Non-intrusive Load Monitoring Method Considering the Time-segmented State Probability

Lina Liu1, Fangshuo Li1, Zhijiong Cheng1, Yifei Zhou1, Jie Shen1, Ruizhao Li1, Siyu Xiong2

1 State Grid Sichuan Electric Power Company Metering Centre, Chengdu 610031, China
2 School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China

Corresponding author: Lina Liu (e-mail: 610139482@qq.com).

This work was supported by the State Grid Sichuan Electric Power Company Science and Technology Project (52199720005Z).

ABSTRACT Appliance-level data is a prerequisite for establishing friendly two-way interactions between customers and the power company, and this data is now mainly obtained by non-intrusive load monitoring. However, as the number of loads increases, the number of possible appliances state combinations tends to grow exponentially, leading to a significant increase in the time of load identification. In order to reduce the search range of the load state combinations and shorten the algorithm response time, a non-intrusive load monitoring method based on the time-segmented state probability is proposed in this paper. Firstly, the affinity propagation (AP) clustering algorithm is introduced to obtain the power templates of the load, and then the power templates are used to count the time-segmented state probabilities. Secondly, a number of appliance state matrices are generated using the probabilities, and the optimal matrix is selected by the function as the identification result of the appliance state. Finally, the performance of the algorithm is tested on the public NILM dataset and compared to several state-of-the-art techniques. The results illustrate that the proposed method achieves an accuracy of 96% for load state identification and 89% for power decomposition of the load, and is able to meet the real-time application requirements.

INDEX TERMS Time-segmented state probability, load disaggregation, non-intrusive load monitoring, affinity propagation.

I. INTRODUCTION

The feedback to the power company on the customer’s electricity consumption can not only allow the customer to understand their electricity consumption habits but also help the power company to improve its supply strategy for bilateral reciprocity [1]. Current smart meters only provide total power data, while the electricity consumption of individual devices is not included in it, making it difficult to meet the two-way interaction between the customers and the power company. In addition, the monitoring of loads also helps in the detection of abnormal appliances and the identification of inefficient loads, which will contribute to power savings[2]. Among the methods of load monitoring, non-intrusive load monitoring (NILM) is widely used for appliance-level data acquisition due to its cost-effectiveness [3].

Hart at MIT first proposed NILM, which has received a lot of attention from researchers [3]. Compared to intrusive load monitoring, NILM does not require intrusion into the customer’s premises. It uses a monitoring device installed at the customer’s electricity entry point or relies directly on existing smart meters to obtain the customer's bus data. The bus data is then decomposed by a series of algorithms to obtain the usage of the various loads within the customer. Since NILM does not require the installation of sensors for each appliance, it has low economic and labor costs for installation and maintenance.

The NILM problem can be considered as a combination optimization (CO) problem. It uses the characteristics of each load to build a model, and then compares the total characteristics corresponding to the possible combinations of loads with the actual total characteristics to derive the operating state of the appliance. In recent years, NILM has gained a great deal of attention. Some scholars have provided publicly available datasets on this field, such as REDD [5], BLUED [6] and AMPds [7]. In addition, methods such as machine learning, neural networks and Hidden Markov Models are constantly being updated,
providing researchers with algorithmic tools for solving load identification problems. Based on the sampling frequency of the observation, NILM can be divided into two main methods. One is based on high-frequency data and one is based on low-frequency data.

The NILM methods based on high-frequency data use the characteristics of the load in the high-frequency data as the basis for load identification. Among them, literature [8] proposes to use current phase, amplitude and frequency as load features and then combine them with Long Short-Term Memory (LSTM) networks for load identification. Literature [9] uses voltage, active and reactive currents, Fourier transformed and coded to form a color signature, and then uses two-stream convolutional neural networks for load identification. Literature [10] proposes to use reconstructed voltage and current trajectory features and use convolutional neural networks for device classification. Although these methods based on the unique features have achieved good results in terms of recognition accuracy, the extraction of features requires the use of high-frequency data from the load. However, such a high sampling rate is not the standard feature for currently available smart meters. For example, according to the minimum functional specification of the Australian National Smart Metering Program, smart meters only provide power consumption data once every 5 seconds [11].

The NILM methods based on low-frequency data infer the most likely combination of states from the observed power. In [12], a factorial Hidden Markov Model is established to achieve NILM, which is then solved using the Viterbi decoding framework and integer quadratic constraint programming. A convolutional neural network-based energy decomposition method based on a multiple-input multiple-output regression model is proposed in [13], which considers the previous state of the device and the dependencies between devices for load decomposition. Graph signal processing (GSP) is used for NILM in [14]. The GSP-based method achieves load decomposition by minimizing the total variation of the graph as the starting condition and combining it with a simulated annealing algorithm. An improved cross-entropy algorithm is introduced in [15]. Under the constraint of the penalty function, it searches for the best combination of states by iteratively updating the operation probability of the equipment to achieve load decomposition. Generally, methods based on low-frequency data achieve load decomposition by finding the combination of states that makes the minimum error. The computational complexity of these methods tends to increase exponentially with the number of devices and the time of the calculation increases significantly. Therefore, it is beneficial to find fast and efficient solvers in order to suit the application conditions of existing grids.

Based on the above analysis, a NILM method based on the time-segmented state probability is proposed in this paper. The sample data is used to cluster the load power states, and the power template of the household load in various operating states is obtained. Then the operating probability of the load under each time period is calculated, which is used to reduce the solution space and generate the load state for power decomposition. Finally, the feasibility of the method is experimentally verified on the public data set AMPds, which has a sampling frequency similar to that of smart meters.

