

NILM techniques applied to a real-time monitoring system of the electricity consumption

B. Cannas¹, S. Carcangiu¹, D. Carta², A. Fanni¹, C. Muscas¹, G. Sias¹, B. Canetto³, L. Fresi³, P. Porcu³

¹ DIEE, University of Cagliari, Via Marengo 2, 09123 Cagliari, Italy

² IEK-10, Forschungszentrum Jülich, 52425 Jülich, Germany

³ Bithiatech Technologies, Elmas, (CA), Italy

ABSTRACT

Non-Intrusive Load Monitoring (NILM) allows providing appliance-level electricity consumption information and decomposing the overall power consumption by using simple hardware (one sensor) with a suitable software. This paper presents a low-frequency NILM-based monitoring system suitable for a typical house. The proposed solution is a hybrid event-detection approach including an event-detection algorithm for devices with a finite number of states and an auxiliary algorithm for appliances characterized by complex patterns. The system was developed using data collected at households in Italy and tested also with data from BLUED, a widely used dataset of real-world power consumption data. Results show that the proposed approach works well in detecting and classifying what appliance is working and its consumption in complex household load dataset.

Section: RESEARCH PAPER

Keywords: Non-Intrusive Load Monitoring (NILM); electricity consumption; energy disaggregation; features extraction; smart metering; BLUED dataset; proprietary dataset

Citation: Barbara Cannas, Sara Carcangiu, Daniele Carta, Alessandra Fanni, Carlo Muscas, Giuliana Sias, Beatrice Canetto, Luca Fresi, Paolo Porcu, NILM techniques applied to a real-time monitoring system of the electricity consumption, Acta IMEKO, vol. 10, no. 2, article 20, June 2021, identifier: IMEKO-ACTA-10 (2021)-02-20

Section Editor: Giuseppe Caravello, Università degli Studi di Palermo, Italy

Received January 21, 2021; **In final form** March 16, 2021; **Published** June 2021

Copyright: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was partially funded by Sardinian Region under project "RT-NILM (Real-Time Non-Intrusive Load Monitoring for intelligent management of electrical loads)", Call for basic research projects, year 2017, FSC 2014-2020.

Corresponding author: Daniele Carta, e-mail: d.cart@fz-juelich.de

1. INTRODUCTION

Knowing how electric appliances are used and how different appliances contribute to the aggregate total consumption could help users to have a better understanding on how the energy is consumed, thus leading possibly to a more efficient management of their loads. Non-Intrusive Load Monitoring (NILM) is an area of computational sustainability research, and it presently identifies a set of techniques that can disaggregate the power usage into the individual appliances that are functioning and identify the consumption of electricity for each of them [1].

In residential buildings, where it is impractical to monitor single appliances, or even groups of appliances, through specific meters, NILM techniques are a low-cost and not invasive option for electric consumption monitoring, considering a single monitoring point where a smart meter is installed.

Literature reports several papers that applied different methods throughout the years to solve this problem. A first classification of NILM techniques could be done in supervised and unsupervised methods [2]. Supervised methods require labelled data of consumption to train the model. Unsupervised methods are targeted to extract the information to operate directly from the measured aggregate data consumption profiles. Due to their better performance, most of the approaches are based on supervised algorithms and they require appliance data for model training to estimate the loads number, type and power by analysing the aggregate consumption signal.

Solutions based on machine learning range from classic supervised machine-learning algorithms (e.g., Support Vector Machines, and Artificial Neural Networks) to supervised statistical learning methods (e.g., K-Nearest Neighbours and Bayes classifiers) and unsupervised method (e.g., Hidden-Markov Models and its variants). A review of these methods is reported in [2]. Recently, Deep Learning (DL) methods were also

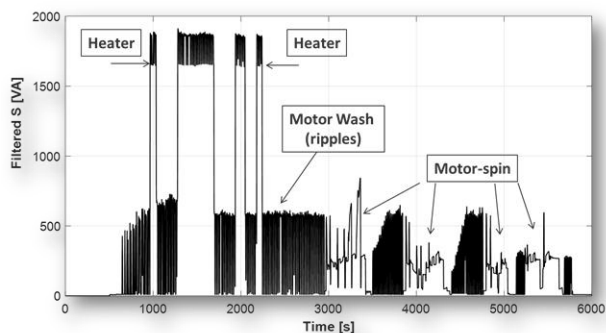


Figure 1. Apparent power for a washing machine (hot water).

employed and seem promising for the most challenging problem posed by the consumption profiles of multi-state appliances [3]-[6].

The frequency of energy data monitoring drives the use of the analysis techniques and the specific tools. Although higher the frequency of energy data monitoring frequency higher could be the accuracy of the NILM disaggregation algorithms, commercial smart meters for homes supply low frequency sampling (less than 60 Hz) of the electric power quantities. In this field, the majority of the research efforts focused on event-based techniques that identify significant variations in the power signals as switching events of appliances. These events must be classified as a state transition related to a specific appliance. For this purpose, electric signal characteristics extracted from measurements in proximity of the events (i.e., signatures) are used as distinctive features, and then labelled with classification procedures.

In this paper, a monitoring system is proposed using a smart meter that supplies low frequency (1 Hz) samples of power consumption. The system is able to disaggregate and keep track of the power consumption of the devices existing in a typical Italian house. The households should follow the proposed procedure to customize the system for their homes, choosing the appliances of interest and collecting the corresponding measurements. The disaggregation algorithm is an improvement of the one proposed in [7]-[8]. The load disaggregation is performed applying a hybrid approach to power data, i.e. event-based techniques and pattern recognition techniques for large household appliances. Moreover, the procedure of the first set-up phase has been simplified: the event detection for Type-I and Type-II appliances that was performed separately for each device is now carried out on the aggregate power signals, strongly reducing the user effort.

