

ORIGINAL RESEARCH

Open Access



# Graph representation learning-based residential electricity behavior identification and energy management

Xinpei Chen<sup>1</sup>, Tao Yu<sup>1</sup>, Zhenning Pan<sup>1\*</sup> , Zihao Wang<sup>1</sup> and Shengchun Yang<sup>2</sup>

## Abstract

It is important to achieve an efficient home energy management system (HEMS) because of its role in promoting energy saving and emission reduction for end-users. Two critical issues in an efficient HEMS are identification of user behavior and energy management strategy. However, current HEMS methods usually assume perfect knowledge of user behavior or ignore the strong correlations of usage habits with different applications. This can lead to an insufficient description of behavior and suboptimal management strategy. To address these gaps, this paper proposes non-intrusive load monitoring (NILM) assisted graph reinforcement learning (GRL) for intelligent HEMS decision making. First, a behavior correlation graph incorporating NILM is introduced to represent the energy consumption behavior of users and a multi-label classification model is used to monitor the loads. Thus, efficient identification of user behavior and description of state transition can be achieved. Second, based on the online updating of the behavior correlation graph, a GRL model is proposed to extract information contained in the graph. Thus, reliable strategy under uncertainty of environment and behavior is available. Finally, the experimental results on several datasets verify the effectiveness of the proposed model.

**Keywords** Behavior correlation graph, Graph reinforcement learning, Home energy management system, Multi-label classification, Non-intrusive load monitoring

## 1 Introduction

The energy crisis is a matter of current concern all over the world. The energy consumption of residents and business end-users accounts for more than 40% of the total, and continues to rise [1]. In this context, improving energy efficiency on the demand side is particularly critical to the sustainable development of both economy and society [2, 3]. A home energy management system (HEMS) is one of the most important technologies for energy saving and emission reduction. It achieves the

maximum benefit on the demand side by promoting flexible loads to participate in demand response efficiently [4].

An efficient HEMS is built on two critical issues, i.e., identification of user behavior and energy management strategy. Behavior identification provides accurate input to the optimization model. Thus, developing a practical method which can accurately capture and describe home energy usage is critical to behavior identification. Then, an HEMS strategy can be developed after grasping user behavior. It is desirable that this intelligent strategy can not only deal with the uncertainty of exogenous information, but also achieve rapid self-adaptation for different users.

Existing research either assumes that the usage behavior is known in advance, or additional intrusive devices are used to obtain user behavior. However, considering

\*Correspondence:

Zhenning Pan  
scutpanzn@163.com

<sup>1</sup> School of Electric Power Engineering, South China University of Technology, Guangzhou 510640, China

<sup>2</sup> China Electric Power Research Institute, Nanjing 210003, China

the dynamic changes of the behavior and the heavy deployment cost of intrusive devices, these methods are not practical [5]. Using non-intrusive load monitoring (NILM) to assist behavior identification is a feasible alternative. It does not require additional investment and equipment transformation [6]. However, traditional NILM methods only disaggregate the load, but cannot realize behavior identification. In addition, current NILM methods face low disaggregation accuracy and high equipment requirement [7–9]. Thus, developing a practical and accurate online energy behavior identification method for HEMS input is still a challenging task.

For energy management strategy, some studies assume that behavior is known without specifying the source of behavioral information. Thus, these studies fail to effectively consider the dynamic uncertainty of user behavior and result in suboptimal management strategy. Reference [10] describes user satisfaction according to the time difference between strategy decisions and habits. However, the optimization decision-making process of each appliance is independent of each other in this study. This may cause unsatisfactory decisions that do not meet the expectations of users. Reference [11] considers the interdependence of specific appliances, such as the dependence between washing machines and washer dryers. However, this dependence cannot reflect the correlation between all electrical appliances. While there is a certain correlation between different usage habits of appliances, this kind of behavior correlation contains complex and unstructured data. Traditional methods, e.g., LSTM or classic Q-learning, cannot make full use of such unstructured data to make effective decisions. Existing methods for exploiting behavioral information in NILM are limited [12]. Reference [13] proposes a graph-based representation of the temporal features of appliance activities, but it fails to capture the dependencies among electricity usage patterns. In [14], label correlation is incorporated into behavior recognition, but it relies on the time series signal of appliances to capture their correlation. This may introduce errors or miss some information.

To compensate for the above-mentioned shortcomings, including the insufficient use of behavior information, inefficiency of behavior identification, and inadaptability of strategy, NILM-assisted graph reinforcement learning (GRL) is proposed for an intelligent HEMS strategy. The main contributions of this study can be summarized as follows.

- (1) A behavior correlation graph is constructed to represent the complex behavior correlation. The dynamically updated behavior correlation graph can effectively represent the dynamic habits of users, and directly provide the necessary behavior

information for an intelligent decision by the HEMS.

- (2) A behavior identification method based on multi-label NILM technology is proposed. NILM is regarded as a multi-label classification question, which can effectively reduce the model scale. This behavior identification method can not only accurately realize the function of load disaggregation, but also realize the online update of the user behavior correlation graph. The generated behavioral features are combined with electrical features to effectively improve the performance of behavior identification.
- (3) A GRL-based adaptive HEMS is proposed for excavating information from the behavior correlation graph and for providing energy management strategy. The GRL model can be adjusted based on dynamic habits and uncertain exogenous information. It continuously adapts to the changes of both internal and external factors and makes decisions that are compatible with user expectations.

## 2 HEMS framework

The proposed HEMS framework is shown in Fig. 1. The left part of Fig. 1 is the learning process of user behavior, and is based on the multi-label NILM model, namely, a multi-label sub-task gated network (ML-SGN). The learning process of user behavior comprises four parts: construction of the behavior correlation graph, model training, appliance disaggregation, and behavior update. First, in order to represent the users' initial behavior, a subgraph is extracted from the prior graph based on users' appliance information. Next, after data preprocessing such as normalization, the graph and the data are provided to the model training proportionally. Finally, when the disaggregation results are obtained, the behavior graph is updated and continuously provides the latest behavior information for accurate disaggregation and effective energy management.

The right part of Fig. 1 shows the learning process of the HEMS strategy according to the updated and learned behavior graph. The strategy generates on–off commands on the basis of load state and optimization objectives, and is a process of exploration. Subsequently, all states, actions, and rewards are sent to the replay buffer, which provides data for the learner to update strategy.

The practical deployment of the proposed method is shown in the middle part of Fig. 1. The ML-SGN model uses the aggregated data provided by an outdoor electricity meter for load disaggregation and behavior identification. The HEMS strategy generates on–off commands based on the objective function, load states, and environment states

