A Novel Approach on Semantic Performance Oriented Radio Resource Allocation

Zhanjun Liu, Kaixuan Zhao, Ao Xiao, Chengchao Liang(✉), and Rong Chai

School of Communications and Information Engineering,
Chongqing University of Posts and Telecommunications, Chongqing, China
{liuzj,chairong}@cqupt.edu.cn, liangcc@cqupt.edu.cn.com

Abstract. Recent research has made considerable progress in the field of Semantic Communication (SemCom), demonstrating its effectiveness in reliable transmission under low signal-to-noise ratio (SNR) conditions. The latest technique considers the semantic spectral efficiency (S-SE) of text information and related resource allocation problems. However, in radio access network (RAN) environments, it is usually necessary to allocate physical resource blocks (PRBs) to users rather than simply assigning channels. Therefore, further research is needed to address the resource allocation problem for SemCom in RAN environments. This paper reconsiders and solves the resource allocation optimization problem based on SemCom, optimizing resource allocation from two perspectives: allocating PRBs to users and transferring semantic symbols. Additionally, two conversion methods are proposed to compare the performance of different systems and allocation methods. The simulation results validate the proposed resource allocation method and the effectiveness of SemCom in terms of S-SE.

Keywords: semantic communication · resource allocation · semantic spectral efficiency · physical resource block

1 Introduction

From the first generation of wireless communication (1G) to the current fifth generation (5G), the transmission rate has increased tens of thousands of times and the system capacity has come close to the Shannon limit [1]. In the modern era of machine intelligence, where the Internet of Things and Artificial Intelligence are developing in tandem, Shannon's classical information theory is beginning to show its limitations, as it does not take into account the semantics and meaning of information in communication [2,3]. With the increasing demand

This work was supported by the Natural Science Foundation of Chongqing under Grant (cstc2020jcyj-msxmX0878, CSTB2023NSCQ-MSX0726).

F. Gao et al. (Eds.): Chinacom 2023, LNICST 590, pp. 21–35, 2024.
https://doi.org/10.1007/978-3-031-67162-3_2

for intelligent wireless communication applications, the future communication network will move away from the traditional architecture that simply pursues a high transmission rate to a new architecture that focuses on the intelligent connection of everything [4]. Semantic communication systems use a background knowledge base or a neural network model to transmit semantic information, which has the potential to reduce network traffic and alleviate spectrum shortage [5,6]. The semantic information of different types of information sources has its own characteristics and processing methods, such as text [6–8], image [9–11], audio [12], etc. A variety of Semantic Communication (SemCom) systems are being studied for different sources of information to ensure that different types of semantic information can be accurately and efficiently transmitted.

Currently, only a few studies have been conducted on the resource allocation of SemCom. In particular, the channel allocation problem based on semantic awareness in the radio access network (RAN) was discussed in depth in [6]. However, in the current deployment, RAN does not just assign channels to users, but allocates network slices to them and assigns physical resource blocks (PRBs) to them in the slices to transmit data, in order to achieve a more flexible and efficient deployment [13,14]. Therefore, it is essential to investigate the allocation of SemCom resources and enhance communication efficiency by using SemCom while ensuring reliable transmission. Based on [6], this paper further explores the semantic performance-oriented PRB assignment problem in RAN.

The challenge of allocating resources in wireless communication is to measure the amount of information and spectral efficiency (SE). Traditional communication relies on bit streams to transmit data, which is based on statistical knowledge of the source symbols rather than the semantic information of the source. Therefore, SemCom cannot use bit streams to measure information content and SE. To address this issue, we must reconsider the allocation of resources from the perspective of semantic theory. Existing basic research of semantic theory provides a great help in this regard, but the past work is mainly based on abstract models, which cannot quantify semantic information. Fortunately, the development of deep learning can make up for the lack of research on SemCom theory and mathematical models, and the emergence of SemCom systems based on deep learning makes it possible to measure SE in the semantic domain. In particular, the Deep learning-based Semantic communication system (DeepSC) [7] and its variants [15] can effectively extract semantic information from text and deliver meaning successfully to the recipient. This paper will explore the problem of SE and the problem of resource allocation for semantic information based on DeepSC. The main contributions of this paper are as follows.

