IEICE TRANSACTIONS

on Fundamentals of Electronics, Communications and Computer Sciences

DOI:10.1587/transfun.2024EAP1062

Publicized:2024/08/05

This advance publication article will be replaced by the finalized version after proofreading.

A PUBLICATION OF THE ENGINEERING SCIENCES SOCIETY

The Institute of Electronics, Information and Communication Engineers

Kikai-Shinko-Kaikan Bldg., 5-8, Shibakoen 3 chome, Minato-ku, TOKYO, 105-0011 JAPAN

PAPER

Multidimensional Tensor-Aware GAN based Pseudo Measurement Data Deduction in IoT-Empowered Distribution Station

Jie REN[†], **Minglin LIU**[†], **Lisheng LI**[††], **Shuai LI**[††a)], **Mu FANG**[†], **Wenbin LIU**[††], **Yang LIU**[††], **Haidong YU**[††], *and* **Shidong ZHANG**[††], ***Nonmembers***

SUMMARY The distribution station serves as a foundational component for managing the power system. However, there are missing data in the areas without collection devices due to the limitation of device deployment, leading to an adverse impact on the real-time and precise monitoring of distribution stations. The problem of missing data can be solved by the pseudo measurement data deduction method. Traditional pseudo measurement data deduction methods overlook the temporal and contextual correlations of distribution station data, resulting in a lower restoration accuracy. Motivated by the above challenges, this paper proposes a novel pseudo measurement data deduction model for minimal data collection requirements in distribution stations. Compared to the traditional GAN, the proposed enhanced GAN improves the architecture by decomposing the input tensor of the generator, allowing it to handle high-dimensional and intricate data. Furthermore, we enhance the loss function to accelerate the model's convergence speed. Our proposed approach allows GAN to be trained within a supervised environment, effectively enhancing the accuracy of model training. The simulation result shows that the proposed algorithm achieves better performances compared with existing methods.

key words: *IoT, generative adversarial network, pseudo measurement data deduction, distribution station, multidimensional tensor awareness*

1. Introduction

The distribution station is a crucial component of the power system that maintains close contact with the users[1], [2]. It is widely distributed, equipped with numerous devices, and possesses a complex structure. Simultaneously, the distribution station plays a pivotal role in power transmission, distribution, and control, ensuring the utmost reliability, stability, and safety of power supply [3], [4]. The integration of internet of things (IoT) and distribution station can maximize the distribution network's overall efficiency by real-time data collection and data processing. Consequently, it is imperative to collect various data through IoT devices such as sensors, directing the operation and management of the network [5]. Due to minimal data collection requirements in the low-voltage distribution station area, collection devices are deployed only in some areas of the station, and the data in the areas without collection devices are missing. Pseudo measurement data are artificially generated through processes such as simulation, synthesis, or other methods. These generated data typically exhibit features akin to authentic measurement data but are not derived from actual observations. Pseudo measurement data deduction can effectively solve the problem of missing data in the areas without collection devices [6]–[8].

The current methods for deducting pseudo measurement data can mainly be classified into two categories: statistical-based approaches [9] and machine learning-based approaches [10], [11]. Among them, the machine learning-based methods have gained widespread application in the field of pseudo measurement data deduction by considering the interrelationships among data[12]. One such machine learning method is the generative adversarial network (GAN), which was introduced in 2014 [13]. This network model operates without reliance on any prior assumptions and possesses the capability to learn high-dimensional and intricate data distributions in an unsupervised manner[14]. By mapping noise to the sample space through neural networks, GAN generates data that conforms to the distribution patterns of real samples. This remarkable data generation capability has made GAN a hot topic of research in recent years. Therefore, in this paper, we adopt a GAN-based approach to infer and deduce pseudo measurement data in IoT-empowered distribution station. However, despite GAN's ability to learn the temporal characteristics and interdependencies of the collected data in the station, there are still several challenges in the research of pseudo measurement data deduction in IoT-empowered distribution station.

Firstly, traditional GANs have limited capabilities in processing high-dimensional and complex data. On one hand, the inclusion of high-dimensional data increases the complexity and computational burden of the model, thereby impacting the discriminative power of the network and making the training process more challenging and time-consuming [15]. On the other hand, high-dimensional data also diminishes the generative capacity of the model, making it difficult to generate samples of high quality and diversity. Secondly, traditional GANs are susceptible to mode collapse, wherein the discriminator fails to cover all the categories present in the data distribution during training. Therefore, the generated samples from the generator tend to be overly similar or lack diversity [16]. Finally, the loss function employed by traditional GANs can lead to issues such as gradient vanishing or exploding, resulting in unstable training [17]. Moreover, the conventional loss functions for the generator and discriminator lack explicit metrics to quantify the differences between the generated completion data and

[†]The author is with State Grid Shandong Electric Power Company, Jinan, 250001, China
[††]The author is with Distribution Technology Center, State Grid Shandong Electric Power Research Institute, Jinan, 250003, China
a) E-mail: ls_19922023@163.com

the real data. Thus, GANs struggle to generate accurate completion data for the distribution station by incorporating contextual information from multidimensional data.

Several studies have investigated issues of data deduction in distribution station. In [18], Zhang *et al.* proposed a generative adversarial imputation nets-based method to achieve data deduction for the power systems. In [19], Yan *et al.* designed an implicit generative model with Wasserstein GAN objectives, namely Unbalanced Graph Generative Adversarial Network (UG-GAN), to mimic almost all features of real-world networks. However, the aforementioned article overlooks the potential issue of prolonged model training time and even convergence failure when the inputs of the GAN are high-dimensional and intricate data. In [20], Liu *et al.* proposed a completion method based on a tensor-assisted GAN to generate high-accuracy location data, the objective of which is to improve the location accuracy and storage consumption. In [21], Zhang *et al.* proposes mixed generative adversarial networks (mixed-GANs) as a practical way to provide additional data, ensuring data reliability. However, the aforementioned article fails to address the issue of mode collapse in GAN discriminators, which leads to the generation of only a singular type of data. In [22], Zhang *et al.* combined reinforcement learning with GAN and designed a novel deduction model, which aims to improve the integrity of surface deduction. In [23], Kang *et al.* proposed a novel cross-modal generative adversarial network (CM-GAN) to combine the cross-modal data fusion technique with the deep adversarial generation technique in order to construct a cross-modal data generator. However, the aforementioned article overlooks the potential issues of gradient vanishing caused by the inadequate design of the loss function.

To address the above challenges, this paper proposes a multidimensional tensor-aware GAN-based pseudo-measurement data deduction in IoT-empowered distribution station. Firstly, the measurement data tensor model in IoT-empowered distribution station and traditional GAN model are constructed. Secondly, pseudo measurement data processing is realized by abnormal data elimination. Then, the improved GAN based on tensor decomposition is formulated to create a supervised learning environment for the GAN and improve the convergence speed of the GAN, which ultimately solves deduction of pseudo measurement data in distribution station. The main contributions of this work are summarized as follows.

- ***Multidimensional tensor decomposition-based input of GAN generator:*** This paper takes the expansion matrices of multidimensional tensor, expansion matrix of mask tensor, and a set of noise vectors as inputs to the generator. The dimensionality of the input data can be efficiently reduced through tensor expansion, allowing the generator to more effectively capture the data's diversity. Additionally, leveraging the expansion matrix of the mask tensor aids the generator to focus on generating the portion of the samples that are relevant to useful information. This approach mitigates the instability often encountered by GAN when handling high-dimensional and complex data.
- ***Improved GAN discriminator:*** In the improved GAN model, the task of the discriminator is to further distinguish whether the input discriminator data come from measurement data collected by deployed collection devices or data derived from generators. The generated adversarial network can be trained in a supervised environment and improve model training accuracy.
- ***Optimization of loss function:*** We improve the generator loss function based on the multidimensional data repair loss and the multidimensional context loss to ensure that the overall maximization of the inferred data follows the true data distribution. In addition, the improved discriminator loss consists of the original loss, the gradient penalty term, and the multidimensional context loss. The multidimensional context loss is introduced to help the discriminator better understand the input samples, thus improving the discriminator's ability to discriminate the samples generated by the generator. The gradient penalty term can effectively prevent the gradient explosion problem and make GAN training more stable.