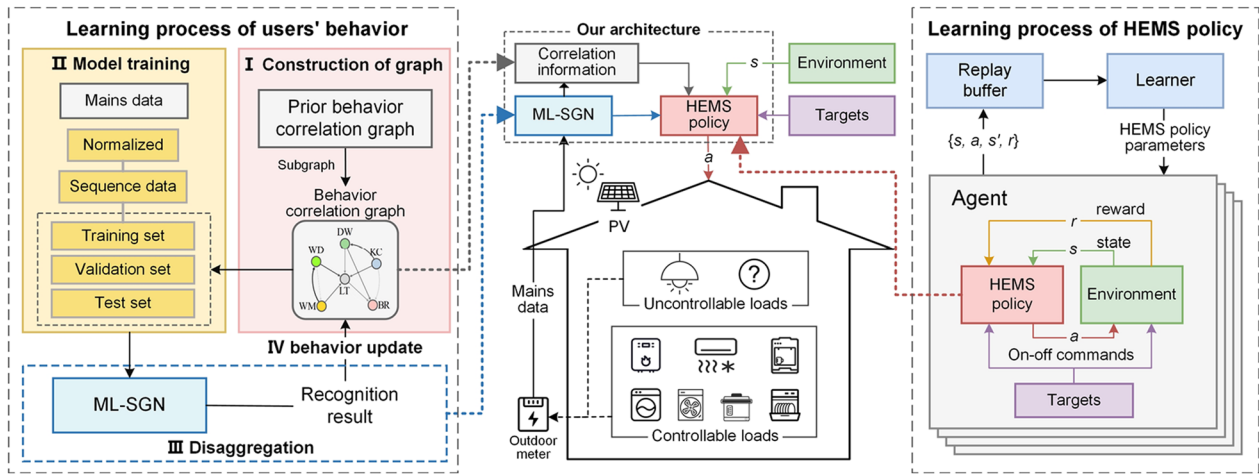


Fig. 1 HEMS framework

to guide users to manage home energy consumption. Specifically, the load state refers to the online state of loads, the environment state refers to the exogenous information, e.g., outdoor temperature and electricity price, and the correlation state refers to the correlation information in the graph.

### 3 Online behavior monitoring method

#### 3.1 Behavior representation and updating method

In this study, a graph is used to represent users' usage behavior. Essentially, behavior information consists of the usage habits of appliances in each period of time and the correlations of usage habits of different applications [15]. Among them, the usage habits of different periods can be expressed by the use probability in each period, and the behavior correlation can be expressed by the probability that an appliance is used after other appliances. Behavior correlation is a kind of complex unstructured data, which is represented by graphs in this paper. Generally, graph data consists of nodes and weighted edges. It can describe the association relationship more intuitively, and is a promising way of dealing with complex data relations.

The nodes of the behavior correlation graph represent the appliances, the edges represent the correlations between the appliances, and the weights of the edges reflect the strength of the correlations. The weight matrix is called the behavior correlation matrix in this study, which can be calculated as:

$$p_{i,j} = p(a_i|a_j) = \frac{p(a_i, a_j)}{p(a_j)} = \frac{N_{i,j}}{N_j} \quad (1)$$

where  $p_{i,j}$  represents the probability that appliance  $a_i$  works after appliance  $a_j$ .  $N_j$  is the total number of times that appliance  $a_j$  is on, and  $N_{i,j}$  represents the number of times that appliance  $a_i$  works after  $a_j$ .

A priori behavior correlation matrix is used to represent the habits of mass users, and the corresponding graph is the users' initial behavior correlation graph. The prior behavior correlation graph can be derived from other sources of data, such as may be publicly available or institutionally collected large-scale user datasets. In order to avoid the influence of signal noise and the overfitting of the priori correlation matrix,  $p$  is smoothed by threshold  $\tau$ , as:

$$\bar{A}_{i,j} = \begin{cases} 1, & p_{i,j} \geq \tau \\ 0, & p_{i,j} < \tau \end{cases} \quad (2)$$

Clearly,  $\bar{A}$  is not a symmetric matrix and the behavior correlation graph is directed, as shown in Fig. 2.

The behavior correlation matrix is updated by the prior behavior correlation matrix and the posterior behavior correlation matrix, as:

$$A^{(n)} = \lambda \times A^{(n-1)} + (1 - \lambda) \times \frac{\bar{A} + p}{2} \quad (3)$$

where the posterior matrix  $p$  refers to the correlation matrix calculated based on the online behavior data of specific users. This comes from the results of load disaggregation.  $\lambda$  is the retention ratio of the historical behavior at each iteration.

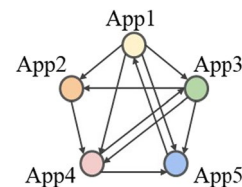


Fig. 2 Behavior correlation graph

### 3.2 NILM-based behavior identification model

The usage probability of appliances in each period and behavior correlation can be obtained by behavior identification. Different from the traditional NILM, behavior identification needs not only to identify the type of the appliance, but also to extract the behavior of appliance usage.

Early NILM studies do not consider the information contained in user habits [16, 17]. In order to make efficient use of behavior correlation information, recent research uses multi-label classification technology to consider label correlation [18, 19]. However, the traditional multi-label classification methods are either not competent for the analysis of unstructured behavior data, or the extracted behavior data cannot be effectively used for decision-making in an HEMS. Thus, this study improves the single label NILM model sub-task gated network (SGN) that is in [7], and proposes a high precision multi-label behavior identification model ML-SGN.

The structure of ML-SGN is shown in Fig. 3. The network takes the aggregated power sequence and the behavior correlation graph as input and outputs the online probability of each appliance through a sigmoid function. At the same time, from the output load disaggregation results, the behavior information of specific users is learned and updated to ensure that the dynamic behavior can be described accurately.

To make sure that the correlation of appliance behavior can be effectively learned during feature extraction, the model provides the behavior information extracted by a Graph Convolutional Network (GCN) layer to each process of electricity feature extraction. The extraction method in the orange dotted frame in Fig. 3 is the feature extraction layers in SGN, which consist of one-dimensional convolutional layers and dense layers.

## 4 Management strategy based on GRL

### 4.1 Problem formulation of HEMS

In general, residential load is divided into thermostatically controlled loads (TCL), interruptible loads (IL), transferable loads (TL), uncontrollable loads (UL), and distributed photovoltaic (PV) [20]. TCL mainly includes appliances where short-term interruptions have almost

no impact on the comfort of users, e.g., air conditioners and water heaters. IL includes appliances where users often have no usage needs but still consume electricity, e.g., water dispensers. TL includes washing machines, washer dryers, dishwashers, and cookers, whose task can be delayed. UL includes lighting, etc., and unknown types of appliances also belong to UL. Although the energy consumption of these appliances cannot be adjusted, their information reveals user behavior and this can help manage the energy of other appliances more efficiently.

*Thermostatically controlled load (TCL)* For air conditioners and water heaters, an equivalent thermodynamic model is used to reflect the state transfer process. The equivalent thermodynamic model of air conditioners [21] and water heaters [22] can be expressed as:

$$T_{n+1}^{AC} = T_{n,env} + x_n P_{AC} R_{AC} - (T_{n,env} + x_n P_{AC} R_{AC} - T_n^{AC}) \exp \frac{-\Delta t}{R_{AC} C_{AC}} \quad (4)$$

$$T_{n+1}^{WH} = [T_{n,inject}^{WH} + x_n P_{WH} R_{WH}] (1 - \exp \frac{-\Delta t}{R_{WH} C_{WH}}) + T_n^{WH} \exp \frac{-\Delta t}{R_{WH} C_{WH}} \quad (5)$$

$$T_{n+1}^{WH} = \frac{[T_{n+1}^{WH'} (V - V_{n,demand}) + T_{n,inject}^{WH} V_{n,demand}]}{V} \quad (6)$$

