A Feature Extraction Method for Non-intrusive Load Identification

Wei Wei¹,a, Tao Peng²,b, Li Ye³,c, Xianyu Feng⁴,d*, Zisong Jiang⁵,e, Heyang Yu⁶,f

¹State Grid Hubei Marketing Service Center Hubei Electric Power Co., Ltd Hubei, China
²State Grid Hubei Marketing Service Center Hubei Electric Power Co., Ltd Hubei, China
³State Grid Hubei Marketing Service Center Hubei Electric Power Co., Ltd Hubei, China
⁴Zhejiang Chint Instrument & Meter Co., Ltd Chint Group Hangzhou, China
⁵Zhejiang Chint Instrument & Meter Co., Ltd Chint Group Hangzhou, China
⁶College of Electrical Engineering Zhejiang University Hangzhou, China

¹1036243440@qq.com, ²2661900807@qq.com, ³21515996@qq.com, ⁴fxy@chint.com, ⁵zisong.jiang@chint.com, ⁶yuheyang@zju.edu.cn

Abstract—When some household appliances are running, the electrical quantities such as active power and reactive power will change slowly for a long time, which will affect the feature extraction of event-based non-intrusive load identification, resulting in low accuracy of load identification results. In view of this situation, a feature extraction method is proposed. Its core idea is that there is a certain law in the change of active power and reactive power in a very short time, and the curve can be fitted to speculate the values of these electrical quantities in the next time, so as to eliminate the influence. Based on this, a feature extraction method applied to non-intrusive load identification is realized. The test results show that this method can effectively improve the accuracy of non-intrusive load identification feature extraction data, and has good application value.

1. INTRODUCTION
With the development of the national economy, “power demand side management” has become increasingly important, which can give full play to the value of power companies and avoid unnecessary waste [1-2].

The use of load identification technology can finely analyze the characteristics of power consumption of power users, formulate energy-saving and supply-demand interaction strategies for the power grid, and provide a reference for the development of “power demand side management” [3-5].

The current load identification technology is divided into Intrusive Load Monitoring (NLM) technology and Non-Intrusive Load Monitoring (NILM) technology [6]. Compared with intrusive load identification technology, non-intrusive load identification technology has the characteristics of low cost and easy installation [7-10]. Only a device for collecting electricity consumption data is installed at
the power entrance, and the electricity consumption data is analyzed through algorithms, thus acquiring
the composition and operation of electrical equipment. In recent years, with the development of
microelectronics technology, sampling sensing technology, and machine algorithms, the research on
non-intrusive load identification technology has become increasingly widespread, mainly based on
Hidden Markov Model (HMM) [11-13], deep learning neural network [14-16], event method [17-20], etc.
Although there are many methods of NILM, event-based NILM, due to its strong operability, has been
widely used in recent studies. The load decomposition of the event-based framework includes steps
such as data measurement, data processing, event detection, feature extraction, and load identification.
The effectiveness of the feature extraction step plays a key role in the accuracy of load identification. If
the home appliances are running smoothly, when the NILM system detects the occurrence of an
event (in this case, the start-up, state switching, and shutdown of an appliance are called events), the
difference between the appliance feature quantity after the event ends and the appliance feature quantity
before the event occurs can be extracted directly as the appliance feature extraction value. When
household appliances are not all running smoothly, such as air conditioners, refrigerators, and other
appliances, there will be a long-term slow increase or slow decrease in the amount of active power,
reactive power, and other appliances (hereinafter collectively referred to as slow change). This will lead
to deviations in the extracted values of the feature quantities, which will reduce the accuracy of
subsequent steps such as load identification and matching.

Therefore, how to solve the problem of accurate feature extraction of non-intrusive load
identification appliances with a slowly changing background has become an urgent problem. This paper
introduces the compensation algorithm in the feature extraction step for the framework described above.
The core idea is that there is a certain law of in the change of active and reactive power changes in a
very short period of time when these slow transformers are operati

2. Feature Extraction Principle of Slow Transformers

It is assumed that the feature extraction method of event-based non-intrusive load identification is to
calculate the feature group $\Delta \Omega_1$ of electrical appliance feature changes from the beginning to the end of
the event, which mainly includes feature quantities such as active power, reactive power, and harmonic
content.

Among them, the feature group $\Delta \Omega_1$ is calculated by the following formula:

$$\Delta \Omega_1 = \Omega(t_{\text{event-off}}) - \Omega'(t_{\text{event-off}})$$

(1)

$\Omega(t)$ refers to the feature group used for load identification that changes over time.

