IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON THE INTERNET OF ENERGY: ARCHITECTURES, CYBER SECURITY, AND APPLICATIONS (PART II)

Received January 18, 2018, accepted February 15, 2018, date of publication March 5, 2018, date of current version April 4, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2812211

Efficient Forecasting Scheme and Optimal Delivery Approach of Energy for the Energy Internet

LIUFENG DU[1,2], LINGHUA ZHANG[1,3], XIYAN TIAN[2], AND JINHUI LEI[2]

[1]College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[2]School of Mechanical and Electrical Engineering, Henan Institute of Science and Technology, Xinxiang 453003, China
[3]Jiangsu Engineering Research Center of Communication and Network Technology, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Corresponding author: Linghua Zhang (zhanglh@njupt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771258 and in part by the Landmark Innovation Project of Henan Institute of Science and Technology.

ABSTRACT The energy Internet (EI) is an important infrastructure for effectively utilizing and intelligently managing renewable energy sources (RES). In this paper, we study the architecture design of the EI under the backdrop of large-scale RES grid connection and the efficient forecasting and optimal utilization of energy. The contribution of this paper is threefold. First, we design a hierarchical integration architecture for the EI and attempt to solve the issues of energy and information management that stem from large-scale RES grid connection. Second, we propose a novel energy forecasting scheme that significantly reduces the amount of effort and ensures the accuracy of formulating the energy forecasting as an instance of the matrix completion issue. Third, we take electric vehicle charging as a typical case and propose the use of reinforcement learning to achieve optimal energy delivery. An experimental evaluation of real-world data sets validates the expectations of the study and highlights the superiorities of our proposed approaches.

INDEX TERMS Energy internet, energy forecasting and delivery, matrix completion, reinforcement learning.

I. INTRODUCTION

To achieve the efficient exploitation and intelligent management of large-scale renewable energy sources (RES), the Energy Internet (EI), which integrates new energy technology and IT, was proposed [1], [2]. The EI controls various energy flows (mainly electrical energy, including thermal, chemical and electromagnetic energy) with information flows while ensuring the efficient management of the dual-flow of energy and information. More importantly, it enables the open interconnection and peer-to-peer sharing of energy. The EI is regarded as an update of the existing smart grid (SG) [3], and it can make up for the deficiencies of the SG in response to numerous complex generations and variable loads. Currently, research on the EI has drawn widespread attention from industry and academia [4], [5].

RES is the main energy source for the EI [6]; however, with the increases in RES generation units, the EI is bound to face two potential problems: 1) RES is scattered geographically and the output is uncontrollable; thus, from-top-to-bottom, unified energy management is unsuitable for the scenario of large-scale RES grid connection. 2) The demand for a highly intelligent EI will result in explosive growth in energy data, and a centralized information management mode is affected by data acquisition, storage and transmission issues [7]. Determining how to safely and steadily connect RES to the EI for efficient management and real-time energy sharing is the main problem facing EI architecture design, and our architecture is proposed to address this issue. First, by the support of advanced power electronics and energy storage technologies [8] and based on the principle of ''local consumption first surplus grid-connection,'' we design a structure of RES-based generation units to ensure the stability and controllability of dual-flow and achieve the distributed management of energy. Second, referencing a distributed computational framework from the Internet of things [9], [10], an architecture of hierarchical integration is designed to manage the dual-flow hierarchically and attempt to resolve the problems resulting from unified and centralized management. Next, under the proposed EI, we study the forecasting and delivery of energy.

Efficient and accurate energy forecasting is the premise of ensuring safe EI operation and intelligent energy management. Hence, we first explore the output power (OP)

forecasting of the E-microgrid (shown in Fig. 2, where ''E-'' denotes ''energy'') in our proposed EI. The E-microgrid OP in the paper refers to the generation surplus energy that meets the load requirements of the E-microgrid; briefly, it is the surplus power of total generation minus the total load within a E-microgrid. However, as the number of RES generation units increases, the corresponding number of E-microgrids also increases, which makes OP forecasting a difficult effort. Motivated by [11]–[13], we argue that the geographically close E-microgrids in our EI must exhibit potential spatiotemporally correlated behaviors. Therefore, based on the current advanced load and power forecasting technologies (see section II in detail), we first characterize the OPs of the E-microgrids for the same control center (called E-fog, Fig. 2) as a two-dimensional energy matrix, or E-matrix. Then, we fully demonstrate that the E-matrix has an approximate low-rank property via real-word traces [14]. Finally, we achieve efficient and reliable OP forecasting for the E-microgrids in the same E-fog by introducing matrix completion (MC) [15]. Ulteriorly, guided by the above forecasting scheme, we explore the optimal delivery of energy.

From the perspectives of both EI operators and energy consumers, the optimum utilization of energy is a win-win strategy [5]. By leveraging efficient OP forecasting, E-fog can improve green energy saving by making day-ahead energy delivery schedules for consumers. However, the randomness of some users' energy utilization, such as electric vehicle (EV) charging behaviors, can cause the actual demand trends to mismatch the delivery schedules, which can result in the frequent use of storage facilities or requests for backbone grid response. To achieve optimal energy delivery and take the following factors into account. 1) The EVs can serve as important energy storage facilities for the EI to balance RES generation fluctuations [16]. 2) The randomness of driver behavior leads to the extreme flexibility of charging activities and difficult in precisely modeling these systems, and the design of optimal energy use based on the EV charging control is challenging [17]. Therefore, we take the EV charging as a case study and propose a model-free control approach based on reinforcement learning (RL) [18] for accurate energy delivery. Compared to traditional approaches, our proposed method does not require much prior knowledge and precise modeling, and the control terminal can converge to the optimal solution through heuristic learning with a few parameters.

The main contributions of our work can be summarized as follows:

- We investigate the impact of large-scale RES grid connection on the EI, and in response to emerging energy and information management and transmission challenges, a hierarchical integration architecture for the EI is presented.
- Under our architecture, we propose a novel energy forecasting scheme based on MC, which efficiently generates a complete energy matrix based on partially acquired observations from various E-microgrids.
- We explore the challenges of massive EV charging in the EI and employ the RL-based technique to capture the optimal energy delivery, which minimizes the energy cost based on heuristic learning.
- We evaluate our claims based on real-world load and RES generation data sets, and the experimental results meet expectations and reflect the superiorities of our proposed scheme and approach.

The rest of this paper is organized as follows. Section II describes the related work. In Section III, we design a hierarchical integration architecture for EI. Section IV introduces a power forecasting scheme based on MC. In Section V, we take the EV charging control as an object and study the optimal delivery issue of energy using RL technique. Section VI presents the evaluation results by utilizing real-world data sets. Finally, we conclude in Section VII.

II. RELATED WORK

We divide the related work into three parts, the first part is the existing achievements about the EI architecture, the second investigates the application of spatio-temporal correlations in energy forecasting, and the last part, a survey of the literatures on employing RL for charging control is given.

A. EI ARCHITECTURE DESIGN

The concept of EI has been brought forward for more than ten years, but it is still in the exploration stage now. In [19], Huang *et al.* proposed an architecture, named as FREEDM, for RES generation based on the energy router (E-router) [20], which followed the core router of internet. In the case of large-scale RES being integrated into the EI, Bui *et al.* [21] studied the problem of improving the efficiency and optimizing the allocation of energy demand. Wang *et al.* summarized the security issues [22] and communication standards of the energy information network [23], and also explored the challenges about the design of communication system of EI. However, these efforts do not study the related problems of energy big data in depth, which stemmed from a large number of RES generation units and our architecture will fill up the deficiency.

B. ENERGY FORECASTING SCHEME

Promoted by the development of machine learning and big data analytics, the accuracy and precision of load or power forecasting are substantially improved [24]–[27]. These contributions, which are the great foundations for our scheme, adequately utilize the power of intelligent algorithm to solve nonlinear problems and reduce the dependence on historical data along with improving the adaptability for the small sample sets. In fact, a substantial amount of attention has been drawn to study power or load forecasting by using spatio-temporal property. Tastu *et al.* [11] studied spatio-temporal correlation effect on wind speed and direction in western Denmark, and proposed a nonlinear model to measure the forecasting errors for wind power. In [28], Tascikaraoglu and Sanandaji studied residential power demand using spatial and temporal information respectively, and implemented

the predicted objects recovery via the compressive sensing. In order to reduce the data volume of power load samples, Carreno *et al.* [29] proposed an allocation method which using the spatial correlation of new loads for load growth of city. Inspired by these works, we take full advantage of the spatio-temporal property of the E-matrix and reduce the amount of forecasting effort by exploiting the power of the MC.

C. EV CHARGING CONTROL APPROACH

The studies of charging approach for the EV can be broadly split into two groups: one is based on the deterministic algorithms and the other depends on the heuristic ones [30].

In the first group, Fan [31] modeled the user's charging preferences which based on dynamic real-time electricity prices, and minimized the cost of EV charging. Rücker *et al.* [32] used the battery lifetime as an objective function and minimized the total operating cost of a fleet employing the nonlinear programming. However, the above efforts did not consider the impact of the strong randomicity and uncertainty of EV charging behavior, and the optimization effect is overly relay on the accuracy of the build model.

