A Multi-objective Non-intrusive Load Monitoring Method Based on Deep Learning

Hua Yu1,3, Zhiwei Jiang1, Yaping Li2, Jing Zhou2, Ke Wang2, Zifeng Cheng1 and Qing Gu1

1 Department of Computer Science and Technology, Nanjing University, Nanjing 210023, China
2 China Electric Power Research Institute Co., Ltd. (Nanjing), Nanjing 210003, China
3 Corresponding author

E-mail address: mg1733085@smail.nju.edu.cn; Telephone number: +86 18860902811

Abstract. Fine-grained power data are used to characterize user’s behavior in smart grid. Non-intrusive load monitoring can effectively separate the load of a single electrical appliance from the whole energy consumption of a dwelling, which is beneficial to fully exploit the load potential. In this paper, considering the relationship on usage habits among electrical appliances, the characteristics of energy consumption, we propose a multi-objective modeling method based on deep learning. The multi-objective model is constructed by CNN and LSTM, and the neural network is collaboratively optimized by multi-objective outputs. The experimental results show that the multi-objective model performs better than the other models on the five appliances, reducing the absolute error to less than 10 and the total error of the standardized signal to less than 0.1.

1. Introduction

The construction of smart grid emphasizes the construction of interactive service system, and the demand side user participation in interactive response is important. Therefore, fully grasping the user's power consumption status is the basis for mining their interactive potential. Fine-grained user power consumption information can also help to better grasp the user's power usage dynamics. How to use load monitoring technology to collect and analyze user electricity data in real time is the key to solve this problem.

Non-Intrusive Load Monitoring (NILM) was proposed by George Hart in the mid-1980s [1-2], which is a single-channel blind source separation problem that requires the identification and separation of multiple target consumers from the whole house electricity consumption. In order to achieve energy saving and reduce the pressure of grid, these data is used to help analyze the user's electricity usage behavior to work out a demand response strategy. Researches have shown that detailed electricity consumption data can reduce the average annual electricity consumption of users by 12% [3].

The issue of NILM has been an active area of pattern recognition and machine learning [4]. For the NILM problem, Some researches builds the model based on the Hidden Markov Model (HMM), and regards the load data of the electrical appliance as the observable sequence, while using the electrical operating state as the hidden state[5-8]. The hidden state sequence is calculated and derived from the observable sequence. However, these methods require a relatively stable operation state of the...
electrical appliance, and the load fluctuation characteristics are obvious when the state is switched. But modern control appliances are embedded with a control processor, so that the state of the appliance is no longer a simple finite state. There is a limit to the actual use of such methods in the current scenario. Based on HMM method performs well on controllable multi-state appliances, but their performance is degraded for non-controllable multi-state appliances[9]. NILM can be regarded as a classification problem[10]. Firstly, the load curve is monitored for the event. Secondly, train the Bayesian classifiers for every electrical appliance, classifiers identify each monitored event to the state of the applied electrical device. Finally, obtain a complete sequence of electrical usage. Graph signal processing methods are also can be used to deal with NILM problem[11-12]. All of the above methods have high complexity, lack of flexibility, and the response speed is not suitable for real-time application scenarios.

The neural network has achieved many achievement in many fields under the support of powerful computing power [13-16]. In literature[17], a Convolutional Neural Network (CNN), a Recurrent Neural Networks (RNN) and a Auto Encoder(AE) are respectively constructed to solve the NILM problem, which proves the applicability of the neural network model to the NILM problem. Deep neural networks are combined with HMM to solve the NILM problem[18]. Literature [19] proposed a different architecture called sequence-to-point learning to improve the model identification results. These methods are limited by the single output neural network used, and are affected by the characteristics of different networks. So only some of the information on the load curve can be utilized, the final identification result is still unsatisfactory.

In this paper, we propose a multi-objective non-intrusive load monitoring method based on deep learning. Considering the local characteristics of the load curve and the relation of long-term time series information, we also propose a multi-objective neural network structure that can fully learn and utilize the load curve information. Two weak constraints are given for the multi-output characteristics of the proposed multi-objective neural network structure. A naive and universal objective function is designed for multi-objective method. Multi-objective method avoids the repeated modeling of single output models, reduces the cost of model training and saves computation time. Finally, the effectiveness of the method is verified by experiments.