The remainder of this article is organized as follows. Section II introduces the system model. Section III presents the improved GAN based on tensor decomposition for data deduction. Simulation results are given in Section IV. Section V provides the conclusion.

2. Pseudo Measurement Data deduction Model

2.1 Measurement Data Tensor Model in Low-Voltage Distribution Station

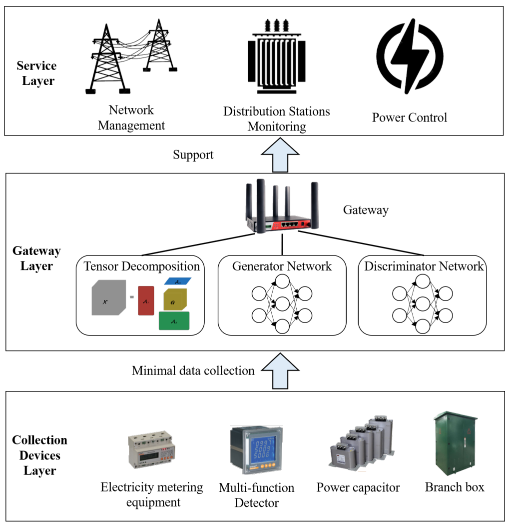


Fig. 1 The minimal data collection model for IoT-empowered distribution station area.

In IoT-empowered distribution station area, the electricity consumption data of each user are collected through collection devices and then transmitted to the gateway for management [24]. Minimal data collection means reducing collection costs and resource consumption through effective methods and strategies, while ensuring the accuracy and observability of collected data. In order to meet the minimum requirements for accurate data collection within the distribution area, IoT collection device is selectively deployed only in certain zones of the distribution station area. Therefore, it is necessary to use the measured data collected by the deployed collection devices to perform pseudo measurement data deduction for the data in the area without the deployed collection devices. The minimal data collection model for IoT-empowered distribution station area is shown in fig. 1. In order to fulfill the demand for minimal precise data collection within the distribution station area, IoT collection device is selectively deployed only in certain zones of the distribution station area. Simultaneously, the measured data collected by the deployed devices are utilized to extrapolate and supplement the data in regions where collection device is not deployed, thus facilitating a variety of computational applications within the power distribution station. This methodology facilitates the execution of diverse computational applications within the power distribution station area. The collected measurement data includes voltage, current, electrical energy, etc [25].

A tensor is a high-dimensional array, whose spatial dimension is usually referred to as the order of the tensor. Defined $\mathcal{Z}^N \in \mathbf{R}^{q_1 \times q_2 \times \cdots q_n \times \cdots \times q_N}$ as a tensor of N orders, where q_n is the n-th dimension[26]. In order to deduce missing data of the areas without collection devices, taking current data as an example, a three-dimensional tensor $\mathcal{L}^3 \in \mathbf{R}^{q_1 \times q_2 \times q_3}$ is first constructed, including current data in the areas with collection devices and missing data in the areas without collection devices. In the three-dimensional tensor, q_1 is the user information dimension, which represents the number of users in the station. q_2 is the date dimension, which represents the number of days included in each collection cycle. q_3 is the collection frequency dimension, which represents the number of times of data collected per day in the collection cycle [27]. Each element in this tensor $\mathcal{L}^3$ represents the current value of the user at each measurement time.

2.2 Traditional GAN Model

GAN possesses the capacity to acquire knowledge about the temporal patterns and correlations within the data gathered from the low-voltage distribution station area. Based on GAN, it can achieve the deduction of missing data in the areas without collection devices, supporting various computing applications in the low-voltage distribution station area. GAN consists of two neural networks. One is called the discriminator D and the other one is called the generator G[28]. The goal of the discriminator is to accurately determine whether the input samples are derived from deployed collection devices or the generator as much as possible. The generator expects its generated data to deceive the discriminator as much as possible. These two networks with opposite objectives continuously undergo alternating training, and when they finally converge, the network reaches Nash equilibrium [29], [30]. The traditional generator model takes a stochastic sample of noise as its input, while the discriminator generates a probability value indicating the likelihood that the input data originate from the measured data acquired by deployed collection devices [31]. GAN makes the generator generate deduction data tensors $\mathcal{G}^3$ through adversarial training. The distribution of data p_g in the $\mathcal{G}^3$ follows the distribution of data collected by deployed devices, i.e., p_r. On the one hand, the discriminator D uses a label $y = 1$ to indicate that the input data came from the deployed collection device, and $y = 0$ to indicate that the input data came from the generator. The output of the discriminator D is the probability that the input data come from the measurement data obtained by the deployed collection devices. By optimizing the discriminator parameters θ_D, it ensures that the discriminator D can identify as much as possible the measurement data obtained by the deployed collection devices and the measurement data deduced by the generator. The optimization objective of the discriminator is expressed as

$$\max_{D} \mathop{\mathbb{E}}_{l \sim p_r} [\ln D(l)] - \mathop{\mathbb{E}}_{j \sim p_j} [\ln D(G(j))], \tag{1}$$

where $\mathbb{E}[\cdot]$ represents the expectation, l is the measurement data obtained from the deployed collection devices in $\mathcal{L}^3$, j is the input data of the generator, $G(j)$ is the measurement data deduced by the generator, p_r is the distribution of measurement data obtained from the deployed collection devices, p_j is the distribution of data input to the generator, $D(l)$ is the probability that the discriminator recognizes the measurement data obtained from the deployed collection devices as the measurement data obtained from the deployed collection devices, and $D(G(j))$ is the probability that the discriminator recognizes the measurement data deduced by the generator as the measurement data obtained by the deployed collection devices. On the other hand, the generator D continuously optimizes its parameters θ_G during the training process, hoping that the deduced data can confuse the discriminator D as much as possible, i.e. minimizing the discrimination probability of the deduced data $G(j)$. The optimization objective of the generator is represented as

$$\min_{G} - \mathop{\mathbb{E}}_{j \sim p_j} [\ln D(G(j))]. \tag{2}$$

The training process of D and G is essentially a zero-sum game, with the goal of obtaining the discriminator parameter θ_D that maximizes the classification accuracy of the discriminator and the generator parameter θ_G that deceives discriminator to the greatest extent possible. Define $V(G, D)$as a value function, which contains generator and discriminator parameters. The game objective of this process can be expressed as

$$\min_G \max_D V(G,D) = \mathop{\mathbb{E}}_{l\sim p_r}[\ln D(l)] - \mathop{\mathbb{E}}_{j\sim p_j}[\ln D(G(j))]. \tag{3}$$

3. Improved GAN Based on Tensor Decomposition

The traditional GAN framework is known to exhibit instability when confronted with high-dimensional complex datasets. Furthermore, the presence of multidimensional attributes in pseudo-measurement data collected from important nodes within the low-voltage distribution station area results in diminished network training speed and accuracy for GANs. In response to the above issues, we propose an improved GAN based on tensor decomposition. The improved GAN network model architecture is shown in Fig. 2. Firstly, the expansion matrices of tensor $\mathcal{L}^3$, the expansion matrices of mask tensor $\mathcal{X}^3$, and a set of noise vector $\mathbf{k}$ are used as inputs to the generator to alleviate the instability of GAN when processing high-dimensional complex data and improve GAN's discriminative ability. Secondly, within the framework of the improved GAN model, the goal of the discriminator is to differentiate between data originating from deployed collection devices or generated by the GAN's generator. This adaptation creates a supervised learning environment for the GAN, which ultimately results in the enhancement of model training accuracy. Finally, by optimizing the loss functions for both the generator and discriminator components, the convergence speed of the GAN model is further improved.

3.1 Model comparison

GAN is an unsupervised generative model with "two-person zero-sum game" as the core idea, and the two players in the game are composed of generators and discriminators. The input of traditional GAN model is only a set of random noise, so the samples generated by GAN model are random, and cannot meet the needs of high-dimensional data deduction. In addition, the traditional GAN model is trained without supervision, and the model loss function is relatively simple. Aiming at the requirement of high-dimensional data inference, the improved GAN model preprocesses the data input into the model based on data dimension reduction, so as to avoid the instability of traditional GAN frameworks in the face of high-dimensional complex data sets. In addition, the mask matrix is added to the training of the adversarial network, so that the GAN model can be trained under supervision. At the same time, the improved GAN improves the convergence speed and accuracy of the model by designing a more complex loss function. The detailed comparison between traditional GAN and improved GAN is shown in Table 1.

