Real and Fake Channel: GAN-based Wireless Channel Modeling and Generating

Wenwu Xie[a], Ming Xiong[a], Zhihe Yang[a], Wei Liu[b], Lvrong Fan[b] and Jian Zou[c,d]

[a] *School of Information Science and Engineering, Hunan Institute of Science and Technology, Yueyang 414006, China*
[b] *College of Physical Science and Technology, Central China Normal University Hubei, Wuhan 430079, China*
[c] *College of Applied Technology, Shenzhen University, Shenzhen 518061, China*
[d] *Graduate School, Shenzhen Technology University, Shenzhen 518118, China*

ARTICLE INFO

Keywords:
Multiple-input multiple-output (MIMO)
channel modeling
deep learning
generative adversarial network (GAN)

ABSTRACT

With the increasing complexity of wireless environment, scene, frequency band, antenna scale and other factors, it brings new challenges to wireless channel modeling. On the other hand, wireless communication solutions based on artificial intelligence have been constantly proposed in recent years, which are highly dependent on the quality and quantity of channel data. However, in the actual communication system, the acquisition of real channel data is restricted by the high cost of channel data acquisition. In this paper, an improved wasserstein generative adversal network with gradient penalty (WGAN-GP) is proposed to solve the channel modeling and generation problem in MIMO communication system. The experimental results show that the improved WGAN-GP can generate fake channel data with the same distribution as the real channel data. In addition, the validity of the fake channel data is cross-verified by a channel feedback scheme based on deep learning.

1. Introduction

The 5th generation (5G) of mobile communication technology to realize the Internet of everything has gradually begun large-scale commercial use, and the sixth mobile communication technology has also started exploration and research. In order to realize the application scenarios and key performance indicators of 6th generation (6G), some key wireless frontier technologies have been introduced [1, 2]. This makes the wireless environment, scene, frequency band, antenna scale and other factors more and more complex, which brings great challenges to the modeling and generation of wireless channel. At the same time, the organic combination of artificial intelligence technology and wireless communication, namely intelligent communication, has become the focus of future communication research. However, wireless communication solutions based on artificial intelligence technology are built on a large amount of high-quality channel data, and the high cost of channel data acquisition hinders the practical application of intelligent communication [3, 4]. Therefore, how to construct wireless channel data quickly and efficiently at a small cost is another difficult problem faced by wireless communication.

In [5], a multi-input multi-output (MIMO) channel characterization and modeling method based on principal component analysis (PCA) is proposed. This method can extract some hidden features and structures from the measured channel data, and combine the detailed information of the measured environment and antenna configuration to effectively reconstruct the amplitude and phase of the channel impulse response (CIR). In [6], the spatial correlation and channel capacity of two-dimensional and three-dimensional MIMO channel models are compared. The results demonstrate the importance and challenge of channel modeling in the performance evaluation of 3D MIMO systems. In [7], some clustering techniques and big data algorithms are analyzed and applied to the clustering channel modeling of sea volume measurement data. Traditional channel modeling methods, such as ray-tracing and geometry-based stochastic channel models, require in-depth domain specific knowledge and technical expertise in radio signal propagation across electromagnetic fields. Moreover, the channel model is very complex and has many parameters, it is not suitable to predict the statistical characteristics of wireless channel, and it is not flexible and cannot be transferred to other communication environment.

In recent years, the generative adversarial network (GAN) in deep learning (DL) has been effectively applied to the generation of high-dimensional images [8], and the channel data is similar to one high-dimensional image after another. The generator network is used to extract the channel data distribution, and the discriminator network is used to determine the authenticity and falsity of the generated channel data. The generator and discriminator are opposed to each other during training, and finally the generated channel data distribution is infinitely close to the real channel distribution. In [9], generative adversarial network is used for the first time to solve the wireless channel modeling problem without complex theoretical analysis or data processing, and its good performance and effectiveness are verified. In [10], an end-to-end wireless communication system is proposed based on deep learning, and conditional GAN is adopted to conduct channel modeling in a data-driven way, providing a new solution for the construction of data-driven end-to-end deep neural network. Aiming at the channel modeling problem of RIS communication system, literature [11]