The main contributions of this paper are summarized as follows:

1. The time-segmented state probability is introduced to narrow the search set of the solution, which leads to improved recognition accuracy and efficiency.
2. AP clustering strategy is adopted to establish the power template of the load, which makes the power template more accurate.
3. The proposed NILM method can greatly reduce the time of load identification with limited data training. Meanwhile, the proposed method does not limit the number of load state changes over a single sampling cycle.

II. HOME LOAD POWER TEMPLATE

In order to correspond the operating state of the load to the power, it is assumed in this paper that all appliances are capable of being represented by a finite number of states. As proposed by [16], we focus on three types of loads, Type-I, Type-II and Type-VI.

A. POWER TEMPLATE CONSTRUCTION

Since a single feature is prone to feature overlap, we choose easily available steady-state active power and the steady-state reactive power as the load features. Assuming that there are \( M \) appliances, and the corresponding states of
appliance $m$ have $N$ modes, the templates of the appliance $m$ ($Te_m$) are established as:

$$Te_m = \{[p_{m1}^n, q_{m1}^n] \ldots [p_{mn}^n, q_{mn}^n] \ldots [p_{mN}^n, q_{mN}^n]\}$$

(1)

where $p_{mn}^n$ and $q_{mn}^n$ represent the active power and reactive power, respectively, when the appliance $m$ is operating at mode $n$.

### B. Power Template Acquisition Based on AP Clustering Algorithm

In order to reduce the influence of the human factor, we introduce the AP algorithm to cluster the power template. The AP algorithm was proposed by Brendan J. Frey and Delbert Dueck in 2007 [17]. It does not need to predetermine the number of categories. It directly treats each data point as a possible cluster center (exemplars) and clusters according to the similarity between objects and information transmission. The basic idea of the AP algorithm is to treat all data points as network nodes and find cluster sets by continuously transmitting information between nodes. The two kinds of information transmitted in the iterative process of the AP algorithm are responsibility and availability. Suppose the appliance $m$ has $L$ instances, then the similarity matrix $O=(o_{ik})_{L \times L}$ is defined through negative Euclidean distance. The $o_{ik}$ that means the similarity between the instance $i$ and the instance $k$ is formulated as follows:

$$o_{ik} = -\sqrt{(p_{mi} - p_{mk})^2 + (q_{mi} - q_{mk})^2}$$

(2)

where $p_{mi}$ and $q_{mi}$ represent the active and reactive power of appliance $m$ at the $i^{th}$ instance, respectively. The diagonal elements of the similarity matrix are generally set to the median or minimum value of the non-diagonal elements and defined as preference. The preference degree affects the number of exemplars obtained by the clustering result.

The responsibility and availability matrices are defined as $R=(r_{ik})_{L \times L}$, $A=(a_{ik})_{L \times L}$, respectively. The $r_{ik}$ that means the responsibility between the $i^{th}$ and the $k^{th}$ instances is formulated as follows:

$$r_{ik} = o_{ik} - \max_{s \neq k} (a_{is} + o_{is})$$

(3)

The $a_{ik}$ that means the availability between the $i^{th}$ and the $k^{th}$ instances is formulated as follows:

$$a_{ik} = \min\{0, r_{ik} + \sum_{j \neq k} \max(0, r_{ij})\}$$

(4)

The update formulat for $A$ is:

$$A_{i+1} = \alpha A_{i} + (1 - \alpha)A_{i-1}$$

(5)

where $A_{i-1}$ and $R_{i-1}$ are the output of the previous iteration, $A_{i}$ and $R_{i}$ are the output of this iteration.

Step 2: Calculate $R$ and $A$ successively by (3) and (4). In order to suppress data oscillations during the iteration process, the damping coefficient $\alpha$ is added in each iteration. The update formula is as follows:

$$R_{i+1} = \alpha R_{i} + (1 - \alpha)R_{i-1}$$

Step 3: Repeat step 2 until the matrices are stable or the maximum number of iterations is reached.

Step 4: Then, the instance $e$ that satisfies the following condition is the exemplar:

$$A(e,e) + R(e,e) > 0$$

(6)

Step 5: Calculate which exemplar each instance belongs to.

Step 6: Determine whether the number of instances belonging to each exemplar meets more than a predefined threshold. If there are, remove the exemplar and return to step 5 to recalculate the attribution of each point; otherwise, output the exemplar.

### III. THE NILM METHOD BASED ON THE TIME-SEGMENTED STATE PROBABILITY

#### A. THE TIME-SEGMENTED STATE PROBABILITY

Different appliance has different electricity consumption rules, so the probability of each appliance in various operating states is different. At the same time, combined with the characteristics of household electricity consumption habits, the operating state of the household load reflects a unique law within 24 hours. Taking the four loads in AMPs as an example, the active and reactive power dispersions are drawn for a typical 24h period, as shown in Fig.2. It can be seen from Fig.2 that different home appliances exhibit different distributions over time due to different functions and the characteristics of user behaviour habits. For example, as a device for maintaining indoor temperature, the heat pump frequently runs at night when the temperature is low to maintain a suitable indoor temperature. Related entertainment appliances such as TV are mostly used before dinner and before going to bed, while dishwashers are generally used at dinner time.

Step 1: Initialize $R$ and $A$ as all-zero matrices, and calculate the similarity matrix $O$ by (2).
Effective division of the load's typical state behaviour time period can improve the effect of load disaggregation. Taking into account the general residential electricity consumption behaviour, the division of the load operation time period in this paper is shown in Fig. 3.