2. A multi-objective non-intrusive load monitoring method based on deep learning

2.1. Formalization of the problem

The purpose of NILM is to decompose the total load data into specific electrical load data. Let the total load data time series of length N be \( X = \{ x_0, x_1, \cdots, x_{N-1} \} \). The total load data is superimposed by the electrical load, and there are a total of M corresponding target appliances. And the data time series of the s electrical appliance is \( Y_s = \{ y_{s,0}, y_{s,1}, \cdots, y_{s,N-1} \} \), \( s=1,2,\ldots,M \). The load data generated by the non-target electrical appliance is regarded as noise and is represented by e. The total load can be expressed as:

\[
x_n = \sum_{s=1}^{M} y_{s,n} + e_n, n = 0, 1, \ldots, N-1
\]

(1)

The goal of this paper is to identify and separate the electrical load from X, denoted as \( Y_s^* \), and the separation result should be as close as possible to the target real data \( Y_s \). From the perspective of probability theory, the essence of the problem is the posterior distribution estimation, which can be directly learned and fitted through the neural network. It is not necessary to use the pre-processing algorithm such as event detection to perform complex processing on the total load data. The following is a simple proof.

The total load X and the electrical appliance load Y are generally subject to the distribution \( \pi ( Y | X ) \), X and Y are vectors of length N and aligned in the time dimension, this distribution cannot be directly observed. The NILM solves a set of parameters \( \theta \) such that the probability distribution \( p(Y | X, \theta) \).
\( \theta \) approximates \( \pi(Y|X) \) as much as possible. In other words, we can think it finds a set of parameters \( \theta \) that minimize the cross entropy, as shown in equation (2):

\[
\min_{\theta} KL(\pi \| \rho) = \min_{\theta} \int \pi(Y|X) \log \frac{\pi(Y|X)}{\rho(Y|X,\theta)} dY
\]

(2)

The standardization of this problem is interpreted as minimizing a Monte Carlo approximation of the cross entropy. This paper assumes such a decomposable form:

\[
p(Y|X,\theta) = \prod_{n=1}^{N} p_n(y_n|X,\theta)
\]

(3)

Then the target becomes the form of equation (4):

\[
KL(\pi \| p) = \sum_{n=1}^{N} KL(\pi(y_n|X) \| p(y_n|X,\theta))
\]

(4)

Let the total load \( X \) be given, the conditional probability of \( \theta \) is:

\[
\phi_{n}(\theta|X) = KL(\pi(y_n|X) \| p(y_n|X,\theta))
\]

(5)

And assume a normal distribution like (6):

\[
p(Y|X,\theta) = \mathcal{N}(\mu(\theta),\Sigma) = \prod_{n=1}^{N} \mathcal{N}(\mu(\theta),\Sigma) = \prod_{n=1}^{N} p_{n}(y_n|X,\theta)
\]

(6)

Where \( \mu(\theta) = (\mu_1(\theta),...,\mu_N(\theta))^T \), \( \Sigma \) is a constant, \( I \) is unit matrix, all distributions \( p_n \) are share the same set of parameters \( \theta \). To find the optimal solution, that is, the \( \theta \) that minimizes the cross entropy. We need to consider the performance of \( p_n(n=1,2,...,N) \) over the entire data sequence, as in equation (7):

\[
\min_{\theta} \sum_{n=1}^{N} \phi_{n}(\theta|X)
\]

(7)

From the above process, it is seen that the essence of the NILM problem is the posterior distribution estimation, which requires the parameter \( \theta \). Therefore, the neural network can be trained directly based on the total load data, the parameter \( \theta \) is calculated and the identification decomposition is completed.

2.2. Framework

For the formalization in Section 2.1, this paper proposes a multi-objective non-intrusive load monitoring framework (MLM) based on deep learning. MLM takes mixed load data as input, and is multi-output model. Firstly, a set of sliding window length and sliding step length are determined according to the type of electrical appliance, and the mixed load data is processed; Then, the deep neural network is trained by using the set of network structure parameters and the objective function. Finally, the load data of each electrical appliance is obtained from the total load data through the deep neural network at the same time. The overall framework of the MLM method is shown in Figure 1.

