_{i,j}^g, r_{i,jk}^g, e_{i,k}^g), \forall e_{i,j}^g, e_{i,k}^g \in \mathcal{E}_i, j \neq k, \forall r_{i,jk}^g \in \mathcal{R}_i$ is a semantic triple and G_i is the number of semantic triples in $\mathcal{G}_i$. Since each entity (e.g., $e_{i,j}^g$ or $e_{i,k}^g$)

of text data to a set $\mathcal{U}$ of U users. Hereinafter, the meaning of the text data transmitted over wireless links is referred to *semantic information*. To achieve semantic communications, the BS extracts the semantic information from the original text and sends it to the corresponding user.[1] Then, each user recovers the text based on the received semantic information. In particular, the considered semantic communication process consists of four phases (shown in Fig. 1): a) semantic information extraction, b) semantic-oriented resource allocation and semantic information selection, c) original text data recovery, and d) semantic similarity evaluation. Next, the process of semantic communications is first introduced. Then, a semantic similarity model is proposed to measure the performance of semantic communications.

A. Semantic Information Extraction

A *token* $w_{i,n}$ is used to represent a word, a symbol, or a punctuation in the text data. As a result, the text data that the BS needs to transmit to user i consists of a sequence of tokens, as follows:

$$L_i = \{w_{i,1}, w_{i,2}, \ldots, w_{i,n}, \ldots, w_{i,N_i}\}, \quad \forall w_{i,n} \in \mathcal{V}, \quad (1)$$

where $\mathcal{V}$ is the set of tokens in a corpus and N_i is the number of tokens in L_i. For example, if the BS requires to transmit "*Little girls are playing.*" to user i, then, we have

---

[1]In this work, we only consider the textual data transmission. One can easily extend the proposed model to other types of data such as audio data and image data [20]–[22].
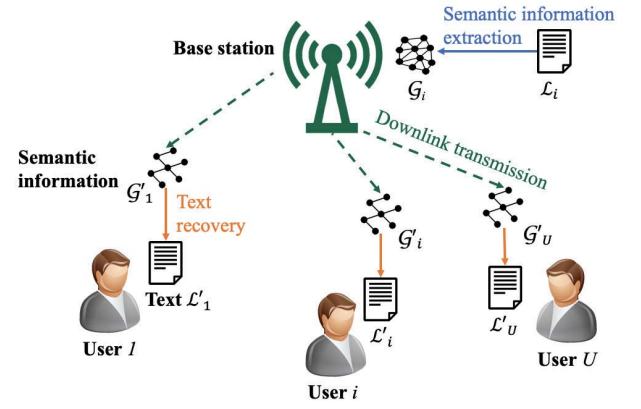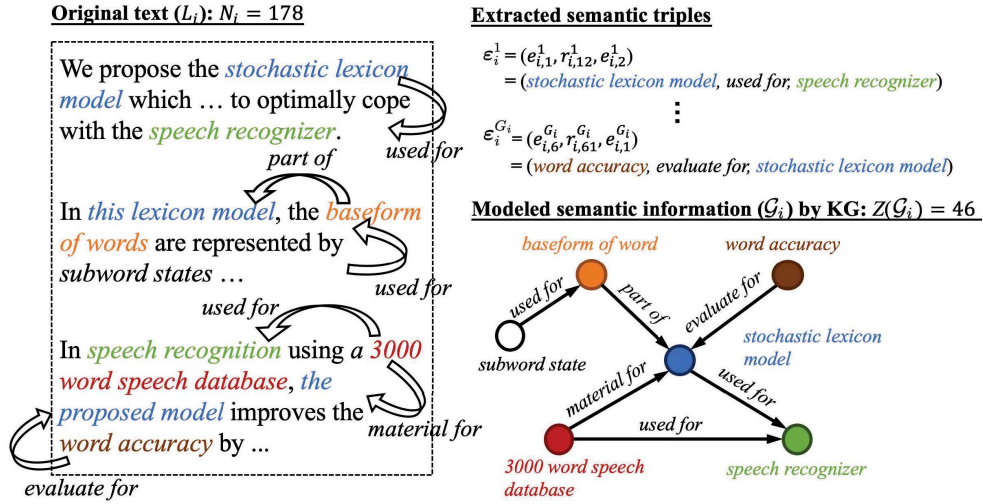
Fig. 2. An example of an original text and the extracted semantic information.

and each relation $r_{i,jk}^g$ consist of a sequence of tokens (e.g., $e_{i,j} = \{[\text{baseform}], [\text{of}], [\text{word}]\}$, $e_{i,k} = \{[\text{stochastic}], [\text{lexicon}], [\text{model}]\}$, and $r_{i,jk} = \{[\text{part}], [\text{of}]\}$), triple ε_i^g can be expressed as $\varepsilon_i^g = \{v_{i,1}^g, \ldots, v_{i,b}^g, \ldots, v_{i,B_i^g}^g\}$, where $v_{i,b}^g \in \mathcal{V}$ is token b in semantic triple ε_i^g and $B_i^g = S_{i,j}^g + S_{i,k}^g + S_{i,jk}^g$ with $S_{i,j}^g$ and $S_{i,jk}^g$ being the number of tokens in entity $e_{i,j}^g$ and relation $r_{i,jk}^g$, respectively. Then, the number of tokens in semantic information $\mathcal{G}_i$ is

$$Z(\mathcal{G}_i) = \sum_{g=1}^{G_i} \left(S_{i,j}^g + S_{i,k}^g + S_{i,jk}^g \right). \tag{3}$$

From Fig. 2, we see that the data size of the extracted semantic information is much smaller than the data size of the original text (i.e., $Z(\mathcal{G}_i) \ll N_i$). The reason is that the two-token relations between the entity pairs in $\mathcal{G}_i$ can reduce the redundant context in original text L_i.

B. Transmission Model

An orthogonal frequency division multiple access (OFDMA) technique is used for semantic information transmission. In our model, a set $\mathcal{Q}$ of Q downlink orthogonal resource blocks (RBs) can be allocated to serve the users. The RB allocation vector of user i is $\boldsymbol{\alpha}_i = [\alpha_{i,1}, \ldots, \alpha_{i,q}, \ldots, \alpha_{i,Q}]$, where $\alpha_{i,q} \in \{0, 1\}$. Here, $\alpha_{i,q} = 1$ implies that RB q is allocated to user i; otherwise, we have $\alpha_{i,q} = 0$. In our model, we assume that each RB can only be allocated to one user and each user can only occupy one RB. Then, we have

$$\sum_{q=1}^{Q} \alpha_{i,q} \leqslant 1, \quad \forall i \in \mathcal{U}; \quad \sum_{i=1}^{U} \alpha_{i,q} \leqslant 1, \ \forall q \in \mathcal{Q}. \tag{4}$$

The downlink channel capacity of the BS transmitting semantic information $\mathcal{G}_i$ to user i is given as [26]

$$c_i(\boldsymbol{\alpha}_i) = \sum_{q=1}^{Q} \alpha_{i,q} W \log_2 \left(1 + \frac{P\phi_i}{I_q + W N_0} \right) \tag{5}$$

where W is the bandwidth of each RB, P is the transmit power of the BS, I_q represents the interference caused by BSs that are located in other service areas and use RB q, N_0 is the noise power spectral density, and $\phi_i = \gamma_i d_i^{-2}$ is the channel gain between the BS and user i with γ_i being the Rayleigh fading parameters and d_i being the distance between the BS and user i. Here, we assume that the transmit power P of each user i is a constant and one can easily consider the optimization of transmit power by extending the studied model. We also assume that the transmission delay between the BS and each user i is limited to D. Given the data rate $c_i(\boldsymbol{\alpha}_i)$, the maximum number of tokens that can be transmitted within the transmission delay D is determined. Hence, to satisfy the transmission delay constraint, the BS must determine partial semantic information (i.e., a subset of triples) to be transmitted.

C. Text Recovery

The partial semantic information that the BS transmits to user i is given as

$$\mathcal{G}_i' = \{\varepsilon_i'^1, \ldots, \varepsilon_i'^h, \ldots, \varepsilon_i'^{H_i}\} \subset \mathcal{G}_i, \tag{6}$$

where $\varepsilon_i'^h = (e_{i,j}'^h, r_{i,jk}'^h, e_{i,k}'^h)$ and H_i is the number of selected semantic triples in $\mathcal{G}_i'$. Given the transmission delay threshold D, the selected semantic information $\mathcal{G}_i'$ should satisfy the delay constraint as follows:

$$\frac{Z(\mathcal{G}_i')O}{c_i(\boldsymbol{\alpha}_i)} \leqslant D, \tag{7}$$

where $Z(\mathcal{G}_i') = \sum_{h=1}^{H_i} \left(S_{i,j}^h + S_{i,k}^h + 2 \right)$ and O is the number of bits used to represent each token. After each user i receives the partial semantic information $\mathcal{G}_i'$, a graph-to-text generation model, such as the graph transformer in [27], can be used to recover the coherent multi-sentence text from $\mathcal{G}_i'$. We assume that the graph-to-text generation model [27] is well-trained and

IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. 40, NO. 9, SEPTEMBER 2022

shared among all users. The text recovered by user i based on $\mathcal{G}'_i$ is

$$L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) = \{w'_{i,0}, w'_{i,1}, \ldots, w'_{i,m}, \ldots, w'_{i,M_i}\}, \quad (8)$$

where M_i is the number of tokens in the recovered text L'_i.