* Corresponding author
zoujian250@gmail.com (J. Zou)
ORCID(s):

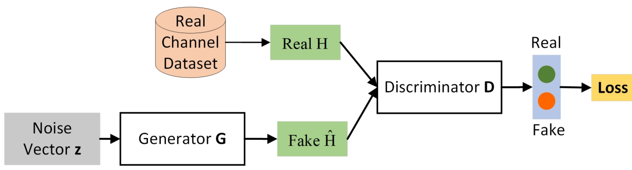


Figure 1: GAN Architecture.

proposes a channel modeling framework based on model-driven generative adversarial network (GAN). Aiming at the modeling problem of link-level MIMO channel of 3rd generation partnerships project, literature [12] proposes the ChannelGAN modeling scheme and cross-verifies the availability of the generated data.

To reduce the cost of data acquisition and complex channel modeling, a simple channel modeling scheme based on WGAN-GP is proposed without complex theoretical analysis or data processing. Similarity and diversity indexes, normalized channel power spectrum and CSI feedback based on deep learning are used to verify the cross-performance of channel modeling. Experiments show that the improved WGAN-GP in this paper has better performance than the existing ChannelGAN, and the network complexity and number of parameters are lower than ChannelGAN.

2. System Model

A classic massive MIMO-OFDM system is considered in this letter, wherein, the channel model adopts cluster delay line (CDL) model in 3GPP TR 38.901[13], and the CDL model represents the spatial characteristics of the channel model by adding azimuth of arrival angle, azimuth of departure angle, zenith of arrival angle and zenith of departure angle. N_t antennas are set at the BS, N_r antennas are set at the UE, and the number of delay extensions is set at N_d. Therefore, the full CSI data can be expressed as $\mathbf{H} \in \mathbb{R}^{N_t \times N_r \times N_d \times 2}$. The dimension 2 represents the real and imaginary parts of the complex channel data.

3. Channel Modeling and Generating

This section introduces the channel modeling and generation scheme based on improved WGAN-GP, which is mainly divided into three parts: the principle of improved WGAN-GP, the network structure of improved WGAN-GP and the algorithm implementation process of improved WGAN-GP.

3.1. Preliminaries of WGAN-GP

The architecture of a GAN is shown in Fig.1, where $\mathbf{H}$ is the known real channel data, $\hat{\mathbf{H}}$ is the fake channel data generated by noise vector $\mathbf{z}$ through the generator. Assuming $\mathbf{H} \sim p_1$ and $\hat{\mathbf{H}} \sim p_2$, the goal is to generate fake channel data similar to real channel data. In this paper, wasserstein distance is used to calculate the direct distance between two distributions, which can be obtained according to [14]

$$\begin{aligned} W\left(p_1, p_2\right) &= \inf_{\pi \in \Pi\left(p_1, p_2\right)} \mathrm{E}_{(\mathbf{H}, \hat{\mathbf{H}}) \sim \pi}\left[\left\|\mathbf{H}-\hat{\mathbf{H}}\right\|\right] \\ &= \sup_{D \in \mathrm{T}} \mathrm{E}_{\mathbf{H} \sim p_1}\left[D(\mathbf{H})\right]-\mathrm{E}_{\hat{\mathbf{H}} \sim p_2}\left[D\left(\hat{\mathbf{H}}\right)\right], \end{aligned} \tag{1}$$

where $\Pi\left(p_{1,} p_2\right)$ denotes the set of all joint distributions $\pi\left(\mathbf{H}, \hat{\mathbf{H}}\right)$ whose marginals are p_1 and p_2, respectively. T is the set of 1-Lipschitz functions, $D(\bullet)$ represents the discriminator, and $\hat{\mathbf{H}} = G(\mathbf{z})$.

To generate realistic enough fake channel data, then minimize $W\left(p_1, p_2\right)$. According to Kantorovich-Rubinstein duality [15], we can get

$$\min_G \max_{D \in \mathrm{T}} \mathrm{E}_{\mathbf{H} \sim p_1}\left[D(\mathbf{H})\right]-\mathrm{E}_{\mathbf{z} \sim p_z}\left[D(G(\mathbf{z}))\right], \tag{2}$$

where $G(\bullet)$ represents the generator, p_z follows the standard normal distribution.

