Load Disaggregation Using Microscopic Power Features and Pattern Recognition

Wesley Angelino de Souza 1, Fernando Deluno Garcia 2, Fernando Pinhabel Marafão 2,* , Luiz Carlos Pereira da Silva 3 and Marcelo Godoy Simões 4,*

1 Department of Computer Science, Federal University of São Carlos (UFSCar), Sorocaba SP 18052-780, Brazil
2 Institute of Science and Technology of Sorocaba, São Paulo State University (UNESP), Sorocaba SP 18087-180, Brazil
3 School of Electrical and Computer Engineering (FEEC), University of Campinas (UNICAMP), Campinas SP 13083-852, Brazil
4 Department of Electrical Engineering, Colorado School of Mines, Golden, CO 80401, USA
* Correspondence: fernando.marafao@unesp.br (F.P.M.); msimoes@mines.edu (M.G.S.)

Received: 29 May 2019; Accepted: 5 July 2019; Published: 10 July 2019

Abstract: A new generation of smart meters are called cognitive meters, which are essentially based on Artificial Intelligence (AI) and load disaggregation methods for Non-Intrusive Load Monitoring (NILM). Thus, modern NILM may recognize appliances connected to the grid during certain periods, while providing much more information than the traditional monthly consumption. Therefore, this article presents a new load disaggregation methodology with microscopic characteristics collected from current and voltage waveforms. Initially, the novel NILM algorithm—called the Power Signature Blob (PSB)—makes use of a state machine to detect when the appliance has been turned on or off. Then, machine learning is used to identify the appliance, for which attributes are extracted from the Conservative Power Theory (CPT), a contemporary power theory that enables comprehensive load modeling. Finally, considering simulation and experimental results, this paper shows that the new method is able to achieve 95% accuracy considering the applied data set.

Keywords: load disaggregation; artificial intelligence; cognitive meters; machine learning; state machine; NILM

1. Introduction

Electricity metering has undergone significant technological progress over the last 30 years, from electromechanical to electronic metering. An essential stage of this evolution arose with Automatic Meter Reading (AMR) [1], which includes the following main features [2,3]:

- Improved accuracy in terms of energy readings when compared to the preceding electromechanical meters, in which the measuring errors were quite susceptible and dependent on human operators;
- Automatic and frequent meter readings;
- Customer support improvement;
- More consumption information;
- Support to hourly price charges.

Moreover, the AMR was the basis for the succeeding evolution called Advanced Metering Infrastructure (AMI), which includes the following characteristics [4–6]:

- Bidirectional communication;
- A power consumption monitoring system;
Advanced and accurate sensors;

The embedded system is responsible for data collection and manages the required information between meter and utility.

AMI is related to the entire metering infrastructure and ‘smart meter’ is the popular name for the power metering device in this infrastructure. The term “smart” makes sense in the data processing approach, that is, the meter might process input data (voltage and current), transforming it into useful output information (e.g., energy consumption, power quality indicators, efficiency, and others). However, the concept of “smart” is not usually well defined, since most of the smart meters on the market do not have any smart functionality. Thus, there is a current demand for innovative and intelligent techniques to provide different sorts of information to utilities and consumers, improving their knowledge about energy use, efficiency, costs, consequently improving energy management. Therefore, researchers all over the world are proposing new tools and methods to provide further information about energy consumption [7,8], as well as proposing innovative ways to save energy [9–12].

In this context, a new generation of smart meters are called cognitive meters, which propose to use artificial intelligence and load disaggregation methods, also known as Non-Intrusive Load Monitoring (NILM). They recognize appliances connected to the grid during certain periods, while providing much more information to consumers than the traditional monthly consumption. Consequently, consumers’ operations can be inspected to provide them with detailed information about their electrical consumption [13] so they can make better decisions concerning saving electricity, as well as implementing energy management systems for automatic generation/consumption regulation. This is certainly a meaningful advance concerning the relationship between utilities and consumers [11–13]. Moreover, “c-meters” can provide advanced functionalities, such as detailed recommendation of when to use some particular appliances (according to statistical behavior, real-time consumption or hourly-energy price) and they can also suggest some tip to save electricity over weeks, months and years.

Hart [14] initially introduced the NILM method, considering active power levels and distributing them into individual appliance data. With such a type of cognition, the consumer profile can be mapped and by using artificial intelligence techniques, new methodologies can be proposed for modern smart meters [7,8,10,14–22]. However, although NILM is quite a good approach to detect home appliances, it does not always present a reasonable accuracy, demanding other signal analysis or AI to improve detection accuracy.