D. Semantic Similarity Model

To measure the quality of semantic communications, we propose a metric of semantic similarity (MSS). Compared to the existing cosine similarity metric [4], the proposed MSS can capture semantic similarity while avoiding semantic errors caused by word vectorization. Hereinafter, we define the degree of the information in the recovered text being correct as *semantic accuracy* and the degree of the recovered text containing the information of the original text as *semantic completeness*. Compared to the existing bilingual evaluation understudy (BLEU) metric [28] that only captures the semantic accuracy of the recovered text, the proposed MSS jointly captures the semantic accuracy and completeness of the recovered text. A method based on token matching is introduced to calculate the semantic accuracy and completeness [29]. The semantic accuracy of recovered text $L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)$ is defined as

$$A_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) = \frac{\sum\limits_{m=1}^{M_i} \min\left(\sigma\left(L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i), w'_{i,m}\right), \sigma(L_i, w'_{i,m})\right)}{\sum\limits_{m=1}^{M_i} \sigma\left(L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i), w'_{i,m}\right)}, \quad (9)$$

where $\sigma(L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)_i, w'_{i,m})$ and $\sigma(L_i, w'_{i,m})$ is the number of occurrences of token $w'_{i,m}$ in recovered text $L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)$ and in original text L_i, respectively, and $\min(\sigma\left(L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i), w'_{i,m}\right), \sigma(L_i, w'_{i,m}))$ indicates the number of correct occurrences of token $w'_{i,m}$ in recovered text $L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)$. For example, if token $w'_{i,m} = $ [you] occurs twice in the recovered text and once in the original text, only one of the two occurrences of $w'_{i,m} = $ [you] is correct in the recovered text. In (9), $\sum\limits_{m=1}^{M_i} \min\left(\sigma\left(L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i), w'_{i,m}\right), \sigma(L_i, w'_{i,m})\right)$ represents the sum of the number of correct occurrences of each token in the recovered text and $\sum\limits_{m=1}^{M_i} \sigma\left(L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i), w'_{i,m}\right)$ represents the sum of the number of occurrences of each token in the recovered text. Hence, $A_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)$ is defined as the ratio of the sum of the number of correct occurrences of each token to the sum of the number of occurrences of each token.

The semantic completeness of the recovered text is defined as

$$R_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) = \frac{\sum\limits_{m=1}^{M_i} \min\left(\sigma\left(L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i), w'_{i,m}\right), \sigma(L_i, w'_{i,m})\right)}{\sum\limits_{n=1}^{N_i} \sigma(L_i, w_{i,n})}. \quad (10)$$

In (10), $\sum\limits_{n=1}^{N_i} \sigma(L_i, w_{i,n})$ represents the the sum of the number of occurrences of each token in the original text. Hence,

$R_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)$ is defined as the ratio of the sum of the number of correct occurrences of each token in the recovered text to the sum of the number of occurrences of each token in the original text. Next, we use an example to explain the differences between the semantic accuracy and the semantic completeness more clearly. For example, $L_i = \{[\text{little}], [\text{girls}], [\text{are}], [\text{playing}]\}$. For a recovered text $L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) = \{[\text{girls}], [\text{are}], [\text{playing}]\}$, we have $A_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) = \frac{1+1+1}{1+1+1} = 1$ and $R_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) = \frac{1+1+1}{1+1+1+1} = \frac{3}{4}$, respectively.

Based on (9) and (10), the MSS of recovered text $L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)$ can be given as

$$E_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) = \xi_i \frac{A_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) R_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)}{\varphi A_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) + (1-\varphi) R_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)}, \quad (11)$$

where $\varphi \in (0,1)$ is a weight parameter used to adjust the contributions of semantic accuracy and completeness to the MSS. ξ_i is an additional penalty for short text and can be represented by [28]

$$\xi_i = \begin{cases} 1, & M_i \geqslant N_i, \\ e^{1-\frac{N_i}{M_i}}, & M_i < N_i. \end{cases} \quad (12)$$

Changing the value of ξ_i, each user can efficiently use of the received partial semantic information for text recovery. From (11), we see that the proposed MSS used to evaluate the recovered text can control the tradeoff between the semantic accuracy and the semantic completeness. In particular, for a given token $w'_{i,m}$ in the recovered text $L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)$, if it appears more times in $L'_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)$ than L_i, then $A_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)$ decreases; otherwise, $R_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i)$ decreases.

E. Problem Formulation

Given the defined system model, our goal is to maximize the total MSS of all the texts recovered by the users while satisfying the transmission delay requirement. This maximization problem includes the RB allocation optimization and the partial semantic information determination. The MSS maximization problem is formulated as follows:

$$\max_{\boldsymbol{\alpha}_i, \mathcal{G}'_i} \sum_{i=1}^{U} E_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i), \quad (13)$$

$$\text{s.t. } \alpha_{i,q} \in \{0,1\}, \quad \forall i \in \mathcal{U}, \ \forall q \in \mathcal{Q}, \quad (13a)$$

$$\sum_{q=1}^{Q} \alpha_{i,q} \leqslant 1, \quad \forall i \in \mathcal{U}, \quad (13b)$$

$$\sum_{i=1}^{U} \alpha_{i,q} \leqslant 1, \quad \forall q \in \mathcal{Q}, \quad (13c)$$

$$\frac{Z_i(\mathcal{G}'_i)O}{c_i(\boldsymbol{\alpha}_i)} \leqslant D, \quad \forall i \in \mathcal{U}, \quad (13d)$$

where the constraints in (13a), (13b), and (13c) guarantee that each user can only occupy one RB and each RB can only be allocated to one user for semantic information transmission. The constraint in (13d) is the delay requirement of semantic information transmission. Problem (13) is challenging to solve by the traditional algorithms such as greedy algorithms due to the following reasons. First, from (11), we see that the

Authorized licensed use limited to: Princeton University. Downloaded on November 14,2022 at 17:10:52 UTC from IEEE Xplore. Restrictions apply.

objective function of problem (13) depends not only on the transmitted partial semantic information $\mathcal{G}_i'$ and the RB allocation $\boldsymbol{\alpha}_i$, but also on the text generation model which is implemented by neural networks. Second, the relationship between the non-convex objective function (i.e., the total MSS) and the optimization variables (i.e., $\mathcal{G}_i'$ and $\boldsymbol{\alpha}_i$) cannot be characterized accurately. Third, the coupling between the semantic information selection and the RB allocation further complicates the MSS maximization problem. In particular, the size of the selected partial semantic information for each user depends on the allocated RB. Meanwhile, the RB allocation depends on the importance of different semantic information in terms of the MSS improvement. To solve (13), we propose an attention network based reinforcement learning (RL) algorithm that enables the BS to evaluate the importance of semantic triples and optimize the RB allocation and the partial semantic information selection based on the importance of triples so as to improve the total MSS of all recovered texts.

III. ATTENTION RL FOR SEMANTIC INFORMATION SELECTION AND RESOURCE ALLOCATION

We now introduce a proximal policy optimization (PPO)-based RL algorithm [30] integrated with an attention network [31], called attention proximal policy optimization (APPO). First, for each semantic triple ε_i^g, we use an attention network to calculate the corresponding importance value. Since the triples that are highly correlated with the original text are important for text recovery, the importance of a semantic triple is defined as the correlation between the triple and the original text. Hereinafter, we use an importance vector $\boldsymbol{f}_i(\mathcal{G}_i)$ to represent the *importance distribution* of the triples in each semantic information $\mathcal{G}_i$. Based on the importance evaluation, the proposed APPO algorithm enables the BS to analyze the relationship between the importance distribution $\boldsymbol{f}_i(\mathcal{G}_i)$ and the total MSS so as to optimize the *policy* for the RB allocation and semantic information selection. Here, since the trained text recovery model can be considered as the prior knowledge of the texts to transmitted, it is shared by the users and the BS. Hence, the BS can obtain the total MSS once the RB allocation and the partial semantic information to transmit are determined. We first introduce the use of the attention networks to calculate the importance of the triples in each semantic information. Then, we explain the components of the APPO algorithm and the process of using our proposed APPO algorithm to optimize the RB allocation for each user and determine the partial semantic information to be transmitted. Finally, we show the complexity and convergence of the proposed APPO algorithm.

A. Attention Network for Importance Evaluation

Given semantic information $\mathcal{G}_i$ that consists of G_i triples, an attention network is used to capture the importance of each triple $\varepsilon_i^g \in \mathcal{G}_i$. Then, the importance distribution $\boldsymbol{f}_i(\mathcal{G}_i)$ of semantic information $\mathcal{G}_i$ can be obtained by normalizing the values of importance. In particular, an attention network consists of an input layer, a hidden layer, an output layer, and
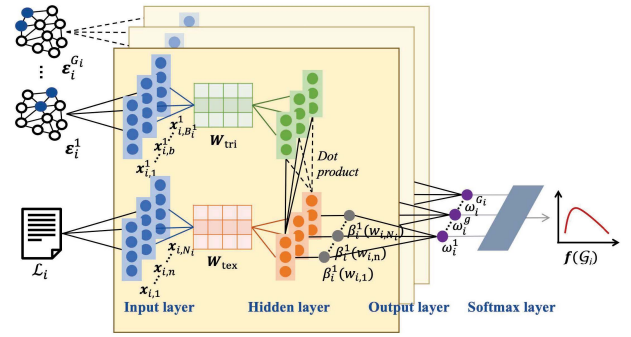

Fig. 3. The architecture of the attention networks.

a softmax layer, as shown in Fig. 3. Next, we introduce each layer of an attention network.

- *Input layer*: To obtain the importance of semantic triple ε_i^g, the BS first needs to vectorize each token $v_{i,b}^g$ in triple ε_i^g and each token $w_{i,n}$ in original text L_i as done in [32]. We define the vector used to represent token $v_{i,b}^g$ as $\boldsymbol{x}_{i,b}^g \in \mathbb{R}^{D_x}$ and the vector used to represent token $w_{i,n}$ as $\boldsymbol{x}_{i,n} \in \mathbb{R}^{D_x}$, where D_x is the dimension of each token vector. The input of an attention network is a sequence of token vectors $\boldsymbol{X}_i^g = (\boldsymbol{x}_{i,1}^g, \ldots, \boldsymbol{x}_{i,b}^g, \ldots, \boldsymbol{x}_{i,B_i^g}^g)$ that represent semantic triple ε_i^g and a sequence of token vectors $\boldsymbol{X}_i^L = (\boldsymbol{x}_{i,1}, \ldots, \boldsymbol{x}_{i,n}, \ldots, \boldsymbol{x}_{i,N_i})$ that represent original text L_i.