3.2 Pseudo measurement data processing

In the context of minimal collection, only some collection devices are deployed in the low-voltage distribution station area. In order to conveniently represent the deployment of collection devices in the low-voltage distribution station area, a mask tensor $\mathcal{X}^3$ that is consistent with the collection data dimension is established. When its element value is 0, it indicates that the collection device has not been deployed, and state pseudo measurement data deduction is required. Otherwise, its element value is 1. Simultaneously, a tensor expansion technique is applied to reduce the dimensionality of the input data.During the matrix expansion process of the tensor, it samples all tensor orders in a staggered manner. Instead of simply taking eigenvalues from one order and then another, it conducts modal expansion of eigenvalues from different orders in a staggered sampling fashion. This enables the transfer and fusion of eigenvalues between different orders in the process of sampling. In the process of tensor unfolding, a three-dimensional tensor is unfolded into three two-dimensional arrays, thus realizing the dimensionality reduction of the data.

In order to reduce the impact of anomalous data in the tensor on data deduction, we combine the seasonal component of the time series decomposition and the absolute median deviation for anomalous data detection. Since the tensor contains the pseudo measurement data of each measurement site over a period of time, the time series of the detected site is defined as Y. Firstly, we perform a time series decomposition to obtain the periodic component S_Y of the time series and compute the residual component Y_1, which is givn by

$$Y_1 = Y - S_Y - y', \tag{4}$$

where y' is the median of the time series Y.

The absolute median deviation has a strong robustness to outliers. When there are extreme values or deviated data points, theabsolute median deviation has a lower sensitivity to these anomalies and can more accurately reflect the variability of the time series. Define d_{MAD} as the absolute median deviation, which is given by

$$d_{\text{MAD}} = \text{median}(|Y - \text{y}'|), \tag{5}$$

where median$(\cdot)$ denotes taking the median of the variable.

Then, define R_j as the maximum residual between the residual component and the absolute median deviation, which is given by

$$R_j = \frac{\max|Y_1 - d_{\text{MAD}}|}{d_{\text{MAD}}}, 1 \le j \le Q, \tag{6}$$

where j is sample label, $Q = q_2 \times q_3$ is the total number of samples. After completing the residual computation, anomalous data detection can be performed by comparing the critical value of the t-distribution and the maximum residual. Define λ_j as the critical value of the corresponding t-distribution, which is given by

$$\lambda_j = \frac{(Q-j)t_{p,Q-j-1}}{\sqrt{(Q-j-1+t^2_{p,Q-j-1})(Q-J-1))}}, \tag{7}$$

$$p = 1 - \frac{\sigma}{2(Q-j-1)}, \tag{8}$$

Table 1 Model Comparison

	Traditional GAN	Improved GAN
The imput of model	Random noise	The expansion matrices of tensor, the expansion matrices of mask tensor, and noise vector
The type of model	Unsupervised model	Supervised model
The loss function of the generator	Discrimination probability	Multidimensional data repair loss and multidimensional contextual loss
The loss function of the discriminator	Original loss	Original loss, gradient penalty term, and multidimensional contextual loss

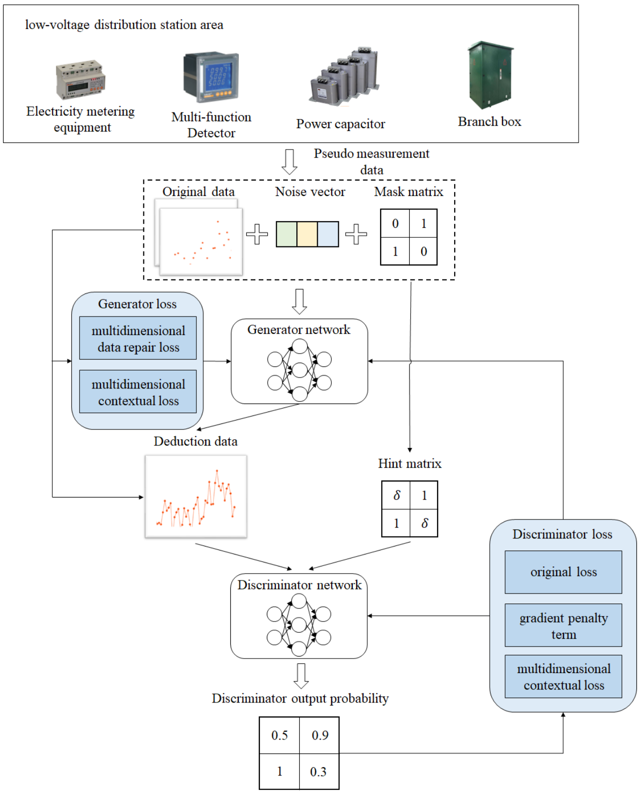


Fig. 2 Improved GAN network model architecture based on tensor decomposition.

where $t_{p,Q-j-1}$ is the critical value of the t-distribution with significance equal to p and degrees of freedom equal to $Q-j-1$. When $R_j \geq \lambda_j$, the data point is defined as an anomalous data point, and it is removed from the tensor, and the removed position is set to zero.

3.3 Improved Generator Network

In the context of minimal collection, only some collection devices are deployed in IoT-empowered distribution station area. In order to conveniently represent the deployment of collection devices in the low-voltage distribution station area, a mask tensor $\mathcal{X}^3$ that is consistent with the collection data dimension is established. When its element value is 0, it indicates that the collection device has not been deployed, and state pseudo measurement data deduction is required. Otherwise, its element value is 1. Simultaneously, a tensor expansion technique is applied to reduce the dimensionality of the input data. Defined $\mathbf{X}^3_{(i)}$ as the i-order expansion matrix of the mask tensor $\mathcal{X}^3$. In order to further utilize the known information in the measurement data obtained by deployed collection devices and reduce the dimensionality of the input data, the expansion matrices of tensor $\mathcal{L}^3$, the expansion matrices of mask tensor $\mathcal{X}^3$, and a set of noise vector $\mathbf{k}$ are used as inputs to the generator. The measurement data in the expansion matrices of tensor $\mathcal{L}^3$ follow the data distribution p_j, and noise vector $\mathbf{k}$ is a random vector with sampling interval $(0, 0.01)$. From this, defined $\bar{\mathcal{L}}^3$ as the output of the generator, which is represented as

$$\bar{\mathcal{L}}^3 = G\left(\mathbf{L}^3_{(i)}, \mathbf{X}^3_{(i)}, \left(1-\mathbf{X}^3_{(i)}\right) \odot \mathbf{k} \mid i = 1,2,3\right), \quad (9)$$

where $\odot$ represents the Hadamard product between two matrices. $\mathbf{L}^3_{(i)}$ is the i-order expansion matrix of the tensor $\mathcal{L}^3$.

Since the generator network generates and replaces the measurement data obtained from the deployed collection devices while deducting the missing data from the regions where no collection devices have been deployed, $\mathbf{L}^3_{(i)}$ is not the final extrapolated data. Define $\hat{\mathcal{L}}^3$ as the deduced data tensor and the three expansion matrices of the deduced data tensor $\hat{\mathcal{L}}^3$ as $\hat{\mathbf{L}}^3_{(1)}$, $\hat{\mathbf{L}}^3_{(2)}$, and $\hat{\mathbf{L}}^3_{(3)}$, which are represented as

$$\begin{aligned}
\hat{\mathbf{L}}^3_{(1)} &= \mathbf{X}_{(1)} \odot \mathbf{L}^3_{(1)} + \left(1-\mathbf{X}_{(1)}\right) \odot \overrightarrow{\mathbf{L}}^3_{(1)}, \\
\hat{\mathbf{L}}^3_{(2)} &= \mathbf{X}_{(2)} \odot \mathbf{L}^3_{(2)} + \left(1-\mathbf{X}_{(2)}\right) \odot \overrightarrow{\mathbf{L}}^3_{(2)}, \\
\hat{\mathbf{L}}^3_{(3)} &= \mathbf{X}_{(3)} \odot \mathbf{L}^3_{(3)} + \left(1-\mathbf{X}_{(3)}\right) \odot \overrightarrow{\mathbf{L}}^3_{(3)},
\end{aligned} \quad (10)$$

