Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Model division multiple access for semantic communications*

Ping ZHANG[†§1,2,3], Xiaodong XU[†§1,2,3], Chen DONG[†‡1], Kai NIU[†‡1,2], Haotai LIANG[1], Zijian LIANG[1], Xiaoqi QIN[1], Mengying SUN[1], Hao CHEN[2], Nan MA[1,2], Wenjun XU[1], Guangyu WANG[1], Xiaofeng TAO[2,4]

[1]*State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China*

[2]*Department of Broadband Communication, Peng Cheng Laboratory, Shenzhen 518066, China*

[3]*ZGC Institute of Ubiquitous-X Innovation and Applications, Beijing 100876, China*

[4]*School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China*

[†]E-mail: pzhang@bupt.edu.cn; xuxiaodong@bupt.edu.cn; dongchen@bupt.edu.cn; niukai@bupt.edu.cn

Received Mar. 21, 2023; Revision accepted May 15, 2023; Crosschecked May 19, 2023

Abstract: In a multi-user system, system resources should be allocated to different users. In traditional communication systems, system resources generally include time, frequency, space, and power, so multiple access technologies such as 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) are widely used. In semantic communication, which is considered a new paradigm of the next-generation communication system, we extract high-dimensional features from signal sources in a model-based artificial intelligence approach from a semantic perspective and construct a model information space for signal sources and channel features. From the high-dimensional semantic space, we excavate the shared and personalized information of semantic information and propose a novel multiple access technology, named model division multiple access (MDMA), which is based on the resource of the semantic domain. From the perspective of information theory, we prove that MDMA can attain more performance gains than traditional multiple access technologies. Simulation results show that MDMA saves more bandwidth resources than traditional multiple access technologies, and that MDMA has at least a 5-dB advantage over NOMA in the additive white Gaussian noise (AWGN) channel under the low signal-to-noise (SNR) condition.

Key words: Model division multiple access (MDMA); Semantic communication; Multiple access
https://doi.org/10.1631/FITEE.2300196 **CLC number:** TN92

1 Introduction

Semantic communication is considered a new paradigm with great potential for the next-generation communication system (Zhang P et al., 2022a, 2022b, 2022c). The work on semantic communication is still carried out under the framework proposed by Shannon (1948). Artificial intelligence (AI) technology is used to achieve the following: establish the knowledge base at both the receiver

§ These two authors contributed equally to this work
‡ Corresponding authors
* Project supported by the National Key R&D Program of China (No. 2022YFB2902102)
ORCID: Ping ZHANG, https://orcid.org/0000-0002-1485-5849; Xiaodong XU, https://orcid.org/0000-0003-4245-5989; Chen DONG, https://orcid.org/0000-0002-3443-1453; Kai NIU, https://orcid.org/0000-0002-8076-1867

and the transmitter; provide modular and highly abstract methods to extract semantic information; make full use of prior knowledge and channel environment information, making semantic communication more efficient.

Since the first generation of the text-based semantic communication system, namely the deep learning based semantic communication system (DeepSC) (Xie et al., 2021), there have been a large number of semantic communication systems for various sources (Table 1), such as text (Xie et al., 2021), speech signal (Weng and Qin, 2021), image (Dai JC et al., 2022; Dong et al., 2023), video (Jiang et al., 2023; Wang SX et al., 2023), and three-dimensional (3D) point cloud (Zhu et al., 2022). DeepSC (Xie et al., 2021) aims to maximize system capacity and minimize semantic errors by recovering the meaning of sentences rather than bit or symbol errors in traditional communication. Compared with the traditional communication system, it has obvious semantic transmission advantages under the low signal-to-noise ratio (SNR) condition. DeepSC-S (Weng and Qin, 2021) uses a compression and stimulation network to provide higher weights to recognize important speech semantic information through the attention mechanism when training the neural network. Nonlinear transform source-channel coding (NTSCC) was applied in Dai JC et al. (2022) and Wang SX et al. (2023) to map image and video sources into a nonlinear potential space to achieve more efficient semantic extraction and transmission. Dong et al. (2023) believed that the layer-based semantic communication system for image (LSCI) model is the carrier of intelligence, and that semantic transmission is essentially the flow of the AI model. Therefore, the semantic slice model (SeSM) was designed to realize the flow of semantic intelligence. In the face of 3D point cloud sources, AITransfer (Zhu et al., 2022) implements end-to-end source compression from the semantic level, providing ideas for the semantic transmission of 3D point cloud sources.

The construction and verification of the semantic communication system in the physical layer have gradually broadened the level of Shannon's theoretical communication, and the AI model has become an important carrier for building semantic information. The semantic communication systems mentioned above are point-to-point ones, and the multi-user communication system has always been the research focus (Luo et al., 2022; Zhang Y et al., 2022; Li et al., 2023). In the 40 years of development from first-generation (1G) to fifth-generation (5G) communication systems (Mao et al., 2022), each generation has had representative and breakthrough multiple access (MA) technology. The development of MA technology has been directed mostly toward the allocation of orthogonal wireless resources to different physical resources to avoid multi-user interference, for example, the time domain of time division MA (TDMA) in the first generation, the frequency domain of frequency division MA (FDMA) in the second generation, the code domain of code division MA (CDMA) in the third generation, and scheduling of both time and frequency domains in the fourth generation of orthogonal frequency division MA (OFDMA). In addition, the development of the multiple-input multiple-output (MIMO) technology has created a space domain, which has spawned the space division MA (SDMA) technology. In the 5G era, to further improve the spectrum efficiency (SE), MA technology is developing towards non-orthogonal MA (NOMA) by allocating resources in the power domain (Saito et al., 2013; Dai LL et al., 2015; Ding et al., 2016). Each user's information is allocated a different power and then superimposed for transmission. At the receiver, interference cancellation technology is used to separate the signals of different users.

Table 1 Semantic communication system for different sources

Reference(s)	Source	Model
Xie et al., 2021	Text	DeepSC
Weng and Qin, 2021	Speech signal	DeepSC-S
Dai JC et al., 2022; Dong et al., 2023	Image	NTSCC, LSCI
Jiang et al., 2023; Wang SX et al., 2023	Video	DVST (NTSCC-based), SVC-HARQ
Zhu et al., 2022	3D point cloud	AITransfer

NTSCC: nonlinear transform source-channel coding; LSCI: layer-based semantic communication system for image; DVST: deep video semantic transmission; SVC-HARQ: semantic video conferencing - hybrid automatic repeat-request framework for varying channels

In particular, there is a sixth-generation wireless communication (6G) candidate MA scheme, namely rate splitting multiple access (RSMA) (Mao et al., 2022). The main idea of RSMA is to divide the user message into common and private parts, and thus to enable the capability of partially decoding the interference. When the interference is weak or strong, RSMA automatically reduces to SDMA or NOMA, respectively, by adjusting the power and content of the common and private streams.

