IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 28, 2020, accepted January 19, 2021, date of publication January 25, 2021, date of current version May 18, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3054034

Temporal Graph Super Resolution on Power Distribution Network Measurements

ZHISHENG WANG[1], YING CHEN[1], (Member, IEEE), SHAOWEI HUANG[1], (Member, IEEE), XUEMIN ZHANG[1], (Member, IEEE), AND XIAOPENG LIU[2,3]

[1]Department of Electrical Engineering, Tsinghua University, Beijing 100084, China
[2]National Active Distribution Network Technology Research Center (NANTEC), Beijing Jiaotong University, Beijing 100044, China
[3]Collaborative Innovation Center of Electric Vehicles in Beijing, Beijing Jiaotong University, Beijing 100044, China

Corresponding author: Ying Chen (chen_ying@tsinghua.edu.cn)

This work was supported by the National Key Research and Development Program of China (Basic Research Class) under Grant 2017YFB0903000.

ABSTRACT The applications of super-resolution (SR) technology in the field of image completion are successful. Nevertheless, industry applications demand not only image completion but also the topology and time-series completion. In this article, the SR technology on a topology graph is studied in the scenario of recovering measurements in power distribution systems for cost saving and security & stability improvement. The power flow and voltage magnitude measurements on feeders are reported at different frequencies. In this article, a new data completion method considering distribution system topology is proposed. Firstly, the graph convolutional neural network (GCN) is used for spatial-temporal convolution on a graph, and then the power system state estimation (SE) is used introducing the physical constraints. This method realizes the super-resolution of distribution system measurements, improves the state awareness of distribution systems. Hence, it helps to improve the efficiency of distribution network operation and to reduce equipment failures.

INDEX TERMS Super resolution, graph convolution, power distribution network, state estimation.

I. INTRODUCTION

The super-resolution (SR) has satisfying applications in the field of image completion. Moreover, the emerging deep learning methods are capable of recovering the high resolution (HR) images from the low resolution (LR) ones to an impressive extent [1], [2]. Medical imaging super-resolution also draws a lot of attention [3]. However, the researchers focus more on images. This article tries to perform temporal SR on power distribution system measurements. The power distribution system [4] is an important part of the power system. It takes electricity from power transmission systems and distributes it to end-users (factories, dwellings). Because of historical reasons, its measurement quality is significantly lower than that of transmission systems [5]. However, due to the demand of power supply quality, development of electrical vehicles, and penetration of renewable energy, the state awareness of the power distribution systems becomes increasingly important [6]. Hence, this article tries to perform graph temporal SR, in order to recover the high temporal resolution (HTR) measurements from its current low temporal resolution (LTR) measurements. It will help to improve the efficiency of distribution network operation and to reduce equipment failures.

The super-resolution, that is, obtaining HR data from LR observations, is an attractive field of research. The field of image SR is especially attractive. The image super-resolution has many applications, such as medical image processing [7], [8], facial image improvement [9], [10], enhancement of compressed images/videos [11], [12], and thermal image enhancement [13]. Nevertheless, the SR is a broader concept. To some extent, the power system state estimation (SE) [14]–[16] is also a kind of SR. The SE takes the redundant amount of low accuracy measurements and generates more precise system states. As for methods of the SR, they range from Fourier frequency domain methods [17], [18], wavelet methods [19], [20], to deep learning methods [1], [21]–[25]. The first attempt of applying deep learning on SR is SRCNN [24]. A very simple three-layer convolutional neural network (CNN) reports impressive performance. SRCNN [24] divides SR into three steps: Patch extraction and representation, non-linear mapping,

The associate editor coordinating the review of this manuscript and approving it for publication was Zhanpeng Jin.

and reconstruction. Then, the mean square error (MSE) is selected as the loss function. This article still follows some basic ideas of SRCNN [24]. Later, more deep learning structures are applied on SR, e.g., SR with very deep residual channel attention networks [26] and SRGAN [1]. SRGAN stands out for its novel structure and novel assessment of loss function. Generative adversarial network (GAN) is used for generating HR images while the loss function is not MSE or other analytical functions, but another CNN. The works mentioned above in this paragraph all focus on image SR. However, industrial applications demand more than this. SR on industrial sensors are also valuable and in need. There are an enormous amount of sensors installed, including thermal sensors, voltage sensors, current sensors, power sensors, vibration sensors, and so on. These sensors create time series data at different frequencies, from different installed locations. Chen *et al.* [27] performs temporal super-resolution on smart meters. It uses GAN with 1-D convolution layers and reports impressive results. However, it deals with only one smart meter, neglecting the correlations. As far as we know, there are no published works on temporal SR on multi-sensors on a graph.

On the power system measurement completion problem, researchers also proposed methods based on the matrix completion [28] and tensor completion [29], [30]. These methods generally perform the tensor/matrix decomposition followed by tensor/matrix reconstruction, exploiting the low-rank property of measurements. They perform well on PMU missing data completion and power distribution network measurement recovering and enhancing. These tensor/matrix completion methods generally do not require topology parameters.

This work tries to implement temporal SR on a graph with graph neural networks and SE. Deep learning on a graph is also an attractive research topic. There are applications ranging from rational reasoning [31], modeling molecular fingerprints [32], to image recognition [33]. The initial of learning on a graph can be dated back before the emergence of deep learning [34], [35]. Frequency domain methods are used in these researches. The graph convolutional network (GCN) follows the basic idea of CNN: sharing weights and extracting structural features. The graph convolution layer evolves from the intuitive spectral way [36], to Chebyshev polynomials [37] then to first-order approximation [38]. There are node level learning and graph-level learning in works about GCN. In the node level learning, the GCN treats one node as a sample, e.g., classifying whether a person is emotional in a social network, and classifying whether an article is valuable in a citation network. The graph level GCN, i.e., treating one graph as a sample, is less common. For example, classifying whether a social network is healthy, and generating a fake citation network. The graph level GCN is helpful for SR in power distribution systems.

Contributions: This article presents a temporal graph super-resolution method for measurement completion in power distribution systems. As far as we know, there are no existing works about graph level temporal SR on industrial sensors. It will help to improve power distribution network security and stability, and to reduce monitoring costs.

1) This article presents a framework about applying SR on power distribution systems, including feature transformation, graph level GCN, and state estimation.
2) This article solves the problem that mixed node features and edge features are hard to train. This is accomplished by transforming edge features to node features using properties of the power distribution system.
3) This article modifies the first-order Chebyshev polynomials GCN for graph level regression.
4) This article considers physical constraints (results of GCN is tuned by an SE)

Tests show the superiority of the proposed SR method over different interpolation methods for more than 20% (overall MSE). Moreover, it can also perform spatial-temporal SR when data of some nodes are completely missing.

II. PROBLEM DESCRIPTION

A. INTRODUCTION OF POWER DISTRIBUTION SYSTEMS

Power distribution systems along with power transmission systems are significant parts of the power system. The power transmission system transmits electricity for long distances and acts as the bone of power systems, while the power distribution system takes power from the transmission system and distributes the electricity to end-users. For brevity, the power distribution system and power transmission system are called distribution system and transmission system for short.