where $\overrightarrow{\mathbf{L}}^3_{(1)}$, $\overrightarrow{\mathbf{L}}^3_{(2)}$, and $\overrightarrow{\mathbf{L}}^3_{(3)}$ are the three expansion matrices of $\mathbf{L}^3_{(i)}$. It can be seen that $\hat{\mathbf{L}}^3_{(1)}$, $\hat{\mathbf{L}}^3_{(2)}$, and $\hat{\mathbf{L}}^3_{(3)}$ retain the measurement data obtained by the deployed collection devices in the original data, and the missing data in the undeployed collection device area are filled in with the corresponding values of the deduced data expansion matrices $\overrightarrow{\mathbf{L}}^3_{(1)}$, $\overrightarrow{\mathbf{L}}^3_{(2)}$, and $\overrightarrow{\mathbf{L}}^3_{(3)}$, respectively. $\hat{\mathbf{L}}^3_{(1)}$, $\hat{\mathbf{L}}^3_{(2)}$, and $\hat{\mathbf{L}}^3_{(3)}$ are the final and complete expansion matrix of the deduced data.

3.4 Improved Discriminator Network

In the improved GAN model, the task of the discriminator is no longer to identify the source of input data, but to further distinguish whether the input data of the discriminator are measurement data obtained by deployed collection devices or deduced by the generator. Therefore, the output of the discriminator in the improved GAN is a three-dimensional tensor $\mathcal{H}^3$, where each value h represents the probability that the data in the corresponding input data are the measurement data obtained by the deployed collection devices. When $h = 1$, it means that the discriminator determines that the data is measurement data obtained by the deployed collection devices. When $h = 0$, it means that the discriminator determines that the data are deduced data. At this point, the discriminator is equivalent to predict the mask tensor $\mathcal{X}^3$, because in the mask tensor $\mathcal{X}^3$, the value corresponding to the deployed collection device is 1, and the value corresponding to the undeployed collection device is 0. Since $\mathcal{X}^3$ is a tensor that can be determined in advance, the discriminator is in a supervised state, which can further improve its level. To help the discriminator D better distinguish between the original collection data and the deduced data, a hint mechanism is designed to provide an additional hint tensor $\mathcal{V}^3$ for the discriminator D. Firstly, define a random binary tensor $\mathcal{X}'^3$ with dimensions consistent with the mask tensor $\mathcal{X}^3$ and a sampling space of $\{0, 1\}$. The three expansion matrices of the tensor $\mathcal{V}^3$ can be represented as

$$\begin{aligned}\mathbf{V}^3_{(1)} &= \mathbf{X}'^3_{(1)} \odot \mathbf{X}^3_{(1)} + \delta\left(1 - \mathbf{X}'^3_{(1)}\right),\\ \mathbf{V}^3_{(2)} &= \mathbf{X}'^3_{(2)} \odot \mathbf{X}^3_{(2)} + \delta\left(1 - \mathbf{X}'^3_{(2)}\right),\\ \mathbf{V}^3_{(3)} &= \mathbf{X}'^3_{(3)} \odot \mathbf{X}^3_{(3)} + \delta\left(1 - \mathbf{X}'^3_{(3)}\right),\end{aligned} \tag{11}$$

where $\delta \in (0, 1)$ is a random number and $\mathbf{X}'^3_{(i)}$ is the expansion matrix of the tensor $\mathcal{X}'^3$. The hint tensor retains partial information of the mask tensor. When all the elements of $\mathcal{X}'^3$ are 1, the hint tensor $\mathcal{V}^3$ is consistent with the mask tensor $\mathcal{X}^3$. When all the elements of $\mathcal{X}'^3$ are 0, all the elements in the hint tensor $\mathcal{V}^3$ are δ. Therefore, the sampling space for the elements in the hint tensor $\mathcal{V}^3$ is $\{0, \delta, 1\}$. When the element v in the hint tensor $\mathcal{V}^3$ is set to 0 or 1, the hint tensor $\mathcal{V}^3$ prompts the discriminator that the corresponding value in the input data is from the measurement data obtained by the deployed collection devices or the deduced data from the undeployed collection device area. When $v = \delta$, the hint tensor does not provide additional information for the discriminator, and the discriminator needs to make its own judgment.

3.5 Data Deduction Model for Area without Collection Devices

The flowchart of the proposed improved GAN based on tensor decomposition is shown in Fig. 3. Firstly, fix the generator parameter θ_G and optimize the discriminator parameter θ_D. Due to the presence of the hint mechanism, the discriminator is in a supervised state at this time, and the data that need to be judged by the discriminator itself are the data domain with the corresponding hint tensor $\mathcal{V}^3$'s information amount of 0, i.e. $v = 1$. This article further improves the GAN model by improving the loss function of the discriminator and generator. The improved discriminator loss consists of original loss, gradient penalty term, and multi-dimensional contextual loss. The original loss refers to the

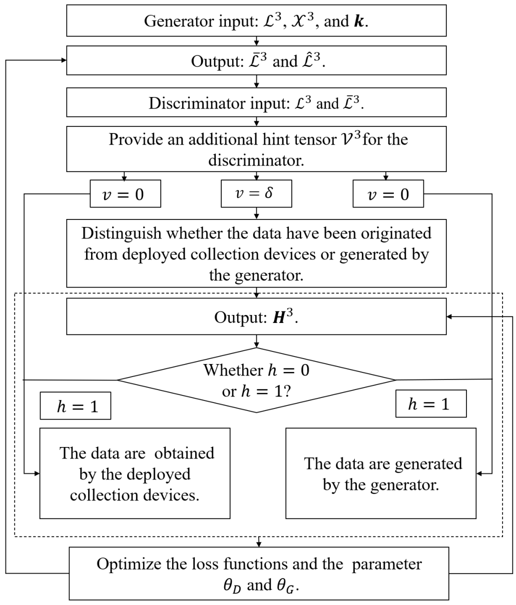


Fig. 3 Flowchart of the proposed improved GAN based on tensor decomposition.

difference between the deduced data and the real data. The gradient penalty term is used to measure the change in the output of the discriminator in the gradient direction, thereby avoiding the disappearance or explosion of the gradient, and the multidimensional contextual loss refers to the difference between the hint tensor and the discriminator output. The loss function of the discriminator can be expressed as

$$\begin{aligned} L_D = & \underset{l\sim p_r}{\mathbb{E}}[\ln D(l)] - \underset{l\sim p_j}{\mathbb{E}}[\ln D(l)] \\ & + \underset{l\sim p_{pen}}{\mathbb{E}}\left[\left(\left\|\nabla D\left(\hat{l}\right)\right\|_2 - 1\right)^2\right] \\ & + \alpha_i \sum_{i=1,v=1}^{N} \left\|\mathbf{V}_{(i)}^3 - \mathbf{H}_{(i)}^3\right\|, \end{aligned} \tag{12}$$

where the distribution p_{pen} is obtained by uniformly sampling the corresponding data points during the training process. The generated data distribution approximates the real data distribution. $\hat{l}$ is the sampling point on this distribution. Then further update the generator parameter θ_G. To ensure that the overall deduced data follow the true data distribution to the maximum extent, the generator network loss function is composed of multidimensional data repair loss and multidimensional contextual loss. The multidimensional data repair loss refers to the difference between the missing data in the areas without collection devices and the deduction data. The multidimensional contextual loss refers to the difference between the measurement data obtained by the deployed collection devices and the deduced data. The loss function of the generator can be expressed as

$$\begin{aligned} L_G = & -D\left(G\left(\mathbf{L}_{(i)}^3 + \left(1 - \mathbf{X}_{(i)}^3\right) \odot \vec{\mathbf{L}}_{(1)}^3 \mid i = 1,2,3\right)\right) \\ & - \alpha_i \sum_{i=1}^{N} \left\|\mathbf{X}_{(i)} \odot \mathbf{L}_{(i)}^3, \mathbf{X}_{(i)} \odot \overline{\mathbf{L}}_{(1)}^3\right\|. \end{aligned} \tag{13}$$