![Figure 1. The framework of the MLM method](image)

2.2.1. Data Processing. Sliding window unit: This unit uses the sliding window to process the long-term sequence data of the input and output to simplify the computational complexity and speed up the calculation. Using sliding window processing is to train the neural network each time with fixed length data mapped by a window (the window moves according to certain rules during training), instead of
directly using the complete data for neural network training. Therefore, design such a sliding window unit in the data processing module. And its input is the complete mixed load data and use the "window" for segmentation for subsequent processes. Assume data is \( L \), the window function \( w(i,t) \) is as in equation (8). After sliding the window unit, data \( L^s \) is obtained. The sliding window unit repeatedly calls the window function to adjust the value of \( i \) by adjusting the value of \( i \) and setting the sliding step size to complete the cutting of the data.

\[
w(i,t) = \begin{cases} \frac{|t-i|s}{2} & i \leq 0, \\
0 & \text{other} \end{cases}
\]

\[i \] is the data point index at the center of the window, \( t \) is the time point index in the current data sequence, and \( s \) is the set window length. After the window function, a data sequence \( L^s \) of length \( s \) is cut out from the original data sequence \( L \). The sliding window unit repeatedly calls the window function to adjust the value of \( i \) by adjusting the value of \( i \) and setting the sliding step size to complete the cutting of the data.

**Separation unit:** This unit performs necessary denoising on the mixed load data, and separates the total load data from the source data of the target electric appliance for training the deep neural network.

2.2.2. Multi-objective Deep Neural Network Structure. After the above data processing, the data is obtained. The multi-objective deep neural network is used to model the NILM problem.

Convolutional Neural Network (CNN) is a deep feedforward neural network with good feature learning and feature extraction capabilities. The basic structure of CNN consists of two layers: one of which is a convolutional layer, the input of each neuron of the layer is connected to the local receptive field of the previous layer, and the local features are extracted. Once the local feature is extracted, its positional relationship with other features is also determined; the other one is the pooling layer, where each computing layer is composed of multiple feature maps. Each feature map is a plane with equal weights for all neurons on the plane. These two layers have different functions and can be used separately according to the needs. For example, each household appliance has certain electrical characteristics that reflect the change in its electrical load and can be used to extract and identify transient changes in the load through the CNN. Although the CNN structure has a good ability to recognize local features, it is limited by its structure and cannot learn long-term dependent information. For example, washing machine runs continuously for a period of time. The width of CNN layer is not wide enough to cover the time interval between start and stop, so CNN can’t learn the relation between this pair of state.

The Recursive Neural Network (RNN) structure is good at analyzing the context of data. Long short-term memory Neural Network (LSTM) is a variant of RNN. Because of its hidden layer memory cell, LSTM can retain long-term dependence information and avoid the vanishing gradient problem that is easy to occur in traditional RNN.

Specifically, for the \( n \)-th data in the input data \( L^n \), we define the LSTM unit at time \( t \), with forgetting gate \( f_{n,t} \), input gate \( i_{n,t} \), output gate \( o_{n,t} \), a memory unit \( C_{n,t} \) and a hidden state \( h_{n,t} \), where \( f_{n,t} \), \( i_{n,t} \) and \( o_{n,t} \) have values in the range \([0, 1]\). The entire LSTM unit transformation process is as follows [12]:

\[
\begin{align*}
    i_{n,t} &= \sigma(W_i[L^n_t, h_{n,t-1}] + b_i) \\
    f_{n,t} &= \sigma(W_f[L^n_t, h_{n,t-1}] + b_f) \\
    o_{n,t} &= \sigma(W_o[L^n_t, h_{n,t-1}] + b_o) \\
    C_{n,t} &= \tanh(W_C[L^n_t, h_{n,t-1}] + b_C) \\
    h_{n,t} &= o_{n,t} \odot \tanh(C_{n,t})
\end{align*}
\]

Where \( W \) is the weight of each gate, \( b \) is the offset of each gate, \( \sigma \) is the Sigmoid function, and its data output range is \([0, 1]\).
Although LSTM can fully learn long-term dependency information, the ability to identify local features is still less than CNN. So this paper designs a multi-objective neural network structure. The network structure is shown in figure 2.