- *Hidden layer*: The hidden layer is used to find the correlation between each token $v_{i,b}^g \in \varepsilon_i^g$ and each token $w_{i,n}$ in original text L_i. Given the token vectors, the correlation between $v_{i,b}^g$ and $w_{i,n}$ can be given as

$$\psi(\boldsymbol{x}_{i,b}^g, \boldsymbol{x}_{i,n}) = (\boldsymbol{W}_{\text{tri}}\boldsymbol{x}_{i,b}^g)^{\mathsf{T}}(\boldsymbol{W}_{\text{tex}}\boldsymbol{x}_{i,n}), \qquad (14)$$

where $\boldsymbol{W}_{\text{tri}} \in \mathbb{R}^{D_a \times D_x}$ and $\boldsymbol{W}_{\text{tex}} \in \mathbb{R}^{D_a \times D_x}$ are both parameter matrices of the attention network with $D_a \times D_x$ being the size of the parameter matrices. Then, the correlation between triple ε_i^g in $\mathcal{G}_i$ and token $w_{i,n}$ in L_i can be given as

$$\beta_i^g(\boldsymbol{x}_{i,n}) = \sum_{b=1}^{B_i^g} \frac{\psi(\boldsymbol{x}_{i,b}^g, \boldsymbol{x}_{i,n})}{B_i^g}, \qquad (15)$$

where B_i^g is the number of tokens in triple ε_i^g.

- *Output layer*: The output of an attention network is the importance of semantic triple ε_i^g, as follows:

$$\varsigma_i^g = \sum_{n=1}^{N_i} \beta_i^g(\boldsymbol{x}_{i,n}). \qquad (16)$$

- *Softmax layer*: Since the importance ω_i^g obtained by an attention network can be any value, the BS cannot directly determine the resource allocation and the partial semantic information to be transmitted based on ω_i^g. A softmax layer is used to normalize each importance ω_i^g so as to obtain the importance distribution of the triples in

semantic information $\mathcal{G}_i$ as follows

$$\boldsymbol{f}(\mathcal{G}_i) = \frac{e^{\varsigma_i}}{\sum\limits_{g=1}^{G_i} e^{\varsigma_i^g}} = [y_i^1, \ldots, y_i^g, \ldots, y_i^{G_i}], \quad (17)$$

where e is the Euler number, $\varsigma_i = [\varsigma_i^1, \ldots, \varsigma_i^g, \ldots, \varsigma_i^{G_i}]$. Here, note that $\boldsymbol{f}(\mathcal{G}_i)$ is an importance vector whose length is the number of the semantic triples in $\mathcal{G}_i$. From (17), we see that each element $y_i^g \in [0, 1]$ represents the normalized importance of the corresponding semantic triple ε_i^g and $\sum\limits_{g=1}^{G_i} y_i^g = 1$.

Here, we note that, given the importance distribution $\boldsymbol{f}(\mathcal{G}_i)$, the BS still cannot use the traditional algorithms such as greedy algorithms to directly obtain the optimal RB allocation and determine the partial semantic information to be transmitted. The reason is that the MSS of each user depends on the received semantic information and the partial semantic information that can be transmitted depends on the RB allocation. The BS is not aware of the relationship between each importance distribution $\boldsymbol{f}(\mathcal{G}_i)$ and the optimal RB allocation.

B. Components of the APPO Algorithm

Next, we introduce the components of the proposed APPO algorithm. The APPO algorithm consists of five components: a) agent, b) actions, c) states, d) policy, and e) reward, which are specified as follows:

- *Agent*: Our agent is the BS that determines the RB allocation and the partial semantic information to be transmitted for each user.
- *Actions*: We define the action of the agent as a vector $\boldsymbol{a} = [\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_i, \ldots, \boldsymbol{\alpha}_U]$ that represents the RB allocation for all users. The action space $\mathcal{A}$ is the set of all optional actions that satisfy the constraints in (13). Here, once the RB allocated for each user is determined, the BS can transmit the triples in the corresponding semantic information in the order of the importance until the delay threshold is reached. Hence, given the RB allocation $\boldsymbol{\alpha}_i$, the partial semantic information $\mathcal{G}_i'$ to be transmitted to user i consists of the most important triples in $\mathcal{G}_i$ while satisfying $\frac{Z_i(\mathcal{G}_i')O}{c_i(\boldsymbol{\alpha}_i)} \leqslant D$. For example, we consider $\mathcal{G}_i = \{$("*stochastic lexicon model, used for, speech recognizer*"), ("*baseform of word, part of, stochastic lexicon model*")$\}$ and hence, $\varepsilon_i^1 = ($"*stochastic lexicon model, used for, speech recognizer*") and $\varepsilon_i^2 = ($"*baseform of word, part of, stochastic lexicon model*"). Based on (17), we have $\boldsymbol{f}(\mathcal{G}_i) = [0.7, 0.3]$. We also set that only 10 tokens can be transmitted over the wireless network given the determined RB allocation $\boldsymbol{\alpha}_i$ and the delay constraint D. Then, the BS will transmit the most important triple ε_i^1 to the user i. Hence, each $\mathcal{G}_i'$ is determined based on each $\boldsymbol{\alpha}_i$ in the selected action $\boldsymbol{a}$.
- *States*: The state defined as $\boldsymbol{s} = [\boldsymbol{f}(\mathcal{G}_1), \ldots, \boldsymbol{f}(\mathcal{G}_U)]$ is the importance distribution of all semantic information. The state space $\mathcal{S}$ is a continuous space whose size depends on the number of users and the number of triples in each semantic information.

- *Policy*: The policy is the probability of the agent choosing each action given state $\boldsymbol{s}$. The APPO algorithm uses a deep neural network (DNN) parameterized by $\boldsymbol{\theta}$ to build the relationship between the input state $\boldsymbol{s}$ and the output policy that can achieve the maximum total MSS. Then, the policy can be expressed as $\boldsymbol{\pi_\theta}(\boldsymbol{s}, \boldsymbol{a}) = P(\boldsymbol{a}|\boldsymbol{s})$.
- *Reward*: The reward of choosing action $\boldsymbol{a}$ based on state $\boldsymbol{s}$ is $R(\boldsymbol{a}|\boldsymbol{s}) = \sum\limits_{i=1}^{U} E_i(\boldsymbol{\alpha}_i, \mathcal{G}_i')$ which is the total MSS resulting from action $\boldsymbol{a}$ at state $\boldsymbol{s}$. Since the reward function of the proposed APPO algorithm is equivalent to the objective function of problem (13) and the APPO algorithm aims to maximize the reward, the proposed APPO algorithm can solve the total MSS maximization problem (13).

C. APPO for Total MSS Maximization

Next, we introduce the entire training process for the proposed APPO algorithm that allows us to solve problem (13). The proposed APPO algorithm is trained offline, which means that the RB allocation and semantic information selection policy is trained using the historical data of semantic communications. Different from the attention policy gradient (APG) algorithm [1] that uses a static learning rate for policy update, the proposed APPO algorithm can dynamically adjust the learning rate according to the difference between the old policy and the updated policy. Hence, the proposed APPO algorithm can improve the total MSS over the training process and is guaranteed to converge to a locally optimal solution of problem (13). In addition, different from the APG algorithm of [1] that samples new training data based on the policy updated at each iteration, the proposed APPO enables the BS to perform the RB allocation sampling and the policy update asynchronously. In other words, the proposed APPO algorithm can use historical sampled actions and their rewards for model training thus reducing the overhead caused by continuously sampling new actions when the policy is iteratively updated. In particular, the BS stores a historical policy $\boldsymbol{\pi_{\theta^*}}(\boldsymbol{s}, \boldsymbol{a})$ (i.e., the policy updated at a past iteration) and samples K RB allocation vectors according to $\boldsymbol{\pi_{\theta^*}}(\boldsymbol{s}, \boldsymbol{a})$. The set of collected actions is $\mathcal{K} = \{\boldsymbol{a}_1^*, \ldots, \boldsymbol{a}_k^*, \ldots, \boldsymbol{a}_K^*\}$ that is used to train the policy $\boldsymbol{\pi_\theta}(\boldsymbol{s}, \boldsymbol{a})$ for total MSS maximization. The expected reward of policy $\boldsymbol{\pi_\theta}(\boldsymbol{s}, \boldsymbol{a})$ that the BS aims to optimize is defined as

$$\begin{aligned}
\bar{A}(\boldsymbol{\theta}) &= \mathbb{E}_{\boldsymbol{a} \sim \boldsymbol{\pi_\theta}(\boldsymbol{s}, \boldsymbol{a})} \left(R(\boldsymbol{a}|\boldsymbol{s}) \right), \\
&= \int R(\boldsymbol{a}|\boldsymbol{s}) \boldsymbol{\pi_\theta}(\boldsymbol{s}, \boldsymbol{a}) d\boldsymbol{a}, \\
&= \int R(\boldsymbol{a}|\boldsymbol{s}) \frac{\boldsymbol{\pi_\theta}(\boldsymbol{s}, \boldsymbol{a})}{\boldsymbol{\pi_{\theta^*}}(\boldsymbol{s}, \boldsymbol{a})} \boldsymbol{\pi_{\theta^*}}(\boldsymbol{s}, \boldsymbol{a}) d\boldsymbol{a}, \\
&= \mathbb{E}_{\boldsymbol{a} \sim \boldsymbol{\pi_{\theta^*}}(\boldsymbol{s}, \boldsymbol{a})} \left(R(\boldsymbol{a}|\boldsymbol{s}) \frac{\boldsymbol{\pi_\theta}(\boldsymbol{s}, \boldsymbol{a})}{\boldsymbol{\pi_{\theta^*}}(\boldsymbol{s}, \boldsymbol{a})} \right), \\
&\approx \frac{1}{K} \sum\limits_{k=1}^{K} R(\boldsymbol{a}_k^*|\boldsymbol{s}) \frac{\boldsymbol{\pi_\theta}(\boldsymbol{s}, \boldsymbol{a}_k^*)}{\boldsymbol{\pi_{\theta^*}}(\boldsymbol{s}, \boldsymbol{a}_k^*)}, \quad (18)
\end{aligned}$$

