计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.220400002

# Clustering-based Demand Response for Intelligent Energy Management in 6G-enabled Smart Grids

**Ran WANG[1,2], Jiang-tian NIE[3], Yang ZHANG[1,2] and Kun ZHU[1,2]**

1 College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

2 Collaborative Innovation Center of Novel Software Technology and Industrialization, Nanjing 211106, China

3 School of Computer Science and Engineering, Nanyang Technological University, 639798, Singapore

(wangran@nuaa.edu.cn)

**Abstract** As a typical industrial Internet of things (IIOT) service, demand response (DR) is becoming a promising enabler for intelligent energy management in 6G-enabled smart grid systems, to achieve quick response for supply-demand mismatches. However, existing literatures try to adjust customers' load profiles optimally, instead of electricity overhead, energy consumption patterns of residential appliances, customer satisfaction levels, and energy consumption habits. In this paper, a novel DR method is investigated by mixing the aforementioned factors, where the residential customer cluster is proposed to enhance the performance. Clustering approaches are leveraged to study the electricity consumption habits of various customers by extracting their features and characteristics from historical data. Based on the extracted information, the residential appliances can be scheduled effectively and flexibly. Moreover, we propose and study an efficient optimization framework to obtain the optimal scheduling solution by using clustering and deep learning methods. Extensive simulation experiments are conducted with real-world traces. Numerical results show that the proposed DR method and optimization framework outperform other baseline schemes in terms of the system overhead and peak-to-average ratio (PAR). The impact of various factors on the system utility is further analyzed, which provides useful insights on improving the efficiency of the DR strategy. With the achievement of efficient and intelligent energy management, the proposed method also promotes the realization of China's carbon peaking and carbon neutrality goals.

**Keywords** Demand response (DR), Customer clustering, Deep learning, 6G-enabled industrial Internet of things (IIOT), Smart srid (SG)

**Chinese Library Classification** TP393

## 1 INTRODUCTION

The 6G network will become a basis of intrinsic intelligence, ubiquitous connectivity and multidimensional integration, and an important foundation for various future industrial vertical application developments. 6G network will be embedded with artificial intelligence (AI) capabilities to achieve architecture-level intelligence. Internally, intelligence can be used to optimize network performance, enhance user experience, and automate network operation, laying a solid foundation for the cross-border convergence of different industrial networks. Externally, it can extract and encapsulate network intelligence, and provide communication and computing services combining network and AI for industrial Internet of things (IIOT), for instance, the future smart grid (SG).

As an important 6G-enabled IIOT scenario, SG is a promising solution for conventional power networks by leveraging information and communication technologies, where power supply status can be detected and collected with digital or analog signals when electricity is supplied, and power usage information can be properly analyzed when electricity is used. Demand response (DR) is becoming a novel enabler in SG by exploring its potential benefits to achieve quick response for supply-demand mismatches[1].

Nowadays, different types of DR programs have been implemented to realize efficient energy consumption management for residential customers and achieve carbon peaking and carbon neutrality goals[2-3]. Generally speaking, there are two main types of DR programs: control-based and motivation-based DR programs. However, these two types of program are not completely mutually exclusive. We will study the state of art concerning the DR programs in 6G enabled SG, and then analyze our research motivations in the following subsections.

We organize the rest of this paper as follows. In Section 2, we review the state of the art. Section 3 studies the system operation details and proposes an optimization problem for the

Submission date: 2022-04-01 Modification date: 2022-05-01

This work was supported by the National Natural Science Foundation of China (62171218).

Corresponding author: Jiang-tian NIE (jnie001@e.ntu.edu.sg)

DR problem. The DR scheme based on the cluster of residential customers is introduced in Section 4. In Section 5, the optimization framework is presented. Section 6 presents the baseline methods in simulation. The simulation experiments are included and discussed in Section 7. Finally, this paper is concluded.

## 2 RELATED WORK

### 2.1 Features of DR Program

In the control-based DR programs, centralized and distributed schemes are defined according to where the decisions are made. For the former, utility companies with centralized infrastructure make the decisions according to the information collected from all appliances and provide optimal energy consumption solutions for customers by considering limited grid technical constraint[4]. Nevertheless, the complexity of the centralized scheme increases significantly with the increase in the number of residential appliances due to centralized decision-making. It is extremely difficult to collect information accurately from massive appliances. More flexibility can be achieved by the distributed scheme due to scalability and customer involvement. Efficient coordination is realized among different customers according to the load demands and received price signals. Although the decentralized one can provide ownership authority to customers, the optimal operation of appliances cannot be guaranteed and new security issues of power grids arise.

Unlike the aforementioned DR programs, motivational DR programs can be mainly divided into the following two types.

With China is already committed to peak carbon dioxide emissions before 2030 and achieve carbon neutrality before 2060, the Central Economic Work Conference urges quicker steps to come up with an action plan that enables the peaking of emissions.

Incentive-based and price-based DR programs. In the former one, customers are encouraged to decrease their energy consumption upon the request offers or based on their contractual agreements reached by both customers and companies[6-7]. Under such an incentive scheme, some customers will be rewarded due to actions taken by the program administrator, e. g. appliance rescheduling, reduction, and disconnection. However, it may bring new challenges and issues regarding customers' privacy and security when accessing appliances directly. In the latter one, time-varying electricity prices are offered to customers[8]. The electricity price is increased when peak comes and reduced during less congested periods. With the price signal, the peak-to-valley gap of the power load can be reduced, and the energy consumption behaviors of customers can be changed to achieve peak shaving and valley filling. The examples of the price-based scheme mainly include the time of use price(ToU), critical-peak price(CPP), and real-time price(RTP)[9]. With the usage of the price-based DR program, customer privacy and security can be enhanced dramatically. However, the customers should respond to the varying price signal quickly. At the same time, the unfairness among consumers may incur due to the different pricing mechanisms. To summarize, the performance metrics, e. g. overall energy overhead and peak-to-average ratios (PARs), can be improved significantly by leveraging appropriate DR schemes.

### 2.2 Optimization Methods of DR

To optimize the cost minimization-based DR problems, fuzzy and programming methods as well as game theory have been widely utilized[10-12]. In addition, machine learning-based methods, such as clustering, have been used to further improve the performance. Kwac et al.[13] propose a stable coding mechanism that uses an adaptive $K$-means algorithm to find the representative load shape from the daily power data of consumers and uses hierarchical clustering to summarize, and calculates various indexes according to the coding mechanism. Kang et al.[14] propose a $k$-sliding distance between two electricity customers for customer clustering. In literature [15], the density-based spatial clustering of applications with noise is applied to explore end-customers' inherent electricity consumption patterns from historical load data. With the complementarity of different consumers, utility companies can develop more appropriate DR schemes for different types of consumers[16]. Haben et al.[17] present a comprehensive study of customers' smart meter data for analyzing peak demand and main factors of variability in their behavior. A finite mixture model-based clustering method is also presented for discovering different behavior groups based on demand and variability. Zhou et al.[18] develop an improved fuzzy clustering model for the mining of household monthly electricity consumption patterns. Melicio et al.[19] employs $K$-means clustering algorithm to predict daily load profile of typical electricity usage.