By introducing Wasserstein distance and Lipschitz function, WGAN solves the problems of difficult training, gradient disappearance and model instability, but the introduction of Lipschitz function also brings gradient explosion problem. According to [14], this shortcoming is effectively avoided by adding gradient penalty to the discriminator loss function. Then, we can obtain

$$\begin{aligned} \min_G \max_D \underset{\mathbf{H} \sim p_1}{\mathrm{E}}\left[D(\mathbf{H})\right] - \underset{G(\mathbf{z}) \sim p_2}{\mathrm{E}}\left[D(G(\mathbf{z}))\right] \\ +\lambda \underset{\bar{\mathbf{H}} \sim p_3}{\mathrm{E}}\left[\left(\left\|\nabla_{\bar{\mathbf{H}}} D(\bar{\mathbf{H}})\right\|_2-1\right)^2\right], \end{aligned} \tag{3}$$

where $\bar{\mathbf{H}} = \alpha \mathbf{H} + (1-\alpha) G(\mathbf{z})$. p_3 is defined as uniformly sampling along a line between pairs of points sampled from p_1 and p_2. a random number $\alpha \sim U[0,1]$, λ is penalty parameter, $\|\bullet\|_2$ denotes the l_2 norm.

3.2. Improved WGAN-GP Structure

The improved WGAN-GP structure is shown in Fig. 2. The network structure consists of generator (G) and discriminator (D). The generator is mainly used to generate fake channel data, and the discriminator is mainly used to judge the real and fake channel data.

(1) *Generator.* The generator consists of a LinearBlock and five ConvTBlocks, in which hidden dimension of the linear layer in the LinearBlock is 1024. The ConvTranspose2d filters in the five ConvTBlocks are successively (1024, 512, 256, 128, 1), and the kernel size and stride are respectively 2, 1. The activation function of the remaining ConvTBlocks is LeakyReLU except for the last ConvTBlock, which is Tanh.

(2) *Discriminator.* The discriminator consists of five ConvBlocks and a Linear layer. Among the five ConvBlocks, the filters of Conv2d are successively (128, 256, 512, 1024, 128), and the kernel size, stride and padding are 3, 2, 1, respectively. The hidden dimension of linear layer is 1.

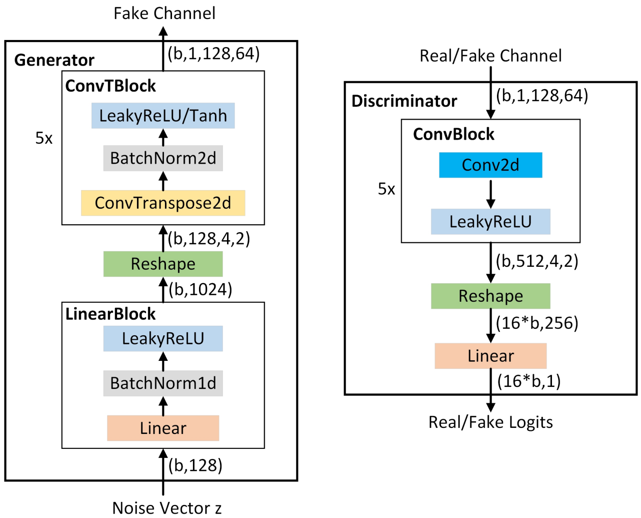


Figure 2: Improved WGAN-GP structure.

Through continuous training of the generator and discriminator, the generator can continuously improve its ability to generate fake channel data similar to real channel data, and the discriminator can continuously improve its ability to distinguish real channel data and fake channel data. In the ideal case, the generator and discriminator reach the optimal solution through the mechanism of cooperation and confrontation, and generate the best fake channel data.

3.3. Improved WGAN-GP Algorithm

In Algorithm 1, the process of wireless channel modeling and generating based on Improved WGAN-GP is explained in detail. The algorithm is divided into two stages: online training and offline generation. In the online training stage, Adam optimizer is used to train the generator $G_\xi(\bullet)$ and discriminator $D_\eta(\bullet)$ of optimal model parameters. The online generation phase uses the optimal parameter generator $G_\xi(\bullet)$ model to generate fake channel data. Where, $-\frac{1}{b}\sum_{k=1}^{b} D_\xi\left(G_\eta\left(z^{(k)}\right)\right)$ represents the loss function of generator and $\frac{1}{b}\sum_{k=1}^{b} L^{(k)}$ represents the loss function of discriminant.