The new paradigm of the semantic communication system also creates a new allocation resource, namely, the model information space. The generation of model information space resources informs the establishment of a new MA technology, namely model division multiple access (MDMA). Traditional MA technologies consider physical resources, such as time, frequency, power, and space, to realize multi-user differentiation, while MDMA is a new MA method that uses the characteristics of the source semantic domain from a higher information dimension. Semantic features extracted from information sources with the help of the AI model enable the modeling of user's personality characteristics, which provides a knowledge basis for the differentiation of multi-user semantic information. As shown in Fig. 1, combining traditional physical resources and the model information space, the channel capacity can be expressed as follows:

$$C = TB\sum_{k=0}^{n}\log_2\left(\boldsymbol{I}+\frac{P_k}{\sigma^2}\boldsymbol{H}\boldsymbol{\Phi}_k\boldsymbol{H}^{\mathrm{H}}\right), \quad (1)$$

where $T, B, P, \boldsymbol{H}$, and $\boldsymbol{\Phi}$ represent the time-domain resource, frequency-domain resource, power-domain resource, space-domain resource, and model information space resource, respectively, $\boldsymbol{I}$ is the identity matrix, and σ^2 is the noise variance. The semantic information extracted by an AI model from different users may contain both shared and personalized information in the high-dimensional space, but this distinction may be lost or obscured when the information is compressed into a lower-dimensional space. In the MDMA technology, we identify, analyze, and extract the shared and personalized information represented in the semantic information extracted by the model from multiple users, and then superimpose and reuse the signals bringing the shared information to achieve the effect of MA and reduce the transmission bandwidth.

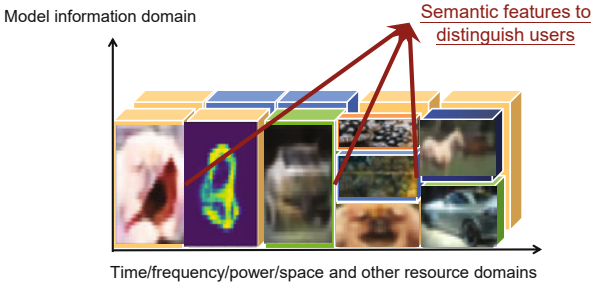


Fig. 1 The new paradigm of semantic communication creates a model information space, which is different from traditional physical resources (MDMA realizes user differentiation through semantic information features extracted by the model)

The main contributions of this paper are summarized as follows:

1. Through a deep understanding of the semantic information extracted by the AI model, the semantic information extracted from different users is characterized as shared and personalized information, and the MA technology for the semantic communication paradigm, namely, MDMA, has been proposed for the first time.

2. From the perspective of information theory, we define a new metric called "feasibility" to evaluate the systematic service capability for each user in MA systems, and construct a feasible region based on the upper limit of the feasibility for each user. We prove that MDMA schemes have greater feasible regions than the traditional MA methods, which means that more performance gains can be attained by MDMA schemes compared with the traditional schemes.

3. By conducting simulations using the datasets of CIFAR-10 (Krizhevsky, 2009) and OpenImage (Kuznetsova et al., 2020), multi-user transmission with MDMA can save an additional 25% bandwidth when compared to image-based semantic communication (Dong et al., 2023). This demonstrates that the use of multi-user transmission is a more efficient and effective approach for transmitting image data across multiple users.

2 MDMA: uplink and downlink designs

When semantic signals are extracted, they are mapped into a high-dimensional space to express the source in another form, and we call the high-dimensional space the semantic model space. We demonstrate in this section that there is a large amount of shared information in semantic signals

between different users in the semantic model space. We begin by presenting the semantic information that is shared and personalized to the two users, followed by a detailed description of MDMA uplink and downlink processes based on these shared and individual semantic features.

2.1 Multi-user semantic information analysis

The LSCI system (Dong et al., 2023) is used to extract the semantic information from MNIST and OpenImage datasets. As shown in Fig. 2, the blue and orange lines represent different sets of image-related semantic information extracted by user 1 and user 2, respectively.

From Fig. 2a, we observe that the semantic information sets of user 1 and user 2 are quite similar, despite the fact that the signal maps are extracted from different-numbered images; further, there are only minor variations in amplitude and phase. Fig. 2b displays the semantic signals of two more complex images. There is a significant similarity between the two signals in the first 30 dimensions. Notably, in dimensions 60–80, the signal amplitudes are similar but the polarities are opposite, indicating that the amount of information shared between the two signals is also minimal.

To further illustrate the existence of shared and personalized information among different users, we give the statistical distribution on each semantic feature dimension extracted from the MNIST test set, as shown in Fig. 3. The statistical distribution wrapped in a red box shows that the shared information is not a random phenomenon, but generally exists in certain feature dimensions.

Based on the results of the preliminary experiments, it is evident that there is a significant amount of shared information between the semantic signals of multiple users. The proposed MDMA technology aims to leverage this shared information to prevent bandwidth waste due to redundant transmissions.

Although the performance of extracting shared information is important for the capacity domain of MDMA, we focus only on the discovery of common information. The method we use to extract common information is relatively simple. We compare the variance of each dimension of two semantic vectors and set a threshold for the percentage of commonality.

The concept of our proposed MDMA should have a generalized definition. MDMA does not specifically differentiate between the semantic signals of different users in the semantic signal space. Instead, we mine the shared and personalized semantic information of different users in the model information space. Among them, the shared semantic information is transmitted within the same time-frequency resources, while personalized semantic information is transmitted separately. The gain of MDMA, compared to those of other MA techniques, comes mainly from the reuse of shared information among different users in the model information space.