In the second, Shi and Wong [33] formulated the real-time vehicle-to-grid control problem as a Markov decision processes (MDP), and took the uncertain electricity price as a Markov chain with unknown transition probabilities, but they considered the MDP with only one EV. In a scenario with RES supplies, Vandael *et al.* [17] considered the unknown EV charging flexibility and built a model via the MDP, and took the state of charge (SoC) of the battery of the EV as the state space, but the optional actions were only classified into three categories. Our work propose an overall control approach for charging EVs, which achieves a fine control for the charging behavior via designing the state and action space carefully. A multi-agent-based distributed Q-Learning algorithm [34] is employed to search for the optimal control policy.

III. EI ARCHITECTURE DESIGN

Electric power is the main carrier of energy, and the large-scale RES will provide most of the energy for the future EI. To steadily leverage the large-scale RES and account for the balance between the energy supply and demand and the flexibility of energy interactions, we first design a structure for RES-based generation. This structure is known as an E-edge unit, and it uses the E-router as the dual-flow master node, as shown in Fig. 1.

In Fig. 1, E-edge unit generation occurs through its own RES supply side, and after being processed by the terminal controller, the generated energy is fed to the local demand side, and the surplus is stored in the local E-storage facility or transported to E-microgrids via bidirectional converters [35] and the E-router to participate in management and interactions. Bidirectional converters have dual roles of input rectifiers or output inverters, and the E-router, which is based on FREEDM [19], is the key hub of dual-flow interactions. The associated information module enables two-way

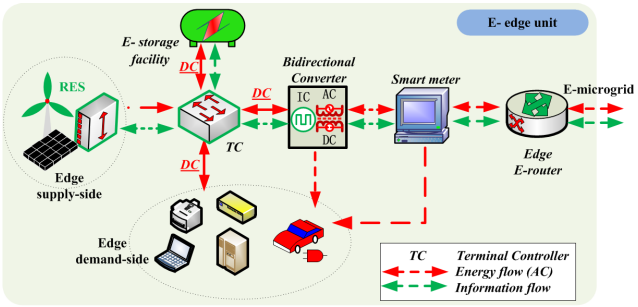


FIGURE 1. Structure of the E-edge unit.

communication and data storage, and the core electrical module, a solid-state transformer, enables the intelligent conversion and control of multiple power forms at multiple voltage levels, which ensures energy-flexible grid connection.The designed E-edge unit enables "plug and play" for the internal supply and demand side and storage facility, allows energy transmission from the E-microgrid and guarantees the free sharing of energy. Furthermore, the output or input quantity of energy can be metered using the smart meter [36] and settled based on real-time price.

The essence of the EI, which is different from the SG, is that it enables open, peer-to-peer power sharing between distributed generation units; thus, it is especially important to achieve efficient information flow and processing in the grid. Considering the large size of the energy field and the increase in the number of E-edge units, the information processing system of the EI will face pressure from big data [7]. Traditional centralized tightly coupled architecture inhibits the real-time management and sharing experiences of users due to data congestion and delayed response.

To overcome these disadvantages, we present an architecture of hierarchical integration and divide the overall architecture of the EI into three layers: the E-edge layer, regional E-fog and cloud control center at the top. The energy interaction and part of the data stored within the first two levels are completed on the spot, which reduces the data storage of the cloud center and the data circulation of the main communication link and increases the real-time efficiency throughout the network. The architecture is shown in Fig. 2.

In Fig. 2, as the most basic energy unit, the E-edge is both the demand side and the most important supply side of energy in the EI. Due to low productivity and scattered and huge numbers of E-edges, before integrating into the upper layer, multiple geographically proximal E-edges form a group to connect the upper layer, i.e., the upper layer regional E-fog through an E-microgrid. The E-microgrid is a dual-flow management center in the E-edge layer and is responsible for the following tasks: 1) operating in grid-connected mode, managing the external energy supply and demand of each unit in the group, and the in situ storing of information associated with the E-edge units. This approach ensures internal consumption and sharing, and the surplus assists in the energy distribution under the command of E-fog.

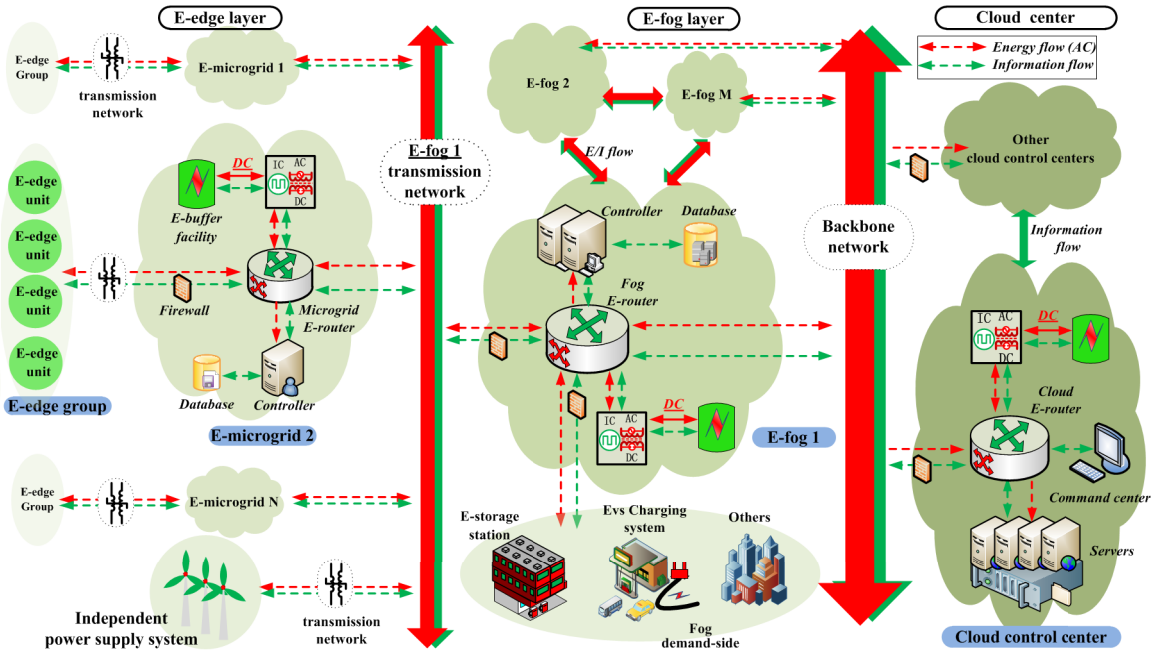


FIGURE 2. The hierarchical integration architecture of the EI.

2) Island mode coordinates the demand of group members to ensure an internal supply-demand balance of energy by leveraging controllers and the storage facility. In addition, an independent power generation system (IPGS), which consists of a large power generation plant, is connected to E-fog to ensure safe and stable operation.

The E-fog layer is a regional dual-flow control center and is responsible for geographically proximal E-microgrids. First, it has a function similar to the E-microgrid in the management of the dual flow but on the E-microgrid level; thus, is associated with a stronger control center and storage facility. Second, E-fog supplies gathered energy to regional large-scale demand sides, such as office buildings, factories, charging stations, etc., and it is the main place for EI earnings. Finally, the scheduling and interconnection of energy between E-fogs is controlled by the cloud center through the backbone network, and it is responsible for continuously accepting and integrating most of the traditional generation systems with the evolution of the EI.

As a top command and information processing center, cloud control centers only conduct information-level interactions with others. The cloud center is responsible for regional energy distribution and coordination and for the completion of storage and analysis of most energy big data. Additionally, it provides computational resource support for dynamic assessment, the balance of the supply and demand, grid monitoring and maintenance, assessment and adjustment of the operating service quality, and the security and protection of the network.

Our constructed EI architecture is based on the concept of open and peer-to-peer information/energy integration, which attempts to alleviate the issues caused by the large-scale RES grid connection and solve the real-time and dynamic utilization issues of distributed RES. In the background of such an architecture, we further conduct research on energy forecasting and optimal delivery.

IV. EFFICIENT ENERGY FORECASTING SCHEME

In this section, with the analysis of the low-rank property of the E-matrix, we propose an output power forecasting method.

A. LOW-RANK PROPERTY OF THE E-MATRIX

Under the same E-fog, we infer that the load and power generation in different E-microgrids have spatio-temporal correlation behaviors because of their geographical proximity, which must result in a potentially low-rank property of the E-matrix. To support the above hypothesis, the following works were carried out.

First, for the one-day OP of each E-microgrid in the same E-fog, values are assigned in a two-dimensional spatio-temporal energy matrix, which is the E-matrix and can be expressed as:

$$E^{mat} = \begin{bmatrix} p^{grd}_{1,1} & p^{grd}_{1,2} & \cdots & p^{grd}_{1,\mathcal{D}} \\ \vdots & \vdots & p^{grd}_{s,d} & \vdots \\ p^{grd}_{\mathcal{S},1} & p^{grd}_{\mathcal{S},2} & \cdots & p^{grd}_{\mathcal{S},\mathcal{D}} \end{bmatrix} \tag{1}$$

The rows and columns of E^{mat} characterize the space and time behaviors of the E-matrix, respectively. $\mathcal{S}$ represents that an E-fog manages $\mathcal{S}$ E-microgrids, and $\mathcal{D}$ means that a day is divided into $\mathcal{D}$ measurement stages. Each element $p^{grd}_{s,d}$ in the E-matrix indicates the current OP value of E-microgrid s

at time d, which can be defined as:

$$p_{s,d}^{grd} = p_{s,d}^{gen} - p_{s,d}^{loa} \quad \begin{matrix} \forall s \in \{1, 2, \ldots, \mathcal{S}\} \\ \forall d \in \{1, 2, \ldots, \mathcal{D}\} \end{matrix} \tag{2}$$

where $p_{s,d}^{gen}$ is the total power generation of the RES and $p_{s,d}^{loa}$ is the total electricity demand loads of the E-microgrid. Under our proposed architecture, because of the existence of the IPGS (reference Fig. 2), $p_{s,d}^{grd}$ can be set to greater than 0. Therefore, E-fog only considers the case of the E-microgrid as a net OP source.