In order to validate the method, the procedure is also applied to Building-Level fully-labeled (BLUED) dataset for Electricity Disaggregation [9]. BLUED is a publicly available big dataset consisting of real voltage and current measurements for a single-family residence in the United States, sampled at 12 kHz for a whole week. BLUED has been applied as benchmark dataset in several recent papers on NILM [5], [10], [11], [12]. However, few contributions have been proposed with low frequency samples. Among them, [13] and [14] will be considered in this work as comparison.

2. THE HOME ENERGY MANAGEMENT SYSTEM

The Home Energy Management System (HEMS) is used to provide comfortable life for consumers as well as to save energy. This can be obtained using a home's smart meter to monitor the electricity consumption of the devices existing in a household

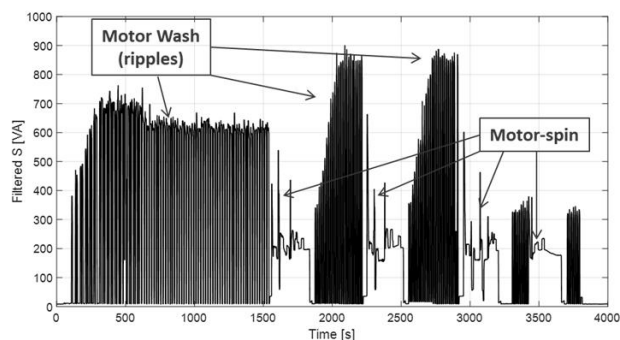


Figure 2. Apparent power for a washing machine (cold water).

applying NILM techniques. The HEMS should identify the appliances that are active at any time, disaggregate the energy and estimate consumptions of the single device. In this work, the chosen low-frequency smart meter is a sensor belonging to a pre-commercial prototype [7] which provides steady-state signatures, such as real and reactive power time series. It is important that the HEMS set-up is easy to understand and interact with. It should have features, like auto-configuration, which make the set-up process very easy. The non-intrusive technique, resorting to only one smart meter for the household, requires some effort in the first set-up phase, but it can be sometimes the only possible choice because installing a specific monitoring infrastructure, including new devices cables, may result in high implementation cost to the user.

In this work the appliances are categorized into four types based on their operation states [15]: 1) Type-I, appliance with ON/OFF states (binary states of power); 2) Type-II, finite state machines with a finite number of operation states (multiple states of power); 3) Type-III, continuously varying devices with variable power consumption, as a washing machine and a light dimmer (infinite middle states of power); 4) Type-IV, appliances that stay ON for days or weeks and with a constant power consumption level.

In case of Type-II, data for all the transitions between possible states should be acquired and labelled (manual set-up). In case of Type-III devices, data from each cycle are characterized by complex patterns. In this paper, a technique to deal with such devices is proposed, by considering the washing machine as a case study. Figure 1 and Figure 2 show the filtered apparent power consumption typical of different models of washing machines with different washing cycles. As can be noted, the power consumption fluctuates while heating/washing or rinsing/drying the laundry. Thus, the events do not correspond to simple steps in the power consumption, but characteristic complex patterns appear in the power time series. As it can be noted, heating water accounts for about 90% of the energy needed to run a washer and in both figures the washing machine typically has electrical components which turn ON and OFF in sequence.

The proposed technique facilitates the training process by pre-populating the training data set with signatures of some Type-III devices showing typical patterns (automatic set-up).

3. THE PROCEDURE

In the following, the implemented procedure is presented showing how the variations of apparent (S), real (P) and reactive powers (Q), the oscillation frequencies of the signals and the varying patterns of Type-III appliances are used in the training

