

Research Article

A Novel Feature Extraction Method for Nonintrusive Appliance Load Monitoring

Khaled Chahine¹ and Khalil El Khamlichi Drissi²

¹ Department of Electrical and Electronics Engineering, Lebanese International University, Mazraa, Beirut 146404, Lebanon

² Pascal Institute, UMR 6602, 24 Avenue des Landais, 63177 Aubière Cedex, France

Correspondence should be addressed to Khaled Chahine; khaled.chahine@liu.edu.lb

Received 30 March 2013; Accepted 12 April 2013

Academic Editor: Baoding Liu

Copyright © 2013 K. Chahine and K. El Khamlichi Drissi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Improving energy efficiency by monitoring household electrical consumption is of significant importance with the climate change concerns of the present time. A solution for the electrical consumption management problem is the use of a nonintrusive appliance load monitoring (NIALM) system. This system captures the signals from the aggregate consumption, extracts the features from these signals and classifies the extracted features in order to identify the switched-on appliances. This paper focuses solely on feature extraction through applying the matrix pencil method, a well-known parametric estimation technique, to the drawn electric current. The result is a compact representation of the current signal in terms of complex numbers referred to as poles and residues. These complex numbers are shown to be characteristic of the considered load and can thus serve as features in any subsequent classification module. In the absence of noise, simulations indicate an almost perfect agreement between theoretical and estimated values of poles and residues. For real data, poles and residues are used to determine a feature vector consisting of the contribution of the fundamental, the third, and the fifth harmonic currents to the maximum of the total load current. The result is a three-dimensional feature space with reduced intercluster overlap.

1. Introduction

The reason behind the drive for the installation of smart meters in homes and businesses is that they facilitate for consumers to monitor their energy consumption, thereby making it easier for them to save energy, carbon emissions, and money. To help customers as well as utilities in the monitoring process, researchers have been studying load disaggregation schemes for more than two decades.

One method of load disaggregation is distributed direct sensing which requires a sensor at each device or appliance in order to measure consumption. The one-sensor-per-device requirement is both the blessing and the curse of this method, for it is highly accurate but expensive. To overcome the limitations associated with the direct sensing approach, researchers have explored methods to infer disaggregated energy usage via a single sensor. Pioneering work in this area is non intrusive appliance load monitoring (NIALM), first introduced by Hart in the late 1980s [1]. In contrast to the direct sensing

methods, NIALM relies solely on single-point measurements of voltage and current on the power feed entering the household. NIALM consists of four steps: data acquisition, event detection, feature extraction, and event classification. The raw current and voltage waveforms are transformed into a feature vector, that is, a more compact and meaningful representation that may include real power, reactive power, current-voltage phase difference, and harmonics (e.g., [2]). These extracted features are monitored for changes, identified as events (e.g., an appliance turning “on” or “off”), and classified down to the appliance or device category level using a classification algorithm, which usually compares the features to a preexisting database of signatures. Several reviews of feature extraction methods for electric loads in residential and commercial buildings can be found in the literature [3, 4].

Based on the degree of nonintrusiveness, the literature draws a distinction between manual-setup NIALM (MS-NIALM) and automatic-setup NIALM (AS-NIALM) systems. While the former requires switching individual appliances on

and off manually to learn their signatures, the latter sets itself up using prior information about potential appliances. AS-NIALM hence extracts the signatures and labels them without any sort of manual intervention which would greatly facilitate mass installation of smart meters. To the authors' knowledge, no AS-NIALM system has hitherto been implemented. It is hence the main goal of this work to pave the way for such a solution.

In this paper, the matrix pencil method, a well-known parametric estimation technique, is applied to the electric current drawn by some elementary linear and nonlinear electric loads driven by a sinusoidal voltage source as well as real loads. The result is a compact representation of the current in terms of complex numbers referred to as poles and residues [5, 6]. These complex numbers are shown to be characteristic of the considered load and thus can serve as features for the subsequent classification phase [7]. For both synthetic and real data, results indicate that poles and residues extracted by the MPM allow an almost perfect reconstruction of drawn electric currents. Results obtained from a database of a household indicate that the extracted features succeed in reducing the intercluster overlap of different appliances.