The update of network parameters can be realized by using the improved Adam algorithm based on the above loss function. Improved Adam dynamically adjusts the learning rate of each parameter with the first-order moment estimation and second-order moment estimation of the gradient, and introduces momentum and adaptive learning rate to make the algorithm speed up the model learning, improve the network recognition accuracy, and reduce oscillations at convergence. The network update process as follow. First, first-order moment estimation is performed based on the gradient of the generator loss function as well as the discriminator loss function, which are given by

$$m_o^D = \rho_1 m_{o-1}^D + (1-\rho_1)\nabla_{\vartheta_D} L_D, \tag{14}$$
$$m_o^G = \rho_1 m_{o-1}^G + (1-\rho_1)\nabla_{\vartheta_G} L_G, \tag{15}$$

where ρ_1 is the exponential decay rate of the first-order moment estimation, m_o^G is the first-order moment estimation of generator networks in the o-th iteration, m_o^D is the first-order moment estimation of discriminator networks in the o-th iteration, $\nabla_{\vartheta_D} L_D$ is the gradient of the generator loss function, and $\nabla_{\vartheta_G} L_G$ is the gradient of the discriminator loss function. Second, second-order moment estimation is performed based on the gradient of the generator loss function as well as the discriminator loss function, which are given by

$$v_o^D = \rho_2 v_{o-1}^D + (1-\rho_2)(\nabla_{\vartheta_D} L_D)^2, \tag{16}$$
$$v_o^G = \rho_2 v_{o-1}^G + (1-\rho_2)(\nabla_{\vartheta_G} L_G)^2, \tag{17}$$

where ρ_2 is the exponential decay rate of the second-order moment estimate, v_o^G is the second-order moment estimation of generator networks in the o-th iteration, v_o^D is the second-order moment estimation of discriminator networks in the o-th iteration,

Then, we introduce a learning rate decay strategy based on Adam algorithm, which can speed up the updating of parameters, make Adam algorithm converge faster in the early stage, and can improve the accuracy of the model. The network updates are given by

$$\vartheta_{D,o} = \vartheta_{D,o-1} - \frac{\psi_o \frac{m_o^D}{1-\rho_1}}{\sqrt{\frac{v_0^D}{1-\rho_2}} + \zeta}, \tag{18}$$

$$\vartheta_{G,o} = \vartheta_{G,o-1} - \frac{\psi_o \frac{m_o^G}{1-\rho_1}}{\sqrt{\frac{v_o^G}{1-\rho_2}} + \zeta}, \tag{19}$$

where ζ is a positive number close to 0, which prevents the denominator from being 0 in the formula calculation. ψ_o represents the learning rate, which decreases as the number of iterations increases. The formula for ψ_o is given by

$$\psi_o = \frac{\psi_{o-1}}{1+\xi o}, \tag{20}$$

where ξ is the decay factor.

4. Simulation Result

In practice, the data in the areas without collection devices are not available, which means that data deduction is an unsupervised learning problem. To facilitate validation, we uses a complete dataset as the basis for generating a dataset that contains missing data in the areas without collection devices. The experimental dataset comes from 3 days of current data from 200 users in a distribution station in a city in China, where the data collection frequency is once every 15 minutes. Gaussian perturbations with a standard deviation of 0.02 are added to the data set to further simulate the volatility present in the actual collected data [32]. At the same time, in order to ensure the comprehensiveness of the features of the training data and the strong correlation between the features, the order of the sample data is disrupted. Fig. 4 shows the heat map of the original current data set, with the high current values in red and the low current values in blue. In this figure, the labels denote the user numbers, and we have considered the current data of 200 users during a day. There are differences in the characteristics of electricity consumption of different users under the same station, and their peaks and valleys of electricity consumption are different. In addition, the user data are temporal in nature and are collected, transmitted and stored sequentially at equal time intervals. The data analyzed in this paper were collected at 15-minute sampling intervals, with 96 data points collected in one day. Therefore, the total size of the experimental dataset is 57,600 data points. The dataset exhibits a time series nature, reflecting the electricity usage patterns of users at different time intervals within a day. Preprocessing steps include adding Gaussian noise with a standard deviation of 0.02 to simulate inherent variability in real-world data. To ensure comprehensive features and enhance feature relationships, sample data order was randomized to eliminate temporal effects, enabling the model to better learn inter-feature correlations.

Due to the presence of missing data in the areas without collection devices and the lack of complete data for training network parameters, a multi-layer perceptron is chosen to construct the generator and discriminator network. The input data dimensions of the network are the original data dimensions $(dim) \times 2$. The generator network consists of a three-layer fully connected networks, with the activation functions of the first and second layers selected as $relu$ function, and the activation function of the third layer selected as $tanh$ function. The discriminator network structure is similar to the generator network structure, and is also composed of a three-layer fully connected network[33]. The difference is that the activation function of the third layer is selected as a $sigmoid$ function, and the output is mapped to the interval of $[0, 1]$, representing the discriminator's discrimination probability[34]. The detailed network parameters are shown in Table 2 and Table 3.

The performance of the proposed algorithm is compared with two existing algorithms, which are introduced below.

- ***GAN-based data deduction algorithm[35]:*** Traditional GAN requires a large amount of data and computing resources to support the functions of generators and discriminators. Compared with the algorithm proposed in this paper, traditional GAN is trained in an unsupervised environment with low model convergence speed as well as convergence accuracy. In addition, traditional GAN does not consider the utilization of known information and takes only a set of noise vectors as input to the generator.
- ***WGAN (Wasserstein generative adversarial network)-based data deduction algorithm [36]:*** WGAN introduces the Wasserstein distance to measure the difference between the generated samples and the real samples, so as to improve the stability of training. In addition, in order to ensure that the discriminative network satisfies the Lipschitz continuity condition, WGAN adopts the weight clipping technique to limit the weights of the discriminative network to be fixed within a reasonable range, which ensures the effective calculation of the Wasserstein distance. However, WGAN ignores the utilization of known information and only adopts a set of noise vectors as input, which is more random.
- ***Sequential tensor completion algorithm (STCA) [37]:*** STCA models the data as a three-dimensional tensor, thus formulating the data deduction problem as a low-rank tensor complementation problem. The method can effectively capture the change of data correlation over time, thus improving the efficiency of data deduction.
- ***MissForest algorithm [38]:*** MissForest is a missing data deduction method based on random forest. The method utilizes non-missing data as a training set and missing data as a test set, fits a random forest through the training set and predicts the missing data. In each iteration, MissForest updates the predicted values of the missing values to further improve the repair performance until the convergence condition is satisfied or the maximum number of iterations is reached.

Table 2 Generator Network Structure

Layer	name	parameters	value
0	Input layer	-	$dim \times 2$
1	Fully connected layer	Number of neurons	$dim \times 4$
		Activation function	$relu$
2	Fully connected layer	Number of neurons	$dim \times 3$
		Activation function	$relu$
3	Fully connected layer	Number of neurons	dim
		Activation function	$tanh$

Table 3 Discriminator Network Structure

Layer	name	parameters	value
0	Input layer	-	$dim \times 2$
1	Fully connected layer	Number of neurons	$dim \times 4$
		Activation function	$relu$
2	Fully connected layer	Number of neurons	$dim \times 3$
		Activation function	$relu$
3	Fully connected layer	Number of neurons	dim
		Activation function	$sigmoid$

Fig. 4 and Fig. 5 show the deat map of the original current data set and the current data heat map after data deduction, respectively. The data heat map deduced by the proposed algorithm is almost consistent with the original data heat map, which demonstrates that the proposed algorithm is capable of accurately inferring all the original data from the existing partial data, achieving a relatively good data deduction effect under supervision. The reason for this is that the proposed algorithm utilizes tensor expansion to effectively reduce the dimensionality of the input data, allowing the generator to capture the diversity of the data more effectively. Additionally, the algorithm enhances the generator loss function by integrating multidimensional data repair loss and multidimensional context loss to ensure that the

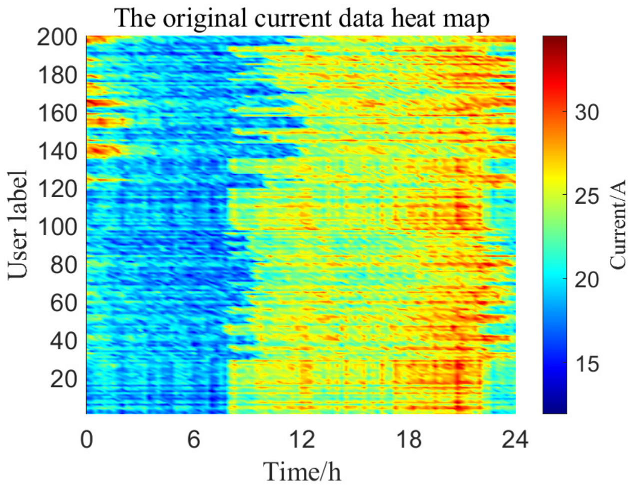


Fig. 4 The heat map of the original current data set.