The temperature of air conditioners and water heaters should be controlled within specific ranges, as:

$$T_{min}^{AC} \leq T_n^{AC} \leq T_{max}^{AC} \quad (7)$$

$$T_{min}^{WH} \leq T_n^{WH} \leq T_{max}^{WH} \quad (8)$$

For  $\forall m \in M_{TCL}$ , the closer the indoor or water temperature  $T_n$  to the expected temperature  $T_{set}$ , the higher the comfort, as:

$$C_{m,n} = -|T_{set} - T_n| \quad (9)$$

*Interruptible load (IL)* For IL, user comfort is related to the difference between management strategy and usage habits. The greater the difference, the lower the comfort, and it can be formulated as:

$$C_{m,n} = -x_n p_n - \sum_{m'=1}^{|M_{UL}+M_{IL}+M_{TL}|} x_{m',n-1} A_{m',m} x_n \quad (10)$$

where  $m \in M_{IL}$ . The first item of (10) represents the difference between the decisions on which appliances are used in each period of time and the habits of users, and

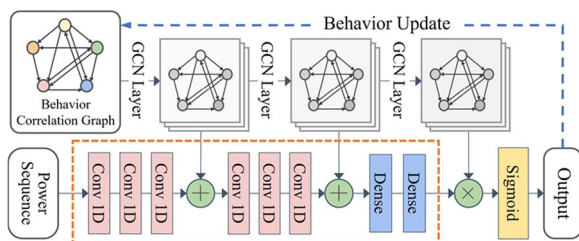


Fig. 3 Structure of ML-SGN

the second item represents the behavior correlation difference between the strategy and the previous habit.

**Transferable load (TL)** The usage comfort of TL can also be calculated by the difference between management strategy and usage habits according to (10). The work of these appliances is expected to be finished within the set time. Once the appliance is turned on, it cannot be interrupted until the work is finished. These requirements are formulated respectively, as:

$$t_{m,\min} \leq t_{m,\text{start}} \leq t_{m,\max} - \Delta t_m \quad (11)$$

$$\sum_{n=t_{m,\min}}^{t_{m,\max}-\Delta t_m} x_{m,n} \Delta t = c \times \Delta t_m \quad (12)$$

$$\sum_{n=t_{m,\text{start}}}^{t_{m,\text{start}}+\Delta t_m} x_{m,n} \Delta t = \Delta t_m \quad (13)$$

where  $m \in M_{TL}$ .

In this paper, the binary decision vector  $x_{m,n}$  is used to represent the working condition of appliances.  $x_{m,n} = 1$  indicates that the appliance is on, otherwise it is off. Therefore, a set of decision variables are defined as:

$$X = \{x_{m,n}, m \notin M_{UL}; n = 1, 2, 3, \dots, N\} \quad (14)$$

#### 4.2 Design of state

The state variables of TCL, IL, and TL are formulated respectively as follows:

$$S_n^{TCL} = \{s_n^{MP}, x_{n-1}, s_n^{TC}, s_n^{RM}, V_{n,\text{demand}}\} \quad (15)$$

$$S_n^{IL} = \{s_n^{MP}, x_{n-1}, s_n^{RM}, p_n\} \quad (16)$$

$$S_n^{TL} = \{s_n^{MP}, x_{n-1}, s_n^{CP}, s_n^{RM}, p_n\} \quad (17)$$

where these variables mainly include the working status of the appliance itself, user comfort, and corresponding usage habit information. The state of temperature comfort  $s_n^{TC}$  and the completion progress of work  $s_n^{CP}$  can be calculated as:

$$s_n^{TC} = T_n - T_{\text{set}} \quad (18)$$

$$s_n^{CP} = \frac{\sum_{n'=t_{m,\min}}^n x_{m,n'} \Delta t}{c \times \Delta t_m} \quad (19)$$

In addition to the state of various loads, system state  $S_n$  also includes the state of the associated load  $S_{n,\text{cor}}$ , the predicted power of PV  $P_n^{PV}$ , the electricity price  $\rho_n$ , and the outdoor temperature  $T_{n,\text{env}}$ , as:

$$S_n = \{S_n^{TCL}, S_n^{IL}, S_n^{TL}, S_{n,\text{cor}}, P_n^{PV}, \rho_n, T_{n,\text{env}}\} \quad (20)$$

$S_{n,\text{cor}}$  includes the correlation information of the two appliances with the largest correlation coefficient in controllable loads and the correlation information of the uncontrollable loads, as:

$$S_{n,\text{cor}} = \{S_{n,i,\text{cor}}^{(1)}, S_{n,i,\text{cor}}^{(2)}, S_{n,j,\text{cor}}^{UL}\} \quad (21)$$

where  $i \in M_{IL} \cup M_{TL}$ , and  $j \in M_{UL}$ . All this correlation information includes the corresponding appliance's id, correlation coefficient, and on-off state of the previous period.

#### 4.3 Design of reward function

The optimization target of the HEMS is to minimize the cost of energy consumption with respect to user comfort and the constraints of appliance operation. Therefore, the reward function includes three parts: the cost of energy consumption, user comfort, and the penalty for violating the constraints, which is formulated as

$$r_n = \begin{cases} R_o + C_n - \alpha E_n S_n \in S_c \\ -F_n S_n \notin S_c \end{cases} \quad (22)$$

where  $E_n$ ,  $C_n$ , and  $F_n$  denote the energy consumption cost, comfort, and penalty, respectively.  $\alpha$  is the energy consumption coefficient,  $R_o$  is a positive reward offset, and  $S_c$  is the feasible region that satisfies the constraints.

The electricity cost can be expressed as:

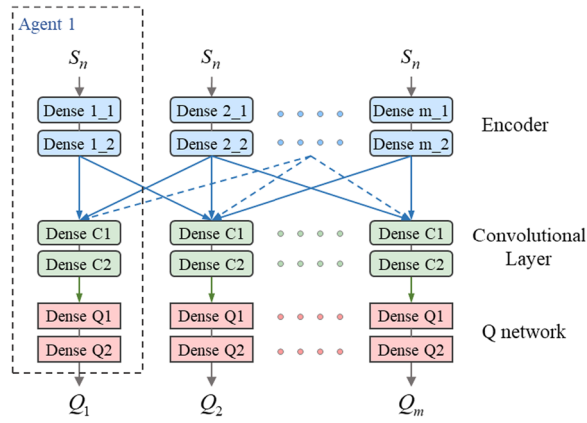
$$E_n = \rho_n \left( \sum_{m=1}^M P_m x_{m,n} - P_n^{PV} \right) \Delta t \quad (23)$$

Comfort  $C_n$  includes temperature in accordance with (9) and (10). If the mentioned constraints (7)–(8) and (11)–(13) are violated, a penalty of  $-F_n$  is added to the reward function.

#### 4.4 Design of GRL model

In order to effectively use the behavior information in the behavior correlation graph, an HEMS strategy based on GRL is proposed. The designed GRL model structure is shown in Fig. 4, where each agent corresponds to one controllable appliance in the home and manages its optimal decision. The input of the model is the state of each appliance, and the output is the Q value of each action. The observation encoder layer consists of two dense layers. Since the states of different appliances are different, independent dense layers are employed to code for different types of appliances. The convolutional layer and the Q network are also made up of two dense layers, which try to collect the





**Fig. 4** Structure of GRL model

features of appliances and obtain the action value. Since the features have been extracted by the encoder, the parameters of the convolutional layer and Q network can be shared among different appliances. The sharing of parameters ensures the model size will not surge with an increase in the number of loads.

To ensure the effective exploration of the action to obtain strategy improvement, a random decision is performed with probability  $\epsilon$ , and the optimal decision  $\hat{X}_n$  of the model is performed with probability  $1 - \epsilon$ . The exploration rate decreases  $\Delta\epsilon$  after each epoch of training until it reaches the lowest value  $\epsilon_d$ . The optimal decision  $\hat{X}_n$  can be obtained by:

$$\hat{X}_n = \arg \max G(S_n) \quad (24)$$

where  $G(\cdot)$  represents the GRL model.