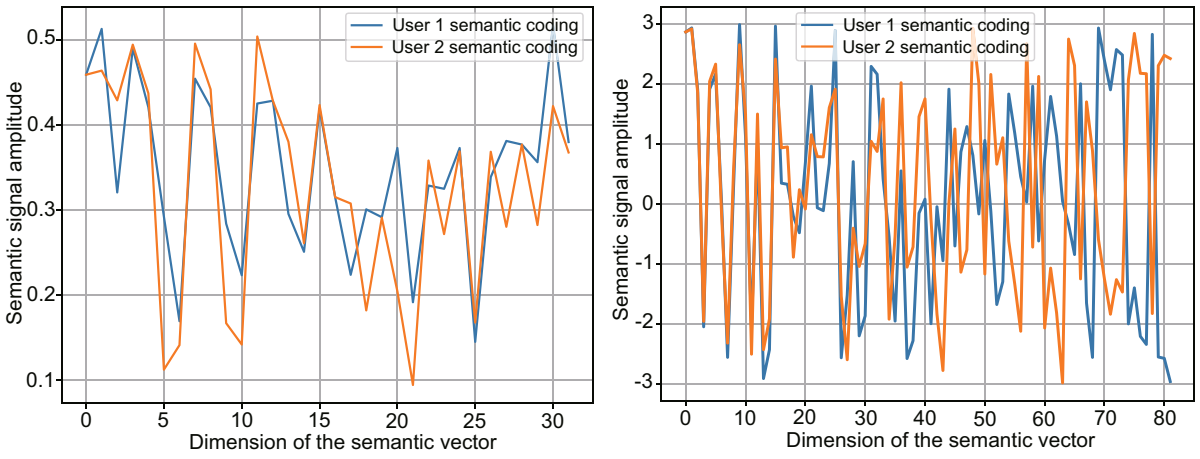


Fig. 2 Semantic signals of different images extracted by two users for the MNIST dataset (a) and the OpenImage dataset (b) (References to color refer to the online version of this figure)

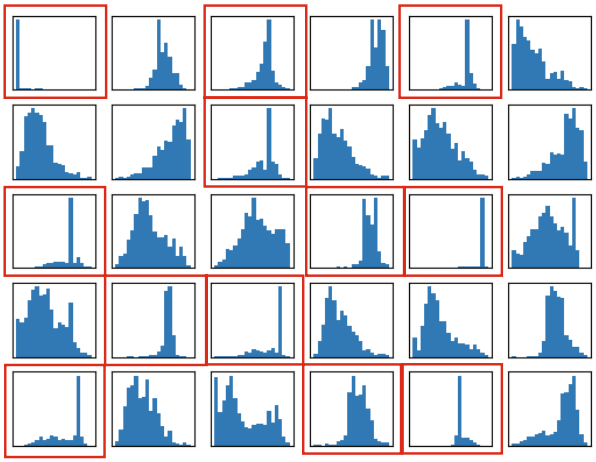


Fig. 3 Statistical distribution on each semantic feature dimension extracted from the MNIST test set (References to color refer to the online version of this figure)

2.2 Design of the uplink channel based on MDMA

Fig. 4 depicts an uplink scenario in which multiple users simultaneously transmit semantic information to the base station. The base station uses a semantic decoder to reconstruct the original source. First, the base station matches two users who have initiated uplink transmission instructions. User 1 and user 2 extract their source semantic information S_x and S_y, respectively, as follows:

$$S_x = \phi_1(X), S_y = \phi_2(Y), \tag{2}$$

where S_x and S_y represent the semantic information extracted from X and Y, respectively.

Then user 1 and user 2 send the shared semantic information $S_{x_\mathrm{s}} + S_{y_\mathrm{s}}$ to the base station at time 1 (frequency 1). Thereafter user 1 and user 2 send the personalized semantic information S_{x_p} and S_{y_p} to the base station at time 2 (frequency 2). Arbitrary channel coding modules can be configured to transmit the above semantic information.

After receiving the shared and personalized information from user 1 and user 2, the base station can recover the original semantic signals of each user, and recover the source of information using the semantic decoder.

2.3 Design of the downlink channel based on MDMA

The downlink design, as depicted in Fig. 5, allows for the base station to distribute semantic signals from multiple source data to multiple users simultaneously. First, the base station searches for matching users and extracts the semantic information using AI-based semantic models based on users' requirements.

By comparing the similarity between the two semantic features, the shared information S_{x_s} and S_{y_s} and the personalized information S_{x_p} and S_{y_p} can be extracted. The signals carrying the shared information S_{x_s} and S_{y_s} are combined and sent together, while the signals carrying personalized information S_{x_p} and S_{y_p} are transmitted separately to the user. Thus, user 1 and user 2 receive the shared information and their own personalized information, and the original images X and Y can be restored using the semantic decoder.

3 System model and theoretical analysis

3.1 Uplink MDMA system

We consider a K-user ($K \geq 2$) uplink MDMA system with a Gaussian MA channel. Each user i ($1 \leq i \leq K$) attempts to send the source $\boldsymbol{x}_i$ to the base station, by using a model-based semantic-channel encoder $\phi_i(\cdot)$ to map the source $\boldsymbol{x}_i$ to the channel-input sequences $\phi_i(\boldsymbol{x}_i)$. Assume that each sender i transmits its signal with power P_i on a series of shared time-frequency resources. After transmitting in the MA channel, the signal received at the base station can be expressed as follows:

$$\boldsymbol{y} = \sum_{i=1}^{K} \sqrt{P_i}\phi_i(\boldsymbol{x}_i) + \boldsymbol{n}, \tag{3}$$

in which $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$ respresents the white Gaussian noise received at the base station, and σ^2 refers to the power of noise $\boldsymbol{n}$.

According to the analysis in the previous section, much shared semantic information is extracted by users, so the bandwidth and time used by users can overlap. Let $\mathcal{T}_i$ and $\mathcal{F}_i$ be the time and bandwidth resources occupied by user i ($1 \leq i \leq K$), respectively, and let $T_i = |\mathcal{T}_i|$ and $B_i = |\mathcal{F}_i|$ indicate

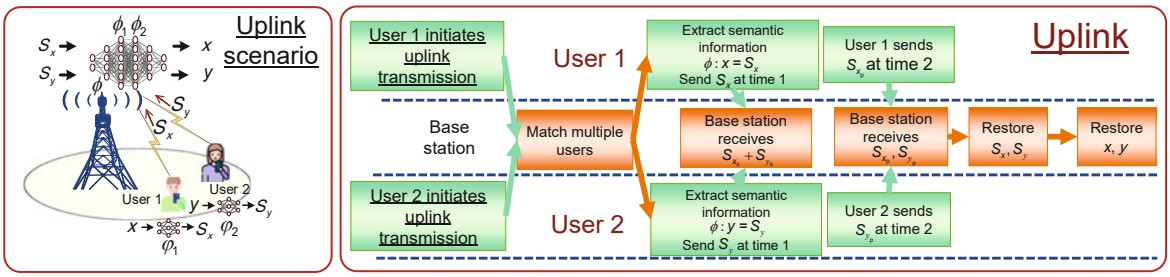


Fig. 4 Schematic of uplink MDMA. The base station matches multiple users, receives the users' shared information and personalized information, restores the original semantic information of each user, and then restores the source through the semantic decoder