The objectives of this paper are summarized in the following two points:

- (1) show that the reduced number of poles and residues estimated by MPM enable an accurate reconstruction of synthetic and real signals;
- (2) show that the fundamental and higher harmonic currents determined from poles and residues yield a feature space with reduced intercluster overlap.

The rest of the paper is organized as follows. Section 2 presents the signal model and the principle of the MPM. Sections 3 and 4 show the validation on simulated and real data, respectively. Finally, Section 5 provides the summary and conclusion.

2. Feature Extraction

2.1. Signal Model. For a sinusoidal driving voltage of the form $v(t) = V\sqrt{2}\sin(\omega t)$, the drawn electric current can be modeled as a linear combination of d cisoids (complex-valued sinusoidal signals) weighted by complex residues according to the following signal model:

$$i(t) \approx \sum_{m=1}^d r_m \exp\{(\alpha_m + j2\pi f_m)t\} + b(t), \quad (1)$$

where r_m is the residue of the m th cisoid, α_m is its attenuation factor, f_m is its frequency, and $b(t)$ is additive white Gaussian noise. After sampling, the time variable, t , is replaced by $t_k = kt_s$, where $t_s = 6.25 \times 10^{-4}$ is the chosen sampling period. The discrete current signal becomes

$$i(k) \approx \sum_{m=1}^d r_m z_m^k + b(k) \quad k = 1, 2, \dots, N, \quad (2)$$

where

$$z_m = \exp\{(\alpha_m + j2\pi f_m)t_s\} \quad m = 1, 2, \dots, d \quad (3)$$

is the m th complex pole. Under matrix form, the signal model is expressed by

$$\mathbf{i} = \mathbf{A}\mathbf{r} + \mathbf{b} \quad (4)$$

with the following notational definitions:

$$\begin{aligned} \mathbf{i} &= [i(1) \ i(2) \ \dots \ i(N)]^T, \\ \mathbf{A} &= [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_d], \\ \mathbf{a}_m &= [z_m \ z_m^2 \ \dots \ z_m^N]^T, \\ \mathbf{r} &= [r_1 \ r_2 \ \dots \ r_d]^T, \\ \mathbf{b} &= [b(1) \ b(2) \ \dots \ b(N)]^T. \end{aligned} \quad (5)$$

The superscript T denotes the transpose operator.

The feature extraction problem can now be stated as follows. Given the electric current data sequence $\{i(k)\}_{k=1}^N$, use a feature extraction method to extract the complex poles $\{z_m\}_{m=1}^d$ and residues $\{r_m\}_{m=1}^d$ of the load.

2.2. Matrix Pencil Method (MPM). This section briefly recalls the principle of MPM which is a linear prediction method tailored to the parameter estimation of the damped/undamped exponential model. Starting from the signal model given in (1), MPM chooses a free parameter, L , known as the pencil parameter such as $d \leq L \leq N - d$. The proper choice of L results in significant robustness against noise. The next step is to construct a Hankel data matrix:

$$\mathbf{H} = \begin{bmatrix} i(1) & i(2) & \dots & i(L+1) \\ i(2) & i(3) & \dots & i(L+2) \\ \vdots & \vdots & \ddots & \vdots \\ i(N-L) & i(N-L+1) & \dots & i(N) \end{bmatrix}. \quad (6)$$

Two matrices are then obtained by removing the last and first columns of \mathbf{H} . In MATLAB notation, they are given as follows:

$$\begin{aligned} \overrightarrow{\mathbf{H}} &= \mathbf{H}(:, 1:L), \\ \overleftarrow{\mathbf{H}} &= \mathbf{H}(:, 2:L+1). \end{aligned} \quad (7)$$