Usually, in a distribution system, there is only one node connecting to the transmission system. That node is called the parent node. The parent node is equipped with supervisory control and data acquisition (SCADA) devices or even phasor measurement units (PMU), it has reliable and HTR measurements. Several feeders connect loads or distributed energy resources to the parent node. To measure the states of the distributed system, measuring devices are installed along the feeders. However, the measurement quality is significantly lower than that of transmission systems, this is because of considerations of costs. Although the scale (number of nodes) of a distribution system is small, there are too many distribution systems in one power system. The low measurement quality reflects in two folds:

1) Temporal resolution of measurements in distribution systems are low and uneven. The parent node is equipped with SCADA devices and therefore is HTR and reliable. The measurements of other nodes or lines are LTR. The temporal resolution of measurements along different feeders are also different.
2) There are also fewer types of measurements in distribution systems. Typical available types of measurements are P_{lf} (active power flow of lines) and V_{m} (voltage magnitude of nodes). Usually, there are no reactive power measurements.

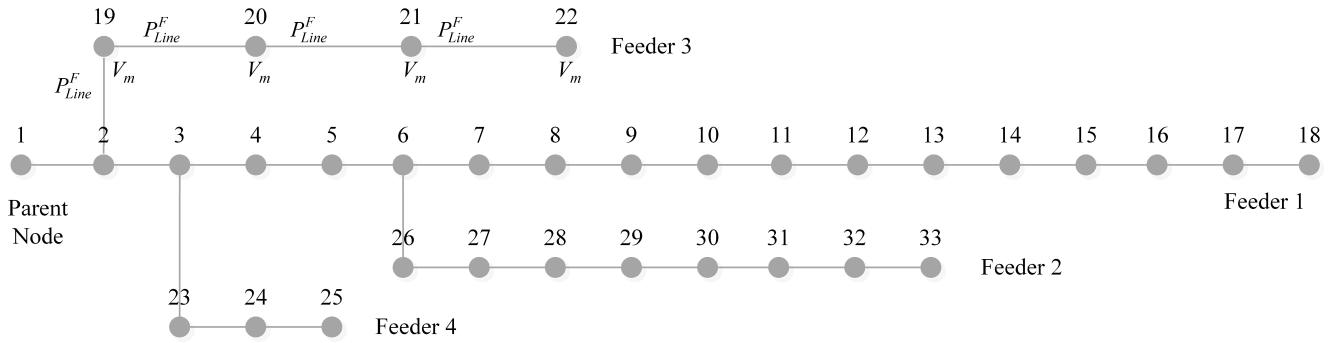


FIGURE 1. IEEE 33-node distribution system. Line measurements P_{lf} (active power flow on the transmission line) and node measurements V_m (voltage magnitude of the node) of the feeder 3 are labeled, measurements of other feeders are omitted in the figure.

FIGURE 2. Temporal resolution of measurements associated with different nodes.

An illustration of the IEEE 33-node distribution system is shown in Fig. 1. There are 4 feeders in this system. The number #1-#33 represent for the nodes. These nodes have loads, distributed energy resources, or both. Node 1 is the parent node that connects with the transmission system. Each node is equipped with devices measuring voltage magnitude V_{m} as shown in Fig. 1. The power flow on the feeder is measured at different positions, denoted as P_{lf}. In normal operation, the topology of distribution systems are radial, that is, it is an acycline graph.

B. SUPER RESOLUTION OF TEMPORAL MEASUREMENTS

Due to considerations of costs, most measurements of the distribution system are LTR, resulting in low state awareness. However, with more penetration of distributed (renewable) energy resources and rapid development of the electricity market, state awareness becomes important. Limited by the cost of communication bandwidth, the capacity of the data center and measuring devices, measurement quality is less likely to upgrade in a short time. Therefore, a new demand rises in the power industry: Using LTR equipment to accomplish HTR measuring and enhance state awareness.

The temporal resolution, that is, the data reporting frequency, follows Fig. 2. The parent node (node 1 in Fig. 2) has the highest temporal resolution, while the temporal resolution of other nodes varies. Usually, the temporal resolutions along a feeder are the same. Denote the frequency of parent node reporting data as f_p. As shown in Fig. 2, node 2 (on feeder 1) reports data at frequency $f_p/2$, while node 23 (on feeder 4) reports data at frequency $f_p/4$.

For illustration, suppose the HTR of measurement of node 1, 2, 23 are as follows:

$$\begin{aligned}
\text{Node 1}: y_1 &= y_2 + y_{23} = \sin(\frac{2\pi}{16}t) + \sin(\frac{2\pi}{8}t) \\
\text{Node 2}: y_2 &= \sin(\frac{2\pi}{8}t) \\
\text{Node 23}: y_{23} &= \sin(\frac{2\pi}{16}t)
\end{aligned} \tag{1}$$

The original HTR measurements are the right three subfigures (d)-(f) in Fig. 3, while the left three subfigures (a)-(c) report the LTR measurements. It can be observed that LTR causes distortion. This article tries to restore the HTR measurements from the LTR ones. By doing so, the state awareness of distribution systems will be enhanced.

There are several obstacles for this task:

1) The measurements are on a graph, not the 2-D image. Topological connection of the distribution system holds important information, and should not be abandoned.
2) We have to perform learning on graph level, yet researchers focus more on node level graph learning.

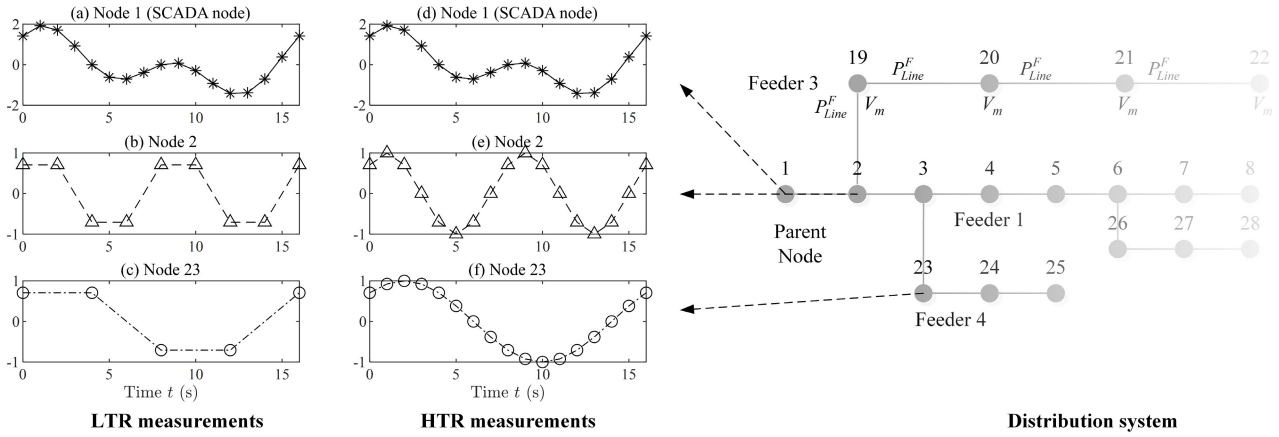


FIGURE 3. Illustration of LTR and HTR measurements of node 1,2, and 23 in a distribution system. The left three subfigures (a)-(c) give the LTR measurements, while the right three subfigures (d)-(f) report the HTR measurements.

Resulted by graph level learning and considering time-series data, the scale of the input data is relatively high.

3) Both node features, e.g., voltage magnitude V_{m} and edge features, e.g., power flow on the feeder P_{lf} are involved.
4) How to take advantage of the physical constraints.

