Semantic-Aware Sensing Information Transmission for Metaverse: A Contest Theoretic Approach

Jiacheng Wang, Hongyang Du, *Graduate Student Member, IEEE*, Zengshan Tian, *Member, IEEE*,
Dusit Niyato, *Fellow, IEEE*, Jiawen Kang, *Member, IEEE*, and Xuemin Shen, *Fellow, IEEE*

Abstract—With the advancement of network and computer technologies, virtual cyberspace keeps evolving, and Metaverse is the main representative. As an irreplaceable technology that supports Metaverse, the sensing information transmission from the physical world to Metaverse is vital. Inspired by emerging semantic communication, in this paper, we propose a semantic transmission framework for transmitting sensing information from the physical world to Metaverse. Leveraging the in-depth understanding of sensing information, we define the semantic bases, through which the semantic encoding of sensing data is achieved for the first time. Consequently, the amount of sensing data that needs to be transmitted is dramatically reduced. Unlike conventional methods that undergo data degradation and require data recovery, our approach achieves the sensing goal without data recovery while maintaining performance. To further improve Metaverse service quality, we introduce contest theory to create an incentive mechanism that motivates users to upload data more frequently. Experimental results show that the average data amount after semantic encoding is reduced to about 27.87% of that before encoding, while ensuring the sensing performance. Additionally, the proposed contest theoretic based incentive mechanism increases the sum of data uploading frequency by 27.47% compared to the uniform award scheme.

Index Terms—Semantic-aware transmission, contest theory, metaverse, wireless sensing.

I. INTRODUCTION

IN 1992, Neal Stephenson proposed the concept of Metaverse in its novel *Snow Crash* for the first time. As an imagined virtual environment parallel to the physical world, in Metaverse, users can build their own avatars, analogous to their physical selves, to experience an alternate life [1]. Due to limited technology, resources, and other factors, at that time, Metaverse appeared to exist only in fantasy, which was out of reach. However, as time goes by, the economies have thrived and various technologies have evolved, bringing the once-imagined world into reality. In 2021, Meta (formerly Facebook) released Horizon Worlds [2], a free virtual reality online video game with an integrated game creation system, which allows users to build their own virtual characters and explore the space around them within the confines of their physical floor-space. In addition to giving people a better understanding of Metaverse, similar releases also mark the transformation of Metaverse's from fantasy to reality.

As a novel type of Internet application and social form, Metaverse's transformation from a concept to reality is inextricably linked to innovative technologies. Digital twin-based data replication, for example, supports continuous data synchronization for Metaverse, that is, obtaining live data from the physical world and transmitting it to Metaverse to keep it updating in real time [3].The extended reality technologies can render Metaverse and create an immersive experience for users [4]. Moreover, blockchain technology, which is capable of tightly integrating the virtual world and the real world into the economic, social, and the identity system [5], is also essential.

While eye-catching, the implementation of the above technologies rests on a fundamental premise, i.e., the data portraying the physical world can be effectively (accurately) and efficiently (fast) transmitted to Metaverse, which is of great challenge due to the large amount and variety of data. Taking the popular wireless sensing as an example, a pair of sensing devices can generate 3.072 megabytes of sensing data per second.[1] Even, this number could grow further as sensing

Jiacheng Wang is with the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China, and also with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: jiacheng.wang@ntu.edu.sg).

Hongyang Du and Dusit Niyato are with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: hongyang001@e.ntu.edu.sg; dniyato@ntu.edu.sg).

Zengshan Tian is with the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: tianzs@cqupt.edu.cn).

Jiawen Kang is with the School of Automation and the Key Laboratory of Intelligent Information Processing and System Integration of IoT, Guangdong University of Technology (GDUT), Ministry of Education, Guangzhou 510006, China (e-mail: kavinkang@gdut.edu.cn).

Xuemin Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: sshen@uwaterloo.ca).

[1]Consider a transmitter equipped with one antenna sends signal with bandwidth of 80 Mhz and packet transmission rate of 500 Hz, and the receiver captures signal with three antennas, under the IEEE 802.11ac protocol. Then, the amount of sensing data produced per second can reach $3 \times 256 \times 500 \times 2 \times 4$ bytes, where 256 is the number of Orthogonal Frequency Division Multiplexing (OFDM) subcarriers, 2 represents I and Q sampling.
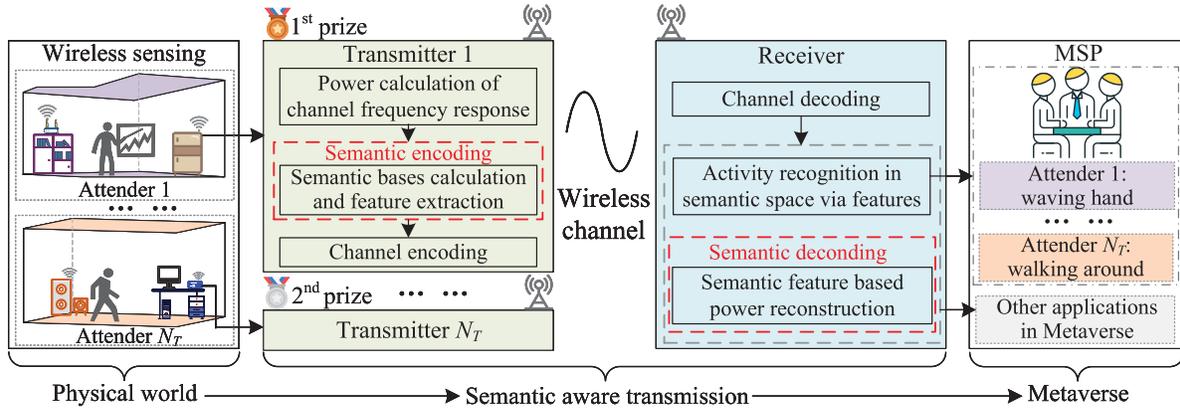
Fig. 1. Framework of the proposed system. In the physical world, wireless smart devices are used to sense the users by sending and receiving wireless signals, and then the transmitter encodes the sensing information semantically and sends it out through wireless channels. In the MSP, the receiver captures the semantic information and realizes activity recognition in the defined semantic space. Meanwhile, through the semantic decoding, the CFR power can be fully recovered, thereby supporting other Metaverse applications. Furthermore, a contest theory-based incentive mechanism is introduced to motivate the transmitter to upload data more frequently, improving the quality of MSP services.

granularity and the number of users increase, triggering not only an unbearable burden on data transmission and storage but also a non-negligible data processing delay, affecting the quality of service (QoS) of the MSP. For example, equipment providers may struggle to support high refresh rate and rendering due to storage and computing overhead. Therefore, the effective and efficient transmission of sensing data is a urgent issue that must be addressed.

Semantic communication, which intends to extract the meaning of data, filter out the irrelevant information, and transmit it to the receiver, has natural advantages in reducing data storage and transmission load and improving data transmission efficiency [6]. Semantic communication is task-oriented, i.e., aims to convey the information relevant to the transmission goal. In image transmission, for instance, the information carried in the picture is essential, instead of the picture itself. With this in mind, semantic communication extracts information from the picture according to the transmission goal and gives it to the receiver to realize higher transmission efficiency. Therefore, if sensing data can be transmitted semantically, the burden of data communication and storage would be greatly reduced, laying the foundation for improving the QoS of MSP.

Inspired by this, in this paper, we propose a semantic-aware sensing data transmission framework, including a contest theory based incentive mechanism, to deliver the wireless sensing data, i.e., channel frequency response (CFR) power, from the physical world to Metaverse. As shown by the proposed framework in Fig. 1, the transmitter focuses on data semantic coding and transmitting, the receiver works on data receiving, human activity recognition, and CFR power recovery to support other applications, while the contest theory based mechanism concentrates on encouraging the transmitter to upload sensing data at a higher frequency via award setting. Through such a collaboration, the amount of sensing data transmitted each time decreases significantly while the data upload frequency increases remarkably, providing support for MSP to further improve its QoS. The contributions of this paper are summarized as follows:

- We define a novel semantic base to encode the information of human activity in the CFR. On this basis, the semantic features are further extracted and transmitted to

the receiver. Unlike transmitting the original CFR, our result shows the average data amount after semantic encoding is compressed to about 27.87% of that before encoding.
- We propose a semantic space, in which receiver realizes human activity recognition via the received semantic features, without the requirement for recovering the entire CFR power. We build an IEEE 802.11ac protocol based platform and collect data in real-world scenarios to verify the effectiveness of semantic coding and activity recognition in semantic space.
- We design a contest theory based strategy to incentivise the uploading of sensing data, to further improve the users experience in Metaverse. We derive the optimal award setting that maximises the total data uploading frequency, which directly affects the refresh rate of MSP, and investigate the advantages of semantic encoding in the contest. Numerical analysis proves and demonstrates the effectiveness of our proposed strategy.

II. RELATED WORK

This section investigates the major works related to the system proposed in this paper.

A. Compressed Sensing

The channel state information (CSI) compression has always been a major topic in mobile wireless network. Previous attempts [7], [8], [9] use quantization and general techniques to compress the CSI. For example, the authors in [8] quantize CSI from time, frequency and numerical values domains, and adapt compression intensity to balance the capacity loss and overhead reduction. Other works [10], [11] reduce the request rate and amount of CSI to achieve the compression. The authors in [10] evaluate the required CSI repetition rate for multiple-input multiple-out put (MIMO) systems and provide a theoretical support for reducing CSI transmission in time domain. Besides, the sparse nature of the wireless channel is always used to underpin CSI compression [12], [13]. The authors in [13] analyze CSI compression via the sparsity obtained by modeling feedback distortion and rate loss. CsiNet [14] learns a transformation from CSI to a near-optimal number of representations via deep learning to realize CSI sensing

and recovery. Moreover, CsiNet-LSTM [15] improves the CSI recovery quality and trade-off between compression ratio and complexity by learning spatial structures from training samples.

B. Semantic Communications

The semantic communication was first introduced in [16], then researchers present the semantic information theory [17] and a universal semantic communication model [18]. Following that, the authors in [19] develop a compression theory at the semantic level to reduce the amount of data. The authors in [20] use the semantic redundancy and ambiguity to achieve compression. Using AI techniques, semantic communication is further developed both in technology and application. DeepSC [21] utilizes Transformer and transfer learning to maximize the capacity and minimize the errors in sentence meaning recovery, so as to perform better when the signal-to-noise ratio (SNR) is poor. Through the adversarial training and masked auto-encoder, the authors in [22] present the first end-to-end semantic communication system to improve the performance against noise. A Transformer based semantic communication system is presented in [23], which introduces a circulation mechanism to transmit sentences more flexibly. Besides, semantic transmission of speech and images has also received extensive attention [24], [25].