![Figure 2](image)

**Figure 2.** The framework of the MLM method

In figure 2, conv2D/ReLU represents a two-dimensional convolutional network layer, followed by a ReLU activation function, and LSTM + Drop out represents an LSTM layer followed by a Drop out layer.

The multi-objective deep neural network proposed in this paper contains a local feature extraction sub-network, a long-term dependence extraction sub-network, and an information integration analysis sub-network. The local feature extraction sub-network is a full convolutional neural network, all composed of convolutional layers, and the total load data is mapped to low-dimensional features through a series of convolutional layers; The long-term dependence information extraction sub-network learns and transmits the long-term dependent information through the LSTM layer, and adds a Drop out layer between the LSTM layers to reduce the transmission of redundant information; Finally, local features and long-term dependence information are input into a multi-layer perceptron composed of a fully connected network. And this network integrates and analyzes information, and output the load decomposition identification results of each target appliance.

2.2.3. **Objective Function.** In this paper, the final output is \( \hat{Y} \) through the deep neural network, which is used to feedback and train the entire deep neural network. In the past, the model used in the research work was a single output model [19], and the model can only be adjusted by the single target output. So the commonly used objective function is:

\[
J = \frac{1}{N} \| Y - Y^* \|^2
\]  

(10)

Where \( Y^* \) is the actual load data of the appliance \( s \).

The multi-objective neural network structure proposed in this paper is a multi-output network. We can learn parameters of the neural network by coordinating the output of multiple electrical appliances. We give two weak constraints: 1) the load of a single electrical appliance will not exceed the total load; 2) the sum of the load of each electrical appliance is equal to the total load.

For the first weak constraint, the output of multi-target neural network is \( Y^s \) of the total load of each appliance. And the value in the range[0,1]. Then Calculate the electrical load process according to \( Y^s \) as shown in equation (11):

\[
\hat{Y} = Y^s \otimes X
\]  

(11)

According to the two weak constraints, the multi-objective neural network optimization objective function is equation (12):
Where $Y^C$ is the total load after noise is removed at the current time. This formula applies to all such multiple output models under the NILM problem.

2.2.4. Performance Evaluation. This paper uses the following two metrics to compare models from both single-point performance and overall performance. Assume $Y^* = \{y_1^*, y_2^*, ..., y_T^*\}$ is the decomposition result of the model output, $Y = \{y_1, y_2, ..., y_T\}$ is the actual value, $T$ is the length of the data sequence.

Mean absolute error (MAE) is used to measure the load error at each point:

$$MAE = \frac{1}{T} \sum_{t=1}^{T} |\hat{y}_t - y_t|$$  \hspace{1cm} (13)

Signal aggregate error (SAE) is used to measure the overall error over time:

$$SAE = \frac{\hat{r} - r}{r}$$  \hspace{1cm} (14)

$\hat{r}$ and $r$ are respectively the total amount of load and the actual load obtained by the model decomposition over a period of time. The equation is (15) and (16):

$$\hat{r} = \sum_{t=1}^{T} y_t$$  \hspace{1cm} (15)

$$r = \sum_{t=1}^{T} y_t$$  \hspace{1cm} (16)

3. Experiment

This article conducted experiments on six NILM models: AFHMM [6], seq2seq (CNN) [17], seq2point (CNN) [19], MLM, MO-cnn (multi-output model that use only CNN) and MO-lstm (multi-output model that use only LSTM). These models are implemented in the Python language and are based on TensorFlow which is an open source machine learning framework.

3.1. Dataset

This paper conducts an experiment on the UK-DALE dataset [20] contributed by Kelly et al., which records load data for five households in the UK. The load data includes the total load and the load of each appliance. The data set is sampled every 6 seconds, and the recording interval is from November 2012 to January 2015. The data set contains more than 10 household appliances, but for the convenience of comparison with other models, we will only focus on five types of appliances such as electric kettle, microwave oven, refrigerator, dishwasher and washing machine. The electrical parameters are listed in Table 1. In this paper, 95% of the data of 5 households is used as training data, and the remaining data is test data.