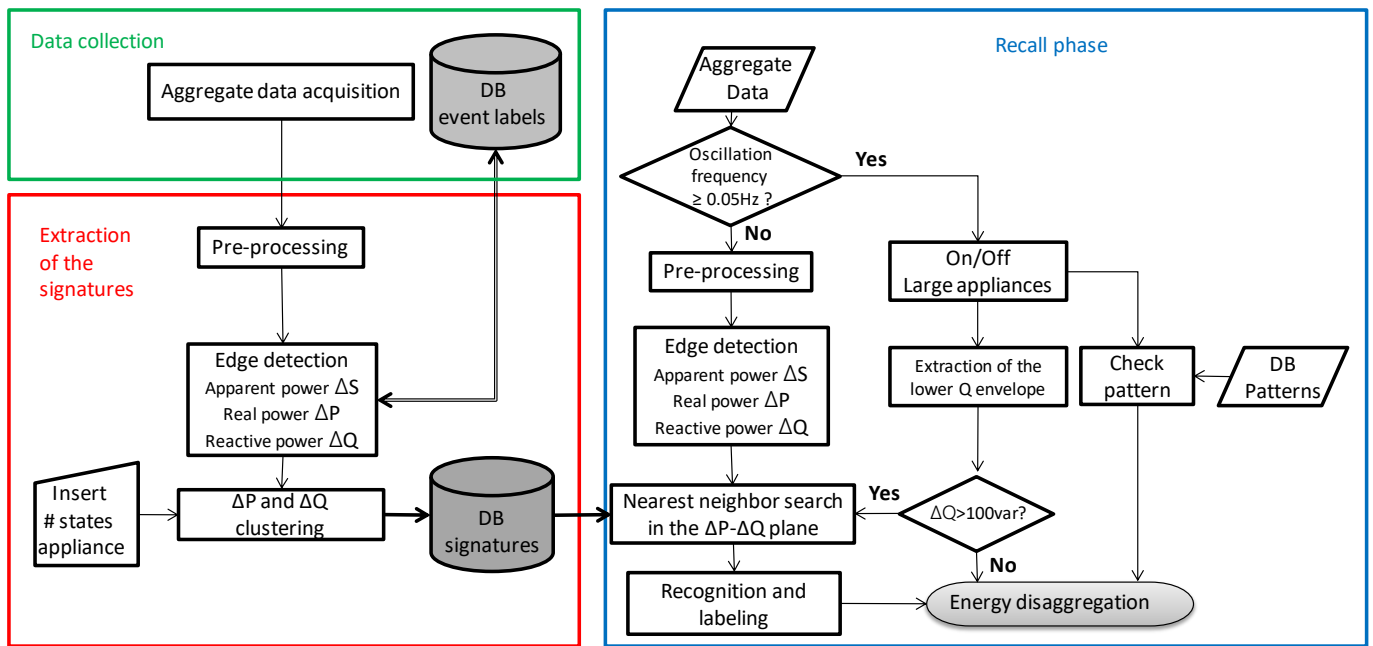


Figure 3. Flow chart of the procedure.

phase for the creation of the signatures database, whereas during a recall phase the appliances are recognized inside an aggregated signal. In order to better understand the whole procedure, Figure 3 shows its flow chart with the main phases of Data collection and Extraction of the signatures.

3.1. Data collection

The automatic procedure to collect the aggregate data of electric consumption consists of the following two steps, one for Type-I and II appliances, one for Type-III appliances.

In the first step, Type-III appliances (such as washing machine or dishwasher) are switched off, whereas Type-I and Type-II appliances could regularly work. The user must switch on and off Type-I and II devices several times for each event of interest. This allows to increase data robustness to the noise, e.g., small fluctuations in appliance consumption, electronics constantly on, and appliances turning ON/OFF with consumption levels too small to be detected.

Multi-state devices, such as kitchen ovens, stoves, clothes dryers, etc., go through several states where heating elements and fans can be switched in various combinations. Thus, to collect all the events, the user must test all the possible transitions from one state to another. For instance, for an electric stove with three power levels (states), the user should trigger twelve possible switching events, among OFF/ON states.

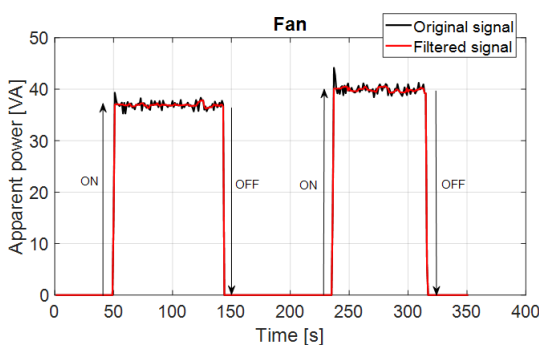


Figure 4. Original and filtered fan power consumption.

In the second step, large Type-III appliances work alone and their data of consumption are recorded.

Note that, for this work, also data from Type-I and Type II appliances working alone have been collected, in order to combining them and create several synthetic aggregate data series to test the system.

3.2. Extraction of the signatures

Once the aggregate data have been collected, the database of the signatures of each switching event is built. These data are normalized with respect to the constant voltage of 230 V in such a way that the voltage drops due to the load insertion do not influence the result. Moreover, a causal filtering is applied to apparent, real, and reactive power signals. In this way, possible spikes and outliers can be discarded or smoothed. With such a low sampling frequency, fast transients should be removed, since they could be sometimes recorded, and other times missed during the acquisition. Figure 4 shows the changes in electricity consumption due to the switching ON and OFF of a fan before and after the filtering.

For the Type-I and II devices, an edge detector finds the switching events in the apparent power data when the absolute value of the difference ΔS between two consecutive values is larger than 20 VA. The sign of ΔS identifies the start up or the shutdown of the appliance. Then, the real ΔP and reactive ΔQ power variations in each edge must be determined and candidate as one of the signatures of the individual load.

Within the time interval between two consecutive events there is an almost constant *power level* consumption. In order to find a candidate signature of the appliance-switching events, the difference between the mean values of the measurements before and after each edge of real and reactive power is evaluated. Among all the candidate signatures, the *k-medoids* clustering method [16] is applied to partition the set of switching events into a set of clusters whose number k depends on the possible states of the single appliances and must be set by the user. This clustering method is robust to noise and outliers and it chooses data points as centers of the clusters. At the end of the clustering

procedure, these centers will form the signatures associated with each transition.

Note that, as appliances with small power consumptions are not interesting from the point of view of energy savings, and hardly distinguished, loads with $P < 20 \text{ W}$ are discarded and loads with $20 \text{ W} < P < 50 \text{ W}$ are associated to a unique “low consumption” cluster. For the same reason, switching events lasting less than 5 s are not taken into account. Note that, the threshold of 50 W is implemented by default, but it can be modified by the user, in case only consumptions greater than a predetermined power are of interest.

At the end of the training phase, all the collected data are given as input to the monitoring system; in case of errors in the classification, new common clusters are created for those devices characterized by close consumptions of real and reactive power.

For large Type-III appliances, an *ad hoc* procedure has been implemented. For the washing machine, the start and the end of the cycle and the motor-spin events are detected. To this aim, the peak values of the real power oscillations that identify the heating and washing phases are identified. In order to avoid peaks due to noise or other events not characterizing this device, maximum relevance peak values are selected, i.e., those that drop at least 30 W on either side before the signal attains a higher value. A statistical analysis of the time distance of such peaks shows that the typical time distance between the peaks is close to 20 s. Thus, the oscillations in the active power signal show a frequency greater than or equal to 0.05 Hz.