Therefore, this paper introduces the Power Signature Blob (PSB)—a novel methodology that correlates a hybrid load disaggregation technique, which is based on feature extraction from current and voltage waveforms with power signatures. In this context, a predefined threshold level is compared to the active power variation, in addition to the power signatures. The procedure uses the difference between the actual active power and the last active power value used to define the step level direction—when the appliance is turned on or turned off. Considering that every step level detection is a new event, the novel NILM method calculates the proper features to classify the load and then, applies machine learning for appliance recognition. For classification, the NILM dataset from [15] was used, including instances of 35 appliances (In this paper, the term sample is defined as a signal acquisition from the voltage and current waveforms, and the term instance represents a dataset sample with the respective features.). Each appliance instance comprises active power and other power components (power factor, reactivity factor, and distortion factor). The features to classify and from the dataset are calculated using a contemporary power theory called Conservative Power Theory (CPT) [23,24]. Hence, the novelty of this paper is a new state-machine NILM that uses and acknowledges a dataset of 35 appliances with features based on power components by the CPT.

The next section discusses the main concept of load disaggregation and different techniques from the literature. Afterwards, the PSB method, using the power decompositions from CPT [23–25], is presented. Finally, simulations and experimental results depict the performance of the proposed approach.
2. NILM State-of-the-Art Review

Considering current and voltage acquisition and processing, the literature on NILM can be divided into two main categories—“high-frequency” and “low-frequency”. The low-frequency is the category in which the features are extracted at 1 kHz or less. The high-frequency is the category in which features are extracted at kHz or MHz [15,19]. Table 1 presents a list of NILM studies and the features extracted from the point of view of the attributes.

### Table 1. Summary and categorization of some Non–Intrusive Load Monitoring (NILM) techniques—relevant works and their main features.

<table>
<thead>
<tr>
<th>Sampling</th>
<th>Signal Sampling Technique</th>
<th>Features (Attributes)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low-frequency</td>
<td>Active/reactive power signature and variation</td>
<td>P, Q</td>
<td>[14,26–30]</td>
</tr>
<tr>
<td>(Macroscopic)</td>
<td>Active power signature and variation</td>
<td>p</td>
<td>[31–34]</td>
</tr>
<tr>
<td></td>
<td>Power signature and power indicators</td>
<td>p, I_{RMS}, U_{RMS}, I_{MAX}, U_{MAX}, PF(λ)</td>
<td>[35,36]</td>
</tr>
<tr>
<td></td>
<td>Load on or off probability</td>
<td>p_{ON}, p_{OFF}</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>Temporal discrete power pulses</td>
<td>P_{pulse}(t)</td>
<td>[38]</td>
</tr>
<tr>
<td></td>
<td>Harmonic decomposition</td>
<td></td>
<td>[39]</td>
</tr>
<tr>
<td></td>
<td>Power signature and harmonic decomposition</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Power signatures and Power Theory</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>W_{coeff}, W_{c}</td>
<td></td>
<td>[19,47–53]</td>
</tr>
<tr>
<td></td>
<td>Voltage and current trajectory</td>
<td>V_{traj}</td>
<td>[18,54–56]</td>
</tr>
<tr>
<td></td>
<td>Harmonic noise in power transients</td>
<td>H_f, 36 ≤ f ≤ 500Hz</td>
<td>[57,58]</td>
</tr>
<tr>
<td></td>
<td>Combination of independent features</td>
<td>V_{traj}, i(t), p(t), and others</td>
<td>[17]</td>
</tr>
<tr>
<td>High-frequency</td>
<td>Harmonic decomposition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Microscopic)</td>
<td>Power signature and harmonic decomposition</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Power signatures and Power Theory</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wavelets</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Voltage and current trajectory</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Harmonic noise in power transients</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Combination of independent features</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** P: active power; Q: reactive power; I_{RMS}: effective current; U_{RMS}: effective voltage; U_{MAX}: maximum voltage; I_{MAX}: maximum current; PF(λ): power factor; p_{ON}: load on probability; p_{OFF}: load off probability; P_{pulse}(t): time power pulse; H_{1,3,5,…,N}: nth harmonic components (n = 1, 3, 5,…, N); THD: voltage or current total harmonic distortion; S or A: apparent power; H: harmonic power; I_f: inactive current; CPT_{components}: CPT power components; W_{coeff}: wavelet coefficients; W_{c}: wavelet equivalent coefficients; V_{traj}: voltage vs current trajectory; H_f: non-fundamental power components; i(t): instantaneous current; p(t): instantaneous power.