C. Contest Theory

The contest is a game, in which contestants spend resources to win prizes with some probability. This definition can be applied to many situations to solve different problems. Early attempts, such as rent-seeking model [26], all-pay auction [27], and others [28], [29], focused on economic problems. Besides economics, contest theory also plays a significant role in other areas. The authors in [30] use contest to balance request and service among users, so as to solve the problem of low willingness of service in a service exchange application. In [31], the authors use the contest theory in crowd sourcing to study the reward distribution strategies and estimate users' capabilities via the observed competition results. The authors in [32] show requester can maximize contributions by rewarding only the top contributors, whereas the low-capability participants become risk-averse and are unwilling to join the contests. To handle this problem, the following work [33] introduces a lottery mechanism based on Tullock contest, so that every player has a positive chance of winning if they participate.

Motivated by the above studies, we design a semantic transmission system for wireless sensing data. Our system uses the proposed semantic encoding to reduce the amount of sensing data that needs to be transmitted. Besides, we propose a competition based mechanism to stimulate the data upload frequency, which further improves the system overall performance.

III. SYSTEM OVERVIEW

A. Wireless Sensing in Physical World

As shown in Fig. 1, without loss of generality, we use the activity recognition of users in virtual conference service as an example to illustrate our work. In the physical world, smart devices sense the attendee by transmitting and receiving signals and perform semantic encoding and other processing on the obtained sensing data. Here, the data refers to the CFR that is widely available in existing wireless communication systems and its power describes the physical channel fluctuation induced by attendee activity. Therefore, the CFR power, dented as $|H(f, t)|^2$, where f and t represent the frequency and time of CFR extraction, respectively, holds the potential for the pervasive sensing in the indoor environment [34]. Once the CFR obtained, the transmitter needs to process it. However, there are two issues need to be addressed.

- Due to the limited storage and computing resources, sensing devices need to send the $|H(f, t)|^2$ via the transmitter module to the receiver, e.g., edge cloud nodes or fog nodes close to the sensing devices, for activity recognition. However, as mentioned previous, the sensing incurs a large amount of data, resulting in significant communication overheads.
- One receiver needs to serve multiple transmitters. Therefore, if all transmitter upload their data simultaneously to the receiver, the receiver will face significant costs due to the high demands on processing power and storage space.

These issues motivate us to develop a simple but effective semantic encoder to reduce the amount of data and the processing burden of receiver while maintaining activity recognition accuracy.

B. Semantic-Aware Encoding at Transmitter

In this paper we focus on extracting and preserving semantic information related to attendee activity from CFR power, which allows us to transmit and store less data. Recall that the sinusoidal wave, the most natural representation of a signal, can be stored with only three parameters, i.e., amplitude, frequency, and initial phase. Although an infinite number of sinusoidal waves may be required to accurately describe a signal, we show later that a few waves are sufficient to maintain the CFR semantic properties required for human activity recognition. Therefore, we define semantic base as the sinusoidal wave, and semantic features as the corresponding parameters, including amplitude, frequency, and phase. Through the semantic bases, CFR can be semantically encoded. The theoretical support is presented in Section IV-A. Given this, we consider the case that one receiver with N_R antennas serves N_T transmitters, as shown in Fig. 1. Each transmitter only needs to transmit semantic features to the receiver through the wireless channel, after CFR power calculation, semantic encoding, and channel encoding. By doing so, the transmission and storage resources are greatly reduced.

C. Activity Recognition at Receiver With Contest Incentive

With respect to the transmitter, the receiver first conducts channel decoding to extract semantic features from the received signal. The features are directly related to the activities of the conference attendees. Therefore, the k-nearest neighbor (kNN) is applied to classify the semantic features in

semantic space (detailed in Section IV-C) to achieve activity recognition. After that, the recognition results are fed to the MSP to complete the virtual conference function in Metaverse. Besides, receiver can restore the entire CFR power by combining sine waves described by the semantic features, so as to support other services in Metaverse.

The receiver serves multiple transmitters at the same time, each of which has different sensing, processing, and data transmission capabilities. As mentioned previously, the data transmission rate directly influences the refresh rate and rendering of the MSP, which further affects the user experience. Therefore, we introduce a reward mechanism (in Section IV-D) for transmitters of the service framework according to the contest theory. To win the reward, each transmitter needs to upload sensing data with a higher frequency. Therefore, with this incentive, the overall rendering and refresh effects at the MSP are boosted, which further enhances the user experience.

IV. System Model

A. Wireless Sensing and Semantic-Aware Encoding

In this part, the theoretical support of semantic-aware encoding is introduced. As the framework shown in Fig. 1, various wireless smart devices in the physical world are used to sense the conference attendee by sending and receiving wireless signals. Since the surrounding objects, the transmitted signal reaches at the receiver via multiple propagation paths. Assuming there are L different propagation paths, then the CFR of the wireless channel can be denoted as

$$H(f,t) = \sum_{l=1}^{L} a_l(f,t) \exp\left(-j2\pi f \frac{d_l(t)}{c}\right)$$
$$\times \exp\left(-j2\pi \left(f(\varepsilon_1 + \varepsilon_2) + \varepsilon_3\right)\right), \quad (1)$$

where $a_l(f,t)$ represents the attenuation and the initial phase of the lth propagation path, $\exp(-j2\pi f d_l(t)/c)$ is the phase introduced by propagation delay corresponding to the distance $d_l(t)$, c is the signal propagation speed, and ε_1, ε_2, and ε_3 are the time offset introduced by the symbol timing offset (STO), sampling frequency offset (SFO), and carrier frequency offset (CFO), respectively. As shown in (1), the change of $d_l(t)$ triggers a phase shift to the signal corresponding to the lth propagation path. In the indoor environment, the propagation path length of reflections induced by static objects, such as walls and furniture, are constant, while the path length of the conference attendee induced reflection varies with time. Therefore, the propagation paths can be categorized into static and dynamic parts. Leveraging this fact, we can rewrite (1) as

$$H(f,t) = \exp(-j2\pi\varepsilon) \times (H_s(f,t) + H_d(f,t)), \quad (2)$$

where $\varepsilon = f(\varepsilon_1 + \varepsilon_2) + \varepsilon_3$, constant vector $H_s(f,t)$ is the sum of static reflections, and

$$H_d(f,t) = \sum_{l_d \in P_d} a_{l_d}(f,t) \exp(-j2\pi f d_{l_d}(t)/c) \quad (3)$$

is the sum of dynamic reflections, P_d is the set of the dynamic reflections. Without loss of generality, we can assume that the l_d^{th} propagation path is introduced by the conference attendee
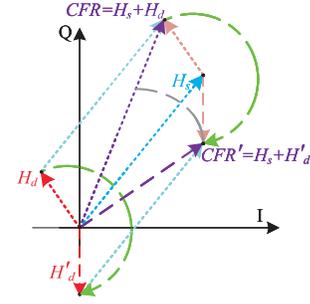

Fig. 2. Impact of the H_d phase on the overall CFR power.

and path length changes at a constant rate of v_k in a short period of time. Then, we have

$$H_d(f,t) = \sum_{l_d \in P_d} a_{l_d}(f,t) \exp(-j2\pi f (d_{l_d}(0) + v_k t)/c),$$
$$(4)$$

where $d_{l_d}(0)$ is the initial path length of attendee induced reflection. Apparently, the phase of the CFR contains information about v_k, which has been proven to be directly related to the activity of the human body [34]. Due to the phase error $\varepsilon = f(\varepsilon_1 + \varepsilon_2) + \varepsilon_3$, it is hard to directly calculate v_k from the CFR. Fortunately, besides the phase, the v_k manifests itself via the power change of CFR, which can be illustrated via the example in Fig. 2. One can see from the figure that the overall CFR power undergoes a significant decrease after H_d shift to H_d', as evidenced by the shortening of CFR' in comparison with CFR, even if the power remains unchanged. Therefore, we calculate the CFR power and determine the semantic base by analyzing the relationship between power and v_k. Concretely, according to the (2), (3), and (4), we have

$$|H(f,t)|^2 = H_D + H_S + H_{DC}, \quad (5)$$

where

$$H_D = \sum_{l_d \in P_d} 2\,|H_s(f,t)\,a_{l_d}(f,t)|$$
$$\times \cos\left(\frac{f}{c}(2\pi v_k t + 2\pi d_{l_d}(0)) + \varphi_{sl_d}\right) \quad (6)$$

is the cross-term,

$$H_S = \sum_{l_d,k \in P_d, l_d \neq k} 2\,|a_{l_d}(f,t)\,a_k(f,t)|$$
$$\times \cos\left(\frac{f}{c}2\pi((v_k - v_l)t + d_{l_d}(0) - d_k(0)) + \varphi_{l_d k}\right) \quad (7)$$

is the self-term,

$$H_{DC} = \sum_{l_d \in P_d} |a_{l_d}(f,t)|^2 + |H_s(f,t)|^2 \quad (8)$$

is the DC-term, and φ_{sl_d} and $\varphi_{l_d k}$ are the initial phase. From the above derivation, the following conclusions can be drawn.

- The power of CFR consists of three parts. Sorted by the power in a descending order, these components are DC term, cross-term, and self-term. Excepting the DC term,

the rest two terms are composed of multiple sinusoidal waves.

- The frequencies of the sinusoidal waves that make up cross-term and self-term are determined by the v_k, which, as mentioned previously, is directly related to the activity of human body. Moreover, the overall initial phases of these two terms are constants, which are determined by $d_{l_d}(0)$, φ_{sl_d}, $d_k(0)$, and φ_{l_dk}.

From the foregoing derivation and conclusions, it can be seen that the CFR power of different activities can be represented by the sum of a constant component and a set of sinusoidal waves.

Algorithm 1 Semantic-Aware Encoding

Input: The original CFR obtained by sensing the physical world

Output: The amplitude, frequency, and initial phase of semantic bases

1: $\#(Iteration) = 0$, $order = 0$
2: The CFR power calculation $|H(f,t)|_i^2$, $i = 1, \ldots, I$
3: The PCA based denoising and H_{DC} removal
4: Fast Fourier transform based order extraction
5: Initializing A, F, and θ
6: Levenberg-Marquardt (LM) based fitting:
7: $\min\limits_{A,F,\theta} \sum\limits_{i=1}^{I} \left[|H(f,t)|_i^2 - f(A,F,\theta) \right]^2$
8: Fitting error calculation
9: **while** Fitting error $> T_1$ **do**
10:     $order = order + 1$
11:     $\#(Iteration) = \#(Iteration) + 1$
12:     Updating A, F, and θ
13:     LM based fitting
14:     Fitting error updating
15:     **if** $\#(Iteration) > T_2$ **then**
16:         break
17: **return** $[A_r, F_r, \theta_r]$, $r = 1 : order$.