Next, to acquire the E-matrix for the scenario we established, we combined the real-world traces, and those from PJM's Data Miner [14] contain metered load and RES generation data from 2016 and geographic location information. These data are used to create 366 one-day energy matrices according to (1) and (2). The reason that these matrices can be used to represent the E-matrix of the E-fog is that the sources of the above data are geographically close [37]. We take these E-matrices as the object based on their spatio-temporal correlation behaviors, and the assumption that the E-matrix has potential low-rank properties will be gradually confirmed below.

We randomly choose one matrix from the created 366 E-matrices and plot its color map, as shown in Fig. 3 (a). The size of the matrix is 69×24, i.e., in our scenario, there are 69 E-microgrids, and the one-day OP forecasting occurs at an hourly time scale.

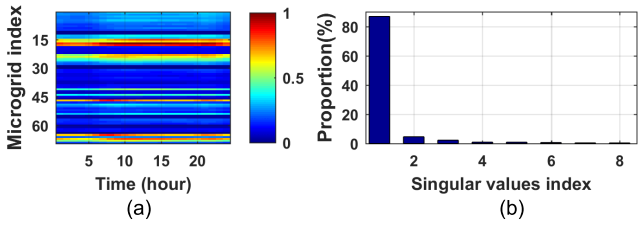


FIGURE 3. Color map and singular values of the E-matrix. (a) Color map. (b) Normalized singular values.

In Fig. 3 (a), it can be seen that 1) the OPs from the same E-microgrid but at different times are highly correlated and 2) the different E-microgrids tend to generate similar $p_{s,d}^{grd}$ values at the same time. The correlation behaviors among the OPs of different E-microgrids indicate that the degrees of freedom of the E-matrices are much smaller than their dimensions, and low degrees of freedom lead to the E-matrix exhibiting a low-rank property, which is the necessary condition for applying matrix theory. Furthermore, we employ the singular-value decomposition (SVD) to confirm the low-rank property of the E-matrix with sparse prior information. The process of SVD is described as:

$$E^{mat} = \sum_{i=1}^{\min(\mathcal{S},\mathcal{D})} u_i \cdot \sigma_i \cdot (v_i)^{\mathrm{H}} \tag{3}$$

where σ_i denotes the singular value of E^{mat}, u_i and v_i are the left- and right-singular row vectors of σ_i respectively. The result of the decomposition of the above E-matrix by SVD is shown in Fig. 3 (b).

Fig. 3 (b) illustrates that the energy of the E-matrix is mainly concentrated on the first few singular values, and the ratio is over 80%. Additionally, the noise interference factors contribute to the remaining singular values that exhibit low proportions. The approximate low-rank property of each E-matrix will play a significant role in our work. A low-rank property is universal among our E-matrices; therefore, we decomposed all 366 E-matrices by leveraging SVD and calculated the ratio of the sum of the first two singular values $(\sigma_1 + \sigma_2)$ for each matrix, The results are shown in Fig. 4.

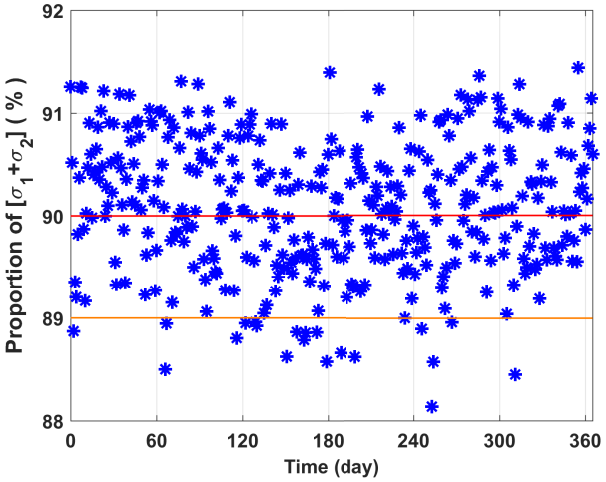


FIGURE 4. The proportion of the sum of the first two singular values.

Fig. 4 illustrates that the number of matrices with a proportion of more than 0.90 accounts for approximately 58% of all matrices, and those with a proportion greater than 0.89 account for over 94%. This result suggests that the approximate low-rank property is ubiquitous in our E-matrices. Based on this conclusion, we discuss the proposed efficient OP forecasting scheme.

B. MC-BASED EFFICIENT FORECASTING SCHEME

According to the contributions of known works [15], [38], [39], if a matrix has a sparse priors of low-rank property and satisfies the constraints of restricted isometry property and uniform random sampling, then the exact reconstruction of the matrix can be achieved. The reconstruction process, i.e., matrix completion is an affine rank minimization problem subject to random constraints, which can be described as:

$$\begin{aligned} \underset{\hat{E}^{mat}}{\textbf{minimize}} \quad & \mathrm{rank}(\hat{E}^{mat}) \\ \textbf{subject to} \quad & \hat{E}_{s,d}^{mat} = E_{s,d}^{mat}, \quad (s,d) \in \Omega \\ & \text{or} \ \ \mathbf{F}_{\Omega}(\hat{E}^{mat}) = F_{\Omega}(E^{mat}) \end{aligned} \tag{4}$$

where $\hat{E}^{mat} \in \mathbb{R}^{\mathcal{S}\times\mathcal{D}}$ is the target matrix and we suppose $\mathcal{S} \geq \mathcal{D}$, E^{mat} represents the actual matrix to be completed, $\Omega \subset [\mathcal{S}] \times [\mathcal{D}]$ denotes a set of observation coordinates for

E^{mat} and $|\Omega| \ll (\mathcal{S} \times \mathcal{D})$ denotes the number of observations. $\mathbf{F}_\Omega(\cdot)$ denotes the projection mapping as follows:

$$\mathbf{F}_\Omega(E^{mat}) = \begin{cases} E^{mat}_{s,d} & (s,d) \in \Omega \\ 0 & (s,d) \notin \Omega \end{cases} \tag{5}$$

In our scenario, E^{mat} represents the E-matrix containing only $|\Omega|$ E-microgrids' OP, and $\hat{E}^{mat}$ denotes the reconstruction E-matrix using the MC; the values of $\mathcal{S}$ and $\mathcal{D}$ are 69 and 24, respectively.

Problem (4) is an NP-hard problem due to the non-convex property of the operation of the low-rank matrix. From the point of nuclear norm minimization, Candès and Recht [15] and Cai *et al.* [39] performed leveraged convex relaxation to formulate the problem as follows:

$$\begin{aligned} &\underset{\hat{E}^{mat}}{\textbf{minimize}} \quad \left\| \hat{E}^{mat} \right\|_* \\ &\textbf{subject to} \quad \mathbf{F}_\Omega(\hat{E}^{mat}) = \mathbf{F}_\Omega(E^{mat}) \end{aligned} \tag{6}$$

where $\left\| \hat{E}^{mat} \right\|_* = \sum_{i=1}^{\mathcal{D}} \sigma_i(\hat{E}^{mat})$ and $\|\cdot\|_*$ denotes the nuclear norm, σ_i is the ith singular value of $\hat{E}^{mat}$. In addition, Candès et al. proved that when the number of observations is satisfied, $|\Omega| \geq c\mathcal{S}^{1.2} r \log \mathcal{S}$, the original matrix E^{mat} can be accurately reconstructed with high probability via equation (6).

As the nuclear norm is non-linear and non-smooth but convex, equation (6) can be formulated as a semi-definite programming (SDP) [40]. Moreover, different iterative methods, such as singular value thresholding (SVT) [39] and atomic decomposition for minimum rank approximation (ADMiRA) [41], are also employed to seek the optimum solution. In the experimental part of our paper, we evaluate some typical algorithms for matrix reconstruction, and a suitable algorithm is selected.

V. OPTIMAL ENERGY DELIVERY APPROACH

In this section, we take the EV charging control as a case study and propose an RL-based energy delivery approach. We ensure that the aim of the agent agrees with the objective, which minimizes the deviation between planned energy delivery and actual usage by devising the state and action space along with the reward function carefully, and a training method based on distributed Q-Learning is proposed. We group the design process into four parts.

A. DAY-AHEAD ENERGY DELIVERY SCHEDULE

On the forecasting day, according to the one-day OP forecasting results in section IV, E-fog aggregates the OP of each E-microgrid within its jurisdiction as follows:

$$\mathbb{E}_{total} = \{\sum_{s=1}^{\mathcal{S}} p^{grd}_{s,1}, \sum_{s=1}^{\mathcal{S}} p^{grd}_{s,2}, \ldots, \sum_{s=1}^{\mathcal{S}} p^{grd}_{s,\mathcal{D}}\} \tag{7}$$

Then, E-fog's charging control center (CCC), which is capable of obtaining the charging data, including the number of requests, initial SoC and battery parameters of the EV, combines $\mathbb{E}_{total}$ and the history of charging information to plan the energy delivery schedule for the EVs. Note that an accurate delivery schedule is not essential in our work, and many previous research efforts have yielded excellent results in this area [17], [33]. Hence, our work does not discuss the delivery schedule in detail and in the experimentation section, we only give approximate delivery curves based on different scenarios to evaluate the proposed approach. An illustration of the energy delivery schedule is shown in Fig. 5.

