A Novel Procedure for Virtual Measurements Generation suitable for Training and Testing in the context of Non Intrusive Load Monitoring

Alaa Saleh, Pirmin Held and Dirk Benyoucef
Furtwangen University, Germany
Mechanical and Medical Engineering
Email:{alaa.saleh, prmin.held, dirk.benyoucef}@hs-furtwangen.de

Djaffar Ould Abdeslam
Universite de Haute-Alsace, France
IUT de Mulhouse
Email: djaffar.ould-abdeslam@uha.fr

***Abstract*— Getting "smarter" Energy by means of advanced electrical and computer engineering tools is the theme of our digital age. This paper presents a new concept for virtual data generation in the context of non-intrusive load monitoring, where the goal is to fill in the gap when aggregate measurements are needed along with individuals ones. We develop a method to generate aggregate measurements starting from single original measurements. The performance of standard NILM tools and algorithms on both original and "virtual" data is compared using own lab measurements.**

Keywords- Smart meters; Non-intrusive load monitoring; Event detection; feature exrtraction; classification; signal processing.

I. INTRODUCTION

Self-healing reliable smart Grids are sought-after all over the world. Non-Intrusive load monitoring is one of the major tools for such ambition; as it delivers detailed information of the energy consumption in a certain facility/building, which can be very useful for various objectives.

Ever since Hart's pioneering work[1] , several approaches were presented over the years [2]; but they were mostly tailored to specific measurements, be it self-collected or from publically available data sets; such measurements were made with a specific application in mind, making them not entirely suitable for extracting and processing of newly adapted features, e.g., harmonical properties in transient states of the current signal [3]. Machine learning algorithms along with advanced signal processing techniques are ready to exploit such features to deliver the best possible identification of electrical appliances.

While several datasets were made publically available in the last decade, they were still not optimal for extracting important features such as harmonical properties. The majority of them were limited to aggregate measurements only or single measurements only, or lacking enough sampling rate in one of them, as shown in Table1.

The question to be addressed in this work is: if only single measurements are available, can we use them for 'synthesizing' whole measurements without actually measure again? it will suffice then for producers to provide interested customers with a list of stand-alone measurements of their device, enabling them to test many scenarios of consumption before actually putting any constellation of devices together,

TABLE I. PUBLIC DATASETS WITH MEASUREMENTS TYPES

Dataset	Type of measurements
REDD	whole-house measurements only [4]
BLUED	whole-house measurements only [5]
PLAID	individual measurements only [6]
UKDALE	whole-house measurements, single measurements with low sampling rate [7]
WHITED	individual measurements only [8]
COOLL	individual measurements only [9]

which is of high importance for big factories and sensitive facilities.

This paper introduces a procedure to generate an approximately identical version of the real whole measurements using single measurements only. The similarity of the resulted signal to the original one will be judged according to the perspective of NILM algorithms.

The remainder of the paper will be organized as follows; In section II, the general data-construction scheme is presented and basic concepts behind it are explained, Section III is devoted to numerical comparisons and tests. Finally, Section IV briefly presents the main conclusions of our study.

II. GENERAL SCHEME AND BASIC CONCEPTS

A flow chart for the proposed procedure is depicted in Figure 1.

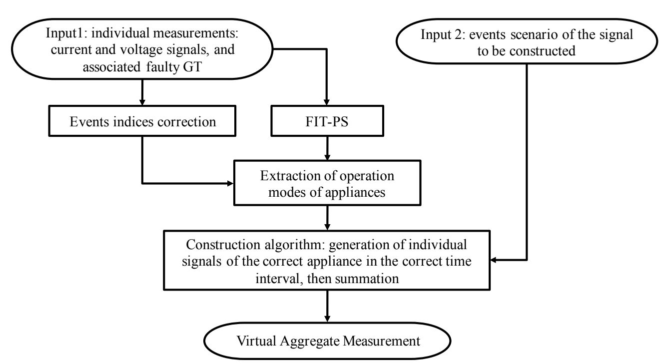

Figure 1. General scheme for virtual aggregation procedure

Figure 1 can be translated to a step-by-step plan for constructing the desired virtual aggregate signal out of real individual measurements:

Step 1. Convert real single measurements to " Frequency Invariant Transformation of Periodic Signals " representation (FIT-PS) [10].

Step 2. Correct event indices.

Step 3. Extract operational modes for each appliance.

Step 4. Generate the targeted current signals according to the desired scenario, adding steady-state periods to meet the desired length, as in Figure 2.

