

# A non-intrusive load decomposition method for residents

**Chengjian Wang\* and Mingyue Zhai**

School of Electrical and Electronic Engineering, North China Electric Power University, Beijing, 102206, China

\*Corresponding author's e-mail: qq2323454983@163.com

**Abstract.** Aiming at the problem that the large amount of data involved in the existing load decomposition method leads to low decomposition efficiency and high hardware requirements, a non-intrusive load decomposition method based on hidden Markov model (HMM) and improved Viterbi algorithm is proposed. Monitor steady state current of each load and quantify the data obtained and then establish the probability distribution of quantized monitoring values by statistics. The states that a load has can be identified from the probability distribution. The running states of each load at the same moment are arranged in sequence to form a combination state. HMM is created based on the correlation of combination states in the transition and the correlation between the combination state and the quantized monitoring value of the total current. Based on matrix sparsity, a compression algorithm is used to reduce data storage, and a query algorithm and an improved Viterbi algorithm are used to avoid unnecessary calculations about zero-probability terms to further improve the decomposition efficiency. Experimental results show that this method can effectively improve the decomposition efficiency and need lower hardware requirements while accurately obtaining the running state of various loads.

## 1. Introduction

With the development of society, the scale of power grids and the demand for electrical energy are increasing[1], the production, transmission and utilization of electric energy are of great significance to the sustainable development of the economy and the environment[2]. At present, China is building a strong smart grid, and its core is to realize intelligent services. In the smart grid, users can also participate in generation, transmission, and management of electric energy. Load decomposition is one of the key technologies of the smart grid, which allows users to know the energy consumption information of loads within the home, actively respond to energy-saving policies, and promote rational use of electricity[3]. Studies have shown that giving residents more information on electricity consumption helps residents to adopt energy-saving measures, which can reduce energy consumption by an average of about 10%[4]. Load decomposition can be divided into two categories: intrusive and non-intrusive. Intrusive load decomposition needs to be equipped with monitoring sensors for each load. Although the data obtained is more accurate, it is costly and is not suitable for promotion. Non-intrusive load decomposition only installs the sensor at the user's electrical inlet, by collecting the total power or total current of the user's electricity to decompose the power consumption and working state of each type of load, it is simple, economical, reliable[5].

Existing decomposition methods typically either need to run in high performance hardware devices or have insufficient precision. For example, literature[6] adopts fuzzy clustering method to determine the characteristic parameters of the steady-state current or voltage data, and then the load decomposition is carried out by the differential evolution algorithm, but the identification device type



is not precise enough; literature[7] adopts an evolutionary algorithm to realize non-intrusive load decomposition, but the single load feature is used as the optimal target, and the recognition effect is not ideal, especially when the household load power is small or close; literature[8] proposed a load transient event detection method based on the bilateral accumulation of sliding windows; literature[9] proposed an instantaneous multi-label classifier for load decomposition. When detecting a transient event of a device, if multiple power loads are simultaneously switched, the load characteristics generated may be superimposed, existing methods are difficult to handle, and the load monitoring device is required to continuously capture every change information of the electric load, so the high-frequency sampling of the load data is required to obtain a more differentiated load characteristic. That leads to greater data storage and processing costs. The high performance hardware itself consumes high energy, reduces the role of the decomposer in promoting the user to save electricity, so it is not suitable for the residents.

In order to make the splitter more suitable for residential users, it should be cheap and low power consumption. To solve this problem, according to the law of residents' electricity consumption, the correlation of load state transition is statistically analyzed, and a non-intrusive load decomposition method based on HMM and improved Viterbi algorithm is proposed. Since the internal state of each load cannot be directly monitored from the total energy consumption data, HMM can model the time series data and present the internal state of each load unobservable. Therefore, HMM is suitable for creating a load decomposition process model[10]. And HMM will be decoded by the improved Viterbi algorithm to obtain the decomposition result.