According to the life pattern of the general family, 1:00 to 5:00 is bedtime, the load usage is simple, and the load changes less, classified as a section. Except for this section, dividing the other time periods into two-hour intervals is reasonable. This division will not make the section too complicated to process, and it can also fully reflect the difference in the load operation status of different sections. For example, 7:00-9:00, 11:00-13:00, and 17:00-19:00 are generally mealtime, and the operation of the cooking-related appliance is very different from other time periods. 19:00-21:00 is generally in the relaxation time after a meal, and the entertainment load usage is quite different from the rest of the time period.

We assume that the training set contains $D$ days, the load category is $M$, and data for each load is sampled once a minute. Then the operating probability factors of the load appliance in different time periods can be expressed as follows:

$$f_{i,n}^m = \frac{T_{i,n}^m}{D \cdot T_i} \times 100\%$$

where $f_{i,n}^m$ represents the probability that the appliance $m$ is in the state $n$ during $t_i$; $T_{i,n}^m$ represents the total number of minutes that the device $m$ is in the state $n$ during $t_i$ in the entire training set, and $T_i$ is the length of the time period $t_i$.

As shown in Fig. 3, except that the $t_1$ time period is 240 minutes, each of the remaining segments is 120 minutes. For load $m$, there should be one and only one state in $t_i$, which satisfies formula (15):

$$\sum_{n=1}^{N} f_{i,n}^m = 1$$

Then the time-segmented probabilities of each load are stored as 11 probability matrices according to time segments so that they can be easily called. Each matrix $F_i$ is represented as follows:

$$F_i = \begin{bmatrix}
    f_{i,1}^1 & \cdots & f_{i,1}^M \\
    \vdots & \ddots & \vdots \\
    f_{i,1}^N & \cdots & f_{i,M}^N
  \end{bmatrix}$$

B. NILM MATHEMATICAL MODEL

The essence of the NILM problem is to decompose the total power consumption collected from the user's power inlet into a combination of various appliance power consumption. Assuming that the total active power recorded at a certain time is $P_{\text{total}}$ and the total reactive power is $Q_{\text{total}}$, they can be expressed by the sum of the power of each load:

$$P_{\text{total}} = \sum_{m=1}^{M} \sum_{n=1}^{N} s_m^n \cdot p_m^n + p_{\text{noise}}$$

$$Q_{\text{total}} = \sum_{m=1}^{M} \sum_{n=1}^{N} s_m^n \cdot q_m^n + q_{\text{noise}}$$

where $p_{\text{noise}}$ and $q_{\text{noise}}$ represent background noise; $s_m^n$ represents the state of the appliance $m$, and the value is 0 or 1. When the appliance $m$ is running in mode $n$, the $s_m^n$ is 1; otherwise, it is 0. The appliance state matrix $S$ can be expressed as:

$$S = (s_1 \cdots s_m \cdots s_N)$$

$$s_m^n = [s_1^n \cdots s_m^n \cdots s_N^n]^T$$

where $N$ represents the maximum possible number of modes among all $M$ types of appliances. Since any load has only one
The power template obtained in the previous section is represented as:
\[
P = (p_1 \cdots p_m \cdots p_M)
\]
\[
Q = (q_1 \cdots q_m \cdots q_M)
\]
where \( p_m = [p_{m1} \cdots p_{mn} \cdots p_{Mn}]^T \) and \( q_m = [q_{m1} \cdots q_{mn} \cdots q_{Mn}]^T \) represent the active and reactive power templates of the load \( m \), respectively.

The purpose of NILM algorithm is to disaggregate the measured signal \( P_{total} \) into its various components \( p_m \). Based on these definitions above, the optimization problem can be formulated as:
\[
\hat{S} = \arg \min_{S_1 \cdots S_M} \left\{ |\Delta P| + (1 - \lambda) |\Delta Q| \right\}
\]
\[
\Delta P = P_{total} - \sum_{m=1}^{M} (\hat{S}_m^T p_m)
\]
\[
\Delta Q = Q_{total} - \sum_{m=1}^{M} (\hat{S}_m^T q_m)
\]
where \( \lambda \) represents the weighting factor of power. The number of load status types in \( S, P, \) and \( Q \) are all marked as \( N \). When the load type of \( m \) is less than \( N \), the excess is made up by 0.

C. NILM BASED ON THE TIME-SEGMENTED STATE PROBABILITY

We regard load disaggregation as integer programming, and all possible solutions have been determined when the number and state of the appliance are determined. The proposed NILM algorithm is based on the time-segmented state probability (TSSP). It uses the probabilities to generate load state combinations. The purpose is to reduce the range of the search set and accelerate the speed and accuracy of recognition. The optimal combination of states for the load is obtained by Algorithm 1.

In the initialization phase, the time-stamped aggregated data are input and the TSSP matrix \( F_t \) is selected based on the time. In the main loop, set a random value \( \hat{s}_m \) for each appliance \( m \) in the solution \( \hat{S} \) according to the probability \( f_{i,m}^{n} \), which is the probability that appliance \( m \) will operate as mode \( n \) in time segment \( t \). For the state estimation matrix \( \hat{S} \), an appropriate number is chosen, taking into account both time cost and accuracy. Finally, the optimal state estimation matrix is filtered by the function (12) and (13).