According to these observations, the principal component analysis (PCA) is first employed to denoise the obtained CFR power [2] and eliminate the constant component by subtracting the mean of the power.[3] After that, a series of sinusoidal waves with different amplitudes, frequencies, and initial phases are selected to fit the CFR power to achieve semantic encoding. Here, the selected waves are defined as the semantic bases of the CFR and the corresponding wave parameters are the semantic features. This semantic encoding process can be achieved via the Levenberg-Marquardt (LM) algorithm [35] and we summarize it in Algorithm 1, where $\#(Iteration)$ denotes the number of iterations, $order$ means the number of

---

[2]There are two main reasons for choosing PCA. First, the CSI contains burst noise introduced by internal state transitions. For cancelling such noise, PCA outperforms the traditional filtering-based algorithms [34]. Second, through matrix decomposition and mapping, PCA can convert the CFR matrix (corresponding to the subcarrier and time domain) into vectors, reducing the data dimension, while cancelling the noise and preserving the signal characteristics.

[3]The DC term is a constant that does not contain any activity-related information, so it can be removed by subtracting the mean of the CFR power.

bases required for semantic encoding, A, F, and θ represent the amplitude, frequency, and initial phase of the semantic bases, respectively, T_1 is the fitting error threshold, T_2 is the threshold for the number of iterations, $|H(f,t)|_i^2$ is the ith element of CFR power, $f(A,F,\theta)_i$ is the ith element of fitted value, and I is length of CFR for semantic encoding.

Taking the power of CSI (a sampled version of CFR collected via the Nexmon toolkit [36]) for two common activities, including walking and sitting, as an example, Fig. 3 presents the semantic encoding results. From the original CSI power and encoding results, one can see that walking contains richer components than sitting. Thus, eight semantic bases are used to encode it and the corresponding the semantic features, include the amplitude, frequency, as well as the initial phase, are shown as f1 to f8 in the figures. Relatively speaking, sitting down holds fewer components than walking. Therefore, six semantic bases are sufficient to realize semantic encoding.

So far, the rationality of the semantic bases and the feasibility of semantic encoding are proved by the theoretical derivation and the encoding examples. Clearly, semantic encoding reduces the amount of data that needs to be transmitted each time, while preserving the activity related information, providing more flexibility in increasing data transmission frequency and activity recognition. To translate the gains obtained in semantic encoding into QoS improvement, we define the semantic space, in which the activity recognition is realized without reconstructing the CFR power. Moreover, an incentive mechanism is proposed in Section IV-C to increase the transmission frequency, providing support for boosting the refresh rate for MSP.

B. Wireless Network

Note that the wireless environment between the transmitter and receiver could impact the performance of the system. For example, the bit error probability (BEP) caused by wireless transmission can affect the accuracy of the activity recognition. To investigate this, we consider a generalised network architecture, where the distance between the transmitter and the receiver is D_w, and the large-scale path loss exponent is α_w. Considering the small-scale shadow fading and multipath fading, a generalized small-scale fading model is used to characterize the fluctuations in the amplitude of the transmitted signal. Since the wireless environment is dynamic and volatile, we use H-fading channel model [37], which includes most of typical models such as Rayleigh, Nakagami-m, Weibull, α-μ, N*Nakagami-m, generalized K-fading, and Weibull/gamma fading as its special cases. Specifically, the H-fading can be converted to the various commonly used small-scale fading models as [38, Table 1]. The PDF of H-fading is given as [38]

$$f_X(x) = \kappa H_{p,q}^{m,n} \left[\lambda x \left| \begin{array}{c} (a_i, A_i)_{i=1:p} \\ (b_l, B_l)_{l=1:q} \end{array} \right. \right], \tag{9}$$

where $H_{p,q}^{m,n}(\cdot)$ is the Fox's H-function [39, eq. (1.2)], $\lambda > 0$ and κ are the constants and satisfy $\int_0^\infty f_X(x)\,dx = 1$, $(a_i, A_i)_{i=1:p}$ means $(a_i, A_i)_{i=1:p} = (a_1, A_1), \ldots, (a_p, A_p)$. Assuming that maximum-ratio combining reception in the nth transmitter [40], the output SNR at the receiver, i.e., Z_n, can
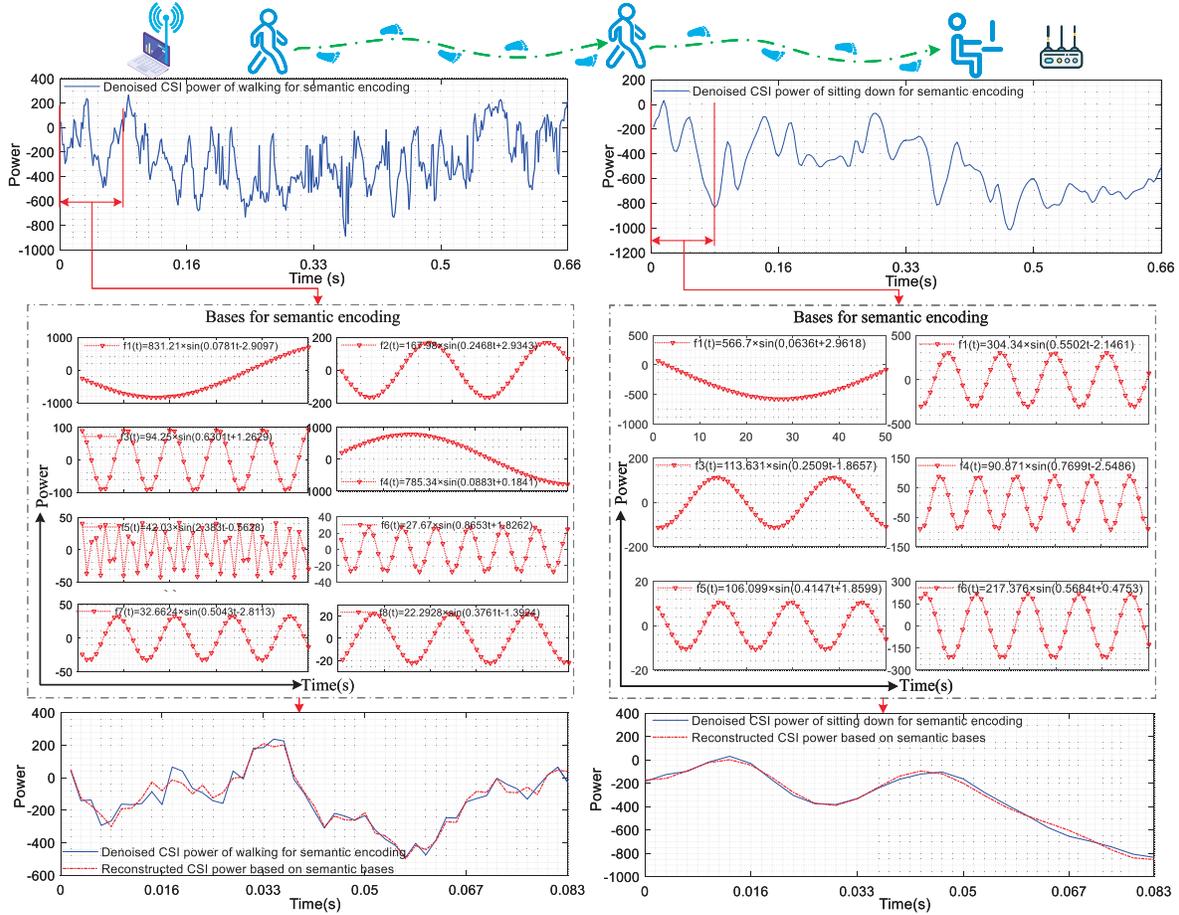
Fig. 3. Example of semantic encoding for walking and sitting.

be expressed as the sum of the individual branches [41]. Thus, we have $Z_n = X_1 + \cdots + X_{N_R}$, where X_j $(j = 1, \ldots, N_R)$ denotes the SNR between the transmitter and the jth antenna in the receiver. The probability density function (PDF) and cumulative distribution function (CDF) of Z_n are derived as in [41, eq. (8)] and [41, eq. (9)], respectively.

1) Data Rate Evaluation: The ergodic data rate (or Shannon capacity) for the nth transmitter, which is known to be the maximum data rate that the channel can support, is defined as

$$C_{\text{n}} = B_n \int_0^\infty \log_2(1 + \gamma) f_{Z_n}(\gamma) d\gamma, \qquad (10)$$

where B_n is the available channel bandwidth of the nth transmitter. By substituting [41, eq. (8)] into (10), the data rate can be derived as [41, eq. (28)].

2) Bit Error Rate (BER) Evaluation: The average BEP for the nth transmitter under a variety of modulation formats is given by [42]

$$E_{\text{n}} = \int_0^\infty \frac{\Gamma(\tau_2, \tau_1\gamma)}{2\Gamma(\tau_2)} f_{Z_n}(\gamma) d\gamma, \qquad (11)$$

where τ_1 and τ_2 are modulation-specific parameters for several modulation and detection scheme combinations, respectively, $\Gamma(\tau_2, \tau_1\gamma)/2\Gamma(\tau_2)$ represents the conditional bit-error probability, different modulation and detection scheme combinations

TABLE I
DIFFERENT MODULATION AND DETECTION SCHEME COMBINATIONS
WITH CORRESPONDING PARAMETERS τ_1 AND τ_2

τ_1	τ_2	**Modulation & Detection Scheme**
0.5	0.5	Orthogonal coherent binary frequency-shift keying (BFSK) scheme
1	0.5	Antipodal coherent binary phase-shift keying (BPSK) scheme
0.5	1	Orthogonal non-coherent BFSK scheme
1	1	Antipodal differentially coherent BPSK (DPSK) scheme

with corresponding parameters τ_1 and τ_2 are shown in Table I, and $\Gamma(\cdot, \cdot)$ denotes the upper incomplete Gamma function [43, eq. (8.350.2)]. Substituting [41, eq. (8)] into (11), we can obtain the close-form of BER as [41, eq. (25)]. With the modeling ability of the H fading model for various channel environments, we analyze the impact of bit errors on semantic information transmission in Section V.