The matrix pencil for the two matrices $\overrightarrow{\mathbf{H}}$ and $\overleftarrow{\mathbf{H}}$ is defined as their linear combination $\overleftarrow{\mathbf{H}} - \lambda \overrightarrow{\mathbf{H}}$, with λ a scalar parameter. In the absence of noise and owing to the assumed signal model, it is easily verified that $\overrightarrow{\mathbf{H}}$ and $\overleftarrow{\mathbf{H}}$ admit the following Vandermonde decomposition:

$$\begin{aligned} \overrightarrow{\mathbf{H}} &= \mathbf{Z}_1 \mathbf{R} \mathbf{Z}_2, \\ \overleftarrow{\mathbf{H}} &= \mathbf{Z}_1 \mathbf{R} \mathbf{Z}_0 \mathbf{Z}_2, \end{aligned} \quad (8)$$

where

$$\mathbf{Z}_1 = \begin{bmatrix} z_1 & z_2 & \cdots & z_d \\ z_1^2 & z_2^2 & \cdots & z_d^2 \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{N-L} & z_2^{N-L} & \cdots & z_d^{N-L} \end{bmatrix},$$

$$\mathbf{Z}_2 = \begin{bmatrix} 1 & z_1 & \cdots & z_1^{L-1} \\ 1 & z_2 & \cdots & z_2^{L-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_d & \cdots & z_d^{L-1} \end{bmatrix}, \quad (9)$$

$$\mathbf{Z}_0 = \text{diag} \{z_1, z_2, \dots, z_d\},$$

$$\mathbf{R} = \text{diag} \{r_1, r_2, \dots, r_d\},$$

revealing the fundamental shift-invariance property in the column and row spaces. The matrix pencil can then be written as

$$\underline{\mathbf{H}} - \lambda \underline{\mathbf{H}} = \mathbf{Z}_1 \mathbf{R} [\mathbf{Z}_0 - \lambda \mathbf{I}] \mathbf{Z}_2, \quad (10)$$

where \mathbf{I} is the identity matrix. Hence, each value of $\lambda = z_m$ is a rank-reducing number of the pencil. The estimates of z_m are, therefore, the generalized eigenvalues (GEVs) of the matrix pair $[\underline{\mathbf{H}}, \underline{\mathbf{H}}]$.

Once the complex poles $\{z_m\}_{m=1}^d$ are determined, the complex residues can be estimated using a least squares fit having the following solution:

$$\mathbf{r} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{i}. \quad (11)$$

For noisy data, total least squares matrix pencil method (TLSMPM) is usually preferred in which the singular value decomposition is used to prefilter the complex signals, and then conventional procedures follow. For more details, the reader can refer to [8].

3. Validation on Synthetic Data

3.1. Linear Loads. To validate MPM as a feature extraction method, we shall first compare its poles and residues with those obtained from the theoretical expressions of the following linear elementary loads: series RC, series RL, parallel RL, and series RLC. The RC and RL circuits lead to first order differential equations in time whereas the RLC circuit leads to a second-order differential equation. Using Euler's formula and rearranging allow rewriting the current expression obtained from the solution of the differential equation characterizing the load in the form of (1). The poles and residues of each elementary load can then be readily identified. Tables 1, 2, 3, and 4 give the residues, attenuation factors, and frequencies of the four studied elementary loads. As can be seen from these tables, first-order circuits (RL and RC) are characterized by two pure imaginary conjugate poles representing their forced response and one real pole representing their natural response, whereas the second-order circuit (RLC) has, besides the two pure imaginary conjugate poles of its forced

TABLE 1: The residues, attenuation factors, and frequencies of the series RC load.

m	r_m	α_m	f_m (Hz)
1	$\frac{V\sqrt{2}}{2jR} \cos(\phi) e^{-j\phi}$	0	+50
2	$-\frac{V\sqrt{2}}{2jR} \cos(\phi) e^{j\phi}$	0	-50
3	$-\frac{1}{R} (v_{C_0} - V\sqrt{2} \sin(\phi) \cos(\omega t_0 - \phi)) e^{t_0/\tau}$	$-\frac{1}{\tau}$	0