Algorithm 1: Improved WGAN-GP Algorithm

Offline Training:

Initialization: The discriminator and generator parameters are η and ξ, respectively. The adam hyperparameters $lr = 2\mathrm{e}^{-4}$, $\beta_1 = 0.5$, $\beta_2 = 0.9$. The gradient penalty coefficient $\lambda = 10$.

Input: The number of discriminator iterations per generator iteration $n_{critic} = 3$. The epochs and batch size are e and b, respectively.

for $i \leftarrow 1$ **to** e **do**
  for $j \leftarrow 1$ **to** n_{critic} **do**
    for $k \leftarrow 1$ **to** b **do**
      Randomly sample $\mathbf{H}'$ from real channel $\mathbf{H}$, noise vector $\mathbf{z} \sim p_z$, random number $\alpha \sim U[0,1]$.
      $\hat{\mathbf{H}}' \leftarrow G_\xi(\mathbf{z})$
      $\bar{\mathbf{H}} \leftarrow \alpha\mathbf{H}' + \left(1 - \alpha\hat{\mathbf{H}}'\right)$
      $gp \leftarrow \lambda\left(\left\|\nabla_{\bar{\mathbf{H}}} D_\eta\left(\bar{\mathbf{H}}\right)\right\|_2 - 1\right)^2$
      $L^{(k)} \leftarrow D_\eta\left(\hat{\mathbf{H}}'\right) - D_\eta\left(\mathbf{H}'\right) + gp$
    end
    $\eta \leftarrow Adam\left(\nabla_\eta \frac{1}{b}\sum_{k=1}^{b} L^{(k)}, \eta, lr, \beta_1, \beta_2\right)$
  end
  Sample a batch of noise vector $\left\{\mathbf{z}^{(k)}\right\}_{k=1}^{b} \sim p_z$.
  $\xi \leftarrow Adam\left(\frac{-\nabla_\xi}{b}\sum_{k=1}^{b} D_\eta\left(G_\xi\left(\mathbf{z}^{(k)}\right)\right), \xi, lr, \beta_1, \beta_2\right)$
end
output: $G_\xi(\bullet)$, $D_\eta(\bullet)$

Online Generating:

Input: The number of the test set N_s.
Sample N_s noise vector $\left\{\mathbf{z}^{(k)}\right\}_{k=1}^{N_s} \sim p_z$.
$\hat{\mathbf{H}} \leftarrow G_\xi\left(\mathbf{z}^{(k)}\right)$
output: The fake channel $\hat{\mathbf{H}}$.

4. Simulation Results

In this section, extensive simulation experiments are provided to verify the effectiveness of the proposed algorithm. Firstly, the consistency performance of real channel data and fake channel data is demonstrated numerically by using similarity and diversity indexes. Secondly, the normalized power of real channel data and fake channel data is visually demonstrated and compared. Thirdly, a channel feedback scheme based on deep learning is used to cross-verify the performance of fake channel data in AI algorithm. Finally, the complexity of the deep learning algorithm is analyzed. Table 1 shows the setting of relevant simulation parameters. Wireless-Intelligence is a public channel dataset library, experimental data from Wireless-Intelligence [16].

4.1. Similarity and Diversity Analysis

To evaluate the similarity between the channel data distribution generated by the network and the real channel data distribution, the similarity can be expressed as

$$\mathrm{sim} = \frac{1}{N_{fake}}\sum_{j=1}^{N_{fake}} \max_{1\le i\le N_{real}} \frac{\left\|\hat{\mathbf{H}}_j^{\mathrm{H}}\mathbf{H}_i\right\|^2}{\left\|\hat{\mathbf{H}}_j\right\|^2\left\|\mathbf{H}_i\right\|^2}, \tag{4}$$

where, N_{fake} is the number of generated channel samples, N_{real} is the number of real channel samples, and $\hat{\mathbf{H}}_j$ and $\mathbf{H}_i$ are vectorized channels of the j-th generated channel sample and the i-th real channel sample respectively.