If we can overcome these obstacles, the benefits will be promising: 1) **Cost saving**: The power system consists of transmission networks and distribution networks. The state monitoring techniques and devices, e.g., SCADA, SE, and PMU, are well developed and installed in transmission networks. However, state monitoring in the distribution system is rather primitive. The most important reason is the quantity and therefore the price: Usually, every node in the transmission network branches out a distribution network. The PJM transmission network in the U.S. has thousands of nodes. That is, a single transmission network branches out thousands of distribution networks, and therefore make it almost impossible to install the same monitoring devices like those in transmission networks. Nonetheless, power grids in both China and the U.S. are trying to develop smart distribution networks, which requires sophisticated monitoring. This article tries to solve this problem. The low frequency and low precision monitoring devices are much cheaper than high frequency and high precision ones, not to mention the cost of network bandwidth and storage. The method proposed in this article restores the high temporal resolution (HTR) data from the low temporal resolution (LTR) ones. And uses the distribution network state estimation to increase precision. The proposed method managed to acquire acceptable monitoring results with much cheaper devices. 2) **Power system security and stability**: More importantly than all of that, power system failures are vital for modern societies. Apart from reducing costs, this method can obtain HTR data from LTR observations, which will contribute to making power systems more secure, stable, and economical.

III. TEMPORAL SUPER RESOLUTION ON A GRAPH

This article solves the distribution system super-resolution problem with methods involving GCN and SE. Firstly, we perform the data preprocessing. The edge features are converted to node features according to properties of the radial graph. Later, the features and labels are normalized to the interval [0, 1] to meet the sigmoid output functions. Secondly, the features and labels are sent to train a graph-level GCN. Finally, the trained data are recovered from normalization and tuned by the SE.

A. NORMALIZATION AND TRANSFORMATION

The raw data are preprocessed to meet the demand for training. Firstly, we define the data format. The raw data are voltage magnitude V_{m} for all nodes and power flow on the feeder P_{lf} at all edges. At the time t, the measurement vectors are denoted as:

$$\boldsymbol{V}_{\mathrm{m}}^{(t)} = \begin{bmatrix} V_{\mathrm{m},1}^{(t)} \\ V_{\mathrm{m},2}^{(t)} \\ \vdots \\ V_{\mathrm{m},n_N}^{(t)} \end{bmatrix} \in \mathbb{R}^{n_N \times 1}, \quad \boldsymbol{P}_{\mathrm{lf}}^{(t)} = \begin{bmatrix} P_{\mathrm{lf},1}^{(t)} \\ P_{\mathrm{lf},2}^{(t)} \\ \vdots \\ P_{\mathrm{lf},n_E}^{(t)} \end{bmatrix} \in \mathbb{R}^{n_E \times 1}, \tag{2}$$

where n_N represents for the number of nodes in the distribution system, while n_E represents for the number of edges.

Unfortunately, $\boldsymbol{V}_{\mathrm{m}}^{(t)}$ is a node feature but $\boldsymbol{P}_{\mathrm{lf}}^{(t)}$ is an edge feature. To make the proposed SR method capable of dealing with both node features and edge features, we firstly perform a transformation. This transformation is available because of properties of the distribution system. The key idea is to transform line power flows to node power injections. The transformation is carried out using incidence matrix (node-edge matrix), which stores the association information between nodes and edges [39]. Associated with a direct graph $\mathcal{G} = \mathcal{G}(V, E)$, the incidence matrix is defined as $\mathbf{A} = \mathbf{A}(\mathcal{G}) \in \mathbb{R}^{n_N \times n_E}$, where $V = \{v_1, v_2, \cdots, v_m\}$ is the node set and

$E = \{e_1, e_2, \cdots, e_n\} \in V \times V$ is the edge set:

$$a_{ij} = \begin{cases} 1, & \text{Edge } e_j \text{ starts from node } v_i \\ 0, & \text{Edge } e_j \text{ is not directly connected to node } v_i \\ -1, & \text{Edge } e_j \text{ ends to node } v_i. \end{cases} \tag{3}$$

The distribution system is a radial graph, i.e., there are no circles in the graph. This property enables an invertible transformation:

$$\boldsymbol{P}_{\mathrm{i}}^{(t)} = \mathbf{A}\boldsymbol{P}_{\mathrm{lf}}^{(t)}, \quad \boldsymbol{P}_{\mathrm{i}}^{(t)} \in \mathbb{R}^{n_N \times 1}, \tag{4}$$

where the $\boldsymbol{P}_{\mathrm{i}}^{(t)}$ is called the pseudo-injection, and it is associated with nodes rather than edges. It should be noted that such transformation is *not* available in transmission systems because they have loops and may have power flow circulations.

After transformation, $\boldsymbol{V}_{\mathrm{m}}^{(t)}$ and $\boldsymbol{P}_{\mathrm{i}}^{(t)}$ are linearly normalized to the interval [0, 1] to meet the demand of GCN training. Then, time series of $\boldsymbol{V}_{\mathrm{m}}^{(t)}$ and $\boldsymbol{P}_{\mathrm{i}}^{(t)}$ are grouped as training samples: Denote the length of the sample as l_{s}. Every continuous l_{s} features are grouped as a sample $\mathbf{S} = (\mathbf{F}, \mathbf{L})$, $\mathbf{F}$ represents for features for the GCN, and $\mathbf{L}$ represents for labels for the GCN.

$$\begin{aligned} \mathbf{L}_v &= \begin{bmatrix} \boldsymbol{V}_{\mathrm{m}}^{(1)} & \boldsymbol{V}_{\mathrm{m}}^{(2)} & \cdots & \boldsymbol{V}_{\mathrm{m}}^{(ls)} \end{bmatrix} \\ &= \begin{bmatrix} V_{\mathrm{m},1}^{(1)} & V_{\mathrm{m},1}^{(2)} & \cdots & V_{\mathrm{m},1}^{(l_{\mathrm{s}})} \\ V_{\mathrm{m},2}^{(1)} & V_{\mathrm{m},2}^{(2)} & \cdots & V_{\mathrm{m},2}^{(l_{\mathrm{s}})} \\ \vdots & \vdots & \vdots & \vdots \\ V_{\mathrm{m},n_N}^{(1)} & V_{\mathrm{m},n_N}^{(2)} & \cdots & V_{\mathrm{m},n_N}^{(l_{\mathrm{s}})} \end{bmatrix} \in \mathbb{R}^{n_N \times l_{\mathrm{s}}}, \end{aligned} \tag{5}$$

$$\begin{aligned} \mathbf{L}_p &= \begin{bmatrix} \boldsymbol{P}_{\mathrm{i}}^{(1)} & \boldsymbol{P}_{\mathrm{i}}^{(2)} & \cdots & \boldsymbol{P}_{\mathrm{i}}^{(ls)} \end{bmatrix} \\ &= \begin{bmatrix} P_{\mathrm{i},1}^{(1)} & P_{\mathrm{i},1}^{(2)} & \cdots & P_{\mathrm{i},1}^{(l_{\mathrm{s}})} \\ P_{\mathrm{i},2}^{(1)} & P_{\mathrm{i},2}^{(2)} & \cdots & P_{\mathrm{i},2}^{(l_{\mathrm{s}})} \\ \vdots & \vdots & \vdots & \vdots \\ P_{\mathrm{i},n_N}^{(1)} & P_{\mathrm{i},n_N}^{(2)} & \cdots & P_{\mathrm{i},n_N}^{(l_{\mathrm{s}})} \end{bmatrix} \in \mathbb{R}^{n_N \times l_{\mathrm{s}}}. \end{aligned} \tag{6}$$