$\Omega(t_{\text{event-off}})$ is the actual value of the feature group at the end of the event $t_{\text{event-off}}$.

$\Omega'(t)$ represents the value of the feature group predicted by fitting the curve to the feature group data
before the event that changes with time for load identification.

$\Omega'(t_{\text{event-off}})$ is the predicted value of the feature group at the end of the event $t_{\text{event-off}}$.

That is to say, in the context of a slowly-varying feature, when the event occurs, each appliance
feature quantity is predicted by curve fitting of all appliance feature quantities to obtain the predicted
value of each appliance feature quantity at the end of the event if the event does not occur. Finally, the
actual operating feature quantity of the electrical appliance at the end of the event is compared with the
predicted value of the appliance feature quantity, so as to compensate for the change of the electrical
feature quantity during the event period under the background of the feature quantity when the slow
transformer is operating, thereby improving the non-intrusive load identification of the electrical feature
quantity and enhancing the identification accuracy of the overall algorithm.
3. Feature Extraction Algorithm to Overcome the Influence of Slow Transformers

Based on the above principles, this paper proposes the following algorithm:

Step (1): Sample the voltage and current at a rate of 256 points per cycle and calculate according to the sampling results. Thus, a feature sequence of active power, reactive power, and harmonic content is formed, which is combined into the above-mentioned $\Omega(t)$.

Step (2): Detect whether the feature sequence $R_i(t)$ is slowly rising or slowly falling (hereinafter collectively referred to as slow change). If it is determined that the feature sequence $R_i(t)$ is slowly changing, it is needed to record the starting time $t_{slow-on}$ of the feature sequence slowly changing, go to Step (3), otherwise go to Step (1).

Among them, the method of judging that the feature sequence $R_i(t)$ is slowly changing is as below: calculate the average value $R_n$ of all the feature sequences of increasing for a period of time, calculate the increment sequence $\Delta R_n$ of the average value, and when the average increment $\Delta R_n$ is greater than a certain threshold $D_1$, and if the number $num$ is greater than a certain threshold $D_2 (num > D_2)$, it is judged that the feature quantity is slowly increasing. The calculation method of the average value $R_n$ of the feature quantity sequence and the increment sequence $\Delta R_n$ of the average value is as follows:

$$R_n = \frac{1}{t_{increase}} \sum_{t=n*t_{increase}+1}^{(n+1)*t_{increase}} R(t)$$

$$\Delta R_n = R_{n+1} - R_n$$

Step (3): Calculate the change $\Delta P_1(t)$ of active sequence $P_1(t)$ at time $t$, detect whether $\Delta P_1(t)$ is greater than the active event starting threshold $P_{start}$, if $\Delta P_1(t) \geq P_{start}$, it is needed to record the current moment as the event start moment $t_{event-on}$, and at the same time the moment is the end moment $t_{slow-off}$ of collecting the slowly-varying data interval of the active sequence. After the event starts, when detecting whether the active power sequence changes $\Delta P_1(t)$, $\Delta P_1(t-1)$, and $\Delta P_1(t-2)$ at time $t$, $t-1$, and $t-2$ are all less than the active event end threshold $P_{end}$, if yes, it is required to record the current time as the event end time $t_{event-off}$ and the actual value $\Omega(t_{event-off})$ of the feature group at this time, go to Step (4), otherwise continue to Step (3).

Step (4): Set the feature group of the slow change interval of the collected active power sequence to $\Omega_{slow}(t)$.

$$t \in \{t_{slow-on}, t_{slow-on}+1, \ldots, t_{slow-off} - 1\}$$

The variables of the feature are fitted with the following curve to obtain the curve group $\Omega(t)$.

$$\hat{y} = \hat{a} + \hat{b}t$$

Where the total time length $T_{slow}$ is calculated by the following formula:

$$T_{slow} = t_{slow-off} - t_{slow-on} + 1$$

According to the least square method of curve fitting, it is needed to set a characteristic in the feature group $\Omega_{slow}(t)$ as $\omega(t)$, which is its corresponding fitting curve.
\[ \hat{y} = p_{\omega}(t) = \hat{a}_{\omega} + \hat{b}_{\omega} t \]  

(7)

\[ \hat{b}_{\omega} = \frac{\sum_{i=t_{\text{slow}}-on}^{T_{\text{slow}}-sum} t_i y_i - \sum_{i=t_{\text{slow}}-sum}^{T_{\text{slow}}-sum} t_i y_i}{\sum_{i=t_{\text{slow}}-on}^{T_{\text{slow}}-sum} t_i^2 - \left( \sum_{i=t_{\text{slow}}-on}^{T_{\text{slow}}-sum} t_i \right)^2} \]

