SEQUENTIAL CLUSTERING-BASED EVENT DETECTION FOR NON-INTRUSIVE LOAD MONITORING

Karim Said Barsim and Bin Yang

Institute of Signal Processing and System Theory,
University of Stuttgart, Germany
karim.barsim@iss.uni-stuttgart.de
bin.yang@iss.uni-stuttgart.de

ABSTRACT

The problem of change-point detection has been well studied and adopted in many signal processing applications. In such applications, the informative segments of the signal are the stationary ones before and after the change-point. However, for some novel signal processing and machine learning applications such as Non-Intrusive Load Monitoring (NILM), the information contained in the non-stationary transient intervals is of equal or even more importance to the recognition process. In this paper, we introduce a novel clustering-based sequential detection of abrupt changes in an aggregate electricity consumption profile with accurate decomposition of the input signal into stationary and non-stationary segments. We also introduce various event models in the context of clustering analysis. The proposed algorithm is applied to building-level energy profiles with promising results for the residential BLUED power dataset.

KEYWORDS

Event detection, change-interval detection, density-based clustering, DBSCAN, non-intrusive load monitoring, NILM, BLUED, energy disaggregation

1. INTRODUCTION

Non-Intrusive Load Monitoring (NILM), also known as electricity disaggregation, is an energy monitoring technique that aims at inferring the energy consumption profiles of individual electrical loads merely from a single or a limited number of aggregate measurement points in a building [1]. Recently, NILM has witnessed a rapidly increasing progress in both academic and commercial research due to its promising applications in energy conservation, activity monitoring [2], dynamic pricing [3], demand forecasting [4], and home automation [5]. Currently, the majority of NILM systems are event-based approaches in the sense that they rely on the detection of abrupt changes occurring in the aggregate signal which indicate state-changes of the monitored appliances. It was observed that events attain distinctive features according to the physical properties of their appliances such as energy storage elements, counter-electromotive force in induction motors, or striking voltages in fluorescent lamps. Features extracted from steady and transient intervals (such as power surges, overshoot currents, decay rate, etc) are utilized in event clustering or classification stages of the disaggregation system. Consequently, a robust detection

Jan Zizka et al. (Eds) : CCSIT, SIPP, AISC, CMCA, SEAS, CSITEC, DaKM, PDCTA, NeCoM - 2016
pp. 77–85, 2016. © CS & IT-CSCP 2016 DOI : 10.5121/csit.2016.60108

and accurate segmentation of such change-intervals is of particular importance for event-based NILM systems.

Basseville and Nikiforov [6] described various detection algorithms from which two approaches have been utilized in event-based NILM systems, namely the Generalized Likelihood Ratio (GLR) test [7, 8] and the CUmulative SUM (CUSUM) filtering [9]. Jin et al. [10] proposed a more robust change-point detection approach based on a Goodness-of-Fit (GoF) test. In addition, various machine learning tools such as kernel clustering [11], Hidden Markov Models (HMM) [12], and Support Vector Machines (SVMs) [13], have been proposed as solutions to address the change point detection problem.

Even though many previous works on NILM proposed utilizing features extracted from the transient intervals, only few event detection approaches consider accurate segmentation of the transient periods for the extraction of more stable transient features [9, 14]. Moreover, many approaches need a probabilistic model for the sample distribution in the stationary segments which is often difficult to obtain from aggregate consumption profile of several, simultaneously operating appliances. The result is that the current event detection algorithms are not robust and fail sometimes provide reliable event-based feature for appliance recognition in practice. In this paper, we propose a novel clustering-based event detection algorithm for event-based NILM systems. In contrast to other event detection algorithms, the proposed approach features accurate segmentation of the input signal into stationary (steady) and non-stationary (transient) segments. Such accurate segmentation is crucial for the extraction of more stable and repeatable features from both transient and steady-state intervals. Moreover, the utilized density-based clustering scheme does not impose any probabilistic models on the sample distribution in either of the stationary segments and supports arbitrarily shaped, weakly stationary segments leading to an enhanced robustness to noise. In addition, the proposed algorithm features a sequential (instead of batch) clustering that is more efficient for real-time NILM systems.

The presented approach is modular in the sense that it can combine any clustering-based event detection algorithm with any event model. For this purpose, we also introduce different event models at different complexity- and robustness-levels. This paper is organized as follows. In section 2, we introduce different event models in the context of spatial and time-series clustering. In section 3, we describe the proposed sequential event detection algorithm in which the Density-Based Spatial Clustering for Applications with Noise [15] is assumed and utilized sequentially in spatial and temporal analysis of the input power signals. Section 4 shows results of application of the proposed algorithm on the publicly available, residential BLUED [16] dataset. Finally, section 5 concludes this paper.

2. EVENT MODELS

Event models will be introduced in the order of their increasing coverage of real events, robustness, and complexity.