Algorithm 1 The NILM Algorithm Based on the TSSP

**Inputs:**
\( P_{total}, Q_{total}, P, Q, \) real-time \( t, Th \) (the number of preset estimation matrices)

**Procedure:**
- Select \( F_t \) from time \( t \)
- Init: \( j = 1 \)
- While \( j \leq Th \)
  - For each load \( m (m=1,2,\ldots M) \)
    - Generate a random number \( Y \) between 0 and 1
    - If \( Y \leq f^1_{i,m} \)
      - \( \hat{s}_m = [1,0,\ldots,0]^T \)
    - If \( f^1_{i,m} < Y \leq (f^1_{i,m} + f^2_{i,m}) \)
      - \( \hat{s}_m = [0,1,\ldots,0]^T \)
    - ...... 
    - If \( \sum_{n=1}^{N-1} f^n_{i,m} < Y \leq 1 \)
      - \( \hat{s}_m = [0,0,\ldots,1]^T \)
  - End
  - Select the best \( \hat{S} \) by (12) and (13)

**Output:** \( \hat{S} \)

The steps of the overall proposed approach for NILM are as follows:

Step 1: Use the data in the training set and the AP algorithm to obtain the appliance's power template matrix \( P \) and \( Q \).

Step 2: The data in the training set is divided into sub-datasets according to time segments, and the status of the appliance is tagged according to the power template obtained in step 1.

Step 3: Count the TSSP for each appliance and save the probability matrix of each time segment.

Step 4: Time-label each data set in the testing set (Existing smart meters often come with a time scale).

Step 5: Solve the state combination of the load using Algorithm 1.

Unlike [18], the proposed algorithm does not require information from the previous moment. Furthermore, it does not limit the number of state changes between two adjacent moments.

IV. SIMULATION EXPERIMENT AND ANALYSIS

A. EVALUATION METRICS

Many different kinds of metrics exist in the study of NILM algorithms. In this paper, the NILM problem is classified as integer programming, and the load state is compared with a binary classification task. There are four classification results, namely: 1) the number of times an appliance was correctly identified as turned on (TP); 2) the number of times an appliance was correctly identified as turned off (TN); 3) the number of times an appliance was incorrectly identified as turned on (FP); 4) The number of times that a device was wrongly identified as closed (FN). Based on the above four recognition results, the following metrics can be used to evaluate the performance of NILM algorithms:
\[ S_{ac} = \frac{TP + TN}{TP + TN + FP + FN} \]  
\[ FM = \frac{2 \cdot Pre \cdot Re}{Pre + Re} \]

where \( FM \) represents F-measure, which indicates the harmonic average of recall rate \( Re \) and precision rate \( Pre \) [19]. \( Re \) and \( Pre \) are calculated as follows:

\[ Pre = \frac{TP}{TP + FP}, \quad Re = \frac{TP}{TP + FN} \]

where \( Pre \) is defined as the precision classification in all positive estimations; \( Re \) is expressed as the percentage of correctly recognized appliances. In addition, the method mentioned in [20] is also used to evaluate the accuracy of power decomposition. For \( T \) samples, the total power decomposition accuracy is defined as \( PD_{acv} \): the power decomposition accuracy for each load is defined as \( PD_{(m)acv} \):

\[ PD_{acv} = 1 - \frac{\sum_{t=1}^{T} \sum_{m=1}^{M} |\hat{p}_m^t - p_m^t|}{2 \cdot \sum_{t=1}^{T} P^t} \]  
\[ PD_{(m)acv} = 1 - \frac{\sum_{t=1}^{T} |\hat{p}_m^t - p_m^t|}{2 \cdot \sum_{t=1}^{T} p_m^t} \]

where \( \hat{p}_m^t \) represents the active power estimated by the NILM method, \( p_m^t \) is the real active power of appliance \( m \) at \( t \), and \( P^t \) is the total power at \( t \).

## B. SIMULATION RESULTS ON AMPDS

Our research work is to realize the identification and intelligent control of user loadshed for the demand-oriented response. In order to reduce the additional economic burden, the data available on the smart meter is used for identification. For this reason, AMPds2 is chosen to test the proposed method. AMPds2 records 2 years of electricity consumption data for a Canadian household, including electrical characteristics such as voltage, current, active and reactive power, collected at 1-minute intervals. We choose eight kinds of loads as test objects: basement plugs & lights (BME), dishwasher (DWE), clothes washer (CWE), entrainment TV/PVR/AMP (TVE), kitchen fridge (FGE), heat pump (HPE), HVAC furnace fan & thermostat (FRE), kitchen well oven (WOE). Three test cases will be carried out, and all three cases use the first two weeks of data to obtain TSSP. The first case (CASE I) decomposes the data of eight appliances for 30 days to verify the algorithm of this paper. The second case (CASE II) adds to the first one the data obtained from the actual detection of electric vehicle (EV) and is used to test the effectiveness of our method for the recognition of electric vehicles. The third case (CASE III) chooses the literature [21] as a reference to compare the accuracy of the proposed method in terms of power decomposition. The fourth case (CASE IV) chooses the literature [22] as a reference to compare the accuracy of the algorithm in terms of load states identification.

