

# Wireless semantic communication based on semantic matching multiple access and intent bias multiplexing

Ren Chao<sup>1</sup>, He Zongrui<sup>1</sup>, Sun Chen<sup>2</sup>, Li Haojin<sup>2</sup>, Zhang Haijun<sup>1</sup> (✉)

1. University of Science and Technology Beijing, Beijing 100083, China

2. Research and Development Center, Sony (China) Limited, Beijing 100027, China

## Abstract

This paper proposes a multi-access and multi-user semantic communication scheme based on semantic matching and intent deviation to address the increasing demand for wireless users and data. The scheme enables flexible management of long frames, allowing each unit of bandwidth to support a higher number of users. By leveraging semantic classification, different users can independently access the network through the transmission of long concatenated sequences without modifying the existing wireless communication architecture. To overcome the potential disadvantage of incomplete semantic database matching leading to semantic intent misunderstanding, the scheme proposes using intent deviation as an advantage. This allows different receivers to interpret the same semantic information differently, enabling multiplexing where one piece of information can serve multiple users with distinct purposes. Simulation results show that at a bit error rate (BER) of 0.1, it is possible to reduce the transmission by approximately 20 semantic basic units.

**Keywords** semantic communication, multiple access, multiplexing, multimodal communication

## 1 Introduction

As the number of wireless users continues to grow and the amount of data demanded by each user increases, wireless communication spectrum is becoming increasingly congested, leading to inadequate network connectivity and speed experience. Wireless multi-access and multi-user sharing technologies can allow for more users or links to be supported per unit bandwidth, making them a key technology for improving last-mile connectivity. In addition,

multimedia data such as text, images, voice, and video, as well as communication, sensing, and computing, are often mixed in current wireless networks and require heterogeneous services. Therefore, fully distinguishing among multiple sources of information and providing precise services while addressing the challenges of data heterogeneity, protocol heterogeneity, and transmission media heterogeneity is a significant challenge<sup>[1-2]</sup>.

Traditional multi-access technologies focus on user capacity, which involves encoding and decoding different users to ensure that their signals do not interfere with each other in the time, frequency, space, or power domains using the same radio resources. Common multi-access technologies include time-division multiple access (TDMA), frequency-

division multiple access (FDMA), space-division multiple access (SDMA), code-division multiple access (CDMA), and non-orthogonal multiple access (NOMA). Multiplexing, on the other hand, refers to a technique for combining multiple signals into a single wireless resource to improve system efficiency. Existing researches have explored these multiple access technologies in semantic communication, which share the common goal of improving multi-user transmission efficiency and network performance while avoiding multiple access conflicts, while ensuring semantic interpretation.

Benefiting from the ability of orthogonality to effectively reduce interference and the significant reduction in data volume provided by semantic communication, joint use of semantic communication and orthogonal multiple access (OMA) techniques, as well as various multiple access optimization techniques for resource management and user collaboration methods, have been applied first. To support more semantic users to access simultaneously, Ref. [3] uses a joint optimization-based OMA technique for uplink communication. Classic TDMA has been emphasized in Ref. [4], and it can also assist in resolving multiple access conflicts based on semantic requirements. In intelligent transportation systems, TDMA techniques can be used to achieve efficient priority management of information flow in the network to meet security requirements. However, conventional bit-based TDMA systems have a problem of high average information freshness and long wait times for users to obtain the next update opportunity. To address this issue, Ref. [5] proposed a knowledge graph-assisted semantic communication (KGSC) technology called KGSC-TDMA to reduce the amount of transmitted information and the time required to receive update information.

Compared with conventional OMA techniques, NOMA technologies can better adapt to communication requirements between heterogeneous devices, support simultaneous multi-user access, and improve network capacity and efficiency. Ref. [6] utilizes NOMA technology to assign users to different resource blocks and avoid interference between users by assigning

multiple users to the same resource block. Ref. [7] combines semantic communication technology with NOMA communication, focusing on the most relevant and important information to further reduce data traffic. This indicates that NOMA has a unique advantage in multi-user semantic communication. Ref. [8] proposed a semantic communication framework for multimedia data that explores the coexistence of traditional bit transmission and semantic information transmission with NOMA and resource allocation in semantic communication. The use of artificial intelligence has accelerated the updating of NOMA based semantic communication systems. Ref. [9] uses NOMA based on deep learning interference management for a semantic communication network that supports communication between heterogeneous devices. Ref. [10] uses different source information modalities to map discrete features to self-learning symbols and completes intelligent multi-user detection (MUD) at the NOMA receiver. In addition, there are also novel NOMA architectures based on semantic communication, such as Ref. [11], where secondary users (F-users) use SemCom while primary users (N-users) use BitCom. F-users can participate in NOMA in each fading state through the most suitable communication method while ensuring their own performance while reducing interference to the main N-users. In Ref. [12], semi-NOMA implements a point that simultaneously sends semantic and bit streams to one semantic interested user (S-user) and one-bit interested user (B-user).

There are also multiple access technologies that lie between OMA and NOMA that have been applied to semantic communication. For example, Ref. [13] proposes two new joint source and channel coding (JSCC) schemes based on deep learning for searching similar images on shared OMA and NOMA multiple access channels. Ref. [14] proposes a novel semantic information framework based on rate-split multiple access (RSMA) technology, which only transmits semantic information related to each user's interests to reduce the burden of data transmission and processing.