TABLE 2: The residues, attenuation factors, and frequencies of the series RL load.

m	r_m	α_m	f_m (Hz)
1	$\frac{V\sqrt{2}}{2jR} \cos(\phi) e^{-j\phi}$	0	+50
2	$-\frac{V\sqrt{2}}{2jR} \cos(\phi) e^{j\phi}$	0	-50
3	$i_{L_0} - \frac{V\sqrt{2}}{R} \cos(\phi) \sin(\omega t_0 - \phi) e^{t_0/\tau}$	$-\frac{1}{\tau}$	0

TABLE 3: The residues, attenuation factors, and frequencies of the parallel RL load.

m	r_m	α_m	f_m (Hz)
1	$\frac{V\sqrt{2}}{2jR \cos(\phi)} e^{-j\phi}$	0	+50
2	$-\frac{V\sqrt{2}}{2jR \cos(\phi)} e^{j\phi}$	0	-50
3	$i_{L_0} - \frac{V\sqrt{2}}{R \cos(\phi)} \sin(\omega t_0 - \phi)$	0	0

TABLE 4: The residues, attenuation factors, and frequencies of the series RLC load.

m	r_m	α_m	f_m (Hz)
1	$\frac{V\sqrt{2}}{2jR} \cos(\phi) e^{-j\phi}$	0	+50
2	$-\frac{V\sqrt{2}}{2jR} \cos(\phi) e^{j\phi}$	0	-50
3	$Ae^{-k_1 \omega_0 t_0}$	$k_1 \omega_0$	0
4	$Be^{-k_2 \omega_0 t_0}$	$k_2 \omega_0$	0

response, two conjugate complex poles related to its natural response. The expressions of the dependent parameters are given in the appendix.

3.2. Nonlinear Loads. A nonlinear load is one for which the relationship between the current through the load and the voltage across the load is a nonlinear function. A simple view of the nature of nonlinear loads can be presented using Ohm's Law, which states that the voltage is the product of the load resistance and the current ($V = RI$). For a linear load, the resistance (R) is a constant; for a nonlinear load, the resistance varies. When AC power is supplied to a nonlinear load,

TABLE 5: Current composition of a nonlinear load.

I_1	I_5	I_7	I_{11}	I_{13}
100%	18.9%	11%	5.9%	4.8%

the result is the creation of currents that do not oscillate at the supply frequency. These currents are called harmonics. Harmonics occur at multiples of the supply (fundamental) frequency. For instance, if the fundamental frequency is 50 Hz, the so-called second harmonic is 100 Hz, the third harmonic is 150 Hz, and so on. Any number of harmonics can be created by a particular piece of equipment depending on that equipment's electrical characteristics. Therefore, the current drawn by nonlinear loads can still be represented by (1) where harmonics appear in the form of pole-residue couples at frequency multiples of 50 Hz.

3.3. Results. Assuming zero initial conditions ($i_{L_0} = 0$ and/or $v_{C_0} = 0$), the following numerical values were used to determine the electric current data sequence from which MPM extracted poles and residues: $\{R = 100 \Omega, C = 0.1 \text{ mF}\}$ for the series RC circuit, $\{R = 10 \Omega, L = 100 \text{ mH}\}$ for both the series and parallel RL circuits, and $\{R = 1 \Omega, L = 20 \text{ mH}, C = 60 \text{ mF}\}$ for the series RLC circuit. A duration of ten periods or 0.2 seconds was chosen for the current which at $t_s = 6.25 \times 10^{-4}$ is equivalent to 320 samples, and MPM was applied at each period. Figures 1, 2, 3, and 4 show the current obtained from the analytic expression of poles and residues in the tables above and its reconstruction obtained from the poles and residues extracted by MPM. An almost perfect agreement can be seen between the two curves indicating the accuracy of the characteristic complex numbers extracted by MPM. In addition, the figures show the forced and natural responses of each of the four elementary circuits.