With the development of computer technology, deep learning plays an increasingly important role in SG. Shi et al.[20] propose a novel pooling-based deep recurrent neural network which batches a group of customers' load profiles into a pool of inputs to directly learn the uncertainty of load profiles. Sun et al.[21] propose a novel probabilistic baseline estimation framework, which employs a deep learning-based clustering method to process a large quantity of daily patterns for improving estimation performance. Fan et al.[22] exploit and compare the performance metrics of cooling load prediction with deep learning in building field with two manners, which are methods with typical feature extraction as well as popular prediction. Cai et al.[23] propose two classical deep neural network models in both recursive and direct multi-step manners to forecast building loads. Kong et al.[24] propose a long short-term memory (LSTM) recurrent neural network-based framework to forecast the electric load of a single energy customer. Kim et al.[25] study a CNN-LSTM neural network that extracts both spatial and temporal features for developing an effective housing energy consumption prediction.

### 2.3 Main Contributions

The above-mentioned works mainly studied the optimization problems for load profiles of customers by smoothing the demand profile and decreasing the PARs. None of them considers the satisfaction degree, fairness and energy consumption preferences of customers, which are very important for the interactions between utility companies and appliances since customers are the main roles in the system. For example, the electricity prices are highlighted by some customers and others focus more on the comfort of power consumption. Thus, a novel DR scheme is proposed in this work based on the cluster of residential customers, which is achieved by the energy consumption habits and preferences. The main contributions can be summarized as follows:

(1) An optimization problem for the proposed DR scheme is formulated by considering the energy consumption patterns, electricity prices, customer satisfaction, as well as fairness. To the best of our knowledge, the aforementioned impact factors are not comprehensively investigated in a DR scheme in existing literatures.

(2) Clustering methods are leveraged to investigate the electricity consumption features of customers. Based on the clustering results, residential appliances are scheduled more flexibly and effectively.

(3) An optimization framework based on clustering and deep learning is proposed to solve the problem of huge computation in large-scale customer scenarios.

(4) Extensive simulation experiments are conducted with real-world traces. The numerical results show that our proposed DR method and optimization framework outperform other baseline schemes in terms of the system overhead and PAR of the electricity grid.

## 3 DR PROBLEM FORMULATION

As shown in Fig. 1, an SG system is composed of a power supplier, i. e., a utility company, and also $N$ residential customers. We denote $\mathcal{U}=\{1,\cdots,N\}$ as the set of customers. The 6G communication network is assumed to connect the power supplier and numerous residential customers as power users. It is assumed that the residential customers have signed relevant authorization documents with the utility company. Then, the power data can be downloaded from the utility company to the residential customers so that the customer energy consumption can be monitored and the remote control can be achieved to realize the required demand response strategy. A time slot-based system is considered, where there are $T$ time slots with equal length. The length of a time slot can be set to 15 min. $\mathcal{T}=\{1,\cdots,T\}$ is denoted as the index set of time slots. Since each residential customer has several appliances, we assume that the appliance set is denoted by $\mathcal{A}_u=\{1,\cdots,A_u\}$ in household $u\in\mathcal{U}$. The total number of appliances is set to $A_u$.

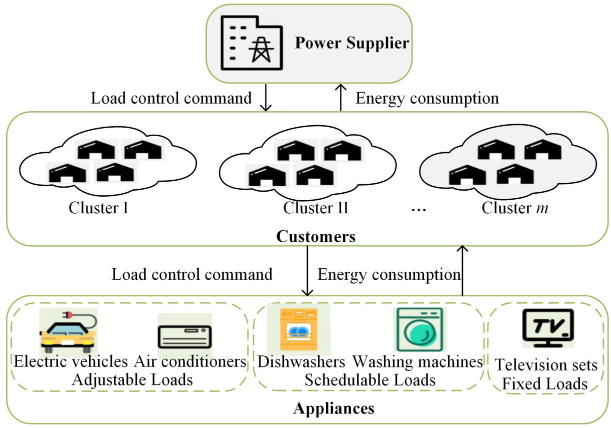



Fig. 1 Illustration of the proposed system

### 3.1 Model of Residential Appliances

According to the appliance features, the loads can be categorized into three different types, including adjustable, schedulable and fixed loads.

(1) Adjustable Loads: Loads can be considered as adjustable loads given the loads are able to be adjusted in each time slot. For instance, electric vehicles and air conditioners can control their energy consumption on a time slot basis. The set of adjustable loads used by customer $u$ is denoted by $\mathbb{C}_u$. By dynamically altering the charging rates of electric vehicles as well as the current temperatures of air conditioners, the control of appliance operations can be realized. In time slot $t\in\mathcal{T}$, a variable $I_{u,a,t}\in\{0,1\}$ is introduced to depict the the state of customer $u$'s appliance $a$ (0 for the state off and 1 for on). The active power consumption is defined as $E_{u,a,t}$ when appliance $a$ is switched on in time slot $t$. The consumed energy of customer $u$'s appliance $a$ during time slot $t$ is derived as:

$$E^{\mathrm{eca}}_{u,a,t}=I_{u,a,t}\cdot E_{u,a,t},\ \forall u\in\mathcal{U},\forall a\in\mathbb{C}_u,\forall t\in\mathcal{T} \tag{1}$$

$E^{\mathrm{eca}}_{u,a,t}$ denotes the baseline of energy consumption for customer $u$'s app-liance $a$. Let $\gamma^{\mathrm{upper}}_{u,a}$ and $\gamma^{\mathrm{down}}_{u,a}$ denote the elasticities of ramping up and ramping down for customer $u$'s appliance $a$. We have the following constraint for the potential appliance's scheduling strategy:

$$\gamma^{\mathrm{down}}_{u,a}\cdot E^{\mathrm{base}}_{u,a}\leqslant\sum_{t=1}^{T}E^{\mathrm{eca}}_{u,a,t}\leqslant\gamma^{\mathrm{upper}}_{u,a}\cdot E^{\mathrm{base}}_{u,a},\ \forall u\in\mathcal{U},\forall a\in\mathbb{C}_u \tag{2}$$

The total energy consumption amount by the adjustable appliances of customer $u$ can be achieved as:

$$E^{\mathrm{adj}}_{u,t}=\sum_{a\in\mathbb{C}_u}E^{\mathrm{eca}}_{u,a,t} \tag{3}$$

(2) Schedulable Loads: The operation time of particular

loads can be shifted backward or forward. Nevertheless, these loads should be completed without a break once the system runs. Assuming that $\mathbb{S}_u$ is the category of schedulable loads generated by customer $u$. Typical appliances with schedulable loads include pumps, dish washers, and washing/laundry machines. The dedicated energy consumption pattern is another feature of the appliances generating schedulable loads. For example, as illustrated in Fig. 2, the operation state of a washing machine can be divided into fine-grained stages, including soaking, washing, and drying. Energy consumption for each stage shows different patterns. The key procedure of scheduling for schedulable loads is the translation of energy consumption on the time axis. The DR scheme is expected to properly consider the energy consumption patterns of such appliances during the scheduling process. Let $E^{\text{pa}}_{u,a}=(e^{\text{pa}}_{u,a,1}, e^{\text{pa}}_{u,a,2},\cdots,e^{\text{pa}}_{u,a,T^{\min}_{u,a}})$ denote the energy consumption pattern vector of customer $u$'s app-liance $a$, where $T^{\min}_{u,a}$ is the minimum required number of time slots to finish the operation process for the schedulable appliance. The pattern for energy consumption of $u$'s appliance $a$ is formulated as:

$$f^{\text{pa}}_{u,a}(t)=\begin{cases} e^{\text{pa}}_{u,a,t} & 1\leqslant t\leqslant T^{\min}_{u,a} \\ 0 & \text{otherwise}\end{cases} \tag{4}$$