Step 5. Sum the generated current signals according to the desired scenario.

In the following, we will clarify the concepts behind those steps:

A. Phase preserving reconstruction

The general model for reconstructed signal can be written as:

$$\mathrm{I_t} = \sum_{\mathrm{i}=1}^{\mathrm{N}} \delta_{\mathrm{i}} \mathrm{I_i} + v \tag{1}$$

where $\mathrm{I_t}$ is the aggregate current signal,
N is the number of the appliances,
$\mathrm{I_i}, \mathrm{i} = 1, \cdots, \mathrm{N}$: are the current signals of the individual appliances
$\delta_{\mathrm{i}}, \mathrm{i} = 1, \cdots, N$ indicates the state of the corresponding appliance as:

$$\delta_{\mathrm{i}} = \begin{cases} 1, & \text{if the } \mathrm{i}^{\mathrm{th}} \text{ appliance is 'ON'} \\ 0, & \text{if the } \mathrm{i}^{\mathrm{th}} \text{ appliance is 'OFF'} \end{cases} \tag{2}$$

v is the additive noise term, which will be assumed to be AWGN with zero mean and appropriate variance.

the loads of the appliances are complex in general, which creates a phase shift between the current signal and the respective voltage signal, this shift can take different values for different appliances.

If we assume the total signal to be composed of two appliances only, and their current signals are given by:

$$\mathrm{I_1} = \mathrm{A}e^{\mathrm{j}\theta_1},\ \mathrm{I_2} = \mathrm{A}e^{\mathrm{j}\theta_2} \tag{3}$$

Then the total current signal will be obtained as:

$$\mathrm{I} = 2\mathrm{A}e^{\mathrm{j}\left(\frac{\theta_1+\theta_2}{2}\right)} \cos\left(\frac{\theta_1 - \theta_2}{2}\right) \tag{4}$$

$$|\mathrm{I}| = 2\mathrm{A}\left|\cos\left(\frac{\theta_1 - \theta_2}{2}\right)\right| \tag{5}$$

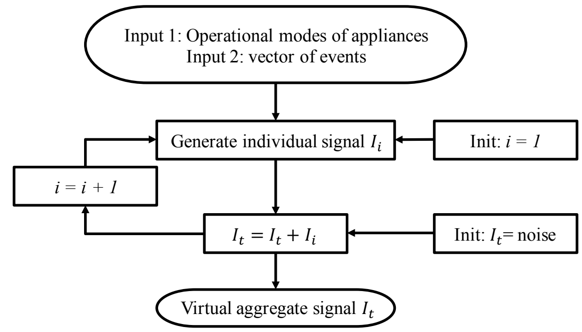

Figure 2. Construction algorithm

Equation (5) shows the dependence of the total signal amplitude on the phase difference between the signals $\mathrm{I_1}$ and $\mathrm{I_2}$, which implies that a simple addition as in (1) will not be correct unless all current signals are added with emphasis on their relative phase.

Fortunately, voltage signal stays almost intact and can be considered a reliable phase reference, so current signals of the individual appliances are converted to FIT-PS representation, which takes the zero-crossing point (from negative to positive) in the voltage signal as a phase reference[10].

B. operational modes extraction

Some appliances have more than one operational mode in general, where we define an operational mode as a distinctive subset of the current signal of a certain device; that can comprise both transient state and part of a steady state. An example is depicted in Figure 3 for a refrigerator.

These operational modes differ from each other and must be taken into account when constructing the virtual aggregate signal. Each operational mode has a different power level in its transient or steady state or both, which can be used to choose the most appropriate one in the construction procedure. Simple structure appliances have only one operational mode in general, e.g., Lamp.

The individual measurements of a certain appliance should be long enough to be able to extract all of their operational modes. The extraction is done as follows:

- For each event, we take a window starting at the event index and containing both transient state and a part of the steady state.
- The power level in the steady state of that window is calculated, then compared with the levels of the previously extracted operational modes of the appliance; it will be added as a new mode if it differs from those previously detected.
- a new sequence number is assigned to the new mode and will be checked again in other occurrences (as the current signal drawn by some appliances will change its operational modes in a sequential manner) .