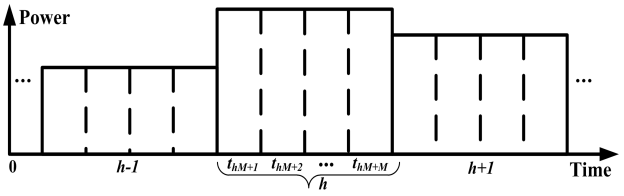


FIGURE 5. Time division of an energy delivery schedule.

In Fig. 5, the energy delivery stage is split into H supply periods and the time interval Δh between adjacent supply periods is set to 1 hour. The power schedule for each supply period is described as:

$$\begin{aligned} \mathbb{P}_{FS} &= \sum_{x=0}^{H-1} \{\mathrm{P}^x_{FS}[\varepsilon(h-x) - \varepsilon(h-x-1)]\} \\ &= \{\mathrm{P}^0_{FS}, \mathrm{P}^1_{FS}, \ldots, \mathrm{P}^h_{FS}, \ldots, \mathrm{P}^{H-1}_{FS}\} \end{aligned} \tag{8}$$

where $\forall h \in \mathbb{H} = \{0, 1, \ldots, H-1\}$, the $\boldsymbol{\varepsilon}(\cdot)$ denotes a step function and can be expressed as:

$$\boldsymbol{\varepsilon}(x) = \begin{cases} 1, & x \geq 0 \\ 0, & others \end{cases} \tag{9}$$

Furthermore, each h is segmented into M charging time slots t, $\forall t \in \mathbb{T} = \{t_1, t_2, \ldots, t_{H \cdot M}\}$, and the temporal resolution of each slot is defined as $\Delta t = \Delta h / M = 1/M$. In this way, a whole energy delivery stage can be discretized into $H \times M$ charging slots. At the start of every slot, the charging control decisions are made, and the charge piles are assumed to have constant OP throughout every slot.

B. CHARGING PROCESS FORMULATION

According to the currently known scenarios of EVs that have been applied for charging, the CCC of an E-fog develops a power distribution planning for each EV before starting each charge slot as follows:

$$\mathrm{P}^t_{pla} = \{p^t_{1,pla}, p^t_{2,pla}, \ldots, p^t_{n,pla}, \ldots, p^t_{\tilde{N}_t,pla}\} \tag{10}$$

where $p^t_{n,pla}$ denotes the plan charging power. Additionally, $\forall n \in \mathbb{V} = \{1, 2, \ldots, \tilde{N}_t\}$, where $\tilde{N}_t$ denotes the number of charging EVs in slot t. Considering that the fully charged EVs leave the charging queue or the EVs with low battery join after each charging slot, $\tilde{N}_t$ is a dynamic value and is subject to the constraint of $\tilde{N}_t \in [1, N_{\max}]$, where $N_{\max}$ denotes

the total number of charging piles within the control range of the CCC. The initial charging power for each EV can be expressed as:

$$p^t_{n,pla} = \mathbf{F}_{pla}(e^t_n, \tilde{N}_t)\mathrm{P}^t_{FS} = \frac{\mathbf{sgn}(e_{\max} - e^t_n)}{[\sum_{n=1}^{\tilde{N}_t} \mathbf{sgn}(e_{\max} - e^t_n)/e^t_n]e^t_n}\mathrm{P}^t_{FS}\Bigg|_{t=0} \quad (11)$$

We consider that the EVs are homogeneous and have the same maximum usable battery capacity $B_{\max}$, and the $e^t_n \in [0.2B_{\max}, 0.8B_{\max}]$ in general [30]. For convenience, we map the interval $[0.2B_{\max}, 0.8B_{\max}]$ to $[0, e_{\max}]$. The function $\mathbf{sgn}(\cdot)$ which is defined as follows, is used to eliminate the fully charged EV.

$$\mathbf{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1 & x < 0 \end{cases} \quad (12)$$

When slot t ends, the SoC of EVs can be increased as:

$$e^{t+1} = e^t + p^t_{act}/M \quad (13)$$

the notation p^t_{act} represents the actual charging power of these EVs:

$$\mathrm{P}^t_{act} = \mathbf{F}_{act}(\mathrm{P}^t_{FS}, \mathrm{P}^t_{pla}, \Omega_{ev}, \tilde{N}_t) = \{p^t_{1,act}, p^t_{2,act}, \ldots p^t_{n,act} \ldots, p^t_{\tilde{N}_t,act}\} \quad (14)$$

where p^t_{act} is calculated based on the current values of P^t_{FS} and P^t_{pla} and the number of charging EVs $\tilde{N}_t$, in addition to the parameters of the batteries $\Omega_{ev} = \{e_{\max}, p_{\min}, p_{\max}\}$; $p_{\min}$ and $p_{\max}$ denote the lower and upper bounds of the EV charging power, respectively; and $\mathbf{F}_{act}(\cdot)$ is a policy function included in **Algorithm 1** to ensure that each request based on the SoC is considered based on the supply power P^h_{FS}. Based on the above stipulations, before starting the next charging slot, the CCC updates the distribution plan according to the following formula:

$$\mathrm{P}^{t+1}_{pla} = \mathbf{F}_{upd}(\mathrm{P}^{t+1}_{FS}, \mathrm{P}^t_{pla}, e^t, \Omega_{ev} \Big| t \in \mathbb{T}) \quad (15)$$

where $\mathbf{F}_{upd}(\cdot)$ represents an update function for the plan power, and its calculation is given in **Algorithm 1**. We ignore the time required to execute the updated plan because it is much smaller than the time slot length, and when the SoC of an EV reaches the maximum in slot t, the charging channel is automatically cut off.

The purpose of formulating the control problem is to minimize the energy cost caused by the deviation between the planned and actual power, which we call the charging power deviation (CPD). Then, we discuss how to formulate the mathematical model for this purpose.

During the EV charging stage, the CPD in a slot is directly related to the E-fog's cost, and the actual deviation comes from two factors.

Algorithm 1 Plan Power Update Function

Input: Parameters p^t_{pla}, $\tilde{N}_t$, Ω_{ev}, P^t_{FS} and M.
Output: Parameter p^{t+1}_{pla}.
Initialization: p^0_{pla} and loop count $Lc = M$ for each h.

```
for t = 1, 2, ..., Lc do
    Compute: p^t_pla
    if p^t_pla ∉ [p_min, p_max] then
        p^t_act ← [p_min, p_max]
    else p^t_act ← p^t_pla
    end if
    Compute: e^{t+1} by using equation (13).
    Taking into account the real world scenario:
    if e^{t+1} > e_max then
        Removing the EV and Ñ_t ← Ñ_t − 1
    end if
    Compute: p^{t+1}_pla by using equation (11).
end for
```

- If the plan charging power is greater than the actual power, the surplus energy must be stored in the storage station or exported to the grid, which creates storage or transmission costs. We denote the unit price of surplus energy as β^-_{sur}/kWh, and the total loss is given as:

$$C^-_{sur} = \{-[\sum_{t\in\mathbb{T}}\sum_{n\in\mathbb{V}}(p^t_{n,act} - \mathrm{P}^t_{FS})]\Delta t\}\beta^-_{sur} \quad (16)$$

- Conversely, when the planned generation is less than the actual generation, E-fog must bear the cost of making up for the energy shortage, and the unit price of an energy shortage is denoted as β^-_{sht}/kWh; therefore, the following equation can be formed as:

$$C^-_{sht} = \{[\sum_{t\in\mathbb{T}}\sum_{n\in\mathbb{V}}(p^t_{n,act} - \mathrm{P}^t_{FS})]\Delta t\}\beta^-_{sht} \quad (17)$$

Obviously, maximizing revenues is achieved by minimizing C^-, and our aim is to find a charging control approach to minimize the CPD of EVs. Before exploring this approach, we first calculate the theoretical optimal value of the deviation, which can be used as a baseline to evaluate the performance of our proposed approach. Substituting $\Delta t = 1/M$ into (16) and (17), and let $\Delta\mathrm{P}^t \triangleq \mathrm{P}^t_{act} - \mathrm{P}^t_{FS}$ denotes the CPD. Thus, the optimization problem can be formulated as follows.

$$\underset{\Delta\mathrm{P}^t}{\mathbf{minimize}} \ \frac{1}{2M}\sum_{t\in\mathbb{T}}\sum_{n\in\mathbb{V}}\{[1 - \mathrm{sgn}(\Delta p^t_n)]\beta^-_{sur} + [1 + \mathrm{sgn}(\Delta p^t_n)]\beta^-_{sht}\}\Delta p^t_n \quad (18)$$

subject to

$$p_{\min} \le p^t_{n,act} \le p_{\max}, \quad \forall n \in \mathbb{V}, \forall t \in \mathbb{T} \quad (19a)$$
$$0 < e^t_n \le e_{\max}, \quad \forall n \in \mathbb{V}, \forall t \in \mathbb{T} \quad (19b)$$
$$1 \le |\mathbb{V}| \le N_{\max}, \quad H \le |\mathbb{T}| \le HM \quad (19c)$$

In our optimization problem, which is described in (18), **sgn**$(\cdot)$ guarantees that both surplus and shortage situations do not occur simultaneously, which makes the formulation agree with real-world scenarios. The optimization problem can be solved after being relaxed as $H \times M$ integer linear programming subproblems [43], and the smallest one of these $H \times M$ solutions is the optimal solution of the problem. We can employ an off-the-shelf and free solver (e.g., **CVX** [44] or **Ip_solve** [45]) to solve the objective function of the optimization.

