

Joint active and passive beamforming optimization of IRS-assisted vehicular system with IA-GNN

Liang Xiaolin, Chen Wenhua (✉), Liu Qianlong, Li Caixia, Huang Yongping

School of Electronic and Information Engineering, Hebei University, Baoding 071002, China

Abstract

The implementation of an effective intelligent reflecting surface (IRS) assisted millimeter-wave (mmWave) vehicular system critically depends on the proper beamforming design exploiting accurate channel state information (CSI). However, it is difficult to obtain accurate CSI due to the passive characteristics of reflective elements and the high-speed movement of vehicle user equipment (VUE). Furthermore, the channel estimation requires a tremendous amount of pilot overhead. In this paper, an instantaneous CSI aware graph neural network (IA-GNN) model based on unsupervised deep learning (DL) was proposed to implicitly learn the channel features and capture the relationship between the IRS and VUEs for IRS-assisted vehicle-to-infrastructure (V2I) links. The graph attention (GAT) mechanism is introduced to better integrate the correlation of node features into the IA-GNN model and improve the accuracy of mapping relationships. By using the IA-GNN model, different utility functions are taken as system objectives for joint active and passive beamforming optimization. Numerical results demonstrate that the proposed beamforming optimization scheme using the IA-GNN model has significant advantages over the existing schemes in terms of generalization and low pilot overhead.

Keywords intelligent reflecting surface (IRS), vehicle-to-infrastructure (V2I), beamforming, instantaneous CSI aware graph neural network (IA-GNN), graph attention (GAT)

1 Introduction

As the crown jewel of intelligent transportation, vehicle-to-everything (V2X) covers a wide range of wireless technologies. V2I is one of the fundamental communication technologies. V2X communication will be an important part of future connected autonomous vehicles, which are expected to have a profoundly transformative impact on intelligent transportation and

intelligent life^[1]. In emerging vehicle-mounted applications, mmWave has abundant bandwidth and spectrum resources, which can effectively support higher transmission rate and lower transmission latency^[2]. However, mmWave has serious path loss during transmission and is easily affected by obstacles, which is more serious in high-speed mobile scenarios. Therefore, high interrupt probability and low coverage are two urgent problems to be solved in mmWave V2X communication systems^[3].

The IRS breaks the uncontrollable limitation of the traditional wireless channels and realizes the intelligent control of the wireless propagation environment^[4]. IRS does not actively configure radio frequency links to

receive, transmit, and process signals, but uses its own unique electromagnetic characteristics to reflect signals, thus realizing coherent superposition of the reflected signal at the target receiver^[5]. Therefore, the IRS can effectively improve the data rate and transmission reliability of communication systems, as well as provide technical support for solving the bottleneck problem of mmWave applications in V2X systems^[6]. In actual traffic scenarios, V2I communication links are blocked by buildings or obstacles in complex intersections and crowded vertical streets, making the strength of the received signal rapidly decline with the increase of distance. In order to solve this problem, IRS is installed on the surface of buildings around the road to improve the reliability and coverage of vehicular communication. In addition, IRS can be used to mitigate the Doppler effect and multipath fading in V2I communication links due to high-speed mobility at the receiving or transmitting end^[7]. By designing an IRS reflection coefficient, i. e., passive beamforming, IRS can be widely deployed in mmWave V2X systems to achieve long distance, high-rate, and ultra-reliable communication.

The beamforming design in the IRS-assisted system encompasses two interdependent components: active beamforming at the base station (BS) or VUE and passive beamforming at the IRS. Active beamforming involves the design of complex-valued precoding vectors at the BS (or VUE) equipped with active radio frequency chains. Its function is to perform coherent signal combination at the transmitter, to maximize the signal strength at the intended receivers, and to mitigate inter-user interference. In contrast, passive beamforming refers to the optimization of the phase shifts introduced by the passive elements of the IRS. Without amplifying or generating signals, the IRS passively shapes the wireless environment by deliberately altering the phase of the impinging electromagnetic waves, thereby constructively combining signal paths at the receivers. The joint optimization of these active and passive beamforming vectors is paramount to unlocking the full potential of IRS-assisted systems. Despite the appealing advantages and great potential of IRS, there are new practical

issues and challenges that need to be addressed in IRS-assisted wireless communications. IRS passive beamforming design under imperfect CSI is one of the main challenges.

IRS passive beamforming optimization highly depends on the available CSI. Some of the current work on the passive beamforming design assume that the perfect CSI can be obtained^[8-9], however, the CSI is required to estimate to achieve joint optimization of active and passive beamforming in practice^[10]. Most of the researches on IRS use the estimated CSI for beamforming design on time-invariant systems^[11-12]. Recently, a small amount of work has begun to investigate IRS beamforming configuration in mobile scenarios^[13]. Considering the communication interruption probability, the problem of maximizing the average received signal power by using statistical CSI is proposed in Ref. [11] through the joint optimization of active and passive beamforming of access points. The estimated CSI is used to analyze the strongest signal path of the orthogonal frequency division multiplexing (OFDM) system in the time domain to achieve IRS reflection coefficient optimization in Ref. [12]. A joint beamforming optimization algorithm for mobile scenarios is proposed in Ref. [13], which utilizes a two-tier composite neural network based on an attention mechanism to improve communication quality in mobile environments.

Due to the lack of active components for baseband signal processing and the substantial increase in system overhead of IRS channel estimation, obtaining accurate CSI is formidable. Moreover, due to the high mobility of vehicles at the receiving and transmitting ends of V2I links, serious Doppler effect and multipath fading will be caused, preventing the accurate estimation of channels.

Besides the CSI-based IRS passive beamforming design, it has been used to design the IRS reflection coefficient via passive beamforming training based on DL, thus alleviating the need for explicit CSI. And DL can alleviate computational complexity caused by the optimization of active and passive beamforming, which is a high-dimensional nonconvex problem.

Given perfect channels, a multi-layer fully

connected neural network is proposed in Ref. [14] to learn the reflection coefficients of a single-user system. Similarly, an unsupervised two-stage neural network is proposed in Ref. [15] to predict the active and passive beamforming of an IRS-assisted multiple-input single-output (MISO) system in real-time. A heterogeneous-based multi-agent deep deterministic policy gradient approach is proposed in Ref. [16] for joint active and passive beamforming optimization without channel estimation overhead. The above work is mainly for IRS-assisted time-invariant systems. The active and passive beamforming for the multi-user mobile systems, such as V2I links, is required to be investigated in depth.

The graph neural network (GNN) linear minimum mean square error (LMMSE) model has permutation invariance and permutation equivariance, which allows it to be generalized to a communication system with an arbitrary number of users. It is worth noting that GNN was proposed mostly to solve the problem of system resource allocation in Refs. [17–19], and the GNN designed for the IRS-assisted mmWave V2I system needs to be further explored.

To overcome the practical issues and challenges mentioned above, an IA-GNN model based on unsupervised DL is proposed, which bypasses the explicit channel estimation. And the GAT mechanism is introduced into the IA-GNN model. The problem of designing the active and passive beamforming using the IA-GNN model for IRS-assisted V2I links is considered as a variational optimization problem with taking active and passive beamforming vectors as optimization variables. The IA-GNN model explains how to analyze and grasp the active and passive beamforming patterns from training data. Based on the proposed IA-GNN model, different utility functions are taken as system objectives for joint active and passive beamforming optimization. The main contributions can be summarized as follows.

1) An IA-GNN based on unsupervised DL is proposed. The proposed IA-GNN can accurately capture the relationship between the IRS and VUEs to reasonably design the IRS reflection coefficient. The IA-GNN can easily share parameters with different

VUEs and take different data types as inputs, thus it can be applied to multi-user V2I vehicular systems.

2) The GAT mechanism is introduced into the IA-GNN model to further capture the connections between IRS and VUEs, and make the update of node features more comprehensive for IRS-assisted V2I links. It makes IA-GNN better realize the link prediction of nodes and achieve target prediction for IRS-assisted V2I links.

3) The pilot transmission strategy for the IRS-assisted V2I mmWave system is proposed to train the internal parameters of the IA-GNN model and directly map received pilots to the targeted passive beamforming at the IRS and targeted active beamforming at the BS or VUEs. By bypassing the explicit channel estimation, it can significantly reduce the training overhead.

4) Based on the proposed IA-GNN model, the active and passive beamforming are optimized for different network utility functions. And the numerical results are compared with traditional beamforming optimization schemes. It reveals that the proposed IA-GNN model has better generalization and can achieve better system performance with lower pilot overhead.

The rest of the paper is organized as follows. Sect. 2 presents the channel description for V2I links. The downlink pilot transmission scheme and system utility problem formulation are introduced in Sect. 3. Sect. 4 describes the proposed IA-GNN model, and Sect. 5 presents the numerical results. Sect. 6 draws conclusions.

The notations used in this paper are as follows. $(\cdot)^T$ and $(\cdot)^H$ represent the transpose and Hermitian transpose of matrices, respectively. $\mathcal{CN}(\cdot)$ represents a complex Gaussian distribution. $E\{\cdot\}$ is the statistical expectation operation. $\text{diag}(\cdot)$ denotes generating a diagonal matrix. $\odot$ represents the Hadamard product. $\|\cdot\|_2$ and $\|\cdot\|_F$ represent the spectral norm and Frobenius norm, respectively.