- The existing work is used to define the semantic spectrum efficiency (S-SE), which is an expression used to measure communication efficiency from a semantic perspective. On the basis of this, a new formula is proposed and solved to maximize the S-SE of the system. This formula takes into account the allocation of physical resource blocks and the number of transmitted semantic symbols.
- A comparison of the performance of various SemCom systems and traditional communication systems is conducted by introducing a bit-based SE conver-

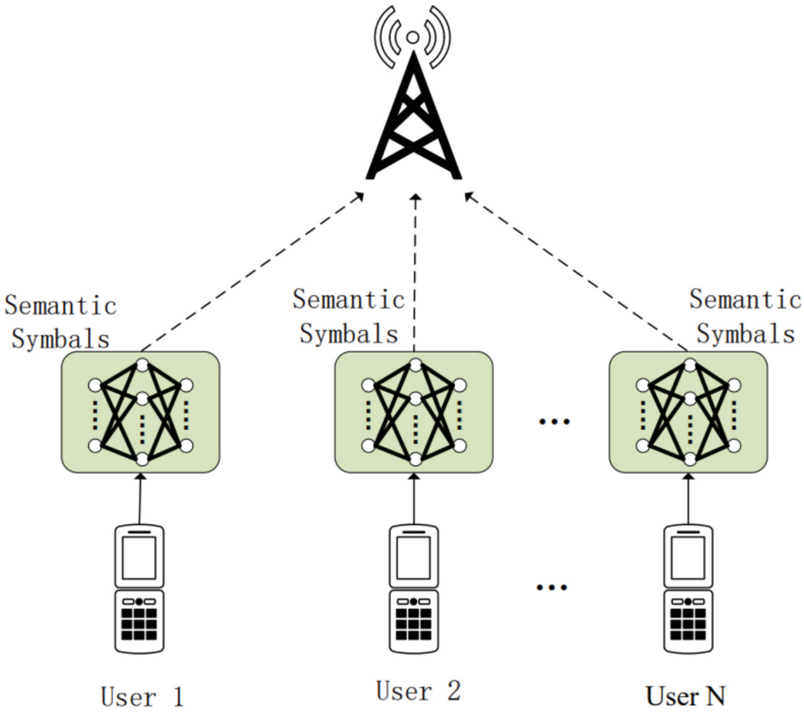

Fig. 1. System Model

sion approach to S-SE and a performance index conversion technique between the allocated channel and the allocated PRB resource allocation methods.
- The proposed resource allocation model has been confirmed to be valid through the simulation results, and the SemCom system has been demonstrated to be superior in S-SE.

The structure of this paper is organized as follows. We provide a detailed description to the system model in Sect. 2. Section 3 describe the optimization problem in detail, along with the optimal solution based on heuristics and matching algorithms. We demonstrate the results of our solution and comparisons with traditional schemes in Sect. 4. Finally, we discuss our work and conclude our paper in Sect. 5.

2 System Model

We analyze a cellular network composed of a base station (BS) and a set of users $\mathcal{N} = \{1, 2, \dots, n, \dots, N\}$ as depicted in Fig. 1. This network uses DeepSC [7] as a SemCom model for the transmission of textual data to each user. The DeepSC model is pre-trained on the BS, edge server, or cloud platform, and the trained semantic transmitter model is broadcasted to the users in the area through the base station when necessary. We will divide the entire system into three components: the DeepSC transmitter at the user, the transmission model, and the DeepSC receiver at the BS.

2.1 DeepSC Transmitter

We must first take into account user behavior. The user n needs to send a sentence $s_0 = [a_{n,1}, a_{n,2}, \ldots, a_{n,l}, \ldots, a_{n,L_n}]$. At the user n, $a_{n,l}$ is the l^{th} word and L_n is the length of the sentence. This sentence is then put into the DeepSC emitter and converted into a semantic symbol vector $Z_n = [z_{n,1}, z_{n,2}, \ldots, z_{n,k_n L_n}]$. This vector is a real matrix of size $k_n L_n \times 2$, where $k_n L_n$ is the length of the vector of semantic symbols created by the sentence in the n^{th} user, and k_n is the average number of semantic symbols used per word at the n^{th} user. We believe that the SemCom system transmits semantic symbols directly.