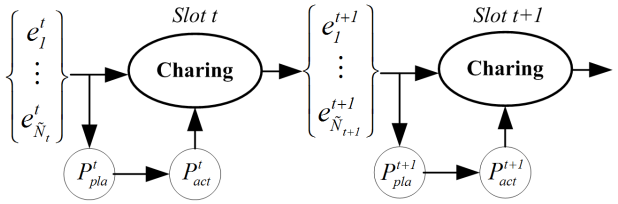


FIGURE 6. The EVs charging process.

C. MDP MODEL

According to the above sections, the charging process of the EVs can be described in Fig. 6. This process can be modeled as a finite discrete-time MDP defined by four tuples $< \mathbb{S}, \mathbb{A}, \Pr, R >$. $\mathbb{S}$ and $\mathbb{A}$ denote the state and action space, respectively; Pr is the probability of transition; and the reward function is represented by R.

1) STATE SPACE

During the EV charging stage, the number of charging EVs changes, as do the transitions of charging slots, i.e., parameter $\tilde{N}_t$ changes. These variations in turn lead to changes in the dynamic parameters of the charging model, so that $\tilde{N}_t$ directly affects the match between the energy delivery curve and the actual consumption curve, i.e.,. the minimization of the energy cost. Hence, we choose the number of charging EVs as the state space $\mathbb{S}$ of our MDP. Moreover, another advantage is that such a state space is simple, discrete and finite. The state space is defined as follows.

$$\mathbb{S} \triangleq \langle s_i \in \{\tilde{N}_t\} \Big| 1 \le \tilde{N}_t \le N_{\max} \rangle, \quad |\mathbb{S}| = N_{\max} \tag{20}$$

2) ACTION SPACE

We use π to denote the policy in the current state s_i. Combined with the function $\mathbf{F}_{pla}(\cdot)$ in equation (11), the charging action for an EV can be described as follows:

$$\begin{aligned} a_n^t &= \pi(s_i | p_{\min}, p_{\max}, \Psi) = \mathbf{F}_{pla}(e_n^t, s_i) \odot \Psi \\ &= \begin{cases} \{a_n^{t,j}\} & p_{\min} < a_n^{t,j} < p_{\max} \\ \boldsymbol{\varepsilon}(a_n^{t,j} - p_{\max})p_{\max} & \\ \quad +\boldsymbol{\varepsilon}(p_{\min} - a_n^{t,j})p_{\min} & a_n^{t,j} \ge p_{\max} \mathbf{or} a_n^{t,j} \le p_{\min} \end{cases} \end{aligned} \tag{21}$$

where $\Psi = (\psi_1, \psi_2, \ldots \psi_J | 0.5 \le \psi_j \le 1.5, \psi \in \mathbb{R})$ is the fine-grained control factor (FCF), which reduces the CPD and promotes the minimization of the energy cost. Combining equations (11), (14) and **Algorithm 1**, we know that a_n^t is a refined distribution of the actual charging power of the EV, and $|a_n^t| = J$. The action space is described as:

$$\begin{aligned} \mathbb{A} \triangleq \langle a_j = \{a_n^{t,j}, p_{\min}, p_{\max}\} \Big| & p_{\min} \le a_n^{t,j} \le p_{\max}, \\ & n \in \mathbb{V} \rangle, \quad |\mathbb{A}| = \tilde{N}_t \times J \end{aligned} \tag{22}$$

3) REWARD FUNCTION

Based on the smallest CPD required to obtain the maximum reward value, combining equations (8), (18) and (21), we can derive the reward function of the MDP as follows.

$$\begin{aligned} R(s_i^t, a_j^t) &= \exp(\frac{-\xi \left|\Delta \mathrm{P}^t\right|}{M}) \\ &\Rightarrow \exp[-\xi \left| \mathrm{P}_{FS}^t - \sum_{n \in |s_i^t|} (a_n^t | \psi_j = a_j^t) \right|] \end{aligned} \tag{23}$$

In (23), the exponential function ensures that the smallest CPD gets the maximum reward. ξ denotes the convergence rate control factor and by adjusting its value, a balance can be achieved between the speed and accuracy of the algorithm. Note that because the price parameter and the total number of slots are constant, they have no effect on the optimization; therefore, we ignore them in equation (23).

D. LEARNING ALGORITHM

In our MDP, both the state and action space dimensions directly rely on the number of charging EVs, thus a large number of EVs will incur the curse of dimensionality. To alleviate this problem, we leverage the multi-agent-based distributed-Q learning algorithm to improve the efficiency. First, the agent is decomposed into J independent learners with the same state space and reward function, and the corresponding action space is as:

$$\begin{aligned} \mathbb{A} &= \{\mathrm{A}_1, \mathrm{A}_2, \ldots, \mathrm{A}_j, \ldots, \mathrm{A}_J\} \\ \mathrm{A}_j &= \{a_n^{t,j}, p_{\min}, p_{\max}\}, \quad |\mathrm{A}_j| = \tilde{N}_t \end{aligned} \tag{24}$$

Each agent explores the environment and updates its own state-action value function $Q(s, a)$ independently in a supply period h. At the end of h, the optimal exploitation policy is selected through "exchange experience" (i.e., compare $Q(s, a)$ among agents). The update rule for each Q_j is as follows:

$$\begin{aligned} {Q_j}^{t+1}(s^t, a^t) = (1-\alpha) Q_j(s^t, a^t) + \alpha [R(s^t, a^t) \\ + \gamma \Pr(s^{t+1} \Big| s^t, a^t) \max {Q_j}^{t,\pi}(s^{t+1}, a^{t+1})] \end{aligned} \tag{25}$$

where $a^t, a^{t+1} \in \mathrm{A}_j$; $\gamma \in (0, 1]$ is the discount factor and $\alpha \in (0, 1]$ is the learning rate, Pr obeys the Poisson distribution with the rate parameter λ_{arr}; and $\max(\cdot)$ indicates that the convergence of the Q-matrix using ε–greedy policy. Q_j is described as:

$$Q_j \in \mathbb{Q}_{mul} = \{Q_1, Q_2, \ldots, , Q_j \ldots, Q_J\} \tag{26}$$

For the partially state cooperation RL, the greatest mass strategy is employed to select the optimal action, which is given as:

$$a^h = \arg\max_{a_s \in \mathrm{A}} \mathbf{E}[\sum_{esp=1}^{M} \mathbb{Q}_{mul}^{esp}(s, a_s)] \tag{27}$$

where $\mathbb{Q}_{mul}^{esp}$ denotes the Q value of each slot for each agent within supply period h. Exploiting the optimal action and combined with the reward of the current state space to update the global $Q(s^h, a^h)$, the final policy is described as:

$$\begin{aligned} \pi^*(s^t) &= \arg\max_{a_s \in \mathrm{A}} \mathbb{Q}_{mul}(s^t, a^t) \\ &= \arg\max_{a_s \in \mathrm{A}} \mathbb{Q}_{mul}(s^t, p_{act}^t \psi_j) \end{aligned} \tag{28}$$

Algorithm 2 Multi-Agent-Based Distributed-Q Learning

Input: EV parameters, power schedule set $\{\mathrm{P}_{FS}^t\}$, $H, M, \Psi_J, \xi, \alpha, \gamma$ and the end condition: ***tol***.
Output: Action-value function $Q(s, a)$, joint optimal policy π^*.
Initialization: Q, Q_j, SoC of the EVs.
Generate $\{\mathrm{A}_j\}$ by using Ψ_J.
while TURE **do**
  Initialize state space S and for each agent j :
  for $h = 1, 2, \ldots, H$ **do**
  $P_{FS} \leftarrow \mathrm{P}_{FS}^h$
  Multi-agent Q-learning algorithm for agent j :
    for $t = 1, 2, \ldots, M$ **do**
      Select an action value a_j^t randomly from A_j.
      Update Q_j by using equation (24).
      Update e^{t+1} according to **Algorithm 1**.
      $s^{t+1} \leftarrow s^t$
    end for
    Select optimal action by using equation (25) and get the reward
    R by equation (23).
    $Q^{h+1} \leftarrow Q^h$
  end for
  if $|Q^{h+1} - Q^h| \leq$ ***tol*** **then**
    break
  end if
end

In the algorithm execution (shown in **Algorithm 2**), each agent has a common goal and reward function. Because the delivery energy in each h is constant and cooperation among agents occurs at the end of each supply period, the agent's communication overhead is reduced, and cooperation occurs based on the exploration of the environment. For a certain state space, the computation complexity is reduced from $O(N_{\max} \times \tilde{N}_t \times J)$ to $O(N_{\max} \times \tilde{N}_t)$ with J parallel computing.