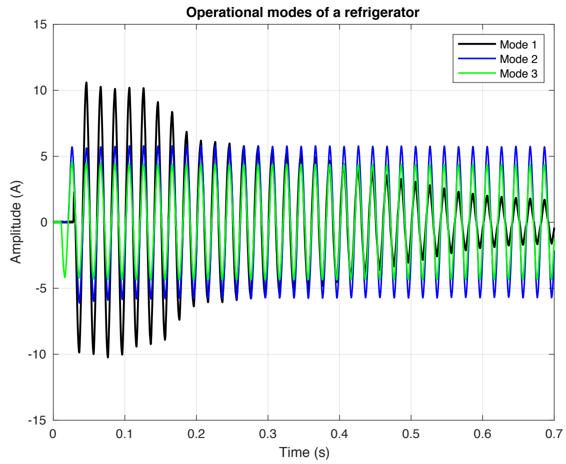

Figure 3. Example of 3 operational modes of a refrigerator

C. *Event indices correction*

Dealing with faulty reference data is an inevitable problem with the vast majority of datasets, so we need to be sure that all events are precise to avoid errors when extracting operational modes.

To correct an event index:

- A window around the possible inaccurate index is taken, it should contain only one event, and be long enough to account for the supposed maximum error (to be set according to training data if possible).
- The event type is checked, if it is an 'off' event, it will be flipped to unify the correction procedure.
- The average between the maximum amplitude of the first few periods in the window and its global maximum is calculated and set as new threshold (yielding better accuracy for low power appliances compared to averaging on the peaks of the last periods in the window).
- The precise event index is the first index in the window where the amplitude exceeds the calculated threshold.

III. NUMERICAL EXAMPLES AND TESTS

In this section, Numerical experiments are to be conducted to test the hypothesis of "suitability of the resulted virtual aggregate signal for training and testing NILM algorithms ".

To do that, the performance of standard NILM algorithms on such generated data will be compared against their performance on an original aggregate signal that has the same scenario of events and appliances. Figure 4 shows a classical chain of processes for typical NILM system.

The comparison will be conducted between the results of standard event detection, feature extraction and classification algorithms on both original aggregate measurements and constructed aggregate measurements. we will use own measurements since the available public datasets are incomplete for our purpose, and to avoid any internal source of error in their reference data.

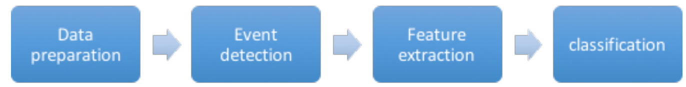

Figure 4. Typical sequence of processes in NILM system

A. *Collecting single & aggregate measurements*

The measurement system developed previously at our lab [11] was used to collect individual and aggregate measurements; as it enables us to define the set of switching events beforehand. The sampling frequency is set to 4 KHz for both current and voltage signals. Those measurements are part of a public Dataset to be made available in the near future.

Tab. II contains a list of all appliances considered for our tests, this chosen group includes high and low power appliances.

The real measurements were conducted as follows:

- Single measurements for each appliance were taken with 80 pairs of On-Off events, each active cycle was kept running 10 seconds to get transient state and part of the steady state.
- Aggregate measurements, where appliance were switched On and Off randomly, but only 4 appliances can be active on the same time at most.
- A minimum temporal distance of 3 seconds was kept between events.

The virtual aggregate measurements were then generated according to the same scenarios of events followed in the real ones. In Figure 5, an example of those measurements is depicted.

Now that both types of measurements are available, we move to testing NILM tools.

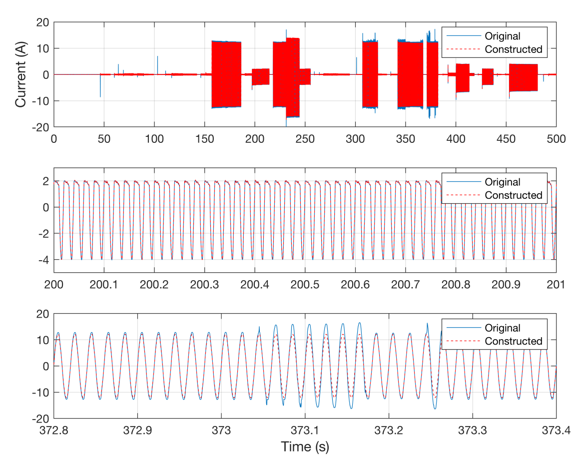

Figure 5. Original and constructed signals together (top), with 2 magnified reigns where they were identical (middle), and most different (bottom)