2.2 Transmission Model

The number of PRBs X_n required by each user to send semantic information is counted, and then the total number X of PRBs needed by all users is calculated. The set of channels available on the network is denoted as $\mathcal{M} = \{1, 2, \ldots, m, \ldots, M\}$, where M is the number of channels. This leads to the number of time slots T occupied by PRBs required to send information, and PRBs are allocated in these time slots to users. The PRB assignment matrix for the n-th user is denoted as $\omega_n = [\omega_{n,1}^1, \ldots, \omega_{n,m}^1, \ldots, \omega_{n,M}^1; \omega_{n,1}^t, \ldots, \omega_{n,m}^t, \ldots, \omega_{n,M}^t; \omega_{n,1}^T, \ldots, \omega_{n,m}^T, \ldots, \omega_{n,M}^T]$, where $\omega_{n,m}^t \in \{0, 1\}, t \in (0, T), m \in (0, M)$. This is a matrix of dimension $M \times T$. If the t-th slot of the m -th channel is assigned to the user n, then $\omega_{n,m}^t = 1$, and $\omega_{n,m}^t = 0$, otherwise. Each PRB can be assigned to only one user, and each user can only use the required number of PRBs X_n.

In addition, we believe that all channels are composed of large-scale fading and small-scale Rayleigh fading. The signal-to-noise ratio(SNR) of all PRBs for the n-th user on the m-th channel is

$$\gamma_{n,m} = \frac{p_n g_n |h_{n,m}|^2}{W N_0} \tag{1}$$

where p_n is the transmitting power of n-th user, g_n is the large-scale channel gain of n-th user, including path loss and shadow effect, $h_{n,m} \sim \mathcal{CN}(0, 1)$ is the Rayleigh fading coefficient transmitted by n-th user on the m-th channel, and N_0 is the noise power spectral density.

2.3 DeepSC Receiver

The signal received by the base station from the n-th user can be expressed as $\mathbf{Y}_n = \sqrt{g_n} h_{n,m} \mathbf{X}_n + \mathbf{z}$, where $\mathbf{z}$ is additive white Gaussian noise (AWGN) and each element of $\mathbf{z}$ follows a $\mathcal{CN}(0, N_0)$ distribution. The received signal is first decoded by the channel decoder and then by the semantic decoder to recover the sentence. We use semantic similarity [6] as the performance metric to evaluate the SemCom performance of text transmission.

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

where $\mathbf{B}(\cdot)$ is the Sentence-Bidirectional Encoder Representations from Transformers(BERT) model. Compared to other metrics for assessing semantic performance, such as BLEU [16], semantic similarity can be a more precise way to measure the disparity in semantic content between two sentences, particularly for synonyms and near synonyms. According to [6], semantic similarity is measured on a scale of $0 \leqslant \varepsilon \leqslant 1$, where $\varepsilon = 1$ implies the highest similarity between two sentences and $\varepsilon = 0$ implies no similarity.

3 Problem Formulation

In this section, S-SE [6] is defined as a new performance metric for SemCom systems. The SemCom resource allocation is reduced to the S-SE maximization problem based on PRB allocation and the number of semantic symbols per word. Finally, the solution of the optimization problem is obtained.

3.1 Semantic Spectral Efficiency

In order to assess the transmission performance of SemCom, a new performance metric, S-SE, has been defined. Unlike traditional communication systems, which measure spectral efficiency in terms of bits per second per hertz, this metric is designed to measure the transmission rate of semantic information. This is because the sequence of bits generated in traditional communication systems is based on statistical knowledge of the source, and is independent of the meaning of the source. To measure semantic information, we assume that it can be expressed in terms of a semantic unit (sut), which is the basic unit of semantic information. Based on this, two key semantically based performance metrics have been defined.

- *Semantic transmission rate* (*S-R*): the effective transfer of semantic information per second, expressed in $suts/s$.
- *Semantic spectral efficiency* (*S-SE*): the rate at which semantic information can be successfully transmitted within a unit bandwidth, expressed in $suts/s/Hz$.

We calculate the expressions for S-R and S-SE. Let $\mathcal{D} = \{s_j\}_{j=1}^{D}$ be the text information database of size D, where s_j is the j-th sentence and its length is L_j. The semantic information of s_j is denoted as I_j and its probability is $p(\mathbf{s}_j)$. The expected semantic information of each sentence is $I = \sum_{j=1}^{D} I_j p(\mathbf{s}_j)$ and the expected number of words is $L{=}\sum_{j=1}^{D} L_j p(\mathbf{s}_j)$. For statistical convenience, we take the expected values I and L instead of random values to obtain more efficient data and conclusions. In the n-th user, the average semantic information carried by each semantic symbol $k_n L$ is I and the average semantic information carried by each symbol is $I/(k_n L)$. The symbol rate of the passband transmission is equal to the channel bandwidth and the SemCom system transmits semantic symbols directly, so the total semantic information transmitted on the channel with bandwidth W is WI/(knL). Therefore, the S-R of the n-th user on the m-th channel is given by WI/(knL).