## 2. Establishing load model

### 2.1. State identification of a single load

With the vigorous development and popularization of smart meters, the load characteristics such as power, current or voltage under steady state conditions are easy to obtain, and have lower hardware requirements than transient switching characteristics obtained at high sampling frequencies. In this paper, to establish a data set, each load is equipped with a sensor to monitor and collect the steady-state current data of the load at a sampling frequency of 1/60 Hz. Due to noise interference and voltage fluctuations, an infinite number of discrete sampling values can be obtained within the maximum allowable current value range. To facilitate statistical analysis, the data needs to be quantized, and then the probability distribution of quantized monitoring values can be statistically calculated[11]. In this paper, the rounding criterion is used, and the unit of quantization is 0.1A. For example, if the monitored value is 1.63A, the quantified monitored value is 16; if the monitored value is 16.7A, the quantified monitored value is 17. When the load is running in a certain state, the actual steady-state current value fluctuates up and down at the rated current value in the state, so that the probability of monitoring value near the rated current value is large, while the probability of monitoring value away from the rated current value is low. According to this feature, the states of a load has can be identified from the probability distribution of the quantized monitored value, and then index the state of load. As shown in table 1, a load has a total of  $N=7$  quantized monitoring values,  $n$  is the quantized monitoring value,  $p(n)$  is the probability of  $n$ ,  $k$  is the index number of the identified state, and  $n_k$  is the quantized monitoring value of the state with index number  $k$ .

Table 1. The probability distribution of quantized monitoring values.

$n$	0	1	2	3	4	5	6
$p(n)$	0.45	0.05	0.06	0.33	0.07	0.04	0.0
$k$	0			1			
$n_k$	0			3			

Definition: When  $p(n)-p(n-1)>0$  and  $p(n)-p(n+1)>0$ ,  $p(n)$  is a maximal value; when  $p(n)-p(n-1)<0$  and  $p(n)-p(n+1)<0$ ,  $p(n)$  a minimum value. When  $p(n)$  is a maximum value and  $p(n)>e$  (to avoid

identifying the maximum value with less probability as states), it is considered that a load state is identified.

### 2.2. Creating the combination state

The running states of each load within a home at the same moment is arranged in sequence to form a combined state  $S$ ,  $S=k^{(1)}k^{(2)}k^{(3)}\dots k^{(M)}$ , where  $k^{(M)}$  is the internal state index of the  $m$ -th load, and  $M$  is the total number of loads within the home.  $S$  is indexed according to the following formula, where  $K^{(i)}$  is the total number of states that the  $i$ -th load has. The index number is  $s$ .

$$s = k^{(M)} + \sum_{m=1}^{M-1} \left( k^{(m)} \times \prod_{i=m+1}^M K^{(i)} \right) \quad (1)$$

### 2.3. Creating the HMM

Assuming that there are  $M$  loads within a home, these loads can form  $T$  combination states. Building HMM as  $\lambda = \{P_0, A, B\}$ , Where  $P_0$  is a row vector consisting of the initial probabilities of the combination states, its length is  $T$ ,  $A$  is a  $T \times T$  transition probability matrix,  $B$  is a  $T \times N$  emission matrix,  $N$  is the number of possible observations.  $A, B$  is defined as follows, where  $S_t$  is the combination state at time  $t$ , and  $y_t$  is the quantized monitoring value at time  $t$ :

$$A[i, j] = p\{S_t = j | S_{t-1} = i\} \quad (2)$$

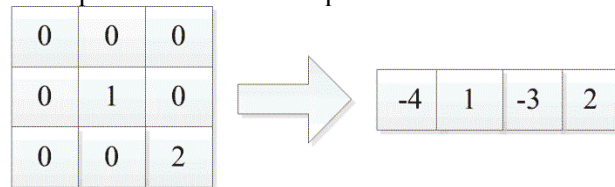
$$B[j, n] = p\{y_t = n | S_t = j\} \quad (3)$$

## 3. Data compression algorithm

It is found in experiments that as the number of loads increases, the amount of data in the matrix of HMM will increase sharply, but the matrix has strong sparsity. It turns out that  $A$  is 45.3% sparse on average while  $B$  is 97.5% sparse on average. Therefore, data compression is very valuable. Based on the non-negative nature of the data in HMM, the data compression steps are as follows:

- (1) Create an empty row vector  $V$  and a variable  $r = 0$ ;
- (2) Start from the first column of the matrix and read the matrix data by column;
- (3) According to the value of the matrix element, if it is positive, it is added to  $V$ ; if it is 0, the variable  $r$  starts counting, and then continues to read data from the matrix until the next positive number or the end of the matrix, so the value of  $r$  is the number of consecutive zeros, take the negative value of  $r$  added to  $V$ ;
- (4) Reset  $r$  to 0 and continue reading the data in the matrix;
- (5) Repeat step (3) and (4) until the end of data reading of the matrix.

The row vector  $V$  is the compressed form of the sparse matrix. As shown in figure 1.