2 Channel description

As illustrated in Fig. 1, an IRS-assisted downlink mmWave V2I communication system is considered. The IRS is installed on the surface of buildings around the road intersection with $N = N_y \times N_z$ passive reflecting

elements, where N_y and N_t represent the number of elements in the vertical and horizontal directions of the IRS, respectively. It is assumed that the BS with J -element uniform plane array (UPA) performs local V2I data exchange with K VUEs via the IRS, while $U_k, k \in \{1, 2, \dots, K\}$, represents the k th receiving VUE (r-VUE).

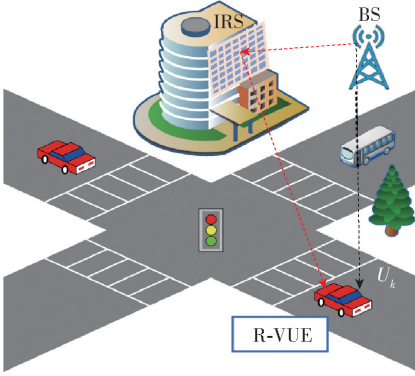


Fig. 1 IRS-assisted mmWave vehicular communication system

In order to describe the realistic channel, both direct and IRS-assisted indirect links are considered for V2I links. It is assumed that the vehicular system operates in time-division duplex (TDD) mode, and each VUE is mobile, and the BS/IRS are static. Therefore, the location of BS and IRS can be depicted by the three-dimensional (3D) coordinates $\mathbf{L}_B = (x_B, y_B, z_B)$ and $\mathbf{L}_I = (x_I, y_I, z_I)$ respectively, where x_B, y_B , and z_B (or x_I, y_I , and z_I) are the coordinates on x -axis, y -axis, and z -axis, respectively. Considering the mobility of VUEs, the locations of r-VUE at the i th time slot are $\mathbf{L}_{\text{r-VUE}}^i = (x_{\text{r-VUE}}^i, y_{\text{r-VUE}}^i, 1)$, where $x_{\text{r-VUE}}^i$ and $y_{\text{r-VUE}}^i$ represent the coordinates on x -axis and y -axis and 1 is expressed as the height of the vehicle on z -axis. Therefore, the location of each time-varying VUE is defined as

$$\mathbf{L}_{\text{VUE}}^{i+1} = \mathbf{L}_{\text{VUE}}^i + \mathbf{v}_{\text{VUE}}^i \Delta T \quad (1)$$

where $\mathbf{L}_{\text{VUE}}^{i+1}$ can be expressed as the location information of r-VUE for the next time slot. $\mathbf{v}_{\text{VUE}}^i = (v_{\text{VUE},x}^i, v_{\text{VUE},y}^i, 0)$ is the instantaneous speed in a unit time slot, where $v_{\text{VUE},x}^i$ and $v_{\text{VUE},y}^i$ are indicated as the velocity component on the x -axis and y -axis, respectively. ΔT is the duration of each time slot.

2.1 V2I channel model

Taking into account V2I links, the Rician fading

model is utilized to imitate the V2I channel. Let denote the channel matrices of the BS – IRS fixed link existing line-of-sight (LoS) and non-line-of-sight (NLoS) components, which can be depicted as

$$\mathbf{h}_{\text{BI}} = \sqrt{\frac{\varepsilon^{\text{BI}}}{\varepsilon^{\text{BI}} + 1}} \bar{\mathbf{h}}_{\text{BI}} + \sqrt{\frac{1}{\varepsilon^{\text{BI}} + 1}} \tilde{\mathbf{h}}_{\text{BI}} \quad (2)$$

where, $\varepsilon^{\text{BI}} \geq 0$ is the Rician factor with respect to the BS – IRS channel. $\bar{\mathbf{h}}_{\text{BI}}$ and $\tilde{\mathbf{h}}_{\text{BI}}$ define LoS component and NLoS component, respectively. $\bar{\mathbf{h}}_{\text{BI}}$ can be obtained as

$$\bar{\mathbf{h}}_{\text{BI}} = \sqrt{\frac{1}{L_b}} \sum_{l_b=1}^{L_b} \beta_{l_b}^{\text{BI}} \mathbf{a}_{\text{IRS}}(\varphi_l^{\text{BI}}, \vartheta_l^{\text{BI}}) \mathbf{a}_{\text{BS}}^H(\varphi_B^{\text{BI}}, \vartheta_B^{\text{BI}}) \quad (3)$$

$\tilde{\mathbf{h}}_{\text{BI}}$ is modeled as standard Gaussian distribution, i.e., $\tilde{\mathbf{h}}_{\text{BI}} \sim \mathcal{CN}(0, 1)$. L_b defines the number of propagation paths about the BS – IRS link, while $\beta_{l_b}^{\text{BI}}$ represent the l_b th path gain. φ_B^{BI} and ϑ_B^{BI} denote the azimuth angles of departure (AAOD) and elevation angles of departure (EAOD) from BS to IRS. φ_I^{BI} and ϑ_I^{BI} denote the azimuth angles of arrival (AAOA) and elevation angles of arrival (EAOA) from BS to IRS. Then the steering vectors about the BS – IRS channel are derived as

$$\mathbf{a}_{\text{BS}}(\varphi_B^{\text{BI}}, \vartheta_B^{\text{BI}}) = (1, e^{j\frac{2\pi d_{\text{BS}}}{\lambda_c} \cos \varphi_B^{\text{BI}} \cos \vartheta_B^{\text{BI}}}, \dots, e^{j\frac{2\pi (J-1)d_{\text{BS}}}{\lambda_c} \cos \varphi_B^{\text{BI}} \cos \vartheta_B^{\text{BI}}}) \quad (4)$$

$$\mathbf{a}_{\text{IRS}}(\varphi_I^{\text{BI}}, \vartheta_I^{\text{BI}}) = (1, e^{j\frac{2\pi d_{\text{IRS}}}{\lambda_c} \sin \varphi_I^{\text{BI}} \cos \vartheta_I^{\text{BI}}}, \dots, e^{j\frac{2\pi d_{\text{IRS}}}{\lambda_c} [r_1(n) \sin \varphi_I^{\text{BI}} \cos \vartheta_I^{\text{BI}} + r_2(n) \sin \vartheta_I^{\text{BI}}]}, \dots, e^{j\frac{2\pi d_{\text{IRS}}}{\lambda_c} [r_1(N) \sin \varphi_I^{\text{BI}} \cos \vartheta_I^{\text{BI}} + r_2(N) \sin \vartheta_I^{\text{BI}}]}) ; N = N_y \times N_z \quad (5)$$

$$\left. \begin{aligned} \cos \varphi_B^{\text{BI}} \cos \vartheta_B^{\text{BI}} &= \frac{x_I - x_B}{d^{\text{BI}}} \\ \sin \varphi_I^{\text{BI}} \cos \vartheta_I^{\text{BI}} &= \frac{y_B - y_I}{d^{\text{BI}}} \\ \sin \vartheta_I^{\text{BI}} &= \frac{z_B - z_I}{d^{\text{BI}}} \end{aligned} \right\} \quad (6)$$

where, d^{BS} is the distance between two adjacent BS antenna elements, λ_c is the wavelength of carrier frequency. d^{IRS} is the distance between two adjacent IRS reflection elements. d^{BI} is the distance between BS and IRS. $r_1(n) = \text{mod}(n-1, 10)$, $r_2(n) = (n-1)/10$, and n indicates the sequence of corresponding reflection elements.

The channel of IRS – U_k link at the i th time slot can be denoted as $\mathbf{h}_{\text{IU}_k}^i \in \mathbb{C}^{1 \times N}$, which is similarly depicted as

$$\mathbf{h}_{\text{IU}_k}^i = \sqrt{\frac{\varepsilon_k^{\text{IU}}}{\varepsilon_k^{\text{IU}} + 1}} \bar{\mathbf{h}}_{\text{IU}_k}^i + \sqrt{\frac{1}{\varepsilon_k^{\text{IU}} + 1}} \hat{\mathbf{h}}_{\text{IU}_k}^i \quad (7)$$

where, $\varepsilon_k^{\text{IU}} \geq 0$ is the Rician factor with respect to IRS – U_k link. $\bar{\mathbf{h}}_{\text{IU}_k}^i$ and $\hat{\mathbf{h}}_{\text{IU}_k}^i \sim \mathcal{CN}(0, 1)$ are defined as LoS and NLoS components, respectively. $\bar{\mathbf{h}}_{\text{IU}_k}^i$ can be expressed as

$$\bar{\mathbf{h}}_{\text{IU}_k}^i = \sqrt{\frac{1}{L_r}} \sum_{l_r=1}^{L_r} \beta_{k,l_r}^{\text{IU}} \mathbf{a}_{\text{IRS}}(\varphi_k^{l_r}, \vartheta_k^{l_r}) e^{j2\pi f_{k,l_r} T_s \cos \vartheta_k^{l_r} \cos \varphi_k^{l_r}} \quad (8)$$