Moreover, in order to avoid spurious detections, the pattern of the motor spin pattern is isolated. The motor-spin pattern shown in Figure 5 is identified in the individual appliance signals. The pattern identified in the individual appliance signals is then extracted and included in the database of the system.

3.3. The recall phase

During the recall phase, first of all, a check is carried out to verify if a washing machine is running: as described in section 3.2, the switching ON (OFF) of the washing machine operation is identified when the oscillations, characterized by a frequency greater than or equal to 0.05 Hz, start (end).

When no washing machine cycle is detected, the following procedure is applied. When an edge in the aggregated signal is detected, the corresponding point in the ΔP - ΔQ plane is evaluated. Then, a nearest neighbour search in the ΔP - ΔQ plane is performed, and the event is classified and associated to the appliance event with the nearest signature vector.

Moreover, a check on the sign of Q is considered as further information, in addition to the cluster center distance, to identify the proper cluster in order to increase the discrimination capabilities. See as an example the signatures of the hairdryer and

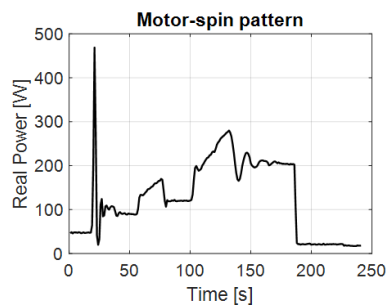


Figure 5. Motor-spin pattern.

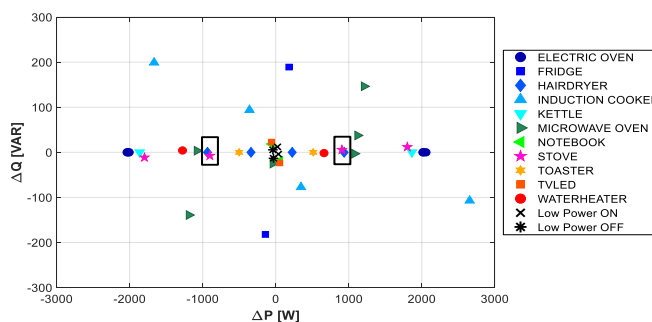


Figure 6. Custom dataset: appliance signatures in ΔP - ΔQ plane. The overlapping signatures of hairdryer and stove are highlighted with black rectangles [7].

stove in Figure 6, where the signatures are very close but the former, unlike the latter, is characterized by zero reactive power.

If no association is performed, the event is labelled as unidentified.

The detections of events of Type-I and II appliances switched ON and OFF during a washing machine cycle become quite challenging with such a low frequency and many false switching could be triggered directly applying the described procedure. Thus, if the washing machine is running in its heating phase, the lower envelope of Q is extracted. Since the washing-machine Q lower envelope during the heating phases is equal to zero, the presence of the intervals of not-null constant values indicates the switching of an appliance. In this case a threshold of 100 var is considered. When an edge is detected in a flat-top interval of the Q lower envelope, the nearest neighbour search is then applied in the ΔP - ΔQ plane in order to associate the event to the appliance with the nearest signature vector. This procedure allows one to improve the detection in case of reactive loads switched ON and OFF during the washing machine heating phases.

Finally, using a similarity search algorithm, it is possible to identify, in the aggregate signal, the operating phase that best matches with the reference associated to the spin motor functioning. This procedure could be applied for other devices with characteristic patterns, e.g. microwave ovens, to distinguish their operating conditions.

4. CASE STUDY

In this section the experiments carried out to verify the validity of the methodology proposed are reported.

The algorithms are implemented in MATLAB.

Performance have been evaluated both on a custom dataset collected by the authors in some Italian houses and on a public dataset, BLUED [9], which has already been extensively analysed in the literature.

4.1. Custom dataset

In order to create our dataset, several domestic consumption data have been acquired by installing an energy meter between the investigated appliances and the domestic network.

The implemented acquisition system of the electricity consumption data consists of:

- one EASTRON SDM220 single-phase energy meter [17], for residential and industrial applications at a rated voltage of 230 V (range 176 V to 276 V) and current of 10 A (range 0.5 A to 100 A). The accuracy requirements of the meter are reported in Table 1;

Table 1. Accuracy requirements of EASTRON SDM220 [17].

Parameter		Accuracy
Voltage		0.5% of range maximum
Current		0.5% of nominal
Power	Active	1% of range maximum
	Reactive	
	Apparent	
Energy	Active	Class 1 IEC 62053-21
		Class B EN 50470-3
	Reactive	1% of range maximum

- a PC on which the measurement software and the load disaggregation algorithm are implemented;
- a MODBUS/RS485 Serial interface, including a Serial Port Converter Adapter Cable USB-RS485, allowing the remote communication between the energy meter and the PC.

The dataset has been created with an acquisition frequency equal to 1 Hz that represents the maximum value allowed by the used meter. The dataset consists of three parts: individual loads, aggregate loads and synthetic aggregate loads. The data were acquired in single and aggregate manner during the actual operation of the devices or were generated by simulating conditions corresponding to actual user behaviour.

The electricity consumptions of the individual loads reported in Table 2 have been acquired by connecting the meter directly to the device plugs. The dataset was obtained acquiring multiple appliances simultaneously via a multi-socket. In order to increase the amount of the data corresponding to aggregate consumption, an aggregate synthetic dataset was obtained by combining the consumption data of the individual appliances, by summing the measurements of P and Q and averaging the values of V.

4.2. BLUED dataset

This dataset was built in 2011 by monitoring a whole house located in Pennsylvania (US), for 8 days. In US there are three electricity feed lines for ordinary houses: two firewires and one neutral line. The two firewires have 120 V amplitude of voltage and they are named as phase A and phase B. Usually, small 120 V-rated appliances are connected between one firewire and one neutral while larger 240 V-rated appliances such as heaters and air conditioners are connected between two firewires. In this work only phase A data have been used. The appliances connected to this phase are reported in Table 3.