In 1992, Hart [14] did some pioneering work on load disaggregation, in which he defined Nonintrusive Appliance Load Monitoring (NALM)—NILM is a derivative term from NALM. Such work showed that it is possible to separate power consumption by appliances observing the collective power consumption. To do this, it is necessary to discover the power behavior of each load or appliance. For a long time, this research did not draw attention due to digital device limitations for these transient statuses. Biansongnem and Plungklang [29] created a NILM method with 90% of accuracy when testing air conditioning and refrigerators. Using deep learning to detect operational load changes, Xiao and Cheng proposed a method and validated it using the Reference Energy Disaggregation Dataset (REDD) [59].

Powers et al. [31] applied a rule-based algorithm to detect loads with high consumption, such as air conditioning, water heaters and electric space heaters. Likewise, rule-based algorithms were proposed by Farinaccio and Zmeureanu [32] to detect power load behavior, as well as by Marceau and Zmeureanu [33] with the indication of 90% accuracy. With regards to genetic algorithms, Baranski and Voss [34] proposed their employment to detect patterns based on the use of loads frequency.
Ruzzelli et al. [35] created a dataset with low-frequency features of voltages and currents ($I_{RMS}$, $V_{RMS}$, $I_{MAX}$, $V_{MAX}$) and the power factor (PF) to describe load behavior, and with the P-Q analysis they carried out load disaggregation. Kelly and Knottenbelt [36] used a deep neural network for load disaggregation, as well as the UK Domestic Appliance-Level Electricity (UK-DALE) dataset [60], achieving an excellent performance for that dataset. Figueiredo et al. [37] used the load step changes in active power and the PF to create a dataset and concluded that there is a need to extract other attributes that can better detail the loads, especially those that have the same power and the same behavior as the equivalent circuit.

Kim et al. [38] combined frequency independent features, such as the distribution probability of ON/OFF duration, frequency of appliance usage and the correlation between the usage of various appliances with the active power feature, achieving between 64% to 99.8% accuracy in terms of load disaggregation. Koutitas and Tassiulas [39] replaced time-series power analysis for a set of discrete pulses. They created features based on pulses, such as variance, spike, slope, periodicity, multi-state, and sequence of operation. They reached an accuracy of 85%.

Sultanem [40] is the pioneer of the high-frequency use on NILM applications. Srinivasan et al. [41] used machine learning to recognize the harmonic signatures of 8 loads, and they obtained an accuracy of 99% or more for the load disaggregation. Laughman et al. [42] used the P-Q analysis for similar loads and increased the 3rd harmonic to distinguish them. Bouhouras et al. [43] used harmonic components to create a dataset for load disaggregation with stand-alone loads or combined loads, and the accuracy was between 85% and 95%.

Dong et al. [7] adopted Total Harmonic Distortion (THD) of current waveforms and P-Q analysis for load discrimination, and they used some pulse-based features. Lin et al. [17] created a NILM using features including the THD, P-Q analysis, voltage-current trajectory, current indicators and quadratic programming. The accuracy of this work is generally more than 90%.

Using power theory concepts, Teshome et al. [44] proposed a NILM with components of active, reactive, apparent power, and nonactive currents. Nguyen et al. [45] created a NILM method based on active, reactive and apparent power and used a decision tree (DT) to disaggregate five loads, achieving more than 98.8% accuracy. Huang et al. [46] pointed out that the application of power theories can be a useful tool for load disaggregation, especially for loads that have the same value of active power. In 2003, Tenti and Mattavelli [23,24] proposed the Conservative Power Theory, which allows load modelling in terms of power components. This work was the basis for the load characterization in the NILM dataset proposed by Souza et al. [15,25].

Considering the Continuous Wavelet Transform (CWT), Chan et al. [47] applied it up to the 4th level for the loads, and used Daubechies as mother wavelets, achieving an accuracy of 70%. Su et al. [48] and Duarte et al. [49] compared the Fast Fourier Transform (FFT) with CWT, pointing out the advantages of the CWT during load transients and recommended using CWT for feature extraction for load disaggregation. Chang et al. [52,53] implemented a NILM using active and reactive power with CWT. The authors extracted load features from five filters and applied a genetic algorithm for load identification. They reached almost 100% accuracy but the studies were carried out considering situations of significant discrepancy in power levels.