$\mathbf{L}_v$ and $\mathbf{L}_p$ are all HTR. By element-wise multiplying a mask $\mathbf{M}$ with the HTR labels, we have the LTR features.

$$\mathbf{F}_v = \mathbf{L}_v \odot \mathbf{M}, \quad \mathbf{F}_p = \mathbf{L}_p \odot \mathbf{M} \tag{7}$$

where $\odot$ represent for element-wise product, and the mask $\mathbf{M}$ is defined as:

$$m_{i,j} = \begin{cases} 1, & \text{if there are data reporting from node } i \text{ at time} j \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

So far we have transformed edge features to node features, constructed LTR features $\mathbf{F}_v$, $\mathbf{F}_v$ and HTR labels $\mathbf{L}_v$, $\mathbf{L}_p$

B. THE GRAPH CONVOLUTIONAL NETWORK

The graph convolutional network (GCN) acts as the bone of the proposed SR method. The GCN deals with data on a graph. Its typical application includes node-level classification, graph-level classification, node-level regression, and graph-level regression [40]. In this article, we use the GCN for graph level regression. The GCN used in this article is inspired by Kipf. *et al.* [38]. The original model is developed for node-level semi-supervised classification. This article modified the model for graph-level regression.

We take the re-normalized linearized convolution layer. It is defined as below:

$$\mathbf{Z} = \widetilde{\mathbf{D}}^{-\frac{1}{2}}\widetilde{\mathbf{A}}\widetilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{X}\Theta \tag{9}$$

where the symbols are defined as below:

1) $\mathbf{X} \in \mathbb{R}^{n_N \times n_{\mathrm{in}}}$ is the input of this layer.
2) $\mathbf{Z} \in \mathbb{R}^{n_N \times n_{\mathrm{out}}}$ is the output of this layer.
3) $\widetilde{\mathbf{A}} \in \mathbb{R}^{n_N \times n_N}$ is the re-normalized adjacent matrix. $\widetilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$, where $\mathbf{A}$ is the adjacent matrix of the graph, and $\mathbf{I}$ is the unit matrix.
4) $\widetilde{\mathbf{D}} \in \mathbb{R}^{n_N \times n_N}$ is the re-normalized degree matrix. defined as: $\widetilde{\mathbf{D}}_{i,i} = \sum_j \widetilde{\mathbf{A}}_{i,j}$
5) $\Theta \in \mathbb{R}^{n_{\mathrm{in}} \times n_{\mathrm{out}}}$ is the a matrix of trainable filter parameters.

Equation (9) is derived from $g_\theta \star x \doteq \theta(\mathbf{I} + \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2})x$, the spectral graph convolution equation approximated with Chebyshev polynomials [37]. Rewrite the equation above from element-wise presentation to matrix presentation with a re-normalization trick: $\widetilde{\mathbf{D}}^{-1/2}\widetilde{\mathbf{A}}\widetilde{\mathbf{D}}^{-1/2} = \mathbf{I} + \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$, we will get (9).

In this article, we perform the training of $\mathbf{S}_v = (\mathbf{F}_v, \mathbf{L}_v)$ (voltage) and $\mathbf{S}_p = (\mathbf{F}_p, \mathbf{L}_p)$ (active power flow) separately. In power systems, the voltage and active power are weakly correlated. In high-voltage power systems, the voltage magnitude V_{m} and reactive power Q are in a strongly correlated group. The angle of voltage θ and active power P are strongly correlated. However, the correlation between these two groups is weak. This is called the PQ decoupled property. Because of such property, we can separate them to reduce unnecessary trainable parameters. In this article, the V-network and the P-network are of the same structure. As shown in Fig. 4, the GCN has 6 layers. The number of input filters increases from 64 to 512 and then decreases back to 64. Detailed input and output filters are illustrated in Fig. 4. The first 5 layers all use rectified linear unit (ReLU) as the activation function, while the last layer uses the sigmoid activation function. Batch normalization is inserted before layer 4 to accelerate the training. The GCN is trained with minibatch (batch size = 32), by Adam optimizer (with decreasing learning rate) for 200 epochs (approximately 800 batches for an epoch). It takes approximately 40 s for training one epoch on a server with one NVIDIA Tesla P100 GPU.

The number of trainable parameters for one graph convolution layer is approximately $n_{\mathrm{in}}n_{\mathrm{out}} + n_N n_{\mathrm{out}}$. The float point operations (FLOPs) for one graph convolution layer

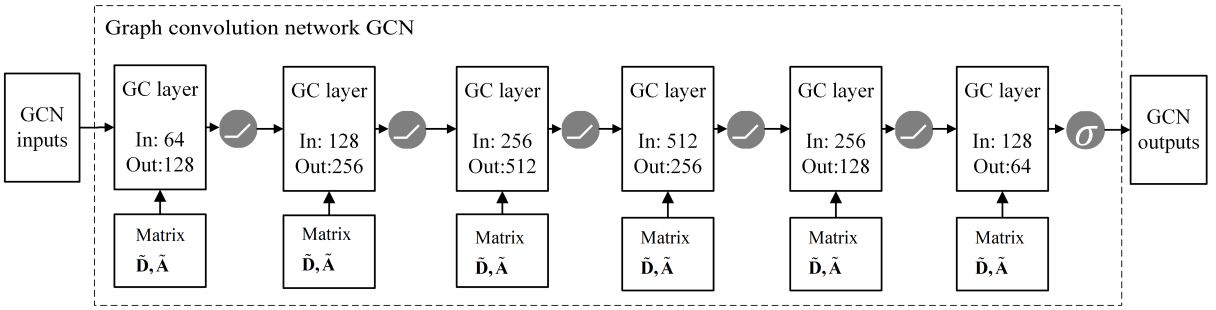


FIGURE 4. Sturcture of the GCN.

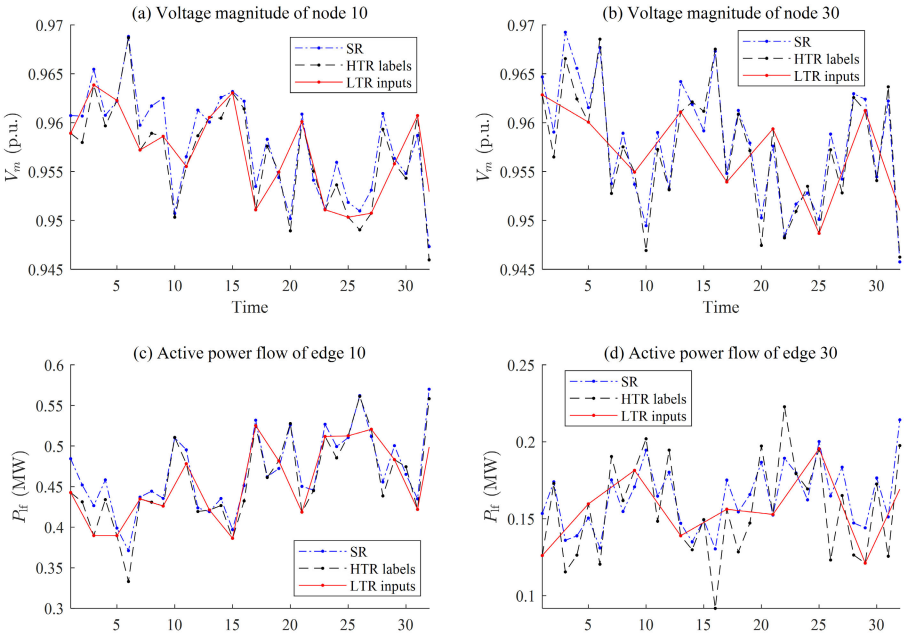


FIGURE 5. Performance of the proposed SR method on different nodes/edges: SR results, HTR labels and LTR inputs.