Let the matrix

$$\mathbf{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N], \;\; \boldsymbol{x}_n \in \mathbb{R}^l \qquad (1)$$

contain a time series of N consecutive l-dimensional data samples (feature vectors). Typically, $\boldsymbol{x}_n$ contains the measured real P and reactive Q powers at time instance n. Assume that all N samples have been clustered into m non-empty, disjoint clusters (sets) $C_1, C_2, \dots, C_m$. In addition, we assume that a noise-aware clustering algorithm assigns un-clustered samples (i.e. outliers or noisy samples) to the set C_0. Clearly, $\sum_{i=0}^{m}|C_i| = N$ where $|C_i|$ is the cardinality of the cluster $|C_i|$. Let

$$y_n = \omega(\boldsymbol{x}_n) \in \{0,1,2,\dots,m\} \tag{2}$$

be the corresponding cluster index of $\boldsymbol{x}_n$ (i.e. $\boldsymbol{x}_n \in C_{y_n}$). We then introduce the following definitions for two metrics of a cluster and three different event models:

Definition 1: The *temporal length* $\mathrm{Len}(C_i)$ of cluster C_i is defined as the minimum window size that contains all its elements. If

$$\exists u\colon\ \boldsymbol{x}_u \in C_i \ \text{ and }\ \boldsymbol{x}_n \notin C_i \ \ \forall(n < u) \tag{3}$$

$$\exists v\colon\ \boldsymbol{x}_v \in C_i \ \text{ and }\ \boldsymbol{x}_n \notin C_i \ \ \forall(n > v) \tag{4}$$

Then $\mathrm{Len}(C_i)$ is defined as

$$\mathrm{Len}(C_i) = v - u + 1 \quad \geq \ |C_i| \tag{5}$$

Here u and v denote the time instances of the first and last samples belonging to C_i, respectively.

Definition 2: The *temporal locality* ratio $\mathrm{Loc}(C_i)$ of cluster C_i is defined as

$$\mathrm{Loc}(C_i) = \frac{|C_i|}{\mathrm{Len}(C_i)} \quad \in\ ]\,0,1\,] \tag{6}$$

The temporal locality ratio is a measure of how a cluster is spreading over time domain. A value of one ($\mathrm{Loc}(C_i) = 1$) refers to the maximum temporal locality where the cluster is represented by a single segment of consecutive observations. This measure is utilized later in the event models as a means to control the amount of noisy samples permitted in the stationary segments.

Event model $\mathcal{M}_1$: In this event model, a sequence of samples **X** is defined as an event if

(a) it does not contain any noisy samples (i.e. $C_0 = \phi$),

(b) it contains two clusters C_1 and C_2 (i.e.$m = 2$),

(c) both clusters do not interleave (overlap) in the time domain[1],
(i.e. $\exists u : \boldsymbol{x}_n \in C_1\ \forall(n \leq u)$ and $\boldsymbol{x}_n \in C_2\ \forall(n > u)$).

This is the simplest event model without any outliers. It consists of two stationary segments $\mathbf{X}_{s1} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_u]$ and $\mathbf{X}_{s2} = [\boldsymbol{x}_{u+1}, \boldsymbol{x}_{u+2}, \dots, \boldsymbol{x}_N]$. The segment $\mathbf{X}_t = [\boldsymbol{x}_u, \boldsymbol{x}_{u+1}]$ (including the last sample of $\mathbf{X}_{s1}$ and the first one of $\mathbf{X}_{s2}$) is called the change-interval of the event and u is the change point. In other words, an event $\mathcal{M}_1$ is a change interval of length two surrounded by two noise-free weakly stationary segments. This model is valid for switch-off events of most loads as well as switch-on events of resistive ones in a noise-free power signals.

[1] For simplicity, and without loss of generality, we assume that the first and second stationary segments of an event are assigned to the cluster sets $C_1$and C_2, respectively.

Figure 1(a) shows an example of a signal segment matching the first event model $\mathcal{M}_1$ where the scalar samples $x_n \in \mathbb{R}$ and their corresponding cluster indices $y_n = \omega(x_n) \in \{1, 2\}$ are plotted over time. The signal represents a step-like event that consists of two stationary segments (red, solid) and a change interval (blue, dashed).

Event model $\mathcal{M}_2$: A sequence of samples $\mathbf{X}$ is defined as an event if

a) it contains two clusters C_1 and C_2 (i.e. $m = 2$) and the outliers set C_0 is not necessarily empty allowing noisy samples,

b) both clusters C_1 and C_2 show a high temporal locality ratio, i.e. $Loc(C_i) \geq 1 - \epsilon,\ \text{for}\ i = 1, 2$

c) both clusters do not interleave in the time domain, i.e.

$$\exists u, v > u\text{: } \boldsymbol{x}_n \in C_0 \cup C_1\ \forall(n < u)\ \text{and}\ \boldsymbol{x}_u \in C_1, \text{and}$$
$$\boldsymbol{x}_v \in C_0 \cup C_2\ \forall(n > v)\ \text{and}\ \boldsymbol{x}_v \in C_2$$

Compared with $\mathcal{M}_1$, this event model permits noisy samples (i.e. outliers) as well as a lengthy transient interval. This, however, requires the utilization of a noise-aware clustering algorithm. By definition, $\boldsymbol{x}_n \in C_0, \forall(u < n < v)$. In this case, the event contains two stationary segments $\mathbf{X}_{s1}$ and $\mathbf{X}_{s2}$ consisting of samples belonging to C_1 and C_2, respectively, and a change-interval $\mathbf{X}_t = [\boldsymbol{x}_u, \boldsymbol{x}_{u+1}, \ldots, \boldsymbol{x}_{v-1}, \boldsymbol{x}_v]$.