In addition, the variance is used to calculate the diversity of generated channel data, and the diversity is expressed as

$$\mathrm{multi} = \sqrt{Var(\mathbf{I})}, \tag{5}$$

Table 1
Simulation parameter setting

Parameter	Value
Carrier frequency	3.5GHz
Bandwidth	10MHz
Subcarrier spacing	15KHz
Tx antennas N_t	32
Rx antennas N_r	4
Delay paths N_d	32
Subband number N_{sb}	12
Learning Rate lr	2e-4
Batch Size b	4
Epochs e	500
Number of training set	500
Number of test set N_{real}	10000
Number of fake channel set N_{fake}	10000

Table 2
Similarity And Diversity Of Different Channel Modeling And Generating Schemes

Schemes / sim and multi	ChannelGAN	Improved WGAN-GP
sim	0.403314	0.420951
multi	2.038140	1.918272
multi/sim	5.053481	4.557001

where $Var(\cdot)$ represents the calculation of variance against vector $\mathbf{l}$, and each element $\mathbf{l}_{i_0}$ in $\mathbf{l} = \left[\mathbf{l}_1, \mathbf{l}_2, ..., \mathbf{l}_{N_{real}}\right]$ is represented by

$$\mathbf{l}_{i_0} = N\left(\underset{1\leq i\leq N_{real}}{\arg\max} \frac{\left\|\hat{\mathbf{H}}_j^{\mathrm{H}}\mathbf{H}_i\right\|^2}{\left\|\hat{\mathbf{H}}_j\right\|^2\left\|\mathbf{H}_i\right\|^2} = i_0\right), \tag{6}$$

where $\arg\max$ represents the real channel sample index corresponding to the maximization of the above equation, and $N(\cdot)$ represents the number of times that the real channel sample index corresponding to i_0 is taken for all N_{fake} generated channel samples.

Table 2 shows the numerical results of similarity and diversity between fake channel data and real channel data when the Improved WGAN-GP and ChannelGAN [10] model is trained to the optimal parameters. When the similarity between fake channel data and real channel data is greater, the diversity is smaller, that is, the smaller multi/sim is, the distribution of generated fake channel data is closer to that of real channel data. According to the numerical results in Table 2, the scheme proposed in this paper is superior to ChannelGAN in similarity and diversity indexes, indicating that the fake channel generated by the scheme proposed in this paper is closer to the real channel.

4.2. Comparison of Real and Fake Channels

Fig. 3 intuitively shows the normalized power spectrum comparison between real channel data and fake channel data in the time delay domain and the antenna domain. Fig. 3 (a), (b) and (c) are respectively drawn from the average 1000 real channel data, the average 1000 fake channel data generated by improved WGAN-GP and the average 1000 fake channel data generated by ChannelGAN. By comparing Fig. 3 (a), (b) and (c), it can be found that the improved WGAN-GP and ChannelGAN modeling and generation schemes can generate fake channel data with the same distribution as the real channel data. However, a careful observation of the power spectrum of the time delay domain between 12 and 32 shows that the graphs drawn by real channel data are very smooth, while those drawn by fake channel data are all somewhat smooth, and the graphs drawn by fake channel data generated by ChannelGAN are the least smooth. It not only verifies the effectiveness of the improved WGAN-GP algorithm, but also verifies the correctness of the similarity and diversity indexes.