VI. SIMULATION AND ANALYSIS

A. MC-BASED FORECASTING SCHEME

We randomly choose one of the 366 created E-matrices to evaluate the completion effect of one-day energy forecasting, and the matrix size is 69×24. According to the formula $|\Omega| \geq c\mathcal{S}^{1.2} r \log \mathcal{S}$, the lower bound of the observation is 818 (where $c = 0.4$, $r = 3$), and the four to-be-completed matrices have proportions of observations of 50% (828), 60%, 70% and 80%. Additionally, the non-sampling points are set to NaN. Furthermore, two common criteria, relative reconstruction error (RRE) and total reconstruction error (TRE), are adopted to evaluate the complete accuracy:

$$\boldsymbol{RRE} = \frac{\left|\hat{E}^{mat} - E^{mat}\right|}{E^{mat}} \tag{29}$$

$$\boldsymbol{TRE} = \frac{\left\|\hat{E}^{mat} - E^{mat}\right\|_F}{\|E^{mat}\|_F} \tag{30}$$

where $\|\cdot\|_F$ is the Frobenius norm, $\hat{E}^{mat}$ and E^{mat} denote the completed and actual matrices, respectively. The reconstruction results are shown in Fig. 7.

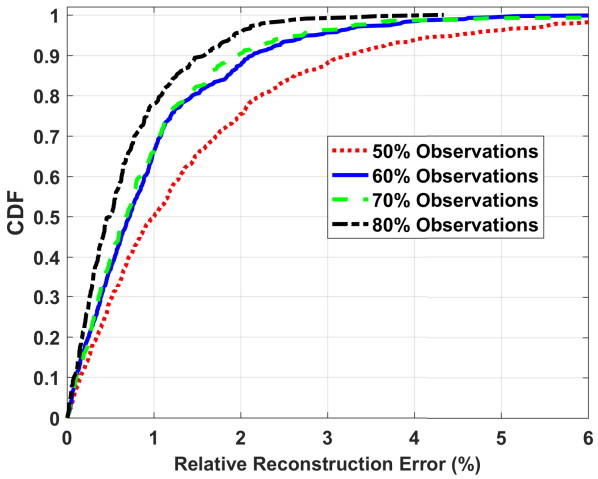


FIGURE 7. The CDF of the RRE for different observations.

Fig. 7 illustrates the cumulative distribution function (CDF) of the RRE for the completed E-matrix. The reconstruction algorithm selected in this paper is SDP-based convex optimization, and a performance comparison of several typical algorithms will be conducted later. In Fig. 7, when the observation proportion is 50%, the reconstructed values with the RREs less than 0.02 account for 78% of all values. Increasing the proportion to 80%, the proportion of reconstructed elements with the RREs less than 0.01 is over 80%.

To highlight the superiority of our proposed, we introduce two common metrics for load or power forecasting, the mean absolute percentage error (MAPE) and root mean square error (RMSE), to compare the order of magnitude of the error with [24] and [25]. The results are shown in Table 1.

TABLE 1. Metrics comparison.

Method		MAPE(%)	RMSE
Ours	50%	1.69	2.11
	60%	1.25	1.71
	75%	1.10	1.45
	80%	0.87	1.18
Load Fore. in **[24]**		1.30	2.87
Power Fore. in **[25]**		4.7	6.53

From Table 1, at greater than 60% of the observations, the errors of our method are less than those of the state-of-the-art forecasting approaches based on the order of magnitude. This result indirectly reflects the feasibility and accuracy of the proposed method.

To compare the effects of different algorithms on E-matrix reconstruction, we perform the reconstruction of 366 E-matrices. We choose four typical completion algorithms: the SDP-based algorithm, the ADMiRA, the SVP and the SVT. The evaluation criterion adopts the TRE, and the results are shown in Fig. 8.

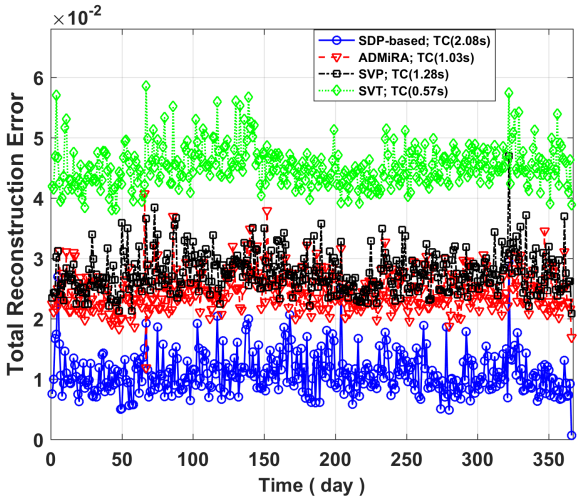


FIGURE 8. The TRE of the completion algorithms, "TC" is time-consuming.

In Fig. 8, the SDP-based algorithm exhibits optimal performance, and SVT is the worst because the E-matrix does not have an absolute low-rank property due to noise. ADMiRA's time-consuming is the lowest, and the accuracy is comparable to that of SVP, but it fluctuates more. Considering that our reconstruction is offline and the E-matrix is small in size, we select the SDP-based method, which utilizes excessive iterations in exchange for accuracy.

From the E-fog point of view, it requires the total output energy from each E-microgrid in its jurisdiction at the same time, which is the sum of the columns of the E-matrix and is denoted in equation (6). Fig. 9 illustrates the comparison between the sum of each column of 31 reconstructed and actual matrices. In Fig. 9, the actual and reconstructed curves are highly consistent because superposition can counteract the partial effects of positive and negative errors.

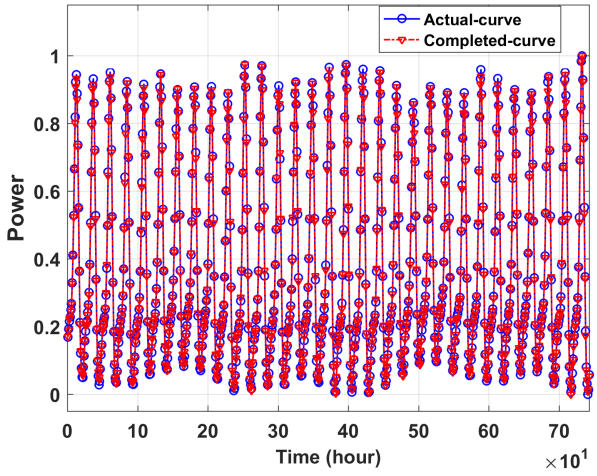


FIGURE 9. Comparison of the actual and completed aggregation values.

The above numerical results suggest that our proposed method, which reduces the workload, is effective and accurate, and they validate the effective implementation of the novel approach.

B. RL-BASED DELIVERY CONTROL APPROACH

In this section, we combine the current charging technology and the parameters of the EV in the market and set the charging scenario as follows: E-fog has several DC fast-charging stations, and the total number of charging piles is $N_{\max}$, the parameters of battery are summarized in Table 2.

TABLE 2. Battery parameters.

Nominal capacity	$B_{\max}$	$e_{\max}$	$p_{\min}$	$p_{\max}$
80kWh	72kWh	43.2kWh	60kW	120kW

Then, we create the following settings for the major related parameters in the charging stage. According to [42], the starting time of charging behavior has a Gaussian distribution, and more than 90% of charging events occur between 13:00 and 23:00. The charging peak appears at approximately 18:00 during a day. In addition, combining other information, we develop the energy delivery curve, as shown in Fig. 10 (a); Combining our charging scenario and the mobility analysis of the EVs in [46], we design the arrival number distribution of the EVs over time, shown in Fig. 10 (b), which is used to assist in calculating changes in the state space; Fig. 10 (c) illustrates the probability density function (PDF) of the initial battery SoC:

$$\mathbf{F}_{SoC}(s,\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\{-\frac{(x-\mu)}{2\sigma^2}\} \qquad (31)$$

Equation (31) takes μ for 0.4 ($0.4B_{\max}$), and a standard deviation of $\sigma = 0.1$ denotes that the initial SoC is mainly concentrated within $[0.3B_{\max}, 0.5B_{\max}]$.

Following the above scenario, we implement the proposed RL-based charging control approach, and the effectiveness of

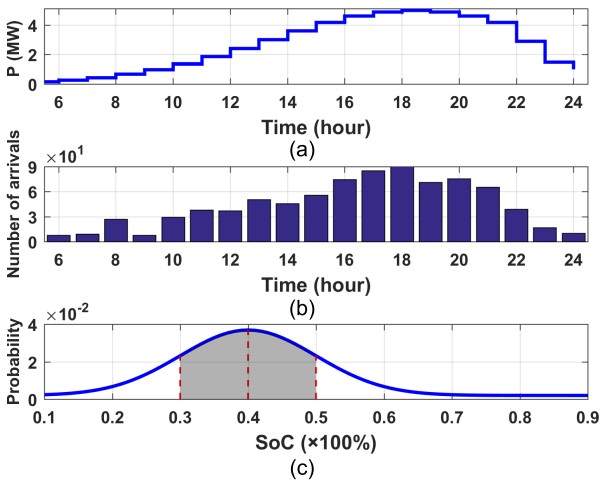


FIGURE 10. Related parameters in the charging stage($N_{max} = 50$, $\lambda_{arr} = 10$). (a) Energy delivery curve. (b) Distribution of arrival number of the EVs. (c) PDF of initial battery SoC of the EVs.