is approximately $n_N^2\ n_{\text{in}} + n_N n_{\text{in}} n_{\text{out}}$, where n_{in} is the layer input dimension, n_{out} is the layer output dimension, and n_N is the number of nodes in the graph. There are six layers in the GCN, the number of nodes of the graph is $n_N = 33$ and is the same for each layer. The n_{in} and n_{out} can be found in the network structure in Fig. 4. The total number of trainable parameters is:

$$\sum_{i=1}^{6} n_{\text{in},i} n_{\text{out},i} + n_{N,i} n_{\text{out},i} \doteq 3.88 \times 10^5 = 0.388\ \text{M} \tag{10}$$

The FLOPs:

$$\sum_{i=1}^{6} n_{N,i}^2 n_{\text{in},i} + n_{N,i} n_{\text{in},i} n_{\text{out},i} \doteq 1.2 \times 10^8 = 0.12\ \text{G} \tag{11}$$

C. THE STATE ESTIMATION

The power system state estimation (SE) tunes the outputs of the GCN with physical constrains of the distribution system. The output data are firstly restored from normalization and transformation. The SE takes redundant measurements to estimate the truth value of the power system. In this process the constrain of power flow equation is embedded:

$$\begin{cases} \Delta P_i = P_i^{SP} - V_i \sum_{j \in i} V_j \left(G_{ij} \cos\theta_{ij} + B_{ij} \sin\theta_{ij} \right) \\ \qquad\qquad \text{with } i = 1, 2, \cdots, n_N \\ \Delta Q_i = Q_i^{SP} - V_i \sum_{j \in i} V_j \left(G_{ij} \sin\theta_{ij} - B_{ij} \cos\theta_{ij} \right) \\ \qquad\qquad \text{with } i = 1, 2, \cdots, n_N - r \end{cases} \tag{12}$$

where, P_i^{SP} and Q_i^{SP} represents for node active and reactive power injections; V_i represents for the voltage magnitude of node i; $\theta_{ij} = \theta_i - \theta_j$, θ_i represents for the voltage phase of node i; G_{ij} and B_{ij} are the real part and image part of the admittance matrix. For brevity, the detailed information of SE is omitted here, and can be found in [41]–[43].

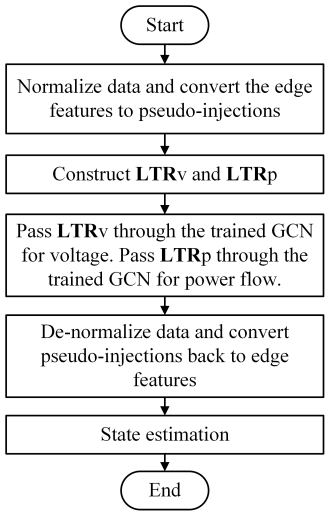


FIGURE 6. Workflow of the proposed SR method.

D. SUPER RESOLUTION ALGORITHM

The workflow of the proposed SR method is reported in Fig. 6, while Algorithm 1 is a more detailed version. The workflow can be divided into four parts: 1) Process data for GCN representation; 2) GCN inference; 3) Restore data from GCN representation; 4) State estimation. The code of the proposed SR algorithm is available at https://github.com/DelbertWang2/TemporalGraphSR.

IV. EXPERIMENTAL EVALUATION

In this part, the proposed SR method is applied to IEEE 33-node distribution system with distributed solar generation considered, whose topology is already shown in Fig. 1. Solar generators are installed on nodes 7, 8, 24, 25, 30, 32. Datasets are generated through simulation. Day-level load curve, hour-level load curve, and solar generation curve are used to adjust the load and generation level. The noises and detail patterns are also added. The real load curve and solar generation curve are not perfectly smooth. Therefore a small amount of noise (0.5 %, normal distribution) is added when generating samples. Then the system is simulated by solving power flow equations. For brevity, the detailed process of dataset generation is omitted here. It can be found in the attached code.[1] This article uses the power flow and SE module in a popular MATLAB toolbox MATPOWER. The GCN is implemented via PyTorch.[2] In all cases below, node 1 is the parent node equipped with a SCADA device. Its data reporting frequency is $f_p = 1/300$ Hz, that is, reporting data every 5 minutes. Node 2-18, which are the node along the feeder 1, report data at frequency $1/2\,f_p$. The other nodes (node 19-33) report data at frequency $1/4\,f_p$. All test samples are not involved in training.

[1]Data: / MatScripts / gen_data /; Generated dataset: / Data / rawdata.mat
[2]ATTACHMENT / PythonScripts /

Algorithm 1 Super Resolution Algorithm

Require: LTR measurements: voltage magnitude $\boldsymbol{V}^{(t)}_{\mathrm{m,\,LTR}}$, power flow $\boldsymbol{P}^{(t)}_{\mathrm{lf,\,LTR}}$, Adjacent matrix of the graph $\mathbf{A}$;

Ensure: HTR measurements: voltage magnitude $\boldsymbol{V}^{(t)}_{\mathrm{m,\,HTR}}$, power flow $\boldsymbol{P}^{(t)}_{\mathrm{lf,\,HTR}}$

Process data for GCN representation

1: Normalize data to [-1, 1].
2: Convert power flow $\boldsymbol{P}^{(t)}_{\mathrm{lf,\,LTR}}$ (edge feature) to pseudo-injection $\boldsymbol{P}^{(t)}_{\mathrm{i,\,LTR}}$
3: Combine the sequence of LTR measurements.

$$\mathbf{LTR}_{\mathrm{v}} = \left[\boldsymbol{V}^{(1)}_{\mathrm{m,\,LTR}}\ \boldsymbol{V}^{(2)}_{\mathrm{m,\,LTR}}\ \cdots\ \boldsymbol{V}^{(ls)}_{\mathrm{m,\,LTR}}\right]$$

$$\mathbf{LTR}_{\mathrm{p}} = \left[\boldsymbol{P}^{(1)}_{\mathrm{i,\,LTR}}\ \boldsymbol{P}^{(2)}_{\mathrm{i,\,LTR}}\ \cdots\ \boldsymbol{P}^{(ls)}_{\mathrm{i,\,LTR}}\right] \tag{13}$$

GCN inference

4: Pass $\mathbf{LTR}_{\mathrm{v}}$ to the trained GCN for voltage magnitude with graph parameters $\mathbf{A}$, and get the corresponding output matrix $\mathbf{HTR}_{\mathrm{v}}$.
5: Pass $\mathbf{LTR}_{\mathrm{p}}$ to the trained GCN for pseudo-injection with graph parameters $\mathbf{A}$, and get the corresponding output matrix $\mathbf{HTR}_{\mathrm{p}}$.