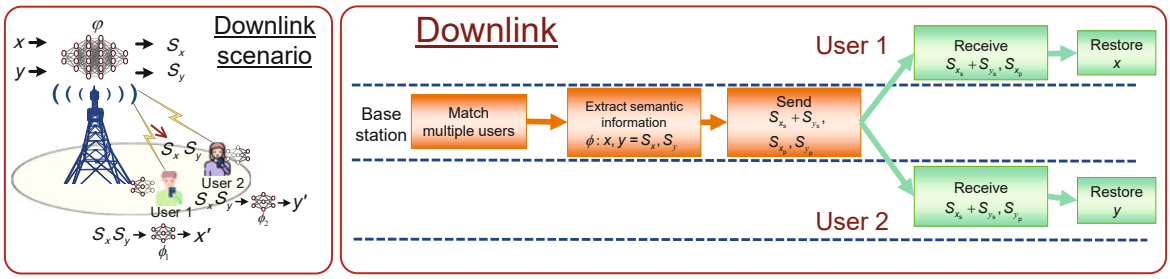


Fig. 5 Schematic of downlink MDMA. After multiple users extract the source semantic information separately, they transmit their semantic information to the base station at the same time and save bandwidth by reusing the shared information

the corresponding time and bandwidth overheads of user i ($1 \leq i \leq K$), respectively. We define a metric called the semantic overlap rate (Sor) to characterize the resource reuse rates in the time or frequency domain, which can be expressed as follows:

$$\text{Sor} = \frac{|\mathcal{T}_1 \cap \mathcal{T}_2 \cap \cdots \cap \mathcal{T}_K|}{|\mathcal{T}_1| + |\mathcal{T}_2| + \cdots + |\mathcal{T}_K|}, \tag{4}$$

or

$$\text{Sor} = \frac{|\mathcal{F}_1 \cap \mathcal{F}_2 \cap \cdots \cap \mathcal{F}_K|}{|\mathcal{F}_1| + |\mathcal{F}_2| + \cdots + |\mathcal{F}_K|}. \tag{5}$$

By evaluating the relationship between the reuse rates Sor and the performance of multiple-user signal detection, the resource-saving ability of uplink MDMA schemes can be shown in comparison to traditional multiple access methods.

In addition, based on a theoretical analysis, we provide an explanation for why the uplink MDMA system has performance advantages over the traditional uplink MA methods.

Evidently, as shown in Eq. (3), the system model of uplink MDMA is similar to that of the uplink power domain NOMA (PD-NOMA) in terms of form. Hence, it can be easily proved that the uplink MDMA system can achieve the limit of the achievable capacity region of the MA channel (Cover and Thomas, 2001) as the PD-NOMA scheme does (Wu et al., 2018).

For further analysis, we draw our inspiration from the source-channel coding theorem of the information theory (Cover and Thomas, 2001), which is also suitable for the MA system:

Theorem 1 (Source-channel coding theorem) If a finite alphabet stochastic process $\boldsymbol{X}$ satisfies that its entropy $H(\boldsymbol{X})$ is less than the channel capacity C, i.e., $H(\boldsymbol{X}) < C$, there exists a source-channel coding $\Phi(\boldsymbol{x}), \boldsymbol{x} \in \boldsymbol{X}$ with the probability of error $\Pr(\hat{\boldsymbol{x}} \neq \boldsymbol{x}) \to 0$.

Apparently, the above theorem holds true for the semantic-channel encoder $\phi_i(\cdot)$ in the MDMA system, as it does not differentiate between separate and joint source-channel coding.

To facilitate a theoretical comparison of the performance of the uplink MDMA system with that of the NOMA system, we set the ratio of the channel transmission rates R_{cc} (upper-limited by the channel capacity C of the MA system) to the source coding rates R_{sc} (lower-limited by the entropy $H(\boldsymbol{X})$ of source $\boldsymbol{X}$) as a new metric called feasibility (labeled as F), which is used to characterize the service capabilities that the channel can provide for the MA

systems, i.e.,

$$F = \frac{R_{\mathrm{cc}}}{R_{\mathrm{sc}}}. \tag{6}$$

Based on this new metric, we expand the capacity region to an achievable feasible region for the uplink MDMA and other compared schemes with a two-user MA channel, as shown in Fig. 6. To achieve the upper bound of the achievable feasible region, the base station is assumed to detect both users' signals to achieve the limit of the achievable capacity region, e.g., by using successive interference cancellation (SIC). C_i^{u} represents the achievable channel capacity for user i, i.e.,

$$C_i^{\mathrm{u}} = \frac{1}{2}\log_2\left(1+\frac{P_i}{\sigma^2}\right),\ i=1,2, \tag{7}$$

and C_{12}^{u} means the total channel capacity for this two-user MA channel, i.e.,

$$C_{12}^{\mathrm{u}} = \frac{1}{2}\log_2\left(1+\frac{P_1+P_2}{\sigma^2}\right), \tag{8}$$

in which P_1 and P_2 are the powers of the signals transmitted by user 1 and user 2, respectively. Besides, $R_{\min}^{\mathrm{NOMA}}$ and $R_{\min}^{\mathrm{MDMA}}$ express the minimum source encoding rates of both users using the traditional source encoder in uplink NOMA and the semantic encoder in uplink MDMA, respectively, expressed as follows:

$$R_{\min}^{\mathrm{NOMA}} = \min\left(R_1^{\mathrm{NOMA}}, R_2^{\mathrm{NOMA}}\right), \tag{9}$$

$$R_{\min}^{\mathrm{MDMA}} = \min\left(R_1^{\mathrm{MDMA}}, R_2^{\mathrm{MDMA}}\right), \tag{10}$$

in which R_i^{NOMA} and R_i^{MDMA} represent the source encoding rates of each user i in NOMA and MDMA, respectively.

Fig. 6 shows that, compared with the feasible region of the uplink NOMA scheme, the channel provides lower service capacities for uplink TDMA/FDMA schemes because their sum transmission rates cannot achieve the total channel capacity (Cover and Thomas, 2001; Wu et al., 2018). However, since the semantic encoders typically exhibit lower encoding rates $R_{\min}^{\mathrm{MDMA}}$ compared with traditional encoders $R_{\min}^{\mathrm{NOMA}}$, the feasible region of uplink MDMA can be greater than that of uplink NOMA, i.e.,

$$\frac{C_i^{\mathrm{u}}}{R_{\min}^{\mathrm{MDMA}}} > \frac{C_i^{\mathrm{u}}}{R_{\min}^{\mathrm{NOMA}}},\ i=1,2, \tag{11}$$

so the channel can provide higher service capability for uplink MDMA, which makes uplink MDMA obtain performance benefits compared to uplink NOMA.