To evaluate the performance of MPM on nonlinear loads, we considered the current shown in Table 5. It consists of a fundamental and four harmonics and hence can be represented by ten pairwise complex conjugate pole-residue couples. We then used MPM to extract these ten couples which served to reconstruct the current as shown in Figure 5. As can be seen, MPM is successful in estimating the pole-residue couples of the load.

4. Validation on Real Data

4.1. Reconstruction Results. In this section, the validation of MPM is carried out on currents of three representative loads: a television set, a vacuum cleaner, and an economy lamp. As for the case of synthetic data, MPM was applied at each period. Figures 6, 7, and 8 show the current drawn by the appliances and its reconstruction based on the pole-residue estimates of MPM. The close agreement shown in the figures indicate that the exponential model of (1) and its parameters estimated by the MPM accurately predict the response of the actual loads. It is worth mentioning that the number of pole-residue couples d increases with the nonlinearity of the load. For instance, the current of the vacuum cleaner could be accurately reconstructed from four pole-residue couples,

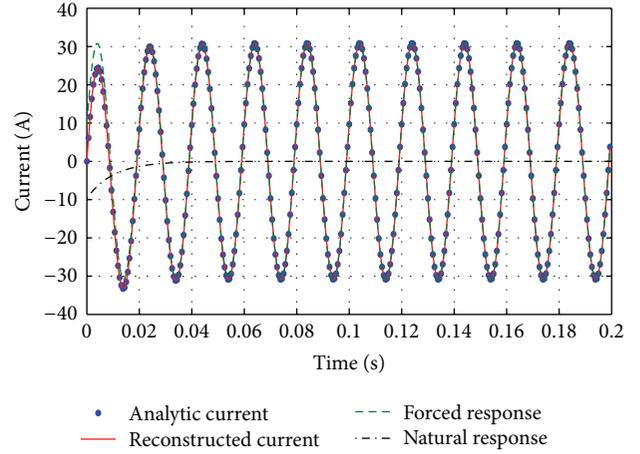


FIGURE 1: The analytic and reconstructed currents of the series RC circuit along with its forced and natural responses.

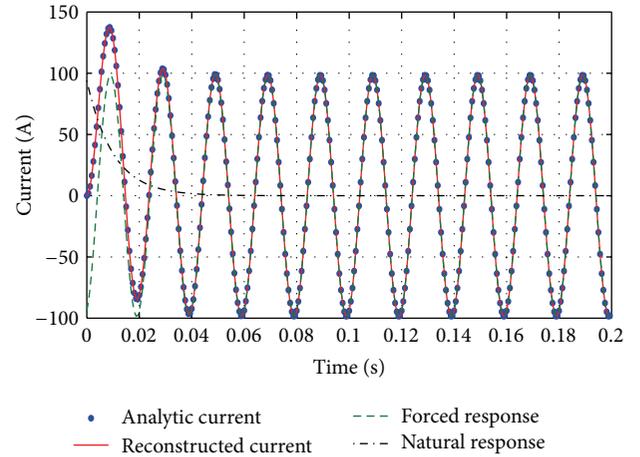


FIGURE 2: The analytic and reconstructed currents of the series RL circuit along with its forced and natural responses.

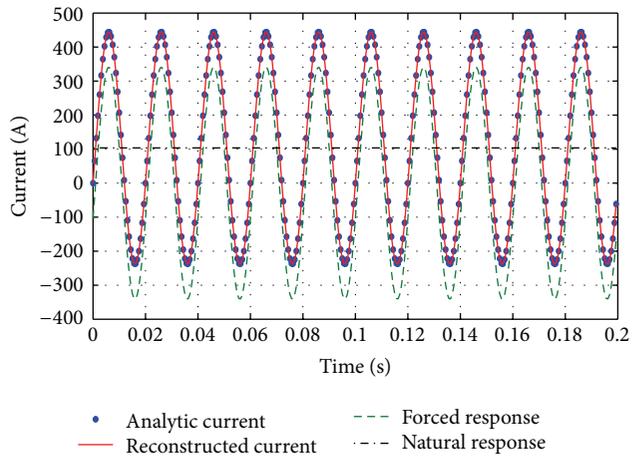


FIGURE 3: The analytic and reconstructed currents of the parallel RL circuit along with its forced and natural responses.

