Applied Energy 195 (2017) 634–645

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Dynamic time warping based non-intrusive load transient identification

Bo Liu [a], Wenpeng Luan [b], Yixin Yu [a,*]

[a] Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, Tianjin, China
[b] China Electric Power Research Institute, Beijing, China

HIGHLIGHTS

- DTW algorithm is used to measure the similarity of variable-length TPW time-series.
- A transient identification method using DTW-based integrated distance is proposed.
- The integrated distance is designed combining multiple types of TPW signatures.
- Comparison and field tests verify the proposed method's accuracy.
- The proposed method is easy to implement in an embedded system at a reasonable cost.

ARTICLE INFO

Article history:
Received 27 September 2016
Received in revised form 25 February 2017
Accepted 2 March 2017
Available online 27 March 2017

Keywords:
Non-intrusive load monitoring
Transient power waveform
Dynamic time warping
Nearest neighbor classification
Load transient identification

ABSTRACT

Non-intrusive load monitoring (NILM) is a novel and cost-effective technology for monitoring load electricity energy consumption details. In the event-based NILM, transient power waveform (TPW) time-series can be used as signatures to identify the transients of the electrical appliances in the aggregated load, and then to determine their operating states, estimate their power demand and cumulative energy consumption. In this paper, for load transient identification, the dynamic time warping (DTW) algorithm is adopted for the first time to measure the similarity between the variable-length raw TPW sample and template time-series. Accordingly, a nearest neighbor transient identification method is proposed to identify the appliance creating the TPW sample time-series, in which the DTW-based integrated distance is used to measure the similarity of TPW signatures. Three schemes to calculate the integrated distance are designed, combining multiple types of TPW signatures. Comparison tests with existing methods are conducted using public datasets. The comparison test results indicate that the proposed load transient identification method cannot only improve the accuracy of load transient identification, but also is easy to implement at a reasonable cost. Ultimately, the proposed method is implemented in an embedded system. The field test results show that it can identify the operating states of electrical appliances with high accuracy.

1. Introduction

Energy shortage and global warming have motivated unprecedented technology innovation and implementations on energy conservation and emission reduction. While it is projected that electricity energy will gradually become the primary form of end-use energy in modern society [1], awareness of the electricity energy consumption details is crucial for energy efficiency improvement, energy saving and emission reduction [2,3]. The mentioned electricity consumption details include the consumption information of each electrical appliance in the aggregated load, such as operating state, power demand, cumulative electricity energy, and fault diagnosis information.

The availability of electricity consumption details provides significant benefits for optimizing power grid planning, system operation and management, thus improving energy efficiency and utility assets utilization [4–10]. For example, the accuracy of power load forecasting can be improved [6], the load model used for the power system stability assessment can be more enhanced [7,8], and the demand side management can be conducted further efficiently [9,10]. The electricity consumption details can also promote end-user energy efficiency [11–18]. Research has shown that the change in user behavior due to the total electricity information feedback can achieve 5–15% energy saving [16], while comprehensive detail electricity energy consumption information online feedback and automatic energy efficiency audit can produce potential

* Corresponding author at: School of Electrical and Information Engineering, Tianjin University, No. 92 Weijin Road, Nankai District, Tianjin 300072, China.
E-mail address: yixinyu@tju.edu.cn (Y. Yu).

http://dx.doi.org/10.1016/j.apenergy.2017.03.010
0306-2619/

Nomenclature

$DTW(\boldsymbol{x}, \boldsymbol{y})$	the DTW distance between the two one-dimensional time-series $\boldsymbol{x}$ and $\boldsymbol{y}$
$DTW(\boldsymbol{X}, \boldsymbol{Y})$	the DTW distance between the two multi-dimensional time-series $\boldsymbol{X}$ and $\boldsymbol{Y}$
$\boldsymbol{P}_{\boldsymbol{z}}^{t}$	the active TPW template time-series of any known electrical appliance transient $\boldsymbol{z}$
$\boldsymbol{Q}_{\boldsymbol{z}}^{t}$	the reactive TPW template time-series of any known electrical appliance transient $\boldsymbol{z}$
$\boldsymbol{P}_{\boldsymbol{e}}^{s}$	the active TPW sample time-series of unknown load transient $\boldsymbol{e}$
$\boldsymbol{Q}_{\boldsymbol{e}}^{s}$	the reactive TPW sample time-series of unknown load transient $\boldsymbol{e}$
$\boldsymbol{T}_{\boldsymbol{z}}$	the TPW template representing any known electrical appliance transient $\boldsymbol{z}$
$\boldsymbol{T}_{\boldsymbol{e}}$	the TPW sample representing the unknown load transient $\boldsymbol{e}$
$D_{\boldsymbol{e},\boldsymbol{z}}(\boldsymbol{T}_{\boldsymbol{e}}, \boldsymbol{T}_{\boldsymbol{z}})$	the integrated distance between $\boldsymbol{T}_{\boldsymbol{e}}$ and $\boldsymbol{T}_{\boldsymbol{z}}$
$\widehat{\boldsymbol{T}}_{\boldsymbol{z}}$	the TPW template that is the most similar to the TPW sample $\boldsymbol{T}_{\boldsymbol{e}}$
$\boldsymbol{TR}$	the set of TPW templates
Ω	the set of electrical appliance transient classes
c	an electrical appliance transient class
TP_c	the number of samples that are correctly recognized as class c
FP_c	the number of samples that are incorrectly recognized as class c
FN_c	the number of class c samples that are incorrectly recognized as other classes
P_c	*Precision*, the proportion of samples recognized as class c that are correct
S_c	*Sensitivity*, the proportion of all the samples belonging to class c that are correctly recognized
F_c	*F*-measure, a composite index combining the precision and sensitivity