3.2 Downlink MDMA system

Let us shift the focus to the K-user ($K \geq 2$) downlink MDMA system with a Gaussian broadcast channel. Under this circumstance, the base station attempts to send K sources $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_K$ to their corresponding user equipments (UEs), by using a series of model-based semantic-channel encoders $\phi_1(\cdot), \phi_2(\cdot), \cdots, \phi_K(\cdot)$ to map the sources to channel-input sequences $\phi_1(\boldsymbol{x}_1), \phi_2(\boldsymbol{x}_2), \cdots, \phi_K(\boldsymbol{x}_K)$. The channel-input sequences are assigned with the power factors $\alpha_1, \alpha_2, \cdots, \alpha_K$, constrained as

$$\sum_{i=1}^{K}\alpha_i = 1, \tag{12}$$

so that the sequences can be overlaid on the shared time-frequency resources and then broadcast to each user's receiver with the total power P. Hence, the signal received at each receiver can be modeled as

$$\boldsymbol{y}_i = \sum_{i=1}^{K}\sqrt{\alpha_i P}\phi_i(\boldsymbol{x}_i) + \boldsymbol{n}_i,\ 1 \leq i \leq K, \tag{13}$$

in which $\boldsymbol{n}_i \sim \mathcal{N}\left(\boldsymbol{0}, \sigma_i^2 \boldsymbol{I}\right)$ represents the white Gaussian noise received at the i^{th} receiver, and σ_i^2 refers to the power of noise $\boldsymbol{n}_i$.

Similarly, we can use Sor to evaluate the resource-saving ability of downlink MDMA schemes. Besides, from the viewpoint of theoretical analysis, we can build an achievable feasible region to evaluate the service capability with a two-user broadcast channel that the channel can provide for the downlink MDMA scheme, in comparison to those of the other schemes (Fig. 7). With the total transmission power P, the transmission signals for user 1 with the larger power factor α should be directly detected, so the channel capacity for user 1 needs to be expressed as

$$C_1^{\mathrm{d}} = \frac{1}{2}\log_2\left(1+\frac{\alpha P}{(1-\alpha)P+\sigma_1^2}\right), \tag{14}$$

in which σ_1^2 means the power of noise $\boldsymbol{n}_1$ at user 1's receiver. As for user 2 with the smaller power factor $1-\alpha$, to achieve the limit of the achievable capacity region, the interference from user 1 is assumed to be canceled (Cover and Thomas, 2001) (e.g., by using SIC), so that the channel capacity for user 2 can be given by

$$C_2^{\mathrm{d}} = \frac{1}{2}\log_2\left(1+\frac{(1-\alpha)P}{\sigma_2^2}\right), \tag{15}$$

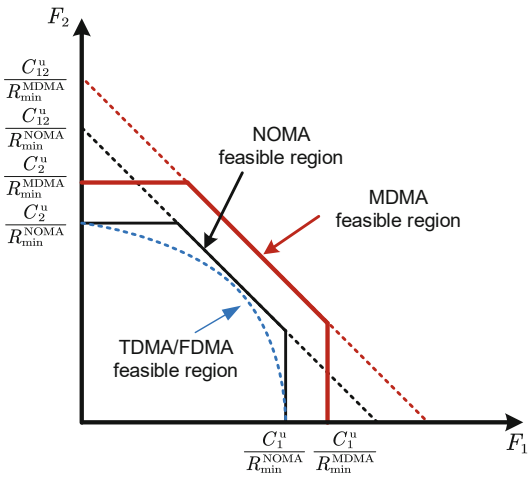


Fig. 6 Feasible region for uplink MDMA and compared schemes (References to color refer to the online version of this figure)

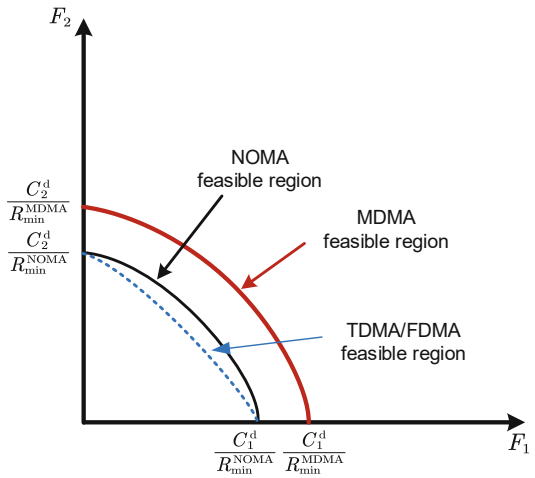


Fig. 7 Feasible region for downlink MDMA and compared schemes (References to color refer to the online version of this figure)

in which σ_2^2 means the power of noise $\boldsymbol{n}_2$ at user 2's receiver. Besides, $R_{\min}^{\mathrm{NOMA}}$ and $R_{\min}^{\mathrm{MDMA}}$ express the minimum source encoding rates for user i using the traditional source encoder in downlink NOMA and the semantic encoder in downlink MDMA, respectively, which are defined in the same form as Eqs. (9) and (10).

Fig. 7 demonstrates that the feasible region of the downlink NOMA scheme is limited by the black curve, and the regions of TDMA/FDMA schemes limited by the blue dotted curve are smaller than those for NOMA, which is similar to their capacity region (Cover and Thomas, 2001; Islam et al., 2017). However, due to the reduction of semantic encoding rates $R_{\min}^{\mathrm{MDMA}}$ compared to the traditional encoding rates $R_{\min}^{\mathrm{NOMA}}$, the feasible region of downlink MDMA is larger than that of downlink NOMA, i.e.,

$$\frac{C_i^{\mathrm{d}}}{R_{\min}^{\mathrm{MDMA}}} > \frac{C_i^{\mathrm{d}}}{R_{\min}^{\mathrm{NOMA}}},\ i=1,2, \tag{16}$$

which results in the same conclusion as in the uplink scenario; that is, the channel can provide higher service capability for downlink MDMA, and thus the downlink MDMA can attain more performance gains than the traditional schemes.

Note that since the fundamental theory of feasibility (i.e., the source-channel coding theorem) does not differentiate between separate and joint semantic (source) channel coding, MDMA schemes based on these two coding frameworks can have the same achievable feasible region in both uplink and downlink systems.

Furthermore, note that the above analysis can only theoretically prove that MDMA schemes have better performance than TDMA/FDMA or NOMA schemes in both uplink and downlink systems from the viewpoint of the achievable feasible region, which comes from the lower rates of the source encoding with semantic encoding caused by reuse of the shared information. The analysis results can be the upper bound of the feasible regions achieved by practical MDMA schemes.

4 Simulation results and analysis

In this section, MDMA simulations are introduced in detail. First, we introduce the setting, including the semantic communication model used, the datasets, and the baseline. Then, we describe the semantic recovery performance of MDMA under different SNRs in an additive white Gaussian noise (AWGN) channel. Finally, we show the performance comparison between the complete semantic communication system combined with MDMA and the traditional semantic communication system.