### 1) TEST CASE I

The training set is the active and reactive power of eight appliances between April 1, 2012, 0:00 to April 14, 2012, 23:59. In order to avoid too much data used for clustering and to improve the efficiency of clustering, a sampling method is adopted for clustering. The training set data is classified by time segments, and the total data is divided into eleven sub-datasets \([r_1, r_2...r_{11}]\). When solving the power template of the appliance, according to the operating conditions of different equipment in different periods, a sub-data set containing more equipment states is selected as the clustering data. After preliminary screening, BME and CDE select \( r_9, r_{10} \), and \( r_{11} \) data sets, FGE and FRE select \( r_8, r_9 \), and \( r_{10} \) data sets, DWE selects \( r_7 \), \( r_8 \), \( r_9 \), and \( r_{10} \) data sets, HPE selects \( r_1, r_2, r_3 \) data sets, TVE selects \( r_7, r_8, r_9, r_{10} \), and \( r_{11} \) data sets, WOE selects \( r_7, r_8 \), and \( r_9 \) data sets. The damping coefficient in (5) is 0.5, and the maximum number of iterations is 200. After the AP algorithm, the power templates of each appliance are shown in Table I.

| Appliance | modes | Power template |
|-----------|-------|----------------|
| BME       | 3     | [8,7], [335,26], [432,29] |
| CDE       | 3     | [0.0], [242,405], [446,406] |
| DWE       | 3     | [21,1], [139,35], [719,30] |
| FRE       | 2     | [108,26], [363,7] |
| FGE       | 2     | [1.6], [137,11] |
| HPE       | 4     | [0.9], [35,15], [1738,321], [1810,327] |
| TVE       | 4     | [39,13], [259,32], [262,37], [374,32] |
| WOE       | 3     | [1.0], [2650,142], [3402,135] |

There is also power consumption when the appliance is on standby. For example, the BME always has 8W active power and 7var reactive power on standby, so it is considered a state. Through the power template of Table I, the samples in the training set are labelled, and the TSSP is counted. The probabilities are shown in Fig.4.
FIGURE 4. The time-segmented state probability of training set samples.

The testing set selects the data from April 15, 2012, 0:00 to May 14, 2021, 23:59 for 30 days. For each moment, we use the probability distribution to generate 100, 150, and 200 state estimation matrices, calculate their F-measure. We calculate the state accuracy rate $S_{\text{acc}}$ when the number of state estimation matrices is 150. The power weighting factor $\lambda$ in (12) is 0.6. The recognition results are shown in Table II and Table III.

TABLE II
CASE I: STATE RECOGNITION ACCURACY AND F-MEASURE.

| Appliance | $FM^{100\%}$ (%) | $FM^{150\%}$ (%) | $FM^{200\%}$ (%) | $S_{\text{acc}}$ (%) |
|-----------|------------------|------------------|------------------|------------------|
| BME       | 96.16            | 95.98            | 96.03            | 97.32            |
| CDE       | 99.24            | 99.34            | 99.34            | 99.56            |
| DWE       | 97.89            | 97.94            | 97.91            | 98.63            |
| FRE       | 99.82            | 99.82            | 99.83            | 99.82            |
| FGE       | 92.89            | 93.39            | 93.74            | 93.39            |
| HPE       | 68.99            | 69.97            | 68.25            | 84.49            |
| TVE       | 92.72            | 91.97            | 92.14            | 95.98            |
| WOE       | 99.72            | 99.77            | 99.77            | 99.85            |
| Average   | 93.43            | 93.40            | 93.38            | 96.13            |

TABLE III
CASE I: POWER DECOMPOSITION ACCURACY.

| Appliance | $PD_{\text{acc}}^{100\%}$ (%) | $PD_{\text{acc}}^{150\%}$ (%) | $PD_{\text{acc}}^{200\%}$ (%) |
|-----------|-------------------------------|-------------------------------|-------------------------------|
| BME       | 74.73                         | 75.07                         | 74.36                         |
| CDE       | 78.44                         | 79.32                         | 80.66                         |
| DWE       | 49.24                         | 50.6                          | 50.9                          |
| FRE       | 97.05                         | 97.04                         | 97.05                         |

The results show that HPE has the lowest score of the $FM$ parameter. On the one hand, HPE has four modes, and it is easier to identify errors than the two-state appliance. On the other hand, in the probability distribution diagram of the testing set in Fig.3(f), HPE may be in operation at any time of the day, increasing the possibility of incorrect judgments. In addition, the $PD_{\text{acc}}$ is mainly affected by the $S_{\text{acc}}$ and the appliance power template. Due to the complex operation of the appliance, the actual power of the load is not based on the ideal power template. Although the AP algorithm is as accurate as possible for the power template of the appliance, the $PD_{\text{acc}}$ is below 90%.

As can be seen from Table I, for the eight appliances of CASE I, there are a total of 5184 possible state combinations. However, the range of solution sets is narrowed down by restricting the TSSP at different time segments so that satisfactory results can be obtained when the estimation matrix is around 150.

In order to observe more details of the estimated decomposition results, the decomposition results are employed for the two days from May 7, 12:00 to May 9, 12:00 in CASE I, which is plotted as Fig.5.
As can be seen from Fig.5, around the 500th minute the decomposition results for CDE, DWE, TVE and HPE appear to be abnormal. This is because all these appliances may be running at this time. And the CDE’s third mode was falsely recognized as a combination of the other appliances. In the time range 700-1100 and 2300-2700 in Fig.5, corresponding to time segments t1, t2 and t3, the use of TSSP allows for the limitation of incorrect load combinations and higher accuracy of load state identification. A better overlap between the real power and the estimated power is shown in the graph.

The method succeeds in preventing the mismatch during most periods. However, there are still cases of mismatching during periods with complex load operating conditions, such as t9.