where $\mathbb{E}_{\boldsymbol{a} \sim \boldsymbol{\pi_\theta}(\boldsymbol{s}, \boldsymbol{a})}(\cdot)$ is the expectation with respect to action $\boldsymbol{a}$ following the policy $\boldsymbol{\pi_\theta}(\boldsymbol{s}, \boldsymbol{a})$. In (18), in order to ensure

that the set of sampled actions $\mathcal{D}$ can be used to evaluate and iteratively update the policy $\pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})$, the difference between the stored policy $\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a})$ and the policy $\pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})$ to be updated must be controlled [33]. Hence, the goal of optimizing the policy $\pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})$ is to maximize the total MSS with the penalty that captures the difference between $\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a})$ and $\pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})$. Therefore, we have

$$\max_{\boldsymbol{\theta}} J(\boldsymbol{\theta}), \tag{19}$$

where $J(\boldsymbol{\theta}) = \bar{A}(\boldsymbol{\theta}) - \lambda f_{\mathrm{KL}}\left(\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a}), \pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})\right)$ with λ being the penalty coefficient and $f_{\mathrm{KL}}\left(\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a}), \pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})\right)$ being the Kullback–Leibler divergence (KLD) that represents the difference between $\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a})$ and $\pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})$. At each iteration t, the policy $\pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})$ will be updated using the standard gradient ascent method so as to improve the total MSS. The policy update rule is given as

$$\boldsymbol{\theta}^{(t)} \leftarrow \boldsymbol{\theta}^{(t-1)} + \delta \nabla_{\boldsymbol{\theta}} J(\boldsymbol{\theta}), \tag{20}$$

where $\boldsymbol{\theta}^{(t)}$ is the parameters of the policy at iteration t, δ is the learning rate, and the policy gradient is

$$
\begin{aligned}
\nabla_{\boldsymbol{\theta}} J(\boldsymbol{\theta}) &= \nabla_{\boldsymbol{\theta}} \bar{A}(\boldsymbol{\theta}) - \lambda \nabla_{\boldsymbol{\theta}} f_{\mathrm{KL}}\left(\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a}), \pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})\right), \\
&\approx \frac{1}{K} \sum_{k=1}^{K} R(\boldsymbol{a}_k^* | \boldsymbol{s}) \frac{\nabla_{\boldsymbol{\theta}} \pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a}_k^*)}{\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a}_k^*)} \\
&\quad + \lambda \sum_{k=1}^{K} \pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a}_k^*) \nabla_{\boldsymbol{\theta}} \log \frac{\pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a}_k^*)}{\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a}_k^*)}, \\
&= \frac{1}{K} \sum_{k=1}^{K} \frac{\pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a}_k^*)}{\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a}_k^*)} R(\boldsymbol{a}_k^* | \boldsymbol{s}) \frac{\nabla_{\boldsymbol{\theta}} \pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a}_k^*)}{\pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a}_k^*)} \\
&\quad + \lambda \sum_{k=1}^{K} \nabla_{\boldsymbol{\theta}} \log \pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a}_k^*), \\
&= \frac{1}{K} \sum_{k=1}^{K} \frac{\pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a}_k^*)}{\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a}_k^*)} R(\boldsymbol{a}_k^* | \boldsymbol{s}) \nabla_{\boldsymbol{\theta}} \log \pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a}_k^*) \\
&\quad + \lambda \sum_{k=1}^{K} \nabla_{\boldsymbol{\theta}} \log \pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a}_k^*). \tag{21}
\end{aligned}
$$

From (20) and (21), we can see that the learning rate can be adjusted by the penalty coefficient λ to guarantee the convergence of the APPO algorithm. Hence, after T iterations based on (20), we update the penalty coefficient λ as

$$
\lambda \leftarrow
\begin{cases}
\eta \lambda, & f_{\mathrm{KL}}\left(\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a}), \pi_{\boldsymbol{\theta}^{(T)}}(\boldsymbol{s}, \boldsymbol{a})\right) > \mu_{\mathrm{high}}, \\
\dfrac{\lambda}{\eta}, & f_{\mathrm{KL}}\left(\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a}), \pi_{\boldsymbol{\theta}^{(T)}}(\boldsymbol{s}, \boldsymbol{a})\right) < \mu_{\mathrm{low}}, \\
\lambda, & \text{otherwise},
\end{cases} \tag{22}
$$

where $\mu_{\mathrm{high}} = 1 + \tau$ and $\mu_{\mathrm{low}} = 1 - \tau$ are the thresholds that trigger the update of the penalty coefficient λ and $\eta > 1$ is a coefficient used to adjust the learning rate according to the difference between the stored policy $\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a})$ and the updated policy $\pi_{\boldsymbol{\theta}^{(T)}}(\boldsymbol{s}, \boldsymbol{a})$. Then, the BS replace the stored policy as

$$\boldsymbol{\theta}^* = \boldsymbol{\theta}^{(T)}, \tag{23}$$

and resamples the actions for next T policy update iterations.

Algorithm 1 Training Process of the Proposed APPO Algorithm

1: **Input:** Text L_i required to transmit to each user, delay threshold D, and interference I_q of each RB.
2: **Initialize:** Parameters $\boldsymbol{\theta}^*$ generated randomly, semantic information extraction model, text recovery model, penalty coefficient λ, threshold τ, coefficient η.
3: Obtain the importance distribution $\boldsymbol{f}(\mathcal{G}_i)$ of each semantic information based on (17).
4: **repeat**
5:     Store the policy $\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a})$ and collect K trajectories $\mathcal{K} = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_K\}$ using $\pi_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a})$.
6:     **for** $i = 1 \rightarrow T$ **do**
7:       Update the parameters of the policy $\pi_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})$ based on (20).
8:     **end for**
9:     Update the penalty coefficient λ based on (22).
10:    Replace the stored policy based on (23).
11: **until** the objective function defined in (19) converges.

By iteratively updating the policy until the proposed APPO algorithm converges, the policy parameter $\boldsymbol{\theta}$ can find the relation between the importance distributions of all semantic information and the total MSS. Hence, the policy for RB allocation and semantic information selection that can achieve maximum total MSS of all recovered texts can be obtained [34]. The specific training process of the proposed APPO algorithm is summarized in **Algorithm 1**.

D. Complexity and Convergence of APPO

Next, we analyze the computational complexity and the convergence of the proposed APPO algorithm for resource allocation and partial semantic information determination. The complexity of the APPO algorithm lies in calculating the importance distribution by using the attention network and determining the resource allocation by using the trained policy. We first analyze the complexity of the attention network-based semantic importance evaluator. From (15), the complexity of calculating the correlation between semantic triple ε_i^g and token $w_{i,n}$ in original text L_i is $\mathcal{O}\left(B_i^g(2D_a D_x + D_a^2)\right)$, where $D_a \times D_x$ is the size of parameter matrices $\boldsymbol{W}_{\mathrm{tri}}$ and $\boldsymbol{W}_{\mathrm{tok}}$ of the attention network. We assume that $D_a \ll D_x$ (e.g., $D_a = 64$ and $D_x = 500$ as done in [27]). Hence, we have $\mathcal{O}\left(B_i^g(2D_a D_x + D_a^2)\right) = \mathcal{O}\left(B_i^g D_a D_x\right)$. From (16) and (17), the complexity of using the attention network to calculate the importance distribution for all semantic information can be given as $\mathcal{O}\left(\sum_{i=1}^{U} \sum_{g=1}^{G_i} N_i B_i^g D_a D_x\right)$. Next, we investigate the complexity for resource allocation that depends on the size of the policy parameter $\boldsymbol{\theta}$. The size of the policy parameter $\boldsymbol{\theta}$ depends on the size of action space $\mathcal{A}$ and the size of state space $\mathcal{S}$. The action space $\mathcal{A}$ is a set of all possible resource allocation whose size is $|\mathcal{A}| = \frac{U!}{(U-Q)!}$. Based on (17), the size of state space $\mathcal{S}$ is $|\mathcal{S}| = \sum_{i=1}^{U} G_i$. Then, the complexity of

TABLE II
SYSTEM PARAMETERS [15], [30], [32]

Parameters	Description	Value
Q	Number of RBs	10
W	Bandwidth of each RB	2 MHz
P	Transmit power of the BS	1 W
N_0	Noise power spectral density	-174 dBm/Hz
D	Transmission delay threshold	0.1 ms
O	Data size of each token	80 bits
D_a	Dimension of hidden layer in attention network	64
φ	Weight parameter in (11)	0.5
L	Number of layers in DNN used to output policies	3
K	Batch size of policy update	100
η	Coefficient used to adjust the learning rate	2
D_x	Dimension of token vector	500
τ	Threshold of triggering the penalty coefficient	0.8

using the trained policy to determine the resource allocation is $\mathcal{O}\left(\frac{U!}{(U-Q)!}\left(\sum_{i=1}^{U} G_i\right)\prod_{l=2}^{L-1} H_l\right)$, where L is the number of layers in deep neural network used to train the policy and H_l is the number of the neurons in layer l. In consequence, the computational complexity of the proposed APPO algorithm for resource allocation and semantic information selection will be

$$\mathcal{O}\left(\sum_{i=1}^{U}\sum_{g=1}^{G_i} N_i B_i^g D_a D_x + \frac{U!}{(U-Q)!}\left(\sum_{i=1}^{U} G_i\right)\prod_{l=2}^{L-1} H_l\right). \tag{24}$$

From (24), given the structure of the proposed APPO model (i.e., D_a, D_x, and H_l), the complexity of APPO depends on the number of users, the number of RBs, and the number of triples in each extracted semantic information.

