An online non-intrusive load monitoring method based on
Hidden Markov model

Xianqing Huang¹, Bo Yin¹, Zhiqiang Wei¹*, Xinghao Wei¹, Rui Zhang²

¹College of information science and engineering, Ocean University of China, Qingdao, China
²Qingdao Haier Smart Technology R&D Co., Ltd., Qingdao, China

*Corresponding author e-mail: seiouc16@163.com

Abstract. Non-intrusive load monitoring (NILM) can decompose the total power consumption measured by the smart meter into the power consumed by the individual appliances, so as to achieve the purpose of saving energy. In this paper, an improved method of Daubechies9 (DB9) which is a discrete wavelet is proposed, which can effectively remove the noise of the low-frequency components. On this basis, an online NILM method based on Hidden Markov model (HMM) is proposed. The model of load switching can be built using apparent power of transient-state with this method. Besides, the improved forward algorithm which effectively suppressing the data underflow in load classification is proposed. The proposed methods are embedded in the smart meter and can increase the overall recognition rate of the load over 90% in the experiments which prove that they have good applicability.

1. Introduction

The construction and development of the global energy Internet put forward higher requirements for smart grid. Smart power which connects the grid side and the user side is an important part of the smart grid interactive service system [1-2]. Electrical load monitoring is the first step to realize intelligent power consumption. By sampling and analyzing the total power consumption, it monitors the detailed running state of individual appliances to obtain the detailed data information, such as the power consumption and the behavior of power user [3]. Refined load monitoring data provide more energy saving service innovation projects for the energy efficiency market and provide data support for mining user behavior information. Energy efficiency management services of smart power also help to achieve energy saving, emission reduction and sustainable development in the whole society.

Through the decomposition and recognition of the user's total power consumption, INLM can obtain fine internal load classification and the power consumed by the individual appliances. It is an effective way to solve the problem of load monitoring in smart power. With the increasing number of
users, more and more complex load types, and the growing trend of the more and more sophisticated use of electricity service, the key technology of NILM is studied to solve the core technical problems of data acquisition, feature extraction and load identification for smart power. It can reduce the cost of load monitoring, improve the accuracy of monitoring and identification of the smart meter, promote the depth mining and application of load monitoring, and provide technical support for the realization of flexible and interactive for smart power.

In this paper, we propose an improved DB9 algorithm for data preprocessing, an online NILM method based on HMM and an improved forward classification method for the recognition of the load. The effectiveness of the method is verified by experiments. The rest of this paper is organized as follows. The related work is introduced in section 2. Section 3 describes the improved the DB9 algorithm, the online NILM and the improved forward algorithm. Section 4 gives the experimental results and analysis. The summary is provided in section 5.

2. Related Work
Since the introduction of NILM technology, it has been widely concerned by power companies and research institutes in various countries. In 1992, Professor Hart established the first NILM system for household appliances [4]. A finite state load model which is suitable for double state and polymorphic load is established by analyzing the power flow. The model can accurately identify load under certain conditions, but the performance of the algorithm is greatly affected in the presence of interference such as noise. Tsai proposed NIALM system in [5], which integrates feature extraction and feature optimization methods with pattern recognition techniques, is used to identify the operation status of individual appliances. The method extracts the time-domain characteristics of transient currents, and then classifies the determined eigenvalues. The recognition effect of this method is very serious due to voltage fluctuation and noise interference. K.Basu extracts the characteristics of temporal energy change of total load power over a period of time in [6]. The NILM is implemented by decision tree, support vector machine, hidden Markov chain and K nearest neighbor method. The accuracy of electrical identification depends on the trend of characteristic instead of the eigenvalue. Therefore, this method is seldom affected by voltage fluctuation and noise interference. However, the monitoring which based on the low-frequency sampling rate, or off-line data, is lagging for the classification. And the deviation of energy consumption is not negligible. The same problem occurs in [7]. Aiad considered mutual devices interactions and embedded the information on devices interactions into the Factorial Hidden Markov Model (FHMM) representations of the aggregated data. This method disaggregated and assigned energy consumption for individual devices more accurately in the REDD public data set.

In view of the above problems, we propose an online NILM method based on HMM. This method achieves the purpose of distinguishing different appliances by analyzing the trend of apparent power in transient-state. The method is embedded into the NILM device, thus avoiding the need for mass and fine load data transmission.

3. Methods
The data of current and voltage are obtained from the power line through sensors. First, the data is preprocessed to remove the noise in the signal. Then, the features are extracted from the data which are used as the input of training to optimize the parameters of the model. The processing of data is
shown in Fig. 1.