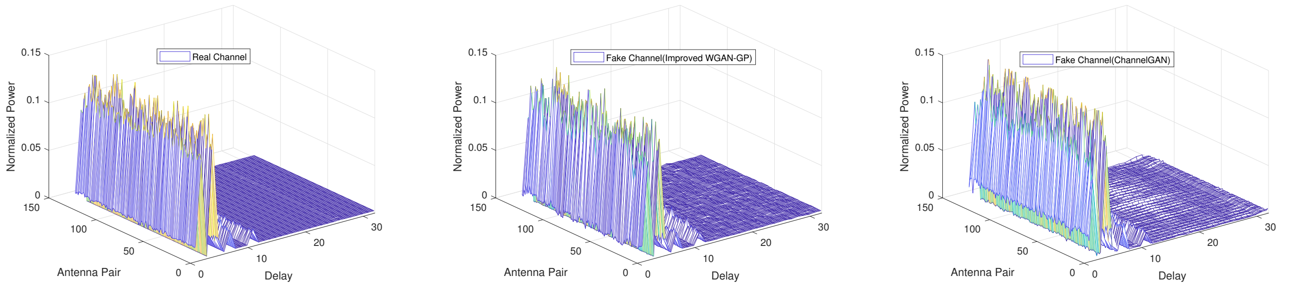
4.3. Cross Validation on DL-Based CSI Feedback

In order to verify the performance of GAN-based fake channel data in deep learning algorithm, this paper uses deep learning-based channel feedback to cross-verify the validity of fake channel data. In recent years, Transformer network has excellent performance in solving the channel feedback problem with its multi-head attention mechanism[17-19]. Therefore, the transformer network is selected as the backbone network of the channel feedback architecture in this paper, and the feedback bit is 128. The network parameters of TransformerEncoder in Encoder and Decoder are d_{model} = 192, $nhead$ = 8, $dim_{feedforward}$ = 256, $dropout$ = 0.0, and $nlayers$ = 3, Fig. 4 in the appendix shows the channel feedback network architecture based on Transformer.. In the data preprocessing part, the channel eigenvector $\mathbf{W} = \left[\mathbf{w}_1, \cdots, \mathbf{w}_{N_{sb}}\right] \in \mathbb{C}^{N_t\times N_{sb}}$ for N_{sb} subbands within 48 resource blocks is calculated by N_{FFT} = 1024 points fast Fourier transform (FFT) and eigenvalue decomposition.

The square of generalized cosine similarity (GCS) is used to evaluate the channel feedback performance based on deep learning. The larger the GCS value is, the better the channel feedback performance is, which can be obtained according to [11]

$$\rho\left(\mathbf{W}, \tilde{\mathbf{W}}\right) = \frac{1}{N_{sp}N_{sb}}\sum_{i=1}^{N_{sp}}\sum_{k=1}^{N_{sb}}\left(\frac{\left\|\mathbf{w}_{k,i}^{\mathrm{H}}\tilde{\mathbf{w}}_{k,i}\right\|_2}{\left\|\mathbf{w}_{k,i}\right\|_2\left\|\tilde{\mathbf{w}}_{k,i}\right\|_2}\right)^2, \tag{7}$$

where $\tilde{\mathbf{W}} = T(\mathbf{W})$ a represents CSI using Transformer network $T(\bullet)$ for channel feedback scheme reconstruction, N_{sp} represents the size of the test sample set, $\mathbf{w}_{k,i}$ and $\tilde{\mathbf{w}}_{k,i}$ respectively represent the corresponding eigenvector in the ideal case and the corresponding eigenvector in the feedback recovery.



(a) Normalized power spectrum of real channel (b) Normalized power spectrum of the fake channel data generated by the improved WGAN-GP (c) Normalized power spectrum of the fake channel data generated by the ChannelGAN

Figure 3: Normalized power spectrum comparison of real channel data and fake channel data in the time delay domain and antenna domain

Table 3
COMPARISON DL-Based CSI Feedback Performance

N_s	Improved WGAN-GP			ChannelGAN	
	$\rho\left(W_r, T_r\left(W_r\right)\right)$	$\rho\left(W_{f1}, T_r\left(W_{f1}\right)\right)$	$\rho\left(W_r, T_{f1}\left(W_r\right)\right)$	$\rho\left(W_{f2}, T_r\left(W_{f2}\right)\right)$	$\rho\left(W_r, T_{f2}\left(W_r\right)\right)$
500	0.665418	0.622264	0.658737	0.594065	0.628852
1000	0.736861	0.695380	0.734362	0.665387	0.703048
2000	0.761418	0.725054	0.75561	0.691617	0.725700
5000	0.809458	0.762492	0.800443	0.726111	0.761734

Note: $\mathbf{W}_r$, $\mathbf{W}_{f1}$ and $\mathbf{W}_{f2}$ are the corresponding eigenvector CSI obtained from the real channel, the fake channel generated by the improved WGAN-GP and the fake channel generated by ChannelGAN, respectively. $T_r(\bullet)$, $T_{f1}(\bullet)$ and $T_{f2}(\bullet)$ respectively represent Transformer network training models in $\mathbf{W}_r$, $\mathbf{W}_{f1}$ and $\mathbf{W}_{f2}$ data sets.