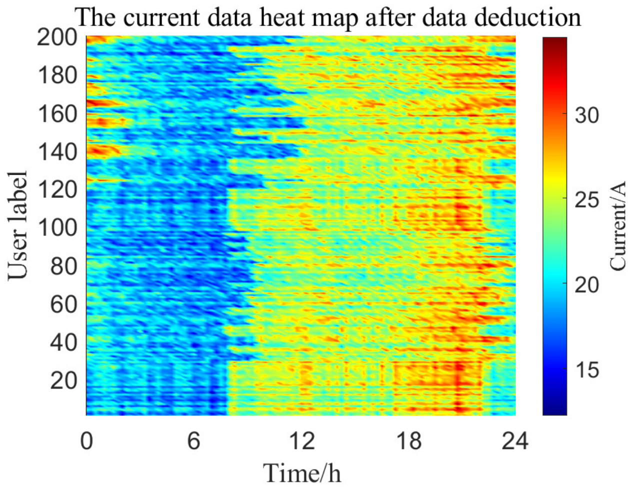


Fig. 5 The current data heat map after data deduction.

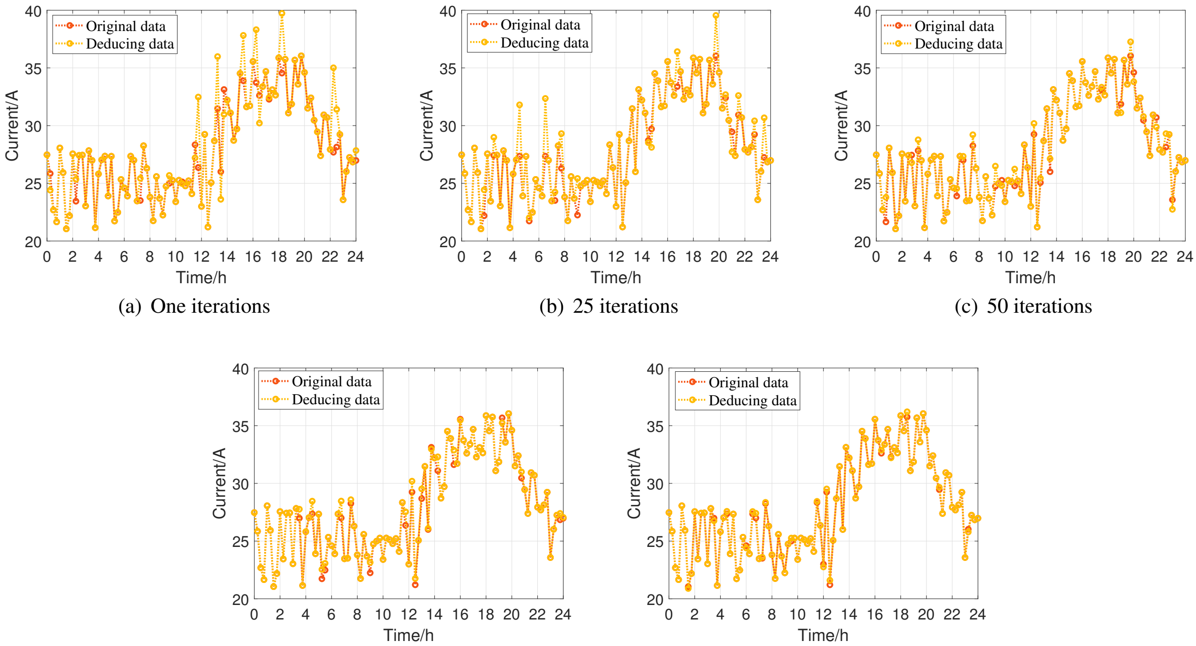
inferred data closely matches the true data distribution.

The measurement data of distribution station in reality has strong randomness and uncertainty, and their missing positions and quantities are uncontrollable. We do not artificially specify fixed missing points. Instead, we randomly generate a mask matrix to ensure a certain range of missing amounts for each data, without limiting the missing position. If a mask is randomly generated with a 20% missing rate as the threshold condition, the missing quantity of 100 sampling points per sample will fluctuate within a small range of 20 sampling points, while the average missing quantity of 2000 data points will remain stable at 20%. It can quantify the deduction effect of the model in areas without deployed collection devices.

Take a load data with a random missing threshold set to 20% in the test data as an example to reflect the performance and process of the model in deducing missing data. Fig. 6 shows data deduction curve with 20% missing rate. Fig. 6(a)-6(e) show the data deduction curve of 1, 25, 50, 100, and 200 iterations respectively. The red curve represents the distribution of the original data, and yellow curve represents the deducing data. The gap between the corresponding points represents the error between the deducing data and the original data. If the two coincide, it means that the error is little, and the data deduced by the model approximate the real data in the area without collection devices. As can be seen from Fig. 6, the gap between the distribution of deduced data and the distribution of real data is gradually narrowing as the number of iterations increases. When the number of iterations is 100, most of the data points of the two curves have overlapped. When the number of iterations is 200, the deduced data have been completely coincident with the original data. The simulation results show that the proposed algorithm performs significantly when the missing rate of data is 20%. It can precisely deduce the missing data in the area without collection devices and performs equally well when endpoint values are missing.

Fig. 7 and Fig. 8 show the variation of root mean square error (RMSE) and mean absolute percentage error with data missing rate. RMSE represents the mean square difference between the deduced value and the original value. It can avoid the problem of errors canceling out each other, thus accurately reflecting the absolute value of the corrected error. Mean absolute percentage error is a comparison between the deduced value and the original value, which better reflects the deduction performance of missing data. The data deduction accuracy of each algorithm is reduced gradually as the data missing rate keeps increasing. When the missing rate is 80%, the RMSE of the proposed algorithm decreases by 25.0%, 47.5%, 65.0%,and 69.6% and the average absolute percentage error of the proposed algorithm is reduced by 9.3%, 12.5%, 24.6%, and 40.2% compared to WGAN, GAN, STCA, and MissForest. The reason for this is that, for time series data or data with temporal relationships, the MissForest algorithm and STCA may not effectively capture data features, leading to inaccurate inference results. Traditional GAN and WGAN operate in an unsupervised environment, resulting in poor discriminator training performance. Additionally, traditional GAN and WGAN do not take into account the utilization of known information, leading to low data inference accuracy of the generator. The discriminator output of the proposed algorithm is a tensor with the same dimension as the input, which puts the discriminator in a supervised environment and effectively improves the training effect of the discriminator. At the same time, the proposed algorithm sets up a hinting mechanism for the discriminator, and it further forces the generator to optimize its own network parameters and improves the generator's data deduction accuracy.

Fig. 9 shows the generator training loss versus iterations. Iteration refers to the process of gradually optimizing the network model by constantly adjusting the model parameters. In each iteration, the model engages in the process of parameter optimization. As the number of iterations increases, the performance of the model will gradually improve. The generator training loss is used to measure the convergence speed and data inference accuracy of GAN and WGAN. The smaller the loss, the higher the inference accuracy. Com-



(a) One iterations (b) 25 iterations (c) 50 iterations

(d) 100 iterations (e) 200 iterations

Fig. 6 Data deduction curve with 20% missing rate.

pared with GAN, the proposed algorithm converges faster and has less loss. When the iteration is 200, the generator training loss of the proposed algorithm is reduced by 61.9% and 75.5% compared to WGAN and GAN. The reason is that the proposed algorithm considers the multidimensional context loss in the design of generator loss functions, and adds a gradient penalty term in the design of generator loss function, which effectively improves the convergence speed of the generator network.