where L_r defines the number of propagation paths about IRS – U_k link. $\beta_{k,l_r}^{\text{IU}}$ represent the l_r th path gain, f_{k,l_r} represent the Doppler shift, and T_s represent the system sampling period. $\varphi_k^{l_r}$ and $\vartheta_k^{l_r}$ denote the AAOD and EAOD from IRS to U_k . The steering vectors about IRS – U_k link can be assigned by

$$\mathbf{a}_{\text{IRS}}(\varphi_k^{l_r}, \vartheta_k^{l_r}) = (1, e^{j\frac{2\pi d_{\text{IRS}}}{\lambda_c} \sin \varphi_k^{l_r} \cos \vartheta_k^{l_r}}, \dots, e^{j\frac{2\pi d_{\text{IRS}}}{\lambda_c} [r_1(n) \sin \varphi_k^{l_r} \cos \vartheta_k^{l_r} + r_2(n) \sin \vartheta_k^{l_r}]}, \dots, e^{j\frac{2\pi d_{\text{IRS}}}{\lambda_c} [r_1(N) \sin \varphi_k^{l_r} \cos \vartheta_k^{l_r} + r_2(N) \sin \vartheta_k^{l_r}]} \quad (9)$$

$$\left. \begin{aligned} \sin \varphi_k^{l_r} \cos \vartheta_k^{l_r} &= \frac{y_{U_k} - y_1}{d_k^{\text{IU}}} \\ \sin \vartheta_k^{l_r} &= \frac{1 - z_1}{d_k^{\text{IU}}} \end{aligned} \right\} \quad (10)$$

where d_k^{IU} is the distance between IRS and r-VUE U_k .

2.2 Signal model

Considering the V2I links described above, IRS is deployed to help transmit information between the transmitter and receiver. The received signal r_k at the receiver U_k can be depicted as

$$r_k = \sum_{k=1}^K ((\mathbf{H}_{\text{BU}_k}^i + \mathbf{h}_{\text{BI}} \text{diag}(\mathbf{\Phi}_i)(\mathbf{h}_{\text{IU}_k}^i)^T)^T \mathbf{w}_{k,i} s_{k,i} + n_{k,i}) \quad (11)$$

where, $\mathbf{H}_{\text{BU}_k}^i \in \mathbb{C}^{J \times 1}$ denotes the direct channel of BS – U_k link at the i th time slot. $s_{k,i} \in \mathbb{C}$ is the symbol from the BS to U_k by the downlink transmission at the i th time slot, and $s_{k,i}$ is controlled by the BS through the beamforming vector $\mathbf{w}_{k,i} \in \mathbb{C}^{J \times 1}$. The phase shift matrix $\mathbf{\Phi}_i$ is used by IRS to reflect symbol $s_{k,i}$ back to

U_k , i. e., $\text{diag}(\mathbf{\Phi}_i) \in \mathbb{C}^{N \times N} = \text{diag}([e^{j\varphi_{1,i}}, e^{j\varphi_{2,i}}, \dots, e^{j\varphi_{N,i}}])$, and $\varphi_{n,i} \in [0, 2\pi)$ is the n th phase shift element. $n_{k,i} \sim \mathcal{CN}(0, \sigma_k^2)$ represents the additive Gaussian noise at U_k , and σ_k^2 is the noise variance.

As described above, the achievable rate of the received signal can be represented as

$$R_k = \text{lb} \left(1 + |(\mathbf{H}_{\text{BU}_k}^i + \mathbf{h}_{\text{BI}} \text{diag}(\mathbf{\Phi}_i)(\mathbf{h}_{\text{IU}_k}^i)^T)^T \mathbf{w}_{k,i}|^2 \cdot \left(\sum_{j=1, j \neq k}^K |(\mathbf{H}_{\text{BU}_k}^i + \mathbf{h}_{\text{BI}} \text{diag}(\mathbf{\Phi}_i)(\mathbf{h}_{\text{IU}_k}^i)^T)^T \mathbf{w}_{j,i}|^2 + \sigma_k^2 \right)^{-1} \right) \quad (12)$$

3 Pilot transmission and problem formulation

3.1 Design of pilot transmission

It is expected to maximize a downlink network utility function by using a neural network at the receiving end in the data transmission phase. Therefore, the network model is considered to be trained by the uplink pilot information strategy. For the V2I links, the k th r-VUE U_k emits a pilot information $\mathbf{z}_k(i)$, $i = 1, 2, \dots, T$ at the i th pilot length, which is transmitted to the BS by the IRS.

The pilot sequence $\mathbf{z}_k(i)$ at the i th pilot length is received can be expressed as

$$\mathbf{y}(i) = \sum_{k=1}^K ((\mathbf{H}_{\text{BU}_k}^i + \mathbf{h}_{\text{BI}} \text{diag}(\mathbf{\Phi}_i)(\mathbf{h}_{\text{IU}_k}^i)^T)^T \cdot \mathbf{w}_k(i) \odot \mathbf{z}_k(i) + \tilde{\mathbf{n}}_k(i)) \quad (13)$$

where $\tilde{\mathbf{n}}_k(i)$ denotes the noise. Specifically, the total pilot length T is divided into g subframes, and each subframe contains T_0 symbols, where the symbol length T_0 is set to the number of VUE k , i. e., $T = gT_0$. The pilot sequences $\mathbf{Z}_k^H = [\mathbf{z}_k(1), \mathbf{z}_k(2), \dots, \mathbf{z}_k(T_0)]$ of length T_0 are transmitted to the BS, and each subframe of the pilot performs the same operation. In order to extract the channel information, the uplink pilots are mutually orthogonal^[20], i. e.

$$\left\{ \begin{aligned} \mathbf{Z}_k^H \mathbf{Z}_m &= 0 \\ \mathbf{Z}_k^H \mathbf{Z}_k &= R \end{aligned} \right\} \quad (14)$$

where $m \neq k$, $R = T_0$. Similarly, the beamforming vector in a unit subframe can be expressed as $\mathbf{w}'_k =$

$[\mathbf{w}'_k(1), \mathbf{w}'_k(2), \dots, \mathbf{w}'_k(T_0)]$, and $\sum_{k=1}^K \|\mathbf{w}'_k(i)\|^2 \leq P_u$, P_u is the uplink pilot transmission power. Since IRS reflection coefficients and channel coefficients are unchanged in each subframe, but are not in agreement with different subframes, the pilot sequences received by each subframe τ are considered as

$$\tilde{\mathbf{Y}}(\tau) = \sum_{k=1}^K ((\mathbf{H}_{\text{BU}_k}^i + \mathbf{h}_{\text{BI}} \text{diag}(\boldsymbol{\Phi}_i) \cdot (\mathbf{h}_{\text{IU}_k}^i)^T)^T \mathbf{w}'_k \odot (\mathbf{Z}_k^H + \tilde{\mathbf{N}}_k(\tau)) \quad (15)$$

where $\tilde{\mathbf{Y}}(\tau) = [\mathbf{y}[(\tau-1)T_0+1], \mathbf{y}[(\tau-1)T_0+2], \dots, \mathbf{y}[\tau T_0]]$ is the received pilots in the τ th subframe, $\tau = 1, 2, \dots, g$, and $\tilde{\mathbf{N}}_k(\tau)$ is a noise matrix with each column independently distributed as $\mathcal{CN}(0, \sigma_k^2(i))$, $i = 1, 2, \dots, T_0$.

Because each pilot is orthogonal, the pilots can be decorrelated at the receiving end. It can be depicted as

$$\tilde{\mathbf{y}}_k(\tau) = \frac{1}{T_0} \tilde{\mathbf{Y}}(\tau) \mathbf{Z}_k = ((\mathbf{H}_{\text{BU}_k}^i + \mathbf{h}_{\text{BI}} \text{diag}(\boldsymbol{\Phi}_i) \cdot (\mathbf{h}_{\text{IU}_k}^i)^T)^T \mathbf{w}_k(\tau) + \tilde{\mathbf{n}}_k(\tau)) \triangleq \mathbf{A}_k(\tau) \mathbf{c}(\tau) \mathbf{w}_k(\tau) + \tilde{\mathbf{n}}_k(\tau) \quad (16)$$

where $\mathbf{A}_k(\tau) \triangleq [\mathbf{H}_{\text{BU}_k}^T, \mathbf{h}_{\text{BI}} (\mathbf{h}_{\text{IU}_k}^i)^T]$ and $\mathbf{c}(\tau) \triangleq [1, \text{diag}(\boldsymbol{\Phi}_\tau)]^T$, and the noise is defined as $\tilde{\mathbf{n}}_k(\tau) = (1/T_0) \tilde{\mathbf{N}}_k(\tau) \mathbf{Z}_k$. The matrix of received pilots by U_k can be defined as $\bar{\mathbf{Y}}_k = [\tilde{\mathbf{y}}_k(1), \tilde{\mathbf{y}}_k(2), \dots, \tilde{\mathbf{y}}_k(g)]$ for g subframes in total, thus, it can be derived as $\bar{\mathbf{Y}}_k = \bar{\mathbf{A}}_k \mathbf{C} \mathbf{W}_k + \bar{\mathbf{N}}_k$ (17)

where $\bar{\mathbf{A}}_k = [\mathbf{A}_k(1), \mathbf{A}_k(2), \dots, \mathbf{A}_k(g)]$, $\mathbf{C} = [\mathbf{c}(1), \mathbf{c}(2), \dots, \mathbf{c}(g)]$, $\mathbf{W}_k = [\mathbf{W}_k(1), \mathbf{W}_k(2), \dots, \mathbf{W}_k(g)]$ is the beamforming matrix for subframes, and $\bar{\mathbf{N}}_k = [\tilde{\mathbf{n}}_k(1), \tilde{\mathbf{n}}_k(2), \dots, \tilde{\mathbf{n}}_k(g)]$.