4.1 Setting and baseline

The datasets considered in the simulations include CIFAR-10 and OpenImage, which consist of 60 000 RGB images of 32×32 resolution and over 100 000 RGB images of large size, respectively. In particular, only 100 000 images are selected as the

training set and 2000 images as the test set.

OFDM (Gao et al., 2018) and NOMA (Ding et al., 2016), which are representative MA technologies of 4G and 5G, are used as the baseline for the simulations. Furthermore, note that in the MDMA simulation, no other MA techniques are combined. The simulations of MDMA are conducted by transmitting the user's personalized information and partially shared information.

The peak SNR (PSNR) is an indicator used for the measurements, defined as follows:

$$\mathrm{PSNR}(X,Y)=10\lg\frac{\mathrm{MAX}^2}{\mathrm{MSE}}, \tag{17}$$

where MSE is the mean squared-error between the reference image X and the reconstructed image Y, and MAX is the maximum possible value of the image pixels, which is equal to one in our simulations.

4.2 Evaluation on the CIFAR-10 dataset

We use the variances of two semantic information sets to express the distinction in the amount of semantic information. Then, an approach is employed to overlay and transmit the signal containing semantic information that has smaller differences in information quantity. The bandwidth occupied by the superimposed signal divided by the bandwidth of the original signal is called the semantic information reuse rate, or Sor.

As shown in Fig. 8a, with the improvement of Sor, the bandwidth occupancy rate naturally decreases gradually. When the quantities of semantic information transmitted are identical, the bandwidth occupation is half of the original bandwidth occupation. The blue and orange lines in Fig. 8a indicate the change in the recovery performance of user 1 and user 2 with the increase of Sor, respectively. The recovery performances of the two users are almost the same and are not biased. When Sor reaches 50%, the performance is almost consistent with the original recovery performance, which means that 25% bandwidth can be saved. When Sor exceeds 50%, the performance decreases gradually.

Fig. 8b shows the change in PSNR performance of MDMA under the AWGN channel with different SNRs when Sor is 50%. To compare the MA technology separately, a channel coding module is not added to the comparison schemes of NOMA and OFDM technologies, and the channel coding SeSM is also not added to LSCI in the MDMA technology in the simulations.

We take bandwidth B_{NOMA} and power P_{NOMA} of NOMA as the benchmark:

$$B_{\mathrm{NOMA}}=1, P_{\mathrm{NOMA}}=1. \tag{18}$$

The bandwidth and power of OFDM and MDMA are as follows:

$$B_{\mathrm{OFDM}}=2, P_{\mathrm{OFDM}}=1, \tag{19}$$

$$B_{\mathrm{MDMA}}=1.5, P_{\mathrm{MDMA}}=1.5. \tag{20}$$

The performance of MDMA in the AWGN channel is evidently better than that of OFDM, despite the unequal conditions. Due to the fact that semantic communication rate adaptation is influenced by multiple factors that are difficult to control, we cannot compare NOMA and MDMA under completely equal bandwidth and power conditions. Therefore, Shannon's formula is used for power normalization and bandwidth normalization.

Specifically, the power normalization operation is to shift the MDMA line to the right by 1.5 dB. The bandwidth normalization operation is to satisfy the following formula:

$$1.5\log_2(1+\gamma_1)=\log_2(1+\gamma_2), \tag{21}$$

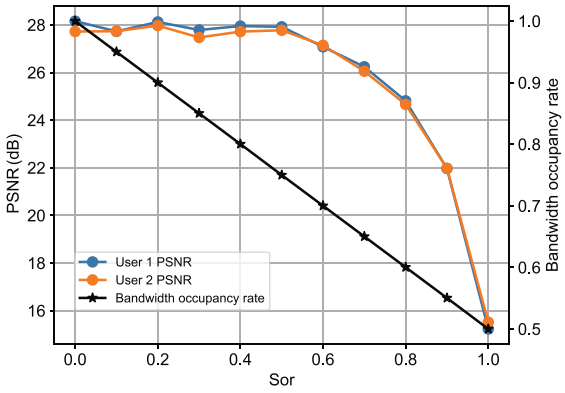
where γ_1 is the SNR of 1.5 bandwidth, and γ_2 is the SNR of the normalized bandwidth. The MDMA line shifts to the right in different degrees according to the SNR.

Fig. 8c shows the PSNR of MDMA, OFDM, and NOMA under different SNRs in the AWGN channel after power and bandwidth normalization. Under a high SNR condition, the performances of the three MA technologies are similar. Under a low SNR condition, MDMA begins to show its superiority, with a 5-dB advantage over NOMA.

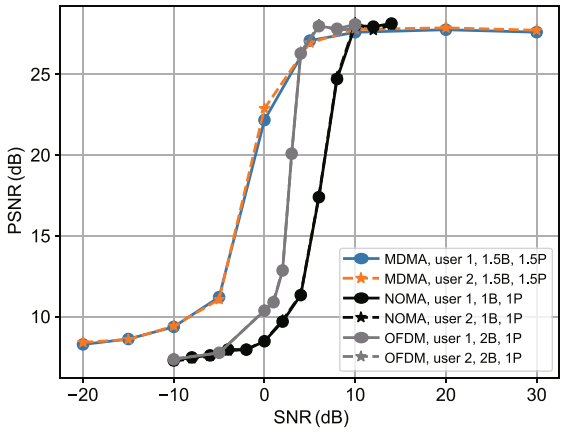
4.3 Evaluation on the OpenImage dataset

The simulation to verify the MDMA technology on the OpenImage dataset, which contains a large number of natural images of large size, is more general.