5. Conclusion

In this paper, we proposed a novel multidimensional tensor-aware GAN algorithm for deducing missing data in IoT-empowered distribution station. The proposed algorithm enhances the generator's input, discriminator's structure, and loss function of the GAN, enabling it to handle complex and high-dimensional data. In addition, by employing historical data as a reference and operating in a supervised environment, the proposed algorithm significantly enhances the accuracy of data deduction and convergence performance compared to existing algorithms. Simulation results demonstrate that compared with MissForest and GAN, the data deduction accuracy of the proposed algorithm is improved by 47.5% and 69.6%. The generator training loss of the proposed algorithm is reduced by 33.5% compared with GAN. In the future, we will explore the integration of unsupervised or semi-supervised learning methods with existing models to address the challenge of data inference in situations with insufficient historical data.

Acknowledgments

This work was supported by the Science and Technology Project of State Grid Corporation of China under Grant Number 5400-202316204A-1-1-ZN.

References

[1] H. Liao, Z. Zhou, Z. Jia, Y. Shu, M. Tariq, J. Rodriguez, and V. Frascolla, "Ultra-low AoI digital twin-assisted resource allocation for multi-mode power IoT in distribution grid energy management," IEEE Journal on Selected Areas in Communications, vol.41, no.10, pp.3122–3132, Oct. 2023.

[2] Y. Tian, L. Shen, L. Shen, G. Su, Z. Li, and W. Liu, "AlphaGAN: Fully differentiable architecture search for generative adversarial networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.44, no.10, pp.6752–6766, Oct. 2022.

[3] X. Yu, C. Wang, L. Xu, C. Wu, Z. Wang, Y. He, and W. Wang, "When connected and automated vehicles meet mobile crowdsensing: A perception and transmission framework in the metaverse," IEEE Vehicular Technology Magazine, vol.18, no.4, pp.22–34, Dec. 2023.

[4] H.T. Nguyen and D.H. Choi, "Decentralized distributionally robust coordination between distribution system and charging station operators in unbalanced distribution systems," IEEE Transactions on Smart Grid, vol.14, no.3, pp.2164–2177, Sept. 2023.

[5] H. Liao, Z. Zhou, N. Liu, Y. Zhang, G. Xu, Z. Wang, and S. Mumtaz, "Cloud-edge-device collaborative reliable and communication-efficient digital twin for low-carbon electrical equipment management," IEEE Transactions on Industrial Informatics, vol.19, no.2, pp.1715–1724, Feb. 2023.

[6] Z. Zhang, S. Zhang, L. Mo, M. Guo, F. Liu, and X. Qi, "Traffic volume estimate based on low penetration connected vehicle data at signalized intersections: A bayesian deduction approach," IEEE Transactions on Intelligent Transportation Systems, vol.23,

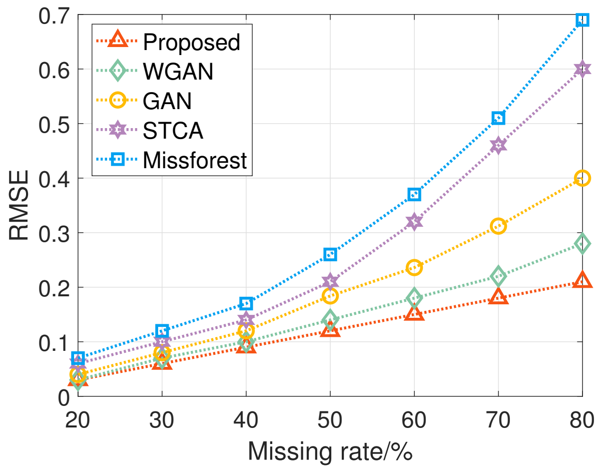


Fig. 7 The RMSE versus the missing rate.

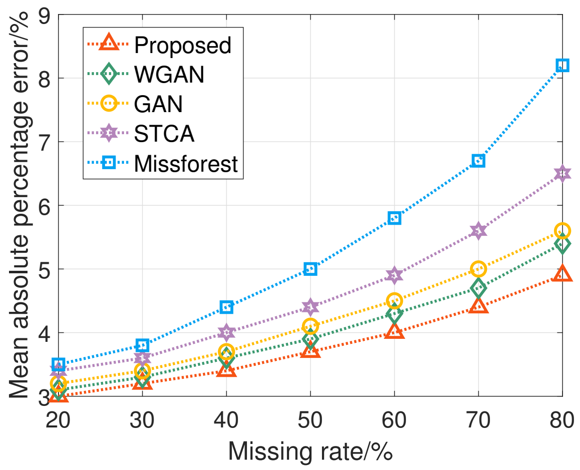


Fig. 8 The mean absolute percentage error versus the missing rate.

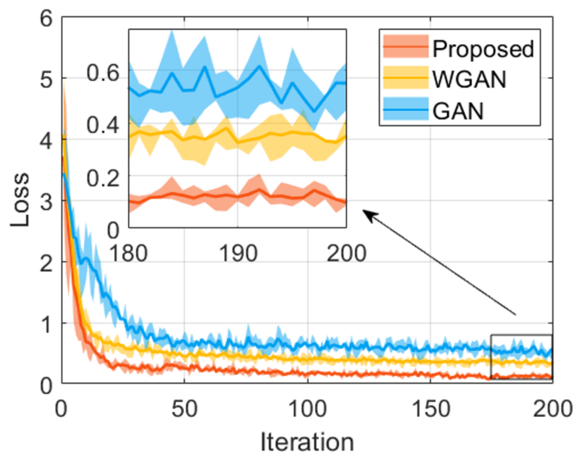


Fig. 9 The generator training loss versus iterations.

no.8, pp.10596–10609, Aug. 2022.

[7] S.J. Siddiqi, B. Tahir, M.A. Jan, and M. Tariq, "Multichain-assisted lightweight security for code mutated false data injection attacks in connected autonomous vehicles," IEEE Transactions on Intelligent Transportation Systems, pp.1–11, 2024.

[8] X. LI, Q. Zhu, W. Xia, and Y. Chen, "A resource-efficient green paradigm for crowdsensing based spectrum detection in internet of things networks," IEICE Transactions on Communications, vol.E106B, no.3, pp.275–286, Mar. 2023.

[9] X. Miao, Y. Wu, L. Chen, Y. Gao, and J. Yin, "An experimental survey of missing data imputation algorithms," IEEE Transactions on Knowledge and Data Engineering, vol.35, no.7, pp.6630–6650, Jun. 2023.

[10] Z. Guo, K. Yu, N. Kumar, W. Wei, S. Mumtaz, and M. Guizani, "Deep-distributed-learning-based POI recommendation under mobile-edge networks," IEEE Internet of Things Journal, vol.10, no.1, pp.303–317, Jan. 2023.

[11] Z. Zhou, Z. Jia, H. Liao, W. Lu, S. Mumtaz, M. Guizani, and M. Tariq, "Secure and latency-aware digital twin assisted resource scheduling for 5G edge computing-empowered distribution grids," IEEE Transactions on Industrial Informatics, vol.18, no.7, pp.4933–4943, Jul. 2022.

[12] S. Khan, K. Muhammad, S. Mumtaz, S.W. Baik, and V.H.C. de Albuquerque, "Energy-efficient deep CNN for smoke detection in foggy IoT environment," IEEE Internet of Things Journal, vol.6, no.6, pp.9237–9245, Dec. 2019.

[13] I. Goodfellow, J. Pouget-Abadie, M. Mirza, and et al., "Generative adversarial nets," Proceedings of Advances in Neural Information Processing Systems, Montréal, Canada, MIT Press, Dec. 2014.

[14] S.J. Siddiqi, M.A. Jan, A.M. Basalamah, and M. Tariq, "Secure teleoperated vehicles in augmented reality of things: A multichain and digital twin approach," IEEE Transactions on Consumer Electronics, vol.70, no.1, pp.956–965, Feb. 2024.

[15] F. Liu, L. Jiao, and X. Tang, "Task-oriented gan for polsar image classification and clustering," IEEE Transactions on Neural Networks and Learning Systems, vol.30, no.9, pp.2707–2719, 2019.