The above existing researches have achieved efficient resource utilization and data transmission under

multiple access technologies through different resource allocation and interference management strategies. However, there are still very few multiple access techniques specifically designed for semantic communication networks. Inspired by conventional multi-access and multiplexing technologies, it is necessary to develop more suitable multi-access and multiplexing schemes for semantic communication to improve reliability and efficiency. For example, dynamic resource allocation and scheduling mechanisms based on semantic understanding, as well as the use of machine learning and artificial intelligence to optimize user multi-access in semantic communication systems, can be considered.

New multiple access techniques that target the characteristics of semantic communication will bring higher efficiency and better performance compared to traditional multiple access techniques. Ref. [15] proposes a type-based multiple access (TBMA) technology that combines source-channel coding to utilize the semantics of observed signals, allowing all sensors measuring the same event to share the same codebook (non-orthogonal code words). All sensors performing local estimation of the same event transmit the same code word, reducing communication overhead and achieving unlicensed random access with semantic awareness in fog computing networks. In Ref. [16], a deep learning-based multiple access (DeepMA) technology consists of multiple independent encoder-decoder pairs (EDPs), where the DeepMA encoder encodes input data into mutually orthogonal semantic symbol vectors (SSVs). The DeepMA decoder can detect and recover its target data from the received mixed SSVs (MSSVs), enabling multiple access.

When considering the use of these existing multiple access technologies in semantic communication, it is possible to achieve higher frequency utilization by prioritizing the reuse of wireless resources within the same frequency band. However, due to differences in the length of semantic information, fixed frame configurations may result in some subframes being empty or insufficient, leading to inefficient use of time resources. In contrast, CDMA and NOMA provide fixed allocation of signal lengths but requires dividing signals

into many narrowband signals and synchronization of signal timing, which can result in semantic decoding errors. Therefore, a time-domain multi-access scheme that is consistent with semantic communication features presents a challenging problem.

Semantic matching is a method that analyzes, compares, and matches content based on semantics. In semantic communication with multiple users forming superframes based on semantic matching, different user data can be segmented into different groups based on their own semantics. This significantly reduces the complexity of time-related signaling and synchronization, allowing users to efficiently access their own data while improving transmission efficiency. However, ensuring consistent semantic databases between senders and receivers is necessary to avoid semantic intent misunderstandings.

To address these challenges, this paper has designed a multi-user superframe based on semantic matching for multi-access and semantic intent misunderstanding multiplexing schemes. The contributions include as follows.

1) A semantic matching multi-access technology based on the purpose of semantics for different semantic classes that may belong to different user groups. By sending long concatenated sequences, different users can obtain their own interested fragments through classification to achieve flexible long-term management of frames.

2) A semantic information multiplexing technology based on semantic deviation due to differences in knowledge databases between senders and receivers. Although this is usually a disadvantage, we consider turning it into an advantage by allowing different receivers to have different understandings of the same semantic information, thereby forming a multiplexing capability where one piece of information can be sent and received by multiple parties with different purposes.

3) Designing a double-layer frame structure based on semantic matching multi-access and semantic deviation.

The rests of the paper are structured as the followings. Sect. 2 describes in the overall semantic

transfer system and focuses on the process of orthogonal semantic categorization multiple access. Sect. 3 introduces semantic similarity and evaluates system performance metrics. Sect. 4 presents the idea of intentional biased multiple access, and analyzes the performance that can be improved. Simulation and analysis performance in Sect. 5. Finally Sect. 6 devotes to the conclusions.

## 2 Orthogonal semantic categorization of semantic matching multiple access

Semantic category multiple access mechanism implements orthogonal semantic categorical multiple access communication by comparing the similarity between the two semantic features extracted at the

transmitter and the received semantic information of the receivers, and we consider analyzing it from the perspective of user group matching.

As shown in Fig. 1, the downlink design allows the base station (BS) to distribute semantic signals from multiple source data to multiple user groups simultaneously.

In BS:

1) First, the user group as the object of transmission, according to the user's needs and thus the formation of user groups with the same semantic categorization needs; For example, for user group  $U_{G1}$ , the data  $x_1$  is semantic encoded as  $v_1$ .

2) Multi-user group with  $i \in K$  is semantic encoded as  $v_i$ , successively spliced as  $V$ .

3) Modulated and send  $X$ .

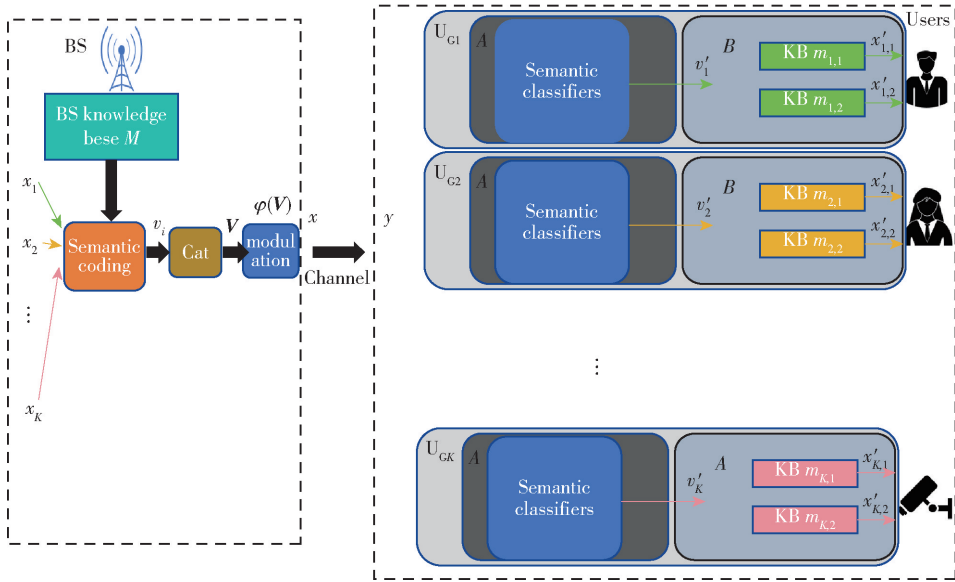


Fig. 1 Semantic matching multiple access and intent bias multiplexing system maps

In the user side:

1) Normal reception demodulation  $X$ ; Obtain semantic information  $V'$ .