3.2 Problem formulation

As mentioned above, the research objective is to jointly optimize the active beamforming matrix at the transmitting end and the IRS reflection coefficient matrix for V2I links, but the indirect cascade channel matrix and direct channel matrix are required to obtain them. However, due to a mass of reflection elements in IRS, it is difficult to obtain accurate CSI for instantaneous channel estimation with low training overhead. Therefore, a DL algorithm is proposed to

map the active and passive beamforming matrix directly, bypassing the channel information estimation step.

The symbol is considered to be transmitted by downlink according to uplink-downlink channel reciprocity. The received symbol information is utilized as input network data through the receiver, where the received symbol information can be expressed as $\mathbf{Y} = [\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(T)] \in \mathbb{C}^{J \times T}$, by referring to Eq. (13). Hence, the active and passive beamforming matrix can be directly optimized based on the input data $\mathbf{Y}$, which can be expressed as

$$f(\mathbf{Y}) = (\mathbf{W}, \boldsymbol{\Phi}) \quad (18)$$

where $f(\cdot)$ is a mapping function, $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k, \dots, \mathbf{w}_K]$ is the active beamforming matrix at BS, and $\boldsymbol{\Phi} = [e^{j\varphi_1}, e^{j\varphi_2}, \dots, e^{j\varphi_N}]^T$ is the passive beamforming matrix at IRS. Therefore, the optimization problem about V2I links can be formulated by using the neural network as

$$\left. \begin{aligned} & \max_{f(\mathbf{Y}) = \mathbf{W}, \boldsymbol{\Phi}} E\{F(R_1(\mathbf{W}, \boldsymbol{\Phi}), R_2(\mathbf{W}, \boldsymbol{\Phi}), \dots, R_K(\mathbf{W}, \boldsymbol{\Phi}), \dots, R_K(\mathbf{W}, \boldsymbol{\Phi}))\} \\ & \text{s. t.} \quad \sum_{k=1}^K \|\mathbf{w}_k\|^2 \leq P_d; \quad 0 \leq \varphi_n < 2\pi, \\ & \quad n = 1, 2, \dots, N \end{aligned} \right\} \quad (19)$$

where $F(\cdot)$ is the utility function of the wireless network, the sum-rate maximization, and the max-min fairness problem are treated as utility functions to

depict in the simulation section. $\sum_{k=1}^K \|\mathbf{w}_k\|^2 \leq P_d$ represents the total power constraint, which means that the total transmission power of all VUEs cannot exceed the downlink transmission power P_d . $0 \leq \varphi_n < 2\pi$ represents the unit phase constraint, which means the range of the phase offset φ_n of each reflecting unit.

4 DL model design

It can be seen that the problem Eq. (19) is a non-convex objective function, which is challenging for practical calculation. In this section, an IA-GNN model based on unsupervised DL is proposed to realize the mapping function Eq. (18) and maximize the corresponding utility function.

4.1 IRS and VUEs node design

GNN is a network model that obtains the corresponding connection relationship through information transferred between nodes. The internal information can be extracted for an actual scenario with abundant relationships between nodes. For this reason, IRS reflection elements and VUEs are treated as nodes in the GNN. A GNN with an IRS element as a central node is established to coordinate the configuration between VUEs. It is assumed that the VUEs are mobile nodes and are connected to the IRS central node, where VUE nodes are represented as $\mathbf{A}_k, k=1, 2, \dots, K$ and $\mathbf{A}_0$ is a single IRS node. For a single channel, each IRS element maps corresponding information to different user nodes in this way, and these corresponding nodes can successfully predict the active beamforming vectors and IRS passive beamforming vectors after mass data training.

For the GNN, it has the characteristics of permutation invariance and permutation equivariance^[19], and it is not necessary to consider the ordering of VUE nodes to update the IRS reflection coefficient, and the VUE nodes can output the beamforming vectors of the corresponding sequence through this feature. Thus, the GNN can accurately capture the relationship between the IRS and VUEs, and reasonably regulate the weight variables between

VUEs. In order to capture edge relationships and make the update of nodes more comprehensive, GAT is introduced. GAT implicitly assigns different weights to different nodes in the neighborhoods. GAT mainly indicates the effect of the objective node $\vec{c}_b$ to node $\vec{c}_e$ through the attention mechanism a_{be} ^[21], which can be expressed as

$$a_{be} = \frac{e^{L(\vec{a}^T[\mathbf{Q}\vec{c}_b||\mathbf{Q}\vec{c}_e])}}{\sum_{e \in B_b} e^{L(\vec{a}^T[\mathbf{Q}\vec{c}_b||\mathbf{Q}\vec{c}_e])}} \quad (20)$$

where $||$ represents concat, $\vec{a}$ is a single layer feedforward neural network (FNN), B_b indicates the set of adjacent nodes of $\vec{c}_b$. $\mathbf{Q}$ is a linear transformation matrix applied to each node, and $L(\cdot)$ is expressed as a nonlinear operation. As above, each node will be updated by the weight coefficient between adjacent nodes, where the multi-head attention mechanism is taken into account to make the weighted results more stable^[21].

4.2 IA-GNN architecture

IA-GNN architecture consists of an input layer, D update layers, and an output layer that includes normalization. As shown in Fig. 2, the linear transformation of input information $\bar{\mathbf{Y}}_k$ is extracted and converted to a feature vector $\mathbf{A}_k^1$ in the input layer, the result $\mathbf{A}_k^D$ is passed into the output layer after update layers, which is mapped to the active beamforming matrix $\mathbf{W}$ and the IRS passive beamforming matrix Φ .

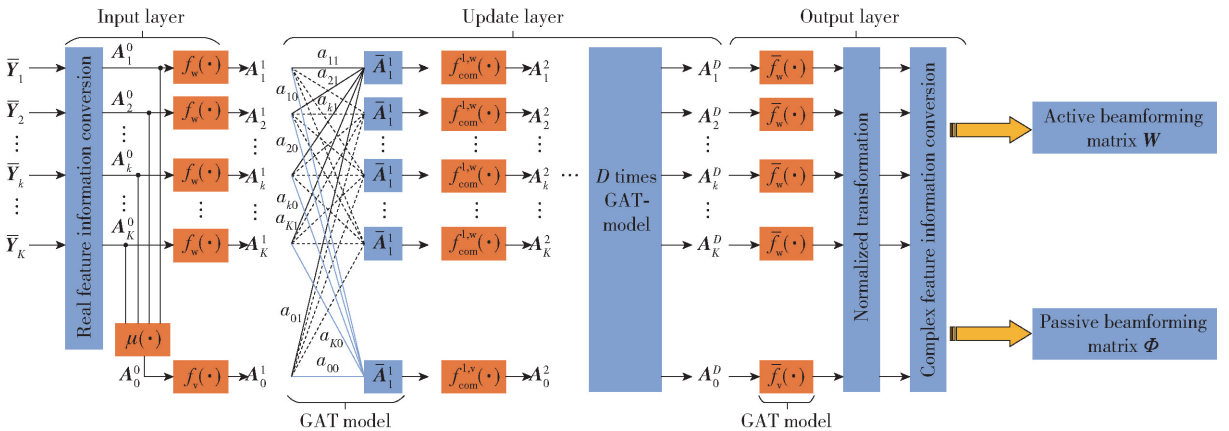


Fig. 2 Structural diagram of IA-GNN model

1) Input layer

Firstly, an initialization operation is performed at the input layer of the model, which converts the received pilot information into real component form $\mathbf{A}_k^0$ as

$$\mathbf{A}_k^0 = (V(\text{Re } \bar{\mathbf{Y}}_k), V(\text{Im } \bar{\mathbf{Y}}_k)) \quad (21)$$

where $V(\cdot)$ denotes the vectorization operation. Note that there is no input information for the IRS node, the mean of VUEs input information is used as the IRS

input feature $\mathbf{A}_0^0$, mainly because the IRS acts as a reflector for all VUEs nodes.

The network model uses layers of fully connected networks to extract the input information features, which define $f_w(\cdot)$ as generating the feature vector $\mathbf{A}_k^1$ for the VUE nodes, and define $f_v(\cdot)$ as generating the feature vector $\mathbf{A}_0^1$ for the IRS node, i. e.

$$\mathbf{A}_k^1 = f_w(\mathbf{A}_k^0); \quad k = 1, 2, \dots, K \quad (22)$$

$$\mathbf{A}_0^1 = f_v(\mathbf{A}_0^0); \quad \mathbf{A}_0^0 = \mu(\mathbf{A}_1^0, \mathbf{A}_2^0, \dots, \mathbf{A}_K^0) \quad (23)$$

where $\mu(\cdot)$ denotes the arithmetic mean of the feature vectors of all VUE nodes.