$$\Gamma_{n,m} = \frac{WI}{k_n L}\varepsilon_{\mathrm{n,m}}, \tag{3}$$

The semantic similarity of the n-th user using the m-th channel, denoted as $\varepsilon_{\mathrm{n,m}}$, is dependent on the neural network structure of DeepSC and the channel conditions, which is a function of k_n and γ. This leads to the expression of S-SE shown below.

$$\Phi_{n,m} = \frac{\Gamma_{n,m}}{W} = \frac{I}{k_n L}\varepsilon_{\mathrm{n,m}}. \tag{4}$$

Based on the S-SE definition, the performance of each PRB is equal to the S-SE of the channel in which it is located.

3.2 Problem Formulation

In this section, a model of resource allocation is proposed to maximize the semantic performance and the overall satisfaction of all users of the system. By representing Φ as the total S-SE for all users in T time slots, we have

$$\Phi = \sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{n,m}^{t}\frac{I}{k_n L}\varepsilon_{n,m}. \tag{5}$$

In order to maximize the potential of DeepSC in low SNR conditions, we first focus on the PRB allocation vector as the optimization target. Additionally, since the value of k_n has an effect on transmission efficiency and accuracy, with larger k_n reducing semantic transmission efficiency and smaller k_n reducing the semantic similarity of transmission, we also optimize k_n of different users. This way, a higher S-SE can be achieved while still ensuring transmission reliability. According to the above analysis, the optimization problem can be expressed as:

$$\text{(P0)} \quad \max_{w_{\mathrm{n},m}^{\mathrm{t}},k_n} \sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{n,m}^{t}\frac{I}{k_n L}\varepsilon_{n,m} \tag{6}$$

$$\text{s.t.} \quad \sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{n,m}^{t} \geq X_n, \forall n \in N, \tag{6a}$$

$$\sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{m=1}^{M} \omega_{n,m}^{t} \leq X, \forall n \in N, \forall m \in M, \forall t \in T, \tag{6b}$$

$$\sum_{n=1}^{N} \omega_{n.m}^{t} \leq 1, \forall m \in M, \forall t \in T, \tag{6c}$$

$$\sum_{n=1}^{N}\sum_{m=1}^{M} \omega_{n,m}^{t} \leq M, \forall t \in T, \tag{6d}$$

$$\omega_{n,m}^{t} \in \{0,1\}, \forall n \in N, \forall m \in M, \forall t \in T, \tag{6e}$$

$$\varepsilon_{n,m} \geq \varepsilon_{th}, \tag{6f}$$

$$\Phi_{n,m} = \frac{\varepsilon_{n,m}}{k_n} \geq \Phi_{\text{th}}, \tag{6g}$$

$$1 \leq k_n \leq K, \tag{6h}$$

where (6a) limits the number of PRBs allocated to a single user to meet the needs of users where X_n represent the number of PRBs required by n-th user, (6b) limits the total number of PRBs that can be allocated to all users to ensure that the transmission is completed in limited time slots, where $X = T \times M$ and the sum of X_n for all users requires a total of T time slots to provide the required PRBs, (6c) states that each PRB can only be assigned to one user. These three restrictions describe the restriction on assigning PRBs to users, with X_n PRBs allocated to the n-th user within a specified T time slot and not repeated, where X_n and T can be calculated in simulation. (6d) sets a limit of the number of PRBs in a time slot, (6e) is a binary allocation variable and it equal to 1 only if n-th user uses PRB in time slot t of m-th channel. Additionally, (6f) and (6g) specify the S-R and S-SE thresholds for the successful transmission of textual semantic information. If the value is lower than the threshold, the receiver is unable to interpret it and considers it a transmission failure. (6h) determines the range of k_n and K is the maximum value.

3.3 The Optimal Solution

For problem (**P0**), we observe that the variable I/L is dependent on the type of source. As per the examination in Section III-A, this term is a fixed value for certain types of source and does not influence resource optimization. Consequently, this term can be ignored when solving (**P0**). Subsequently, the optimization problem can be reformulated as:

$$\text{(P1)} \quad \max_{w^{\mathrm{t}}_{\mathrm{n,m}}, k_n} \sum_{n=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} \omega_{n,m}^{t} \frac{\varepsilon_{n,m}}{k_n} \tag{7}$$
$$\text{s.t.} \quad (6a), (6b), (6c), (6d), (6e), (6f), (6g), (6h)$$