Table 2. List of appliances in the custom dataset.

Appliance	Average power consumption in W	Type
Fridge	180	II
Kettle	1900	I
Lamp	40	I
Notebook	60	I
Stereo	30	I
Toaster	500	I
TV	40	I
Electric oven	2000	II
Hairdryer	300-900	II
Fan	30-40	II
Induction cooker	400-2500	II
Microwave oven	1000-1200	II
Stove	900-1800	II
Water heater	600-1200	II
Washing machine	130-1700	III

Table 3. List of appliances in the BLUED dataset (phase A).

Appliance	Average power consumption in W	Type
Kitchen aid chopper	1500	II
Fridge	120	II
Air compressor	1130	I
Hair dryer	1600	II
Backyard lights	60	I
Washroom light	110	I
Bathroom upstairs lights	65	I
Bedroom lights	190	I

The BLUED dataset contains high frequency (12 kHz) aggregated data of raw current and voltage. During the creation of the dataset, every single switching ON/OFF of any appliance is recorded and called as an event. In particular, all the changes in the state of power consumption higher than 30 W have been considered. Every appliance event in the house was then labelled and time-stamped, providing the necessary event labels database for the evaluation of the proposed procedure. In total, 904 events have been registered in the considered phase A.

In this work, to take into account the technical specifications of the pre-commercial smart meter used in the experiments discussed in the previous section 4.1, power signals evaluated from raw data are down-sampled at 1 Hz. Then, the events identified in BLUED, but unknown, and those with duration less than or equal to 5 s have been discarded obtaining a final database of 662 ground-truth events.

5. RESULTS

Figures 6 and 7 show the signatures of custom and BLUED databases in the ΔP - ΔQ plane obtained by applying the phase of the procedure “Extraction of the signatures”, described in section 3.2. In this section the performances of the recall phase (event detection and appliances identification) over the two datasets are presented.

5.1. Performance measures

In binary classification problems (such as the event detection) there are only two classes, called Positive (ON or OFF event) and Negative (non-event). When a Positive sample is incorrectly classified as Negative, it is called a False Negative (FN); when a Negative sample is incorrectly classified as a Positive, it is called

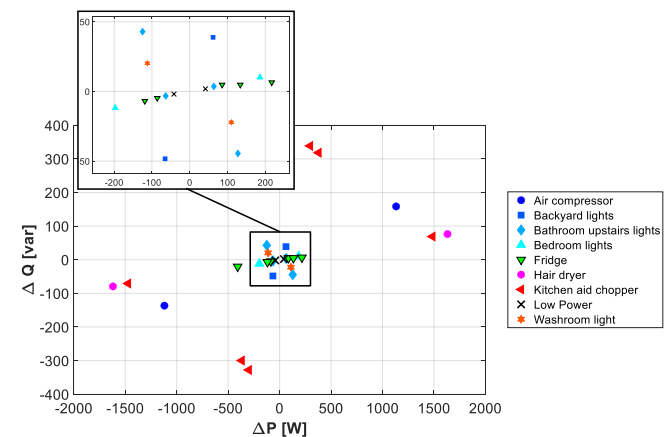


Figure 7. BLUED dataset (phase A): appliance signatures in ΔP - ΔQ plane. The magnification of the part enclosed by the black rectangle shows more in detail the signatures with $|\Delta P| < 250$ W.

a False Positive (FP); when a Positive sample is correctly classified as Positive is called a True Positive (TP).

The *Precision* (Pr) represents what proportion of predicted Positives is truly Positive. It is the ratio between the number of correctly classified Positives and the total number of samples predicted as Positives:

$$Pr = \frac{TP}{TP + FP} \quad (1)$$

The *Recall* (Re) represents what proportion of actual Positives is correctly classified. It is the ratio between the number of correctly classified Positives and the total number of Positives in the dataset:

$$Re = \frac{TP}{TP + FN} \quad (2)$$

The *F₁ - score* combines *Precision* and *Recall* into a single measure. Mathematically it is the harmonic mean of Precision and Recall.

$$F_1 - score = 2 \cdot \frac{Pr \cdot Re}{Pr + Re} \quad (3)$$

In a multiclass classification problem (as the appliance identification) there are no positive or negative classes, but TP, FP and FN and the other performance measures can be evaluated for each individual class (each appliance event). Summing up the single class measures, the total TP, the total FP and the total FN of the model can be obtained; then the global metrics of Precision, Recall and F-score can be evaluated. Note that in this case, all the global metrics become equal, i.e. $Pr = Re = F_1 - score$.

5.2. Performance with custom dataset

The Recall phase has been tested on aggregate signals composed of Type-I and Type-II appliances with and without the washing machine. Using the pattern shown in Figure 5, the operation phases of the motor-spin are individuated even considering washing machines of different brands and with different washing programs.

Figure 8 shows the power demand of a household over a 14-hour period, between 08:00 and 22:00. As it can be observed, the aggregate power demand is generated by the fridge, which shows a periodical power consumption behaviour, and other appliances. The ON and OFF events are shown as markers at the level of $+\Delta P$ and $-\Delta P$ respectively, whereas the motor spin functioning is indicated with red segments.

The results are very satisfactory, especially regarding identification of the activation status of appliances that consume more energy, such as the washing machine. As an example, in Figure 9 the events detection during a washing machine cycle, is shown. Flat-top intervals of the Q lower envelope, denoted by the green segments in Figure 9, identify the switching ON/OFF of the fridge.