Table 3 shows the performance numerical results of real channel data and fake channel data in CSI feedback based on deep learning. N_s represents the training data set size of CSI feedback, and the size of both the test and validation sets is 1000. As can be seen from the table, for the fake channel data generated by the improved WGAN-GP, when $N_s = 5000$, compared with the GCS performance of $\rho\left(W_r, T_r\left(W_r\right)\right)$, the GCS performance of $\rho\left(W_{f1}, T_r\left(W_{f1}\right)\right)$ is about 5.8% lower, and compared with the GCS performance of $\rho\left(W_r, T_r\left(W_r\right)\right)$, the GCS performance of $\rho\left(W_r, T_{f1}\left(W_r\right)\right)$ is about 1.1% lower, indicating that the characteristics of true and fake CSI in the test set are very similar. By comparing the two channel modeling and generation schemes of the improved WGAN-GP and ChannelGAN, the fake channel data generated by the improved WGAN-GP has better GCS performance in CSI feedback based on deep learning.

4.4. Complexity Analyses

Floating point operations (FLOPs) and the number of parameters (Params) are usually used as indicators to evaluate the complexity of the network model. Table 4 compares the FLOPs and Params of the improved WGAN-GP and ChannelGAN network models. It can be seen from the table that the complexity of both generator and discriminator of the improved WGAN-GP is lower than that of ChannelGAN.

Table 4
Flops And Params Of Different Channel Modeling Schemes

Schemes	Improved WGAN-GP		ChannelGAN	
	Generator	Discriminator	Generator	Discriminator
FLOPS($\times 10^9$)	0.828	0.494	3.950	1.047
Params($\times 10^7$)	0.342	1.092	1.146	3.032

Note: The training and testing of the network model are all in the Pytorch framework. During the calculation of FLOPs and Params, the data dimension of the input of the generator network is (1, 128) and that of the input of the discriminator network is (1, 1, 128, 64).

5. Conclusion

Considering classical MIMO-OFDM systems, an improved WGAN-GP algorithm is proposed for wireless channel modeling and generation. In addition, the similarity and diversity indexes are used to achieve numerical channel modeling and generate algorithm performance, and the normalized power of real channel data and fake channel data is compared. Finally, the effectiveness of GAN-based channel modeling and generation is cross-verified by a deep learning-based channel feedback scheme.

6. Appendix

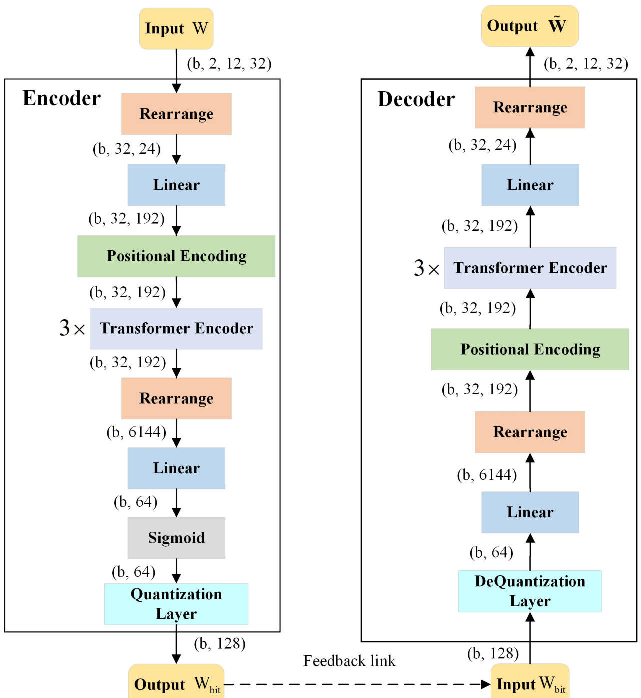


Figure 4: Channel feedback network structure based on Transformer.