We assume that appliance $a$ of customer $u$ begins to operate in time slot $T^{\text{start}}_{u,a}$. Then, the energy consumption of appliance $a$ of customer $u$ in time slot $t$ can be calculated as:

$$E^{\text{ecs}}_{u,a,t}=f^{\text{pa}}_{u,a}(t-T^{\text{start}}_{u,a}+1)\cdot I_{u,a,t}, \forall u\in\mathcal{U}, \forall a\in\mathbb{S}_u, \forall t\in\mathcal{T} \tag{5}$$

Let $T^{\min}_{u,a}$ and $T^{\text{end}}_{u,a}\in\mathcal{T}$ denote the the beginning and ending of a time interval repetitively, in which appliance $a$ can be scheduled. An auxiliary binary variable is introduced such that $z_{u,a,t}\triangleq 1$ if the operation of customer $u$'s appliance $a$ starts in time slot $t$, where $a\in\mathbb{S}_u$. Otherwise, $z_{u,a,t}\triangleq 0$. Then, it holds that:

$$\sum_{t=T^{\text{ini}}_{u,a}}^{T^{\text{end}}_{u,a}-T^{\min}_{u,a}+1} z_{u,a,t}=1 \tag{6}$$

and

$$z_{u,a,t}=0, \forall t\in\mathcal{T}\backslash[T^{\text{ini}}_{u,a}, T^{\text{end}}_{u,a}-T^{\min}_{u,a}+1] \tag{7}$$

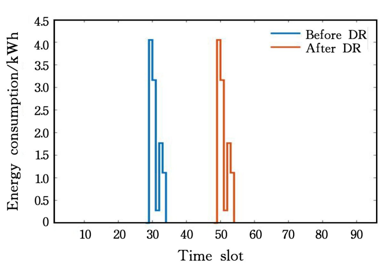



Fig. 2 Energy consumption pattern for a washing machine

Elements in both the state indicator matrix $I_{u\times a\times t}$ and the start time matrix $Z_{u\times a\times t}$ are thus correlated in the following manner:

$$I_{u,a,t}\geqslant z_{u,a,t}, I_{u,a,t+1}\geqslant z_{u,a,t},\cdots, I_{u,a,t+T^{\min}_{u,a}-1}\geqslant z_{u,a,t} \tag{8}$$

which indicates that $T^{\text{start}}_{u,a}=t$ and $I_{u,a,t}=I_{u,a,t+1}=\cdots=I_{u,a,t+T^{\min}_{u,a}-1}=1$ if $z_{u,a,t}=1$. Once a schedulable appliance begins to operate, it keeps operating continuously for a duration of $T^{\min}_{u,a}$ time slots. The total energy consumption of schedulable loads by customer $u$'s appliances is hereof derived as:

$$E^{\text{sch}}_{u,t}=\sum_{a\in S_u} E^{\text{ecs}}_{u,a,t} \tag{9}$$

(3) Fixed Loads: Some appliances are not eligible to be dynamically controlled during operation, which are categorized as appliances with fixed loads. Typical examples of such appliances include microwave ovens, electric stoves, and televisions. Total energy consumption of the appliances with fixed loads by customer $u$ at time slot $t$ is defined as:

$$E^{\text{fix}}_{u,t}=\sum_{a\in\mathbb{F}_u}\widetilde{E}_{u,a,t}, \forall u\in\mathcal{U}, \forall t\in\mathcal{T} \tag{10}$$

where $\mathbb{F}_u$ represents the fixed load category of customer, $\widetilde{E}_{u,a,t}$ is the historical energy consumption of customer $u$'s appliance $a$ at time slot $t$. For fixed loads, no DR strategies are applied since it will cause significant inconvenience to customers. For example, when customers are watching TV, if power companies send DR signals to turn off TVs, it will bring great inconvenience to customers' life.

The total energy consumption of customer $u$ at $t$ is:

$$E^{\text{tot}}_{u,t}=E^{\text{adj}}_{u,t}+E^{\text{sch}}_{u,t}+E^{\text{fix}}_{u,t} \tag{11}$$

### 3.2 Objective Functions

Power suppliers aim at improving satisfaction degree and fairness of their customers, as well as reducing the overall energy cost of the SG system when the system is in operation. Objective functions of the power suppliers are derived in the following.

(1) Degree of Dissatisfaction: To depict a satisfaction measurement of a customer, the degree of dissatisfaction can be defined as an increasing and convex function concerning the energy consumption deviation by leveraging the DR strategy[26]. The convexity indicates that the increase in the energy consumption deviation leads to higher dissatisfaction of customers. Without loss of generality, a quadratic dissatisfaction function is proposed in this work. In time slot $t$, the dissatisfaction degree of customer $u$ can be defined as:

$$f^{\text{uns}}_{u,t}(j_{u,t})=m^{\text{uns}}\cdot(j_{u,t})^2+n^{\text{uns}}\cdot j_{u,t} \tag{12}$$

where $m^{\text{uns}}\geqslant 0$ and $n^{\text{uns}}\geqslant 0$ indicate the coefficients, $j_{u,t}$ is the energy consumption deviation after customer $u$ by leveraging the DR strategy in time slot $t$. It holds that :

$$j_{u,t}=\| E^{\text{tot}}_{u,t}-\widetilde{E}^{\text{tot}}_{u,t}\| \tag{13}$$

where $\widetilde{E}^{\text{tot}}_{u,t}$ denotes the average historical energy consumption of customer $u$ in time slot $t$. It means that once the power utility dispatches the customers' appliances, the customers will be dissatisfied with the DR scheme if the energy consumption after DR is different from the historical energy consumption.

(2) Degree of Unfairness: To measure whether customers are fairly served or not, a generalized increasing and convex

function can be defined for measuring unfairness degree as follows:

$$f_{u,t}^{\text{unf}}(k_{u,t})=m^{\text{unf}}\cdot(k_{u,t})^2+n^{\text{unf}}\cdot k_{u,t} \tag{14}$$

where $m^{\text{unf}}\geqslant 0$ and $n^{\text{unf}}\geqslant 0$ denote the coefficients, $k_{u,t}$ indicates the energy consumption deviation by comparing the average energy consumption of customer $u$ with that of other customers:

$$k_{u,t}=\| E_{u,t}^{\text{tot}}-\frac{1}{n-1}\sum_{u\in\mathscr{U}_u^c\backslash u}E_{u,t}^{\text{tot}}\| \tag{15}$$

where $\mathscr{U}_u^c$ is the customer set that compares with $u$, and $n$ is the size of that set. In the comparison range, only when the energy consumption of each customer after DR is the same in any time slot, the degree of unfairness is 0, and the DR is absolutely fair.

(3) Energy Cost: An increasing and convex function in terms of the amount of consumed energy is defined as the energy cost, indicating the expenditure spent on electricity consumption. The convexity of the cost function indicates that the cost increases significantly when the amount of required energy increases. The quadratic cost function[27] and piecewise linear cost function[28] are widely used in existing works. Therefore, a quadratic cost function is utilized in this work. The energy cost of customer $u$ in time slot $t$ can be achieved as:

$$f_{u,t}^{\text{cost}}(E_{u,t}^{\text{tot}})=m^{\text{cost}}\cdot(E_{u,t}^{\text{tot}})^2+n^{\text{cost}}\cdot E_{u,t}^{\text{tot}} \tag{16}$$

where $m^{\text{cost}}\geqslant 0$ and $n^{\text{cost}}\geqslant 0$ denote coefficients defined by the power supplier. Actually, $m^{\text{cost}}$ is an arbitrarily small number.