2) Update layer

The node feature vectors $\{\mathbf{A}_1^1, \mathbf{A}_2^1, \dots, \mathbf{A}_K^1, \mathbf{A}_0^1\}$ are initialized at the input layer and are passed to the D update layers, which contain features of the IRS element and the VUE channels, respectively. The node features will be updated in this module and produce a new set of node features $\mathbf{A} = \{\mathbf{A}_1^2, \mathbf{A}_2^2, \dots, \mathbf{A}_K^2, \mathbf{A}_0^2\}$. The update of GNN nodes includes an aggregation function and a combining function. $f_{agg}^d(\cdot)$, $d = 1, 2, \dots, D$, is the aggregation function, which can pass the features of adjacent nodes to the central node, emphasizing the aggregation capability of neighbor information. $f_{com}^d(\cdot)$, $d = 1, 2, \dots, D$, is the combining function, which can combine and update the representation of current nodes. In order to possess sufficient expression ability of the updated node feature information, the attention mechanism is utilized as $f_{agg}^d(\cdot)$ to generate higher-level features between all nodes. The feature vector $\bar{\mathbf{A}}_k^d$, $d = 1, 2, \dots, D$, and $\bar{\mathbf{A}}_0^d$, $d = 1, 2, \dots, D$, can be expressed as

$$\left. \begin{aligned} \bar{\mathbf{A}}_k^d &= f_{agg}^d \left(\sum_{\substack{m=1 \\ m \neq k}}^K \mathbf{A}_m^d \right) = o \left(\frac{1}{V} \sum_{v=1}^V \sum_{\substack{m=1 \\ m \neq k}}^K a_{km}^v \mathbf{Q}^v \mathbf{A}_m^d \right) \\ \bar{\mathbf{A}}_0^d &= f_{agg}^d \left(\sum_{m=1}^K \mathbf{A}_m^d \right) = o \left(\frac{1}{V} \sum_{v=1}^V \sum_{m=1}^K a_{0m}^v \mathbf{Q}^v \mathbf{A}_m^d \right) \end{aligned} \right\} \quad (24)$$

where $o(\cdot)$ is expressed as a nonlinear activation function, a_{km}^v and $\mathbf{Q}^v$ represents the v th attention mechanism and linear transformation, V is the number of multi-head attention and $V = 3$. The update of the feature vector $\bar{\mathbf{A}}_k^d$ is to aggregate adjacent feature nodes $\mathbf{A}_m^d$ according to the attention mechanism weights. Otherwise, the IRS node needs to serve all user nodes, unlike user nodes.

Due to the approximation ability of multilayer perception (MLP), which is adopted to get the utmost

out of node features and map the complex relations of optimization problems. For the combining function $f_{com}^d(\cdot)$, which can be expressed as

$$\left. \begin{aligned} \mathbf{A}_k^{d+1} &= f_{com}^{d,1}(\mathbf{A}_k^d, \bar{\mathbf{A}}_k^d) \\ \mathbf{A}_0^{d+1} &= f_{com}^{d,V}(\mathbf{A}_0^d, \bar{\mathbf{A}}_0^d) \end{aligned} \right\} \quad (25)$$

where $f_{com}^{d,1}(\cdot)$ and $f_{com}^{d,V}(\cdot)$ represent the combining function of the IRS node and VUE nodes at the d th update layer, respectively.

The update layers will implement the node transformation of the D times and finally the corresponding parameters for solving the nonlinear problem Eq. (19) are obtained through the framework of the above design.

3) Output layer

The above node features updated by D layers are passed into this layer. After linear transformation, normalization, and complex transformation of the feature vectors, the IRS passive beamforming matrix $\Phi \in \mathbb{C}^N$ and the active beamforming matrix $\mathbf{W} \in \mathbb{C}^{J \times K}$ are output. Meanwhile, the output results satisfy the unit phase constraint of φ_n and the total power constraint of P_d .

For the IRS node, the input feature vectors $\mathbf{A}_0^D$ are first passed through a fully connected layer $\bar{f}_v(\cdot)$ of $2N$ linear units in order to make the output in complex form, which can be expressed as

$$\tilde{\mathbf{A}}_0^D = \bar{f}_v(\mathbf{A}_0^D) \quad (26)$$

where the output vector $\tilde{\mathbf{A}}_0^D \in \mathbb{R}^{2N}$.

Then the output vector $\tilde{\mathbf{A}}_0^D$ is transformed and normalized by real and imaginary components, and the output of the passive beamforming matrix Φ can be expressed as

$$\left. \begin{aligned} \mathbf{A}_\Phi &= [\tilde{\mathbf{A}}_0^D(1:N), \mathbf{A}_0^D((N+1):2N)] \\ \Phi_b &= \frac{\tilde{\mathbf{A}}_\Phi(b,1)}{\sqrt{\mathbf{A}_\Phi(b,1)^2 + \mathbf{A}_\Phi(b,2)^2}} + \\ &\quad \mathrm{j} \frac{\tilde{\mathbf{A}}_\Phi(b,2)}{\sqrt{\mathbf{A}_\Phi(b,1)^2 + \mathbf{A}_\Phi(b,2)^2}} \end{aligned} \right\} \quad (27)$$

where the output vectors $\mathbf{A}_\Phi \in \mathbb{R}^{N \times 2}$, $\mathbf{A}_\Phi(b,1)$ represents the element in the b th row and the 1st column of the coefficient matrix $\mathbf{A}_\Phi$.

For the output of the active beamforming matrix $\mathbf{W}$ in the same way, the input feature vectors $\mathbf{A}_k^D$ are passed

through a layer $\bar{f}_w(\cdot)$ with $2J$ units, which can be represented as

$$\hat{\mathbf{A}}_k^D = \bar{f}_w(\mathbf{A}_k^D) \quad (28)$$

For the active beamforming matrix, the output feature information should satisfy

$$\mathbf{A}_W = [\hat{\mathbf{A}}_1^D, \hat{\mathbf{A}}_2^D, \dots, \hat{\mathbf{A}}_K^D] \in \mathbb{R}^{2J \times K} \quad (29)$$

And through complex transformation and normalization, it can be depicted as

$$\mathbf{W} = \mathbf{A}_W(1:J, :) + j\mathbf{A}_W((J+1):2J, :) \quad (30)$$

$$\mathbf{A}_W \leftarrow \sqrt{P_d} \frac{\mathbf{A}_W}{\|\mathbf{A}_W\|_F}$$

where $\mathbf{A}_W(1:J, :)$ represents extracting all column elements from the 1st row to the J th row of matrix $\mathbf{A}_W$.

According to the design of the IA-GNN model in this paper, the MLP models $f_w(\cdot)$, $f_v(\cdot)$, $f_{\text{com}}^{d,w}(\cdot)$, $f_{\text{com}}^{d,v}(\cdot)$, and the fully connected layers $\bar{f}_w(\cdot)$, $\bar{f}_v(\cdot)$ are not specific to VUEs and IRS reflection elements,

which can promote the communication system with an arbitrary number of VUEs and IRS elements.

4.3 IA-GNN network training

The 3D network diagram of IA-GNN is shown in Fig. 3, which is mainly established by dense layers and one-dimensional (1D) convolution (Conv1D) layers. The dense layers, consisting of 1024 units and 512 units, are jointly adopted for the input layer to extract data features, which are used to implement modules $f_w(\cdot)$ and $f_v(\cdot)$, respectively. And similarly, two dense layers with 512 units are used in the update layers to implement modules $f_{\text{com}}^{d,w}(\cdot)$ and $f_{\text{com}}^{d,v}(\cdot)$. Batch normalization (BN) and dropout layers are also used in networks to prevent training from overfitting. A linear layer with $2J$ units and a dropout layer combine to form the module $\bar{f}_w(\cdot)$ in the output layer, and $\bar{f}_v(\cdot)$ can be represented similarly.

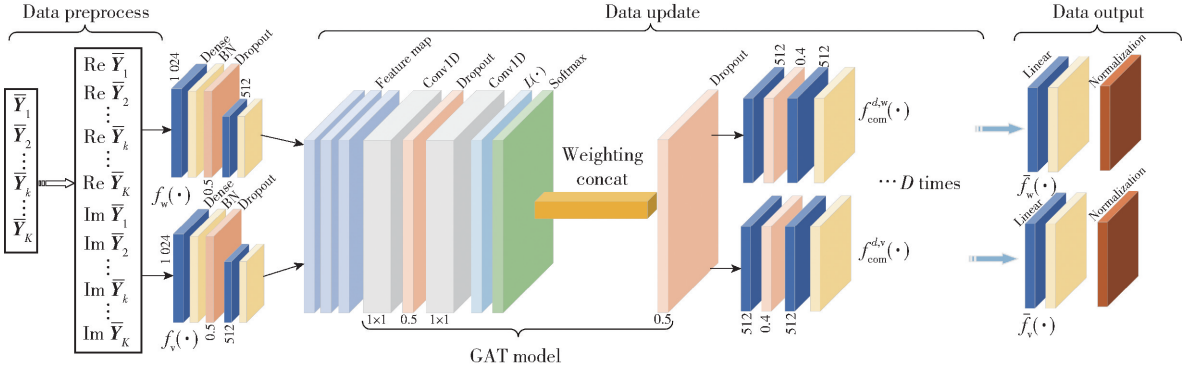


Fig. 3 3D cylinder diagram of the IA-GNN model

For the GAT model of update layers, the weight relationship between the central node and adjacent nodes is calculated after the feature map is entered into the Conv1D layers with 1×1 kernel. The nonlinear operation $L(\cdot)$ is performed after two layers of Cnov1D. The weight relationship is passed into the softmax layer for normalization, and the output of the objective node is weighted through the concat structure. D times of the above operations are included in the update layers.