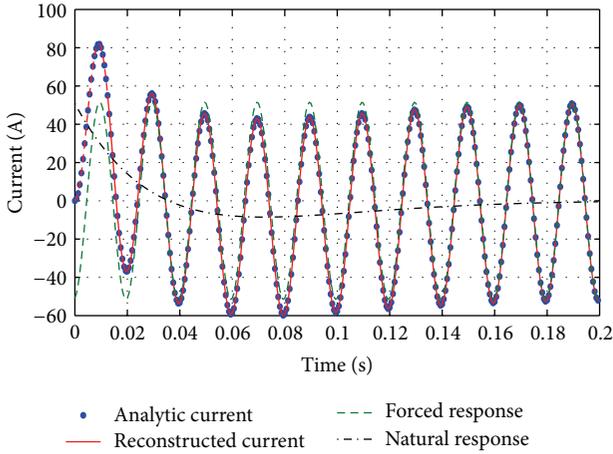


FIGURE 4: The analytic and reconstructed currents of the series RLC circuit along with its forced and natural responses.

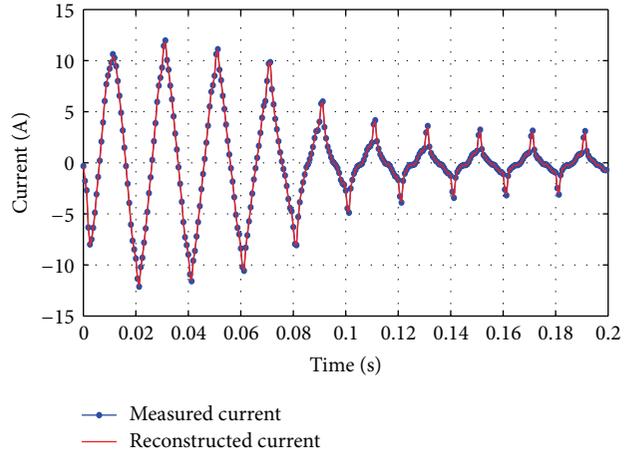


FIGURE 6: The measured and reconstructed currents of the television set.

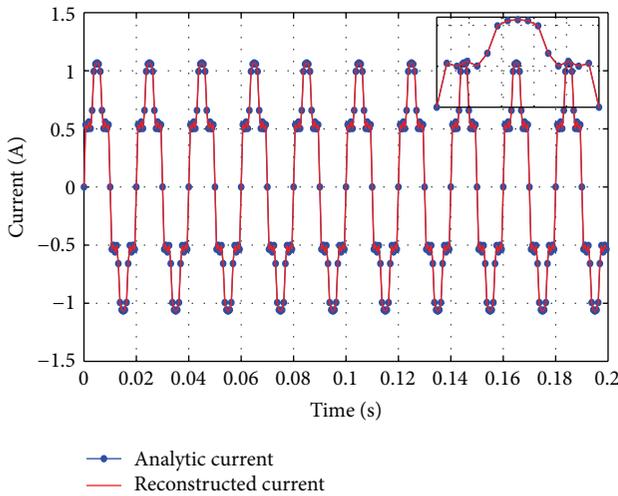


FIGURE 5: The analytic and reconstructed currents of a nonlinear load. The inset zooms in on a half period of the drawn current in order to show the accuracy of reconstruction.

whereas that of the economy lamp needed up to twelve couples.