## 4 CUSTOMER CLUSTER-BASED DR SCHEME

Customers are often featured with various preferences and characteristics. For instance, customers who focus more on the energy costs will actively be a part of the DR scheme to reduce their electricity bills significantly by taking more DR responsibilities. Instead of monetary overhead, the customers who are insensitive to the cost want to achieve convenient and comfortable appliance operations for better experiences. Driven by the aforementioned situations, we propose a customer cluster-based DR scheme by considering different customer preferences to provide flexible scheduling services.

$K$-means algorithm is one of the most popular and used clustering methods for unsupervised learning, where efficient multi-dimensional feature clusters can be obtained in SG[29]. With a given data set, $K$-means can find the globally optimal partitions efficiently. In this work, the customer features are the inputs of the $K$-means algorithm, which is listed in Table I. There are 10 features for every customer, where the first eight features are collected from customers and gathered by smart meters. Moreover, flexibility and interruption tolerance factors are extracted from historical data as two new features. In the following subsections, more details are included and discussed.

Table 1 Cluster features

| Index | Description | Index | Description |
|---|---|---|---|
| 1 | Building type | 6 | Total appliance number |
| 2 | Amount of PV | 7 | Total energy consumption of last month |
| 3 | House construction year | 8 | Power generated by PV in last month |
| 4 | Total square footage | 9 | Flexibility factor |
| 5 | Total room number | 10 | Interruption tolerance factor |

### 4.1 Flexibility Factor

The flexibility factor of appliance $a$ can be defined as the standard deviation of the electricity consumption distribution. Therefore, the flexibility factor of customer $u$ can be achieved as:

$$l(u)=\sum_{a\in A_u}\sigma_{u,a}w_a \tag{17}$$

where $\sigma_{u,a}$ denotes the standard deviation of appliance $a$'s energy consumption distribution extracted from the historical energy consumption data, and $w_a$ indicates the weight appliance $a$, which is calculated as:

$$w_a=\frac{power_a}{\sum_{a'\in A_u}power_{a'}} \tag{18}$$

where $power_a$ is the power of appliance $a$. A smaller flexibility factor means that customer may tend to take fewer DR responsibilities. On the contrary, the customer with a higher flexibility factor will take more DR responsibilities.

### 4.2 Interruption Tolerance Factor

The interruption tolerance factor $inter_{u,a}$ for appliance $a$ denotes the average interruption times per day. As aforementioned, customer's interruption tolerance can be obtained as:

$$i(u)=\sum_{a\in A_u}inter_{u,a} \tag{19}$$

With an interruption tolerance, the customer often wants to use appliances continuously. In contrast, a customer with a larger interruption tolerance factor will be more eager to accept more interruptions during the appliance operation.

### 4.3 Optimization Framework Formulation Based on Customer Clustering

Customers are divided into $C$ clusters by leveraging the $K$-means cluster algorithm. Customers in different clusters are different in the dissatisfaction, unfairness, and electricity cost. With the various parameter settings, for different clusters, the DR scheme should generate different strategies. In cluster $c$, the total expenditure function is defined as:

$$U_{c,t}^{\text{cluster}}=\sum_{u\in c}(\alpha_c\cdot f_{u,t}^{\text{uns}}(j_{u,t})+\beta_c\cdot f_{u,t}^{\text{unf}}(k_{u,t})+\gamma_c\cdot f_{u,t}^{cost}(E_{u,t}^{\text{tot}})) \tag{20}$$

where $\alpha_c$, $\beta_c$ and $\gamma_c$ are the customer coefficients in cluster $c$ for the dissatisfaction level, unfairness level, and electricity cost. It holds that $\alpha_c+\beta_c+\gamma_c=1$. Thus, the objective of the DR scheme is to minimize the total expenditures for customers during the system operation. The energy consumption scheduling problem can be formulated as the following optimization problem:

$$\min\sum_{c=1}^{C}\sum_{t=1}^{T}U_{c,t}^{\text{cluster}},\text{s.t. (1)}-\text{(20)} \tag{21}$$

This problem is a mixed-integer linear programming (MILP) problem. The calculation of the unfairness degree in the DR scheme involves the comparison with other customers' electricity consumption, so the optimal solution to the problem needs to be solved as a whole. When the number of customers increases, it will not be able to solve the optimal solution in a tolerable time. So we propose an optimization framework based on clustering and deep learning to solve it more efficiently in Section 4.

## 5 OPTIMIZATION FRAMEWORK BASED ON CLUSTERING AND DEEP LEARNING

Considering the computational complexity of Eq. 21, we propose an optimization framework based on clustering and deep learning to solve it more efficiently. Fig. 3 demonstrates that the whole optimization framework can be divided into two stages: the data preparation stage and the network training stage. Assuming that there are currently $T_0$ days of customers' electricity data, it is divided into $T_1$, $T_2$ and 1 day ($T_0 = T_1 + T_2 + 1$) for data preparation, network training, and method performance verification respectively.

### 5.1 Data Preparation Stage

In the data preparation stage, each large customer cluster generated in Section 3 is further divided into several small clusters of the same size by using the Same-Size $K$-means Variation (SS $K$-means)[30]. Each small cluster is regarded as a separate virtual optimization group. According to Eq. 21, the global optimal solution is obtained by using the MOSEK toolkit[31], and the optimal virtual energy consumption of customers in each small cluster after DR is obtained. Then the framework calculates and solves the DR scheme again in the large cluster. According to Eqs. 14 and 15, each customer needs to compare with other customers' energy consumption when ca-lculating the degree of unfairness. At this time, the comparison value uses the virtual energy consumption solved for each customer in the small cluster in the above step. In this way, the mixed solution of multiple customers can be divided into unique solutions of single customers. The framework solves the problem vertically, $T_1$ days in turn. After the completion of the data preparation stage, the energy consumption of each customer after DR with a length of $T_1$ days is obtained. It then enters the network training stage when sufficient pre-training data is collected.

### 5.2 Network Training Stage

In the network training stage, the first step is to train an LSTM network for each customer using the energy consumption data collected in the data preparation state. Then we use LSTM to predict the energy consumption of customers after DR. The sum of the results obtained by LSTM may be inconsistent with the total power demand of customers. For example, the total energy consumption of the predicted results is greater or less than the historical energy consumption of customers. Thus it needs to be corrected according to certain rules. After the correction, we take the data to the large cluster for calculation and solution. When calculating the degree of unfairness according to Eqs. 14 and 15, the energy consumption of other comparison customers uses the modified LSTM prediction results corresponding to each customer. After obtaining each customer's optimal DR scheduling results under constraint, the energy consumption data is added to the training data set of the corresponding customers' LSTM networks. Then the LSTM network of each customer is retrained, and the $T2$ days of data is successively added to gradually enhance the prediction ability of the LSTM network. Finally, the network training stage is completed. The LSTM networks of customers can predict the energy consumption of customers after DR.