Assuming that the number of bits generated by wireless sensing for each time is D, then the time required for the nth transmitter to send the sensing data to receiver is D/C_n, and the average number of error bits is $E_{\text{n}}D$. The amount of
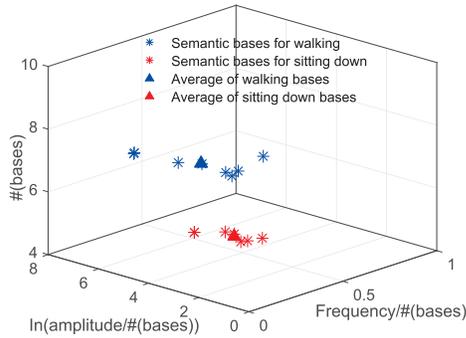
Fig. 4.   Distribution of semantic features in defined 3D semantic space.

raw data generated by wireless sensing is relatively large as discussed above. Fortunately, the proposed semantic encoding greatly reduces the amount of data, enabling it to sense and upload data to complete recognition multiple times per second, which improves synchronization between the physical world and Metaverse.

C. Semantic-Aware Activity Recognition

After receiving the signal, the receiver conducts channel decoding to obtain the semantic features, including the amplitude, frequency, and initial phase of the sine waves, through which the original CFR power can be recovered. Leveraging different signal processing and feature extraction algorithms, coupled with statistical analysis and machine learning methods, many systems can realize activity recognition based on the CFR power [44]. Unlike them, in this work transmitter sends semantic features related to activity to the receiver. Therefore, we define the three dimensional (3D) semantic space, in which the activity recognition is realized by directly classifying the received semantic features, without recovering CFR power. Specifically, the three dimensions of the defined semantic space are

$$
\begin{cases}
x = frequency/\#(bases) \\
y = \mathrm{In}(amplitude/\#(bases)) \\
z = \#(bases),
\end{cases}
\tag{12}
$$

where $\#(bases)$ is the number of semantic bases employed for semantic encoding.[4] Taking the signal in Fig. 3 as an example, we map the semantic features into 3D semantic space and obtain the averages corresponding to both activities, the results are shown in Fig. 4. It is clear that the bases of different activities are distributed in different parts of the 3D semantic space, creating favorable conditions for realizing recognition in the semantic space. This can be interpreted by the fact that different activities trigger different numbers of reflections with varying power and frequency. We collect more sets of data for falling, walking, and sitting down in real-world scenarios and map bases into the 3D semantic space. The distribution of semantic features and corresponding averages, in 3D semantic

space, are shown in Fig. 5 and Fig. 6, respectively, which further verifies the feasibility of achieving activity recognition in the 3D semantic space.

According to this observation, the kNN classifier is applied to classify semantic features in the semantic space. This process consists of two steps: offline training and online recognition. During the offline phase, first, multiple sets of training data corresponding to each activity are collected to extract semantic features. Then, the features are mapped into the semantic space to obtain multiple averages for each activity. Here, the obtained averages are denoted as

$$
\mathbf{T} = \{T_1, \ldots, T_m, \ldots, T_M\},
\tag{13}
$$

where $T_m = (x_m, y_m, z_m, l_m)$, x_m, y_m, and z_m are the coordinates of averages in semantic space, $l_m \in \{A_1, \ldots, A_K\}$ is the label. In this way, we can construct a cluster containing multiple labeled averages for each activity in 3D semantic space. In the online classification process, the test average of the to-be-classified data in the 3D semantic space is first obtained. After that, K training averages closest to the test average in the labeled averages are identified and the neighborhood covering the K training averages is determined as N_k. Finally, the activity recognition can be realized via

$$
l = \arg\max_{A_k} \sum_{T_m \in N_k} \mathbb{I}(l_m = A_k),
\tag{14}
$$

where $\mathbb{I}(\cdot)$ is the indicator function and $k = 1, 2, \cdots, K$.

Considering the limited sensing capacity of a single transmission link and the possibility of numerous smart devices participating in physical world sensing, we combine the recognition results from multiple links via voting to further boost the performance. As Fig. 7 illustrates, each link collects and processes the sensing data independently and transmits it to receiver for recognition and voting. Thereafter, the one with the largest number of votes is selected as the final recognition result[5] and fed to MSP for further processing. The activity recognition process in 3D semantic space is summarized in Algorithm 2.

D. Contest Theory-Based Sensing Data Market

To further enhance the user experience in Metaverse, we design an effective incentive mechanism for the sensing data market to promote transmitters uploading data more frequently. Concretely, under an available bonus amount, our goal is not to promote a specific transmitter to achieve the best possible performance or allocate network resource fairly, but to encourage all transmitters to redouble their efforts. Thus, unlike auction theory, contract theory and game theory, which are already widely used in wireless communication networks, a suitable solution here is to use the bonus to host a contest among transmitters. Here, the contest is defined as a game, in which contestants, i.e., the transmitters, must

---

[4]According to previous derivation, one can see that the initial phase of H_D are determined by $d_{l_d}(0)$ and φ_{sl_d}, which contain no activity-related information. Therefore, the initial phase is not considered in the construction of the semantic space.

[5]Due to transmission noise, packet loss or other causes, multiple activities may receive the same number of votes. For this case, we have two solutions. First, among the activities with the same number of votes, if one of them is the same as the previous recognition result, then this one is picked as the output. Otherwise, We single out the final result by analyzing the distance between the test data and T_m in semantic space.
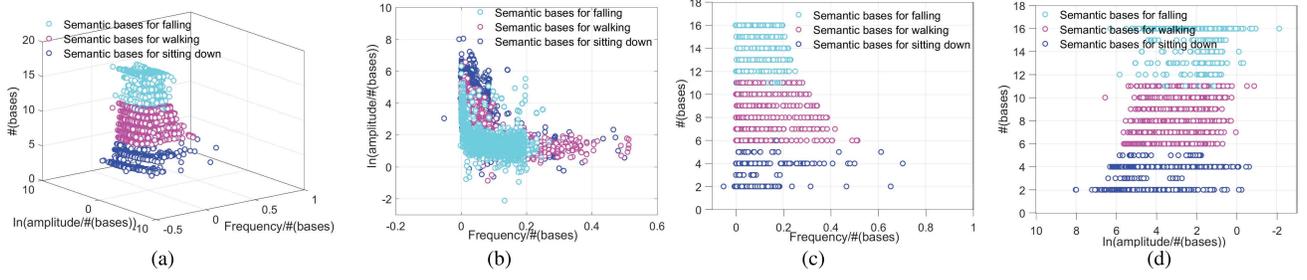
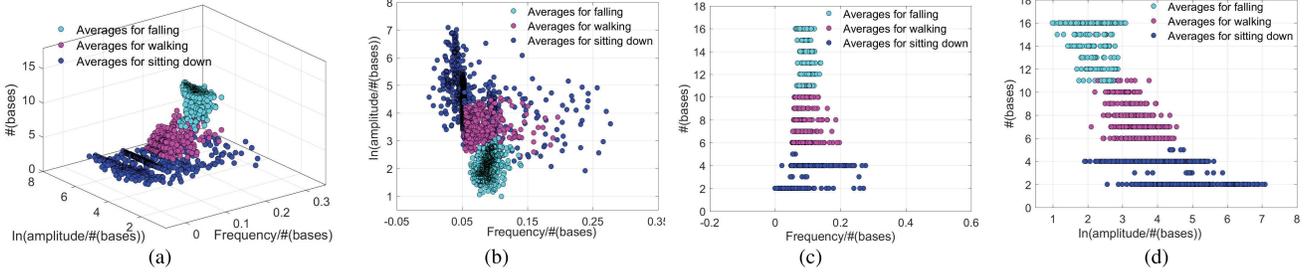Fig. 5. Distribution of features in the 3D semantic space.

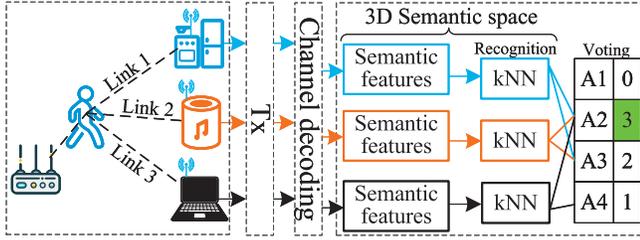Fig. 6. Distribution of features averages in the 3D semantic space.

Fig. 7. Fusion of recognition results from multiple links.

Algorithm 2 Activity Recognition in Semantic Space

Input: The labeled training data, test data, K

Output: The activity recognition result

1: **for** each transmission link **do**

2:     **for** $m = 1{:}M$ **do**

3:         Calculating the distance between the test data and T_m in semantic space

4:     Sort the obtained distances in an ascending order

5:     Determine N_k based on K and sorted distances

6:     For each link do: $l = \arg\max_{A_k} \sum_{T_m \in N_k} \mathbb{I}(l_m = A_k)$

7: Vote to obtain the final recognition result

---

exert irreversible effort to win awards, which are determined according to the quality of the outputs. In this section, we use the contest theory to examine the payoff of transmitters and the payment of receiver [45].

1) Effort, Capability, and Awards: In the context of wireless sensing in Metaverse, contestants are transmitters and their efforts can be regarded as the frequencies of uploading sensing data, i.e., let f_n denote the frequency that the nth transmitter sends the sensing data to the receiver for activity recognition

in a time period T. Thus, the cost function for the nth transmitter can be expressed as a twice differentiable function, i.e., $C_n(a_n, f_n)$, that satisfies [45]

$$\begin{cases} \dfrac{\partial C_n(a_n, f_n)}{\partial f_n} > 0, \\[2mm] \dfrac{\partial C_n(a_n, f_n)}{\partial a_n} < 0, \\[2mm] \dfrac{\partial^2 C_n(a_n, f_n)}{\partial a_n \partial f_n} < 0, \end{cases} \tag{15}$$

where a_n denotes the capability of the nth transmitter. The last inequality in (15) implies that the more capable a contestant is, the easier it is to increase the output.

The capability of transmitter can be measured in terms of the time that it takes to perform each upload. Specifically, a more capable transmitter can achieve faster data uploads because of better channel conditions and faster processing speed. Let T_P and $T_P^{(S)}$ denote the time required for receiver to perform activity recognition using the original sensing data and semantic bases, respectively, and T_S represents the time required to obtain the semantic bases from the sensing data, and a_n can be expressed as

$$a_n = \left(\frac{\mathcal{S}(D)}{C_n} + T_S + T_P^{(S)} \right)^{-1} = \left(\frac{D_S}{C_n} + T_S + T_P^{(S)} \right)^{-1}, \tag{16}$$

where S is the semantic-aware encoding operator, and $S(D)$ is the bit amount of semantic bases for one wireless sensing.