The loss function of the IA-GNN model is depicted by the network, which can be expressed as $-E\{F(R_1(\mathbf{W}, \Phi), R_2(\mathbf{W}, \Phi), \dots, R_k(\mathbf{W}, \Phi), \dots, R_K(\mathbf{W}, \Phi))\}$. Since the existing tools cannot perform

complex valued operations, the achievable rate R_k can be depicted as

$$R_k = \text{lb} \left(1 + \frac{\|\mathbf{v}_k\|_2}{\sum_{m=1, m \neq k}^K \|\mathbf{v}_m\|_2 + \sigma_k^2} \right) \quad (31)$$

where

$$\mathbf{v}_k = |(\mathbf{H}_{\text{BU}_k}^i + \mathbf{h}_{\text{BI}} \text{diag}(\Phi_i)(\mathbf{h}_{\text{IU}_k}^i)^T)^T \mathbf{w}_{k,i}|^2 =$$

$$\begin{bmatrix} \text{Re } \mathbf{w}_{k,i}^T & -\text{Im } \mathbf{w}_{k,i}^T \\ \text{Im } \mathbf{w}_{k,i}^T & \text{Re } \mathbf{w}_{k,i}^T \end{bmatrix} \begin{bmatrix} \text{Re } \mathbf{H}_{\text{BU}_k}^i \\ \text{Im } \mathbf{H}_{\text{BU}_k}^i \end{bmatrix} +$$

$$\begin{bmatrix} \text{Re } \mathbf{h}_{\text{BI}}(\mathbf{h}_{\text{IU}_k}^i)^T & -\text{Im } \mathbf{h}_{\text{BI}}(\mathbf{h}_{\text{IU}_k}^i)^T \\ \text{Im } \mathbf{h}_{\text{BI}}(\mathbf{h}_{\text{IU}_k}^i)^T & \text{Re } \mathbf{h}_{\text{BI}}(\mathbf{h}_{\text{IU}_k}^i)^T \end{bmatrix} \cdot$$

$$\begin{bmatrix} \text{Re } \text{diag}(\Phi_i) \\ \text{Im } \text{diag}(\Phi_i) \end{bmatrix} \quad (32)$$

The instantaneous CSI data is used to train the IA-GNN and calculate the loss function during training. The IA-GNN is set with a fixed number of epochs and carried out using the gradient descent method during training. The loss function is the smallest, i. e., the utility function is the largest, when unsupervised learning is used to select the network weight.

5 Simulation results and analysis

5.1 System simulation setting

An IRS-assisted multi-user mmWave vehicular communication system is considered. The carrier frequency is set as $f_c = 28$ GHz. The number of BS antennas and IRS passive elements is set as $J = 8$ and $N = 100$ ^[22], respectively. The distances between the adjacent BS antenna elements or IRS elements are set as $d^{\text{BS}} = d^{\text{IRS}} = 0.5\lambda_c$. The number of r-VUEs is set as $K = 4$ ^[23]. The coordinates of the BS and IRS locations are $\mathbf{L}_B = (80 \text{ m}, -80 \text{ m}, 50 \text{ m})$ and $\mathbf{L}_I = (0, -10 \text{ m}, 50 \text{ m})$, respectively. Meanwhile, the initial position of r-VUE $\mathbf{L}_{\text{r-VUE}}$ is randomly generated within the region $[40 \text{ m}, 100 \text{ m}] \times [40 \text{ m}, 100 \text{ m}]$ and $[40 \text{ m}, 100 \text{ m}] \times [40 \text{ m}, 100 \text{ m}]$ in the (x, y) -plane with $z = 1 \text{ m}$, respectively. The instantaneous speeds of VUEs v_{VUE}^i are randomly produced within the interval $[7 \text{ m/s}, 14 \text{ m/s}] \times [7 \text{ m/s}, 14 \text{ m/s}]$ in the (x, y) -plane at the i th time slot, and the slot time $\Delta T = 0.02 \text{ s}$ ^[23]. The uplink pilot transmission power is set as $P_u = 15 \text{ dBm}$. The signal noise power is set to $\sigma_k^2 = \sigma_k^2(i) = -96 \text{ dBm}$.

For V2I links, it is assumed that the direct channel $\mathbf{H}_{\text{BU}_k}^i$ between BS and VUEs obeys Rayleigh fading.

$$\mathbf{H}_{\text{BU}_k}^i = \beta_k^{\text{BU}} \tilde{\mathbf{H}}_{\text{BU}_k}^i \quad (33)$$

where $\tilde{\mathbf{H}}_{\text{BU}_k}^i \sim \mathcal{CN}(0, 1)$, β_k^{BU} is the path loss^[24] and can be expressed as

$$\beta_k^{\text{BU}} = 43.8 + 36.8 \lg d_k^{\text{BU}} + 20 \lg \frac{4\pi f_c}{v_c} \quad (34)$$

where d_k^{BU} is the distance from the BS to the k th r-VUE U_k , $c = 3 \times 10^8 \text{ m/s}$ is the speed of light in free space. The IRS-assisted channels $\mathbf{h}_{\text{BI}}$ and $\mathbf{h}_{\text{IU}_k}^i$ are assumed to follow Rician fading^[23]. $\beta_{i_b}^{\text{BI}}$ and $\beta_{k, i_r}^{\text{IU}}$ are the path loss

of $\mathbf{h}_{\text{BI}}$ and $\mathbf{h}_{\text{IU}_k}^i$, respectively.

$$\left. \begin{aligned} \beta_{i_b}^{\text{BI}} &= 30.0 + 22.0 \lg d^{\text{BI}} + 20 \lg \frac{4\pi f_c}{c} \\ \beta_{k, i_r}^{\text{IU}} &= 46.8 + 18.7 \lg d_k^{\text{IU}} + 20 \lg \frac{4\pi f_c}{c} \end{aligned} \right\} \quad (35)$$

where $\mathbf{h}_{\text{BI}}$ and $\mathbf{h}_{\text{IU}_k}^i$ are considered both LoS and NLoS parts, respectively. The corresponding Rician factors are set as $\mathcal{E}^{\text{BI}} = \mathcal{E}_k^{\text{IU}} = 10$ ^[23].

5.2 Neural network parameter setting

The IA-GNN model is mainly built with the TensorFlow framework. An input layer, an output layer, and 3 update layers, i. e., $D = 3$, are concluded.

1) Channel sample

Each training sample generated offline in the model corresponds to an environmental snapshot of a certain time slot, including the random position and speed of the VUE, and the channel matrices $\mathbf{h}_{\text{BI}}$, $\mathbf{h}_{\text{IU}_k}^i$, and $\mathbf{H}_{\text{BU}_k}^i$ obtained by the Rician fading model described in Eqs. (2), (7) and (33), respectively.

2) Pre-training dataset

To make the network weights reach a reasonable level, the epochs are set to 100, and the number of channel samples for each epoch is set to 102 400. The batch size is set as 512 in model training. The adaptive moment (Adam) estimation optimizer is applied to update learnable parameters of the model. The initial learning rate is set to 10^{-3} and will decay to 0.98 times the original value after every 5 epochs. Also, the model is trained with reference to the Sum-Rate optimization problem to define the loss function.

3) Test dataset

The model uses 10 000 groups of channel samples independently generated and without overlapping with the training data to verify the generalization ability of the model.

The IA-GNN model is first trained offline on a pre-training dataset, and the pre-training covers a wide range of changes in vehicle status (location, speed) and channel conditions. After the pre-training is completed, the model weights are fixed. In the online phase, the IA-GNN model directly takes the pilot signals received in real-time as input, and

instantaneously outputs the optimized active and passive beamforming matrices without explicit channel estimation and iterative optimization.

To demonstrate the performance of the proposed IA-GNN, the following benchmarks are chosen for comparison.

Benchmark 1 Perfect CSI with block coordinate descent (BCD)

Perfect CSI combined with the BCD algorithm is used to solve the Sum-Rate maximization problem with active and passive beamforming^[25], and the algorithm is stopped if there is below 10^{-3} increase between the results of successive iterations.

Benchmark 2 LMMSE channel estimation with BCD

The LMMSE algorithm is first used to estimate the channel from the received pilot information. Then the BCD algorithm is utilized to maximize the iterative solution of the Sum-Rate problem with active and passive beamforming^[25].

Benchmark 3 Perfect CSI with random phase shift

Given perfect CSI, the IRS reflection coefficients are randomly generated by a fixed distribution, and the beamforming matrix $\mathbf{W}$ is optimized by using the weighted minimum mean square error (WMMSE) algorithm.