### 5.3 Correction of Prediction Results

For customer $u$, we define the corresponding energy consumption predicted by the LSTM network at time slot $t$ as $E_{u,t}^{\mathrm{lt}}$. Then the energy consumption set of customer $u$ predicted by the LSTM network is $E_{u,t}^{\mathrm{lt}} = \{E_{u,1}^{\mathrm{lt}}, E_{u,2}^{\mathrm{lt}}, \cdots, E_{u,t}^{\mathrm{lt}}\}$. The predicted energy consumption data and the customer's historical energy consumption data are subject to the following constraint:

$$\sum_{t=1}^{T} E_{u,t}^{\mathrm{lt}} = \sum_{t=1}^{T} E_{u,t}^{\mathrm{tot}} \tag{22}$$

If the prediction results of a customer's LSTM network can't meet the above constraints, we need to modify the prediction results before they can be used in the next step. The process of prediction result correction is divided into three steps: 1) correction of abnormal large values, 2) correction at non-scheduling time slots, and 3) correction at scheduling time slots.

(1) Correction of Abnormal Large Values: Correction of abnormal large values is to reduce and correct the data larger than the maximum value of historical energy consumption data in the prediction results. According to the calculation method of energy cost in Eq. 16, additional energy consumption deviation leads to greater expenditure, so it is impossible that the total energy consumption after DR at a certain time slot is greater than the maximum historical energy consumption on the day of dispatching. Let $e_u^{\max}$ indicates the maximum amount of energy consumed by customer $u$ in the historical energy consumption data, then we have:

$$e_u^{\max} = \max_{t \in \{1,2,\cdots,T\}} \widetilde{E}_{u,t}^{\mathrm{tot}} \tag{23}$$

Let $e_{u,t}^{\mathrm{powerMax}}$ denotes the power that customer $u$ receives in the process of correcting abnormal large values of time slot $t$, $e_u^{\mathrm{powerMax}}$ denotes the abnormal power value that customer $u$ obtains in the process of correcting abnormal large values, and $T_u^{\mathrm{powerMax}}$ denotes the time slot set when the predicted result of customer $u$ is greater than the maximum value of historical energy consumption data, so we have:

$$e_{u,t}^{\text{powerMax}}=\max\{0,E_{u,t}^{\text{lt}}-e_{u}^{\max}\}$$
$$e_{u}^{\text{powerMax}}=\sum_{t=1}^{T}e_{u,t}^{\text{powerMax}}\ T_{u}^{\text{powerMax}}=\{t\ e_{u,t}^{\text{powerMax}}>0,\forall t\in\mathcal{T}\} \tag{24}$$

In addition, the elements in the energy consumption set $E_{u,t}^{\text{lt}}$ predicted by the LSTM network are changed to:

$$E_{u,t}^{\text{lt}}=E_{u,t}^{\text{lt}}-e_{u,t}^{\text{powerMax}} \tag{25}$$

(2) Correction at Non-scheduling Time Slots: In Section 2, $I_{u,a,t}$ indicates whether customer $u$'s appliance $a$ is working or not in time slot $t$. It means that there may exist time slots in which only fixed loads are working. The non-schedulable time slots of customer $u$'s adjustable loads and schedulable loads can be defined as:

$$T_{u}^{\text{powerNoTime}}=\{t\mid I_{u,a,t}=0,\forall a\in\mathbb{C}_{u}\cup\mathbb{S}_{u},\forall t\in\mathcal{T}\} \tag{26}$$

Let $e_{u,t}^{\text{powerNoTime}}$ denotes the abnormal power value of customer $u$ at time slot $t$ through Step 2). It is calculated as:

$$e_{u,t}^{\text{powerNoTime}}=E_{u,t}^{\text{lt}}-E_{u,t}^{\text{fix}},\forall t\in T_{u}^{\text{powerNoTime}} \tag{27}$$

and the abnormal power value of customer $u$ in step 2) is calculated as:

$$e_{u}^{\text{powerNoTime}}=\sum_{t\in T_{u}^{\text{powerNoTime}}}e_{u,t}^{\text{powerNoTime}} \tag{28}$$

The result of LSTM network prediction should equal the energy consumption of fixed loads. So the calculation of prediction result after correction at non-scheduling time slots is as follows:

$$E_{u,t}^{\text{lt}}=E_{u,t}^{\text{fix}},\forall t\in T_{u}^{\text{powerNoTime}} \tag{29}$$

(3) Correction at Scheduling Time Slots: Through Step 1), the maximum value of the predicted result is less than the historical maximum value of customer $u$'s energy consumption data. The energy consumption at non-scheduling time slots equals the energy consumption of fixed loads after applying Step 2). For the abnormal power values obtained in Steps 1) and 2), the prediction results need to be further corrected in the scheduling time slots. In Eq. 18, the weight factor $w_a$ is calculated from the power. For customer $u$, the energy weight in time slot $t$can be calculated as:

$$w_{u,t}=\sum_{a\in A_{u}}I_{u,a,t}*w_{a} \tag{30}$$

Then the total energy weight of customer $u$ is defined as:

$$w_{u}=\sum_{t\in\mathcal{T}\setminus(T_{u}^{\text{powerMax}}\cup T_{u}^{\text{powerNoTime}})}w_{u,t} \tag{31}$$

In the correction step, the larger the energy weight of time slot $t$ is, the larger the abnormal power allocated should be. Therefore, the correction of the prediction result of the LSTM network in scheduling time slots is as follows:

$$E_{u,t}^{\text{lt}}=E_{u,t}^{\text{lt}}+\frac{w_{u,t}}{w_{u}}(e_{u}^{\text{powerMax}}+e_{u}^{\text{powerNoTime}}) \tag{32}$$

Upon correction of Steps 1), 2) and 3), the prediction results of the LSTM network corresponding to customer $u$ will meet the constraints of Eq. 22 and DR scheme.

## 6 COMPARISONS

To illustrate the benefits of the proposed optimization framework, a series of baseline methods are compared in this paper.

### 6.1 Random

It does not use the data of previous $T_0-1$ days and is directly verified on the data of the last day generated from the data preparation process in Fig. 3. The customers in large clusters are divided into small equal-size clusters randomly.