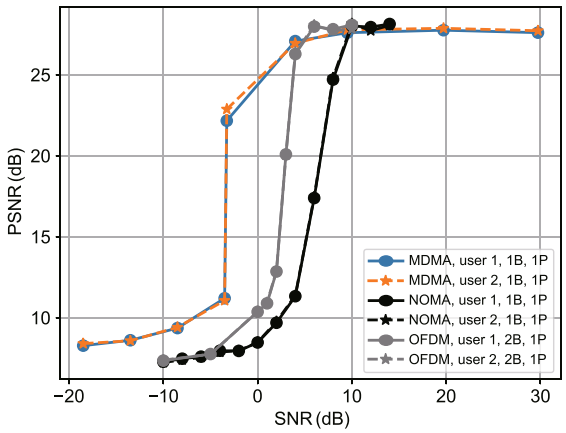
Fig. 9a shows the PSNR performance of the various models for the images under different Sor. Similar to the results of the simulation on the CIFAR-10 dataset, when Sor is more than 50%, the performance



(a)

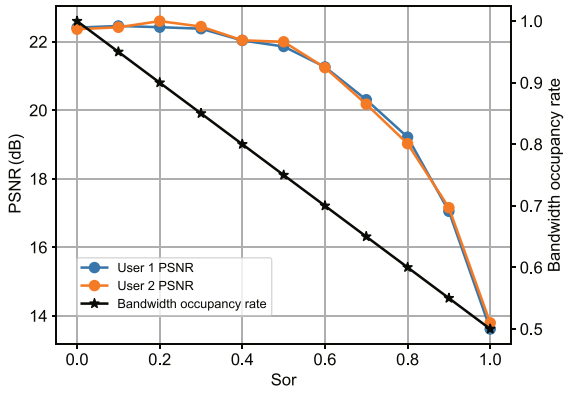


(b)

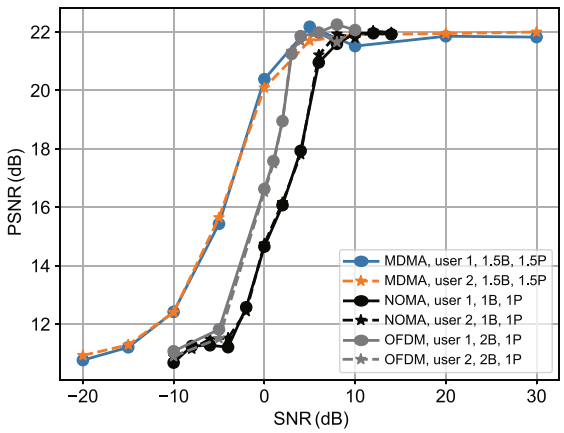


(c)

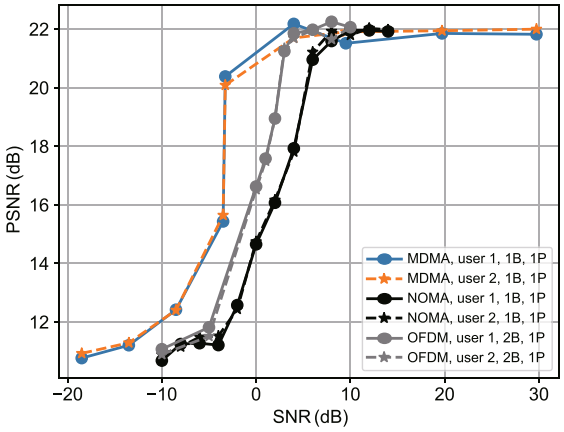
Fig. 8 PSNR performance and bandwidth saving rate under different shared semantic information reuse rates (a), PSNR performance of MDMA, NOMA, and OFDM under different SNRs in the AWGN channel with different bandwidths and powers (b), and PSNR performance of MDMA, NOMA, and OFDM under different SNRs in the AWGN channel after equivalent transformation using the Shannon formula (c) on the CIFAR-10 dataset (References to color refer to the online version of this figure)



(a)



(b)



(c)

Fig. 9 PSNR performance and bandwidth saving rate under different shared semantic information reuse rates (a), PSNR performance of MDMA, NOMA, and OFDM under different SNRs in the AWGN channel with different bandwidths and powers (b), and PSNR performance of MDMA, NOMA, and OFDM under different SNRs in the AWGN channel after equivalent transformation using the Shannon formula (c) on the OpenImage dataset (References to color refer to the online version of this figure)

has an obvious decline trend, and the performance is almost unchanged when Sor is less than 50%.

Fig. 9b shows the change in the PSNR performance of MDMA under the AWGN channel with different SNRs when Sor is 50%. Fig. 9c is the curve after power normalization and bandwidth normalization. Under the condition of high SNR channels, three kinds of MA technologies have strong robustness to weak noise. When SNR is less than 3 dB, the performance of MDMA has at least a 5-dB advantage over that of NOMA.

4.4 MDMA combined with the JSCC-based semantic communication system

To further show the compatibility and advantages of MDMA technology in semantic communication systems, we combine MDMA, LSCI, and the joint source-channel coding (JSCC) based channel slice model, and compare them with traditional communication systems.

Fig. 10 is a comparison curve after power and bandwidth normalization. Under the high SNR condition in the AWGN channel, the performance of MDMA + JSCC is slightly lower than those of NOMA + LDPC (low density parity check) + BPSK (binary phase shift keying) and OFDM + LDPC + BPSK. However, under low SNR, MDMA + JSCC has great advantages. In particular, when SNR is –10 dB, the PSNR of MDMA + JSCC is 18 dB and that of NOMA + LDPC + BPSK is only 11 dB. The simulation shows that MDMA + JSCC has a bandwidth advantage of at least about 7 dB compared with NOMA + LDPC + BPSK.

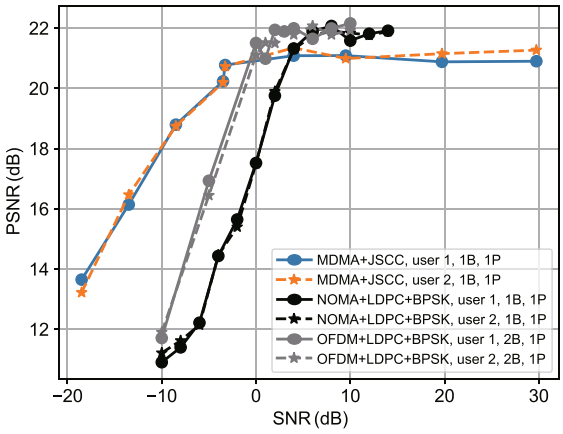


Fig. 10 Performance comparison between the JSCC-based semantic communication system of MDMA and the traditional systems of NOMA and OFDM combined with LDPC coding

MDMA is a multi-user technology based on the resource reuse method in the high-dimensional semantic space, which does not conflict with the time-frequency space in OFDM and the power space in NOMA. We verify only the compatibility and availability of MDMA in the semantic communication system, and do not deny the role of OFDM or NOMA. Just like the combination of NOMA and OFDM, it is worthwhile to explore the MA technology in the time, frequency, power, and semantic space in the future.

5 Conclusions and future work

In this paper, we mine the shared and personalized information between semantic information from different users in the high-dimensional semantic space, and propose a novel MA technology to allocate semantic space resources, named model division MA (MDMA). Specifically, we have proved that the upper limit of the proposed feasible region of MDMA is higher than that of NOMA. Simulation results show that MDMA has at least a 5-dB advantage over NOMA in the AWGN channel under a low SNR condition. Moreover, MDMA can work in a mixed mode with NOMA, which can lead to additional performance advantages while keeping compatibilities. In addition, by combining the upper and lower bounds of the JSCC excess distortion exponent in the MIMO-based semantic-aware communication system (Shi et al., 2023; Wang Z et al., 2023), the performance of multi-user transmission in the MIMO system can be further analyzed.