Gray and Morsi [50] used time-consuming energy to obtain CWT decomposition with Daubechies mother wavelets. The authors presented results comparing the accuracy of applying each order of the Daubechies order and concluded that the higher the Daubechies, the greater the accuracy. Tabatabaei [51] did a similar study, but used power characteristics to create classificatory features and obtained accuracy at around 85%. Gillis et al. [19] proposed a new CWT for NILM applications, obtaining around 94% accuracy for four connected loads at the same time.

Hassan et al. [54] created a NILM method based on V-I trajectory, which used wave-shape features along with the REDD dataset. Similarly, V-I trajectories were mapped to a grid of cells (as a matrix) by Du et al. [18], having a binary value assigned to each of them. On the other hand, 83% of accuracy was achieved by Gao et al. [55] by converting V-I trajectory into a binary image,
while using combined features, and considering 11 appliances. Finally, convolutional neural networks were proposed to be used with V-I trajectory by Baets et al. [56], reaching 77.6% of accuracy for the PLAID dataset [61], and 75.46% for the WHITED dataset [62]. Voltage harmonic (FFT) noise has also been used by Patel et al. [57,58], taking into account noise and electromagnetic interference in the range of 36 to 500 kHz. Nonetheless, such a study only highlighted the types of equipment that present multiple operational stages.

To summarize, there are many other studies regarding NILM with different methodologies, feature extraction, different appliances in the validation and different load disaggregation algorithms. Nevertheless, in Teshome et al. [44], the authors indicate the importance of modern power theories and the lack of these elegant circuit analyses to improve the NILM systems. One of these elegant power theories pointed out in Teshome et al. [44] was the CPT. Thus, such a modern power theory is applied in this work to improve the load disaggregation and present a novel NILM technique.

Accordingly, the next section presents the PSB, a new NILM methodology based on a state machine, which analyzes the active power signature (a low-frequency feature), and on the event detection, which finds features from the CPT and triggers the machine learning algorithm that uses the high-frequency attribute dataset proposed by Souza et al. [15].

3. The Power Signature Blob Method

3.1. Dataset with the Microscopic Features Extraction

In Souza et al. [15] some techniques for appliance disaggregation were evaluated, and the feasibility of identifying home appliances using pattern recognition algorithms was shown. Two pattern recognition algorithms achieved significant results: Optimum-Path Forest (OPF) [63] and K-Nearest Neighbor (KNN) [64] and the KNN (with \( K = 1 \)) was chosen because of its lower computational time. The voltage and current waveforms from several appliances were measured and decomposed in power components using CPT [23,24]. The CPT allows splitting the power into active, reactive, unbalance and residual parts. These power components help to interpret an appliance as an equivalent circuit, as shown in Figure 1, where \( v_m \) is the phase voltage, the current \( i_{Gm} \) coincides with the active current, \( i_{Lm} \) coincides with the reactive current, and the current source \( j_m \) coincides with the void current. \( G_m \) is the equivalent phase conductance and could be represented as a resistance, \( L_m \) is the equivalent phase inductance and could be represented as an inductor. All the mathematical background of the equivalent circuit can be found in Reference [65].

![Figure 1. Load equivalent circuit by Conservative Power Theory (CPT).](image)

Using the CPT power decomposition, Souza et al. [15] created a dataset of 35 home appliances such as irons, microwaves, refrigerators, washing machines, lamps and others. Each appliance features (CPT active power, power factor, reactivity factor, and nonlinearity factor [25]) refer to a set dimension and each collected instance as a point in the multidimensional space. The pattern recognition algorithm uses this dataset for the classification purpose of the appliance.

Figure 2 shows the Voronoi diagram concerning the 1NN results for the appliance dataset of [15], where each class number represents an appliance. It shows three Voronoi diagrams for the four attributes (P—Active power, PF—power factor, QF—reactivity factor, and VF—nonlinearity factor) from the dataset, presented in two dimensions to help visualize the decision boundaries.
Figure 2. K-Nearest Neighbor (KNN) decision boundaries of the CPT appliance dataset.