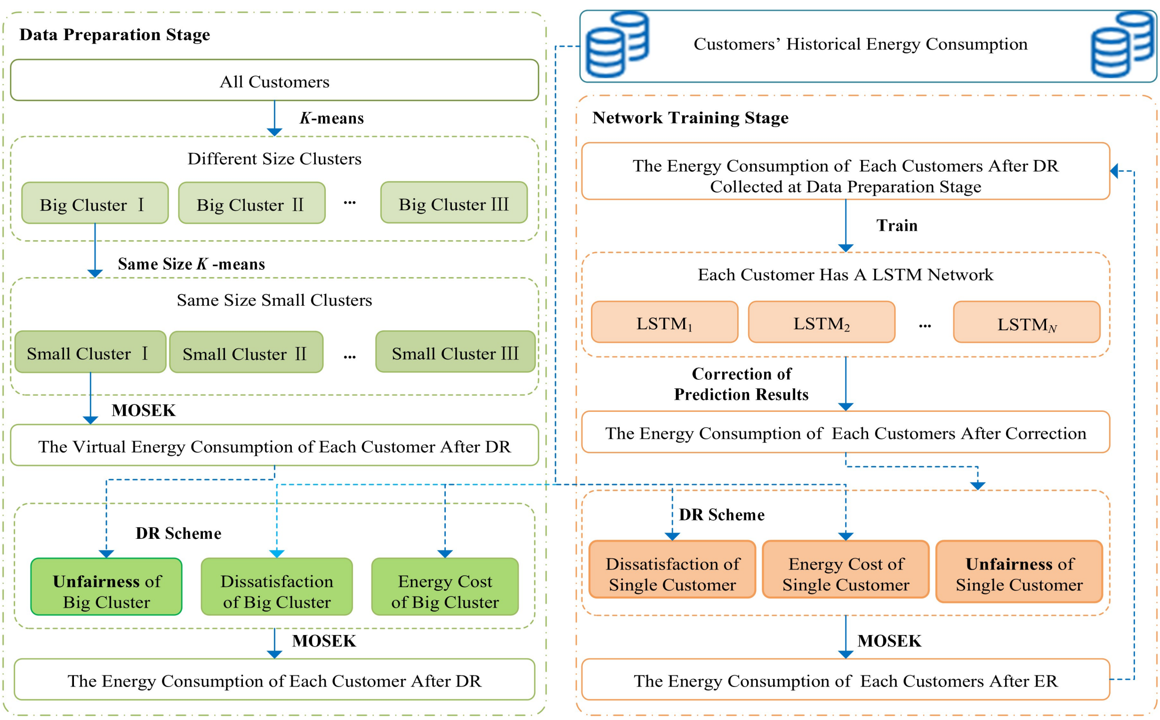



Fig. 3 Optimization framework based on clustering and deep learning

### 6.2 SS *K*-means

This method is directly validated on the data of the last day generated from the data preparation process in Fig. 3. When large clusters are further divided into small clusters, the equal-size clustering algorithm is applied.

### 6.3 SS *K*-means + LSTM

This method makes full use of the data between days of $T_1$ and $T_2$ and follows the whole process described in Section 4. After training each customer's LSTM network, the predicted and corrected results of the LSTM network are used as the energy consumption data of other customers when calculating customers' unfairness degrees. MOSEK solves the energy consumption data of a single customer after the DR scheme.

### 6.4 Best

Based on the data of the last day, we use the MOSEK toolkit to find the optimal solution directly. It is possible to use the MOSEK toolkit to find the global optimal solution in an acceptable time when the cluster is not large.

## 7 SIMULATION RESULTS

Numerical experiments are conducted to show the effectiveness of the proposed DR scheme by setting different system parameters in this section. In this work, the residential energy consumption data is collected from Pecan Street Inc. Dataport 2016[32]. Based on the data set, customer characteristics are extracted. This data set includes the electricity consumption data per minute in each household in a community located in Austin, Texas. Then, the electric vehicles (EVs) are considered as adjustable loads, while dishwashers are treated as schedulable loads.

### 7.1 Clustering Parameter Analysis

The *K*-means algorithm is used to achieve the customer clusters in the community based on 10 features. We first use the Pearson correlation coefficient to measure the selected features[33]. Fig. 4 shows the Pearson correlation coefficient thermodynamic diagram of the features. It can be observed that there is no correlation coefficient above 0.8, which indicates that the features used in this method have no strong correlation and cannot be replaced.

*K*-means is a kind of unsupervised machine learning algorithm. Because the labels of training data are unknown, the clustering effect is difficult to evaluate and display intuitively. In this paper, the effect of clustering is evaluated by the Silhouette Coefficient[34]. It can evaluate the impact of different algorithms or different operation modes on clustering results based on the same data by combining two factors of cohesion and separation[35].

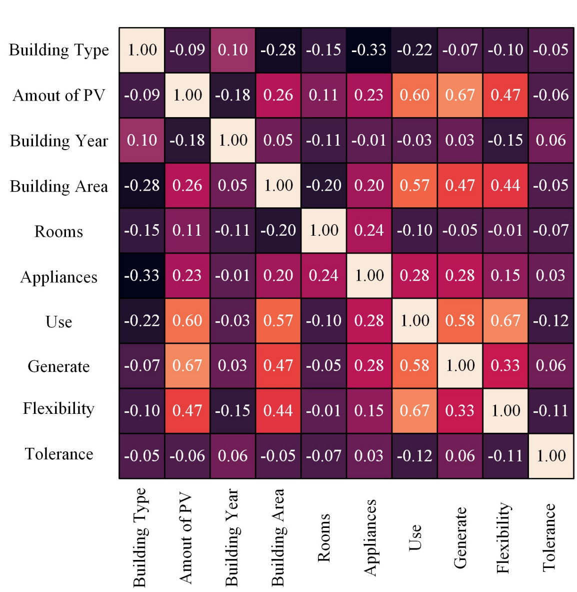



Fig. 4 Pearson correlation coefficients of clustering features

Fig. 5 shows the Silhouette Coefficients of different clusters under the same *K*-means algorithm parameters. The value of the Silhouette Coefficient reaches the maximum when the number of clusters is 3. Therefore, we set the number of clusters to 3 for further analysis.

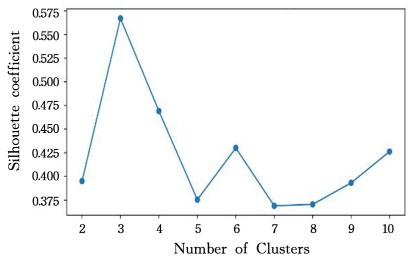



Fig. 5 Silhouette Coefficients with different numbers of clusters

### 7.2 Parameter Setting

The length of each time slot is set as 15 min. The overall length for scheduling is set as one day, i. e., $T=\{1, \cdots, 96\}$. Without loss of generality, the available charging time intervals for electric vehicles range between 12:00 AM to 7:00 AM, and 8:00 PM to 12:00 PM. $E_{u,a}^{\text{base}}$ is collected from the historical data, and the elasticities of all adjustable loads are set as $\gamma_{u,a}^{\text{down}}=1.0$ and $\gamma_{u,a}^{\text{upper}}=1.15$. The operation time interval for the schedulable loads is achieved according to the historical energy consumption data. The parameters of the dissatisfaction degree, unfairness level and total energy cost are $m^{\text{uns}}=m^{\text{unf}}=0.01$, $n^{\text{uns}}=n^{\text{unf}}=0.05$, and $m^{\text{cost}}=0.005$ and $n^{\text{cost}}=0.08$.

The quantity of customer clusters is set as 3, as shown in Table 2. Then, three DR schemes are evaluated against each other with different parameter settings. For Scheme I, customers from different clusters focus on various factors. Scheme II is defined as a baseline scheme, where the cluster is grouped by customers with the same preferences on the dissatisfaction level, unfairness level, and electricity cost. In Scheme III, the customers try to minimize the total amount of energy expenditure and ignore other factors. To evaluate the effectiveness of all schemes, we select two main performance metrics including the system PAR and electricity monetary overhead. In the following subsections, more details are included and discussed.

We first use $K$-means to cluster all customers into 3 clusters and then select 10 customers from each cluster.

### 7.3 Prediction Result Evaluation Based on Deep Learning

For each LSTM network used for training, the input is a 6-dimensional vector: the day of the week, whether the current day is a holiday, the current time slot $t$, the historical energy consumption at time slot $t-1$ of the customer, the historical energy consumption at time slot $t$ of the customer, and the energy consumption of the customer after DR at time slot $t-1$. The output is the predicted energy consumption of customer $u$ after DR at time slot $t$. In the LSTM network structure, the more hidden layers, the better the training effect of the network, but it will take a longer time to train the network. Therefore, this paper uses two hidden layers and either has 50 LSTM units.