[16] W. Shafqat and Y.C. Byun, "A hybrid gan-based approach to solve imbalanced data problem in recommendation systems," IEEE Access, vol.10, pp.11036–11047, 2022.

[17] A.S. Saqlain, F. Fang, T. Ahmad, L. Wang, and Z.u. Abidin, "Evolution and effectiveness of loss functions in generative adversarial networks," China Communications, vol.18, no.10, pp.45–76, 2021.

[18] K. Zhang, X. Dou, and X. Xiao, "Grid defect data completion based on generative adversarial imputation nets," 2021 IEEE Sustainable Power and Energy Conference (iSPEC), pp.952–957, 2021.

[19] R. Yan, Y. Yuan, Z. Wang, G. Geng, and Q. Jiang, "Active distribution system synthesis via unbalanced graph generative adversarial network," IEEE Transactions on Power Systems, vol.38, no.5, pp.4293–4307, Sept. 2023.

[20] X.Y. Liu and X. Wang, "Real-time indoor localization for smartphones using tensor-generative adversarial nets," IEEE Transactions on Neural Networks and Learning Systems, vol.32, no.8, pp.3433–3443, Aug. 2021.

[21] H. Zhang, X. Hu, D. Ma, R. Wang, and X. Xie, "Insufficient data generative model for pipeline network leak detection using generative adversarial networks," IEEE Transactions on Cybernetics, vol.52, no.7, pp.7107–7120, Jul. 2022.

[22] R. Zhang, W. Gao, G. Li, and T.H. Li, "QINet: Decision surface learning and adversarial enhancement for quasi-immune completion of diverse corrupted point clouds," IEEE Transactions on Geoscience and Remote Sensing, vol.60, no.99, pp.1–14, Nov. 2022.

[23] M. Kang, R. Zhu, D. Chen, X. Liu, and W. Yu, "CM-GAN: A cross-modal generative adversarial network for imputing completely missing data in digital industry," IEEE Transactions on Neural Networks and Learning Systems, vol.35, no.3, Mar. 2024.

[24] P.C. Huang, S.J. Huang, and P.H. Lan, "Cooperation between distributed power modules for SoC power management," IEICE Transactions on Electronics, vol.E99C, no.6, pp.606–613, Jun. 2016.

[25] H. Liao, Z. Zhou, X. Zhao, L. Zhang, S. Mumtaz, A. Jolfaei, S.H. Ahmed, and A.K. Bashir, "Learning-based context-aware resource

allocation for edge-computing-empowered industrial IoT," IEEE Internet of Things Journal, vol.7, no.5, pp.4260–4277, May. 2020.
[26] F. Wu, Y. Li, C. Li, and Y. Wu, "A fast tensor completion method based on tensor QR decomposition and tensor nuclear norm minimization," IEEE Transactions on Computational Imaging, vol.7, pp.1267–1277, Nov. 2021.
[27] G. Du, Z. Wang, B. Gao, S. Mumtaz, K.M. Abualnaja, and C. Du, "A convolution bidirectional long short-term memory neural network for driver emotion recognition," IEEE Transactions on Intelligent Transportation Systems, vol.22, no.7, pp.4570–4578, Jul. 2021.
[28] Y. Liu, Y. Wang, and Q. Yang, "Spatio-temporal generative adversarial network based power distribution network state estimation with multiple time-scale measurements," IEEE Transactions on Industrial Informatics, vol.19, no.9, pp.9790–9797, Spet. 2023.
[29] Y. Wang, L. Deng, Y. Wan, Z. Yang, W. Yang, C. Chen, D. Zhao, F. Wang, and Y. Guo, "An intelligent method for predicting the pressure coefficient curve of airfoil-based conditional generative adversarial networks," IEEE Transactions on Neural Networks and Learning Systems, vol.34, no.7, pp.3538–3552, Jul. 2023.
[30] S.J. Siddiqi, F. Naeem, S. Khan, K.S. Khan, and M. Tariq, "Towards ai-enabled traffic management in multipath tcp: A survey," Computer Communications, vol.181, pp.412–427, 2022.
[31] F. Naeem, S. Seifollahi, Z. Zhou, and M. Tariq, "A generative adversarial network enabled deep distributional reinforcement learning for transmission scheduling in internet of vehicles," IEEE Transactions on Intelligent Transportation Systems, vol.22, no.7, pp.4550–4559, Jul. 2021.
[32] A. Goel and P. Moulin, "Fast locally optimal detection of targeted universal adversarial perturbations," IEEE Transactions on Information Forensics and Security, vol.17, pp.1757–1770, Apr. 2022.
[33] R.H. Randhawa, N. Aslam, M. Alauthman, and H. Rafiq, "Evasion generative adversarial network for low data regimes," IEEE Transactions on Artificial Intelligence, vol.4, no.5, pp.1076–1088, Oct. 2023.
[34] S. Guan, "Performance analysis of convolutional neural networks and multilayer perceptron in generative adversarial networks," 2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA), pp.817–821, 2023.
[35] S. DJ and B. P., "MissForest–non-parametric missing value imputation for mixed-type data," Bioinformatics, vol.28, no.1, pp.112–8, Jan. 2012.
[36] Z. Guo and M. Cui, "One software defective data augmentation method based on VAE and WGAN," 2023 International Conference on Frontiers of Robotics and Software Engineering (FRSE), pp.44–49, 2023.
[37] K. Xie, L. Wang, X. Wang, G. Xie, J. Wen, G. Zhang, J. Cao, and D. Zhang, "Accurate recovery of internet traffic data: A sequential tensor completion approach," IEEE/ACM Transactions on Networking, vol.26, no.2, pp.793–806, 2018.
[38] J. Jia, T. Xu, M. Li, X. Li, and W. Pan, "Data recovery of power plant platform based on generative adversarial network," 2023 IEEE 2nd International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), pp.59–61, 2023.

Jie Ren is now working as a senior engineer in State Grid Shandong Electric Power Company. His research direction includes IoT in distribution grid.

Minglin Liu received his B.S. degree from Shandong University of Technology, China in 2003, majoring in power systems and automation. Since July 2003, he has been working in State Grid Shandong Electric Power Company. He is mainly engaged in the research of power distribution digitalization, power distribution IoT, and power distribution terminal detection technology.

Lisheng Li received his B.S. degree from North China Electric Power University, China in 2009, and his M.S. degree from Xi'an Jiaotong University, China in 2013, majoring in power systems and automation. Since July 2018, he has been working in Shandong Electric Power Research Institute of State Grid as director of Distribution Technology Center. His research interests include smart distribution network, distribution automation, distribution network operation and maintenance, and distributed energy grid connection technology.

Shuai Li received his M.S. degree from Shandong University in 2018. He is now working as a senior engineer in Distribution Technology Center of Shandong Electric Power Research Institute of State Grid. His research direction includes cloud-edge-device collaboration and resource management in distribution grid.

Mu Fang received his M.S. degree from Shandong University in 2018. He is now working as a senior engineer in State Grid Shandong Electric Power Company. His research direction includes cloud-edge-device collaboration and resource management in distribution grid.

Wenbin Liu received his B.S. degree from Zhengzhou University, China in 2018 and M.S. degree from North China Electric Power University, China in 2021, majoring in power systems and automation. Since July 2021, he has been working in Distribution Technology Center of Shandong Electric Power Research Institute of State Grid. He is mainly engaged in the research of distributed new energy control and power distribution digitalization.

Yang Liu received his Ph.D. degree from Huazhong University of Science and Technology in 2016. He is now working as a senior engineer in Distribution Technology Center of Shandong Electric Power Research Institute of State Grid. His research direction includes IoT in distribution grid.

Haidong Yu received his B.S. degree from Shanghai Jiaotong University, China in 2017 and M.S. degree from Shanghai Jiaotong University, China in 2020, majoring in power systems and automation. Since September 2020, he has been working in Distribution Technology Center of Shandong Electric Power Research Institute of State Grid. His research interests include grid connection of photovoltaic systems and intelligent distribution network.

Shidong Zhang received his Ph.D. degree from Beijing University in 2012. He is now working as a senior engineer in Distribution Technology Center of Shandong Electric Power Research Institute of State Grid. His research direction includes cloud-edge-device collaboration and resource management in distribution grid.