For a given upload frequency, a more capable transmitter can upload faster, resulting in a smaller cost. Taking (15) into consideration, the cost function can be denoted as $C_n(a_n, f_n) = f_n/a_n$. Sorting N_T transmitters in descending order according to efforts to obtain $\{f_1, \ldots, f_{N_T}\}$, then the

award that the nth transmitter receives can be denoted as r_n ($1 \leq n \leq N_A \leq N_T$), where N_A is the number of awards, $r_1 \geq r_2 \geq \cdots \geq r_{N_A}$, and $r_n = 0$ when $n > N_A$. On this basis, all the awards are summed up to obtain the total amount that receiver pays, which is expressed as $r_T = \sum_{n=1}^{N_A} r_n$.

2) Optimal Effort Analysis: The utility of the nth transmitter participating in the contest can be expressed as

$$\pi(a_n, f_n)$$
$$= \begin{cases} r_m - \mathcal{C}_n(a_n, f_n), & \text{if the transmitter is ranked in } m^{\text{th}}, \\ -\mathcal{C}_n(a_n, f_n), & \text{if the transmitter is not awarded.} \end{cases}$$
$$(17)$$

Then, the expectation of the utility can be derived as follows:

$$\mathbb{E}[\pi(a_n, f_n)] = \mathcal{R}(f_n) - \mathcal{C}_n(a_n, f_n), \tag{18}$$

where $\mathcal{R}(f_n)$ denotes the expectation of the award that the nth transmitter receives. Considering that all transmitters are rational, the following lemma is first derived

Lemma 1: The more capable the transmitter is, the more effort is exerted in the contest. In other words, a transmitter with a higher transmission rate and a shorter activity recognition processing time uploads sensor data more frequently.

Proof: Please refer to Appendix A. ∎

We consider that the CDF of capability in the population is represented by a continuous function $\mathcal{G}(a)$, and each transmitter knows its own capability, but has only probabilistic estimates of the capabilities of fellow potential contestants, i.e., other transmitters. Assuming the transmitter with capability a_n chooses effort f_n and its remuneration consists solely of its prize, so that a transmitter with capability a_n chooses to participate in the contest only if its surplus, $\mathcal{R}(a_n, y(a_n))$, is no less than the opportunity cost of participation.

According to Lemma 1, a transmitter with greater capability will choose a higher level of effort. Hence, the probability that a transmitter with capability a obtains the jth highest prize is simply the probability that it has the jth highest capability among the N_T potential transmitters. On this basis, the expected award won by the nth transmitter with capability a_n is

$$\mathcal{R}(a_n)$$
$$= \sum_{m=1}^{N_A} u(r_m) \binom{N_T - 1}{m - 1} \mathcal{G}^{N_T - m}(a_n)(1 - \mathcal{G}(a_n))^{m-1}, \tag{19}$$

where $u(r_m)$ is the expected utility of the mth award decided by the risk preference [46]. Specifically, $u(r_m) = r_m$ denotes risk neutral and $u(r_m) = \ln(r_m)$ means risk averse.

A transmitter does not know the capabilities of other transmitters, but it will estimate the distribution of the capabilities,

i.e., $\mathcal{G}_X(x)$. Recall that capability is a function of data rate, semantic encoding time, and activity recognition time, here the transmitter will estimate its expected utility in the contest by assuming that the data rate of other transmitters obeys a uniform distribution.[6] To obtain a close-form expression of $\mathcal{R}(a_n)$, the $\mathcal{G}(a_n)$ is analyzed first.

Lemma 2: From the nth transmitter's point of view, the probability that the capabilities of other transmitters is less than a_n, i.e., $\mathcal{G}(a_n)$, is

$$\mathcal{G}(a_n) = \begin{cases} \dfrac{1}{\Delta} \dfrac{a_n D_S}{1 - a_n T_S - a_n T_P^{(S)}}, & \text{Condition 1} \\ 0, & \text{otherwise,} \end{cases} \tag{20}$$

where Condition 1 is

$$0 < a_n < \left(\frac{D_S}{\Delta} + T_S + T_P^{(S)}\right)^{-1}. \tag{21}$$

Proof: Please refer to Appendix B. ∎

With the Lemma 2, the optimal effort for each transmitter is obtained as in Theorem 1.

Theorem 1: The optimal effort of each transmitter is given as

$$f_n^*$$
$$= a_n R(a_n) - \sum_{m=1}^{N_A} \frac{u(r_m) a_n H_f}{\Gamma(N_T - 1)} \binom{N_T - 1}{m - 1} \left(\frac{D_S}{\Delta}\right)^{N_T - m}, \tag{22}$$

where H_f is derived as (23), shown at the bottom of the page.

Proof: Please refer to Appendix C. ∎

3) Optimal Awards Analysis: The receiver adjusts the awards to maximize the sum of transmitter efforts by solving

$$\max_{N_A, r_1, \ldots, r_{N_A}} \sum_{n=1}^{N_T} f_n^*$$
$$\text{s.t.} \sum_{m=1}^{N_A} r_m \leqslant r. \tag{24}$$

With the help of Theorem 1, we can rewrite f_n^* as

$$f_n^* = \sum_{m=1}^{N_A} u(r_m) \mathcal{F}_n\left(a_n, N_T, m, D_S, T_S, T_P^{(S)}, \Delta\right). \tag{25}$$

Theorem 2: By solving the optimization problem (24), we can obtain the optimal award setting as $r_1 = r$ and $r_2 = \cdots = r_{N_T} = 0$, when the transmitters are risk neutral.

Proof: Please refer to Appendix D ∎

---

[6]Here the uniform distribution is considered and, in fact, the proposed analysis method can be easily extended to other distributions.

$$H_f = H_{1,1:3,3;1,1}^{0,1:1,2;1,1}\left(\begin{array}{c} a_n\left(T_S + T_P^{(S)} + \dfrac{D_S}{\Delta}\right) \\ \dfrac{\Delta T_S + \Delta T_P^{(S)}}{-\Delta} a_n \end{array}\middle|\begin{array}{c} (m - N_T; 1, 1) : (0, 1)(m - N_T - 1; 1)(m, 1); (2 - N_T, 1) \\ (m - N_T - 1; 1, 1) : (0, 1)(-1, 1)(m - N_T, 1); (0, 1) \end{array}\right) \tag{23}$$
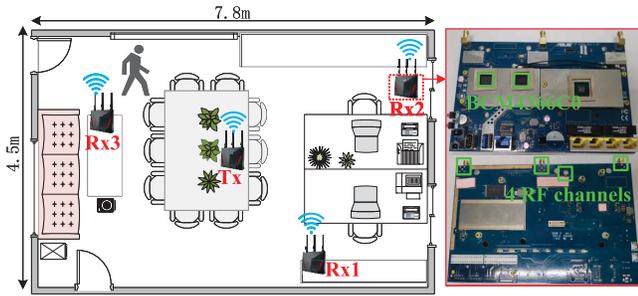
Fig. 8. Test scenario and hardware equipment.

Theorem 3: If the transmitters are risk averse, we can obtain the optimal award setting as follows:

$$r_\ell = \frac{r \sum_{n=1}^{N_T} \mathcal{F}\left(a_n, N_T, \ell, D_S, T_S, T_P^{(S)}, \Delta\right)}{\sum_{m=1}^{N_A} \sum_{n=1}^{N_T} \mathcal{F}\left(a_n, N_T, m, D_S, T_S, T_P^{(S)}, \Delta\right)}, \quad (26)$$

where $\ell = 1, \ldots, N_A$.

Proof: Please refer to Appendix E. ■

V. NUMERICAL ANALYSIS

In this section, we first introduce the experimental configurations. Then, based on the collected data, the effectiveness of activity recognition in semantic space is verified. After that, the impact of wireless transmission on recognition accuracy is simulated. Finally, the effect of semantic transmission and incentive mechanisms on the overall system is investigated.

A. Experimental Platform, Scenario, and Data Collection

An IEEE 802.11ac protocol based experimental platform is built to collect CSI in real-world scenarios to validate the activity recognition in semantic space. The specific hardware and test scenario are shown in Fig. 8. The test scenario is a 7.8m × 4.5m office room with different furniture, such as conference table, chairs, computers, and bookcases. Inside the office room, we placed a transmitter (denoted as Tx) and 3 receivers (denoted as Rx1, Rx2, and Rx3, respectively) to transmit and receive wireless signals, respectively. The experimental equipment is an access point (AP) equipped with a BCM4366C0 chip and the Nexmon toolbox [36], as shown on the right side of Fig. 8. During the data collection process, the frequency of the signal is set to 5.805 GHz, the signal bandwidth is 80 MHz, and the transmission rate of the CSI data packet is 600 Hz. While each AP has multiple antennas, the Tx and Rx are controlled to operate via one antenna.

B. Performance Evaluation

1) The Activity Recognition in Semantic Space: We collected measurement data of four activities and evaluated the recognition performance in semantic space. To demonstrate the effectiveness of our framework, we compare it with CARM [34] and the results are shown in Fig. 9. First, the Fig. 9 (a) shows the performance of the both systems improves

with the increase of the sampling rate and the general trend is nearly identical, as presented by the partially enlarged result. The reason is that the reduce of packet transmission rate makes the obtained original CSI unable to fully describe the high-frequency features, affecting the recognition performance. Second, the recognition accuracy of the proposed system is close to that of CARM. As an example, Fig. 9 (b) indicates that for walking, the maximum recognition accuracy difference between two systems is 0.0365, and the minimum difference is 0.0057. For other activities, the difference in recognition accuracy is smaller. The results reveal that the performance of activity recognition in semantic space is comparable to that of CARM, verifying the effectiveness of semantic encoding, as well as activity recognition in semantic space.

2) The Impact of Wireless Transmission on Recognition Accuracy: We simulated the wireless transmission of the semantic feature to analyze the impact of transmission on activity recognition. Figure 10 (a) gives the activity recognition accuracy in semantic space under different small-scale fading conditions obtained using the H-fading model in Section IV-B. One can see that the BER of Rayleigh is larger than the Nakagami for different transmit power. Meanwhile, for Nakagami, increasing the order can reduce the BER. Therefore, as denoted by black lines, the Nakagami-10 has the highest recognition accuracy, followed by Nakagami-5, then Nakagami-2, and the lowest is Rayleigh. The reason is that smaller order in the Nakagami fading means stronger multipath effect. In addition, increasing the transmit power can reduce the BER, thereby improving the recognition accuracy. When the transmit power reaches 25 dBw, the recognition accuracy of Nakagami-10 is 0.8831, which is comparable to 0.8953 of recognition in semantic space without transmission and 0.9010 of CARM, further demonstrating the feasibility of semantic encoding and activity recognition in semantic space.