![Diagram](image)

**Figure 1.** The data processing in NILM

3.1 Improved DB9 algorithm

In discrete wavelet transform, Daubechies is very sensitive to non-stationary signals. The traditional DB9 algorithm obtains the frequency-domain features of the signal by scaling the scale of the wavelet to filter out the noise. The high-frequency part is processed correspondingly, while the low-frequency component does not take any operation through this method. An improved method based on DB9 is proposed in this paper. The low-frequency components obtained by the DB9 decomposition are added to the median filter algorithm. These residual noises will be filtered under the effect of median filtering, while the detail part will be recovered in the process of wavelet reconstruction. And the process is repeated iteratively to achieve a relatively ideal noise filtering effect. The process of preprocessing is as follows.

\[ H^{(k)} = S - L_M^{(k)} \]  
\[ H^{(k)} \xrightarrow{\text{Wavelet Reconstruction}} \Delta L^{(k)} \]  
\[ L^{(k+1)} = L_M^{(k)} + \Delta L^{(k)} \]  
\[ L^{(k+1)} \xrightarrow{\text{Median Filtering}} L_M^{(k+1)} \]

where, S is the original power signal, H is the high-frequency component, and L is the low-frequency component.

3.2 Application of HMM in NILM

Hidden Markov model (HMM) is a dual stochastic process, which can solve the transformation tracking between states and effectively distinguish short time stationary signal segments with different parameters. Therefore, the HMM can accurately describe the process of the change of the power in the
transient-state, and realize the classification and recognition of the load in different states.

The apparent power of load is calculated by filtered current and voltage data. Taking the features of the apparent power as the observation sequence to build the HMM of load switching. We choose the data which include several periods before and after the load changing point, and add windows to the data in order to extract features from each segment. Some overlaps should be retained between adjacent forms so that the feature changes on each form are smooth. Through the LBG method, these features are generated codebooks, which are used as input to the HMM model. The model is trained by multiple iterations of the Baum-Welch algorithm. When the change of probability during two adjacent parameter revaluations is less than the set threshold, the iteration is terminated, and the HMM is built.

\[ \lambda = (A, B, \pi) \]  

where, \( A \) is the probability distribution of state transfer, \( B \) is the probability distribution of state observation, \( \pi \) is the distribution of the initial state.

3.3 Improved Forward algorithm

For each switch of the load, we build one model of HMM. In the process of recognition, Forward algorithm is used to calculate the similarity probability between the observation sequence and the existing model. This algorithm recursively calculates and outputs the probability of the observation sequence from back to back. It is one of the commonly used methods in HMM for classification and recognition. It is found that both probability matrix of the state transition and probability matrix of the state observation are made up of decimal fractions. When calculating the forward probability, the probability value will overflow after several times of multiplication. This leads to the phenomenon of load identification error. To solve this problem, we improved the algorithm by adding formulas (6) and (7). When calculating the similarity probability between the model and the observation sequence, that is, \( P(O|\lambda) \), where, \( O \) is the observation sequence, the formula (8) is added.

\[ \text{Num}(t+1) = \log \left( \frac{1}{\max \left(a_{t+1}(i)\right)} \right) \quad t = 1, 2, \ldots, T-1 \quad i = 1, 2, \ldots, N \]  

\[ a_{t+1}(i) = a_{t+1}(i) \times 10^{\text{Num}(t+1)} \quad t = 1, 2, \ldots, T-1 \quad i = 1, 2, \ldots, N \]  

\[ \text{Sum}_\text{Num} = \sum_{t=1}^{T} \text{Num}(t) \quad t = 1, 2, \ldots, T \]  

Before judging the \( \max \left(P(O|\lambda)\right) \), the value of the \( \text{Sum}_\text{Num} \) is calculated first. If the value of the \( \text{Sum}_\text{Num} \) is smaller, that means the number of indentation is smaller, that is, the likelihood of similarity is greater. When the number of \( \min \left(\text{Sum}_\text{Num}\right) \) is two or more, the corresponding values of \( P(O|\lambda) \) are calculated respectively. Finally, the result of recognition is defined as the larger corresponding load type.
4. Experimental Results and Discussion

The STMicroelectronics 32-bit chip (STM32), which is used as the core processor, the current sensors and voltage sensors are attached to create a NILM in this paper. The sample frequency is about 4 kHz. The improved filtering algorithm, modeling algorithm of HMM and improved forward algorithm are embedded in the monitoring system. We collect the current and voltage data of six common loads of lamp, television, electric radiator, refrigerator, hair dryer and kettle.