2) TEST CASE II: WITH EV

Due to the popularity of EV, more and more users are charging their EVs at home. Considering this situation, we added EV to the conventional load to test the accuracy of our algorithm in recognizing EV. The EV charging data obtained from the actual test is shown in Fig.6.

Since only the active power of this EV is available, we add the EV load to CASE I and consider only the active power of the appliance as a template. In this case, the state estimation matrix is 150, and only two states are considered for EV: ON (3825W) and OFF (0W).

In the probability matrix of CASE I, we supplement the TSSP for fourteen days of EV. The results obtained from the decomposition of the data for one week during the testing phase are shown in Table IV.

| Appliance | PDacc (%) | Sacc (%) |
|-----------|-----------|----------|
| BME       | 74.46     | 92.68    |
| CDE       | 73.46     | 98.76    |
| DWE       | 47.66     | 95.53    |
| FRE       | 96.93     | 99.87    |
| FGE       | 82.04     | 87.65    |
| HPE       | 88.11     | 92.55    |
| TVE       | 82.87     | 92.47    |
| WOE       | 62.74     | 99.53    |
| EV        | 97.10     | 99.85    |
| **Average** | **87.81** | **95.43** |

From these results, it was observed that, with the proposed method, the average appliance state identification accuracy is more than 95%, and the power decomposition accuracy is more than 87%. In addition, the state identification accuracy for EV exceeds 99% and the power decomposition accuracy reaches 97.1% due to the obvious charging pattern of EV. This suggests that the proposed algorithm is equally applicable to households with EVs.

3) TEST CASE III
CASE III is designed to have the same setups as in [21], which proposes a NILM algorithm based on mixed-integer linear programming (MILP). This case considers a shorter period of one week. The proposed method is also compared with the CO algorithm mentioned in the literature [21]. In order to avoid the effect of power templates, we choose the power template from Table II in [21] and recalculate the TSSP using the data from the first two weeks. In this case, the state estimation matrix is 150, and the load power decomposition accuracy is shown in Table V.

| Appliance | CO(%)  | MILP(%)  | This work(%) |
|-----------|--------|----------|--------------|
| BME       | 81.6%  | 88.9%    | 78.73%       |
| CDE       | 98.6%  | 93.9%    | 96.59%       |
| DWE       | 45.35% | 72.65%   | 64.32%       |
| FRE       | 86.35% | 93.9%    | 95.85%       |
| FGE       | 58.4%  | 80.05%   | 86.34%       |
| HPE       | 97%    | 98%      | 96.54%       |
| TVE       | 52.9%  | 85.2%    | 89.91%       |
| **ALL**   | 74.31% | 87.51%   | 91.45%       |

The optimization objectives of the model proposed in this paper are the same as those of CO and MILP, but both CO and MILP add windows to the signals and solve them within a single window. On the other hand, the proposed model is solved separately at each moment. In this way, even if the identification results of the previous moments are abnormal, they do not affect the subsequent load power decomposition.

Table V shows the most problematic appliances are DWE and BME. This seems to be mainly due to the overlap of the periods of BME, DWE and TVE, and the DWE power values are close to the combination of BME and TVE. Nonetheless, the power decomposition accuracy of the proposed method is still higher than CO and MILP.

4) TEST CASE IV

We select scenario 3 from [22] as the comparison object in this case. The modified cross-entropy (MCE) algorithm uses the previous moment's state and adds a penalty function to achieve the load decomposition in [22]. In contrast to MCE, our model has no additional penalty function. All constraints are embodied through the TSSP and will only come into play during the solution of the optimisation problem. The training set uses data for the two weeks starting April 1. The power templates are the same as [22], containing only active power. The testing set selection starts at 12:18:00 on April 18, 2012, as does the MCE, for a total of 600 hours. The number of the generated matrices is set to 150, and the decomposition results are as follows:

| Appliance | CO(%)  | MILP(%)  | This work(%) |
|-----------|--------|----------|--------------|
| BME       | 97.88/86.24 | 96.50/79.67 |
| CDE       | 93.86/29.28  | 97.02/23.03 |
| DWE       | 98.84/72.38  | 98.84/61.05 |
| FRE       | 82.70/93.16  | 99.98/99.99 |
| FGE       | 92.26/88.27  | 94.67/92.41 |

As can be seen from Table VI and Table VII, compared with MCE, the $S_{acc}$ and $PD_{acc}$ of this paper have improved significantly. Although the proposed method needs to use historical data to obtain the TSSP, it is more accurate than MCE for load recognition. The proposed method does not require iterations and is less time-consuming.

C. PRACTICAL DISCUSSION

Power line communications (PLC) network has been considered as the favourite choice for communicating with electricity meters. In a typical PLC network, smart meters are connected to the data concentrator [23]. A single data concentrator is usually responsible for remote reading of up to 200-500 smart meters. We want to deploy load recognition algorithms in the concentrator, which would be more cost-effective and less invasive of customer privacy than embedding the algorithm at each smart meter. The complexity of the proposed algorithm is $O(M \cdot Th)$, where $M$ is the number of appliances and $Th$ is the number of estimation matrices in algorithm 1. The simulation results as presented in this paper were attained by MATLAB, and the hardware platform is Intel(R) Core(TM) i5-7300HQ CPU @ 2.50GHz 2.50 GHz, 8.00GB RAM notebook computer. The time required for the proposed algorithm to calculate a month's data with a sampling frequency of 1/60Hz is shown in Table VII.

| Number of matrices | 100     | 150     | 200     |
|--------------------|---------|---------|---------|
| CASE I             | 84.57s  | 138.81s | 171.12s |

For a concentrator with 500 smart meters connected, a complete reading cycle is 500s, which means that every second meter's data is transferred to the concentrator[24]. To achieve real-time load disaggregation in the concentrator, and take into account transmission times and other functional time consumption, the algorithm needs to be completed within 0.5s. As can be seen in Table VIII, the time of our algorithm satisfies this requirement. Therefore, our scheme has strong implementability under the existing electricity information collection systems.