Restore data from GCN representation

6: De-normalize data from [-1, 1].
7: Split the output matrix to sequence of measurements.
8: Convert from pseudo-injection representation back to power flow representation. So far we get the HTR measurements $\hat{\boldsymbol{V}}^{(t)}_{\mathrm{m,\,HTR}}, \hat{\boldsymbol{P}}^{(t)}_{\mathrm{lf,\,HTR}}$

State estimation

9: Pass the HTR measurements through SE, and get the final results $\boldsymbol{V}^{(t)}_{\mathrm{m,\,HTR}}, \boldsymbol{P}^{(t)}_{\mathrm{lf,\,HTR}}$

A. BASE CASE

In the base case, we report the performance of the proposed SR method in an intuitive way. In one test sample, the result of SR, the HTR labels, and the LTR inputs for two nodes and two edges are reported in Fig. 5. Fig. 5(a) reports the completion of measurement V_{m} on node 10. Measuring device on that node reports data at $1/2\,f_p$. As can be observed, the SR result recovers many details from the LTR inputs and approximates the HTR labels (the ideal measurements). Fig. 5(b) reports the completion of measurement V_{m} on node 30. Measuring device on that node report data at $1/4\,f_p$. Fig. 5(c) reports the completion of measurement P_{lf} on edge 10 with data reporting frequency $1/2\,f_p$. Finally, Fig. 5(d) reports the completion of measurement P_{lf} on edge 30 with data reporting frequency $1/4\,f_p$. As observed in Fig. 5, the proposed SR method can recover details from the LTR measurements.

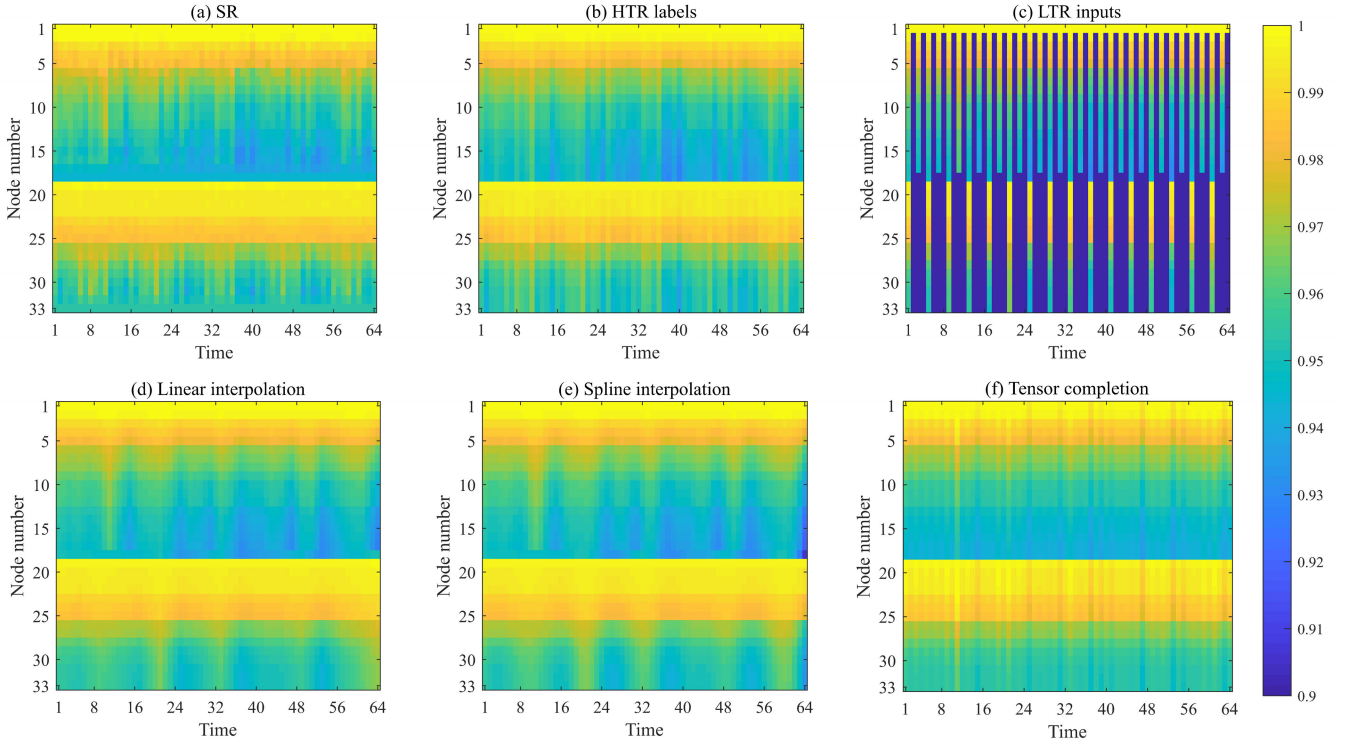


FIGURE 7. Performance of the proposed SR method on all nodes (Measurement type V_{m}): SR results, compared with linear interpolation, spline interpolation, and tensor completion.

TABLE 1. MSE in Tests With SR, Linear Interpolation, Spline Interpolation, and Tensor Completion.

Methods	MSE for completion of measurements V_{m}	MSE for completion of measurements P_{lf}	Overall MSE
SR	**0.0313**	**0.0078**	**0.0196**
Linear interpolation	0.0427	0.0102	0.0264
Spline interpolation	0.1029	0.0266	0.0647
Tensor completion	0.0385	0.0134	0.0259

The performance is impressive on nodes/edges with low data reporting frequency ($1/4\,f_p$).

Then, in Fig. 7, we report the completion of measurement V_{m} of all nodes. Fig. 7 (b) reports the HTR labels (the ideal measurements), Fig. 7 (c) gives the LTR inputs. It is padded with zeros to match the dimension of HTR labels. Compared with the results of the linear interpolation (Fig. 7 (d)) and the spline interpolation (Fig. 7 (e)), the SR results are less smooth but recover more detailed patterns, especially for nodes with low data reporting frequency ($1/4\ f_p$). However, we must report that on node 32 and 33, the SR results have slight deviation from the HTR labels, and perform less satisfying.

B. STATISTICAL RESULTS

In this subsection, we performed 1000 tests, and evaluate the proposed SR methods with normalized mean square errors (MSE),

$$\mathrm{MSE} = \frac{\sum_{i,j}(\hat{\mathbf{X}}_{i,j} - \mathbf{HTR}_{i,j})^2}{\sum_{i,j}(\mathbf{HTR}_{i,j})^2}, \tag{14}$$

where $\hat{\mathbf{X}}_{i,j}$ represents for the data recovered or completed with SR, linear interpolation or spline interpolation; $\mathbf{HTR}_{i,j}$ represents for the HTR labels, i.e., the ideal measurements. There are several popular SR metrics, e.g., PSNR. However, these metrics are suitable only for images. Nevertheless, the data in distribution network SR are not 2d images, but multiple 1d time series on a topological graph. Therefore, these metrics are not appropriate. Therefore, we choose the primitive and intuitive metric: normalized MSE. The MSE in tests with all three methods are reported in Table 1. The SR method out performs linear interpolation and spline interpolation in all three statistic criteria for more than 20%.