Table 2 Different parameters of clusters

| Scheme | Cluster Name | Parameter Set | | |
|---|---|---|---|---|
| | | $\alpha$ | $\beta$ | $\gamma$ |
| Scheme I | Cluster I | 0.3 | 0.3 | 0.4 |
| | Cluster II | 0.2 | 0.2 | 0.6 |
| | Cluster III | 0.1 | 0.1 | 0.8 |
| Scheme II | Cluster I, II, III | 0.2 | 0.2 | 0.6 |
| Scheme III | Cluster I, II, III | 0 | 0 | 1.0 |

Table 3 shows the average mean absolute percentage Errors (MAPE) of the prediction results of each customer's LSTM network on the last day of comparison. Many pieces of research are based on static LSTM systems, that is, the ground truth values are prepared in advance. In this paper, the method of using LSTM is dynamic. That is to say, the ground truth values corresponding to the prediction results of the LSTM network are not prepared in advance, but the data obtained after using the prediction results in the operation. Therefore, the customers' MAPEs of our system are higher than those of the static system. Meanwhile, after the result correction steps, the MAPEs of the modified data are significantly lower than those of the original prediction output.

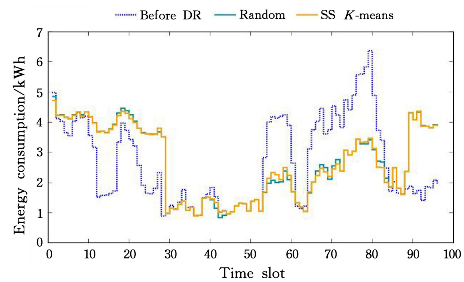
Table 3 Mapes of different methods

(unit: %)

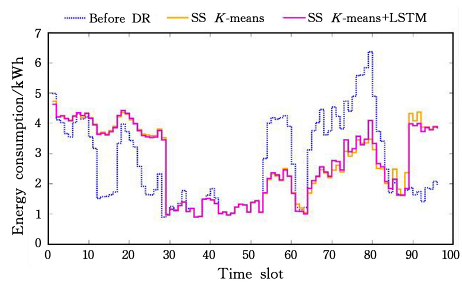
| Method | Min | Max | Average |
|---|---|---|---|
| LSTM | 10.72 | 27.24 | 18.77 |
| LSTM + Correction | 5.94 | 24.31 | 14.59 |

### 7.4 Performance Analysis of DR Scheme and Optimization Framework

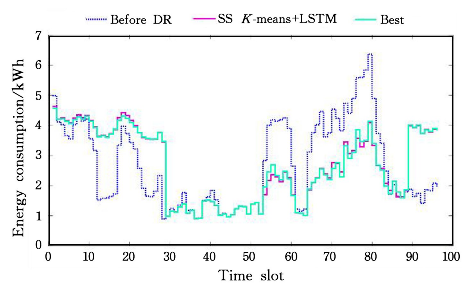
Fig. 6 shows the comparison of the overall energy consumption of the four algorithms before and after DR (Scheme I). Table 4 shows the statistics of PAR and electricity charge changes of the four algorithms under different DR schemes. Based on the results of Fig. 6 and Table 4, PAR and energy consumption cost of the system are significantly reduced under the three schemes. Among them, the random algorithm does not consider customers' personal preferences and randomly divides customers, so its PAR reduction is the smallest among the four comparison methods. SS $K$-means algorithm considers the customers' personal preferences and uses the equal-size clustering algorithm to divide customers again, which reduces the computation scale when customers compare the fairness degree. Its PAR reduction is 2.02% higher than that of the random algorithm. The algorithm of SS $K$-means + LSTM goes further. It not only makes use of the customers' personal preference information but also makes full use of the historical data. Although training the LSTM network is much more complex than the single SS $K$-means algorithm, the PAR reduction of the LSTM network is 1.51% higher than that of the SS $K$-means algorithm.




(a) The energy consumption of the random method and the SS $K$-means method before and after DR (Scheme I)




(b) The energy consumption of the SS $K$-means method and the SS K-means + LSTM method before and after DR (Scheme I)




(c) The energy consumption of the SS $K$-means + LSTM method and the Best method before and after DR (Scheme I)

Fig. 6 Energy consumption of four methods before and after DR (Scheme I)

Although the optimal solution of the SS $K$-means + LSTM algorithm is 0.84% lower than that of MOSEK in PAR, the whole set of data is required for obtaining the optimal solution. As a result, when the scale of customers becomes larger, the solution cannot be obtained in an acceptable time, or even may not be obtained. Although the method of SS $K$-means + LSTM proposed in this paper sacrifices a little precision, it reduces the complexity of computation by decomposing the whole solution into single customer solutions, which can be applied even in large-scale scenarios.

Scheme III of each method adopts the parameter setting of $\alpha=\beta=0$ and $\gamma=1$, that is, all customers only care about energy consumption but not DR. Therefore, the PARs and average prices of the system under Scheme III of the four methods have the same values. According to Eq. 16, a customer's cost is only related to the energy consumption of each time slot. At the same time, the design of the convex function makes each customer reduce the energy consumption of a single time slot as much as possible on the premise of meeting the energy consumption constraints. Therefore, the average prices after DR in Schemes I and II of the four methods are almost the same.

In Table 4, the PAR reduction of Scheme I corresponding to each method is better than that of Scheme II, which proves that for large-scale customer groups, it is necessary to divide customers into different categories according to certain rules, and fully consider the energy consumption preferences of different categories to schedule demand response management.

Table 4 Pars and average cost of different methods under different schemes

| Method | Scheme | PAR | PAR Reduced/% | Average Cost | Cost Reduced/% |
|---|---|---|---|---|---|
| | No DR | 2.381 | — | 0.129 | — |
| Random | Scheme I | 1.814 | 23.81 | 0.099 | 23.26 |
| | Scheme II | 1.840 | 22.72 | 0.099 | 23.26 |
| | Scheme III | 1.855 | 22.09 | 0.097 | 24.81 |
| SS $K$-means | Scheme I | 1.766 | 25.83 | 0.099 | 23.26 |
| | Scheme II | 1.803 | 24.28 | 0.099 | 23.26 |
| | Scheme III | 1.855 | 22.09 | 0.097 | 24.81 |
| SS $K$-means+ LSTM | Scheme I | 1.730 | 27.34 | 0.099 | 23.26 |
| | Scheme II | 1.766 | 25.83 | 0.099 | 23.26 |
| | Scheme III | 1.855 | 22.09 | 0.097 | 24.81 |
| Best | Scheme I | 1.710 | 28.18 | 0.099 | 23.26 |
| | Scheme II | 1.749 | 26.54 | 0.099 | 23.26 |
| | Scheme III | 1.855 | 22.09 | 0.097 | 24.81 |

**Conclusion** In this paper, we propose a centralized residential DR scheme for reducing customer energy cost and PARs in a SG, which is deemed as one of the important 6G enabled IIOT services. The proposed DR scheme is formulated as an optimization problem, where multiple factors are comprehensively considered. The main factors consist of the energy consumption pattern of residential appliances, electricity monetary overhead, customer satisfaction level, and fairness degree. Based on the characteristics and energy consumption behaviors of various customers, the customer features are extracted from the historical energy consumption data. Several customer clusters are obtained according to the achieved features, where different clusters focus on different performance factors. With the cluster results, the residential appliance scheduling can be improved significantly. Next, an optimization framework based on clustering and deep learning is proposed in order to eliminate the computation overhead in large-scale customer scenarios. Finally, extensive simulation experiments are conducted with real-world data sets. The simulation results show that our proposed DR scheme outperforms other baseline algorithms in managing residential appliances in terms of energy cost and PAR reductions. In addition, we also evaluate the impacts of different parameters, which may provide useful insights for the further development of rational DR strategies. With efficient and intelligent energy management, the proposed method lays helpful support for future power networks to achieve carbon peaking and carbon neutrality goals.