The supervised classification algorithm identifies which appliances consume electricity, according to CPT power terms calculation. However, it is likely that some false positives could be detected (when the appliance is identified mistakenly due to the similarity with others, for example, a 100 W bulb lamp and a 100 W LCD TV with high power factor). Another potential issue concerns some appliances with multiple power stages during the power operation, such as washing machines (with washing, spinning and rinsing). This occurs because the classifier needs to observe various levels of “ON” and “OFF” and could not perform the correct classification. Some methodologies [7,8,10,16,22,37,42,66] were created to solve the multiple power steps problem, and both observe the appliance behavior during time operation. The load power signature is relevant because some appliances do not operate in steady-state, and the method needs to disaggregate with high accuracy. Thus, in Reference [15] there was a microscopic feature dataset for the load classification, but it is required to create a method to observe the appliance power behavior before using the appliance classification.

3.2. Load Power Signatures

As initially pointed out in Reference [14], each appliance has a power signature that is not necessarily a steady-state power, thus Figure 3 presents different appliances’ power behaviors, and the appliances can be classified using power signatures [7,10,14,16,22], as shown in Figure 4.
From Figure 4, an appliance can have a power signature with:

(a) Constant behavior: In this case, the appliance has steady power behavior over time. It has a high probability of being an appliance with the resistive equivalent circuit, or the appliance is in steady state;

(b) Multiple power states: The appliance has a high-power peak when starting the operation. This characteristic could be linked to a starting engine, and after such a power step, it presents some power variations without turning off. This case can be related to a motorized appliance such as a washing machine;

(c) All-time constant: in this case, the appliance is always connected as a standby mode device;

(d) Approximately linear variation: in this case, there is a linear power variation over time. This variation corresponds to a transitional period until stabilization, such as by temperature (such as an iron) or gas (such as mercury lamp);

(e) Various operational stages: Some appliances have several power characteristics over time, which can be inductive, resistive or non-linear, etc. Besides, the power could also be switched on and off during the operation. For example, a clothes dryer has a motor that can rotate at different speeds and can also have a heating system to facilitate the drying process;

(f) Power behaviors with hard sequence detection: This type of appliance is usually electronic, and there are several fast operation stages, making it difficult to recognize the power signature.
over time. The noises from current and voltage sensors are also aggregated into this power signature category. The printer is an example of this type of load, which has some power steps that vary and switch very quickly.

Hence, considering the possibility of such different appliances’ signatures, it would be important to have a preliminary filter before using any appliance recognition technique, so as to increase the disaggregation accuracy. Thus, the next section presents the proposed approach, which uses CPT power terms and NILM techniques to detect the power signatures, before using the appliance classification method by means of the KNN algorithm.

The power signature behavior was aggregated into the 35 appliances dataset, as can be seen in Table 2. This characteristic is not used as features into the pattern recognition algorithm [15], but it is used to filter and increase accuracy in load detection.

### Table 2. Household appliance dataset.

<table>
<thead>
<tr>
<th>Order (Class)</th>
<th>Load</th>
<th>Event Type (Based on Figure 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Light bulb</td>
<td>a</td>
</tr>
<tr>
<td>2</td>
<td>Air conditioning</td>
<td>b</td>
</tr>
<tr>
<td>3</td>
<td>Refrigerator</td>
<td>e</td>
</tr>
<tr>
<td>4</td>
<td>Microwave</td>
<td>b</td>
</tr>
<tr>
<td>5</td>
<td>CRT TV</td>
<td>a</td>
</tr>
<tr>
<td>6</td>
<td>LCD TV</td>
<td>a</td>
</tr>
<tr>
<td>7</td>
<td>Plasma TV</td>
<td>a</td>
</tr>
<tr>
<td>8</td>
<td>Electrical shower</td>
<td>a</td>
</tr>
<tr>
<td>9</td>
<td>RL Load 1</td>
<td>a</td>
</tr>
<tr>
<td>10</td>
<td>NL Load 1</td>
<td>a</td>
</tr>
<tr>
<td>11</td>
<td>NL Load 2</td>
<td>a</td>
</tr>
<tr>
<td>12</td>
<td>RL Load 2</td>
<td>a</td>
</tr>
<tr>
<td>13</td>
<td>Iron</td>
<td>d</td>
</tr>
<tr>
<td>14</td>
<td>Washing machine</td>
<td>b</td>
</tr>
<tr>
<td>15</td>
<td>Hairdryer</td>
<td>b</td>
</tr>
<tr>
<td>16</td>
<td>Fluorescent lamp</td>
<td>a</td>
</tr>
<tr>
<td>17</td>
<td>Mix lamp</td>
<td>d</td>
</tr>
<tr>
<td>18</td>
<td>Mercury lamp</td>
<td>d</td>
</tr>
<tr>
<td>19</td>
<td>Sodium lamp</td>
<td>d</td>
</tr>
<tr>
<td>20</td>
<td>ASD Dryer</td>
<td>e</td>
</tr>
<tr>
<td>21</td>
<td>ASD Fridge</td>
<td>e</td>
</tr>
<tr>
<td>22</td>
<td>Blender</td>
<td>b</td>
</tr>
<tr>
<td>23</td>
<td>Bread maker</td>
<td>a</td>
</tr>
<tr>
<td>24</td>
<td>Desktop PC</td>
<td>e</td>
</tr>
<tr>
<td>25</td>
<td>Electronic ball lamp</td>
<td>d</td>
</tr>
<tr>
<td>26</td>
<td>Food processor</td>
<td>b</td>
</tr>
<tr>
<td>27</td>
<td>Freezer</td>
<td>e</td>
</tr>
<tr>
<td>28</td>
<td>Furnace</td>
<td>a</td>
</tr>
<tr>
<td>29</td>
<td>Garage door</td>
<td>a or e</td>
</tr>
<tr>
<td>30</td>
<td>Laptop</td>
<td>e</td>
</tr>
<tr>
<td>31</td>
<td>LCD monitor</td>
<td>a</td>
</tr>
<tr>
<td>32</td>
<td>Regular dryer</td>
<td>b</td>
</tr>
<tr>
<td>33</td>
<td>Regular fridge</td>
<td>e</td>
</tr>
<tr>
<td>34</td>
<td>Vacuum</td>
<td>e</td>
</tr>
<tr>
<td>35</td>
<td>Washer</td>
<td>b</td>
</tr>
</tbody>
</table>