The goal of model training is to minimize the value of the loss function. The loss function of the model training is formulated as:

$$L(\theta) = \sum_{n=1}^{N-1} \frac{1}{|M|} \sum_{m=1}^M (y_{m,n} - G(S_{m,n}; \theta)) \quad (25)$$

$$y_{m,n} = r_{m,n} + \gamma \max G'(S_{m,n+1}; \theta') \quad (26)$$

where  $L(\cdot)$  denotes the loss function, and  $G'(\cdot)$  is the target network.  $\theta$  and  $\theta'$  are the parameters of each network.  $S_{m,n}$  and  $r_{m,n}$  represent the observation and reward values of the appliance  $m$ , respectively.  $\gamma$  denotes the discount rate of the reward.

## 5 Performance evaluation

### 5.1 Datasets

To provide a fair comparison between the proposed behavior identification method and existing methods,

REDD [23] and REFIT [24] datasets are used in the experiments.

The REDD dataset provides energy consumption data of six houses. To avoid the influence of insufficient samples, the data of house 1 and house 3, which is relatively sufficient, is selected in this study. To realize a reasonable comparison, the same preprocessing method in the SGN model [7] is used for the REDD dataset. The appliances in house 1 include dishwasher, fridge, microwave, and washer dryer, while those in house 3 are electronic load, dishwasher, electric furnace, fridge, microwave, and washer dryer.

The REFIT dataset provides electric power measurements from 20 households. The first 10 houses are experimented with in the official preprocessed version. Each house contained energy consumption data of 9 appliances.

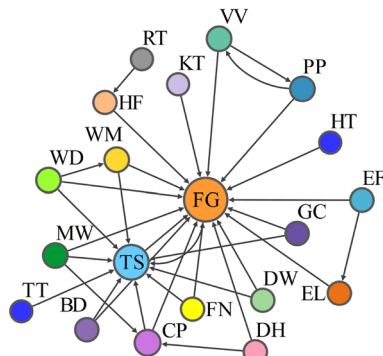
To distinguish the houses in the two datasets, house 1 and house 3 in REDD are abbreviated as B1 and B3, and the first 10 houses in REFIT are denoted by H1–H10, respectively.

For energy management strategy, because neither of these two datasets contains all types of appliances studied in this paper, the behavior information is constructed by the combination of dataset extraction and behavior customization. In order to ensure the authenticity of behavior information, energy consumption data with the house id of "1240" in the Pecan Street dataset [25] is chosen to extract the behavior. This contains the data of all the transferable loads in this study for 6 months. Additionally, the usage probability of a water dispenser in each period is customized based on the habits of most users. The lighting and appliances in the bedroom are regarded as the uncontrollable loads in this study, while it is assumed that water heaters and air conditioners will not be turned off without an HEMS. In addition, it is also necessary to consider the uncertainties of PV output, electricity price, outdoor temperature, and user demand. Thus, some disturbances are added according to [20].

### 5.2 Data preprocess

For the ML-SGN model, the remaining data of REDD and REFIT are used to construct a prior behavior correlation graph as shown in Fig. 5. The threshold  $\tau$  for the process of construction is set to 0.25. The labels in the dataset and their abbreviations are shown in Table 1.

The initial behavior correlation graph for behavior identification is a subgraph extracted by the a priori graph, and the extraction method is to set the related edges of appliances that the user does not have to 0. To compare with the SGN model more fairly, the processing methods



**Fig. 5** A priori behavior correlation graph of ML-SGN

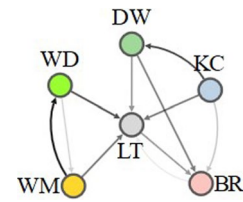
of SGN in data preprocessing and some parameter selections are followed. The inputs of ML-SGN are the power sequence with length of 512 and the behavior correlation graph. The retention ratio  $\lambda$  is set to 0.95. The output of the model is the on–off state of each appliance at the mid-point of the sequence. The working power threshold of each appliance is set as 15 W. Additionally, the aggregated data is normalized by Z-score method before training [26], and the appliances data is normalized by the max–min method [27]. The model is trained by the Adam algorithm [28], and the loss function can be expressed as:

$$L' = - \sum_{m=1}^{|M|} (o_m \log \hat{o}_m + (1 - o_m) \log(1 - \hat{o}_m)) \quad (27)$$

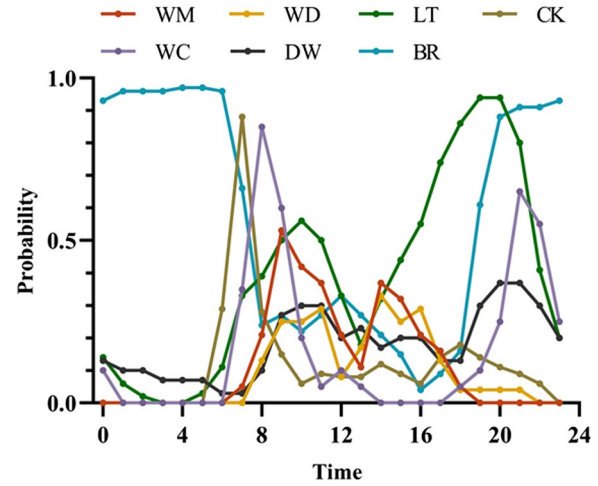
where  $o_m$  and  $\hat{o}_m$  represent the on–off state of appliance  $m$  and the predicted working probability, respectively.

For the GRL model, time step  $\Delta t$  is equal to 1 h and  $N$  is 24. The behavior correlation graph of GRL and the usage probability of each appliance are shown in Figs. 6 and 7, respectively. These are calculated from the data in the Pecan Street dataset.

Load parameters such as air conditioner and water heater are shown in Table 2, while Table 3 shows the



**Fig. 6** Behavior correlation graph of GRL



**Fig. 7** Usage probability

**Table 2** Parameters of WH and AC

	$P$ (kW)	$R$ (°C/kW)	$C$ ((Kw h)/°C)	$T_{\min}$ (°C)	$T_{\max}$ (°C)
AC	1.8	1.8	0.525	23	28
WH	3.6	0.7623	431.701	54	70

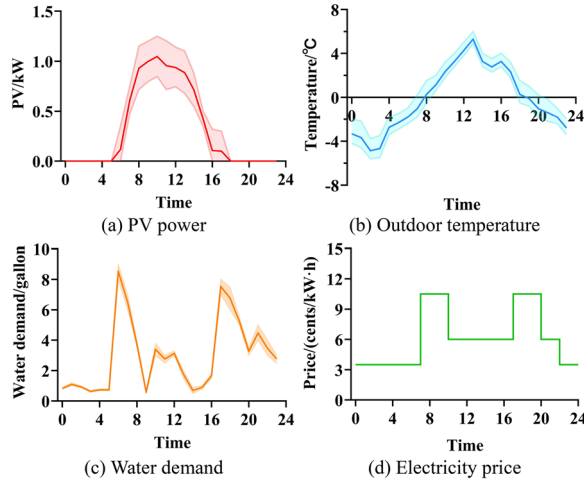
parameters of the transferable loads, in accordance with [20, 29]. The curves of PV output, temperature, water demand and, electricity price used in the simulation are plotted in Fig. 8.