Before Compression

After Compression

Figure 1. The schematic diagram of data compression.

With the increase of the number of loads, the number of combination states will rapidly increase, as shown in figure 2. As the number of combined states increases, the amount of data in the matrix increases exponentially. When the number of loads increases to nine, the amount of data compressed by the data compression algorithm is greatly reduced, so this can effectively reduce the requirements for storage devices, as shown in figure 3.

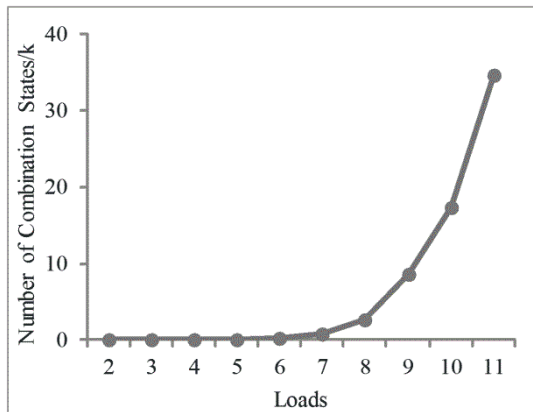


Figure 2. The combination states growth.

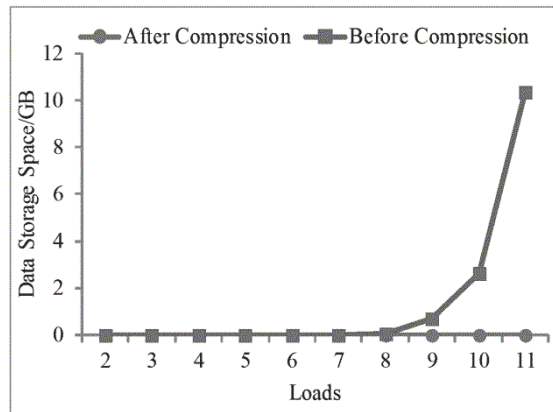


Figure 3. The Space required for data storage.

#### 4. Data query algorithm

In order to improve the computational efficiency, it is necessary to read data directly from  $V$ . Therefore, two data query algorithms are proposed.

##### 4.1. Query data by column

Suppose there is a  $T \times N$  sparse matrix, and its compressed form is row vector  $V$ . Query all non-zero elements in  $i$ -th column of matrix from  $V$ , the steps are as follows:

- (1) Create a variable  $r=0$  and a matrix  $H$  that contains two column vectors;
- (2) Read the data in  $V$  from the first element;
- (3) If the data is negative,  $r$  plus its absolute value; if the read data is positive,  $r$  plus 1;
- (4) When the read data is positive,  $r > T \times i$  and  $r \leq T \times (i+1)$ , add  $r - T \times i - 1$  and the data to  $H$ ;
- (5) When  $r < T \times (i+1)$ , continue to read data from  $V$ ; when  $r \geq T \times (i+1)$ , data reading ends.

##### 4.2. Query single data

Suppose there is a  $T \times N$  sparse matrix, and its compressed form is row vector  $V$ . The steps for querying the data of the  $i$ -th row, the  $j$ -th column of the matrix from  $V$  are as follows:

- (1) Create a variable  $r=0$ ;
- (2) Read the data in  $V$  from the first element;
- (3) If the data is negative,  $r$  plus its absolute value; if the read data is positive,  $r$  plus 1;
- (4) When the data is negative and  $r$  is greater than or equal to  $T \times j + i + 1$  for the first time, the requested data is 0; when the data is positive and  $r = T \times j + i + 1$ , the positive number is the requested data.

#### 5. Improved Viterbi algorithm

Viterbi algorithm is very suitable for decoding of HMM, but due to the strong sparsity of the matrix in HMM, it will generate a large number of calculations on the condition of zero probability. The result of these calculations is zero, and there is no substantial impact on the decomposition result. Therefore, to avoid these unnecessary calculations, this paper has improved the Viterbi algorithm.