In order to address problem (**P1**), we must tackle two issues. The first is how to handle $\omega_{n,m}^{t}$, which is closely related to (6a), (6b), (6c), and (6d). The second is how to select the parameter k_n to maximize the total S-SE of all users while satisfying the constraints (6f), (6g), and (6h). It is important to note that users will select PRBs on different channels and that each user can select only one k_n value for the processing of semantic information in the system. The choice of parameter k_n in the second subproblem will affect the selection result of $\omega_{n,m}^{t}$ in the first subproblem. To solve these two subproblems, we must first determine the relationship between semantic similarity $\varepsilon_{n,m}$ and k_n and $\gamma_{n,m}$. Since $\varepsilon_{n,m}$ depends on the specific SemCom system and the physical channel conditions, we run the DeepSC model on the AWGN channel to obtain the mapping between $\varepsilon_{n,m}$ and $(k_n, \gamma_{n,m})$.

Once the initial k_n is given, we can compute the semantic similarity of semantic transmission through different channels for each user with the current value of k_n and then calculate the semantic spectral efficiency based on the outcomes of the semantic similarity. Subproblem 1 can be viewed as a one-to-many matching problem, which obtains the PRB allocation state with the highest system performance. Subproblem 2 is a global optimization issue with the objective function of the energy of the system of subproblem 1 and the vector k_n as an independent variable.

According to matching theory, subproblem 1 must have a pareto optimal solution where the total allocation results of all users are optimal. This problem is similar to the college admission problem proposed in [17], so the Deferred-Acceptance (DA) algorithm can be used to solve it. For the global optimization problem such as subproblem 2, a heuristic algorithm can be employed to maximize the objective function, that is, the system performance of subproblem 1. In this paper, particle swarm optimization (PSO) is used to solve subproblem 2. Finally, the optimal solution of the optimization problem is obtained by combining the PSO and DA algorithms.

Based on our PSO-DA algorithm, we can get the best PRB allocation scheme for the current task to maximize the semantic performance of the whole system. For semantic communication, our solution is easy to expand, because the existing semantic communication models, including text, voice, image and other source types, are mostly based on neural network to extract semantic symbols for transmission. There is no essential difference in the process of transmission, so this scheme can be easily extended to other source types of transmission tasks.

4 Simulation Results and Analysis

In order to assess the effectiveness of the proposed semantic performance-oriented resource allocation scheme, we conduct a simulation. We compare the proposed semantic performance-oriented resource allocation system for PRB allocation with the reference semantic-aware resource allocation model for channel allocation, and compare the S-SE of the two SemCom systems with the traditional communication system to demonstrate the superiority of SemCom and verify the superiority of this SemCom system. Since traditional systems are usually evaluated in the bit domain, we first use the conversion method proposed in [6] to convert traditional S-E into S-SE, thus making it possible to make a fair comparison between SemCom systems and traditional systems. Secondly, we propose a transformation method to compare the performance of SemCom systems that allocate channels with those that allocate PRBs through quantitative analysis. We then present and analyze the simulation results.

4.1 Compare Method Between Semantic Systems

In traditional communication, a source encoder assigns each letter in a word to a bit. From a semantic perspective, each bit may contain less semantic information than the symbols of DeepSC, but it can be thought of as a semantic symbol. The S-R in one channel can be represented as a sequence of bits in another channel.

$$\Gamma'_{n,m} = C_{n,m} \frac{I}{\mu L} \varepsilon_{n,m}, \tag{8}$$

where $C_{n,m}$ is the transmission rate of the n-th user on the m-th channel, expressed in bit/s. μ is defined as a conversion factor that represents the ability of the source encoding scheme to compress the data, representing the average number of bits per word, in $bits/word$. Specifically, if a word contains an average of five letters and each letter is encoded in ASCII, then $\mu = 40\, bit/word$. Furthermore, when we assume that there are no error codes in traditional communication, then $\varepsilon_{n,m} = 1$. Let us denote S-E as $\mathrm{R_{n,m} = C_{n,m}/W}$, then the equivalent S-SE is

$$\Phi'_{n,m} = R_{n,m} \frac{I}{\mu L}. \tag{9}$$

By taking into account both the source coding and the bit transmission, the S-SE of traditional systems can be determined to make a fair comparison between different systems.

4.2 Compare Method Between Allocation Patterns

The primary distinction between the semantic performance-oriented resource allocation approach presented in this paper and the existing work [6] is that the units of spectrum resources allocated are different. This paper assigns PRBs to users on demand, while the scheme proposed in [6] assigns channels to users.