4.2. Feature Space. The feature space contains 900 signatures uniformly distributed among the following nine appliances: incandescent lamp, halogen lamp, economy lamp, water heater, electric convector, oven, two-burner hot plate, television set, and computer. As shown in Figure 9, each signature (represented by a point in the three-dimensional feature space) is characterized by three pole-residue products corresponding to the maxima of the fundamental, third, and fifth harmonic currents. The restriction to three frequencies has the sole aim of representing the feature space graphically. From the feature space, ten clusters representing the nine appliances can be clearly distinguished. The additional cluster is due to the two-burner hot plate which is represented by two clusters, one for each burner. It can hence be concluded that

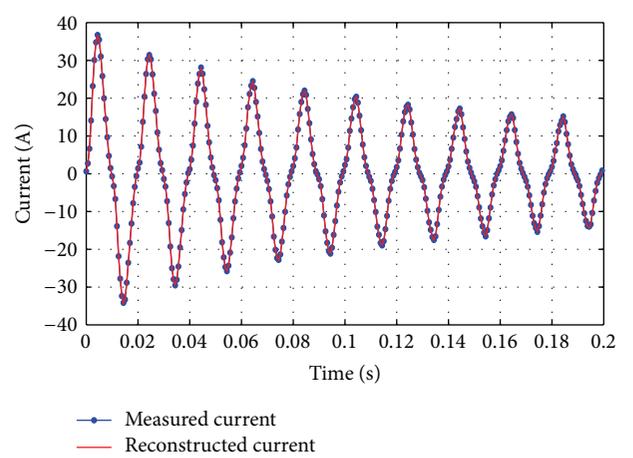


FIGURE 7: The measured and reconstructed currents of the vacuum cleaner.

the studied appliances can be fairly distinguished using the fundamental and higher harmonics.

5. Conclusion

This paper presented a novel feature extraction method for non intrusive appliance load monitoring. First, the poles and residues estimated by the matrix pencil method were shown to enable accurate reconstruction of synthetic and real current signals. Second, these complex numbers were used to determine a three-dimensional feature space with reduced intercluster overlap. Future research will make use of the extracted features for the classification phase.

Appendix

The dependent parameters of first-order circuits are given in Table 6, where τ is the time constant and ϕ is the phase angle.

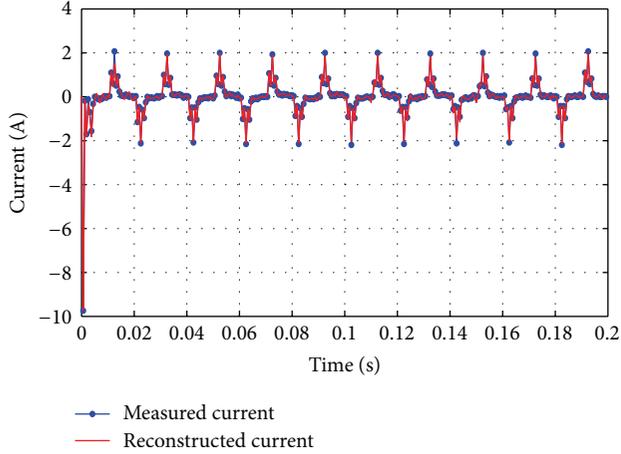


FIGURE 8: The measured and reconstructed currents of the economy lamp.

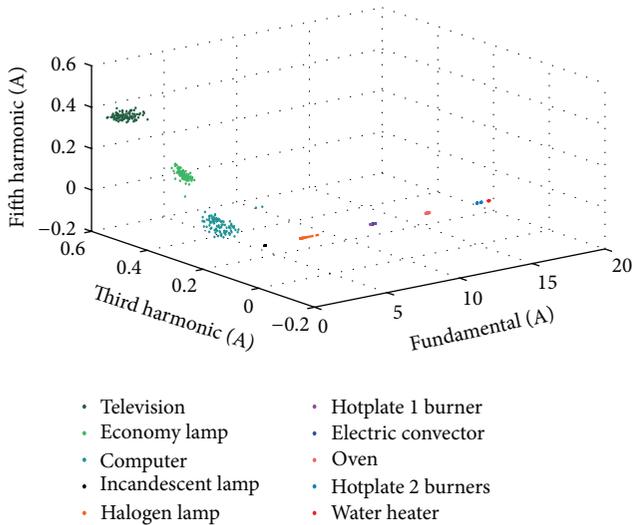


FIGURE 9: The feature space showing the disaggregated contribution of the fundamental and harmonic currents to the maximum of the total current for several appliances.