Besides small-scale fading, the impact of modulation methods on the recognition accuracy is analyzed under the Rayleigh fading condition, and the results are shown in Fig. 10 (b). With the same transmit power, the BPSK has the lowest BER, then BFSK, followed by DPSK, and orthogonal non-coherent BFSK (ON- BFSK) scheme has the largest BER. As a result, BPSK has the best activity recognition performance, while ON-BFSK performs the worst. Meanwhile, similar to the result in Fig. 10 (a), the recognition accuracy improves with the increase of transmit power. When transmit power arrives at 25 dBw, the recognition accuracies corresponding to ON-BFSK and BPSK are 0.8560 and 0.8618, respectively, which are believed to be acceptable.

3) Data Amount Comparison Before and After Semantic Encoding: Let the transmitter upload data 12 times per second,[7] then the data amount of the original CSI, CSI power, and semantic encoded data is illustrated in Fig. 10. (c), where, for example, S-FL refers to the semantic encoding of falling data. As can be seen, the average amount of data uploaded each time after semantic encoding is approximately 27.87%

---

[7] According to [47], the channel coherence time in an indoor environment is around 84 ms and $600 \times 0.084 \approx 50$. Therefore, with a packet transmission rate of 600 Hz, we can assume that the transmitter uploads data 12 times per second, i.e., performing data upload every 50 data packets are collected.
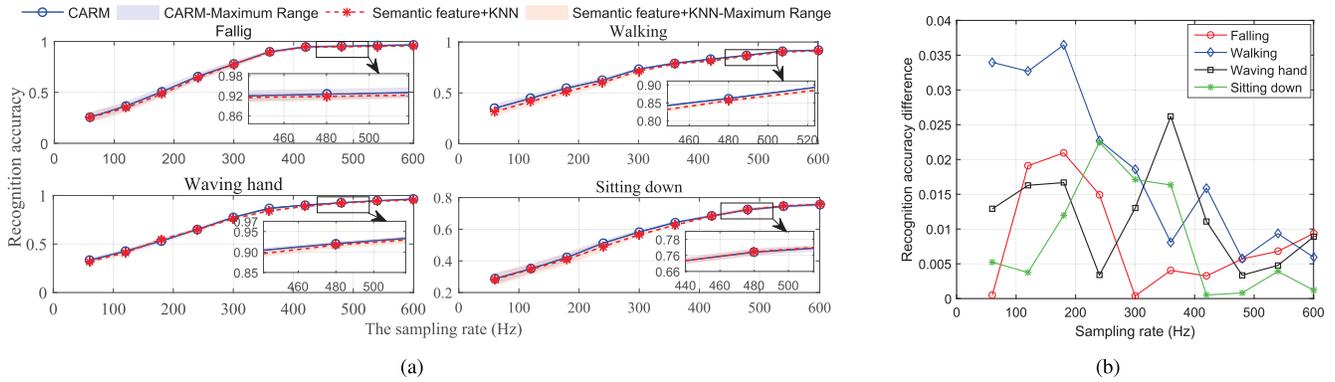
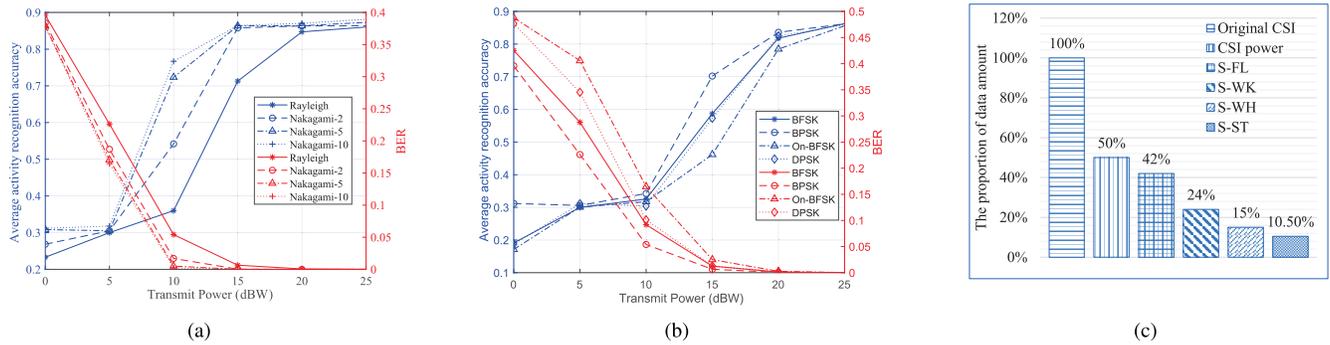Fig. 9. Activity recognition accuracy in the 3D semantic space.

Fig. 10. Impact of wireless transmission on recognition accuracy and the data amount comparison before and after semantic encoding.

of that before encoding, demonstrating the proposed semantic encoding algorithm can greatly reduce the amount of data while ensuring the system performance. This creates excellent space for further establishing the incentive mechanism to promote data uploading. Besides that, the amount of encoded data corresponding to different activities is different. This is understandable, since the part of the body which caused main reflections is different for various activities. For instance, during walking the arms and legs reflect a large number of signals in addition to the torso, while for sitting down the reflections are mainly caused by torso. Therefore, different numbers of semantic bases are needed to encode data corresponding to various activities, leading to differences in the amount of encoded data.

4) The Effects of the Semantic Method and the Award Setting Schemes: We consider the case where there are three risk neutral transmitters and two awards are set in the contest. For a comprehensive analysis, six award setting schemes are compared, i.e., the ratios of the first award to total award are $0.5, \ldots, 1$. Besides that, some other parameter configurations are shown in Table II, where the processing time and the number of bits are obtained from real-world experiments, and the data rate values are typical in wireless networks [48]. Figure 11 (a)-(c) shows the efforts of three transmitters when the semantic encoding is used or not. From the results, one can observe that the use of semantic encoding motivates transmitters to upload sensing data under all award schemes in the contest. The reason is that the data amount of semantic

TABLE II
NETWORK PARAMETERS FOR 3 TRANSMITTERS

Parameters	1_{st} trans-mitter	2_{nd} trans-mitter	3_{rd} trans-mitter
Time to obtain the semantic bases, T_S	9.797 ms		
Time to perform activity recognition, M_U	5 ms		
Estimated maximum data rate, Δ	8 Mbps		
Real data rate, C_n	7 Mbps	6 Mbps	5 Mbps
Number of bits generated in once sensing, D	96000		
Number of bits of semantic bases in once sensing, D_S	7200		

features is much smaller than that of the raw sensing data, making the transmitter using semantic encoding has lower cost and higher capability. Therefore, under the same expected award, each transmitter is willing to put in more effort to win the contest. Another interesting insight from Fig. 11 (d) is that the sum of three transmitters' upload frequencies achieves the maximum when the total award is set as the only first prize, demonstrating the Theorem 2. Compared with the uniform allocation scheme, the optimal award allocation strategy can increase the sum of uploading frequency by 27.47%.
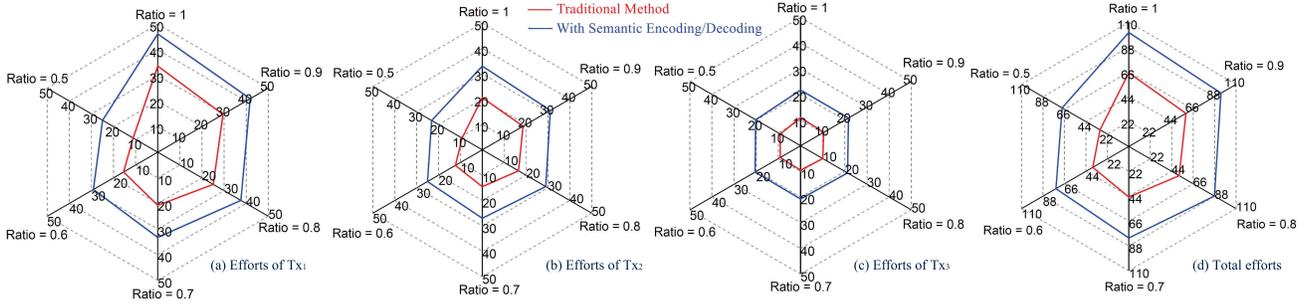
Fig. 11. Efforts of three risk neutral transmitters when the semantic methods are used or not, under six different award setting schemes.
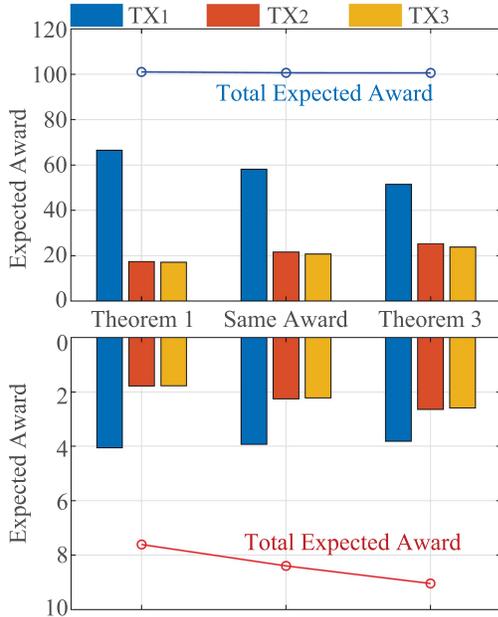

Fig. 12. Expected awards of transmitters under different award scheme settings and risk appetites.

5) The Effects of the Risk Appetite and the Award Setting: Since the optimal award scheme is related to the transmitters' risk appetite, here we further discuss the case where transmitter is risk averse. From (18), one can observe that the greater the expected award, the more willing transmitters are to put effort in the contest. In Fig. 12, we show the awards that transmitters expect to receive under different award scheme settings and risk appetites. Again, there are three transmitters and three awards. The parameters related to transmitters are shown in Table II. Three different award setting schemes are considered, i.e., all awards were allocated to the first prize (Theorem 2), and the awards are divided equally by three prizes, and awards are allocated following Theorem 3. When transmitters are risk neutral, we can see that reducing the proportion of first prize awards results in a significant reduction, i.e., 22%, in the expected award of the most able transmitter. Although the expected awards for the remaining participants increase, the total expected award for the three transmitters decreases slightly. When transmitters are risk averse, our derived optimal award setting scheme, i.e., Theorem 3, brings only a slight decrease in the expected award

for the most capable transmitter, i.e., 7%, compared to the optimal award setting scheme when the user is risk neutral. Meanwhile, the expected awards of the rest of the participants increase obviously. Finally, the total expected award can be increased by 20%.