Benchmark 4 Beamforming neural network (BNN) with channel covariance matrix

BNN uses the channel covariance matrix instead of the instantaneous CSI to design the transmit beamforming and IRS phase shift^[26].

5.3 Sum-Rate of V2I links

The network utility function is defined as Sum-Rate, i. e. , the optimization function Eq. (19) for V2I links can be expressed as

$$\left. \begin{aligned} & \max_{f(\mathbf{Y})=\mathbf{W},\Phi} E \left\{ \sum_{k=1}^K R_k(\mathbf{W},\Phi) \right\} \\ & \text{s. t.} \\ & \sum_{k=1}^K \|\mathbf{w}_k\|^2 \leq P_d; \quad 0 \leq \varphi_n < 2\pi, \quad n = 1, 2, \dots, N \end{aligned} \right\} \quad (36)$$

And the corresponding loss functions for V2I links

can be designed as $-E \left\{ \sum_{k=1}^K R_k(\mathbf{W},\Phi) \right\}$.

Thus, Fig. 4 depicts the Sum-Rate versus the downlink pilot length based on the IA-GNN for V2I links, and the result is compared with Benchmarks 1 – 3. The downlink transmission power is set as $P_d = 20$ dBm. For V2I links, the pilot lengths are set as $T = \{3K, 5K, 7K, 10K, 20K, 30K, 40K, 50K, 60K, 70K, 85K\}$. The IA-GNN model achieves near-optimal Sum-Rate with only 80 pilots, closely matching 95% of Benchmark 1 performance. It surpasses Benchmarks 2 and 3, requiring just 40 pilots for comparable results versus Benchmark 2's 240 pilots. By optimizing IRS phase shifts, it significantly boosts Sum-Rate over Benchmark 3. This demonstrates IA-GNN's ability to cut pilot overhead and improve performance through efficient data adaptation and processing.

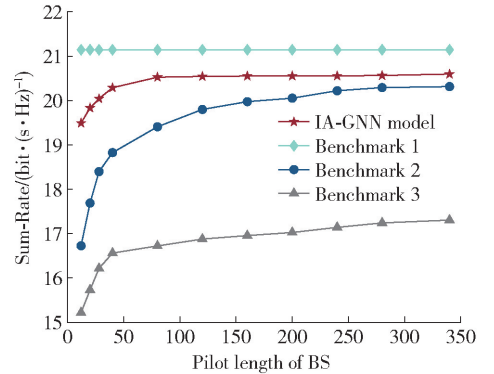


Fig. 4 Sum-Rate with different downlink pilot length for V2I links

Fig. 5 depicts the Sum-Rate versus the number of IRS elements for V2I links. The downlink transmission power is set as $P_d = 20$ dBm, and the pilot length is set as $T = 160$. It can be seen that as the number of IRS elements increases, the Sum-Rate of the IA-GNN model gradually approaches the Sum-Rate of Benchmark 1. As the number of IRS elements increases, Benchmark 2 struggles to estimate accurate CSI with limited pilots, but the IA-GNN exhibits superior performance in processing large-scale data introduced by the IRS. It can also be seen that the proposed IA-GNN model is far superior to Benchmark 3, with the number of IRS elements increasing for V2I links. This is because the IA-GNN model can

effectively adjust the IRS reflection coefficients to improve the Sum-Rate as the number of IRS elements increases. The design complexity of IRS beamforming increases when the number of IRS is larger, and the IRS design is mainly based on the obtained CSI. However, it is more difficult for V2I links to accurately estimate channels when the number of IRS is larger. It can be concluded that the IA-GNN model implicitly learns the channel characteristics and subtly skips the channel estimation stage to design IRS, which has better performance in the design of IRS for high-speed mobile scenarios.

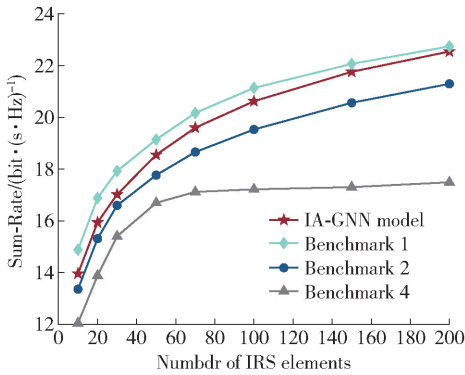


Fig. 5 Sum-Rate with different numbers of IRS elements for V2I links

Fig. 6 depicts the Sum-Rate versus the pilot transmit powers for V2I links, and the fixed pilot length $T = 160$.

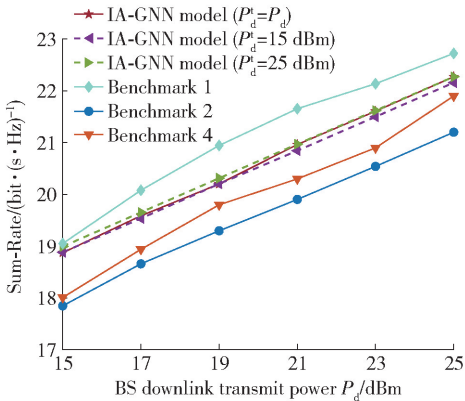


Fig. 6 Sum-Rate with different downlink transmit power for V2I links

In Fig. 6, the Sum-Rate of the IA-GNN model with training pilot transmit power $P_d^t = P_d$ is compared with the fixed training pilot transmit power $P_d^t = 15$ dBm and $P_d^t = 25$ dBm for V2I links. It can be seen that the

Sum-Rates of IA-GNN model with $P_d^t = 25$ dBm and $P_d^t = 15$ dBm approximately equal to those with $P_d^t = P_d$ using different training pilot transmit power for V2I links. It indicates that the IA-GNN model has strong generalization ability with respect to transmission power. It is because the IA-GNN can adaptively adjust network parameters to different test conditions. It can also be seen that the Sum-Rates of the IA-GNN model are significantly better than those of Benchmark 2 and Benchmark 4, and close to Benchmark 1 for V2I links. Moreover, it can be observed that the IA-GNN model is quite robust to the pilot transmit power variation for V2I links.

Fig. 7 depicts the Sum-Rate versus the number of r-VUEs for V2I links. The downlink transmit power is set as $P_d = 25$ dBm. The IA-GNN model is trained under the condition that the number of VUEs $K = 6$ for V2I links.

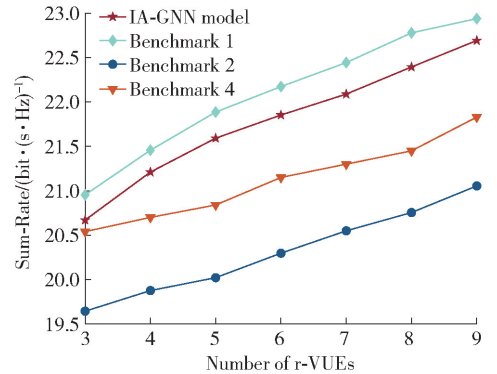


Fig. 7 Sum-Rate with different numbers of r-VUEs for V2I links

From Fig. 7, it can be seen that the IA-GNN model for V2I links is easily scaled to different numbers of VUEs. It is because the parameters of IA-GNN can be tied across the users. Moreover, the performance of the IA-GNN is consistently better than Benchmark 2 and Benchmark 4 for V2I links.

5.4 Network fairness

Although the advantages of the IA-GNN model in solving the system Sum-Rate problems of V2I are described, the Sum-Rate optimization ignores the fairness between VUEs in actual communication systems. Thus, a maximizing the minimum (Max-Min) rate problem is designed as

$$\left. \begin{aligned} & \max_{f(Y)=\mathbf{W}, \Phi} E\{\min_k R_k(\mathbf{W}, \Phi)\} \\ & \text{s. t.} \\ & \sum_{k=1}^K \|\mathbf{w}_k\|^2 \leq P_d; \quad 0 \leq \varphi_n < 2\pi, \quad n = 1, 2, \dots, N \end{aligned} \right\} \quad (37)$$

The expectation is to maximize the minimum rate of a single VUE in the V2I links for the Max-Min problem in Eq. (37), so, the IA-GNN model is used to train, and the loss function is set to $-E\{\min_k R_k(\mathbf{W}, \Phi)\}$. It can be noted that the minimum VUE rate function is differentiable with respect to $\mathbf{W}$ and Φ . Thus, the gradient backpropagation method can be applied for training the IA-GNN model. The number of IRS reflection elements is selected as $N = 30$, because the calculation complexity of semidefinite relaxation is high^[27], which is used to design passive beamforming in Benchmark 2.

In Table 1, the IA-GNN model is trained with a fixed r-VUE number and pilot transmit power, i. e., $K = 4$, $P_d = 20$ dBm. It shows that the Max-Min rate of IA-GNN model with $K = 4$ can reach 97.5% of Benchmark 1 when $T = 60K$ and it can still reach 96.0% of Benchmark 1 with $K = 3$. It implies that the IA-GNN model generalizes well to different VUEs when solving optimization problems in Eq. (37). It can be observed that the Max-Min rate of the IA-GNN model used $T = 20K$ can achieve 98.2% of Benchmark 2 with $T = 60K$, and it can achieve 96.9% of Benchmark 2 when $K = 3$. It can be concluded that the IA-GNN model can still maintain the advantages of low overhead and high reliability in Max-Min rate. Moreover, the generalization of the IA-GNN model performs well with different numbers of VUEs.