With regards to the convergence of the proposed APPO algorithm, we can directly use the result of [33, Theorem 1] which shows that the policy of an APPO algorithm will converge to a locally optimal solution when it satisfies the following conditions.

Lemma 1 (Follows From [33]): The proposed APPO algorithm used to solve the total MSS maximization problem (13) converges if the following conditions are satisfied:

 i) $J(\boldsymbol{\theta})$ defined in (19) is a lower bound function of the objective function $\bar{A}(\boldsymbol{\theta})$ defined in (18) that the BS aims to optimize.
 ii) The actions sampled according to a stored policy $\boldsymbol{\pi}_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a})$ can be used to evaluate and iteratively update the policy $\boldsymbol{\pi}_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})$.
 iii) The value of the objective function $\bar{A}(\boldsymbol{\theta})$ monotonically increases as the policy $\boldsymbol{\theta}$ is iteratively updated based on (20).

Proof: Next, we prove that our proposed algorithm satisfies the conditions from [33]. For condition i), the design of the lower bound function $J(\boldsymbol{\theta})$ follows [33] and, hence, the proof in [33] is still hold. For condition ii), the difference between the stored policy $\boldsymbol{\pi}_{\boldsymbol{\theta}^*}(\boldsymbol{s}, \boldsymbol{a})$ and the policy $\boldsymbol{\pi}_{\boldsymbol{\theta}}(\boldsymbol{s}, \boldsymbol{a})$ to be updated is controlled by the penalty of the Kullback Leibler divergence as we specified in (19). For condition iii), the learning rate of the APPO algorithm can be adjusted based

on (22) so that the value of the objective function $\bar{A}(\boldsymbol{\theta})$ can monotonically increase according to the proof in [33]. ∎

From Lemma 1, we see that, the convergence of the proposed APPO algorithm depends on the lower bound function $J(\boldsymbol{\theta})$ and the dynamic learning rate. In particular, as the policy is iteratively updated, the lower bound function $J(\boldsymbol{\theta})$ locally approximates the objective function $\bar{A}(\boldsymbol{\theta})$ and the learning rate decreases until the APPO algorithm converges to the locally optimal solution of problem (13). According to [33], the proposed APPO algorithm for solving (13) will converge.

IV. SIMULATION RESULTS AND ANALYSIS

In our simulations, a circular network is considered with one BS and $U = 30$ uniformly distributed users. Other parameters are listed in Table II. We use the semantic information extraction model in [24] and the text recovery model in [27]. The text datasets used to train the proposed APPO algorithm are the abstract generation dataset (AGENDA) [35] which consists of 40 thousand paper titles and abstracts from the proceedings of 12 top artificial intelligence conferences and the DocRED dataset [36] which consists of 5053 Wikipedia documents. For comparison purposes, we consider five baselines: baseline a) being the APG algorithm proposed in [1], baseline b) being a deep Q network integrated with attention network (ADQN) algorithm, baseline c) that optimizes RB allocation using the proposed RL solution and directly transmits the original text data, baseline d) that optimizes RB allocation using the proposed RL solution and randomly selects semantic triples to be transmitted, and baseline e) being the semantic framework proposed in [4]. Here, for baseline c), the BS transmits the original texts to the users token by token until the delay threshold is reached [37]. The MSS measures the semantic similarity between the original text and the partial text received by each user.

Figure 4 shows the correlation (as defined in (15)) between a semantic triple and the tokens in the original text. In Fig. 4, the importance of the triple "(*stochastic lexicon model*, *used for*, *speech recognize*)" is shown as the sum of correlations (i.e., (16)). In particular, we use different colors to represent the correlation between the semantic triple ("*stochastic lexicon model*, *used for*, *speech recognize*") and different tokens in the original text. As the correlations between the semantic triple

An semantic triple in the semantic information

$\varepsilon_i^g =$ (*stochastic lexicon model, used for, speech recognizer*)

The correlation $\beta_i^g(x_{i,n})$ between ε_i^g and each token $w_{i,n}$ in the original text L_i

We propose the stochastic lexicon model which represents the pronunciation variations to optimally cope with the continuous speech recognizer. In this lexicon model, the baseform of words are represented by subword states and the probability distribution of subwords as a hidden Markov model. Also, the proposed approach can be applied to a system employing non-linguistic recognition units and the lexicon is automatically trained from training utterances. In speaker independent speech recognition tests using a 3000 word continuous speech database, the proposed system improves the word accuracy by about 27.8% and the sentence accuracy by about 22.4%.
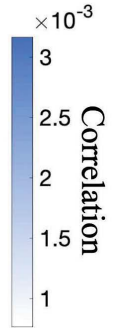
Fig. 4. The correlation between an example of the semantic triple and the original text.

Original text ($N_i = 178$)

Model migration in speaker recognition is a task of converting parametrically-obsolete models to new structures and configurations without the requirement to store the original speech waveforms or feature vector sequences along with the models. This paper reports on further enhancements of a statistical migration technique based on Gaussian mixture models, introduced previously. The present approach is based on a stochastic synthesis of feature sequences from obsolete models that are subsequently used to create the new models. Here, in addition to Gaussian means and priors, as utilized in the previous contribution, also the covariance are included resulting in significant performance gains in the migrated models, compared to the mean-only method. Overall, measured on the NIST 2003 cellular task, the described algorithm achieves a model migration incurring a loss in performance of 8-20% relative to a full re-enrollment from waveforms, dependent on the type of mismatch between the obsolete and the new configuration. The inclusion of the covariance information is shown to reduce the loss of performance by a factor of 3-4 as compared to the baseline mean-only migration technique.

Semantic information ($Z(\mathcal{G}_i) = 86$)

(*statistical migration technique, used for, model migration*),
(*mean-only method, compared with, statistical migration technique*),
(*loss of performance, evaluate for, statistical migration technique*),
(*model migration, used for, speaker recognition*),
(*Gaussian mixture models, used for, statistical migration technique*),
(*feature vector sequences, used for, statistical migration technique*),
(*NIST 2003 cellular task, evaluate for, statistical migration technique*),
(*model migration, conjunction, converting parametrically-obsolete models*),
(*covariance, used for, statistical migration technique*),
(*Gaussian means, used for, statistical migration technique*),
(*priors, used for, statistical migration technique*),
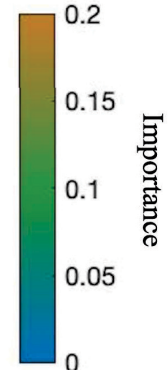(*feature vector sequences, feature of, speech waveforms*),

Fig. 5. An example of the original text and the extracted semantic information.

and the tokens in the original text increase, the color used to mark the tokens changes from white to blue. From Fig. 4, we can see that the expression "*the proposed approach*" in the fourth line of the original text is highly correlated with the example triple "(*stochastic lexicon model, used for, speech recognizer*)". This implies that the attention networks can extract correlations between the triples and the original text according to the meaning of the tokens in the context. This is due to the fact that, compared with the traditional approaches that obtain the correlation between two token sequences by

word frequency statistics, the attention networks can generate the representation for each token according to the context of the token and calculate the correlations based on these representations. In consequence, the attention networks enable the BS to obtain the importance of the triples in the semantic information.

Figure 5 shows an example of the original text, the extracted semantic information $\mathcal{G}_i$, and the importance distribution of the semantic triples in $\mathcal{G}_i$. In Fig. 5, the entities recognized in the original text are shown in red and the extracted

Proposed APPO ($\mathcal{G}'_i$)	*(statistical migration technique, used for, model migration),* *(mean-only method, compared with, statistical migration technique),* *(loss of performance, evaluate for, statistical migration technique),* *(model migration, used for, speaker recognition),* *(Gaussian mixture models, used for, statistical migration technique),* *(feature vector sequences, used for, statistical migration technique),* *(NIST 2003 cellular task, evaluate for, statistical migration technique),*
Recovered text	in this paper, we propose a STATISTICAL MIGRATION TECHNIQUE for MODEL MIGRATION in SPEAKER RECOGNITION. the proposed STATISTICAL MIGRATION TECHNIQUE is based on GAUSSIAN MIXTURE MODELS. FEATURE VECTOR SEQUENCES is included in the proposed STATISTICAL MIGRATION TECHNIQUE. the proposed STATISTICAL MIGRATION TECHNIQUE is evaluated on the NIST 2003 CELLULAR TASK. the experimental results show that the proposed STATISTICAL MIGRATION TECHNIQUE is effective in reducing the LOSS OF PERFORMANCE compared to the MEAN-ONLY METHOD.
MSS	$E_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) = 0.4552$

(a) The solution of the proposed APPO algorithm.

Baseline d)	*(model migration, used for, speaker recognition),* *(model migration, conjunction, converting parametrically-obsolete models),* *(feature vector sequences, feature of, speech waveforms),* *(Gaussian mixture models, used for, statistical migration technique),* *(Gaussian means, used for, statistical migration technique),* *(covariance, used for, statistical migration technique),* *(NIST 2003 cellular task, evaluate for, statistical migration technique),*
Recovered text	in this paper , we propose a MODEL MIGRATION that is a task of CONVERTING PARAMETRICALLY-OBSOLETE MODELS for SPEAKER RECOGNITION. the proposed STATISTICAL MIGRATION TECHNIQUE is based on GAUSSIAN MIXTURE MODELS, GAUSSIAN MEANS and COVARIANCE. The FEATURE VECTOR SEQUENCES are features of SPEECH WAVEFORMS. the proposed STATISTICAL MIGRATION TECHNIQUE is evaluated on the NIST 2003 CELLULAR TASK.
MSS	$E_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) = 0.3324$

(b) The solution of the baseline a).