**Table 1** Labels and abbreviations of appliances

Label name	Abbreviation	Label name	Abbreviation	Label name	Abbreviation
Dishwasher	DW	Kettle	KT	Television site	TS
Fridge, freezer	FG	Overhead fan	FN	Toaster, bread-maker	TT
Microwave	MW	Blender, k mix	BD	Games console	GC
Washer/tumble dryer	WD	Electric heater	HT	Vivarium	VV
Electronics	EL	Router	RT	Pond pump	PP
Electric furnace	EF	Dehumidifier	DH	Hi-fi	HF
Washing machine	WM	Computer	CP	Water dispenser	WC
Light plug	LT	Cooker	CK	Bedroom appliances	BR

**Table 3** Parameters of transferable loads

	$P_m$ (kW)	$t_{m,min}$	$t_{m,max}$	$\Delta t_m$ (h)	$c$
WM	1	7:00	19:00	2	1–2
WD	0.8	8:00	21:00	1	
DW	0.4	7:00	23:59	1	
CK	2.4	6:00	23:00	1	

**Fig. 8** Predicted data

The exploration rate  $\epsilon$  is 0.65,  $\Delta\epsilon$  is 0.02,  $\epsilon_d$  is 0.02, and the discount rate  $\gamma$  is 0.95. These parameters are the optimal values selected after multiple experiments and comparative analyses. The injected water temperature is 8 °C. The temperature ranges of air conditioner and water heater are between 23–28 °C and 54–70 °C, respectively. The volume of the water heater tank  $V$  is 40 gallons, the average rated power of lighting and bedroom appliances are 0.1 kW and 0.8 kW, respectively. The penalty value  $F_n$  is 10, and the reward offset  $R_0$  is 10.

### 5.3 Evaluation method for behavior identification

Hamming loss (HL), accuracy (Acc), and F1-Score are used to evaluate the ML-SGN model [30]. Hamming loss  $L_{HL}$  is a classical evaluation method of multi-label classification, and is used to reflect the misclassification of the model, and can be calculated as:

$$L_{HL} = \frac{1}{N_H} \sum_{n=1}^{N_H} 1(\hat{o}_n \neq o_n) \quad (28)$$

where  $N_H$  is the number of samples. The effect of  $1(\cdot)$  is the logical judgement of whether  $\hat{o}_n \neq o_n$ . If it is true, it equals 1, otherwise it is 0.  $L_{HL}$  is the error rate, and the smaller the value, the more accurate the prediction.

Acc and F1-Score are commonly used in single-label classification. Both values range from 0 to 1, and the larger the value, the better the performance.

### 5.4 Behavior identification result

To demonstrate that the proposed ML-SGN model not only outperforms the traditional multi-label classification methods, but also has stronger recognition ability than the original single-label model, it is compared with the classical multi-label classification model and multi-label k-nearest neighbor algorithm (MLKNN) [31], random k-label sets algorithm (RAKEL) [32], and single-label model SGN. To prove its superior performance, it is also compared with the load disaggregation with attention model (LDWA) [8], which is a more advanced model built on SGN using attention technique.

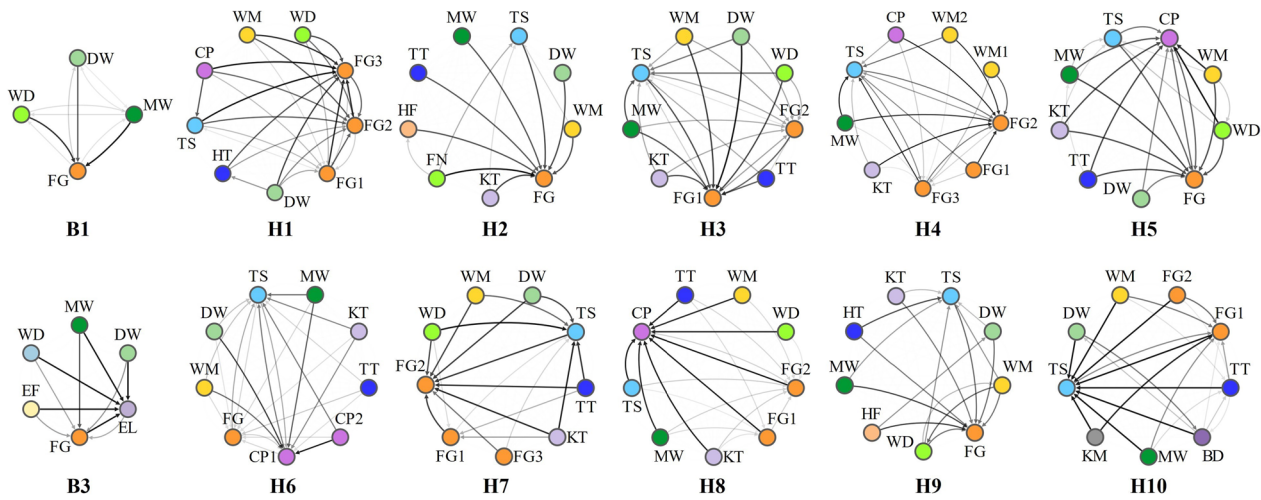
Figure 9 shows the behavior correlation graph of the 12 households obtained when the experiments are performed on the latest data. The thicker edges in the graph represent the greater weights. The experimental results of the proposed behavior identification model on the 12 houses are shown in Table 4, where the best performance of each result is highlighted in bold.

The experimental results show that the performance of the SGN, LDWA and ML-SGN models is significantly better than that of MLKNN and RAKEL. Moreover, except for B3 of REDD and H10 of REFIT, the experimental results of other houses indicate that the improved ML-SGN model achieves better results in terms of hamming loss, accuracy and F1-score. The average recognition accuracy of ML-SGN reaches 93.2%. From the experimental results of B3 and H10, it can be inferred that considering the behavior correlation of appliances does not always improve the recognition accuracy or may even deteriorate performance. This is because of the overfitting of correlation features caused by few appliances in these two houses.

### 5.5 Results evaluation of GRL

This section presents numerical simulation results to evaluate the performance of the NILM-based HEMS. Users' energy consumption is simulated considering uncertainties of environment and usage behavior. The convergence of average Q value and constraints violations in training are shown in Fig. 10. As seen, each agent can converge to the maximum Q value after training. At the beginning of the training, the average Q value of the agent is negative because it is easy to violate the constraints (7)–(8) and (11)–(13). After continuous exploration, the agent gradually learns how to produce more proper actions, and the average Q value gradually increases from negative to positive. Finally, it converges