Table 1 Verification of Max-Min rate of r-VUE

Pilot length T	Number of r-VUEs	Max-Min rate		
		IA-GNN model	Benchmark 2	Benchmark 1
20K	3	6.754	6.340	7.664
	4	5.418	4.889	5.942
40K	3	7.197	6.770	7.664
	4	5.683	5.293	5.942
60K	3	7.356	6.968	7.664
	4	5.794	5.516	5.942

6 Conclusions

In this paper, the IA-GNN model based on DL was proposed to jointly optimize active and passive beamforming bypassing explicit channel estimation for IRS-assisted mmWave V2I links. And the GAT mechanism is introduced into the IA-GNN model to further capture the connections between IRS and VUEs and make the update of nodes' feature more comprehensive for IRS-assisted mmWave V2I links. The pilot transmission strategy is proposed to generate training pilot data for the IA-GNN model. Moreover, the joint active and passive beamforming optimization schemes based on the IA-GNN model for different utility functions of IRS-assisted V2I links are analyzed. And the results are compared with the traditional beamforming optimization schemes. It shows that the IA-GNN model can be applied for IRS-assisted multi-user V2I systems with lower pilot overhead and better generalization. The IA-GNN model does not associate with IRS energy consumption. Future research will focus on adding the IRS energy consumption model to the utility function to achieve the multi-objective joint optimization, so as to explore a more realistic IRS-assisted vehicular system.

Acknowledgements

This work was supported by the Natural Science Foundation of Hebei Province of China (F2024201053), and the S&T Program of Hebei Province (2023KFKT002).

References

- [1] PALA S, SAIKIA P, SINGH S K, et al. Design of RIS-assisted full duplex 6G-V2X communications. Proceedings of the 2023 IEEE Wireless Communications and Networking Conference (WCNC'23), 2023, Mar 26 – 20, Glasgow, UK. Piscataway, NJ, USA: IEEE, 2023: 6p.
- [2] LI C Y, FANG Y F, YU L J, et al. A 5G mmWave communication framework in vehicle-road cooperative system. Proceedings of the 8th Annual International Conference on Network and Information Systems for Computers (ICNISC'22), 2022, Sep 16 – 18, Hangzhou, China. Piscataway, NJ, USA: IEEE, 2022: 46 – 49.
- [3] PAN S, ZHANG X M, Sung D K. Intelligent reflecting surface-aided centralized scheduling for mmWave V2V networks. Proceedings of the 2022 International Conference on Computer Communications and Networks (ICCCN'22), 2022, Jul 25 – 27,

- Honolulu, HI, USA. Piscataway, NJ, USA: IEEE, 2022; 10p.
- [4] WU Q Q, GUAN X R, ZHANG R. Intelligent reflecting surface-aided wireless energy and information transmission; an overview. *Proceedings of the IEEE*, 2022, 110(1): 150–170.
- [5] HUANG C W, HU S, ALEXANDROPOULOS G C, et al. Holographic MIMO surfaces for 6G wireless networks; opportunities, challenges, and trends. *IEEE Wireless Communications*, 2020, 27(5): 118–125.
- [6] TAGLIAFERRI D, MIZMIZI M, AYOUBI R A, et al. Conformal intelligent reflecting surfaces for 6G V2V communications. *Proceedings of the 2022 1st International Conference on 6G Networking (6GNet'22)*, 2022, Jul 6–8, Paris, France. Piscataway, NJ, USA: IEEE, 2022; 8p.
- [7] WU W, WANG H, WANG W N, et al. Doppler mitigation method aided by reconfigurable intelligent surfaces for high-speed channels. *IEEE Wireless Communications Letters*, 2022, 11(3): 627–631.
- [8] DANG J, ZHANG Z C, WU L. Joint beamforming for intelligent reflecting surface aided wireless communication using statistical CSI. *China Communications*, 2020, 17(8): 147–157.
- [9] LI Q, WANG S L, WANG Y Y. Rank-two multicast beamforming for multiuser MISO downlink with intelligent reflecting surface. *IEEE Transactions on Vehicular Technology*, 2022, 71(6): 6722–6726.
- [10] PAPAFAEIROPOULOS A, PAN C H, KOURTESSIS P S. et al. Intelligent reflecting surface-assisted MU-MISO systems with imperfect hardware; channel estimation and beamforming design. *IEEE Transactions on Wireless Communications*, 2022, 21(3): 2077–2092.
- [11] LIN S E, ZHENG B X, Alexandropoulos G C, et al. Reconfigurable intelligent surfaces with reflection pattern modulation; beamforming design and performance analysis. *IEEE Transactions on Wireless Communication*, 2021, 20(2): 741–754.
- [12] ZHENG B X, ZHANG R. Intelligent reflecting surface-enhanced OFDM; channel estimation and reflection optimization. *IEEE Wireless Communications Letters*, 2020, 9(4): 518–522.
- [13] WU T T, ZHU L B, XU W C, et al. IRS aided beamforming design of millimeter wave vehicle networking system. *Proceedings of the 5th International Conference on Artificial Intelligence and Computer Applications (ICAICA'23)*, 2023, Nov 28–30, Dalian, China. Piscataway, NJ, USA: IEEE, 2023; 831–835.
- [14] SONG H B, ZHANG M Y, GAO J B, et al. Unsupervised learning-based joint active and passive beamforming design for reconfigurable intelligent surfaces aided wireless networks. *IEEE Communications Letters*, 2021, 25(3): 892–896.
- [15] POTU V D, SUNKU V S, MOUNIKA M, et al. A deep learning scheme for integrated active and passive beamforming in reconfigurable intelligent surface aided wireless MISO networks. *Proceedings of the 2022 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET'22)*, 2022, Mar 24–26, Chennai, India. Piscataway, NJ, USA: IEEE, 2022; 226–230.
- [16] GAO A, SUN X S, XU Y S, et al. Heterogeneous multi-agent reinforcement learning for joint active and passive beamforming in IRS assisted communications. *Proceedings of the 32nd Wireless and Optical Communications Conference (WOCC'23)*, 2023, May 5–6, Newark, NJ, USA. Piscataway, NJ, USA: IEEE, 2023; 5p.
- [17] XU X, LU Y, FU Q M. Applying graph neural network in deep reinforcement learning to optimize wireless network routing. *Proceedings of the 9th International Conference on Advanced Cloud and Big Data (CBD'22)*, 2022, Mar 26–27, Xi'an, China. Piscataway, NJ, USA: IEEE, 2022; 218–223.
- [18] RAHMAN M M, RAHMAN M S. Performance analysis of classical and graph neural network-based power allocation schemes for single-hop Ad hoc wireless networks. *Proceedings of the 12th International Conference on Electrical and Computer Engineering (ICECE'22)*, 2022, Dec 21–23, Dhaka, Bangladesh. Piscataway, NJ, USA: IEEE, 2022; 449–452.
- [19] EISEN M, RIBEIRO A. Optimal wireless resource allocation with random edge graph neural networks. *IEEE Transactions on Signal Processing*, 2020, 68: 2977–2991.
- [20] CHEN J, LIANG Y C, CHENG H V, et al. Channel estimation for reconfigurable intelligent surface aided multi-user mmWave MIMO systems. *IEEE Transactions on Wireless Communications*, 2023, 22(10): 6853–6869.
- [21] VELIKOVI P, CUCURULL G, CASANOVA A, et al. Graph attention networks. *Proceedings of the 6th International Conference on Learning Representations (ICLR'18)*, 2018, Apr 30–May 3, Vancouver, Canada. 2018; 12p.
- [22] ZHU J Y, HUANG Y M, WANG J H, et al. Power efficient IRS-assisted NOMA. *IEEE Transactions on Communications*, 2021, 69(2): 900–913.
- [23] LIU C, LIU X M, WEI Z Q, et al. Deep learning-empowered predictive beamforming for IRS-assisted multi-user communications. *Proceedings of the 2021 IEEE Global Communications Conference (GLOBECOM'21)*, 2021, Dec 7–11, Madrid, Spain. Piscataway, NJ, USA: IEEE, 2023; 7p.
- [24] PEKKA K, JUHA M, LASSI H, et al. WINNER II channel models. *IST-4-027756 WINNER II D1.1.2 V1.2*, 2008.
- [25] GUO H Y, LIANG Y C, CHEN J, et al. Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks. *IEEE Transactions on Wireless Communications*, 2020, 19(5): 3064–3076.
- [26] KYAW K T, SANTIPACH W, MAMAT K, et al. Optimization of transmit beamforming using channel covariances for MISO downlink assisted by reconfigurable intelligent surfaces. *Proceedings of the 21st International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON'24)*, 2024, May 27–30, Khon Kaen, Thailand. Piscataway, NJ, USA: IEEE, 2024; 6p.
- [27] NADEEM Q U A, ALWAZANI H, KAMMOUN A, et al. Intelligent reflecting surface-assisted multi-user MISO communication; channel estimation and beamforming design. *IEEE Open Journal of the Communications Society*, 2020, 1: 661–680.

(Editor: Wang Xuying)