Contributors

Ping ZHANG, Xiaodong XU, Chen DONG, and Kai NIU proposed the main idea. All the authors designed the research. Ping ZHANG, Xiaodong XU, Chen DONG, Kai NIU, Haotai LIANG, and Zijian LIANG participated in theoretical analysis and simulation verification, and drafted the paper. All the authors revised and finalized the paper.

Compliance with ethics guidelines

Ping ZHANG, Xiaodong XU, Chen DONG, Kai NIU, Haotai LIANG, Zijian LIANG, Xiaoqi QIN, Mengying SUN, Hao CHEN, Nan MA, Wenjun XU, Guangyu WANG, and Xiaofeng TAO declare that they have no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

References

Cover TM, Thomas JA, 2001. Elements of Information Theory. John Wiley & Sons, Ltd., New York, USA, p.374-458. https://doi.org/10.1002/0471200611.ch14

Dai JC, Wang SX, Tan KL, et al., 2022. Nonlinear transform source-channel coding for semantic communications. *IEEE J Sel Areas Commun*, 40(8):2300-2316. https://doi.org/10.1109/JSAC.2022.3180802

Dai LL, Wang BC, Yuan YF, et al., 2015. Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends. *IEEE Commun Mag*, 53(9):74-81. https://doi.org/10.1109/MCOM.2015.7263349

Ding ZG, Adachi F, Poor HV, 2016. The application of MIMO to non-orthogonal multiple access. *IEEE Trans Wirel Commun*, 15(1):537-552. https://doi.org/10.1109/TWC.2015.2475746

Dong C, Liang HT, Xu XD, et al., 2023. Semantic communication system based on semantic slice models propagation. *IEEE J Sel Areas Commun*, 41(1):202-213. https://doi.org/10.1109/JSAC.2022.3221948

Gao S, Zhang M, Cheng X, 2018. Precoded index modulation for multi-input multi-output OFDM. *IEEE Trans Wirel Commun*, 17(1):17-28. https://doi.org/10.1109/TWC.2017.2760823

Islam SMR, Avazov N, Dobre OA, et al., 2017. Power-domain non-orthogonal multiple access (NOMA) in 5G systems: potentials and challenges. *IEEE J Sel Areas Commun*, 19(2):721-742. https://doi.org/10.1109/COMST.2016.2621116

Jiang PW, Wen CK, Jin S, et al., 2023. Wireless semantic communications for video conferencing. *IEEE J Sel Areas Commun*, 41(1):230-244. https://doi.org/10.1109/JSAC.2022.3221968

Krizhevsky A, 2009. Learning Multiple Layers of Features from Tiny Images. Technical Report, TR-2009. University of Toronto, Toronto, Canada.

Kuznetsova A, Rom H, Alldrin N, et al., 2020. The open images dataset V4: unified image classification, object detection, and visual relationship detection at scale. *Int J Comput Vis*, 128:1956-1981. https://doi.org/10.1007/s11263-020-01316-z

Li WZ, Liang HT, Dong C, et al., 2023. Non-orthogonal multiple access enhanced multi-user semantic communication. https://arxiv.org/abs/2303.06597

Luo XW, Gao RB, Chen HH, et al., 2022. Multi-modal and multi-user semantic communications for channel-level information fusion. *IEEE Wirel Commun*, early access. https://doi.org/10.1109/MWC.011.2200288

Mao Y, Dizdar O, Clerckx B, et al., 2022. Rate-splitting multiple access: fundamentals, survey, and future research trends. *IEEE Commun Surv Tutor*, 24(4):2073-2126. https://doi.org/10.1109/COMST.2022.3191937

Saito Y, Kishiyama Y, Benjebbour A, et al., 2013. Non-orthogonal multiple access (NOMA) for cellular future radio access. IEEE 77th Vehicular Technology Conf, p.1-5. https://doi.org/10.1109/VTCSpring.2013.6692652

Shannon CE, 1948. A mathematical theory of communication. *Bell Syst Tech J*, 27(3):379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shi YX, Shao S, Wu YP, et al., 2023. Excess distortion exponent analysis for semantic-aware MIMO communication systems. https://arxiv.org/abs/2301.04357

Wang SX, Dai JC, Liang ZJ, et al., 2023. Wireless deep video semantic transmission. *IEEE J Sel Areas Commun*, 41(1):214-229. https://doi.org/10.1109/JSAC.2022.3221977

Wang Z, Zhang JY, Du HY, et al., 2023. Extremely large-scale MIMO: fundamentals, challenges, solutions, and future directions. *IEEE Wirel Commun*, early access. https://doi.org/10.1109/MWC.132.2200443

Weng Z, Qin Z, 2021. Semantic communication systems for speech transmission. *IEEE J Sel Areas Commun*, 39(8):2434-2444. https://doi.org/10.1109/JSAC.2021.3087240

Wu Z, Lu K, Jiang C, et al., 2018. Comprehensive study and comparison on 5G NOMA schemes. *IEEE Access*, 6:18511-18519. https://doi.org/10.1109/ACCESS.2018.2817221

Xie H, Qin Z, Li GY, et al., 2021. Deep learning enabled semantic communication systems. *IEEE Trans Signal Process*, 69:2663-2675. https://doi.org/10.1109/TSP.2021.3071210

Zhang P, Xu WJ, Gao H, et al., 2022a. Toward wisdom-evolutionary and primitive-concise 6G: a new paradigm of semantic communication networks. *Engineering*, 8:60-73. https://doi.org/10.1016/j.eng.2021.11.003

Zhang P, Peng MG, Cui SG, et al., 2022b. Theory and techniques for "intellicise" wireless networks. *Front Inform Technol Electron Eng*, 23(1):1-4. https://doi.org/10.1631/FITEE.2210000

Zhang P, Xu XD, Dong C, et al., 2022c. Intellicise communication system: model-driven semantic communications. *J China Univ Posts Telecommun*, 29(1):2-12. https://doi.org/10.19682/j.cnki.1005-8885.2022.2002

Zhang Y, Xu W, Gao H, et al., 2022. Multi-user semantic communications for cooperative object identification. IEEE Int Conf on Communications Workshops, p.157-162. https://doi.org/10.1109/ICCWorkshops53468.2022.9814491

Zhu YW, Huang YK, Qiao XQ, et al., 2022. A semantic-aware transmission with adaptive control scheme for volumetric video service. *IEEE Trans Multim*, early access. https://doi.org/10.1109/TMM.2022.3217928