 ISBN: 978-1-61208-638-5

TABLE II. LIST OF APPLIANCES

ID	NAME	NOMINAL POWER (W)
1	Radio	6
2	Heat gun (setting 1)	820
3	Router	9
4	Black desk lamp	20
5	Light bulb box	20
6	Kettle	2100
7	Black hairdryer (setting 1)	500
8	Fan	22
9	Rotary tool (Dremel)	30
10	LED lamp	1

A. Event Detection

Two different event detectors were used here, the first one is based on the work of Hart [1], while the second is an FBE Event detector [12]. they were applied on the original measurements as well as the constructed ones, and the results are given in the tables III and IV respectively

the following performance metrics were used:

$$P_{recall} = \frac{TP}{FN+TP} \quad , \quad P_{precision} = \frac{TP}{FP+TP} \qquad (6)$$

where *TP* is the number of true detected events,
FN is the number of missed events,
FP is the number of false alarms.

The threshold for Hart Event detector was set to 15W as going below that will result in detecting too many false events due to the high variations of the noise.

Appliances below the chosen threshold were treated as noise in all event-detection tests since the comparison is the main goal here. For FBE detector, several thresholds were chosen and the results are shown in Table IV.

from the tables, both event detectors show a slightly better performance on the constructed measurements (less than 5%).

TABLE III. HART EVENT DETECTOR RESULTS

Signal Type	Thr (W)	Total Event	TP	FN	FP	P_{recall} %	$P_{precision}$ %
Original	1	1600	1411	189	3417	88.19	29.23
Constructed	1	1600	1448	152	5542	90.50	20.72
Original	15	1120	958	162	31	85.54	96.87
Constructed	15	1120	959	161	0	85.62	100

TABLE IV. FBE-BASED EVENT DETECTOR RESULTS

Signal Type	Thr (W)	Total Event	TP	FN	FP	P_{recall} %	$P_{precision}$ %
Original	1	1600	1439	161	135	89.94	91.42
Constructed	1	1600	1480	120	97	92.50	93.85
Original	5	1440	1285	155	53	89.24	96.04
Constructed	5	1440	1322	118	7	91.81	99.47
Original	15	1120	1113	7	37	99.38	96.78
Constructed	15	1120	1090	26	0	97.32	100

B. Feature extraction

we started with a set of steady state features at first, as in most classical NILM literature[2,3], active and reactive powers P and Q were chosen, along with the mean power of the harmonics.

if multiple appliance are active at the same time, a subtraction procedure is done at each event to extract a steady state window suitable for calculating those features of interest, as shown in Figure 6.

while taking the phase and the transient state length into account (assuming that events are well separated which is the case for these measurements).

Apparent, active and reactive powers can then be calculated as:

$$P_A = V_{RMS} * I_{RMS} \qquad (7)$$

$$P = P_A * \cos\theta \qquad (8)$$

$$Q = P_A * \sin\theta \qquad (9)$$

where θ is the phase difference between voltage and current signals.

A band-pass filter is applied on the current signal for each harmonic component $f_{c,n} = nf_0$, where n is the harmonic number, and f_0 is the fundamental frequency. Then the mean harmonic power is given by:

$$P_n = V_{RMS} * I_{n,RMS} \qquad (10)$$

In Figure 6, we compare the selected steady-state features: P_A, P, Q, P_2, P_3, P_4 for two different appliances, a kettle which consumes high power (~ 2 KW), and a Radio that consumes 6 W only, where in each sub-plot, the horizontal axis is for event indices while the vertical is for the power.

while the signals (features) are not identical, yet they are very close to each other, which will be reflected in classification results in the next section.

C. classification

Two standard classifiers [13] were implemented to compare their performance on both types of measurements, the first one is based on a feed-forward neural network (FFNN) while the other one is a support vector machine (SVM) classifier.

both classifiers were trained using extracted features from individual measurements, then tested on the original and constructed virtual measurements respectively.

To simplify the classification process, a power threshold of 25 W was applied to divide the appliances into 2 groups: high power appliances and low power appliances; the features are then fed to the corresponding FFNN subnet (or SVM sub-classifier respectively).

In Table V and VI, results of both classifiers are listed.

Both classifiers delivered slightly better results when applied on constructed measurements, which can be attributed to the higher noise level in the original measurements and the fact that their steady-state features may have more variations and outliers.