| appliances   | Maximum power | On power threshold | Mean on power | Min. on duration(secs) |
|--------------|---------------|-------------------|---------------|-----------------------|
| Kettle       | 3948          | 2000              | 700           | 12                    |
| Microwave    | 3138          | 200               | 500           | 12                    |
| Fridge       | 2572          | 50                | 200           | 60                    |
| Dish washer  | 3230          | 10                | 700           | 1800                  |
| Washing m.   | 3962          | 20                | 400           | 1800                  |

3.2. The parameters of network

The network structure used in the experimental is as follows:

The structure of local feature extraction subnetwork is listed in Table 2:

| Network layer | Filter/Step | Number of channel | Activation function |
|---------------|-------------|-------------------|---------------------|

6
The structure of long-term dependence extraction subnetwork is listed in Table 3:

| Network layer | output length | Dropout Rate | Activation function |
|---------------|---------------|--------------|---------------------|
| LSTM          | 512           | -            | Tanh                |
| Dropout       | -             | 20%          | -                   |
| LSTM          | 128           | -            | Tanh                |
| Dropout       | -             | 20%          | -                   |
| LSTM          | 64            | -            | Tanh                |

The multi-layer perceptron structure of the information integration analysis sub-network is three layers. The first layer is 1024 nodes, the activation function is ReLU, the second layer is 64 nodes, the activation function is ReLU, the third layer is 5 nodes, and the activation function is Sigmoid.

3.3. Experiment Result

3.3.1. Result. The results of the six models are shown in the Table 4. It can be seen that 3 multi-output models: MLM, MO-cnn, and MO-lstm are all superior to other single-output models in overall performance. This proves the validity of the multi-output model proposed in this paper. Among the three models, the MLM performance of the multi-target network structure designed in this paper is optimal.

| Metric   | Model        | Kettle | Microwave | Fridge | Dishwasher | Washing m. | Overall     |
|----------|--------------|--------|-----------|--------|------------|------------|-------------|
| MAE      | AFHMM        | 47.38  | 21.18     | 42.35  | 199.84     | 103.24     | 82.79 ± 64.50 |
|          | seq2seq      | 13.000 | 14.559    | 38.451 | 237.96     | 163.468    | 93.488 ± 91.112 |
|          | seq2point    | 7.439  | 8.661     | 20.894 | 27.704     | 12.663     | 15.472 ± 7.718  |
|          | MLM          | 1.866  | 1.179     | 1.779  | 2.577      | 3.131      | 2.106 ± 0.667  |
|          | MO-cnn       | 2.720  | 1.875     | 2.023  | 3.013      | 3.336      | 2.593 ± 0.563  |
|          | MO-lstm      | 3.342  | 2.300     | 2.962  | 2.957      | 3.247      | 2.962 ± 0.364  |
| SAE      | AFHMM        | 1.06   | 1.04      | 0.98   | 4.50       | 8.28       | 3.17 ± 2.88   |
|          | seq2seq      | 0.085  | 1.348     | 0.502  | 4.237      | 13.831     | 4.001 ± 5.124 |
|          | Seq2point    | 0.069  | 0.486     | 0.121  | 0.645      | 0.284      | 0.321 ± 0.217 |
|          | MLM          | 0.006  | 0.032     | 0.002  | 0.012      | 0.018      | 0.014 ± 0.007 |
|          | MO-cnn       | 0.019  | 0.084     | 0.011  | 0.012      | 0.053      | 0.0362 ± 0.028 |
|          | MO-lstm      | 0.023  | 0.285     | 0.006  | 0.013      | 0.011      | 0.064 ± 0.110 |

3.3.2. Model Analysis. Comparing the three multi-output models, the overall performance of MO-lstm in washing machines and dishwashers is better than MO-cnn. On other electrical appliances, MO-cnn performs better. Because the interval between start and stop is long, the CNN is not good at learning such long-term dependence features, while other electrical appliances change state quickly, the period of characteristic change is short, and there are not many long-term dependence features, so characteristics of curve change are concentrated in the local, CNN performs better in this situation.
However, MLM performs better than MO-lstm and MO-cnn on all electrical appliances, because the multi-objective deep neural network used by MLM takes into account the advantages of CNN and LSTM, it can learn the local characteristics of the load curve, and can also learn the relation between features on a longer interval. The experimental results are also in line with our expectations, and the multi-objective deep neural network structure performs optimally.