References

[1] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland and F. Tufvesson, "6G Wireless Systems: Vision, Requirements, Challenges, Insights, and Opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166-1199, July 2021.

[2] H. Zhang, W. Song, X. Liu, M. Sheng, K. Long, and O. A. Dobre, "Intelligent Channel Prediction and Power Adaptation in LEO Constellation for 6G," *IEEE Network*, 2023.

[3] W. Xie, J. Xiao, P. Zhu, C. Yu and L. Yang, "Deep Compressed Sensing-Based Cascaded Channel Estimation for RIS-Aided Communication Systems," *IEEE Wireless Commun. Lett.*, vol. 11, no. 4, pp. 846-850, April 2022.

[4] W. Xie, J. Xiao, P. Zhu and C. Yu, "Multi-Task Learning-Based Channel Estimation for RIS Assisted Multi-User Communication Systems," *IEEE Commun. Lett.*, vol. 26, no. 3, pp. 577-581, March 2022.

[5] X. Ma et al., "A PCA-Based Modelling Method for Wireless MIMO Channel," *Proc. 2017 IEEE Conf. Computer Commun. Wksps.*, Atlanta, GA, 2017, pp. 874–79.

[6] Y. Yu et al., "3D vs. 2D Channel Models: Spatial Correlation and Channel Capacity Comparison and Analysis," *Proc. ICC 2017*, Paris, France, 2017, pp. 1–7.

[7] R. He et al., "Clustering Enabled Wireless Channel Modeling Using Big Data Algorithms," *IEEE Commun. Mag.*, no. 5, May 2018, pp. 177–83.

[8] I. Goodfellow et al., "Generative Adversarial Nets," *Proc. 2014 Conf. Neural Info. Processing Systems*, Montreal, Canada, Dec. 2014, pp. 2672–80.

[9] Y. Yang, Y. Li, W. Zhang, F. Qin, P. Zhu, and C.-X. Wang, "Generative-adversarial-network-based wireless channel modeling: Challenges and opportunities," *IEEE Commun. Mag.*, vol. 57, no. 3, pp. 22–27, Mar. 2019.

[10] F. A. Aoudia and J. Hoydis, "End-to-End Learning of Communications Systems Without a Channel Model," 2018 52nd Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2018, pp. 298-303, Oct. 2021.

[11] Y. Wei, M. -M. Zhao and M. -J. Zhao, "Channel Distribution Learning: Model-Driven GAN-Based Channel Modeling for IRS-Aided Wireless Communication," *IEEE Trans. Commun.*, vol. 70, no. 7, pp. 4482-4497, July 2022.

[12] H. Xiao, W. Tian, W. Liu and J. Shen, "ChannelGAN: Deep Learning-Based Channel Modeling and Generating," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 650-654, March 2022.

[13] 3GPP, "3GPP TR 38.901 v16.1.0 3rd generation partnership project; technical specification group radio access network; study on channel model for frequencies from 0.5 to 100 GHz (release 16)," Tech Rep, 2020.

[14] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein gan," 2017, arXiv:1701.07875.

[15] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of Wasserstein GANs," 2017, arXiv:1704.00028.

[16] OPPO. *Wireless-Intelligence*. Accessed: Sep. 22, 2022. [Online]. Available: https://wireless-intelligence.com/

[17] W. Xie, J. Zou, J. Xiao, M. Li and X. Peng, "Quan-Transformer based Channel Feedback for RIS-Aided Wireless Communication Systems," *IEEE Commun. Lett.*, 2022.

[18] Y. Xu, M. Yuan, and M.-O. Pun, "Transformer empowered CSI feedback for massive MIMO systems," in *Proc. 30th Wireless Opt. Commun. Conf. (WOCC)*, Oct. 2021, pp. 157–161.

[19] Y . Cui, A. Guo, and C. Song, "TransNet: Full attention network for CSI feedback in FDD massive MIMO system," *IEEE Wireless Commun*. Lett., vol. 11, no. 5, pp. 903–907, May 2022.