## REFERRENCES


[1] ESTEBSARI A, MAZZARINO P R, BOTTACCIOLI L, et al. IoT-Enabled Real-Time Management of Smart Grids with Demand Response Aggregators[J]. IEEE Transactions on Industry Applications, 2021, 58(1): 102-112.

[2] HERMANN D T, DONATIEN N, ARMEL T K F, et al. Techno-economic and environmental feasibility study with demand-side management of photovoltaic/wind/hydroelectricity/battery/diesel: A case study in Sub-Saharan Africa[J]. Energy Conversion and Management, 2022, 258: 115494.

[3] PFEIFER A, HERC L, BJELIĆ I B, et al. Flexibility index and decreasing the costs in energy systems with high share of renewable energy[J]. Energy Conversion and Management, 2021, 240: 114258.

[4] TANG R, WANG S, YAN C. A direct load control strategy of centralized air-conditioning systems for building fast demand response to urgent requests of smart grids[J]. Automation in Construction, 2018, 87: 74-83.

[5] BAHRAMI S, AMINI M H, SHAFIE-KHAH M, et al. A decentralized electricity market scheme enabling demand response deployment[J]. IEEE Transactions on Power Systems, 2017, 33(4): 4218-4227.

[6] LIANG H, MA J, LIN J. Robust Distribution System Expansion Planning Incorporating Thermostatically-Controlled-Load Demand Response Resource[J]. IEEE Transactions on Smart Grid, 2021, 13(1): 302-313.

[7] NIKZAD M, MOZAFARI B. Reliability assessment of incentive- and priced-based demand response programs in restructured power systems[J]. International Journal of Electrical Power & Energy Systems, 2014, 56: 83-96.

[8] LIU N, YU X, WANG C, et al. Energy-sharing model with

price-based demand response for microgrids of peer-to-peer prosumers[J]. IEEE Transactions on Power Systems, 2017, 32(5): 3569-3583.

[9] VARDAKAS J S, ZORBA N, VERIKOUKIS C V. A survey on demand response programs in smart grids: Pricing methods and optimization algorithms[J]. IEEE Communications Surveys & Tutorials, 2014, 17(1): 152-178.

[10] ESTHER B P, KUMAR K S. A survey on residential demand side management architecture, approaches, optimization models and methods[J]. Renewable and Sustainable Energy Reviews, 2016, 59: 342-351.

[11] GHORBANIAN M, DOLATABADI S H, SIANO P. Game theory-based energy-management method considering autonomous demand response and distributed generation interactions in smart distribution systems[J]. IEEE Systems Journal, 2020, 15(1): 905-914.

[12] LI Y, WANG C, LI G, et al. Optimal scheduling of integrated demand response-enabled integrated energy systems with uncertain renewable generations: A Stackelberg game approach[J]. Energy Conversion and Management, 2021, 235: 113996.

[13] KWAC J, FLORA J, RAJAGOPAL R. Household energy consumption segmentation using hourly data[J]. IEEE Transactions on Smart Grid, 2014, 5(1): 420-430.

[14] KANG J, LEE J H. Electricity customer clustering following experts' principle for demand response applications[J]. Energies, 2015, 8(10): 12242-12265.

[15] YANG J, ZHAO J, WEN F, et al. A model of customizing electricity retail prices based on load profile clustering analysis[J]. IEEE Transactions on Smart Grid, 2018, 10(3): 3374-3386.

[16] YANG J, ZHAO J, WEN F, et al. A framework of customizing electricity retail prices[J]. IEEE Transactions on Power Systems, 2017, 33(3): 2415-2428.

[17] HABEN S, SINGLETON C, GRINDROD P. Analysis and clustering of residential customers energy behavioral demand using smart meter data[J]. IEEE Transactions on Smart Grid, 2015, 7(1): 136-144.

[18] ZHOU K, YANG S, SHAO Z. Household monthly electricity consumption pattern mining: A fuzzy clustering-based model and a case study[J]. Journal of Cleaner Production, 2017, 141: 900-908.

[19] MELICIO R. Electricity demand profile prediction based on household characteristics[C]// 12th International Conference on the European Energy Market—EEM 2015. 2015: 1-5.

[20] SHI H, XU M, LI R. Deep learning for household load forecasting— A novel pooling deep RNN[J]. IEEE Transactions on Smart Grid, 2017, 9(5): 5271-5280.

[21] SUN M, WANG Y, TENG F, et al. Clustering-based residential baseline estimation: A probabilistic perspective[J]. IEEE Transactions on Smart Grid, 2019, 10(6): 6014-6028.

[22] FAN C, XIAO F, ZHAO Y. A short-term building cooling load prediction method using deep learning algorithms[J]. Applied Energy, 2017, 195: 222-233.

[23] CAI M, PIPATTANASOMPORN M, RAHMAN S. Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques[J]. Applied Energy, 2019, 236: 1078-1088.

[24] KONG W, DONG Z Y, JIA Y, et al. Short-term residential load forecasting based on LSTM recurrent neural network[J]. IEEE Transactions on Smart Grid, 2017, 10(1): 841-851.

[25] KIM T Y, CHO S B. Predicting residential energy consumption using CNN-LSTM neural networks[J]. Energy, 2019, 182: 72-81.

[26] YU M, HONG S H. Incentive-based demand response considering hierarchical electricity market: A Stackelberg game approach[J]. Applied Energy, 2017, 203: 267-279.

[27] MOHSENIAN-RAD A H, WONG V W S, JATSKEVICH J, et al. Autonomous demand-side management based on game-theo-retic energy consumption scheduling for the future smart grid[J]. IEEE Transactions on Smart Grid, 2010, 1(3): 320-331.

[28] ZHAO C, WANG J, WATSON J P, et al. Multi-stage robust unit commitment considering wind and demand response uncertainties[J]. IEEE Transactions on Power Systems, 2013, 28(3): 2708-2717.

[29] JAIN A K. Data clustering: 50 years beyond K-means[J]. Pattern Recognition Letters, 2010, 31(8): 651-666.

[30] SCHUBERT E, ZIMEK A. ELKI: A large open-source library for data analysis-ELKI Release 0. 7. 5" Heidelberg"[J]. arXiv: 1902. 03616, 2019.

[31] APS M. Mosek optimization toolbox for matlab[M]. User's Guide and Reference Manual, Version, 2019.

[32] Pecan street inc. dataport[EB/OL]. https://datapo- rt. pecanstreet. org/data.

[33] WANG Z F, SCAGLIONE A, THOMAS R J. Power Grid Network Analysis for Smart Grid Applications[M]. Smart Grids. 2017: 151-177.

[34] ROUSSEEUW P J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis[J]. Journal of Computational and Applied Mathematics, 1987, 20: 53-65.

[35] AL-JARRAH O Y, AL-HAMMADI Y, YOO P D, et al. Multilayered clustering for power consumption profiling in smart grids[J]. IEEE Access, 2017, 5: 18459-18468.

**Ran WANG**, born in 1989, Ph.D, associate professor. His main research interests include network performance analysis, smart grid and Internet of electric vehicles, etc.

**Jiang-tian NIE**, born in 1994, Ph.D. Her main research interests include network economics, game theory, crowd sensing and learning.

(Responsible editor: Ya-hui LI)