2) First, use artificial intelligence-based semantic model to extract semantic information, and then form semantic segmentation. The semantic information  $V$  from the BS is directly spliced. Thus, it only needs to be directly based on semantic segmentation to form the semantic information required by different user groups.

Specifically, here we specifically consider a downlink semantic matching multiple access system for

$K$  groups of users ( $K \geq 2$ ) with a wireless broadcast channel and the same data sent by the same user group  $U_{Gi}$ .

### 2.1 Transmitting strategy

BS tries to map the source data  $x_1, x_2, \dots, x_K$  into a sequence of semantic vectors  $\varepsilon_1(x_1), \varepsilon_2(x_2), \dots, \varepsilon_K(x_K)$  using a series of model-based semantic coders  $\varepsilon_1(\cdot), \varepsilon_2(\cdot), \dots, \varepsilon_K(\cdot)$ , i. e.

$$v_i = \varepsilon_i(x_i); \forall i \in K \quad (1)$$

Subsequently, the semantic vectors  $v_i, i \in K$  obtained

from these semantic encodings are directly spliced using a semantic splicer, and thus the spliced semantic vectors can be represented as  $\mathbf{V}$

$$\mathbf{V} = \text{cat}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K) = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K\} \quad (2)$$

where  $\text{cat}(\cdot)$  is the semantic information splicing function. The function splices the signals obtained from each semantic code. It is worth noting that here the frame length is flexible and variable, and no synchronization is required. The information is spliced directly, without considering the synchronization of each user's frame. The spliced information is then modulated and the modulated signal  $\mathbf{X}$  is represented as  $\mathbf{X} = \varphi(\mathbf{V})$  (3)

where  $\varphi(\cdot)$  is the modulation function. After the modulation, the transmitter transmits the modulated signal through the wireless channel to reach the receivers of user groups.

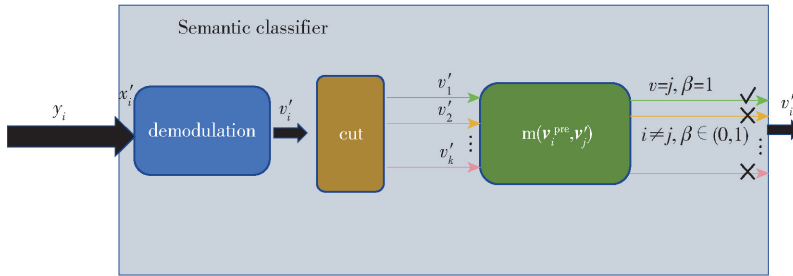
## 2.2 Receiving strategy

Each user group  $\mathbf{U}_{Gi}$  receives a signal differently due to the fact that the transmitted signal arrives over different channels. At the user's receiver, the signal

received for one frame length from the  $i$ th user group  $\mathbf{U}_{Gi}$  can be modeled as  $\mathbf{y}_i = p h_i \mathbf{X} + \mathbf{n}_i$ ,  $i \in K$ , where  $p$  is the power of the transmitted signal,  $h_i$  is the Rayleigh fading coefficient of the channel of the user group, and  $\mathbf{n}_i$  is the additive white Gaussian noise (AWGN).

Unlike NOMA, in this paper, the signals are not temporally collapsed, but rather a TDMA-like transmission method that performs a kind of splicing of the transmitted semantic signals. For semantic matching multiple access, its core detection component at the user side is the semantic classifier, which can detect signals  $\mathbf{X}$  sent to different user groups, and match the decoded signals for semantic similarity. This enables detecting signals from different user groups, distinguishing them from each other and realizing multiple access method.

Fig. 2 shows that at each user group, the core component of the semantic classifier contains three main functions: Decoding, semantic segmentation and semantic matcher.



**Fig. 2** Semantic classifiers at user groups

Here we apply channel estimation using channel gain and forced-zero detector to the transmit signal  $\mathbf{X}'_i$ , and then demodulated it to obtain spliced semantic information  $\mathbf{V}'_i = \varphi^{\text{dem}}(\mathbf{X}'_i)$ , where  $\varphi^{\text{dem}}(\cdot)$  is the demodulation function. Multiple semantic vectors  $\mathbf{v}'_i$  are obtained by segmenting the semantic information  $\mathbf{V}'_i$  into

$$\mathbf{V}'_{i^{\text{cut}}} = \text{cut}(\mathbf{V}'_i) = \{\mathbf{v}'_1, \mathbf{v}'_2, \dots, \mathbf{v}'_K\}; \quad i \in K \quad (4)$$

where  $\text{cut}(\cdot)$  is the semantic segmentation function. This function leads to semantic vectors, and each user group matches based on their own semantic vectors of interest.