### 3.3. The PSB Technique

The appliance dataset helps to detect one appliance by execution but might result in a problem, since more than one appliance may be ON during a certain period of time. In such a case, the active power would contain the aggregated value. Therefore, in order to decompose the consumption of each individual load, it is necessary to know how many loads are operating at that time.
Thus, considering the existence of an algorithm to classify the appliances that generate a power event, the set of instances and the appliance class create a new key-pair value. Therefore, a set of instances may contain several classes associated, and the aggregated set of instances can be decomposed considering the type of appliances.

To obtain a more accurate value of load energy consumption, the PSB takes the mean value of each block associated with each class to evaluate the mean active power during that period. Blocks that contain more than one active appliance should use the historical average value from each of them. This procedure is called disaggregation by blocks.

The load identification algorithm uses classification attributes extracted from the difference between two scenarios: before and after an appliance is turned ON. Hence, the step is just a divider between stable states. This guarantees that classifiers represent the appliances altogether, mitigating the noise generated in the transition state.

Subtracting the waveforms of these steps creates an approximation of the load waveform. Then, the CPT algorithm generates the attributes by processing the waveform. Adopting this approach, four features represent the appliances: active power, power factor, reactive factor, and non-linearity factor. Using these elements as attributes in a four-dimensional space, each load will result in a cluster. Therefore, the classification algorithms, such as KNN, can be used to set the frontiers of each load in the space.

The diagram from Figure 5 shows the state-machine algorithm with the appliance disaggregation dataset from [15]. The existence of appliance power signatures can be observed, which filter and can make the appliance disaggregation more accurate. Moreover, the methodology has algorithms for handling the ON and OFF events, presented in the diagrams from Figures 6 and 7. The methodology stores the fifteen previous cycles (0.25 s of total samples) of voltage and current waveforms to detect the appliance events if there are two or more appliances turned ON.

**Figure 5.** State-machine algorithm of the PSB.
4. Validation Results

In this paper, the PSB evaluation was based on simulations and experimental results, as depicted in the next subsections.

4.1. Simulation Results

Table 3 presents the daily appliance schedule for the simulation according to user behavior in the residence.
Table 3. Simulation: Appliance Schedule Runtime.