Let  $y_{t-1}$  and  $y_t$  be the quantized monitoring values of the total current at  $t-1$  and  $t$  respectively.  $V_A$ ,  $V_B$  is the corresponding row vector for  $A$  and  $B$ . When  $y_{t-1}$  is detected and all non-zero data of column  $y_{t-1}$  is obtained by the query algorithm, the probability of the combination state at  $t-1$  can be calculated by the following formula:

$$P_{t-1}[j] = \begin{cases} P_0[j] \times B[j, y_{t-1}], & B[j, y_{t-1}] > 0 \\ 0, & B[j, y_{t-1}] = 0 \end{cases}, j = 0, 1, \dots, T-1 \quad (4)$$

The calculation involving  $B[j, y_{t-1}] = 0$  will be omitted directly, which can avoid the calculation under the condition of zero probability.  $P_t[j]$  can be obtained by the following formula:

$$P_t[j] = \max_{0 \leq i \leq T-1} \{P_{t-1}[i] \times A[i, j] \times B[j, y_t]\}, j = 0, 1, \dots, T-1 \quad (5)$$

Take the combination state with the highest probability at time  $t$  as the actual state within the home. Its index is:

$$s = \arg \max(P_t) \quad (6)$$

As shown in figure 4, as the number of loads increases, the improved Viterbi algorithm can effectively improve the computational efficiency.

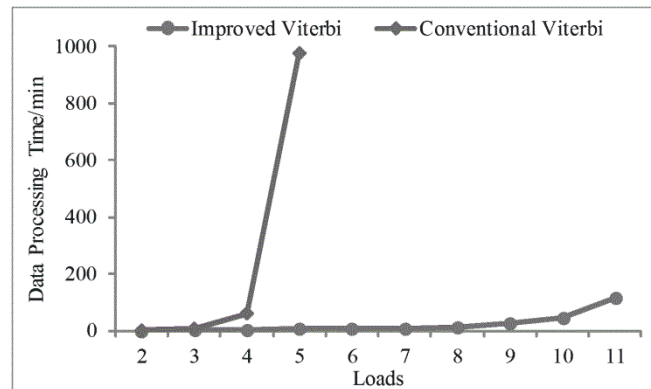


Figure 4. The time required for data processing.

## 6. Summary

By creating HMM and decoding it with an improved Viterbi algorithm, the load decomposition problem is transformed into the process of solving the optimal combination. This method preserves the correlation of the transitions between the combined states, so that the combination state within a home can be accurately inferred based on the the quantized monitoring value of the total current at the previous and the current moment. The accuracy of this method depends primarily on the establishment of the data set. The more comprehensive the collected data, the more accurately the internal electricity consumption law is reflected. In addition, the use of the compression algorithm greatly reduces the data storage, and the use of the query algorithm further improves the computational efficiency. This method provides an idea for making low-cost, low-power and reliable splitter.

## References

- [1] Wu, P., Guo, W., Jin, C.J. (2018) Analysis on the current situation of electricity and energy in China and its solution. *J. Electrical Engineering*, 19: 1–4.
- [2] Xu, J.J., Wang, B.E., Yan, L.M. (2017) The strategy of the smart home energy optimization control of the hybrid energy coordinated control. *J. Transactions of China Electrotechnical Society*, 32: 214–223.
- [3] Chen, S.Y., Gao, F., Liu, T. (2016) Load disaggregation method based on factorial hidden Markov model and its sensitivity analysis. *J. Automation of Electric Power Systems*, 40: 128–136.
- [4] Yan, X.D., Li, Y. (2014) Research on household electricity response mode. *J. Power System Protection and Control*, 42: 51–56.
- [5] Yu, Y.X., Liu, B., Luan, W.P. (2013) Nonintrusive residential load monitoring and decomposition technology. *J. Southern Power System Technology*, 7: 1–5.
- [6] Li, P., Yu, Y.X. (2009) Nonintrusive Method for On-line Power Load Decomposition. *J. Journal of Tianjin University*, 42: 303–308.
- [7] Lou, J.L., Yu, H.T., Qu, C.Y. (2015) Household User Oriented Method of Power Load Decomposition. *J. Journal of Jilin University (Science Edition)*, 53: 744–753.

- [8] Niu, L.L., Jia, H.J. (2011) Transient event detection algorithm for non-intrusive load monitoring. *J. Automation of Electric Power Systems*, 35: 30–35.
- [9] Basu, K., Debusschere, V., Bacha, S. (2015) Nonintrusive load monitoring: a temporal multilabel classification approach. *J. IEEE Transactions on Industrial Informatics*, 11: 262–270.
- [10] Rabiner, L.R., Juang, B.H. (1986) An introduction to hidden Markov models. *J. IEEE ASSP Magazine*, 3: 4–16.
- [11] Zeifman, M. (2012) Disaggregation of home energy display data using probabilistic approach. *J. IEEE Transactions on Consumer Electronics*, 58: 23–31.