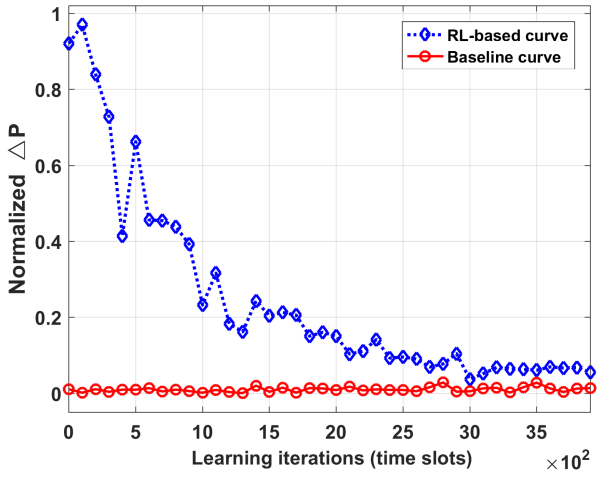


FIGURE 11. Convergence and results of the RL-based approach.

the algorithm is evaluated. The parameters are set as follows: $N_{\max} = 50$, the number of energy supply periods is set to 16 and slot $t = 1/6$ hour; discount factor $\gamma = 0.9$ and the learning rate $\alpha = 0.2$, the FCF $\Psi = [0.8 : 0.1 : 1.2]$ and the convergence rate control factor $\xi = 0.8$. Moreover, we set $\beta_{sur}^{-} = \beta_{sht}^{-} = 1$ and take advantage of the result of (18) as the evaluation baseline. The result is shown in Fig. 11. Moreover, to make our charging scenario more conform to the real-world, we regularize the action a_j when implementing the control algorithm, as follows.

$$\bar{a}_j = 10\mathbf{Fix}(0.1a_j) + 5\mathbf{Round}[0.2\mathbf{Mod}(a_j, 10)] \quad (32)$$

where $\mathbf{Fix}(\cdot)$ indicates that the theoretical value a_j is approximated as an integer towards zero, $\mathbf{Round}(\cdot)$ and $\mathbf{Mod}(\cdot)$ represent rounding and modulo operation, respectively. With equation (32), the actual charging power is normalized to an integer.

The curves in Fig. 11 illustrate the performance of the proposed EV charging control approach. Because the price is set to 1, the CPD $\Delta\mathbf{P}$ is used to represent the cost as a criterion for evaluation. Furthermore, from the perspective of evaluating the algorithm effectiveness, we normalized $\Delta\mathbf{P}$ based on the baseline value [31].

From Fig. 11, the CPD of the RL-based optimization approach decreases with increasing number of iterations, and the value of the CPD tends to stabilize after approximately 3200 iterations in our reference scenario. Meanwhile, the cost of our approach is only 8% more than the offline optimal result in the completion of 3200 explorations. Under the premise of the number of vehicles controlled is threefold more than [17], the equivalent result is achieved in terms of performance and time costs. Specifically, the designed multiagent-based collaborative exploration approach can utilize more exploitation and known dynamic information to further adjust the CDP and capture the optimal action more rapidly in the distributed algorithm.

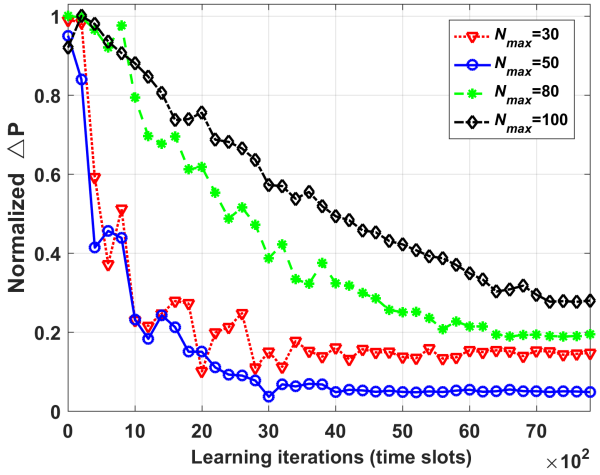


FIGURE 12. Implementations with different N_{max} values.

The state and action space respond to variations in $N_{\max}$, and Fig. 12 illustrates the effect of $N_{\max}$ on algorithm performance. Firstly, as $N_{\max}$ increases, the optimal result that can be initially achieved, but the result gradually worsens, and mainly because the redundant action space causes the cumulative error to increase. Secondly, the overall convergence rate of the algorithm slows and stabilizes after approximately 4000 iterations at $N_{\max} = 50$ and 7000 iterations at $N_{\max} = 100$ due to the enlargement of the space. Finally, as the errors are shared by more actions, the fluctuation is smallest at $N_{\max} = 100$. Moreover, because too few actions lead to insufficient exploration, compared to $N_{\max} = 50$, $N_{\max} = 30$ exhibits strong volatility and poor convergence.

Fig. 13 illustrates the impact of different slots on the algorithm performance in the same reference scenario and for $N_{\max} = 50$. For other conditions being equal, the different slots have little impact on the convergence and effectiveness of the algorithm in theory. However, at real-world charging stations, instead of waiting for the slot to end before unplugging, the EVs usually depart as soon as fully charged, and a

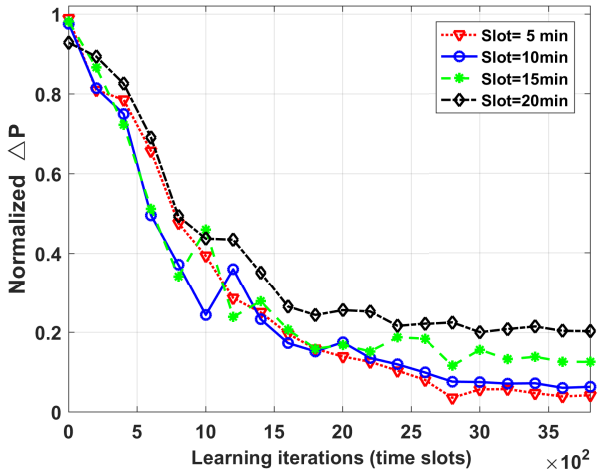


FIGURE 13. Implementations with different slots.

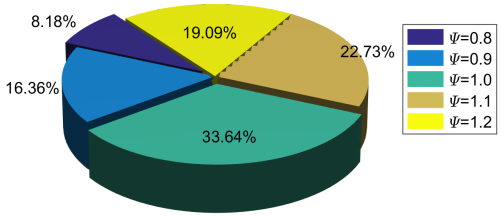


FIGURE 14. Utilization frequency of different FCFs in a completed learning.

larger slot value can lead to our approach to deviate from the real scenario. Accordingly, this paper tends to set the slot to 10 minutes.

Fig. 14 illustrates the number of times different factors are judged as optimal values, which represents their contribution to the final result under the reference scenario, where slot $t = 10$ minutes and $N_{\max} = 50$. The factor of 1.0 has the highest frequency of participation, and 0.8 has the lowest, which indicates that the charging power plan has a high probability of being adopted. Moreover, the factors greater than 1.0 also have a high rate of being selected because the initial SoC of the EV battery is usually small, and selecting large charging powers can match the power forecasting curve faster.

VII. CONCLUSION

The Energy Internet and related technologies are an exciting research filed, with important theoretical significance and practical value. In this paper, based on the architecture of our designed for the EI, we take the energy forecasting and utilization as the cut-in points and carry out the theoretical studies and the case simulations.

To effectively manage the unstable energy flow and the energy big data for the grid connection scenario of large-scale RES, we design an architecture of hierarchical integration for EI and divide the EI into three layers, by which to resolve the problems resulting from unified and centralized management.Based on the proposed EI, an efficient forecasting scheme is proposed, where the estimation of the complete E-matrix can be achieved from a small number of observations by leveraging the power of matrix completion. Furthermore, taking the EVs as the load, we design a model-free charging control approach with the objective of energy optimal delivery, and the reinforcement learning is introduced for the precise delivery of energy during the charging phase. To validate the merits of our proposed scheme and approach, a large number of simulation experiments were carried out by leveraging the real-world traces from PJM. The numerical results indicate that 1) the proposed MC-based forecasting scheme can reliably estimate the E-matrix from partial observations and 2) the RL-based charging approach yields the optimal energy delivery and a significant cost cut without the precise mathematical model and redundant input parameters.

REFERENCES

[1] L. H. Tsoukalas and R. Gao, "From smart grids to an energy Internet: Assumptions, architectures and requirements," in *Proc. 3rd Int. Conf. Electr. Utility Deregulation Restruct. Power Technol.*, Apr. 2008, pp. 94–98.

[2] K. Wang, H. Li, Y. Feng, and G. Tian, "Big data analytics for system stability evaluation strategy in the energy Internet," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1969–1978, Aug. 2017.

[3] S. Aman, Y. Simmhan, and V. K. Prasanna, "Energy management systems: State of the art and emerging trends," *IEEE Commun. Mag.*, vol. 51, no. 1, pp. 114–119, Jan. 2013.

[4] K. Wang *et al.*, "A survey on energy Internet: Architecture, approach, and emerging technologies," *IEEE Syst. J.*, to be published, doi: 10.1109/JSYST.2016.2639820.

[5] K. Zhou, S. Yang, and Z. Shao, "Energy Internet: The business perspective," *Appl. Energy*, vol. 178, pp. 212–222, Sep. 2016.

[6] A. Etxeberria, I. Vechiu, H. Camblong, J. M. Vinassa, and H. Camblong, "Hybrid energy storage systems for renewable energy sources integration in microgrids: A review," in *Proc. IPEC*, Oct. 2010, pp. 532–537.

[7] H. Jiang, K. Wang, Y. Wang, M. Gao, and Y. Zhang, "Energy big data: A survey," *IEEE Access*, vol. 4, pp. 3844–3861, Aug. 2016.