(8)

\[ \hat{a}_{\omega} = \sum_{i=t_{\text{slow}}-on}^{T_{\text{slow}}-sum} y_i - \hat{b}_{\omega} \sum_{i=t_{\text{slow}}-on}^{T_{\text{slow}}-sum} t_i \]

(9)

Step (5): in the curve group \( \Omega(t) \), let \( t = t_{\text{event-off}} \), then the predicted value \( \Omega'(t_{\text{event-off}}) \) of the feature group is obtained if the event does not occur at the end time of the event; the feature group \( \Delta \Omega_i \) is obtained after the electric appliance characteristic extraction and compensation, which is the result.

The changes in active power, reactive power, and current harmonic amplitude changes in the feature group are calculated by the following formulas:

\[ \Delta P_{\Omega-\Omega} = P(t_{\text{event-off}}) - P'(t_{\text{event-off}}) \]

(10)

\[ \Delta Q_{\Omega-\Omega} = Q(t_{\text{event-off}}) - Q'(t_{\text{event-off}}) \]

(11)

\[
\Delta i_{\Omega-\Omega} = \sqrt{\frac{i_{m,t_{\text{event-off}}}^2 + I_{m,t_{\text{event-off}}}^2}{2}} - 2 * I_{m,t_{\text{event-off}}} i_{m,t_{\text{event-off}}} \cos \left( \phi_{k,t_{\text{event-off}}} - \phi_{k,t_{\text{event-off}}} \right)
\]

(12)

The process steps are as follows:
Figure 1. Flow chart of feature extraction of slow transformer
4. Test and Verification
This paper uses Matlab 2019a development platform and m programming language to develop and implement the method and uses a PC with the main frequency of 2.6GHz CPU and memory of 4G to complete the test and verification of this method. The following will carry out technical verification around a household load operation scenario.

The verification background is that a fixed-frequency air conditioner in the home is turned on and the active power is slowly increasing. At the 10th second, the occurrence of the “rice cooker on” event is detected, and this time is recorded. At the 20th second, the “rice cooker on” event ends. The total active power of the household increases rapidly. At this time, it is the algebraic sum of the active power of the fixed-frequency air conditioner and the rice cooker. In the same way, the total reactive power also changes to the algebraic sum of the operating power of the fixed-frequency air conditioner and the rice cooker. Further, the change in the amplitude of the fifth harmonic is considered as the vector sum of the operating power of the fixed-frequency air conditioner and the rice cooker. The curve is used to fit each feature of the household load before the event, and predict the value of each feature at the end of the event, as shown by the thick solid line at the 20th second in Figure 2, Figure 3, and Figure 4 below.

Figure 2. Active power characteristic fitting and actual comparison diagram
The change amount of the feature group after the electrical feature extraction and compensation is calculated, and then it is compared with the data of the rice cooker under stable conditions. The results are shown in Table 1.

**Figure 3.** Reactive power characteristic fitting and actual comparison diagram

**Figure 4.** Current 5th harmonic amplitude fitting and actual comparison diagram

The change amount of the feature group after the electrical feature extraction and compensation is calculated, and then it is compared with the data of the rice cooker under stable conditions. The results are shown in Table 1.

**TABLE 1.** The value of the related variables of the feature group before and after the electrical feature extraction and compensation and the amount of change in the steady-state

| Feature quantity | The actual value when the event starts | The predicted value when the event ends | The actual value when the event ends | Original change | Compensation change | The amount of change in steady state |
|------------------|---------------------------------------|----------------------------------------|-------------------------------------|----------------|-------------------|-------------------------------------|
| a                | 1310.0                                | 1322.4                                 | 1832.9                              | 522.9          | 510.5             | 508.23                              |
| b                | 260.12                                | 257.53                                 | 268.02                              | 7.9            | 10.49             | 13.72                               |
| c                | 1.3690                                | 1.3566                                 | 1.4511                              | 0.102673       | 0.115535          | 0.112467                            |
| d                | 147.8109                              | 147.6028                               | 150.3175                            |                |                   |                                     |
Note: In the above table, a represents active power; b represents reactive power; c represents the amplitude of 5th harmonic of the current; d represents the phase angle of 5th harmonic of the current.

It can be seen from the above table that after compensation, the compensated eigenvalue is close to the eigenvalue in the steady state, which will improve the accuracy of non-intrusive load identification.