**Fig. 9** Final behavior correlation graph

**Table 4** The result of behavior identification

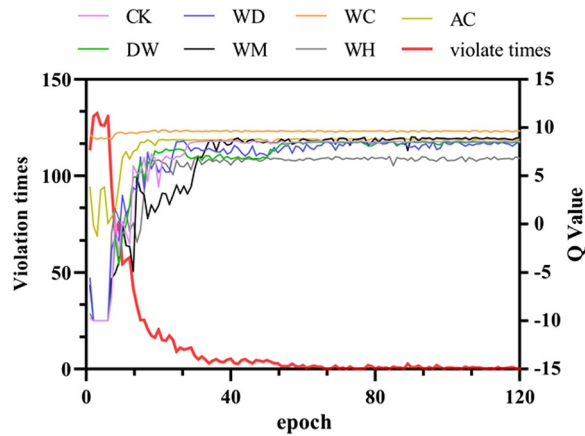
Model	Method	B1	B3	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10
MLKNN	HL	0.024	0.093	0.139	0.059	0.118	0.109	0.174	0.158	0.085	0.131	0.049	0.151
	Acc	0.976	0.906	0.861	0.941	0.882	0.891	0.826	0.842	0.915	0.869	0.951	0.849
	F1-Score	0.825	0.805	0.467	0.568	0.588	0.604	0.415	0.528	0.589	0.527	0.72	0.693
RAKEL	HL	0.02	0.069	0.142	0.06	0.119	0.11	0.171	0.143	0.081	0.129	0.062	0.144
	Acc	0.98	0.931	0.858	0.94	0.881	0.89	0.829	0.857	0.919	0.871	0.938	0.855
	F1-Score	0.84	0.857	0.181	0.54	0.576	0.591	0.4	0.506	0.586	0.502	0.601	0.653
SGN	HL	<b>0.011</b>	0.032	0.078	0.043	0.081	0.075	0.13	0.135	0.05	0.107	0.025	<b>0.131</b>
	Acc	<b>0.989</b>	<b>0.97</b>	0.922	0.957	0.919	0.925	0.87	0.865	0.95	0.893	0.975	<b>0.869</b>
	F1-Score	0.918	<b>0.939</b>	0.698	0.69	0.718	0.72	0.511	0.558	0.765	0.64	0.861	<b>0.728</b>
LDWA	HL	<b>0.011</b>	<b>0.030</b>	0.072	<b>0.041</b>	0.078	0.077	0.117	0.133	0.052	0.101	0.021	<b>0.131</b>
	Acc	<b>0.989</b>	<b>0.97</b>	0.930	<b>0.959</b>	0.922	0.923	0.881	0.865	0.948	0.901	0.975	<b>0.869</b>
	F1-Score	0.918	<b>0.939</b>	0.733	0.706	0.733	<b>0.723</b>	0.537	0.563	0.769	0.653	0.876	<b>0.728</b>
<b>ML-SGN (proposed)</b>	HL	<b>0.011</b>	0.033	<b>0.065</b>	<b>0.041</b>	<b>0.072</b>	<b>0.073</b>	<b>0.102</b>	<b>0.124</b>	<b>0.048</b>	<b>0.097</b>	<b>0.02</b>	<b>0.131</b>
	Acc	<b>0.989</b>	0.967	<b>0.935</b>	<b>0.959</b>	<b>0.928</b>	<b>0.927</b>	<b>0.898</b>	<b>0.876</b>	<b>0.952</b>	<b>0.903</b>	<b>0.98</b>	<b>0.869</b>
	F1-Score	<b>0.92</b>	0.937	<b>0.781</b>	<b>0.72</b>	<b>0.752</b>	0.719	<b>0.574</b>	<b>0.584</b>	<b>0.777</b>	<b>0.684</b>	<b>0.886</b>	0.726

to the maximum Q value, while the violation of constraints also gradually disappears and user comfort is ensured.

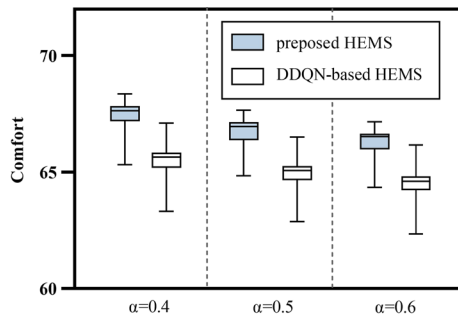
After the model converges, the optimal management strategy of appliances can be carried out. To evaluate the proposed method's performance on user comfort, it is applied with double deep q learning (DDQN) [33] to an HEMS with varying energy coefficients. The comfort can be calculated by (9), (10) and (22). For a more intuitive analysis of comfort level, the electricity cost  $E_n$  is set to 0. Figure 11 shows the comfort results. It can be concluded that the proposed method outperforms DDQN on user comfort across all energy coefficients, and user

comfort declines with increasing energy coefficient, indicating that users can trade off comfort for lower energy consumption.

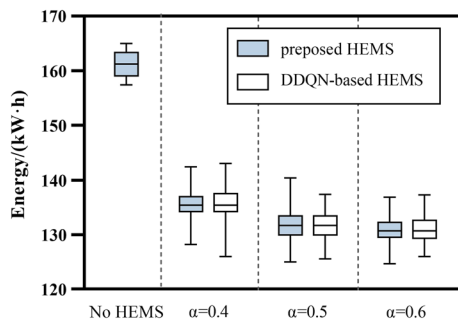
Similarly, the energy consumptions of the proposed method and DDQN under different energy coefficients are compared, as shown in Fig. 12. It can be inferred that the two methods have comparable energy consumption under different energy coefficients. Compared with the case without HEMS, the daily cost of applying the proposed method is significantly reduced, by 15.9%, 18.3%, and 18.7%, respectively. Additionally, the larger the energy coefficient, the smaller the daily cost. The effect of  $\alpha$  is to balance energy saving and



**Fig. 10** Training process of GRL

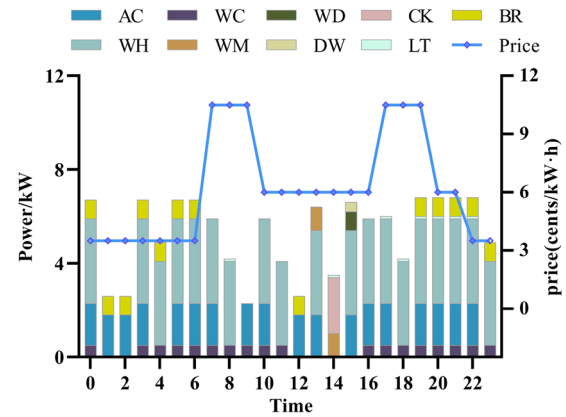


**Fig. 11** Comfort comparison

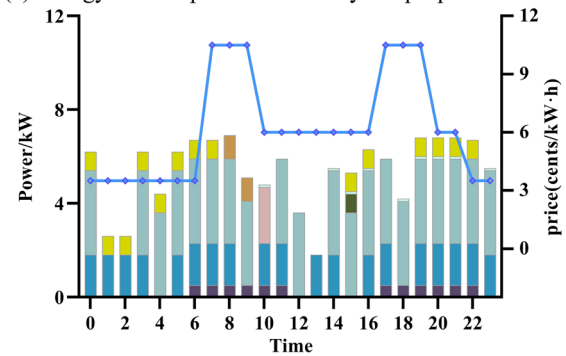


**Fig. 12** Electricity cost comparison

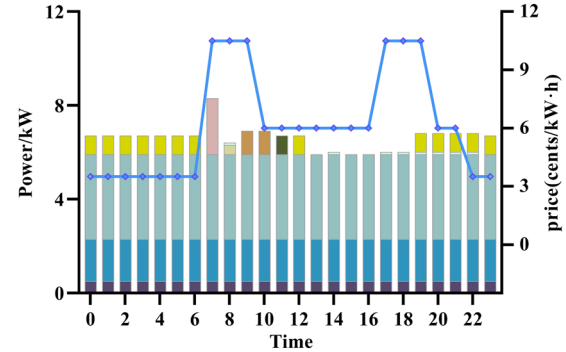
comfort, and greater energy coefficient means that users are more concerned about saving energy than comfort. Therefore, the energy coefficient can be set according to users' usage preference. From the aforementioned experimental results, it can be concluded that the proposed method balances comfort and energy consumption better under uncertainty.