4.1 The effect of improved DB9 algorithm

Figure 2. Comparison of the current waveform of lamp of original data, DB9 filtering and improved DB9 filtering

The comparison of the original current data, the DB9 filtering effect and the improved DB9 filtering effect is shown in Fig.2. The experiment shows that the improved DB9 algorithm can filter the noise to the greatest extent and reduce the distortion of the waveform. The effect is particularly obvious in small power load. The signal to noise ratio of the six loads during different processes is shown in Table 1.

| NO. | Load Type    | Raw Data | DB9 Filtering | Improved DB9 Filtering |
|-----|--------------|----------|---------------|------------------------|
| 1   | Lamp         | 13.4756  | 15.5931       | 18.1903                |
| 2   | TV           | 17.5246  | 20.4267       | 23.1795                |
| 3   | Electric Radiator | 34.7029  | 39.6955       | 44.4578                |
| 4   | Refrigerator | 19.4681  | 23.8815       | 28.5655                |
| 5   | Hair dryer   | 25.1642  | 29.6482       | 34.128                 |
| 6   | Kettle       | 30.2494  | 37.5988       | 45.7033                |

4.2 Performance of proposed algorithm

In this experiment, we take the apparent power of the five periods before and after the load opening time to form the open event data set, and the same method is applied to the closing event. Each
window contains 80 data points, with the moving distance of 70. The data of 10 points are calculated repeatedly. Five features are extracted from the data of each window, which are the maximum of the power, the minimum of the power, the total power, the average power and the peak-to-peak of the power. The size of codebook is 128 and the dimension of code vector is 4 dimensions. We build 12 models of HMM, including each switch of load. We use 100 sets of data to train each model, set the threshold of iteration to 0.015, and use 50 sets of data to test the model. Experiments show that each model can reach convergence state after about 50 times of training. The recognition effect of load opening state is used as a demonstration in this paper.

Because the underflow phenomenon occurs in probability calculation, the probability result is partially zero, and some loads are identified as lamps. The total recognition rate in the opening state of load is 77.33%. By improving the forward algorithm, the data underflow phenomenon is suppressed effectively, and the total recognition rate is increased to 92.67%, as shown in Fig.3. With the HMM based on transient-state apparent power and improved forward algorithm, the classification of loads achieves a good result, as shown in table 2. The accuracy rates of the states of load opening and closing are 92.67% and 93%, respectively.

**Table 2. The statistics of classification accuracy in load opening and closing state**

| NO. | Load Type  | The state of load opening | The state of load closing |
|-----|------------|---------------------------|--------------------------|
|     |            | Correct/Total | Accuracy | Correct/Total | Accuracy |
| 1   | Lamp       | 50/50 | 100.00% | 49/50 | 98.00% |
| 2   | TV         | 46/50 | 92.00% | 45/50 | 90.00% |
| 3   | Electric Radiator | 44/50 | 88.00% | 46/50 | 92.00% |
| 4   | Refrigerator | 45/50 | 90.00% | 44/50 | 88.00% |
| 5   | Hair dryer | 50/50 | 100.00% | 50/50 | 100.00% |
| 6   | Kettle     | 43/50 | 86.00% | 45/50 | 90.00% |
|     | Total      | 278/300 | 92.67% | 279/300 | 93.00% |
5. Conclusion
The experiments show that the improved DB9 algorithm can effectively filter out low-frequency noise and significantly improve the signal-to-noise ratio of monitoring data. The HMM is built online by extracting the features of transient-state apparent power. The improved forward algorithm can effectively suppress data underflow and improve the accuracy of load classification. The load identification rate is over 90% with the algorithm proposed in this paper embedded into the NILM system. To achieve higher accuracy in detection, we will look to evaluate and test this algorithm in different conditions in the future work.

Acknowledgments
This work was supported in part by China International Scientific and Technological Cooperation Special (2015DFR10490), Qingdao national laboratory for marine science and technology Aoshan science and technology innovation project (2016ASKJ07-4) and Qingdao innovation and entrepreneurship leading talent project (13-cx-2).

References
[1] Wan C, Zhao J, Song Y, et al. Photovoltaic and solar power forecasting for smart grid energy management. Csee Journal of Power & Energy Systems, 2016. 1(4):38-46
[2] Kamyab F, Amini M, Sheykhha S, et al. Demand Response Program in Smart Grid Using Supply Function Bidding Mechanism. IEEE Transactions on Smart Grid, 2016, 7(3):1277-1284
[3] Lin Y H, Tsai M S. An Advanced Home Energy Management System Facilitated by Nonintrusive Load Monitoring With Automated Multiobjective Power Scheduling. IEEE Transactions on Smart Grid, 2015, 6(4):1839-1851
[4] Hart G W. Nonintrusive appliance load monitoring. Proceedings of the IEEE, 1993, 80(12):1870-1891
[5] M S Tsai, Y H Lin. Development of a Non-Intrusive Monitoring Technique for Appliance’ Identification in Electricity Energy Management. International Conference on Advanced Power System Automation and Protection, 2012,1:108-113
[6] Basu K, Debusschere V, Bacha S, et al. Nonintrusive Load Monitoring: A Temporal Multilabel Classification Approach. IEEE Transactions on Industrial Informatics, 2015, 11(1):262-270
[7] M Aiad, HL Peng. Unsupervised approach for load disaggregation with devices interactions. Energy and Buildings, 2016, 116:96-103