<table>
<thead>
<tr>
<th>Appliance</th>
<th>“Turn on” Time</th>
<th>Total Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electrical shower</td>
<td>07:15 and 19:40</td>
<td>00:20 and 00:15</td>
</tr>
<tr>
<td>Air conditioner</td>
<td>22:00</td>
<td>08:00</td>
</tr>
<tr>
<td>CRT TV</td>
<td>12:15 and 18:15</td>
<td>01:00 and 05:00</td>
</tr>
<tr>
<td>Refrigerator</td>
<td>All the day</td>
<td>00:30 each cycle</td>
</tr>
<tr>
<td>Iron</td>
<td>18:30</td>
<td>00:15</td>
</tr>
<tr>
<td>Lamp1 (bulb 100 W)</td>
<td>18:30</td>
<td>05:30</td>
</tr>
<tr>
<td>Lamp2 (bulb 60 W)</td>
<td>19:00</td>
<td>04:30</td>
</tr>
<tr>
<td>Notebook</td>
<td>19:00</td>
<td>04:00</td>
</tr>
<tr>
<td>Microwave</td>
<td>07:45</td>
<td>00:05</td>
</tr>
<tr>
<td>Washing Machine</td>
<td>13:00 (Saturday)</td>
<td>02:00</td>
</tr>
</tbody>
</table>

Therefore, an electrical circuit model was created in PSIM software in order to simulate the power behavior of a residence during an entire day, turning on each appliance according to the scheduled runtime of Table 3.

The simulation collects 256 samples per cycle from current and voltage waveforms and sends them to a Dynamic Link Library (DLL) block. The general DLL block in PSIM allows users to write codes in C or C++, compile them as a Windows DLL, and link them to PSIM using the features of input (from PSIM) and outputs (returning to PSIM). Unlike the simple DLL blocks with a fixed number of inputs and outputs, the general DLL block provides more flexibility and capability in interfacing PSIM with custom DLL files. In this paper, the DLL codes were used to implement the CPT, the KNN and state-machine NILM, according to the algorithms from Figures 5–7. The state machine from Figure 5 is a loop responsible for the event decision when there is power consumption changing. If there is power changing, the state machine flows in steps until it detects the power event and triggers the “ON event” (Figure 6) or the “OFF event” (Figure 7). If an “ON event” is detected, the system runs algorithm Figure 6 responsible for classifying the appliance. After that, the algorithm adds the appliance to the turned-on appliance list and saves the waveform status for future comparisons when a new event is triggered. If an “OFF event” is detected, the system runs algorithm Figure 7 that is responsible for classifying the appliance and removing the appliance from the turned on appliance list and saves the waveform status for comparison when a new event is triggered.

Simultaneously, the PSB takes care of four other characteristics:

- A database receives the current power consumption in kWh every minute;
- At every turn ON appliance event, the system makes the pattern recognition and sends all the appliance power information to the database, and then, it starts the detected appliance time operation;
- At every turn OFF event, the system recognizes the appliance that was turned off and updates the database, storing the information from the time the load was switched OFF;
- If there are power changes, the system verifies the power signature of each appliance that is turned ON and recognizes it as appliance power level changes.

Figure 8 presents the simulation results. The state machine with the classifier algorithm accepted all the loads according to Table 3. Figure 8 shows the CPT power decomposition and, consequently, the moments in which the algorithm performed the appliance identification (turn on and turn off triggers).
Figure 8. Daily consumption of electricity by household appliances. (a) Load disaggregation behavior in simulation; (b) CPT power components behavior in simulation.

Figure 9 shows an example of the operation of the PSB between 09:27 and 10:03 from Figure 8. In this interval, there is the operation of a refrigerator, according to the schedule of Table 3. Following the algorithms of Figures 5–7, the state machine has the following behavior:

- **09:27 to 09:30:** Active power is stable, and there is nothing to be done at this time;
- **09:30:** \( \Delta P > \text{Error} \); \( T_{ON} = TN \), goto 1;
- **09:30:** \( \Delta P > \text{Error} \); // waiting for power stabilization;
- **09:31:** \( \Delta P > \text{Error} \); // waiting for power stabilization;
- **09:31:** \( \Delta P > \text{Error} \); // waiting for power stabilization;
- **09:32:** \( \Delta P < -\text{Error} \); goto 5;
- **09:32:** \( \Delta P < -\text{Error} \) and \( \Delta P > -\text{Error} \); ON** Event;
  - CPT Features extraction;
  - KNN classifier;
  - Appliance recognition;
  - Recognized appliance added to the appliance list;
  - goto 0;
- **09:32 to 10:00:** Active power is stable, and there is nothing to be done at this time;
- **10:00:** \( \Delta P < -\text{Error} \); \( T_{OFF} = TN \); goto 3;
- **10:01:** \( \Delta P < \text{Error} \) and \( \Delta P > -\text{Error} \); goto 4;
- **10:02:** \( \Delta P < \text{Error} \) and \( \Delta P > -\text{Error} \); OFF event;
  - Waveform difference;
  - CPT feature extraction;
  - KNN classifier;
  - Appliance recognition;
  - Recognized appliance removed from the appliance list;
  - goto 0;
• **10:02 to 10:03:** Active power is stable, and there is nothing to be done at this time;

a) Load disaggregation behavior between 09:27 to 10:03

![Active power difference](image)

b) Active power behavior between 09:27 to 10:03

![Active power (A)](image)

**Figure 9.** Detailed consumption and state-machine behavior between 09:27 and 10:03. (a) Load disaggregation behavior between 09:27 and 10:03; (b) Active power behavior between 09:27 and 10:03.