Baseline c)	Model migration in speaker recognition is a task of converting parametrically-obsolete models to new structures and configurations without the requirement to store the original speech waveforms or feature vector sequences along with the models. This paper reports on further enhancements of a statistical migration technique based on Gaussian mixture models,
MSS	$E_i(\boldsymbol{\alpha}_i, \mathcal{G}'_i) = 0.1037$

(c) The solution of the baseline b).

Fig. 6. An example of the transmitted data and the corresponding recovered text.

semantic triples are written in different colors according to their importance. In particular, as the importance of a triple increases, the color of that semantic triple changes from blue to yellow. For example, the importance of the yellow triple (i.e., "*statistical migration technique, used for, model migration*") is 0.1923 while the importance of the blue triple (i.e., "*feature vector sequences, feature of, speech waveforms*") is 0.0385. The sum of the importance of all the triples in a semantic information is 1. From Fig. 5, we can see that, using our proposed semantic communication framework, the BS needs to transmit only 86 tokens of the semantic information instead of 178 tokens of the original text data to the user. Therefore, in this example, the proposed framework can reduce the size of the data that needs to be transmitted by 51.7%.
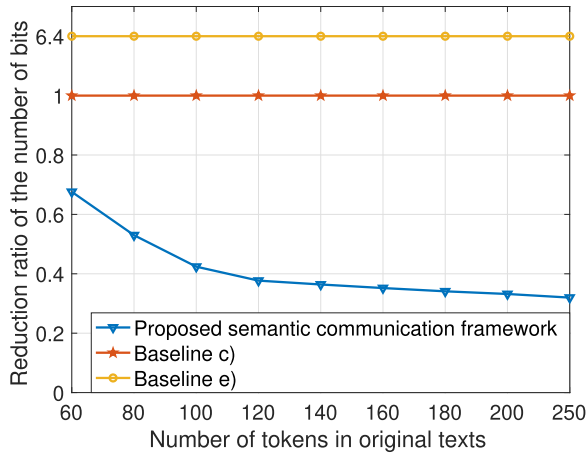
Fig. 7. The reduction ratio of the data size using the semantic communication framework.

Figure 6 shows the partial semantic information that will be transmitted and the corresponding recovered text for the example in Fig. 5. Fig. 6(a) shows the partial semantic information $\mathcal{G}_i'$ selected by the proposed APPO algorithm and the text recovered using $\mathcal{G}_i'$. In Fig. 6(a), the BS uses the APPO algorithm to select the most important triples in the extracted semantic information and transmits them to the user. Figs. 6(b) and 6(c) show the results of baselines a) and b). From Figs. 6(a) and 6(b), we see that the proposed algorithm can improve 36.7% total MSS compared to baseline a). This is because the proposed APPO algorithm can evaluate the importance of the triples in the semantic information and, then determines the RB allocation and partial semantic information to be transmitted. Therefore, the text recovered using the most important semantic triples contains the main meaning of the original text. From Figs. 6(a) and 6(c), we observe that the proposed algorithm can improve 2-fold total MSS compared to baseline b). This gain stems from the fact that the proposed semantic communication framework can extract the meaning of the original data and transmit it to each user.

Figure 7 shows how the ratio of the data size of the semantic information to the data size of the original text changes as the number of tokens in the original text changes. We assume that the data size of an english letter is 8 bits and each token consists of 10 letters. Hence, the number of bits used to encode a token is 80 (i.e., in (7), $O = 80$) [38]. Baseline e) uses 16-dimensional feature vectors extracted by the deep neural networks to represent the semantic information of the tokens in the original text [4]. We also assume that a decimal consists of 32 bits. From Fig. 7 we can see that, compare to baseline c) and baseline e), the proposed semantic communication framework can reduce the size of the data that needs to be transmitted by up to 41.3% and 84%, respectively The reason is that our proposed framework enables the BS to extract the meaning of the texts and model the extracted meaning by a knowledge graph. From Fig. 7, we also observe that, as the number of tokens in the original texts increases, the ratio of the size of the semantic information to the size of the original text decreases. This is due to the fact that the semantic information in multiple sentences in long texts can be expressed by a single

semantic triple. Hence, extracting the semantic information from long texts can significantly reduce the number of tokens that need to be transmitted.

Figure 8 shows examples of the texts recovered by the graph-to-text generation models that are trained via the loss function different metrics. In particular, we include the result of the graph-to-text generation model that aims to maximize the proposed MSS and BLEU in [28]. From Fig. 8, we can see that the text recovered using the MSS-based text generation model contains all semantic triples in the received semantic information and all semantic triples is arranged with articulated textual context. In contrast, the text recovered using the BLEU-based text generation model misses the semantic triple "*feature vector sequences, used for, statistical migration technique*". This is due to the fact that, compared to the BLEU metric that only captures the accuracy of the recovered text, the proposed MSS can capture both accuracy and completeness of the recovered text. From Fig. 8, we can also see that, compared to the BLEU-based text generation model that copies semantic triples from the received semantic information for text recovery, the MSS-based text generation model can generate additional descriptions based on the context of the semantic triples. For example, the MSS-based text generation model generates "*the experimental results show that...*". This is because an additional penalty for short text is considered in the definition of MSS.

Figure 9 shows the convergences of the APPO, APG, and DQN integrated with attention networks (ADQN) algorithms in a semantic communication-enabled network with 10 RBs. The line and shadow are the mean and standard deviation computed over 10 independent runs. From Fig. 9, we observe that, compared to the APG and ADQN algorithms that require about 300 iterations and 250 iterations to reach convergence, respectively, the proposed APPO algorithm converges after 100 iterations. This stems from the fact that the proposed APPO algorithm can adjust the learning rate at each iteration so as to speed up the convergence. The improvement in the convergence speed indicates that the proposed APPO algorithm can reduce the time and energy overhead of the training process in practical semantic communication systems. Figure 9 also shows that, the proposed APPO algorithm can improve the MSS by 11.2% and 18.5%, respectively, compared with the APG algorithm and ADQN algorithm. The reason is that, when the updated policy approximates the locally optimal solution, the APPO algorithm can further improve the total MSS by decreasing the learning rate. In practice, this indicates that the APPO algorithm enables the BS to effectively allocate the limited RBs according to the importance distribution of semantic information so as to improve the performance of the proposed semantic communication framework. In Fig. 9, we can also see that, the APPO algorithm achieves up to 53.3% and 60.1% reduction in the standard deviation of the total MSS compared to the APG algorithm. In practice, this implies that, the performance of the proposed APPO algorithm is more stable than that of the APG and ADQN algorithms.

Figure 10 shows how the total MSS changes as the number of RBs varies. This figure is simulated using the DocRED dataset. From Fig. 10, we can see that the proposed APPO

Received $\mathcal{G}_i'$	*(statistical migration technique, used for, model migration),* *(mean-only method, compared with, statistical migration technique),* *(loss of performance, evaluate for, statistical migration technique),* *(model migration, used for, speaker recognition),* *(Gaussian mixture models, used for, statistical migration technique),* *(feature vector sequences, used for, statistical migration technique),* *(NIST 2003 cellular task, evaluate for, statistical migration technique),*
Text recovered based on MSS($\varphi = 0.5$)	in this paper, we propose a STATISTICAL MIGRATION TECHNIQUE for MODEL MIGRATION in SPEAKER RECOGNITION. the proposed STATISTICAL MIGRATION TECHNIQUE is based on GAUSSIAN MIXTURE MODELS. FEATURE VECTOR SEQUENCES is included in the proposed STATISTICAL MIGRATION TECHNIQUE. the proposed STATISTICAL MIGRATION TECHNIQUE is evaluated on the NIST 2003 CELLULAR TASK. the experimental results show that the proposed STATISTICAL MIGRATION TECHNIQUE is effective in reducing the LOSS OF PERFORMANCE compared to the MEAN-ONLY METHOD.
Text recovered based on BLEU ($\varphi = 0$)	we propose a STATISTICAL MIGRATION TECHNIQUE for MODEL MIGRATION. the MODEL MIGRATION is used for SPEAKER RECOGNITION. the proposed STATISTICAL MIGRATION TECHNIQUE is based on GAUSSIAN MIXTURE MODELS. the STATISTICAL MIGRATION TECHNIQUE is evaluated on the NIST 2003 CELLULAR TASK and achieves higher LOSS OF PERFORMANCE than the MEAN-ONLY METHOD.

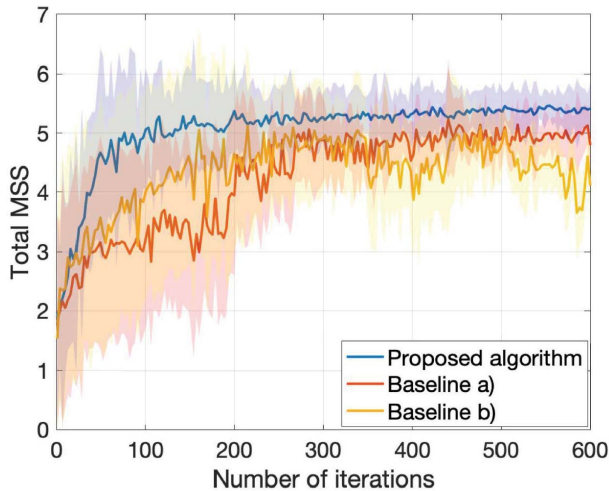Fig. 8.   The text recovered based on different metric.

Fig. 9.   The training process of the considered algorithms.