Table 1. Simulation Parameters

Parameter	Value
Number of users, N	5
Number of Channels, M	10
Channel bandwidth, W	180 KHz
Noise power spectral density, N_0	−174 dBm/Hz
Pathloss model	128.1+37.6lg [d(km)] dB
Shadow effect factor	6 dB
Transmit power, p_n	10 dBm
Maximum number of symbols per word, K	20 symbols/word
Semantic similarity threshold, ε_{th}	0.9
S-SE threshold, Φ_{th}	0.025(I/L) suts/s/Hz
Transforming factor, μ	40 bits/word

To compare the performance difference between two kinds of SemCom system with different resource allocation methods, we propose a S-SE transformation method between the two types of resource allocation methods. Specifically, for the proposed SemCom system to allocate PRBs to users on demand, the S-SE can be directly calculated based on the PRB allocation state. Regarding the resource allocation method proposed in [6], the number of PRBs required by each user to transmit encoded semantic symbols can be determined from the solved channel allocation state and the k_n of each user, and then the S-SE of the system can be obtained. In particular, for the general system, since a semantic symbol can be approximately equivalent to a bit, the number of PRBs occupied by the information of each user after encoding can be calculated from the value of μ, and then the equivalent S-SE, which is related to the value of μ, can be obtained.

It should be noted that the S-SE between these two distributions has distinct significance. The fundamental definition of S-SE is the rate at which semantic information can be effectively transmitted per unit of bandwidth. For traditional channel allocation strategies, the total S-SE of each channel can be used to calculate the rate of semantic information that can be successfully transmitted by the entire system bandwidth, i.e. the sum of S-SE multiplied by the bandwidth of a single channel. For the SemCom system proposed in this paper, the sum of S-SE of each PRB is used to determine the amount of semantic information to be sent that can be successfully transmitted, i.e., the sum of S-SE of each PRB multiplied by PRB duration multiplied by PRB bandwidth.

4.3 Simulation Parameters

We compare our proposed resource allocation scheme, which is designed for the DeepSC semantic system, to three benchmarks: an existing SemCom system, an ideal system, and a system that has been widely implemented in practice.

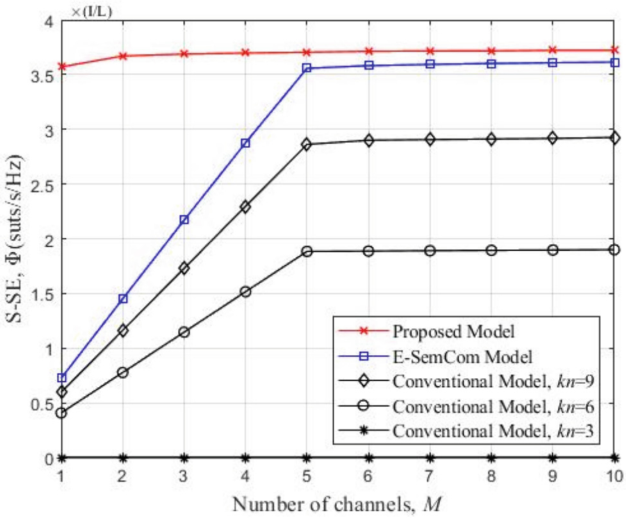

Fig. 2. The S-SE of semantic network between different models.

- *Exist semantic communication(E-SemCom) system*: Referring to the system model proposed in [6], channels are assigned to users.
- *Ideal system*: The Shannon limit can be realized without error, i.e. $R_{n,m} = \log_2(1 + \gamma_{n,m})$.
- *4G system*: Based on the measured SNR, the BS obtains the channel quality indicator (CQI) [18], and the realizable S-E $R_{n,m}$ can be obtained.

We simulated a circular network with a radius of 500 m, in which an equal number of users were placed. Each user sent a random amount of text sentence data. The simulation environment parameters used are listed in Table 1. In addition, there are some parameters that need to be calculated specifically according to the user tasks in the simulation, such as the number of PRBs required by the user tasks, the number of time slots that specify the completion of all transmission tasks, etc. There are no specific values, so they are not listed here.

In this paper, we focus on S-SE to assess its performance. Using the transformed technique developed, we obtain the S-SE optimization problem of the three benchmarks mentioned above and obtain the results reported in [6]. Subsequently, we compare the simulation results of the various systems.