energy saving of approximately 20–35% [17]. According to the project conducted by Bidgely in California, personalized energy efficiency service based on the NILM can contribute to average 14% home electricity energy reduction [18].

NILM is a novel and cost-effective technique for monitoring the load electricity consumption details, which is derived through analysis of the acquired terminal voltage and aggregated load current using pattern recognition technologies and machine learning algorithms. For an NILM system, only one sensor is installed at the electricity service entry. Thus, NILM has the advantages of low cost, high reliability, good data integrity, and easy installation, thus can be implemented quickly [3,4].

Research on NILM has been conducted for nearly 30 years [19]. Based on the electrical appliance operating state monitoring strategy, NILM approaches can be classified into three main categories. The first category implements NILM by classifying and identifying the load-event-related signature patterns [19–32]. It outputs the operating state transition of an appliance, such as turning on or turning off. Since this kind of methods is based on load event detection, they are often referred to as event-based NILM. The second category directly disaggregates the composite signature of the aggregated load through optimization decomposition or pattern recognition [33–38]. It directly determines whether an electrical appliance is in the On or Off state during a certain period. Because this type of methods does not rely on analyzing load events, they are often referred to as non-event-based NILM. The third category integrates the technical ideas of the two previous approaches, and can thus be called hybrid NILM [39–42]. It is noted that, in this paper, the sudden transition between the operating states of an electrical appliance is called transient (or transient process), which corresponds to a load event, so these two terms are used interchangeably.

For the event-based NILM, the transient power waveform (TPW) signatures have been successfully used to monitor the operating states of electrical appliances through transient identification [20–27], especially to distinguish different appliances with similar steady-state signatures [28,29]. On one hand, the operating principles of various types of appliances are different, so there are obvious distinctions among their transient power waveforms (TPWs). For instance, the physical processes of igniting an arc and accelerating a rotor are fundamentally different, so the turn-on transients of a fluorescent lamp and an induction motor are significantly different [20]. ***Namely, the TPW signatures are distinct***. On the other hand, theoretically, each electrical appliance inherently has its deterministic transient pattern, there is a one-to-one correspondence between its TPW and load event (e.g., turn-on or turn-off event). Therefore, the same type of electrical appliance load events can generate repeatedly observable TPWs. ***That is, the TPW signatures are repeatable***. In other words, the TPW can be taken as the signature to describe the electrical appliance transient.

For the load transient identification based on TPW signatures, [20] determined the matching degree between two standardized TPW time-series based on their inner product, while [21] proposed an approaching degree index to measure their similarity, which is based on the fuzzy inner and outer products between two normalized TPW time-series. [22] proposed a load transient identification method based on the optimal matching of two TPWs, and labelled the TPW sample time-series with the appliance type associated with the template time-series that has the minimum least-squares fitting residual with the sample time-series. [23,24] used a linear regression model with appropriately selected basis functions (e.g., Fourier basis functions) to fit the fixed-length raw TPW time-series extracted around the occurrence time of the load event. Then the feature vector composed of the fitting coefficients was input into the pre-prepared classifier. In addition, [24] also presented the results of a transient classification experiment in which the *Minkowski* distance was used to measure the similarity between the fixed-length raw TPW time-series. [25] proposed the use of a piecewise smooth transition regression model to fit the raw TPW time-series. Considering the variability of feature parameters to fit the TPWs within the same type of appliance transients, the hierarchical Bayesian network was applied to establish a probabilistic model to calculate the possibility of the occurrence of different TPW sample time-series. Moreover, the Bayes factor calculated in terms of naive Bayes theory was used as a discriminant index to complete the transient identification. Furthermore, to avoid direct handling of the complexity of TPWs in the time domain, [26,27] separately applied time–frequency analysis techniques (e.g., Short Time Fourier Transform (STFT) and Wavelet Transform (WT)) to the TPWs to derive the signature vector for transient identification. However, there exist the following problems:

First, under the impact of various factors, temporal location shifting and local scaling often occur in TPW time-series. As a