4. Conclusion
This paper proposes a multi-objective NILM method based on deep learning: MLM. Experiments are carried out on the real dataset UK-dale using this method. The experimental results show that the multi-output structure of MLM has better load identification ability than the single output model. The designed objective function can adapt well to multi-output scenes. The multi-target network structure used also achieved the intended purpose and achieved good load identification results. Therefore, based on the non-intrusive load detection technology proposed in this paper, it will provide strong support for mining user interaction potential.

In future work, we will further expanding the data set to verify the universality of MLM in solving the NILM problem. At the same time, mining other intrinsic connections between multi-output appliances and improving the MLM method.

Acknowledgments
This work was supported by Innovation Foundation of EPRI (Essential Theory and Methodology of Multivariate Interactive User Behavior Analysis Based on Weakly Supervised Learning ,Grant No.DZ83-18-008).

References
[1] Hart G W. Prototype nonintrusive appliance load monitor[R]. Cambridge :MIT Energy Laboratory, 1985.
[2] Hart G W. Nonintrusive appliance load monitoring[J]. Proceedings of the IEEE, 1992, 80(12):1870-1891.
[3] Armel K C, Gupta A, Shrimali G, et al. Is disaggregation the holy grail of energy efficiency? The case of electricity ☆[J]. Energy Policy, 2013, 52(1):213-234.
[4] Faustine A, Myungi N H, Kaigaje S, et al. A Survey on Non-Intrusive Load Monitoring Methods and Techniques for Energy Disaggregation Problem[J]. 2017.
[5] Kim H S, Marwah M, Arlitt M F, et al. Unsupervised disaggregation of low frequency power measurements[C]/ Eleventh SIAM International Conference on Data Mining. DBLP, 2011:747-758.
[6] Kolter J Z, Jaakkola T S. Approximate Inference in Additive Factorial HMMs with Application to Energy Disaggregation[C]// Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics. JMLR.org, 2012:1472-1482
[7] Parson O, Ghosh S, Weal M, et al. Non-intrusive load monitoring using prior models of general appliance types[C]/ Twenty-Sixth AAAI Conference on Artificial Intelligence. AAAI Press, 2012:356-362.
[8] Makonin S, Popovich F, Bajić I V, et al. Exploiting HMM Sparsity to Perform Online Real-Time Nonintrusive Load Monitoring[J]. IEEE Transactions on Smart Grid, 2016, 7(6):2575-2585.
[9] Mauch L, Barsim K S, Bin Y. How well can HMM model load signals. Proceeding of the 3rd International Workshop on Non-Intrusive Load Monitoring (NILM 2016), 2016.
[10] Marchiori A, Hakkarinen D, Han Q, et al. Circuit-Level Load Monitoring for Household Energy Management[J]. IEEE Pervasive Computing, 2010, 10(1):40-48.
[11] Zhao B, Stankovic L, Stankovic V. On a Training-Less Solution for Non-Intrusive Appliance Load Monitoring Using Graph Signal Processing[J]. IEEE Access, 2017, 4:1784-1799.
[12] Stankovic V, Liao J, Stankovic L. A graph-based signal processing approach for low-rate energy
disaggregation[C]// IEEE Symposium on Computational Intelligence for Engineering Solutions. IEEE, 2014:81-87.

[13] Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological Cybernetics, 1980, 36(4):193-202.

[14] Hochreiter, S, Schmidhuber, J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8):1735-1780.

[15] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.

[16] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553):436.

[17] Kelly J, Knottenbelt W. Neural NILM:Deep Neural Networks Applied to Energy Disaggregation[C]// Acm International Conference on Embedded Systems for Energy-efficient Built Environments. ACM, 2015:55-64.

[18] Mauch L, Yang B. A novel DNN-HMM-based approach for extracting single loads from aggregate power signals[C]// IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2016:2384-2388.

[19] Zhang C, Zhong M, Wang Z, et al. Sequence-to-point learning with neural networks for non-intrusive load monitoring[C]// Proceedings of the Thirty-Second {AAAI} Conference on Artificial Intelligence, (AAAI-18). AAAI Press,2018:2604-2611.

[20] Kelly J , Knottenbelt W . The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes[J]. 2015. doi:10.1038/sdata.2015.7