VI. CONCLUSION

A semantic aware transmission framework has been proposed in this paper, for transforming and transmitting sensing data from the physical world to MSP in Metaverse. Two breakthroughs lie in paper are semantically aware sensing data transmission and the contest theory based incentive mechanism, respectively. We have proposed the first semantic encoding algorithm for sensing data, reducing the amount of data significantly and lowering the storage and transmission costs, while ensuring the performance of MSP in Metaverse. In addition, we have established a contest-based incentive mechanism to boost all transmitters' data uploading frequency by setting rewards, which provides stronger support for MSP to enhance its QoS. The results show that the average data amount after semantic encoding is only 27.87% of the original sensing data and the data uploading frequency is increased by 27.47% with the help of proposed incentive mechanism. For future work will be extended to semantic aware transmission of other sensing information.

APPENDIX A
PROOF OF LEMMA 1

The transmitter maximizes the expectation of utility by adjusting its effort. Thus, the optimal effort of the nth transmitter, i.e., f_n^*, satisfies the following conditions:

$$\begin{cases} \dfrac{\partial \mathbb{E}\left[\pi\left(a_n, f_n\right)\right]}{\partial f_n}\bigg|_{f_n=f_n^*} = 0 \\ \dfrac{\partial^2 \mathbb{E}\left[\pi\left(a_n, f_n\right)\right]}{\partial f_n^2}\bigg|_{f_n=f_n^*} < 0 \end{cases} \quad \text{(A-1)}$$

Substituting (18) and (19) into (A-1), we have $R\left(f_n\right)$ satisfies

$$\frac{\partial R\left(f_n^*\right)}{\partial f_n^*} = \frac{\partial C_n\left(a_n, f_n^*\right)}{\partial f_n^*} = \frac{1}{a_n} \quad \text{(A-2)}$$

and

$$\frac{\partial^2 R\left(f_n^*\right)}{\partial f_n^{*2}} < 0. \quad \text{(A-3)}$$

Since f_n^* is a function of a_n and r_n, we take the partial derivative of (A-2) with respect to f_n^* to obtain $\frac{\partial^2 R(f_n^*)}{\partial f_n^{*2}} = \frac{-1}{a_n^2}\frac{\partial a_n}{\partial f_n^*}$. With the help of (A-3), we have $\frac{\partial f_n^*}{\partial a_n} > 0$, which means that, under the optimal effort strategy, the more capable transmitter exerts more effort. The proof is completed.

APPENDIX B
PROOF OF LEMMA 2

The CDF of C_n can be expressed as

$$F_{Cn}(x) = \begin{cases} \frac{x}{\Delta}, (\Delta > x > 0) \\ 0. \end{cases} \quad \text{(B-1)}$$

According to (16), we have $C_n = \frac{D_S}{\frac{1}{a_n}-T_S-T_P^{(S)}}$. Thus, the CDF of a_n can be expressed as $F_{a_n}(y) = F_{Cn}\left(\frac{yD_S}{1-yT_S-yT_P^{(S)}}\right)$. Substituting (B-1) into $F_{a_n}(y)$, we can obtain (20) to complete the proof.

APPENDIX C
PROOF OF THEOREM 1

Taking the derivative of $\partial R(a_n)$ with respect to a_n, we obtain $\frac{\partial R(a_n)}{\partial a_n} = \frac{\partial R(a_n)}{\partial f_n^*}\frac{\partial f_n^*}{\partial a_n} = \frac{1}{a_n}\frac{\partial f_n^*}{\partial a_n}$. Therefore, the optimal effort can be expressed as

$$f_n^* = a_n R(a_n) - \underbrace{\int_0^{a_n} R(y)\,\mathrm{d}y}_{I}. \quad \text{(C-1)}$$

Substituting (19) and (20) into I, we have

$$I = \sum_{m=1}^{N_A} \frac{u(r_m)}{\Delta^{N_T-1}}\binom{N_T-1}{m-1}$$
$$\times \int_0^{a_n} \frac{(yD_S)^{N_T-m}\left(\Delta-\left(\Delta T_S+\Delta T_P^{(S)}+D_S\right)y\right)^{m-1}}{\left(1-\frac{\Delta T_S+\Delta T_P^{(S)}}{\Delta}y\right)^{N_T-1}}\mathrm{d}y.$$
$$\text{(C-2)}$$

Note that the integral in (C-2) is difficult to solve directly. To obtain the close-form solution of I, we perform an H-transformation [49, eq. (2.9.6)] of the numerator as follows:

$$(yD_S)^{N_T-m}\left(\Delta-\left(\Delta T_S+\Delta T_P^{(S)}+D_S\right)y\right)^{m-1}$$
$$= D_S^{N_T-m}\Delta^{m-1}\left(T_S+T_P^{(S)}+\frac{D_S}{\Delta}\right)^{m-N_T}$$
$$\times H_{1,1}^{1,0}\left(\left(T_S+T_P^{(S)}+\frac{D_S}{\Delta}\right)y\bigg|\begin{matrix}(N_T,1)\\(N_T-m,1)\end{matrix}\right)$$
$$= \frac{\Delta^{m-1}}{2\pi i(yD_S)^{m-N_T}}$$
$$\times \int_{\mathcal{L}}\frac{\Gamma(-t)}{\Gamma(m-t)}\left(\left(T_S+T_P^{(S)}+\frac{D_S}{\Delta}\right)y\right)^t\mathrm{d}t, \quad \text{(C-3)}$$

where $i = \sqrt{-1}$, $\mathcal{L}$ is the right loop integral path starting at the point $-0.1i + \infty$, terminating at the point $0.1i + \infty$,

and intersecting the x-axis at $(x_s,0)$, where $0 > x_s > -1$. Substituting (C-3) into I, we have

$$I = \sum_{m=1}^{N_A} u(r_m)\binom{N_T-1}{m-1}\left(\frac{D_S}{\Delta}\right)^{N_T-m}\frac{1}{2\pi i}$$
$$\times \int_{\mathcal{L}}\frac{\Gamma(-t)}{\Gamma(m-t)}\left(\left(T_S+T_P^{(S)}+\frac{D_S}{\Delta}\right)\right)^t I_2\mathrm{d}t, \quad \text{(C-4)}$$

where $I_2 = \int_0^{a_n}\frac{y^{N_T-m+t}}{\left(1-\frac{\Delta T_S+\Delta T_P^{(S)}}{\Delta}y\right)^{N_T-1}}\mathrm{d}y$. With the help of [43, eq. (3.194.1)] and [43, eq. (9.113)], I_2 can be solved. Substituting I_2 into I, we obtain (22) to finish the proof. The Fox's H-function in (22) can be calculated efficiently using a Python implementation [50].

APPENDIX D
PROOF OF THEOREM 2

The Lagrangian function associated with problem (24) is given by

$$F_L = -\sum_{n=1}^{N_T}\sum_{m=1}^{N_A} r_m \mathcal{F}\left(a_n, N_T, m, D_S, T_S, T_P^{(S)}, \Delta\right)$$
$$-\lambda\left(\sum_{m=1}^{N_A} r_m - r\right), \quad \text{(D-1)}$$

where λ is the Lagrange multipliers. The Karush–Kuhn–Tucker (KKT) optimality conditions for the optimal solution are

$$\begin{cases} \sum_{m=1}^{N_A} r_m = r \\ \sum_{n=1}^{N_T}\mathcal{F}\left(a_n, N_T, m, D_S, T_S, T_P^{(S)}, \Delta\right) = \lambda, \forall m. \end{cases} \quad \text{(D-2)}$$

Note that if $m_1 \neq m_2$, we have

$$\sum_{n=1}^{N_T}\mathcal{F}\left(a_n, N_T, m_1, D_S, T_S, T_P^{(S)}, \Delta\right)$$
$$\neq \sum_{n=1}^{N_T}\mathcal{F}\left(a_n, N_T, m_2, D_S, T_S, T_P^{(S)}, \Delta\right). \quad \text{(D-3)}$$

Thus, the second condition in (D-2) is only correct when $N_A = 1$. Then, using the first condition, we obtain that $r_1 = r$, which completes the proof.

APPENDIX E
PROOF OF THEOREM 3

When the transmitters are risk averse, the Lagrangian function can be expressed as

$$F_L = -\sum_{n=1}^{N_T}\sum_{m=1}^{N_A}\ln r_m \mathcal{F}\left(a_n, N_T, m, D_S, T_S, T_P^{(S)}, \Delta\right)$$
$$-\lambda\left(\sum_{m=1}^{N_A} r_m - r\right). \quad \text{(E-1)}$$

The KKT conditions then can be written as

$$\begin{cases} \sum_{n=1}^{N_T} \dfrac{1}{r_m} \mathcal{F}\Big(a_n, N_T, m, D_S, T_S, T_P^{(S)}, \Delta\Big) = \lambda, \forall m \\ \sum_{m=1}^{N_A} r_m = r. \end{cases} \quad \text{(E-2)}$$

By solving (E-2), we have

$$\frac{1}{\lambda} \sum_{n=1}^{N_T} \mathcal{F}\left(a_n, N_T, m, D_S, T_S, T_P^{(S)}, \Delta\right) = r_m, \quad \text{(E-3)}$$

and

$$\lambda = \frac{1}{r} \sum_{m=1}^{N_A} \sum_{n=1}^{N_T} \mathcal{F}\left(a_n, N_T, m, D_S, T_S, T_P^{(S)}, \Delta\right). \quad \text{(E-4)}$$

Substituting (E-4) into (E-3), we can derive (26) to complete the proof.

REFERENCES

[1] M. Xu et al., "A full dive into realizing the edge-enabled metaverse: Visions, enabling technologies, and challenges," 2022, *arXiv:2203.05471*.

[2] S. Kraus, D. K. Kanbach, P. M. Krysta, M. M. Steinhoff, and N. Tomini, "Facebook and the creation of the metaverse: Radical business model innovation or incremental transformation?" *Int. J. Entrepreneurial Behav. Res.*, vol. 28, no. 9, pp. 52–77, Dec. 2022.

[3] A. El Saddik, "Digital twins: The convergence of multimedia technologies," *IEEE MultimediaMag.*, vol. 25, no. 2, pp. 87–92, Apr. 2018.

[4] N. Xi, J. Chen, F. Gama, M. Riar, and J. Hamari, "The challenges of entering the metaverse: An experiment on the effect of extended reality on workload," *Inf. Syst. Frontiers*, pp. 1–22, Feb. 2022. [Online]. Available: https://link.springer.com/article/10.1007/s10796-022-10244-x

[5] Q. Yang, Y. Zhao, H. Huang, Z. Xiong, J. Kang, and Z. Zheng, "Fusing blockchain and AI with metaverse: A survey," 2022, *arXiv:2201.03201*.