In the  $A$  process, that is, when considering each

user group as a unit, semantic matching multiple access is applied, and different data information exhibits common semantic features that exist in a large number of similar features. Therefore, there is a high success rate of semantic matching in the process of  $A$

$$m(\mathbf{v}_i^{\text{pre}}, \mathbf{v}'_j) = \beta; \quad i, j \in K \quad (5)$$

where  $m(\mathbf{v}_i^{\text{pre}}, \mathbf{v}'_j)$  is the success rate of matching,  $\mathbf{v}_i^{\text{pre}}$  is the semantic information of interest to the user group  $\mathbf{U}_{Gi}$ , and  $\beta \in (0, 1]$ .

## 3 Semantic similarity and metrics

The evaluation of transmission performance for

semantic communication differs from traditional bit-based communication, such as BER or block error rate (BLR). Commonly used performance metrics for image transmission, such as structural similarity (SSIM) and peak signal-to-noise ratio (PSNR), assess the similarity of two images based on brightness, contrast, structure, and background noise. For sentence transfer performance, bilingual evaluation (BLEU) of alternatives score and sentence similarity are used<sup>[7]</sup>, which measure word-level similarity and take into account contextual information and sentence meaning.

Therefore, semantic similarity, can be mainly analyzed as influenced by two factors.

1) The extraction and matching of semantic information by the knowledge base of the receiver and transmitter. If the knowledge base of the receiver and transmitter is completely consistent, it can be considered that there is no error in the process of transforming the information to the semantics and the semantics to the information.

2) The effect of signal-to-noise ratio in the wireless transmission process.

This paper focuses on semantic signals and does not differentiate the modality of the data source.

The semantic similarity is calculated using the similarity between two values and the BER of the received signal, and depends on the library  $\mathcal{M} = \{M_1, M_2, \dots, M_K\}$  at the transmitter, where the library  $m_i$  is at the receiver, and the signal-to-noise ratio is  $\gamma_i, i \in K$ . In order to reflect the facts that both affect the values of semantic similarity and avoids some large values to take a decisive influence, this paper uses conditions that satisfy normalization. Thus, the first part of the knowledge base difference can be represented by the matching success rate, while the signal-to-noise ratio can be transformed into a BER representation.

Thus, the signal to noise ratio SNR can be expressed as

$$\gamma_i = \frac{p|\mathbf{h}_i|^2}{\sigma_i^2}; \quad i \in K \quad (6)$$

Therefore, it is easy to derive BER  $B_i = f(\gamma_i)$ , and the BER is a function of the signal-to-noise ratio.

The semantic similarity is

$$s(\mathbf{v}_i^{\text{pre}}, \mathbf{v}_j') = m(\mathbf{v}_i^{\text{pre}}, \mathbf{v}_j') (1 - B_j); \quad i, j \in K \quad (7)$$

where  $s(\mathbf{v}_i^{\text{pre}}, \mathbf{v}_j')$  denotes the similarity between the two, and  $B_j$  denotes the BER of the received signal. In process A, this paper assumes that if the transmitter and the user group have identical semantic knowledge bases, there is no error due to the information-to-semantics transformation. Only the effect of noise in the channel is considered.

Specifically,  $m(\mathbf{v}_i^{\text{pre}}, \mathbf{v}_j') = \beta = 1$  when  $i = j$  and  $\beta \in (0, 1)$  when  $i \neq j$ . Thus, the semantic similarity of the A process at successful matching is

$$s(\mathbf{v}_i^{\text{pre}}, \mathbf{v}_i') = (1 - B_i) \quad (8)$$

To measure the performance of this system, this paper will evaluate the effectiveness of the scheme in terms of spectral efficiency (SE). This metric distinguishes it from the matching efficiency of conventional bit transmission.

In conventional communication, SE is typically measured in  $\text{bit}/(\text{s} \cdot \text{Hz})$ . While this is valid for measuring the transmission rate of a sequence of bit, it does not apply to measuring the transmission rate of semantic information. We need to reanalyze SE in the semantic domain. To do this, we define the basic unit of semantic information extraction as a predefined intent in the semantic vector  $\mathbf{v}$  for information extraction with a predefined intent  $e_g$ , where  $1 \leq g \leq G$  is finite and the intent is not subdividable. Each knowledge base consists of a collection of intent-based semantic information units  $e_g$ . Both communicating parties have knowledge bases that comprise finitely many intents  $m = \{e_1, \dots, e_g, \dots, e_G\}$ . A semantic expression corresponds to a collection of intents. For example, the semantic vector  $\mathbf{v}_i = \{e_1, e_2, \dots, e_g\}_i$  for the user group  $U_{G_i}$  contains the  $g$ th intent, denoted as  $e_g$  and  $g \leq G$ . A semantic expression must minimally contain one intention, creating a single-intention semantic vector  $\mathbf{v}_i = \{e_g\}_i$ .

This paper assumes that semantic information can be measured in terms of a semantic unit  $e$ , which represents the basic unit of semantic information. Based on this, this paper defines two key semantic-based performance metrics: The semantic transmission rate  $R_s$ , unit is  $e/\text{s}$ , which measures the semantic

information effectively transmitted per second, and the semantic SE (S-SE), unit is  $e/(s \cdot \text{Hz})$ , which measures the rate of successful semantic information transmission per unit of bandwidth.

Given a set of spliced semantic vectors  $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_K\}$ , where each semantic message has a length of  $l_{v_i}$ , and the average length of each user group's intention  $e_g$  is  $l_{e_i}$ , this paper assumes there are  $D$  semantic vectors. Let the semantic vector  $\mathbf{v}_i$  has semantic information of  $I_i$ , and the probability of its occurrence is  $p(\mathbf{v}_i)$ . The expected amount of semantic information is  $I = \sum_i^D I_i p(\mathbf{v}_i)$ , and its expected length for each vector  $\mathbf{v}$  is  $L = \sum_i^D l_{v_i} p(\mathbf{v}_i)$ .