 ISBN: 978-1-61208-638-5

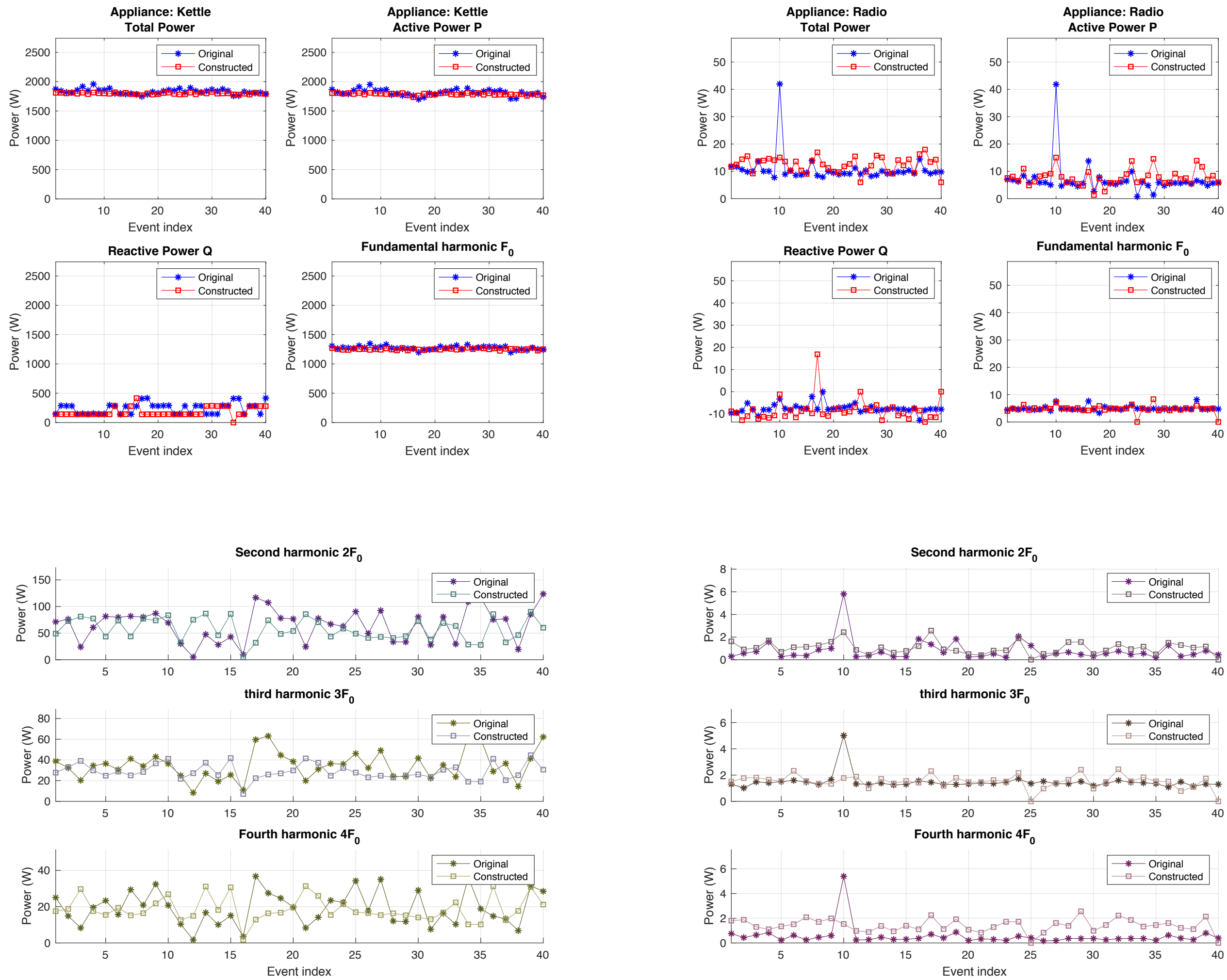

Figure 6. steady state features comparison for : a high power appliance (Kettle, left), a low power appliance (Radio, right)

TABLE V. PERFORMANCE OF FFNN CLASSIFIER

Classifier: FFNN		
%	*High power applinces*	*Low power appliances*
Original	*96.9*	*87.9*
Constructed	*100*	*89.6*

TABLE VI. PERFORMANCE OF SVM CLASSIFIER

Classifier: SVM						
	High power applinces			*Low power appliances*		
%	*RBF*	*Linear*	*Polynomial*	*RBF*	*Linear*	*Polynomial*
Original	*89.7*	*90.6*	*96.3*	*85.4*	*81.9*	*86.3*
Constructed	*99.4*	*82.2*	*100*	*90.6*	*78.8*	*90.6*

In Figure 7, the detailed confusion matrices for FFNN classifier are also shown, where the two tables on the right side are for constructed measurements, and the tables on the left side are for original measurements.