In order to show the effectiveness of the improvements proposed in this paper, in Table 4 the performance for the edge detection as reported in [7], are shown. In Table 5 the event classifier performance is compared with those reported in [7].

As it can be noted, despite the performance index obtained in [7] for the synthetically generated aggregate signals was already very high, a slight improvement has been achieved. A more significant improvement in the performance index can be observed on the experimental data after the introduction of the described changes.

Table 4. Performance metrics for the edge detector tested on synthetic and experimental custom dataset.

Test Set	# ground-truth events	TP	FP	FN	Pr	Re	F ₁ -score
Synthetic Data	140	139	1	1	0.99	0.99	0.99
Experimental Data	137	135	2	2	0.99	0.99	0.99

Table 5. Performance metrics for appliances classification tested on synthetic and experimental custom dataset.

Test Set	# detected events	TP	FP	FN	Pr	Re	F ₁ -score
Synthetic Data [7]	139	133	6	6	0.96	0.96	0.96
Synthetic Data	139	133	3	3	0.97	0.97	0.97
Experimental Data [7]	135	121	14	14	0.90	0.90	0.90
Experimental Data	135	129	6	6	0.96	0.96	0.96

The pie charts in Figure 10 compare the estimated decomposition results with ground-truth of energy consumption. As can be observed, the proposed method is capable of disaggregating energy consumption of the appliances with good accuracy.

5.3. Performance with BLUED dataset

In the BLUED dataset the aggregate signals are composed of Type-I and Type-II appliances. The event detector applied to BLUED time series identifies 647 events characterized by $\Delta S > 20$ var and lasting more than 5 s. Among them, 9 events

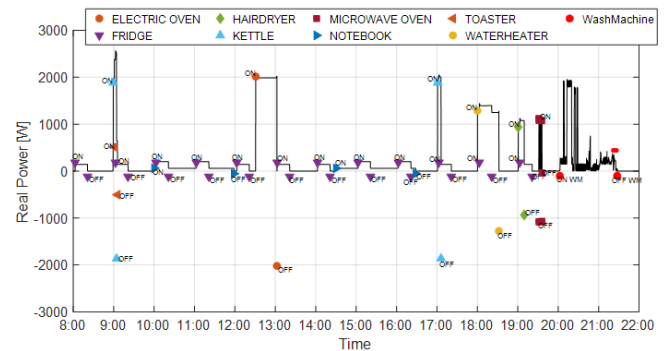


Figure 8. Household power demand over a 14-hour period.

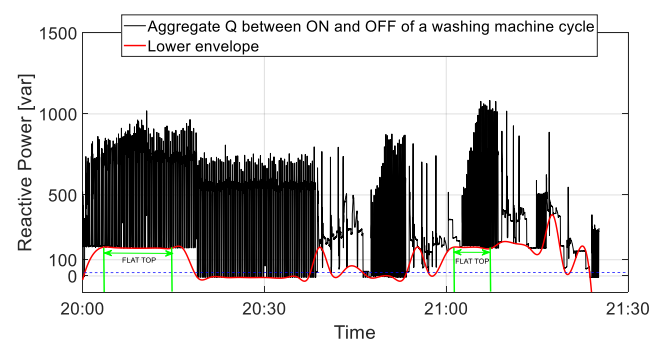


Figure 9. Detections of ON/OFF events for the fridge during a washing machine cycle.

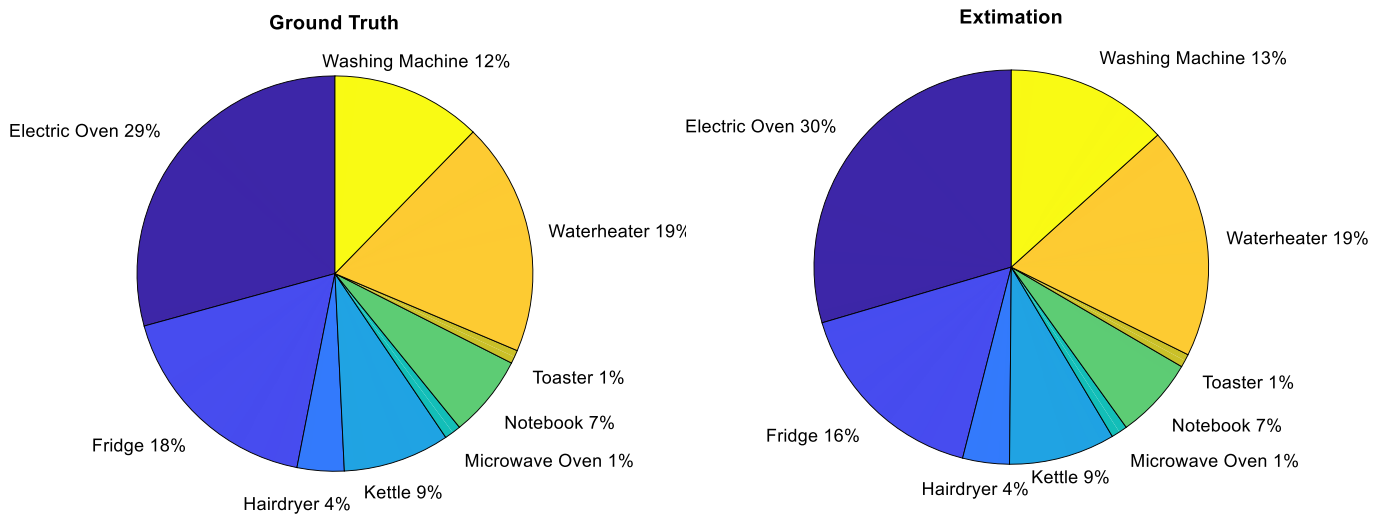


Figure 10. Energy disaggregation results.

refer to appliances characterized by an active power consumption less than 30 W. Since these events are not labeled in BLUED, they have been discarded.