On average, the  $i$ th user group has semantic symbols carrying an amount  $l_{e_i} L$  of semantic information  $I$ , with an average amount per semantic symbol of  $I/(l_{e_i} L)$ . The total semantic information transmitted over a channel with bandwidth is equal to the symbol rate. Furthermore, since the symbol rate is equal to the channel bandwidth used for passband transmission, the total semantic information transmitted over a channel with bandwidth  $W$  is  $WI/(l_{e_i} L)$ . Therefore, the  $R_s$  of the  $i$ th user group can be expressed as

$$R_i = \frac{WI}{l_{e_i} L} s(\mathbf{v}_i^{\text{pre}}, \mathbf{v}_i') ; \quad i \in K \quad (9)$$

where  $\mathcal{R}_i$  the semantic information transfer rate of the  $i$ th user group, and  $s(\mathbf{v}_i^{\text{pre}}, \mathbf{v}_i')$  is the similarity degree of the semantic vector. Thus, the S-SE of the  $i$ th user group can be expressed as

$$\Psi_i = \frac{R_i}{W} = \frac{I}{l_{e_i} L} s(\mathbf{v}_i^{\text{pre}}, \mathbf{v}_i') ; \quad i \in K \quad (10)$$

Then, use the S-SE of all user groups  $U_{Gi}, i \in K$  of the whole system as a measure, i. e.

$$\Psi = \sum_i^K \frac{I}{l_{e_i} L} s(\mathbf{v}_i^{\text{pre}}, \mathbf{v}_i') \quad (11)$$

Since conventional systems are evaluated in the bit domain, this paper develops a conversion method that converts the typical SE into an equivalent S-SE by accounting for the effect of source coding. This allows for fair comparisons between different communication

systems. The simulation results are presented and analyzed based on this conversion method.

In conventional network, each minimal unit of data information is mapped to bit by the source encoder. From a semantic perspective, the original few bit can be interpreted as a semantic symbol. The equivalent  $R_s$  can be expressed as

$$R_i^b = C_i \frac{I}{\rho L} s(\mathbf{v}_i^{\text{pre}}, \mathbf{v}_i') \quad (12)$$

where  $C_i$  is the transmission rate on the channel of the  $i$ th user group in bit/s, and the conversion factor  $\rho$  is the intent-to-bit expression. It indicates the average number of bit per intent. The conventional bit transmission SE is given as  $C_i/W$ , and the equivalent S-SE can be derived as

$$\Psi_i^b = \frac{C_i I}{W \rho L} s(\mathbf{v}_i^{\text{pre}}, \mathbf{v}_i') \quad (13)$$

Thus, the bit transmission process is exported to the S-SE of the conventional system. This enables a comparison between the bit transmission process and the S-SE of the conventional system.

#### 4 Intent bias multiplexing

Above, this paper analyzed semantic matching multiple access in terms of user groups but found that it did not improve transmission efficiency. Therefore, this paper further analyzes and proposes intent bias multiplexing.

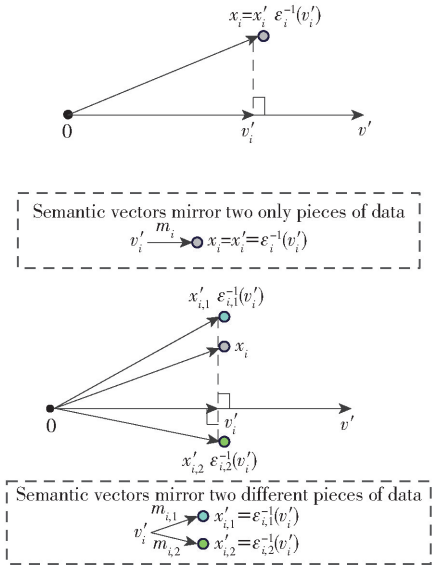
First, let's consider the scenario where the BS sends data  $\mathbf{x}_i$  corresponding to the semantic vector  $\mathbf{v}_i$  for the user group  $U_{Gi}$ . Assuming there are two users, 1 and 2, in the receiving user group without semantic bias, user 1 recovers semantic information as  $\mathbf{x}'_{i,1} = \boldsymbol{\varepsilon}_{i,1}^{-1}(\mathbf{v}_i') = \mathbf{x}_i$ , and user 2 recovers semantic information as  $\mathbf{x}'_{i,2} = \boldsymbol{\varepsilon}_{i,2}^{-1}(\mathbf{v}_i') = \mathbf{x}_i$ .

When there is a difference in the semantic knowledge base between the sender and receiver, resulting in biased understanding, this paper assumes both are biased. For example, the BS sends data  $\mathbf{x}_i$  to  $U_{Gi}$ , and user 1 recovers  $\mathbf{x}'_{i,1} = \boldsymbol{\varepsilon}_{i,1}^{-1}(\mathbf{v}_i')$ , while user 2 recovers  $\mathbf{x}'_{i,2} = \boldsymbol{\varepsilon}_{i,2}^{-1}(\mathbf{v}_i')$ . Neither  $\mathbf{x}'_{i,1}$  nor  $\mathbf{x}'_{i,2}$  is equivalent to  $\mathbf{x}_i$ . This is the image representation of the deviation, where normal semantic communication fails.



However, if this error is mapped, convert the semantic vector corresponding to  $\mathbf{x}_i$  to a modulated symbol, and consider the recovery of  $\mathbf{x}'_{i,1}$  from the BS to the 1 channel and  $\mathbf{x}'_{i,2}$  from the BS to the 2 channel, there is a one-to-one mapping without considering channel errors, only accounting for the knowledge base bias. Exploiting this bias, the system records and analyzes this bias in a database for error correction and system improvement.