FFNN correctly classified nearly all of the samples in the test set on both types of measurements in the high power category. Although the overall accuracy rate is very close in low power category, yet a closer look yields differences in misclassification rate among the individual appliances, for example: the Radio (ID=1) was misclassified as Router only once in the original measurements table, while it was misclassified 8 times as Router in the constructed measurements table. This can be attributed to the difference between the modeled additive noise and the real one.

Figure 7. Confusion matrices for FFNN classification results on original measurements (left), and constructed measurements (right)

IV. Conclusion and Future Work

This paper presents a promising approach for constructing virtual aggregate measurements from original single measurements, enabling better use of available NILM datasets for training and testing Disaggregation algorithms.

Application of standard algorithms for event detection and classification on both types of measurements showed very similar performance. In the future, the noise model will be improved further and virtual single measurements will be tested in a similar manner.

The concept of phase-preserving summation/subtraction will be also used in different disaggregation approaches, e.g., predictive maintenance.

Acknowledgment

This work was created as part of the iMon project (funding number 03FH001IX4).

References

[1] G. Hart, “Residential energy monitoring and computerized surveillance via utility power flows,” IEEE Technology and Society Magazine, vol. 8,no. 2, Jun. 1989, pp. 12–16.

[2] Hosseini, Sayed Saeed, Kodjo Agbossou, Sousso Kelouwani, and Alben Cardenas, "Non-intrusive load monitoring through home energy management systems: A comprehensive review." Renewable and Sustainable Energy Reviews 79 (2017): 1266-1274.

[3] Du, Yi, Liang Du, Bin Lu, Ronald Harley, and Thomas Habetler, "A review of identification and monitoring methods for electric loads in commercial and residential buildings." In Energy Conversion Congress and Exposition (ECCE), 2010 IEEE, pp. 4527-4533.

 ISBN: 978-1-61208-638-5

[4] J. Kolter, and M. Johnson. “REDD: A public data set for energy disaggregation research, ” In Workshop on Data Mining Applications in Sustainability (SIGKDD), 2011

[5] K. Anderson, A. Ocneanu, D. Benitez, A. Rowe, and M. Berges, “BLUED: A fully labeled public dataset for event-based non-intrusive load monitoring research,” in Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD), 2011.

[6] J. Gao, S. Gini E. Kara and M. Bergés, “PLAID: a public dataset of high-resoultion electrical appliance measurements for load identification research: demo abstract,” in Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, 2014.

[7] J. Kelly, W. Knottenbelt, “The UK-DALE dataset domestic appliance-level electricity demand and whole-house demand from five UK homes,” Scientific Data, vol. 2, 2015, [online] Available at: https://arxiv.org/abs/1404.0284.

[8] M. Kahl, A. U. Haq, T. Kriechbaumer, and H. A. Jacobsen. “Whited-a worldwide household and industry transient energy data set,” In Proceedings of the 3rd International Workshop on Non-Intrusive Load Monitoring (NILM), 2016.

[9] T. Picon, M.N. Meziane, P. Ravier, G. Lamarque, C. Novello, J.C.L. Bunetel, and Y. Raingeaud, “COOLL: Controlled On/Off Loads Library, a Public Dataset of High-Sampled Electrical Signals for Appliance Identification,” arXiv preprint arXiv:1611.05803.

[10] P.Held, A. Saleh, D.O. Abdeslam and D. Benyoucef. "Frequency Invariant Transformation of Periodic Signals (FIT-PS) for high frequency Signal Representation in NILM", In BW-CAR| SINCOM, 2016.

[11] T. Bier; D. Ould Abdeslam; J. Merckle; and D. Benyoucef, “Smart meter systems detection & classification using artificial neural networks,” 38th Annual Conference on IEEE Industrial Electronics Society, IECON 2012, 2012, pp. 3324 – 3329.

[12] A.Saleh, P. Held, D. Benyoucef, and D. O. Abdeslam. "EMD inspired filtering algorithm for signal analysis in the context of non intrusive load monitoring." In Industrial Electronics Society, IECON 2017-43rd Annual Conference of the IEEE, pp. 3615-3620.

[13] K.Basu,"Classification techniques for non-intrusive load monitoring and prediction of residential loads." PhD diss., Université de Grenoble, 2014.

 ISBN: 978-1-61208-638-5