(a) Energy consumption obtained by the proposed method



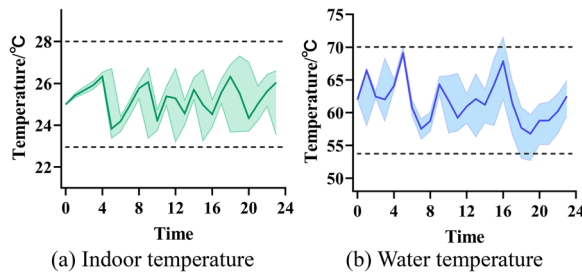
(b) Energy consumption obtained by DDQN-based HEMS



(c) Energy consumption without HEMS

**Fig. 13** Energy consumption of all loads obtained by different methods

The total energy consumption of all loads in three different scenarios is compared when  $\alpha$  is set to 0.5, as shown in Fig. 13. The three scenarios are: using the proposed HEMS method, using the DDQN-based HEMS method, and without HEMS. The time-of-use price for the day is also shown in the figure. As shown in Fig. 13a, when the HEMS method proposed in this work is deployed, the loads consume more energy when the price is low, and reduce the demand when the price is high. Specifically, the transferable loads avoid working



**Fig. 14** Variation of temperature

at peak times, and their energy demand is postponed to the period of moderate electricity price between 10:00 and 16:00. Thermostatically controlled loads are not only expected to reduce the energy cost, but also to ensure that the temperature is kept within the comfort range. The variation curves of the indoor temperature and water temperature of the water heater are shown in Fig. 14. It can be concluded that under the influence of various uncertain factors, the temperature can be maintained within the set range, and the energy consumption during peak times can also be effectively controlled. In contrast, in the case without HEMS, as shown in Fig. 13c, most transferable loads work at peak times, which results in higher cost. In addition, the proposed method also ensures the rationality of the management. Comparing with Fig. 13b, it can be seen that after applying the proposed method, the clothes dryer always works after the washing machine, and the dishwasher generally works after the cooker. This validates that the proposed method combined with behavior correlation can better deal with energy management than the method without incorporating correlations.

## 6 Conclusion

In this study, a novel method for residential electricity behavior identification and energy management based on graph representation learning is presented. The proposed method constructs and updates a graph that captures users' electricity usage habits, and leverages an improved multi-label NILM method to identify their behavior. Moreover, the method proposes an HEMS strategy based on GRL, one which addresses the anomaly management problem arising from ignoring appliance correlation in conventional methods. The proposed method can adapt to users' changing behavior by online updating of the graph, and assist them in continuous energy management.

The proposed method is evaluated through simulations which demonstrate its superior performance in behavior identification and HEMS. The proposed method has two main advantages over existing ones. First, it achieves a

high average recognition accuracy of 93.2% in the experiments, demonstrating its effectiveness in behavior identification. Second, it reduces the average electricity cost for users by 18.3%, while maintaining a high level of user comfort and satisfaction, and making management decisions that match user preferences. Therefore, the method balances user comfort and energy cost better than other methods.

In future work, we will continue to tackle the overfitting caused by the 'few shot' learning problem to further improve the generalization performance of behavior identification. At the same time, it is of great significance to migrate the proposed method to software and hardware systems.

### Abbreviations

AC/WH/PV	Air conditioner/water heater/photovoltaic
DDQN	Double deep q network
GRL	Graph reinforcement learning
HEMS	Home energy management system
IL/TL/UL	Interruptible/transferable/uncontrollable load
LDWA	Load disaggregation with attention
ML-SGN	Multi-label sub-task gated network
MLKNN	Multi-label k-nearest neighbor algorithm
NILM	Non-intrusive load monitoring
RAKEL	Random k-label sets algorithm
SGN	Sub-task gated network
TCL	Thermostatically controlled load

### List of symbols

$A_{m',m}$	Probability of appliance $m$ work after $m'$
$C_{AC}/C_{WH}$	Equivalent thermal capacity of AC/WH
$C_{m,n}$	Comfort value of appliance $m$
$E_n/C_n/F_n$	Electricity cost/comfort/penalty
$M$	Set of all loads
$M_{UL}/M_{TCL}$	Set of UL/TCL
$M_{IL}/M_{TL}$	Set of IL/TL
$N$	Total number of manage steps in one day
$P_m$	Rated power of appliance $m$
$P_{AC}/P_{WH}$	Rated power of AC/WH
$P_n^{PV}$	Predicted power of PV
$R_{AC}/R_{WH}$	Equivalent thermal resistance of AC/WH
$S_n^{TCL}/S_n^{IL}/S_n^{TL}$	Self state of TCL/IL/TL
$S_{n,cor}$	State of the correlated load
$S_{n,i,cor}^{(1)}$	Correlation information of appliance with the largest correlation coefficient
$S_{n,i,cor}^{(2)}$	Correlation information of appliance with the second largest correlation coefficient
$S_{n,i,cor}^{UL}$	Correlation information of appliance with the largest correlation coefficient in UL
$T_n^{AC}/T_{n,env}$	Current indoor/outdoor temperature
$T_n^{WH'}$	Current water temperature without considering the injected water
$T_n^{WH}/T_{n,inj}$	Current heated/injected water temperature
$T_{min}^{AC}/T_{max}^{AC}$	Lower/upper bounds of indoor temperature
$T_{min}^{WH}/T_{max}^{WH}$	Lower/upper bounds of water temperature
$V/V_{n,demand}$	Volume of water heater tank/water demand
$C$	Working times required in one day
$P_n^{MP}$	Usage probability in the $n$ th period
$S_n^{CP}$	State of management permission (1/0)
$S_n^{DC}$	Completion progress of work in one day
$S_n^{TC}$	State of temperature comfort
$S_n^{RM}$	Remain controllable times in one day

$t_{m, \text{start}}$	Start time of appliance $m$
$t_{m, \text{min}}/t_{m, \text{max}}$	Lower/upper bounds of controllable time
$x_n$	On/off decision in the $n$ th period (1/0)
$x_{m, n}$	On/off decision in the $n$ th period of appliance $m$ (1/0)
$\rho_n$	Time-of-use price in the $n$ th period
$\Delta t$	Length of time step
$\Delta t_m$	Length of working time required in one day

## Acknowledgements

Not applicable.

## Author information

### Xinpei Chen

received the B.Eng. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2020, where he is currently pursuing the M. Eng. degree with the School of Electric Power Engineering. His research interests include artificial intelligence techniques and its application in smart grid.

### Tao Yu

received the B.Eng. degree in electrical power system from Zhejiang University, Hangzhou, China, in 1996, the M.Eng. degree in hydroelectric engineering from Yunnan Polytechnic University, Kunming, China, in 1999, and the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2003. He is currently a Professor with the College of Electric Power, South China University of Technology, Guangzhou, China. His research interests include nonlinear and coordinated control theory, artificial intelligence techniques, and operation of power systems.

### Zhenning Pan

received the B.Eng. and Ph.D. degrees in electrical engineering from the South China University of Technology, Guangzhou, China, in 2016 and 2021, respectively. His major research interests include intelligent operation and optimization of smart grid, and demand response.

### Zihao Wang

received the B.Eng. degree in electrical engineering from the Hunan University, Changsha, China, in 2020, where he is currently pursuing the M.Eng. degree with the School of Electric Power Engineering. His research interests include intelligent terminal and topology identification of low-voltage distribution network.