[8] L. Yao *et al.*, "Challenges and progresses of energy storage technology and its application in power systems," *J. Mod. Power Syst. Clean Energy*, vol. 4, no. 4, pp. 519–528, 2013.

[9] P. G. Lopez *et al.*, "Edge-centric computing: Vision and challenges," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 5, pp. 37–42, 2015.

[10] K. Wang, Y. Wang, Y. Sun, S. Guo, and J. Wu, "Green industrial Internet of Things architecture: An energy-efficient perspective," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 48–54, Dec. 2016.

[11] J. Tastu, P. Pinson, E. Kotwa, H. Madsen, and H. A. Nielsen, "Spatio-temporal analysis and modeling of short-term wind power forecast errors," *Wind Energy*, vol. 14, no. 1, pp. 43–60, 2010.

[12] A. Tascikaraoglu *et al.*, "A short-term spatio-temporal approach for photovoltaic power forecasting," in *Proc. 15th IEEE PSCC*, Jun. 2016, pp. 1–7.

[13] K. Wang, C. Xu, Y. Zhang, S. Guo, and A. Zomaya, "Robust big data analytics for electricity price forecasting in the smart grid," *IEEE Trans. Big Data*, to be published, doi: 10.1109/TBDATA.2017.2723563.

[14] PJM Interconnection LLC. *Public Data: Data Miner 2*. Accessed: Feb. 2017. [Online]. Available: http://dataminer2.pjm.com/list

[15] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.

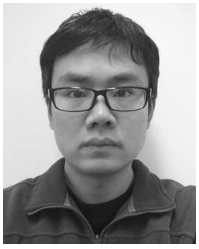
[16] K. Wang *et al.*, "Distributed energy management for vehicle-to-grid networks," *IEEE Netw.*, vol. 31, no. 2, pp. 22–28, Mar./Apr. 2017.

[17] S. Vandael, B. Claessens, D. Ernst, T. Holvoet, and G. Deconinck, "Reinforcement learning of heuristic EV fleet charging in a day-ahead electricity market," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1795–1805, Jul. 2015.

[18] R. S. Sutton and A. G. Barto, *Reinforcement Learning—An Introduction*. Cambridge, MA, USA: MIT Press, 1998.

[19] A. Q. Huang, M. L. Crow, G. T. Heydt, J. P. Zheng, and S. J. Dale, "The future renewable electric energy delivery and management (FREEDM) system: The energy Internet," *Proc. IEEE*, vol. 99, no. 1, pp. 133–148, Jan. 2011.

[20] Y. Xu, J. Zhang, W. Wang, A. Juneja, and S. Bhattacharya, "Energy router: Architectures and functionalities toward energy Internet," in *Proc. IEEE Int. Conf. Smart Grid Commun.*, Oct. 2011, pp. 31–36.
[21] N. Bui, A. P. Castellani, P. Casari, and M. Zorzi, "The Internet of energy: A Web-enabled smart grid system," *IEEE Netw.*, vol. 26, no. 4, pp. 39–45, Jul. 2012.
[22] K. Wang, M. Du, S. Maharjan, and Y. Sun, "Strategic honeypot game model for distributed denial of service attacks in the smart grid," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2474–2482, Feb. 2017.
[23] K. Wang, X. Hu, H. Li, P. Li, D. Zeng, and S. Guo, "A survey on energy Internet communications for sustainability," *IEEE Trans. Sustain. Comput.*, vol. 2, no. 3, pp. 231–254, May 2017.
[24] G. M. U. Din and A. K. Marnerides, "Short term power load forecasting using deep neural networks," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Jan. 2017, pp. 594–598.
[25] A. Fentis, L. Bahatti, M. Mestari, and B. Chouri, "Short-term solar power forecasting using support vector regression and feed-forward NN," in *Proc. 15th IEEE NEWCAS*, Jun. 2017, pp. 405–408.
[26] C.-Y. Zhang, C. L. P. Chen, M. Gan, and L. Chen, "Predictive deep Boltzmann machine for multiperiod wind speed forecasting," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1416–1425, Oct. 2015.
[27] K. Wang *et al.*, "Wireless big data computing in smart grid," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 58–64, Apr. 2017.
[28] A. Tascikaraoglu and B. M. Sanandaji, "Short-term residential electric load forecasting: A compressive spatio-temporal approach," *Energy Buildings*, vol. 111, pp. 380–392, Jan. 2016.
[29] E. M. Carreno, R. M. Rocha, and A. Padilha-Feltrin, "A cellular automaton approach to spatial electric load forecasting," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 532–540, May 2011.
[30] M. Gjelaj, C. Træholt, S. Hashemi, and P. B. Andersen, "Optimal design of DC fast-charging stations for EVs in low voltage grids," in *Proc. IEEE ITEC*, Jun. 2017, pp. 684–689.
[31] Z. Fan, "A distributed demand response algorithm and its application to PHEV charging in smart grids," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1280–1290, Sep. 2012.
[32] F. Rücker, I. Bremer, S. Linden, J. Badeda, and D. U. Sauer, "Development and evaluation of a battery lifetime extending charging algorithm for an electric vehicle fleet," *Energy Proc.*, vol. 99, pp. 285–291, Nov. 2016.
[33] W. Shi and V. W. S. Wong, "Real-time vehicle-to-grid control algorithm under price uncertainty," in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Oct. 2011, pp. 261–266.
[34] M. Lauer and M. A. Riedmiller, "An algorithm for distributed reinforcement learning in cooperative multi-agent systems," in *Proc. 17th Int. Conf. Mach. Learn.*, 2000, pp. 535–542.
[35] H.-S. Kim, M.-H. Ryu, J.-W. Baek, and J.-H. Jung, "High-efficiency isolated bidirectional AC–DC converter for a DC distribution system," *IEEE Trans. Power Electron.*, vol. 28, no. 4, pp. 1642–1654, Apr. 2013.
[36] K. Wang, Z. Ouyang, R. Krishnan, L. Shu, and L. He, "A game theory-based energy management system using price elasticity for smart grids," *IEEE Trans. Ind. Inform.*, vol. 11, no. 6, pp. 1607–1616, Dec. 2015.
[37] PJM Interconnection LLC. *Public Data: Maps*. Accessed: Apr. 2017. [Online]. Available: http://www.pjm.com/library/maps.aspx
[38] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Rev.*, vol. 52, no. 3, pp. 471–501, 2010.
[39] J. F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.
[40] S. Burer and R. D. C. Monteiro, "A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization," *Math. Program.*, vol. 95, no. 2, pp. 329–357, 2003.
[41] K.-C. Toh and S. Yun, "An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems," *Pacific J. Optim.*, vol. 6, no. 3, pp. 615–640, Nov. 2010.
[42] K. Qian, C. Zhou, M. Allan, and Y. Yuan, "Modeling of load demand due to EV battery charging in distribution systems," *IEEE Trans. Power Syst.*, vol. 26, no. 2, pp. 802–810, May 2011.
[43] A. M. Abdelbar, S. H. Gheita, and H. A. Amer, "Exploring the fitness landscape and the run–time behaviour of an iterated local search algorithm for cost–based abduction," *J. Exp. Theor. Artif. Intell.*, vol. 18, no. 3, pp. 365–386, 2006.
[44] M. Grant and S. Boyd. *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.1*. Accessed: Jan. 2016. [Online]. Available: http://cvxr.com/cvx
[45] J. Yang, S. Zhang, X. Wu, Y. Ran, and H. Xi, "Online learning-based server provisioning for electricity cost reduction in data center," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 3, pp. 1044–1051, May 2017.
[46] NREL. (2017). *Fleet DNA Project Data*. Accessed: Jan. 2017. [Online]. Available: www.nrel.gov/fleetdna

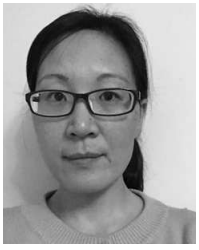


LIUFENG DU received the B.S. degree in electronic information engineering from Zhengzhou University in 2003 and the M.S. degree in signal and information processing from Tianjin University in 2011. He is currently pursuing the Ph.D. degree in signal and information processing with the Nanjing University of Posts and Telecommunications, Nanjing, China.

He is currently a Lecturer with the Henan Institute of Science and Technology, Xinxiang, China. His research interests include energy management technology and signal processing with applications in the Internet of Things.

LINGHUA ZHANG received the M.S. degree in signal and information processing from Southeast University in 1990 and the Ph.D. degree in signal and information processing from the Nanjing University of Posts and Telecommunications,
Nanjing, China, in 2005.

She is currently a Professor and a Ph.D. supervisor with the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, China. Her research interests lie in the area of voice communication, Internet of Things, and signal processing in wireless communication.



XIYAN TIAN received the B.S. degree in applied electronic technology from Henan Normal University in 2002 and the M.S. degree in communication and information system from Zhengzhou University in 2012.

She is currently a Lecturer with the Henan Institute of Science and Technology, Xinxiang, China. Her research interests include sparse representation theory and wireless sensor networks.

JINHUI LEI received the B.S. degree in electronic information engineering and the M.S. degree in communication and information system from Zhengzhou University, in 2003 and 2011, respectively.

He is currently a Lecturer with the Henan Institute of Science and Technology, Xinxiang, China. His research interests include Internet of Things, embedded systems, and wireless sensor networks.