In addition to the phase angle ϕ , the dependent parameters of the series RLC circuit include the resonance angular frequency ω_0 and the damping factor ξ :

$$\begin{aligned}\phi &= \arctan \left[\frac{1}{R} \left(L\omega - \frac{1}{C\omega} \right) \right] \text{ [rad]}, \\ \omega_0 &= \frac{1}{\sqrt{LC}} \text{ [rad/s]}, \\ \xi &= \frac{R}{2} \sqrt{\frac{C}{L}} \text{ [Np]}.\end{aligned}\tag{A.1}$$

TABLE 6: Dependent parameters of first-order circuits.

Load	τ [s]	ϕ [rad]
Series RC	RC	$\arctan \left(\frac{-1}{RC\omega} \right)$
Series RL	$\frac{L}{R}$	$\arctan \left(\frac{L\omega}{R} \right)$
Parallel RL	—	$\arctan \left(\frac{R}{L\omega} \right)$

The roots of the characteristic equation of the second-order differential equation, k_1 and k_2 , are expressed in terms of the the damping factor ξ as follows:

$$\begin{aligned}k_1 &= -\xi - \sqrt{\xi^2 - 1}, \\ k_2 &= -\xi + \sqrt{\xi^2 - 1}.\end{aligned}\tag{A.2}$$

Finally, A and B , shown in Table 4, are expressed as follows:

$$\begin{aligned}A &= \frac{k_1}{k_1 - k_2} (A_2 - k_2 A_1), \\ B &= \frac{k_2}{k_2 - k_1} (A_2 - k_1 A_1),\end{aligned}\tag{A.3}$$

where

$$\begin{aligned}A_1 &= v_{c0} \frac{C}{L} + \frac{V\sqrt{2}}{R} \cos(\phi) \cos(\omega t_0 - \phi) \frac{\omega_0}{\omega}, \\ A_2 &= i_{L0} - \frac{V\sqrt{2}}{R} \cos(\phi) \sin(\omega t_0 - \phi).\end{aligned}\tag{A.4}$$

Acknowledgment

This work was supported in part by Landis+Gyr.

References

- [1] G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.
- [2] C. Laughman, K. Lee, R. Cox et al., "Power signature analysis," *IEEE Power and Energy Magazine*, vol. 1, no. 2, pp. 56–63, 2003.
- [3] Y. Du, L. Du, B. Lu, R. Harley, and T. Habetler, "A review of identification and monitoring methods for electric loads in commercial and residential buildings," in *Proceedings of the 2nd IEEE Energy Conversion Congress and Exposition (ECCE '10)*, pp. 4527–4533, Atlanta, Ga, USA, September 2010.
- [4] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: review and outlook," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 1, pp. 76–84, 2011.
- [5] H. Najmeddine, K. E. K. Drissi, C. Pasquier et al., "Smart Metering by using matrix pencil," in *Proceedings of the 9th International Conference on Environment and Electrical Engineering (EEEIC '10)*, pp. 238–241, Prague, Czech Republic, May 2010.
- [6] H. Najmeddine, K. El Khamlichi Drissi, A. Diop, and T. Jouanet, "Method and device for the non-intrusive determination of the electrical power consumed by an installation, by analysing load transients," French Patent, FR 0856717, October 2008.

- [7] K. Chahine, K. El Khamlichi Drissi, C. Pasquier et al., "Electric load disaggregation in smart metering using a novel feature extraction method and supervised classification," *Energy Procedia*, vol. 6, pp. 627–632, 2011.
- [8] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 5, pp. 814–824, 1990.




Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