### Shengchun Yang

received the B.S. degree from Huazhong University of Science and Technology, Wuhan, China in 1995, M.S. degree from Nanjing Automation Research Institute, Nanjing, China in 1998 and Ph.D. degree from Huazhong University of Science and Technology, Wuhan, China in 2016. He is currently working for China Electric Power Research Institute as associate director. His research interests include demand response, AI applications in power system operations with high penetration of flexible load and renewable generation.

## Author contributions

XC carried out theoretical analysis of the process and performed simulation and experiment to verify the proposed method, TY and ZP offered help in theory and practice, read and put forward suggestions for the paper. ZW contributed to the electrical model simulation experiment. SY guided and assisted the manuscript revision and improvement. All authors read and approved the final manuscript.

## Funding

This work is supported by State Grid Corporation of China Project "Research on Coordinated Strategy of Multi-type Controllable Resources Based on Collective Intelligence in an Energy" (5100-202055479A-0-0-00).

## Availability of data and materials

Not applicable.

## Declaration

### Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received: 23 December 2022 Accepted: 20 June 2023

Published online: 30 June 2023

## References

- Pérez-Lombard, L., Ortiz, J., & Pout, C. (2008). A review on buildings energy consumption information. *Energy & Buildings*, 40(3), 394–398.
- Zhang, D., Yao, L., & Ma, W. (2013). Development strategies of smart grid in china and abroad. *Proceedings of the CSEE*, 31(31), 2–14.
- Liu, S., Zhou, C., Guo, H., Shi, Q., Song, T. E., Schomer, I., & Liu, Y. (2021). Operational optimization of a building-level integrated energy system considering additional potential benefits of energy storage. *Protection and Control of Modern Power Systems*, 6(1), 1–10.
- Zhixin, Fu., Ziyang, Li., Junpeng, Z., & Yue, Y. (2022). Multi-user multi-time-scale power packages and home energy optimization strategies. *Power Systems Protection and Control*, 50(11), 21–31.
- Çimen, H., Çetinkaya, N., & Vasquez, J. C. (2021). A microgrid energy management system based on non-intrusive load monitoring via multitask learning. *IEEE Transactions on Smart Grid*, 12(2), 977–987.
- Lin, Y. H., & Tsai, M. S. (2017). An advanced home energy management system facilitated by nonintrusive load monitoring with automated multiobjective power scheduling. *IEEE Transactions on Smart Grid*, 6(4), 1839–1851.
- Shin, C., Joo, S., Yim, J., Lee, H., & Rhee, W. (2019). Subtask gated networks for non-intrusive load monitoring (Vol. 33, pp. 1150–1157).
- Piccialli, V., & Sudoso, A. M. (2021). Improving non-intrusive load disaggregation through an attention-based deep neural network. *Energies*, 14, 847.
- Xiu, Y., An, Li., Gaiping, S., et al. (2022). Non-invasive load monitoring based on an improved GMM-CNN-GRU combination. *Power Systems Protection and Control*, 50(14), 65–75.
- Lu, R., Hong, S. H., & Yu, M. (2019). Demand response for home energy management using reinforcement learning and artificial neural network. *IEEE Transactions on Smart Grid*, 10(6), 6629–6639.
- Berk, C., Robin, R., Siddharth, S., David, B., & Abdellatif, M. (2017). Electric energy management in residential areas through coordination of multiple smart homes. *Renewable and Sustainable Energy Reviews*, 80, 260–275.
- Zhai, S., Zhou, H., Wang, Z., et al. (2020). Analysis of dynamic appliance flexibility considering user behavior via non-intrusive load monitoring and deep user modeling. *CSEE Journal of Power and Energy Systems*, 6(1), 41–51.
- Peng, B., Pan, Z., Yu T., et al. Graph data modeling and graph representation learning methods and their application in non-intrusive load monitoring problem[J/OJ]. In *Proceedings of the CSEE* (in Chinese).
- Nalmpantis, C., & Vrakas, D. (2020). On time series representations for multi-label NILM. *Neural Computing and Applications*, 32, 17275–17290.
- Kong, W., Dong, Z. Y., Hill, D. J., Ma, J., Zhao, J. H., & Luo, F. J. (2016). A hierarchical hidden Markov model framework for home appliance modeling. *IEEE Transactions on Smart Grid*, 9, 3079–3090.
- He, D., Lin, W., Liu, N., & Harley, R. G. (2013). Incorporating non-intrusive load monitoring into building level demand response. *IEEE Transactions on Smart Grid*, 4(4), 1870–1877.
- Lam, H. Y., Fung, G., & Lee, W. K. (2007). A novel method to construct taxonomy electrical appliances based on load signatures. *IEEE Transactions on Consumer Electronics*, 53(2), 653–660.
- Tabatabaei, S. M., Dick, S., & Xu, W. (2017). Toward non-intrusive load monitoring via multi-label classification. *IEEE Transactions on Smart Grid*, PP(1), 1–1.

19. Singhal, V., Maggu, J., & Majumdar, A. (2018). Simultaneous detection of multiple appliances from smart-meter measurements via multi-label consistent deep dictionary learning and deep transform learning. *IEEE Transactions on Smart Grid*, 10, 2969–2978.
20. Su, Y., Zhou, Y., & Tan, M. (2020). An interval optimization strategy of household multi-energy system considering tolerance degree and integrated demand response. *Applied Energy*, 260, 114.
21. Lei, Y. U., Tang, Q., & Zhang, J. (2015). Optimal operation for residential micro-grids based on load resources classification modelling and heuristic strategy. *Power System Technology*, 39, 2180–2187.
22. Du, P., & Ning, L. (2012). Appliance commitment for household load scheduling. In *Transmission & distribution conference & exposition*. IEEE.
23. Kolter, J. Z., & Johnson, M. J. (2011). REDD: A public data set for energy disaggregation research. In *Artificial intelligence* (Vol. 25).
24. Murray, D. (2015). A data management platform for personalised real-time energy feedback. In *Proc. 8th int. conf. energy efficiency domestic appl. lighting (EEDAL)* (pp. 1–15).
25. Pecan street inc. dataport [EB/OL]. <https://dataport.pecan-street.org/data>.
26. Al Shalabi, L., Shaaban, Z., & Kasasbeh, B. (2006). Data mining: A preprocessing engine. *Journal of Computer Science*, 2(9), 735–739.
27. Xia, M., Liu, W., Wang, K., Zhang, X., & Xu, Y. (2019). Non-intrusive load disaggregation based on deep dilated residual network. *Electric Power Systems Research*, 170, 277–285.
28. Kingma, D., & Ba, J. (2014). Adam: A method for stochastic optimization. In *ICLR 2015*.
29. Wang, J., Li, Y., & Zhou, Y. (2016). Interval number optimization for household load scheduling with uncertainty. *Energy & Buildings*, 130(Oct), 613–624.
30. Lin, W. Z., Fang, J. A., Xiao, X., et al. (2013). iLoc-Animal: A multi-label learning classifier for predicting subcellular localization of animal proteins. *Molecular BioSystems*, 9(4), 634–644.
31. Zhang, M. L., & Zhou, Z. H. (2006). Multilabel neural networks with applications to functional genomics and text categorization. *IEEE Transactions on Knowledge and Data Engineering*, 18(10), 1338–1351.
32. Tsoumakas, G., Katakis, I., & Vlahavas, I. (2010). Random k-labelsets for multilabel classification. *IEEE Transactions on Knowledge and Data Engineering*, 23(7), 1079–1089.
33. Hasselt, H. V., Guez, A., & Silver, D. (2015). Deep reinforcement learning with double Q-learning[J]. In *Computer ence*.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)