4.4 Simulation Results

We initially analyze the traditional resource allocation model in SemCom systems. Through this simulation, the optimal channel assignment of the traditional model is applied to the ideal system with different values of k_n and the E-SemCom system in the network. Subsequently, the S-SE obtained is compared with the S-SE of the proposed model. As illustrated in Fig. 2, firstly, regardless

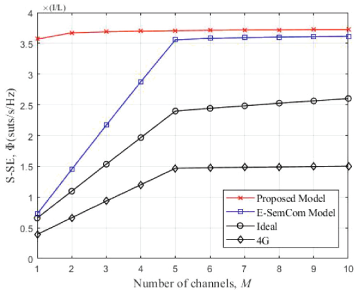

Fig. 3. The S-SE varies with the number of channels between different kinds of systems.

of the value of k_n, the S-SE of the proposed model is higher than the S-SE of the E-SemCom model, both of which are higher than the S-SE of traditional model. Secondly, the system performance of the proposed model does not vary significantly with the number of channels, since the system assigns PRBs to users and does not reject user service requests when the number of channels is less than the number of users. Furthermore, the S-SE of a conventional model with $k_n = 3$ is equal to 0, since in this case the semantic similarity is lower than the threshold and the semantic information is not transmitted successfully.

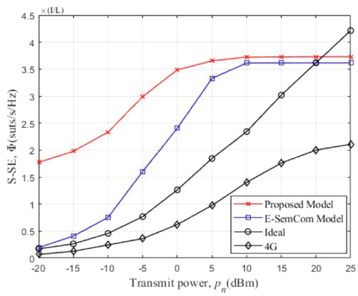

Fig. 4. The S-SE varies with the transmit power between different kinds of systems.

We compare different communication systems and resource allocation methods in the following. Figure 3 illustrates the S-SE in relation to the number of channels for different systems. As the number of channels, M, increases from 1 to 5, the proposed system performance does not vary significantly since no user requests are put on hold, while the S-SE of other systems increases rapidly as more users are served. As M continues to increase from 5 to 10, all system S-SEs grow slowly due to the availability of more channels and the ability of users to select channels with higher SNR. For any value of M, the performance of the proposed model is better than that of other systems. In addition, SemCom systems outperform all traditional communication systems.

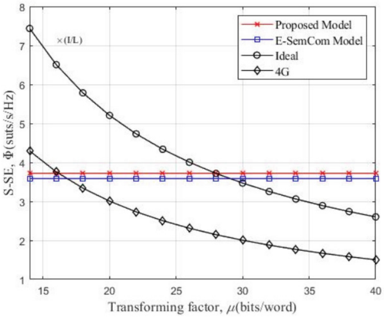

Fig. 5. The S-SE varies with the transforming factor between different kinds of systems.

Figure 4 shows the correlation between S-SE and transmit power. As p_n increases, the S-SE of the ideal system increases rapidly, while the S-SE of the SemCom system and the 4G system initially increases and then stabilizes, indicating that the performance of all practical systems has a ceiling as the SNR increases. The proposed system is able to allocate resources more efficiently when needed, and its performance exceeds that of the E-SemCom system which assigns channels to users. Furthermore, the SemCom system performs better than 4G due to its more powerful data compression ability.

Figure 5 illustrates the correlation between S-SE and the transformation factor μ. Since μ is not dependent on the SemCom systems, its performance remains consistent and the proposed model is more effective than the E-SemCom model. For traditional systems, S-SE decreases as μ increases, since S-SE is the ratio of SE to μ. When μ is less than 28 bits/word, that is, a word can be encoded as less than 28 bits, SemCom systems perform worse than ideal systems. When μ is greater than approximately 16 bits/word, the SemCom systems perform better than the 4G systems. The figure demonstrates that the superiority of the SemCom system over the traditional system is largely determined by the source coding scheme adopted by the traditional system.

5 Conclusion

This paper examines the semantic performance-oriented resource allocation problem for the application of text information SemCom in a radio access network environment. We measure the communication efficiency of the SemCom system based on the DeepSC model according to S-R and S-SE. To maximize the S-SE of all users in the system, we formulate an optimization problem and solve it using the PSO-DA algorithm. The simulation results demonstrate the superiority of the proposed system, showing that it outperforms existing systems when the number of channels is limited and the transmit power is low. In addition, based on the existing neural network model for semantic extraction and transmission of speech and image information, we can easily extend this method to various information sources. However, the resource allocation scheme suitable for the joint transmission of semantic information of different source types needs further research.