D. METHODOLOGY ANALYSIS

Through testing on publicly available datasets and comparative discussions with other methods, our approach has the following advantages: (1) The model is simple, the probabilities and power templates are stored as matrices and can be called up quickly. (2) The use of TSSP limits the
possible combinations of states of the load and reduces the search space. (3) The short computation time and low-rate sampling required for our approach allow deployment in a concentrator, thus reducing economic costs and accommodating the consumer’s privacy and comfort.

However, there are some weaknesses in the proposed approach: (1) The construction of power templates assumes that the load is a finite state machine, so it is poor at identifying appliances with continuously adjustable power during operation. (2) Both the templates and the TSSP acquisition of the appliances require historical data for training. (3) The method may be less effective in identifying when several appliances have similar operating times and similar operating power.

V. CONCLUSIONS

In this paper, a new NILM method is proposed based on the TSSP. The algorithm is easy to implement and consumes a little time. On the one hand, the AP algorithm is introduced to reduce the human interference of load power templates and make the power template more reliable. On the other hand, the time-segmented state probability is proposed in this paper, which reflects the intrinsic connection between the appliance state and time, thus reducing the search space. At the same time, the algorithm considers multi-state appliances and does not limit the number of state-changing devices in a single sampling period.

The proposed algorithm shows good performance when tested on the AMPds dataset. When the number of appliances is eight, for one month's data, the accuracy of load state identification reaches over 96% and the accuracy of the power decomposition is over 89%. Meanwhile, the results show that our method has a significant accuracy advantage over MCE and MILP disaggregation methods. In fact, because of the short time used, our method is suitable to be arranged in a concentrator.

However, the power template acquisition and the time-segmented state probability of appliances need to be trained in advance. Next, we will explore the relationship between appliance operation status and time-segmented probability distribution among different users to simplify the acquisition of prior conditions.

REFERENCES

[1] S. Houde, A. Todd, A. Sudarshan, J. A. Flora, and K. C. Armel, “Real-time Feedback and Electricity Consumption: A Field Experiment Assessing the Potential for Savings and Persistence,” Energy Journal, vol. 34, no. 1, pp. 87-102, 2013.

[2] E. Azizi, M. T. H. Beheshti and S. Bolouki, “Appliance-Level Anomaly Detection in Nonintrusive Load Monitoring via Power Consumption-Based Feature Analysis,” IEEE Transactions on Consumer Electronics, vol. 67, no. 4, pp. 363-371, Nov. 2021.

[3] A. Marchiori, D. Hakkarinen, Q. Han and L. Earle, “Circuit-Level Load Monitoring for Household Energy Management,” IEEE Pervasive Computing, vol. 10, no. 1, pp. 40-48, Jan.-March 2011.

[4] G. W. Hart, “Non-intrusive Appliance Load Monitoring,” Proceedings of the IEEE, vol. 80, no. 12, pp. 1870-1891, Dec 1992.

[5] J. Z. Kolter and M. J. Johnson, “REDD: A public data set for energy disaggregation research,” in Proc. SusKDD Workshop Data Min. Appl. Sustain. San Diego, CA, USA, Aug. 2011, pp. 1-6.

[6] K. D. Anderson, A. Ocneanu, D. R. Carlson, A. G. Rowe, and M. Bergés, “BLUED: A Fully Labeled Public Dataset for Event-Based Nonintrusive Load Monitoring Research,” Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability, Beijing, China: ACM, 2012, pp. 1-5.

[7] S. Makonin, F. Popowich, L. Bartram, B. Gill and I. V. Bajić, “AMPds: A Public Dataset for Load Disaggregation and Eco-Feedback Research,” 2013 IEEE Electrical Power & Energy Conference, 2013, pp. 1-6.

[8] T. -T. H. Le, S. Heo and H. Kim, “Toward Load Identification Based on the Hilbert Transform and Sequence to Sequence Long Short-Term Memory,” IEEE Transactions on Smart Grid, vol. 12, no. 4, pp. 3252-3264, July 2021.

[9] J. Chen, X. Wang, X. Zhang and W. Zhang, “Temporal and Spectral Feature Learning With Two-Stream Convolutional Neural Networks for Appliance Recognition in NILM,” IEEE Transactions on Smart Grid, vol. 13, no. 1, pp. 762-772, Jan. 2022.

[10] D. Jia, Y. Li, Z. Du, J. Xu and B. Yin, “Non-Intrusive Load Identification Using Reconstructed Voltage–Current Images,” IEEE Access, vol. 9, pp. 77349-77358, 2021.

[11] W. Kong, Z. Y. Dong, D. J. Hill, F. Luo and Y. Xu, “Improving Nonintrusive Load Monitoring Efficiency via a Hybrid Programing Method,” IEEE Transactions on Industrial Informatics, vol. 12, no. 6, pp. 2148-2157, Dec. 2016.

[12] W. Kong, Z. Y. Dong, J. Ma, D. J. Hill, J. Zhao and F. Luo, “An Extensible Approach for Non-Intrusive Load Disaggregation With Smart Meter Data,” IEEE Transactions on Smart Grid, vol. 9, no. 4, pp. 3362-3372, July 2018.