Table 6 and Table 7 report the performance obtained by the algorithm for the event detection and the event classification respectively. All the performance indexes are very high confirming the suitability of the approach to detect sudden state changes. In Table 8 the confusion matrix showing per class accuracy (in percent) is reported. In the confusion matrix all the different lights connected to the phase-A have been grouped into a single class. We can observe that the only classification errors are for the fridge class which is confused with that of the lights and vice versa.

Table 6. Performance metrics for the edge detector tested on BLUED dataset.

Test Set	# ground-truth events	TP	FP	FN	Pr	Re	F1-score
Experimental Data	662	638	9	13	0.99	0.96	0.97

Table 7. Performance metrics for appliances classification tested on BLUED dataset.

Test Set	# detected events	TP	FP	FN	Pr	Re	F1-score
Experimental Data	638	625	13	13	0.98	0.98	0.98

Table 8. Confusion matrix showing per class accuracy (in percent) for the appliances connected to phase-A of BLUED; blue: actual device, grey: classified device, green: correct classifications (TP), orange: false positives.

Class	Fridge	Air compressor	Lights	Kitchen aid chopper	Hair dryer
Fridge	0.99	0	0.01	0	0
Air compressor	0	1	0	0	0
Lights	0.05	0	0.95	0	0
Kitchen aid chopper	0	0	0	1	0
Hair dryer	0	0	0	0	1

Literature reports few contributions for NILM algorithms applied to BLUED data at the same frequency of 1 Hz [13], [14], [18].

In [13] a panel of different machine learning algorithms is applied to appliance classification for a subset of BLUED dataset (Air compressor, Basement, Computer, Garage door, Iron, Kitchen light, Laptop, LCD Monitor, Monitor, Refrigerator, TV). Both Precision and Recall are equal to 79%.

In [14], clustering in $\Delta P - \Delta Q$ plane is solved through a hierarchical approach executed with some manual supervision. The paper reports for phase-A BLUED data an F_1 -score for the event detection and for the appliance classification of about 92% and 88% respectively.

In [18] an event detection algorithm is used to identify the time instant when a sudden increase of active power occurs, indicating a possible turn-on event, and a convolutional neural network (CNN) classifier is applied for event classification. Results reported for BLUED are limited to three appliances (washing machine, fridge, microwave).

Table 9 reports the calculated Recall (also called True Positive rate-TPR) and accuracy metrics for event detection and classification respectively for [13], [14] and [18] and also other solutions applied to higher frequency data [19]-[21].

Table 9. Comparison of event detection and classification performance.

Reference	Method	Sampling-rate	Recall (TPR)	Accuracy
-	Proposed	1 Hz	0.96	0.98
[13]	Active machine learning strategy	1 Hz	0.79	0.92
[14]	Hierarchical approach	1 Hz	0.89	0.88
[18]	First and second differences + CNN	1 Hz	0.94	0.98
[19]	Finite-precision analysis	4 kHz	0.91	-
[20]	Extremely Randomized Trees	12 kHz	-	0.94
[21]	Hybrid approach	60 Hz	0.94	-

It can be seen that the proposed algorithm achieves good results while being simple and computationally efficient. In fact, despite the low frequency of data analyzed, the approach shows competitive performance even when it is compared to other more complex methodologies applied to high-sampling-rate signals, from 60 Hz to 12 kHz [19]-[21].

6. CONCLUSIONS

In this paper, a monitoring system has been proposed that is able to disaggregate and keep track of the power consumption of the devices existing in a typical house analysing low frequency aggregate data.

By applying a hybrid approach to power data, i.e. event-based techniques and pattern recognition techniques for large household appliances, the load disaggregation is performed with good performance, even when complex type-III devices, such as the washing machine, are working.

Finally, using the more populated BLUED dataset, we showed that the proposed procedure is able to achieve high performance in two key tasks in energy disaggregation: identifying events from non-events, and identifying which appliance is associated with a specified event.