Fig. 3 illustrates the recovered data bias due to knowledge base differences.



**Fig. 3** Recovered data bias due to knowledge base differences

#### 4.1 Semantic intent bias mapping

For a dataset  $\mathcal{D} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$  of  $N$  data, and each  $\mathbf{x}_i$  is to be transformed into a semantic vector  $\mathbf{v}_i$ , denoted by  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_g\}_i$ . Then, for one of the data  $\mathbf{x}_i \rightarrow \mathbf{v}_i$ , the corresponding semantic vector is  $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N\}$ . Suppose  $\mathbf{M}_i = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ ,  $\mathbf{m}_{i,1} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ ,  $\mathbf{m}_{i,2} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4\}$ . Then, when the semantic vector  $\mathbf{v}_i = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  is sent and  $\mathbf{v}'_i = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  is received, regardless of transmission errors, for user 1,  $\mathbf{v}'_{i,1} = \mathbf{v}'_i$ , and for user 2, who can only translate semantic vectors based on the semantic knowledge base. The received semantic vector is  $\mathbf{v}'_i = \{1, 1, 1\}$ . Due to the difference in the knowledge base  $\mathbf{m}_{i,2} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4\}$ , it is not possible to determine the exact correspondence, resulting in similar semantic

information  $\mathbf{v}'_{i,2} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4\}$ , i. e., the actual understanding is  $\mathbf{v}'_{i,2} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4\}$ . Thus, the deviation semantics  $\mathbf{v}'_{i,2}$  is transformed into the corresponding data  $\mathbf{x}'_{i,2}$  through the knowledge base

$$\mathbf{x}'_{i,2} = F(\mathbf{v}'_{i,2}, \mathbf{m}_{i,2}) \quad (14)$$

where the  $\mathcal{F}$  function is based on the user's knowledge base  $\mathbf{m}_{i,2}$  and corresponds to the recovered data  $\mathbf{x}'_{i,2}$ , i. e., recovered data, through the deviation semantics  $\mathbf{v}'_i$ . Thus, the ability to deviate into  $\mathbf{x}'_{i,1}, \mathbf{x}'_{i,2}$  for user 1 and user 2 occurs precisely. For  $\mathbf{v}'_i$ , due to the deviation of the knowledge base, the reflections, i. e., the semantic  $\rightarrow$  informational posterior, within the same user group  $\mathbf{U}_{G_i}$  differs.

In a user group  $\mathbf{U}_{G_i}$ , statistics and record the data of each deviation semantics. The algorithm is as follows. For each user, Algorithm 1 sends each data of the dataset in turn and statistically records the deviation data of the semantic vectors received by each user.

**Algorithm 1** Statistical record deviation semantic data

Send: Iterate through the dataset  $\mathcal{D} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$  for  $k = 1 : N$

1: Transforming  $\mathbf{x}_k$  into semantic vectors  $\mathbf{v}'_i$  by KB  $\mathbf{M}_i =$

$$\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_g\} \quad \mathbf{x}_k \xrightarrow{\mathbf{M}_i} \mathbf{v}'_i$$

receive: Iterate over each user

for  $j = 1 : n$

2: Receive semantics  $\mathbf{v}'_i$ , based on the deviatoriable KB  $\mathbf{m}_{i,j}$ , to get deviant semantics  $\mathbf{v}'_{i,j}$

$$\text{Save } \mathbf{v}'_i \xrightarrow{\mathbf{m}_{i,j}} \mathbf{v}'_{i,j}$$

3: Then the bias semantics  $\mathbf{v}'_i$  is reflexive based on the KB  $\mathbf{m}_{i,j}$

$$\text{Save } \mathbf{x}'_{i,j} = \mathcal{F}(\mathbf{v}'_{i,j}, \mathbf{m}_{i,j})$$

4: end  $j$

5: end  $k$

#### 4.2 Reliability analysis

Assuming a biased intent probability of  $p_e$ , for  $n$  users, this system  $np_e$  users with biased intents in their knowledge bases. In the case where errors are not considered, the performance of transmitting without error correction is  $c_s = (n - np_e)/n$  and  $q$  users. In practice, there are  $q$  users who need deviation information and  $n - q$  users who need to recover correct information. The performance for system recovery



information for these  $n$  users is

$$c_s = \frac{n - q + np_e}{n}; \quad np_e \leq q \quad (15)$$

Then, consider noise in the semantic judgment. We assume the updated performance is

$$c'_s = c_s (1 - B) \quad (16)$$

where  $B = (1/n) \sum_{k=1}^n B_k$  is the average BER for all users, and  $c'_s$  represents the transmission reliability performance of the system with entrainment noise.

### 4.3 Efficiency analysis

By achieving multiple access through semantic deviation, transmission efficiency is significantly improved. This paper measures this based on the number of intentions sent. Specifically, due to the sender sending semantic vectors  $v_i = \{e_1, e_2, \dots, e_g\}$ , there are  $np_e$  users with deviation intentions due to knowledge base differences. Thus, correction can be made for  $np_e$  users, resulting in a multi-access scenario. Consequently, the number of intentions that need to be sent reduces from  $t'_e = g$  to  $t_e = g(np_e + 1)$  and directly reduces the number of intentions sent  $gnp_e$ . As the number of users increases, the more users that can be corrected by semantic bias, the fewer intentions need to be sent. Therefore, the number of intents sent can be reduced correctly as

$$t'_e = gnp_e S_{\text{sim}} \quad (17)$$

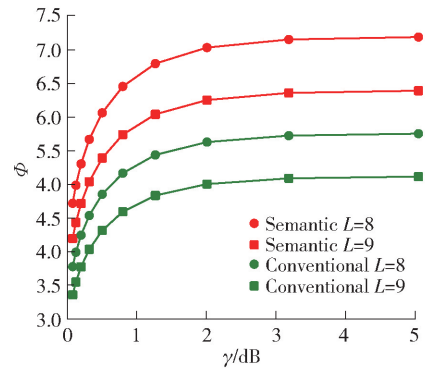
where  $S_{\text{sim}}$  represents the average semantic similarity.