C. SR COMPARED WITH TENSOR COMPLETION

On the power system measurement completion problem, researchers also proposed methods based on the matrix completion [28] and tensor completion [29], [30]. These two methods both consider the low-rank property of measurements. In this subsection, the tensor completion method in [29] is implemented and compared with the SR method

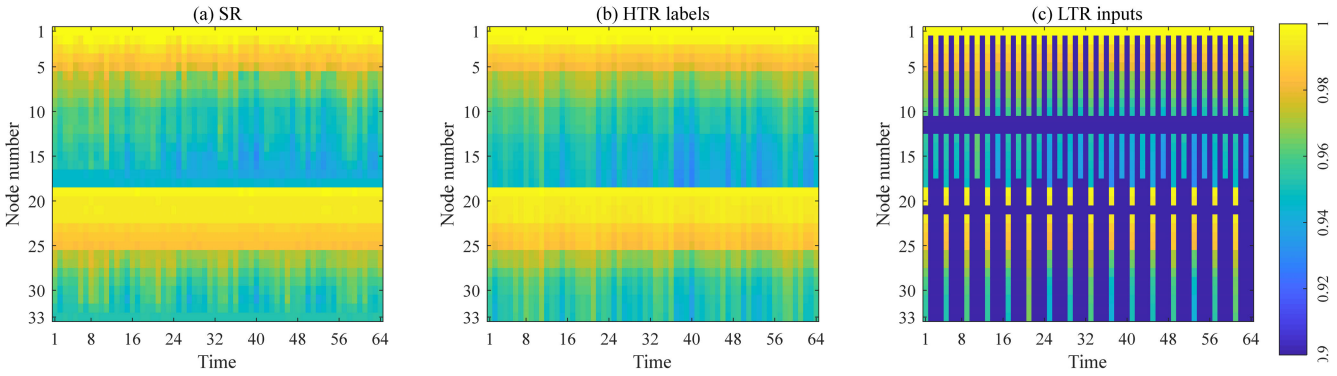


FIGURE 8. Performance of the proposed SR method on temporal-spatial completion when some nodes' data are completely missing (Measurement type V_m).

proposed in this article. The measurements are combined and converted to tensor representation. The weighted missing data polyadic decomposition [44] is performed on the data, i.e., the assembled tensors are decomposed into factor matrices. Finally, new tensors are restored from the factor matrices. So far, the missing data have been recovered in these new tensors.

The result of tensor completion is reported in Fig. 7(f), together with linear interpolation and spline interpolation. The tensor completion method indeed restored more details than linear or spline interpolation, because it exploits the low-rank property of power system measurements. The statistical results of tensor completion are also reported in TABLE 1. For the completion of Vm, the tensor completion method performs better than the interpolation methods. The SR method out performs the tensor completion method in all three statistic criteria for more than 10%. The proposed SR method considers both topology information in GCN and physical constrains in SE. Therefore, it performs better than the tensor completion methods in terms of accuracy. Nonetheless, the tensor completion and matrix completion methods are also highly valuable because of their high flexibility.

D. EXPERIMENT ON TEMPORAL-SPATIAL COMPLETION

Compared to the power transmission network, the distribution network has much more edges and branches. Therefore, it would be expensive to equip each node with highly reliable sensors. The scenario emerges where some nodes are not equipped with sensors or some sensors failed to report data.

In this subsection, we discuss this scenario with the temporal and spatial SR. The previous chapters discuss generating HTR results from LTR inputs, that is, to generate high-quality measurements (in terms of temporal resolution) from low-quality ones. In this subsection, the authors try to generate HTR results from LTR inputs even if there are no sensors at several nodes. The dataset is similar to that in the previous section. Nonetheless, data of several nodes are missing in the training features (inputs), while the training labels remain the same.

The results are reported in Fig. 8. In this case, data from nodes 11, 12, and 21 are set to be unavailable (see Fig. 8(c)). The SR results are reported in Fig. 8(a). As for statistical data in 1000 tests, the MSE for V_m is 0.0340, MSE for P_{lf} is 0.0099 and the overall MSE is 0.0219. It can be seen that with missing data the SR still functions well.

V. CONCLUSION

This article proposes a new temporal data completion method considering distribution system topology, including feature transformation, graph level GCN, and state estimation. 1) This article solves the problem that mixed node features and edge features are hard to train, which is accomplished by transforming edge features to node features using properties of the power distribution system. 2) The authors modify the first-order Chebyshev polynomials GCN for graph level regression. 3) The method considers physical constraints (results of GCN is tuned by an SE)

Tests show its superiority over different interpolation methods for more than 20% (overall MSE). The proposed method also slightly surpluses the tensor completion method. Moreover, it can also perform spatial-temporal SR when data of some nodes are completely missing. This method will help to improve the state awareness of distribution networks. We are also working on further researches about topology-robust graph SR.

REFERENCES

[1] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4681–4690.
[2] X. Wang, K. Yu, C. Dong, and C. Change Loy, "Recovering realistic texture in image super-resolution by deep spatial feature transform," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 606–615.
[3] H. Greenspan, "Super-resolution in medical imaging," *Comput. J.*, vol. 52, no. 1, pp. 43–63, Jun. 2009.
[4] T. Gönen, *Electric Power Distribution System Engineering*, vol. 2. Boca Raton, FL, USA: CRC Press, 2008.
[5] P. P. Barker, J. J. Burke, R. T. Mancao, T. A. Short, C. A. Warren, C. W. Burns, and J. J. Siewierski, "Power quality monitoring of a distribution system," *IEEE Trans. Power Del.*, vol. 9, no. 2, pp. 1136–1142, Apr. 1994.