[13] M. Kaselimi, E. Protopapadakis, A. Voulodimos, N. Doulamis and A. Doulamis, “Multi-Channel Recurrent Convolutional Neural Networks for Energy Disaggregation,” IEEE Access, vol. 7, pp. 81047-81056, 2019.

[14] R. He, L. Stankovic, J. Liao and V. Stankovic, “Non-Intrusive Load Disaggregation Using Graph Signal Processing,” IEEE Transactions on Smart Grid, vol. 9, no. 3, pp. 1739-1747, May 2018.

[15] R. Machlev, Y. Levron and Y. Beck, “Modified Cross-Entropy Method for Classification of Events in NILM Systems,” IEEE Transactions on Smart Grid, vol. 10, no. 5, pp. 4962-4973, Sept. 2019.

[16] A. Zoha, A. Ghuak, M. A. Imran, and S. Rajasegarar, “Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey,” Sensors, vol. 12, no. 12, pp. 16838-16866, Dec 2012.

[17] B. J. Frey and D. Dueck, “Clustering by passing messages between data points,” Science, vol. 315, no. 5814, pp. 972-976, 2007.

[18] M. J. Zhong, N. Goddard, and C. Sutton, “Signal Aggregate Constraints in Additive Factorial HMMs, with Application to Energy Disaggregation,” in 28th Conference on Neural Information Processing Systems (NIPS), Montreal, CANADA, 2014, vol. 27, 2014.

[19] G. Hripcsak and A. S. Rothschild, “Agreement, the F-measure, and reliability in information retrieval,” Journal of the American Medical Informatics Association, vol. 12, no. 3, pp. 296-298, 2005.

[20] M. Z. A. Bhoto, S. Makonin, and I. V. Bajić, “Load Disaggregation Based on Aided Linear Integer Programming,” IEEE Transactions on Circuits and Systems II-Express Briefs, vol. 64, no. 7, pp. 792-796, 2017.

[21] F. M. Wittmann, J. C. López and M. J. Rider, “Non-Intrusive Load Monitoring Algorithm Using Mixed-Integer Linear Programming,” IEEE Transactions on Consumer Electronics, vol. 64, no. 2, pp. 180-187, May 2018.

[22] R. Machlev, Y. Levron and Y. Beck, “Modified Cross-Entropy Method for Classification of Events in NILM Systems,” IEEE Transactions on Smart Grid, vol. 10, no. 5, pp. 4962-4973, Sept. 2019.

[23] R. P. Lewis, P. Igiot and Z. Zhou, “Assessment of communication methods for smart electricity metering in the U.K.,” 2009 IEEE
[24] Y. Ben-Shimol, S. Greenberg and K. Danilchenko, “Application-Layer Approach for Efficient Smart Meter Reading in Low-Voltage PLC Networks,” IEEE Transactions on Communications, vol. 66, no. 9, pp. 4249-4258, Sept. 2018.

LINA LIU received the B.S. degree from the Department of Power System and Automation and the Ph.D. degree from the Department of Electrical Theory and New Technology Professionals, Southwest Jiaotong University, Chengdu, China, in 2004 and 2014, respectively. She is currently working in the Metering Center of Sichuan Electric Power Company. Her research interests are in electric energy metering, smart metering equipment testing technology, and electromagnetic compatibility testing technology.

FANGSHUO LI received the B.S. degree from the Department of Electrical engineering and automation, Nanjing University of Science and Technology, Nanjing, China, in 2008, and the M.S. degree from the Department of Communication and Information System, University of Electronic Science and Technology of China, Chengdu, China, in 2011. He is currently working in the Metering Center of Sichuan Electric Power Company. His research interests are in electric power communication and electric power metering.

ZHIJIONG CHENG received the B.S. degree from the Department of Computer Engineering, University of Electronic Science and Technology of China, China, in 1986, and the M.S. degree from the Department of Project Management Professionals, North China Electric Power University, China, in 2009. He is currently working in the Metering Center of Sichuan Electric Power Company. His research interests include electric energy measurement collection technology and smart metering equipment detection technology.

YIFEI ZHOU received the B.S. degree from the Department of Electrical Engineering and Automation and the M.S. degree from the Department of Electrical Engineering, Southwest Jiaotong University, China, in 2004 and 2016. He is currently working in the Metering Center of Sichuan Electric Power Company. His research interests are in electrical instrumentation, electric energy measurement and its detection technology.

JIE SHEN received the B.E. and the M.E. degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 2006 and 2010, respectively. He is currently working in the Metering Center of Sichuan Electric Power Company. His research interests are in electricity metering and electricity consumption information collection.

RUICHAO LI received the B.S. degree from the Department of Electrical Engineering and Information, Southwest Petroleum University, China, in 2014, and the M.S. degree from the Department of Electrical Engineering, Southwest Jiaotong University, China, in 2016, respectively. He is currently working in the Metering Center of Sichuan Electric Power Company. His research interests are in electrical instrumentation, electric energy measurement and its detection technology.

SIYU XIONG (S’20) received the B.S. degree from the Department of Electrical Engineering and Information, Southwest Petroleum University, China, in 2014, and the M.S. degree from the Department of Electrical Engineering, Southwest Jiaotong University, China, in 2018, where he is currently pursuing the Doctoral degree. His research interests are in power system stability and control, and signal processing and information theory in electrical power system applications.