Fig. 10.   The total MSS as the number of RBs varies.

algorithm can improve 17.8%, 26.8% and 2-fold total MSS on DocRED dataset compared to baselines a), b) and c), respectively. The 17.8% and 26.8% gains stem from the fact that the proposed APPO algorithm dynamically adjusts the learning rate during training process. Hence, the APPO algorithm is guaranteed to obtain a locally optimal RB allocation policy for partial semantic information transmission. The 2-fold% gain stems from the fact that the proposed algorithm enables the BS to transmit the most important semantic triples using the limited wireless resource thus significantly improving the transmission efficiency.
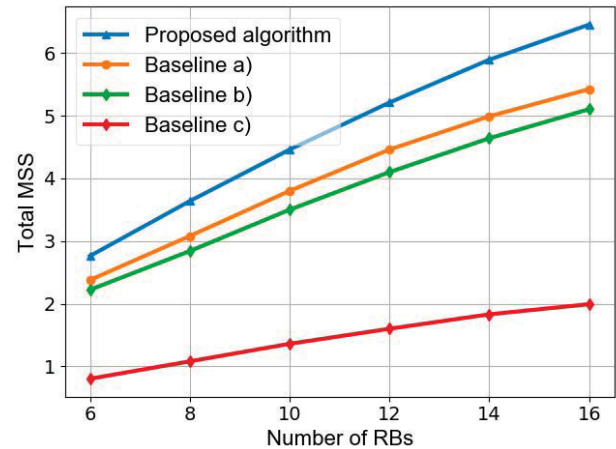
Figure 11 shows how the total MSS changes as the number of users varies. This figure is simulated using the AGENDA dataset. From Fig. 11, we can see that the proposed algorithm can serve 30 users. From Fig. 11, we can also see that, the proposed APPO algorithm can improve 16.9%, 22.3% and 1-fold total MSS compared to baselines a), b) and c), respectively. The 16.9% and 22.3% gains stem from the fact that the proposed APPO algorithm can build the relationship between the importance distribution of the semantic information and the total MSS, thus effectively finding the locally optimal RB allocation policy for partial semantic information transmission. The 1-fold% gain stems from the fact that the proposed
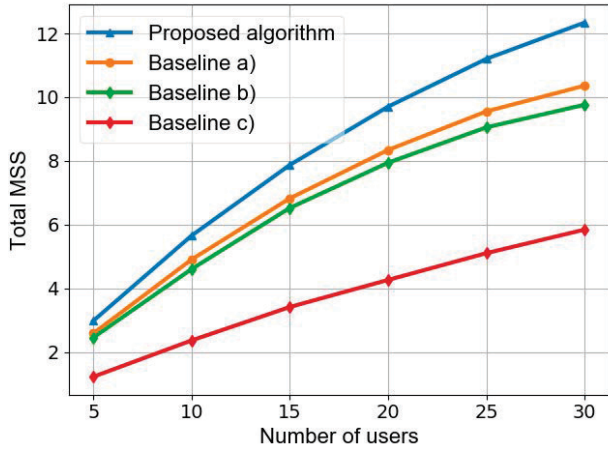
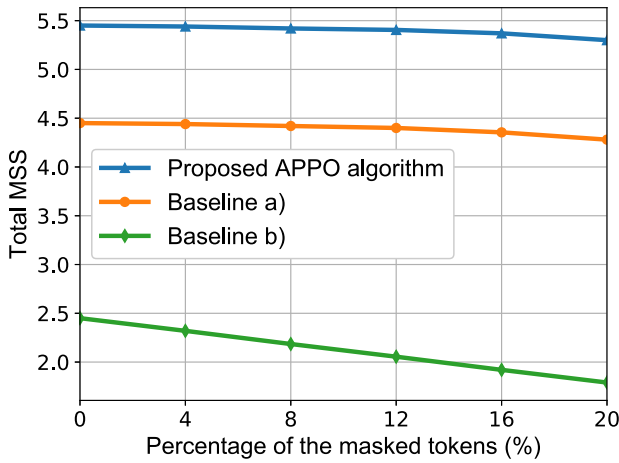Fig. 11. The total MSS as the number of users varies.

Fig. 12. The total MSS as the percentage of the masked tokens varies.

algorithm uses the knowledge graph to model the semantic information and transmit the important semantic triples thus significantly improving the transmission efficiency.

Figure 12 shows how the total MSS changes as the percentage of the masked tokens varies. In Fig. 12, we randomly mask a subset of the tokens in the original texts in the AGENDA dataset to verify the robustness of the proposed semantic communication framework. The obtained texts with masks are used to simulate texts containing typos in actual scenarios. From Fig 12, we can see that, as the percentage of masked tokens increases, the total MSS of the proposed APPO algorithm, baseline a), and baseline b) decreases by 2.75%, 3.8%, and 67%, respectively. The reason why the total MSS of the proposed APPO algorithm and baseline a) remain stable is because the occasional masks do not change the semantic information of the original texts. The semantic information extraction and original text recovery approaches enable the proposed semantic communication framework to ignore the masks in original texts. The 67% MSS reduction of baseline b) stems from the fact that the standard communication networks directly transmit the texts with masks to each user.

V. CONCLUSION

In this paper, we have proposed a semantic driven wireless networks. We have modeled the semantic information of the textual data by a KG. To measure the performance of the semantic communications, we have introduced a new metric, MSS, that captures the semantic accuracy and semantic completeness between the original text and the recovered text. We have jointly considered the wireless resource limitations, transmission delay requirements, and the performance of the semantic communications and formulated an optimization problem whose goal is to maximize the total MSS by optimizing the RB allocation for partial semantic information transmission. To solve this problem, we have developed an APPO algorithm that can obtain the importance distribution of the triples in the semantic information and then build the relationship between the importance distribution and the total MSS. Hence, the proposed APPO algorithm enables the BS to find the policies for RB allocation and semantic information selection for maximizing the total MSS. Compared with the traditional RL algorithms, the proposed APPO dynamically adjusts the learning rate during training process, thus improving the total MSS, the convergence speed, and the stability. Simulation results have demonstrated that the proposed semantic communication framework can significantly reduce the size of data required to transmit and increase the total MSS.

REFERENCES

[1] Y. Wang, M. Chen, W. Saad, T. Luo, S. Cui, and H. V. Poor, "Performance optimization for semantic communications: An attention-based learning approach," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Madrid, Spain, Dec. 2021, pp. 1–6.

[2] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, May 2020.

[3] E. Calvanese Strinati and S. Barbarossa, "6G networks: Beyond Shannon towards semantic and goal-oriented communications," *Comput. Netw.*, vol. 190, May 2021, Art. no. 107930.

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

[5] G. Shi, Y. Xiao, Y. Li, and X. Xie, "From semantic communication to semantic-aware networking: Model, architecture, and open problems," *IEEE Commun. Mag.*, vol. 59, no. 8, pp. 44–50, Aug. 2021.

[6] H. Xie and Z. Qin, "A lite distributed semantic communication system for Internet of Things," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 142–153, Jan. 2021.

[7] E. Uysal *et al.*, "Semantic communications in networked systems: A data significance perspective," 2021, *arXiv:2103.05391*.

[8] M. Kountouris and N. Pappas, "Semantics-empowered communication for networked intelligent systems," *IEEE Commun. Mag.*, vol. 59, no. 6, pp. 96–102, Jun. 2021.

[9] G. Shi *et al.*, "A new communication paradigm: From bit accuracy to semantic fidelity," 2021, *arXiv:2101.12649*.

[10] J. Bao *et al.*, "Towards a theory of semantic communication," in *Proc. IEEE Netw. Sci. Workshop*, West Point, NY, USA, Jun. 2011, pp. 110–117.

[11] A. Maatouk, M. Assaad, and A. Ephremides, "The age of incorrect information: An enabler of semantics-empowered communication," 2020, *arXiv:2012.13214*.

[12] Z. Weng and Z. Qin, "Semantic communication systems for speech transmission," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 8, pp. 2434–2444, Aug. 2021.

[13] B. Guler, A. Yener, and A. Swami, "The semantic communication game," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 4, pp. 787–802, Dec. 2018.

[14] O. Semiari, W. Saad, M. Bennis, and M. Debbah, "Integrated millimeter wave and sub-6 GHz wireless networks: A roadmap for joint mobile broadband and ultra-reliable low-latency communications," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 109–115, Apr. 2019.

[15] M. Chen, Z. Yang, W. Saad, C. Yin, H. V. Poor, and S. Cui, "A joint learning and communications framework for federated learning over wireless networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 269–283, Oct. 2021.

[16] Y. Li, H. Zhang, and K. Long, "Joint resource, trajectory, and artificial noise optimization in secure driven 3-D UAVs with NOMA and imperfect CSI," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 11, pp. 3363–3377, Nov. 2021.

[17] S. Wang *et al.*, "Distributed reinforcement learning for age of information minimization in real-time IoT systems," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 3, pp. 501–515, Apr. 2022.

[18] W. Ahsan, W. Yi, Z. Qin, Y. Liu, and A. Nallanathan, "Resource allocation in uplink NOMA-IoT networks: A reinforcement-learning approach," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5083–5098, Aug. 2021.

[19] Z. Xiong, J. Kang, D. Niyato, P. Wang, and H. V. Poor, "Cloud/edge computing service management in blockchain networks: Multi-leader multi-follower game-based ADMM for pricing," *IEEE Trans. Services Comput.*, vol. 13, no. 2, pp. 356–367, Apr. 2020.

[20] M. Karimzadeh, W. Saad, and M. Debbah, "Common language for goal-oriented semantic communications: A curriculum learning framework," in *Proc. IEEE Int. Conf. Commun.*, Seoul, South Korea, May 2022.

[21] D. Michelsanti *et al.*, "An overview of deep-learning-based audio-visual speech enhancement and separation," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 29, pp. 1368–1396, 2021.

[22] F. Liu, E. Bugliarello, E. M. Ponti, S. Reddy, N. Collier, and D. Elliott, "Visually grounded reasoning across languages and cultures," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, Punta Cana, Dominican Republic, 2021, pp. 10467–10485.

[23] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, "A survey on knowledge graphs: Representation, acquisition, and applications," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 2, pp. 494–514, Feb. 2022.

[24] Y. Luan, L. He, M. Ostendorf, and H. Hajishirzi, "Multi-task identification of entities, relations, and coreference for scientific knowledge graph construction," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, Brussels, Belgium, 2018, pp. 1–14.