REFERENCES

- [1] S. S. Hosseini, K. Agbossou, S. Kelouwani, A. Cardenas, Non-intrusive load monitoring through home energy management systems: A comprehensive review, *Renewable and Sustainable Energy Reviews*, vol. 79, 2017, pp. 1266–1274. DOI: [10.1016/j.rser.2017.05.096](https://doi.org/10.1016/j.rser.2017.05.096)
- [2] A. Ruano, A. Hernandez, J. Ureña, M. Ruano, J. Garcia, NILM Techniques for Intelligent Home Energy Management and Ambient Assisted Living: A Review, *Energies*, vol. 12, June 2019, p. 2203. DOI: [10.3390/en12112203](https://doi.org/10.3390/en12112203)
- [3] W. Kong, Z. Y. Dong, B. Wang, J. Zhao and J. Huang, A practical solution for non-intrusive type II load monitoring based on deep learning and post-processing, *IEEE Trans. Smart Grid*, vol. 11, no. 1, Jan. 2020, pp. 148-160. DOI: [10.1109/TSG.2019.2918330](https://doi.org/10.1109/TSG.2019.2918330)
- [4] A. Harell, S. Makonin, I. V. Bajić, Wavenilm: A causal neural network for power disaggregation from the complex power signal, *Proc. of ICASSP 2019 - IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, UK, 12-17 May 2019, pp. 8335-8339. DOI: [10.1109/ICASSP.2019.8682543](https://doi.org/10.1109/ICASSP.2019.8682543)
- [5] Q. Wu, F. Wang, Concatenate convolutional neural networks for non-intrusive load monitoring across complex background, *Energies*, vol. 12, no. 8, pp. 1572, Apr. 2019. DOI: [10.3390/EN12081572](https://doi.org/10.3390/EN12081572)
- [6] P. Davies, J. Dennis, J. Hansom, W. Martin, A. Stankevicius, L. Ward, Deep neural networks for appliance transient classification, *Proc. of ICASSP 2019 - IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, UK, 12-17 May 2019, pp. 8320-8324. DOI: [10.1109/ICASSP.2019.8682658](https://doi.org/10.1109/ICASSP.2019.8682658)
- [7] B. Cannas, B. Canetto, S. Carcangiu, A. Fanni, L. Fresi, M. Marceddu, C. Muscas, P. Porcu, G. Sias, Non-Intrusive Loads Monitoring Techniques for House Energy Management, *Proc. of the 1st Int. Conf. on Energy Transition in the Mediterranean Area (SyNERGY MED)*, Cagliari, Italy, 28-30 May 2019, pp. 1-6. DOI: [10.1109/SyNERGY-MED.2019.8764104](https://doi.org/10.1109/SyNERGY-MED.2019.8764104)
- [8] B. Cannas, S. Carcangiu, D. Carta, A. Fanni, C. Muscas, G. Sias, B. Canetto, L. Fresi, P. Porcu, Real-Time Monitoring System of the Electricity Consumption in a Household Using NILM Techniques, *Proc. of the 24th IMEKO TC4 International Symposium and 22nd International Workshop on ADC and DAC Modelling and Testing*, Palermo, Virtual, Italy, 14–16 September 2020; pp. 90-95. Online [Accessed 18 June 2021] <https://www.imeko.org/publications/tc4-2020/IMEKO-TC4-2020-18.pdf>
- [9] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, M. Berges, BLUED: A fully labeled public dataset for event-based non-intrusive load monitoring research, *Proc. of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, ACM, Beijing, China, 2012, pp. 1–5.
- [10] H. Liu, Q. Zou, Z. Zhang, Energy Disaggregation of Appliances Consumptions Using HAM Approach, *IEEE Access* 2019, 7, pp. 185977–185990. DOI: [10.1109/ACCESS.2019.2960465](https://doi.org/10.1109/ACCESS.2019.2960465)
- [11] P. Ricardo, P.R.Z. Taveira, C.H.V.D. Moraes, G. Lambert-Torres, Non-Intrusive Identification of Loads by Random Forest and Fireworks Optimization, *IEEE Access* 2020, 8, 75060–75072. DOI: [10.1109/ACCESS.2020.2988366](https://doi.org/10.1109/ACCESS.2020.2988366)
- [12] B. Cannas, S. Carcangiu, D. Carta, A. Fanni, C. Muscas, Selection of Features Based on Electric Power Quantities for Non-Intrusive Load Monitoring, *Applied Sciences*, 2021, 11(2):533. DOI: [10.3390/app11020533](https://doi.org/10.3390/app11020533)
- [13] F. Rossier, Ph. Lang, J. Hennebert, Near Real-Time Appliance Recognition Using Low Frequency Monitoring and Active Learning Methods, *Energy Procedia* 122 (2017), pp. 691–696. DOI: [10.1016/j.egypro.2017.07.371](https://doi.org/10.1016/j.egypro.2017.07.371)
- [14] T. Khandelwal, K. Rajwanshi, P. Bharadwaj, S. Srinivasa Garani, R. Sundaresan, Exploiting Appliance State Constraints to Improve Appliance State Detection, *Proc. of e-Energy '17*, Shatin, Hong Kong, 16-19 May 2017, pp. 111–120. DOI: [10.1145/3077839.3077859](https://doi.org/10.1145/3077839.3077859)
- [15] G. W. Hart, Nonintrusive appliance load monitoring, *IEEE Proc.* 1992, 80, pp. 1870–1891. DOI: [10.1109/5.192069](https://doi.org/10.1109/5.192069)
- [16] Mathworks: k-medoids clustering. Online [Accessed 18 June 2021] <https://it.mathworks.com/help/stats/kmedoids.html>
- [17] Eastrongroup: Energy meters. Online [Accessed 18 June 2021] <http://www.eastrongroup.com>
- [18] C. Athanasiadis, D. Doukas, T. Papadopoulos, A. Chrysopoulos, A Scalable Real-Time Non-Intrusive Load Monitoring System for the Estimation of Household Appliance Power Consumption, *Energies*, vol. 14, Feb. 2021, no. 3. pp. 767. DOI: [10.3390/en14030767](https://doi.org/10.3390/en14030767)
- [19] R. Nieto, L. de Diego-Otón, Á. Hernández, J. Ureña, Finite Precision Analysis for an FPGA-based NILM Event-Detector, *Proc. of the 5th International Workshop on Non-Intrusive Load Monitoring*, Online, 18 November 2020, pp. 30-33. DOI: [10.1145/3427771.3427849](https://doi.org/10.1145/3427771.3427849)
- [20] A. K. Jain, S. S. Ahmed, P. Sundaramoorthy, R. Thiruvengadam, V. Vijayaraghavan, Current peak based device classification in NILM on a low-cost embedded platform using extra-trees, *Proc. of 2017 IEEE MIT Undergraduate Research Technology Conference (URTC)*, Cambridge, MA, 3-5 November 2017, 4 pp. DOI: [10.1109/URTC.2017.8284200](https://doi.org/10.1109/URTC.2017.8284200)
- [21] M. Lu, Z. Li, A Hybrid Event Detection Approach for Non-Intrusive Load Monitoring, *IEEE Transactions on Smart Grid*, vol. 11, no. 1, Jan. 2020, pp. 528-540. DOI: [10.1109/TSG.2019.2924862](https://doi.org/10.1109/TSG.2019.2924862)