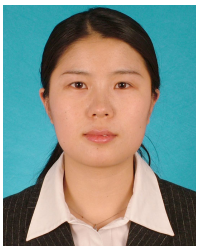
[6] H. Chen, L. Zhang, J. Mo, and K. E. Martin, "Synchrophasor-based real-time state estimation and situational awareness system for power system operation," *J. Mod. Power Syst. Clean Energy*, vol. 4, no. 3, pp. 370–382, Jul. 2016.
[7] Q. Dou, S. Wei, X. Yang, W. Wu, and K. Liu, "Medical image super-resolution via minimum error regression model selection using random forest," *Sustain. Cities Soc.*, vol. 42, pp. 1–12, Oct. 2018.
[8] A. Chaudhari, Z. Fang, H. L. Jin, G. Gold, and B. Hargreaves, "Deep learning super-resolution enables rapid simultaneous morphological and quantitative magnetic resonance imaging," in *Proc. Int. Workshop Mach. Learn. Med. Image Reconstruct.*, 2018, pp. 3–11.
[9] H. Huang and N. Wu, "Fast facial image super-resolution via local linear transformations for resource-limited applications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 10, pp. 1363–1377, Oct. 2011.
[10] T. Uiboupin, P. Rasti, G. Anbarjafari, and H. Demirel, "Facial image super resolution using sparse representation for improving face recognition in surveillance monitoring," in *Proc. Signal Process. Commun. Appl. Conf.*, 2016, pp. 437–440.
[11] L.-W. Kang, B.-C. Chuang, C.-C. Hsu, C.-W. Lin, and C.-H. Yeh, "Self-learning-based single image super-resolution of a highly compressed image," in *Proc. IEEE 15th Int. Workshop Multimedia Signal Process.*, Sep. 2013, pp. 224–229.
[12] Z. Zhang and V. Sze, "FAST: A framework to accelerate super-resolution processing on compressed videos," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Jul. 2017, pp. 19–28.
[13] X. Chen, G. Zhai, W. Jia, C. Hu, and Y. Chen, "Color guided thermal image super resolution," in *Proc. Vis. Commun. Image Process.*, 2017, pp. 1–4.
[14] C. N. Lu, J. H. Teng, and W. H. E. Liu, "Distribution system state estimation," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 229–240, Feb. 1995.
[15] V. Kekatos and G. B. Giannakis, "Distributed robust power system state estimation," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1617–1626, May 2013.
[16] H. Sun, F. Gao, K. Strunz, and B. Zhang, "Analog-digital power system state estimation based on information theory—Part I: Theory," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1640–1646, Sep. 2013.
[17] R. Morin, A. Basarab, and D. Kouame, "Alternating direction method of multipliers framework for super-resolution in ultrasound imaging," in *Proc. 9th IEEE Int. Symp. Biomed. Imag.*, May 2012, pp. 1595–1598.
[18] M. M. Islam, M. N. Islam, V. K. Asari, and M. A. Karim, "Single image super-resolution in frequency domain," in *Proc. Image Anal. Interpretation*, 2012, pp. 53–56.
[19] H. Demirel, S. Izadpanahi, and G. Anbarjafari, "Improved motion-based localized super resolution technique using discrete wavelet transform for low resolution video enhancement," in *Proc. Eur. Signal Process. Conf.*, 2015, pp. 1097–1101.
[20] D. Hasan and A. Gholamreza, "IMAGE resolution enhancement by using discrete and stationary wavelet decomposition," *IEEE Trans. Image Process.*, vol. 20, no. 5, pp. 1458–1460, May 2011.
[21] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.
[22] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 2472–2481.
[23] M. Haris, G. Shakhnarovich, and N. Ukita, "Deep back-projection networks for super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 1664–1673.
[24] D. Chao, C. L. Chen, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Proc. Eur. Conf. Comput. Vis.* Cham, Switzerland: Springer, Sep. 2014, pp. 184–199.
[25] Z. Wang, D. Liu, J. Yang, W. Han, and T. Huang, "Deep networks for image super-resolution with sparse prior," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 370–378.
[26] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 294–310.
[27] G. Liu, J. Gu, J. Zhao, F. Wen, and G. Liang, "Super resolution perception for smart meter data," *Inf. Sci.*, vol. 526, pp. 263–273, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025520302681, doi: 10.1016/j.ins.2020.03.088.
[28] R. Pourramezan, H. Karimi, J. Mahseredjian, and M. Paolone, "Real-time processing and quality improvement of synchrophasor data," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3313–3324, Jul. 2020.
[29] D. Osipov and J. H. Chow, "PMU missing data recovery using tensor decomposition," *IEEE Trans. Power Syst.*, vol. 35, no. 6, pp. 4554–4563, Nov. 2020.
[30] R. Madbhavi, H. S. Karimi, B. Natarajan, and B. Srinivasan, "Tensor completion based state estimation in distribution systems," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf. (ISGT)*, Feb. 2020, pp. 1–5.
[31] A. Santoro, D. Raposo, D. G. T. Barrett, M. Malinowski, R. Pascanu, P. Battaglia, and T. Lillicrap, "A simple neural network module for relational reasoning," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2017, pp. 4967–4976.
[32] D. Duvenaud, D. Maclaurin, J. Iparraguirre, R. Gómezbombarelli, T. Hirzel, A. Aspuru-Guzik, and R. P. Adams, "Convolutional networks on graphs for learning molecular fingerprints," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2015, pp. 2224–2232.
[33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. CVPR*, 2016, pp. 770–778.
[34] M. Gori, G. Monfardini, and F. Scarselli, "A new model for learning in graph domains," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jul. 2005, pp. 729–734.
[35] S. Franco, G. Marco, T. Ah Chung, H. Markus, and M. Gabriele, "The graph neural network model," *IEEE Trans. Neural Netw.*, vol. 20, no. 1, pp. 61–80, Dec. 2009.
[36] J. Bruna, W. Zaremba, A. Szlam, and Y. Lecun, "Spectral networks and locally connected networks on graphs," 2013.
[37] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129–150, Mar. 2011.
[38] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. 5th Int. Conf. Learn. Represent. (ICLR)*. Toulon, France: OpenReview.net, Apr. 2017. [Online]. Available: https://openreview.net/forum?id=SJU4ayYgl
[39] M. Zhao, M. D. Kaba, R. Vidal, D. P. Robinson, and E. Mallada, "Sparse recovery over graph incidence matrices," in *Proc. IEEE Conf. Decis. Control (CDC)*, Dec. 2018, pp. 364–371.
[40] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, "Geometric deep learning: Going beyond Euclidean data," *IEEE Signal Process. Mag.*, vol. 34, no. 4, pp. 18–42, Jul. 2017.
[41] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. Boca Raton, FL, USA: CRC Press, 2004.
[42] A. K. Al-Othman and M. R. Irving, "A comparative study of two methods for uncertainty analysis in power system state estimation," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1181–1182, May 2005.
[43] E. Ghahremani and I. Kamwa, "Local and wide-area PMU-based decentralized dynamic state estimation in multi-machine power systems," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 547–562, Jan. 2016.
[44] E. Acar, D. M. Dunlavy, T. G. Kolda, and M. Mørup, "Scalable tensor factorizations for incomplete data," *Chemometric Intell. Lab. Syst.*, vol. 106, no. 1, pp. 41–56, Mar. 2011.

ZHISHENG WANG received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2017. He is currently pursuing the master's degree with the Department of Electrical Engineering and Applied Electronic Technology, Tsinghua University. His research interests include cyber-physical system modeling and the cyber security of smart grid.

YING CHEN (Member, IEEE) received the B.E. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2001 and 2006, respectively.

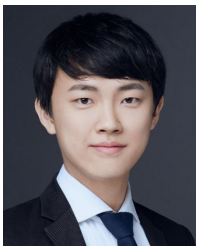
He is currently an Associate Professor with the Department of Electrical Engineering and Applied Electronic Technology, Tsinghua University. His research interests include parallel and distributed computing, electromagnetic transient simulation, cyber-physical system modeling, and the cyber security of smart grid.

SHAOWEI HUANG (Member, IEEE) received the B.S. and Ph.D. degrees from the Department of Electrical Engineering, Tsinghua University, Beijing, China, in July 2006 and June 2011, respectively. From 2011 to 2013, he held a postdoctoral position with the Department of Electrical Engineering, Tsinghua University, where he is currently an Associate Professor. His research interests include power systems modeling and simulation, power system parallel and distributed computing, complex systems and its application in power systems, and artificial intelligence.



XUEMIN ZHANG (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2001 and 2006, respectively.

She is currently an Associate Professor with the Department of Electrical Engineering, Tsinghua University. Her research interests include power system analysis and control, especially stabilization control, and cascading failure modelling and mitigation.



XIAOPENG LIU received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2017. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering and Applied Electronic Technology, Beijing Jiaotong University. His research interests include energy Internet and energy storage.