[25] Y. Lin, S. Shen, Z. Liu, H. Luan, and M. Sun, "Neural relation extraction with selective attention over instances," in *Proc. 54th Annu. Meeting Assoc. Comput. Linguistics*, Berlin, Germany, Aug. 2016, pp. 2124–2133.

[26] M. Chen, N. Shlezinger, H. V. Poor, Y. C. Eldar, and S. Cui, "Communication-efficient federated learning," *Proc. Nat. Acad. Sci. USA*, vol. 118, no. 17, Apr. 2021, Art. no. e2024789118.

[27] R. Koncel-Kedziorski, D. Bekal, Y. Luan, M. Lapata, and H. Hajishirzi, "Text generation from knowledge graphs with graph transformers," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics*, Minneapolis, MN, USA, Jun. 2019.

[28] K. Papineni, S. Roukos, T. Ward, and W. Zhu, "BLEU: A method for automatic evaluation of machine translation," in *Proc. Annu. Meeting Assoc. Comput. Linguistics*, Philadelphia, PA, USA, Jul. 2002, pp. 311–318.

[29] S. Banerjee and A. Lavie, "METEOR: An automatic metric for MT evaluation with improved correlation with human judgments," in *Proc. Annu. Meeting Assoc. Comput. Linguistics Workshop*, Ann Arbor, MI, USA, Jun. 2005.

[30] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017, *arXiv:1707.06347*.

[31] A. Vaswani *et al.*, "Attention is all you need," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, Long Beach, CA, USA, Dec. 2017, pp. 1–11.

[32] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.

[33] J. Schulman, S. Levine, P. Moritz, M. Jordan, and P. Abbeel, "Trust region policy optimization," in *Proc. 32nd Int. Conf. Mach. Learn.*, Lille, U.K., Jul. 2015, pp. 1889–1897.

[34] Y. Hu, M. Chen, W. Saad, H. V. Poor, and S. Cui, "Distributed multi-agent meta learning for trajectory design in wireless drone networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 3177–3192, Oct. 2021.

[35] W. Ammar *et al.*, "Construction of the literature graph in semantic scholar," 2018, *arXiv:1805.02262*.

[36] Y. Yao *et al.*, "DocRED: A large-scale document-level relation extraction dataset," in *Proc. 57th Annu. Meeting Assoc. for Comput. Linguistics*, Florence, Italy, 2019, pp. 1–14.

[37] C. Liu, C. Guo, Y. Yang, and N. Jiang, "Adaptable semantic compression and resource allocation for task-oriented communications," 2022, *arXiv:2204.08910*.

[38] International Organization for Standardization and International Electrotechnical Commission, *Information Technology—ISO 8-bit Code for Information Interchange—Structure and Rules for Implementation*, Standard ISO/IEC 4873:1991, 1991.

Yining Wang (Student Member, IEEE) received the Ph.D. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2022. From 2021 to 2022, she was a Visiting Researcher with the School of Computer Science and Engineering, Nanyang Technological University. Her research interests include semantic communications, virtual reality, wireless networks, machine learning, and reinforcement learning.

Mingzhe Chen (Member, IEEE) received the Ph.D. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2019. From 2016 to 2019, he was a Visiting Researcher with the Department of Electrical and Computer Engineering, Virginia Tech. From 2019 to 2021, he was a Post-Doctoral Research Associate with the Department of Electrical and Computer Engineering, Princeton University. In 2022, he worked as an AI Researcher at Ericsson Research, USA. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering and Institute for Data Science and Computing, University of Miami. His research interests include federated learning, reinforcement learning, virtual reality, unmanned aerial vehicles, and the Internet of Things. He has received from the IEEE Communication Society two paper awards including the 2021 Young Author Best Paper Award and the 2022 Fred W. Ellersick Prize Award, and three Conference Best Paper Awards at IEEE ICC in 2020, IEEE GLOBECOM in 2020, and IEEE WCNC in 2021. He currently serves as an Associate Editor for IEEE TRANSACTIONS ON GREEN COMMUNICATIONS AND NETWORKING. Previously, he guest edited a special issue on Distributed Learning over Wireless Edge Networks for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC).

Tao Luo (Senior Member, IEEE) received the B.S. degree from Shaanxi Normal University in 1993, the M.S. degree from Shanghai University in 1999, and the Ph.D. degree from the Beijing University of Posts and Telecommunications (BUPT), China, in 2002. He is currently a Professor at BUPT. His research interests include mobile communication, semantic communications, the internet of vehicles, and machine learning.
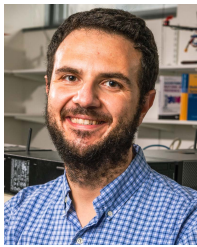
Walid Saad (Fellow, IEEE) received the Ph.D. degree from the University of Oslo in 2010. He is currently a Professor with the Department of Electrical and Computer Engineering, Virginia Tech, where he leads the Network sciEnce, Wireless, and Security (NEWS) Laboratory. His research interests include wireless networks (5G/6G/beyond), machine learning, game theory, security, unmanned aerial vehicles, semantic communications, cyber-physical systems, and network science. He was the author/coauthor of 11 conference best paper awards at WiOpt in 2009, ICIMP in 2010, IEEE WCNC in 2012, IEEE PIMRC in 2015, IEEE SmartGridComm in 2015, EuCNC in 2017, IEEE GLOBECOM in 2018, IFIP NTMS in 2019, IEEE ICC in 2020 and 2022, and IEEE GLOBECOM in 2020. He was a recipient of the NSF CAREER Award in 2013, the AFOSR Summer Faculty Fellowship in 2014, and the Young Investigator Award from the Office of Naval Research (ONR) in 2015. He was also a recipient of the 2015 and 2022 Fred W. Ellersick Prize from the IEEE Communications Society, the 2017 IEEE ComSoc Best Young Professional in Academia Award, the 2018 IEEE ComSoc Radio Communications Committee Early Achievement Award, and the 2019 IEEE ComSoc Communication Theory Technical Committee. He was also the coauthor of the 2019 IEEE Communications Society Young Author Best Paper and the 2021 IEEE Communications Society Young Author Best Paper. From 2015 to 2017, he was named the Stephen O. Lane Junior Faculty Fellow at Virginia Tech and, in 2017, he was named the College of Engineering Faculty Fellow. He received the Dean's Award for Research Excellence from Virginia Tech in 2019. He was also an IEEE Distinguished Lecturer in 2019 and 2020. He currently serves as an Editor for the IEEE TRANSACTIONS ON MOBILE COMPUTING and the IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING. He is an Area Editor of the IEEE TRANSACTIONS ON NETWORK SCIENCE AND ENGINEERING, an Associate Editor-in-Chief of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (JSAC)— Special issue on Machine Learning for Communication Networks, and an Editor-at-Large of the IEEE TRANSACTIONS ON COMMUNICATIONS. He is the Editor-in-Chief of the IEEE TRANSACTIONS ON MACHINE LEARNING IN COMMUNICATIONS AND NETWORKING.

Dusit Niyato (Fellow, IEEE) received the B.Eng. degree from the King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand, in 1999, and the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Canada, in 2008. He is currently a Professor with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interests are in the areas of the Internet of Things (IoT), machine learning, and incentive mechanism design.
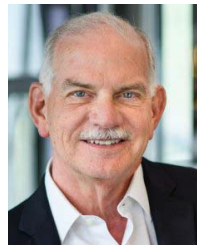
H. Vincent Poor (Life Fellow, IEEE) received the Ph.D. degree in EECS from Princeton University in 1977. From 1977 to 1990, he was on the faculty of the University of Illinois Urbana–Champaign. Since 1990, he has been on the faculty at Princeton, where he is currently the Michael Henry Strater University Professor. From 2006 to 2016, he served as the Dean of Princeton's School of Engineering and Applied Science. He has also held visiting appointments at several other universities, including most recently at Berkeley and Cambridge. His research interests are in the areas of information theory, machine learning and network science, and their applications in wireless networks, energy systems and related fields. Among his publications in these areas is the forthcoming book *Machine Learning and Wireless Communications*. (Cambridge University Press). Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and he is also a foreign member of the Chinese Academy of Sciences, the Royal Society, and other national and international academies. He received the IEEE Alexander Graham Bell Medal in 2017.

Shuguang Cui (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Stanford University, California, USA, in 2005. Afterwards, he has been working as an Assistant Professor, an Associate Professor, a Full Professor, and a Chair Professor in electrical and computer engineering with The University of Arizona, Texas A&M University, UC Davis, and CUHK at Shenzhen, respectively. He has also served as the Executive Dean for the School of Science and Engineering, CUHK, Shenzhen, and the Executive Vice Director for the Shenzhen Research Institute of Big Data. His current research interests focus on data driven large-scale system control and resource management, large data set analysis, IoT system design, energy harvesting based-communication system design, and cognitive network optimization. He was selected as the Thomson Reuters Highly Cited Researcher and listed in the Worlds' Most Influential Scientific Minds by ScienceWatch in 2014. He was a recipient of the IEEE Signal Processing Society 2012 Best Paper Award. He was the Elected Member of IEEE Signal Processing Society SPCOM Technical Committee from 2009 to 2014 and the Elected Chair of IEEE ComSoc Wireless Technical Committee from 2017 to 2018. He is a member of the Steering Committee for IEEE TRANSACTIONS ON BIG DATA and the Chair of the Steering Committee for IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING. He was also a member of the IEEE ComSoc Emerging Technology Committee. He was elected as an IEEE ComSoc Distinguished Lecturer in 2014 and IEEE VT Society Distinguished Lecturer in 2019. In 2020, he won the IEEE ICC Best Paper Award, ICIP Best Paper Finalist, and the IEEE Globecom Best Paper Award. In 2021, he won the IEEE WCNC Best Paper Award. He has served as the general co-chair and the TPC co-chair for many IEEE conferences. He has also been serving as the Area Editor for *IEEE Signal Processing Magazine* and an Associate Editor for IEEE TRANSACTIONS ON BIG DATA, IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Series on Green Communications and Networking, and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.