References

1. Chahbar, M., Diaz, G., Dandoush, A., Cérin, C., Ghoumid, K.: A comprehensive survey on the E2E 5G network slicing model. IEEE Trans. Netw. Serv. Manag. **18**(1), 49–62 (2021). https://doi.org/10.1109/TNSM.2020.3044626
2. Lan, Q., et al.: What is semantic communication? A view on conveying meaning in the era of machine intelligence. J. Commun. Inf. Netw. **6**(4), 336–371 (2021). https://doi.org/10.23919/JCIN.2021.9663101
3. Shi, G., Xiao, Y., Li, Y., Xie, X.: From semantic communication to semantic-aware networking: model, architecture, and open problems. IEEE Commun. Mag. **59**(8), 44–50 (2021). https://doi.org/10.1109/MCOM.001.2001239
4. Xiao, Y., Shi, G., Li, Y., Saad, W., Poor, H.V.: Toward self-learning edge intelligence in 6G. IEEE Commun. Mag. **58**(12), 34–40 (2020). https://doi.org/10.1109/MCOM.001.2000388
5. Luo, X., Chen, H.H., Guo, Q.: Semantic communications: overview, open issues, and future research directions. IEEE Wirel. Commun. **29**(1), 210–219 (2022). https://doi.org/10.1109/MWC.101.2100269
6. Yan, L., Qin, Z., Zhang, R., Li, Y., Li, G.Y.: Resource allocation for text semantic communications. IEEE Wirel. Commun. Lett. **11**(7), 1394–1398 (2022). https://doi.org/10.1109/LWC.2022.3170849
7. Xie, H., Qin, Z., Li, G.Y., Juang, B.H.: Deep learning enabled semantic communication systems. IEEE Trans. Sig. Process. **69**, 2663–2675 (2021). https://doi.org/10.1109/TSP.2021.3071210
8. Xie, H., Qin, Z.: A lite distributed semantic communication system for Internet of Things. IEEE J. Sel. Areas Commun. **39**(1), 142–153 (2021). https://doi.org/10.1109/JSAC.2020.3036968
9. Jankowski, M., Gündüz, D., Mikolajczyk, K.: Deep joint source-channel coding for wireless image retrieval. In: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5070–5074 (2020). https://doi.org/10.1109/ICASSP40776.2020.9054078
10. Bourtsoulatze, E., Burth Kurka, D., Gündüz, D.: Deep joint source-channel coding for wireless image transmission. IEEE Trans. Cogn. Commun. Netw. **5**(3), 567–579 (2019). https://doi.org/10.1109/TCCN.2019.2919300
11. Kurka, D.B., Gündüz, D.: Deep joint source-channel coding of images with feedback. In: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5235–5239 (2020). https://doi.org/10.1109/ICASSP40776.2020.9054216
12. Weng, Z., Qin, Z., Li, G.Y.: Semantic communications for speech signals. In: ICC 2021 - IEEE International Conference on Communications, pp. 1–6 (2021). https://doi.org/10.1109/ICC42927.2021.9500590
13. Papa, A., Jano, A., Ayvaşık, S., Ayan, O., Gürsu, H.M., Kellerer, W.: User-based quality of service aware multi-cell radio access network slicing. IEEE Trans. Netw. Serv. Manag. **19**(1), 756–768 (2022). https://doi.org/10.1109/TNSM.2021.3122230

14. Korrai, P., Lagunas, E., Sharma, S.K., Chatzinotas, S., Bandi, A., Ottersten, B.: A RAN resource slicing mechanism for multiplexing of eMBB and URLLC services in OFDMA based 5G wireless networks. IEEE Access **8**, 45674–45688 (2020). https://doi.org/10.1109/ACCESS.2020.2977773
15. Xie, H., Qin, Z., Li, G.Y.: Task-oriented multi-user semantic communications for VQA. IEEE Wirel. Commun. Lett. **11**(3), 553–557 (2022). https://doi.org/10.1109/LWC.2021.3136045
16. Farsad, N., Rao, M., Goldsmith, A.: Deep learning for joint source-channel coding of text. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2326–2330 (2018). https://doi.org/10.1109/ICASSP.2018.8461983
17. Gale, D., Shapley, L.S.: College admissions and the stability of marriage. Am. Math. Monthly **69**(1), 9–15 (1962). https://doi.org/10.1080/00029890.1962.11989827
18. Ghosh, A., Ratasuk, R.: Essentials of LTE and LTE-A. The Cambridge Wireless Essentials Series. Cambridge University Press, Cambridge (2011). https://books.google.com/books?id=16DYZklHCBUC