## 5 Simulations and analysis

In Sect. 5, numerical results are presented to demonstrate the feasibility of our proposed framework. Assuming  $K = 8$  user groups, we adopt binary phase shift keying (BPSK) for semantic information transmission in downlink transmission. The parameter settings include bandwidth  $W = 1$  MHz,  $N_0 = 10^{-6}$ , fading coefficients  $h_i \sim \mathcal{CN}(0, 1)$  and  $n_i \sim \mathcal{CN}(0, \sigma_i^2)$ , and the BER is expressed as  $B_i = Q(0, \sigma_i^2)$ . Then the BER is expressed as  $B_i = Q(\sqrt{2c\gamma_i})$ , where  $c$  is a constant and  $Q(x) = \int_x^\infty (1/\sqrt{2\pi}) e^{-\frac{t^2}{2}} dt$  is presented in Sect. 3. In

orthogonal semantic categorical multiple access, the data is processed through semantic encoding, splicing, modulation, semantic classifier segmentation, demodulation, and matching with the most suitable semantic vector in each user group  $U_{Gi}$ , achieving orthogonal semantic categorical multiple access. The user group will receive the signal through the semantic classifier segmentation demodulation segmentation, and then match themselves with the most adapted semantic vector  $v_i^{\text{pre}}$ .

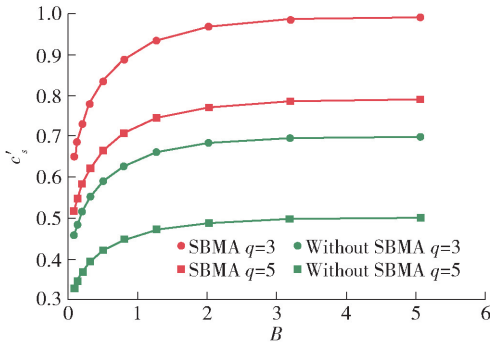
As shown in Fig. 4, when the average semantic vector length is  $L = [8, 9]$ , the amount of semantic information is  $I = 9$ , and the length of the semantic minimum unit is 1, evaluating conventional bit transmission over the semantic domain using transformation methods, we observe that the S-SE remains constant as the signal-to-noise ratio increases from 0 dB to 5 dB. Regardless of whether it is traditional bit transmission or semantic transmission, the effective S-SE improves. Moreover, at the same signal-to-noise ratio, and the smallest unit of semantic vector is constant, the smaller semantic transmission  $L$ , then it represents the better compression of the original data to the semantic information under the same amount of information  $I$ , so the larger the S-SE  $\Phi$  of semantic transmission. Thus, semantic transmission performs better than traditional bit transmission under the same conditions.



**Fig. 4** S-SE  $\Phi$  of semantic vs conventional bit transmission

Semantic matching multiple access enhances S-SE. However, due to the strict requirements of semantic knowledge base alignment between sender and receiver, deviations significantly reduce transmission

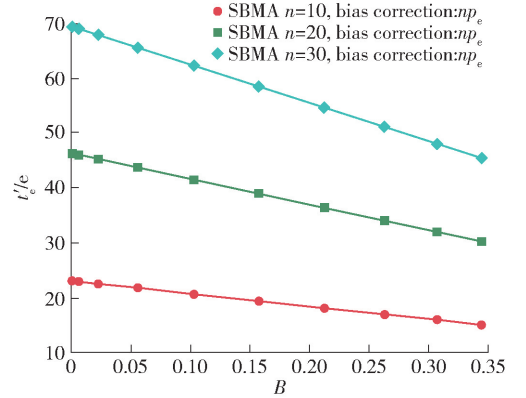
reliability. This transmission does not significantly improve transmission efficiency, but for the third section, its deviation can provide multiple access, improve reliability, and to some extent, improve transmission efficiency. Assuming that each user group contains  $n = 10$  users, the probability of each user's deviation intent occurring is  $p_e \sim U(0.2, 0.4)$ . As shown in Fig. 5, it is actually required that  $q = [3, 5]$  users need deviation information. At the same  $q$  value, the transmission reliability of semantic deviation multiple access increases continuously as the signal-to-noise ratio  $\gamma$  increases, thus reducing interference from noise and fading in the case of entrainment noise. Furthermore, at the same signal-to-noise ratio, the use of semantic deviation multiple access is capable of statistically recording deviations to correct the information sent and thereby improving the reliability of the system. This reduces the need for users to acquire data from the BS and simultaneously enhances overall network utilization.



**Fig. 5** Semantic transmission reliability with and without semantic bias multiple access (SBMA) in the presence of entrainment noise  $c'_s$

Semantic intent bias multiplexing can improve transmission effectiveness by statistically recording deviations of multiple users that are sent simultaneously in a multiple-access mode. As shown in Fig. 6, under a constant number of users and without BER during transmission, the receiver's transmittable performance is entirely determined by the knowledge base bias. As BER  $B$  increases, the transmitter needs to transmit more intents to maintain the same level of performance, resulting in a reduction of transmitted intents  $t'_e$ . On the other hand, with an increasing number of users a

growing number of users  $np_e$  with biasable multiplexing capabilities, fewer intents can be transmitted to satisfy the users' needs.