5. CONCLUSION
With respect to the load identification method based on events and probability, if the cut-off probability is increased, it is prone to misjudgment, and if the cut-off probability is reduced, it is subject to the missed judgment. Thus, this paper provides a method for accurately extracting feature quantities when a slow transformer exists. This method has obvious significance for reducing missed judgments and improving identification accuracy.

ACKNOWLEDGMENT
The authors would like to thank the State Grid Corporation of China for its great support of this research. This paper is funded by the key technology research of smart IoT meter for energy service (52153220001Y) of State Grid Hubei Electric Power CO., LTD.

REFERENCES
[1] Zou H P. Application analysis of power demand side management technology in electrical economic management [J]. China management informatization, 2019, 22 (12): 98-99.
[2] Xia X Z, Chen H, Li W F, et al. Practice of power demand side management based on improving customer response [J]. Electrical technology and economy, 2018 (01): 62-64.
[3] Zhang Y J, Liu S L, Jiang J Q, et al. Research review on low-voltage intelligent distribution network technology[J]. Guangdong Electric power, 2019, 32 (1): 1-12
[4] Zhou M, Song X F, Tu J, et al. Residential electricity consumption behavior analysis based on non-intrusive load monitoring [J]. Power System Technology, 2018, 42(10): 3268-3274.
[5] Li Y, Wang B B, Li F X. Outlook and thinking of flexible and interactive utilization intelligent power[J]. Automation of Electric Power Systems, 2015, 39 (17): 2-9.
[6] Hart G W. Nonintrusive appliance load monitoring[J]. 1992, 80 6(12): 1870-1891.
[7] Shi S B, Zhang H. CUSUM transient event detection algorithm based on compound sliding window[J]. Electrical Measurement & Instrumentation, 2019, 56 (17): 13-18.
[8] Li R Y, Zhang P, Liu Y G, Zhang H, Zhou D G. Non-intrusive household load identification method based on random forest [J]. Electrical Measurement & Instrumentation, 2021, 58 (04): 9-16.
[9] Yu D W, Liu M. A non-invasive load decomposition method based on deep circular convolution model[J]. Electrical Measurement & Instrumentation, 2020, 57 (23): 47-53.
[10] Cheng X, Li L Z, Wu Hao, et al. A survey of the research on non-intrusive load monitoring and disaggregation[J]. Power System Technology, 2016 (10): 3117.
[11] Kolter J Z, Jaakkola T. Approximate inference in additive factorial HMMs with application to energy disaggregation[C], La Palma, Spain, 2012. Micromove Publishing, 2012.
[12] Kong W, Dong Z Y, Hill D J, et al. A Hierarchical Hidden Markov Model Framework for Home Appliance Modeling[J]. IEEE Transactions on Smart Grid, 2018;9; (4);4:3079-3090.
[13] Yu C, Qin Z J, Yang Y D. Non-intrusive Load Disaggregation with Improved Factorial Hidden Markov Model Considering ON-OFF Status Recognition[J]. Power System Technology, 2021, 45(11): 4540-4550.
[14] Kelly J, Knottenbelt W. Neural NILM: Deep neural networks applied to energy disaggregation[C]//, Seoul, Korea, Republic of, 2015. Association for Computing Machinery, Inc, 2015.
[15] Yan X F, Zhai S P, Wang Z H, et al. Application of deep neural network in non-intrusive load disaggregation[J]. Automation of Electric Power Systems, 2019, 43 (1): 126-132.
[16] Meng L, Yu C, Zhang X X, et al. Non-intrusive load disaggregation based on 1D convolutional neural network and self-attention mechanism[J]. Power Systems and Big Data, 2020, 23(10): 1-8.

[17] Zhao B, He K, Stankovic L, et al. Improving Event-Based Non-Intrusive Load Monitoring Using Graph Signal Processing[J]. IEEE Access, 2018, 6: 53944-53959.

[18] Wang J, Geng G, Chen K, et al. Event-Based Non-Intrusive Home Current Measurement Using Sensor Array[J]. IEEE Transactions on Smart Grid, 2018, 9(6): 5878-5886.

[19] Wu X, Han X, Liang K X. Event-based non-intrusive load identification algorithm for residential loads combined with underdetermined decomposition and characteristic filtering[J]. IET Generation Transmission & Distribution, 2019, 13(1): 99-107.

[20] Zhang X X, Bao Y Q, Shao Q, Qian Y J, Jiang D Q. Research on non-Intrusive detection of multi-Rang household appliances[J]. Electric Power Information and Communication Technology, 2021, 19 (05): 89-94.