[6] W. Yang et al., "Semantic communications for 6G future internet: Fundamentals, applications, and challenges," 2022, *arXiv:2207.00427*.

[7] C. Xu, L. Zhang, L. Song, Y. Zhao, and B. Jiao, "Feedback compression for time-correlated MIMO block-fading channels using Huffman coding," in *Proc. IEEE 13th Int. Conf. Commun. Technol.*, Sep. 2011, pp. 411–415.

[8] X. Xie, X. Zhang, and K. Sundaresan, "Adaptive feedback compression for MIMO networks," in *Proc. 19th Annu. Int. Conf. Mobile Comput. Netw. MobiCom*, 2013, pp. 477–488.

[9] D. J. Love, R. W. Heath Jr., V. K. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[10] K. Huang, R. W. Heath Jr., and J. G. Andrews, "Limited feedback beamforming over temporally-correlated channels," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1959–1975, May 2009.

[11] C. Carbonelli, S. Vedantam, and U. Mitra, "Sparse channel estimation with zero tap detection," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1743–1763, May 2007.

[12] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.

[13] H. Son and Y.-H. Cho, "Analysis of compressed CSI feedback in MISO systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 6, pp. 1671–1674, Dec. 2019.

[14] C.-K. Wen, W.-T. Shih, and S. Jin, "Deep learning for massive MIMO CSI feedback," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 748–751, Oct. 2018.

[15] T. Wang, C.-K. Wen, S. Jin, and G. Y. Li, "Deep learning-based CSI feedback approach for time-varying massive MIMO channels," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 416–419, Apr. 2019.

[16] W. Weaver, "The mathematics of communication," *Sci. Amer.*, vol. 181, no. 1, pp. 11–15, 1949.

[17] R. Carnap et al., "An outline of a theory of semantic information," Res. Lab. Electron., Massachusetts Inst. Technol., Cambridge, MA, USA, 1952.

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

[19] B. Juang, "Quantification and transmission of information and intelligence—History and outlook [DSP history]," *IEEE Signal Process. Mag.*, vol. 28, no. 4, pp. 90–101, Jul. 2011.

[20] P. Basu, J. Bao, M. Dean, and J. Hendler, "Preserving quality of information by using semantic relationships," *Pervas. Mobile Comput.*, vol. 11, pp. 188–202, Apr. 2014.

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

[22] Q. Hu, G. Zhang, Z. Qin, Y. Cai, G. Yu, and G. Ye Li, "Robust semantic communications against semantic noise," 2022, *arXiv:2202.03338*.

[23] Q. Zhou, R. Li, Z. Zhao, C. Peng, and H. Zhang, "Semantic communication with adaptive universal transformer," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 453–457, Mar. 2022.

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

[25] D. Huang, X. Tao, F. Gao, and J. Lu, "Deep learning-based image semantic coding for semantic communications," in *Proc. IEEE Global Telecom. Conf.*, Dec. 2021, pp. 1–6.

[26] G. Tullock, "Rent seeking," in *Property Rights and the Limits of Democracy*. Cambridge, MA, USA: Edward Elgar Publishing, 1981, pp. i–98.

[27] H. Moulin, *Game Theory for the Social Sciences*. New York, NY, USA: NYU Press, 1986.

[28] T. Broecker, "Credit-worthiness tests and interbank competition," *Econometrica, J. Econ. Soc.*, vol. 58, pp. 429–452, Mar. 1990.

[29] J. J. Anton and D. A. Yao, "Coordination in split award auctions," *Quart. J. Econ.*, vol. 107, no. 2, pp. 681–707, May 1992.

[30] J. Lu et al., "Multi-level two-sided rating protocol design for service exchange contest dilemma in crowdsensing," *IEEE Access*, vol. 7, pp. 78391–78405, 2019.

[31] M. Vojnović, *Contest Theory: Incentive Mechanisms and Ranking Methods*. Cambridge, U.K.: Cambridge Univ. Press, 2015.

[32] T. Luo, S. S. Kanhere, S. K. Das, and H.-P. Tan, "Incentive mechanism design for heterogeneous crowdsourcing using all-pay contests," *IEEE Trans. Mobile Comput.*, vol. 15, no. 9, pp. 2234–2246, Sep. 2016.

[33] T. Luo, S. S. Kanhere, H.-P. Tan, F. Wu, and H. Wu, "Crowdsourcing with Tullock contests: A new perspective," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2015, pp. 2515–2523.

[34] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Device-free human activity recognition using commercial WiFi devices," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1118–1131, May 2017.

[35] M. Gao, X. Yang, M. Zhou, J. Wang, and W. Nie, "Channel state information compression based on projection transformation and curve fitting," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.

[36] F. Gringoli, M. Schulz, J. Link, and M. Hollick, "Free your CSI: A channel state information extraction platform for modern Wi-Fi chipsets," in *Proc. Int. Workshop Wireless Netw. Testbeds Experim. Eval. Charact.*, 2019, pp. 21–28.

[37] Y. Jeong, H. Shin, and M. Z. Win, "H-transforms for wireless communication," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3773–3809, Jul. 2015.

[38] L. Kong, G. Kaddoum, and H. Chergui, "On physical layer security over fox's H-function wiretap fading channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 7, pp. 6608–6621, Jul. 2019.

[39] A. M. Mathai, R. K. Saxena, and H. J. Haubold, *The H-function: Theory and Applications*. Berlin, Germany: Springer, 2009.

[40] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Feb. 1999.

[41] Y. A. Rahama, M. H. Ismail, and M. S. Hassan, "On the sum of independent fox's H-function variates with applications," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 6752–6760, Aug. 2018.

[42] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[43] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. New York, NY, USA: Academic, 2007.

[44] J. Liu, H. Liu, Y. Chen, Y. Wang, and C. Wang, "Wireless sensing for human activity: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1629–1645, 3rd Quart., 2020.

[45] L. C. Corchón and M. Serena, "Contest theory," in *Handbook of Game Theory and Industrial Organization*, vol. 2. Cheltenham, U.K.: Edward Elgar, 2018.

[46] J. Tremewan, "Behavioral economics: Toward a new economics by integration with traditional economics," *Econ. Rec.*, vol. 96, no. 313, p. 221, 2020.

[47] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single WiFi access point," in *Proc. USENIX Symp. Networked Syst. Design Implement.*, 2016, pp. 165–178.

[48] P. Yadav, S. Kumar, and R. Kumar, "A review of transmission rate over wireless fading channels: Classifications, applications, and challenges," *Wirel. Pers. Commun.*, vol. 122, pp. 1–57, Sep. 2021.

[49] A. A. Kilbas, *H-Transforms: Theory and Applications*. Boca Raton, FL, USA: CRC Press, 2004.

[50] H. R. Alhennawi, M. M. H. El Ayadi, M. H. Ismail, and H.-A.-M. Mourad, "Closed-form exact and asymptotic expressions for the symbol error rate and capacity of the *H*-function fading channel," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 1957–1974, Apr. 2016.

Jiawen Kang (Member, IEEE) received the Ph.D. degree from the Guangdong University of Technology, China, in 2018. He was a Post-Doctoral Researcher at Nanyang Technological University, Singapore, from 2018 to 2021. He is currently a Professor with the Guangdong University of Technology, China. His research interests include blockchain, security, and privacy protection in wireless communications and networking.

Jiacheng Wang received the Ph.D. degree from the School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China. He is currently a Research Associate in computer science and engineering with Nanyang Technological University, Singapore. His research interests include wireless sensing, semantic communications, and metaverse.

Hongyang Du (Graduate Student Member, IEEE) is currently pursuing the Ph.D. degree with the Interdisciplinary Graduate Program, School of Computer Science and Engineering, Energy Research Institute, Nanyang Technological University (NTU), Singapore. His research interests include semantic communications, metaverse, and communication theory.

Xuemin (Sherman) Shen (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Rutgers University, New Brunswick, NJ, USA, in 1990. He is currently a University Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include network resource management, wireless network security, the Internet of Things, 5G and beyond, and vehicular ad hoc and sensor networks. He is a fellow of the Engineering Institute of Canada, a fellow of the Canadian Academy of Engineering, a fellow of the Royal Society of Canada, and a Foreign Member of the Chinese Academy of Engineering. He is a member of the IEEE Fellow Selection Committee of the ComSoc. He was a recipient of Canadian Award for Telecommunications Research from the Canadian Society of Information Theory (CSIT) in 2021, the R. A. Fessenden Award in 2019 from IEEE, Canada, the Award of Merit from the Federation of Chinese Canadian Professionals (Ontario) in 2019, the James Evans Avant Garde Award in 2018 from the IEEE Vehicular Technology Society, the Joseph LoCicero Award in 2015 and Education Award in 2017 from the IEEE Communications Society, and the Technical Recognition Award from Wireless Communications Technical Committee in 2019 and AHSN Technical Committee in 2013. He was also a recipient of the Excellent Graduate Supervision Award in 2006 from the University of Waterloo and the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada. He was the Technical Program Committee Chair/Co-Chair of IEEE GLOBECOM 2016, IEEE Infocom 2014, IEEE VTC 2010 Fall, and IEEE GLOBECOM 2007, and the Chair of the IEEE Communications Society Technical Committee on Wireless Communications. He was the Editor-in-Chief of the IEEE INTERNET OF THINGS JOURNAL, *IEEE Network*, and *IET Communications*. He is the President of the IEEE Communications Society. He was the Vice President of Technical and Educational Activities, the Vice President for Publications, a Member-at-Large on the Board of Governors, and the Chair of the Distinguished Lecturer Selection Committee. He is a Registered Professional Engineer of ON, Canada. He is a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society.

Zengshan Tian (Member, IEEE) received the bachelor's degree from Zhengzhou University, Zhengzhou, China, in 1990, and the master's and Ph.D. degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 1998 and 2002, respectively. He is currently a Professor with the Chongqing University of Posts and Telecommunications, Chongqing, China. Over the past ten years, he was engaged in two national projects and more than ten enterprise projects. His research interests include localization in the cellular networks, personal communication, precise localization, and attitude measurement with GPS, data compression, and deep learning.

Dusit Niyato (Fellow, IEEE) received the B.Eng. degree from the King Mongkuts Institute of Technology Ladkrabang (KMITL), Thailand, in 1999, and the Ph.D. degree in electrical and computer engineering from the University of Manitoba, Canada, in 2008. He is currently a Professor with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interests include sustainability, edge intelligence, decentralized machine learning, and incentive mechanism design.