Therefore, the PSB worked as expected, and there was no error in the disaggregation process in the simulations. In this study, there was no error because it is used an appliance dataset with high accuracy and the simulated loads operate in a steady state. However, Section 4.2 will present real appliance cases, including power variations.

### 4.2. Experimental Results with Household Appliances

The PSB operates with a main loop of 15,360 collected samples. Hence, with a fixed fundamental frequency, this loop takes a second to perform. Then, a four-step procedure tracks rapid power variations. Each step updates the average of the active power evaluated with the last block of samples: 0.25 s of total samples.

Therefore, a predefined threshold level compares the active power variation during the time. Then, the procedure uses the difference between the current average active power and the last one used to define the step level direction: when there is the “power on” or “power off” state of the appliance. Figure 10 shows the behavior of the algorithms from Figures 5–7 in this experiment.

In the trigger events—from Figures 6 and 7—the algorithm stores current and voltage waveforms of the last 0.25 s, extracts the current to be considered in the event and calculates the four elements that represent the appliance: active power, power factor, reactive factor, and non-linearity factor. With these elements (attributes), the algorithm uses the classifier algorithm (the KNN) by means of the knowledge dataset from [15]. The classifier returns the label of the appliance, i.e., the algorithm predicts which appliance is turned on (algorithm from Figure 6) or turned off (algorithm from Figure 7). If it is an “ON event” the recognized appliance is added to the appliance list. If it is an “OFF event”, the recognized appliance is removed from the appliance list. After this, the system stores the fifteen waveform cycles of the state to use in a new event trigger. Over time, appliances can be identified, as shown in Figure 11.
Considering such experiments, the PSB reached 95% of accuracy. The main problem encountered in the method corresponds to the loads with several power changes without the existence of a zero-power instant (i.e., a real turn OFF event). Figure 12 shows an example of an air conditioner with an adjustable speed driver (ASD). In this case, the method of load disaggregation carries out the load identification in the ON event trigger and, in the course of the operation, the power changes without the activation of a new trigger. When the device is turned off, the power level is different from the start of the operation, and the classifier may incorrectly identify the turned off appliance. If there is more than one appliance that is turned on, the algorithm may remove the wrong equipment from the list. If there is only one appliance, the algorithm empties the list.

To solve this problem, the appliance dataset must have a wide range of appliance measuring times and waveforms should consider the last state (in the OFF event trigger). The algorithm from Figure 5 could also be adapted to improve the accuracy in such situations.
5. Conclusions

This work presented a novel load disaggregation method to be used in cognitive meters, making the “smart” concept from smart meters more reasonable.

The novel methodology called PSB uses the power signature technique with a load recognition from an appliance dataset. The method detects the appliance ON and OFF events during the power signature observation, and the method uses classification algorithms to detect the appliance. The association of the classification algorithm and the power signature recognizes appliances that present power variation during their regular operation. The method also recognizes several loads simultaneously.

When evaluating the proposed method through computational simulations, all the loads were classified correctly. The PSB obtained an accuracy of 95% for real data from a typical residence.

In the future, the authors intend to work on energy efficiency evaluation, associated to identifying appliances. Moreover, the authors would like to apply the PSB to other smart grid applications, such as energy management using the concept of Virtual Power Plants and NILM applied to a group of residential installations. Future papers will deal with such ideas and prominent results.


Funding: This research was supported by FAPESP (Sao Paulo Research Foundation) grants number 2012/19375-1 and 2016/08645-9, as well as by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

8. Fagiani, M.; Bonfigli, R.; Principi, E.; Squartini, S.; Mandolini, L. A Non-Intrusive Load Monitoring Algorithm Based on Non-Uniform Sampling of Power Data and Deep Neural Networks. Energies 2019, 12, 1371. [CrossRef]


60. Kelly, J.; Knottenbelt, W. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Sci. Data* 2015, 2, 150007. [CrossRef]