**Fig. 6** Number of transmission intentions reduced by SBMA as the number of error-correctable users increases  $t'_e$

## 6 Conclusions

This paper proposes a multi-access, multi-user semantic communication scheme based on semantic matching and intent deviation to address the challenges of increasing wireless users and data demands. The scheme utilizes semantic classification to flexibly manage long frames, allowing each bandwidth unit to support more users. Additionally, this scheme does not require modifications to existing wireless communication architectures, enabling independent access to the network for different users through the transmission of long concatenated sequences. To overcome the potential disadvantage of incomplete matching in the semantic knowledge base, which may lead to misunderstandings of semantic intent, the scheme introduces the concept of intent deviation as an advantage. This allows different receivers to interpret the same semantic information differently, enabling multiplexing where one piece of information can serve multiple users with different objectives.

## Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (62201034).

## References

- [1] MERIEME E A, MOHAMED A, ALI C, et al. A survey on the challenges of data integration. *Proceedings of the 9th International Conference on Wireless Networks and Mobile Communications (WINCOM'22)*, 2022, Oct 26 – 29, Rabat, Morocco. Piscataway, NJ, USA: IEEE, 2022; 1 – 6.
- [2] ABDELMONIEM A M, HO C Y, PAPAGEORGIOU P, et al. A comprehensive empirical study of heterogeneity in federated learning. *IEEE Internet of Things Journal*, 2023, 10(16): 14071 – 14083.
- [3] CHEN B C, WANG X M, LI D P, et al. Uplink NOMA semantic communications; Semantic reconstruction for SIC. *Proceedings of the 2023 IEEE/CIC International Conference on Communications in China (ICCC'23)*, 2023, Aug 10 – 12, Dalian, China. Piscataway, NJ, USA: IEEE, 2023; 1 – 6.
- [4] UYSAL E, KAYA O, EPHREMIDES A, et al. Semantic communications in networked systems; A data significance perspective. *IEEE Network*, 2022, 36(4): 233 – 240.
- [5] CHEN J, YANG S, CHAN T T, et al. Enhancing information freshness via knowledge graph-aided semantic communication. *Proceedings of the IEEE 11th International Conference on Information, Communication and Networks (ICIN'23)*, 2023, Aug 17 – 20, Xi'an, China. Piscataway, NJ, USA: IEEE, 2023; 155 – 160.
- [6] ZHANG M, ZHONG R K, MU X D, et al. Resource management for heterogeneous semantic and bit communication systems. *Proceedings of the 2023 IEEE International Conference on Communications Workshops (ICC Workshops'23)*, 2023, May 28 – Jun 1, Rome, Italy. Piscataway, NJ, USA: IEEE, 2023; 1629 – 1634.
- [7] LI W Z, LIANG H T, DONG C, et al. Non-orthogonal multiple access enhanced multi-user semantic communication. *IEEE Transactions on Cognitive Communications and Networking*, 2023, 9(6): 1438 – 1453.
- [8] DUAN Y P, DU Q Y, FANG X, et al. Multimedia semantic communications; Representation, encoding and transmission. *IEEE Network*, 2023, 37(1): 44 – 50.
- [9] MU X D, LIU Y W, GUO L, et al. Heterogeneous semantic and bit communications; A semi-NOMA schemes. *IEEE Journal on Selected Areas in Communications*, 2023, 41(1): 155 – 169.
- [10] CHEN J R, WANG J J, JIANG C X, et al. Age of incorrect information in semantic communications for NOMA aided XR applications. *IEEE Journal of Selected Topics in Signal Processing*, 2023, 17(5): 1093 – 1105.
- [11] MU X D, LIU Y W, LIU Y. Exploiting semantic communication for non-orthogonal multiple access. *IEEE Journal on Selected Areas in Communications*, 2023, 41(8): 2563 – 2576.
- [12] MU X D, LIU Y W. Semi-NOMA enabled coexisting semantic and bit communications. *Proceedings of the 2022 International Symposium on Wireless Communication Systems (ISWCS'22)*, 2022, Oct 19 – 22, Hangzhou, China. Piscataway, NJ, USA: IEEE, 2022; 1 – 6.
- [13] LO W F, MITAL N, WU H T, et al. Collaborative semantic communication for edge inference. *IEEE Wireless Communications Letters*, 2023, 12(7): 1125 – 1129.
- [14] CHENG Y Y, NIYATO D, DU H Y, et al. Interest-based semantic information transmission with RSMA in smart cities. *Proceedings of the 2023 IEEE International Conference on Communications (ICC'23)*, 2023, May 28 – Jun 1, Rome, Italy. Piscataway, NJ, USA: IEEE, 2023; 82 – 87.
- [15] DOMMEL J, UTKOVSKI Z, SIMEONE O, et al. Joint source-channel coding for semantics-aware grant-free radio access in IoT fog networks. *IEEE Signal Processing Letters*, 2021, 28: 729 – 732.
- [16] ZHANG W Y, BAI K Y, ZEADALLY S, et al. DeepMA; End-to-end deep multiple access for wireless image transmission in semantic communication. *IEEE Transactions on Cognitive Communications and Networking*, Early Access Article. DOI: 10.1109/TCCN.2023.3326302.

(Editor: Ai Lisha)