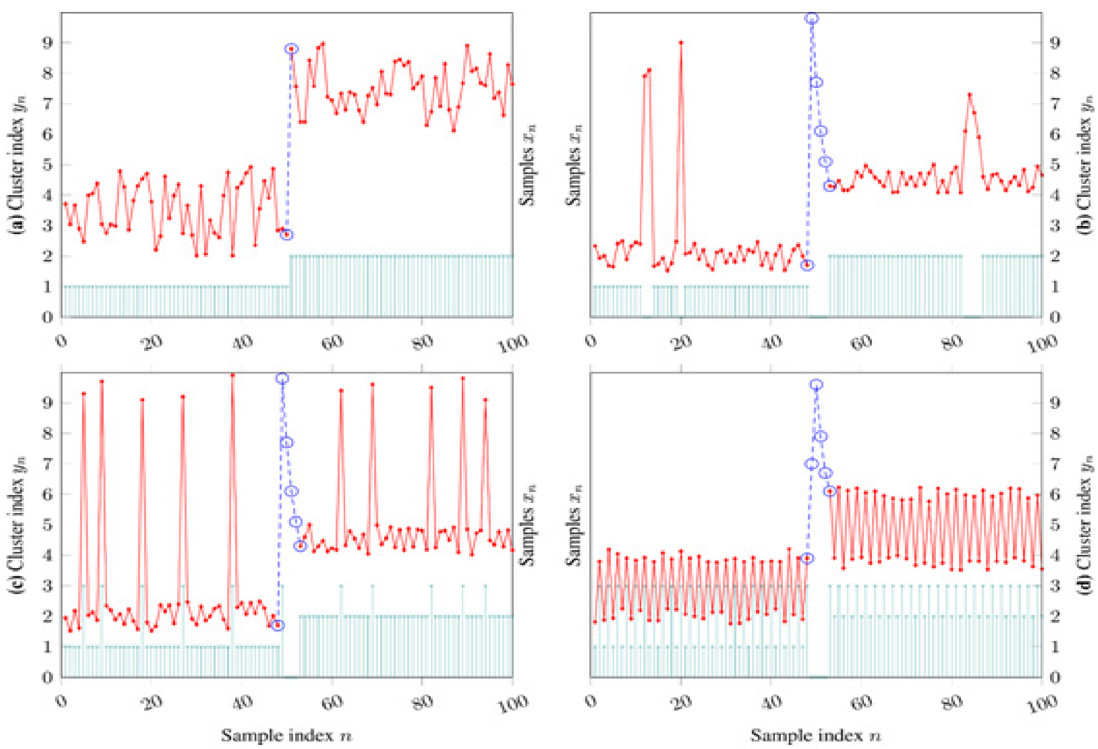

*Figure 1: 1-dimensional signals highlighting differences between the three event models. **(a)** shows a step-like event that is free of both outliers and a transient interval. In **(b)** random outliers as well as a transient interval are permitted. **(c)** shows a repeated pattern of spikes that eventually cluster in C_3. Finally, **(d)** shows high fluctuations in stationary segments leading to the third cluster C_3 as well. The third event model $\mathcal{M}_3$ fits all segments, the second event model $\mathcal{M}_2$ fits only **(a)** and **(b)**, wherease the first model $\mathcal{M}_1$ fits only **(a)**.*

Figure 1(b) shows an example of a signal segment matching the second event model M_2 (but not the first one $\mathcal{M}_1$) where the event contains a slower transient interval in a noisy signal. Even though $\mathcal{M}_2$ is valid for most of the switch-on/off and state-change events within noisy signals, it actually has one implicit assumption on the noise. The assumption that $m = 2$ (maximally two clusters representing two stationary segments) implies that the noise is random and does not contain a repeated pattern that eventually builds up a cluster when projected to the PQ-plane. This is not always the case as shown in the third example in Figure 1(c).

In the aggregate power signal, some appliances trigger a repeated, sometimes periodic, pattern of high fluctuations or spikes. Such repeated patterns tackle the detection of other actual events. This masking behaviour is resolved in the third event model.

Event Model $\mathcal{M}_3$: A sequence of samples **X** is defined as an event if

(a) it contains *at least* two clusters C_1 and C_2 (i.e. $m \geq 2$) and the outliers set C_0 is not necessarily empty,

(b) clusters C_1 and C_2 show a high temporal locality ratio, i.e. $Loc(C_i) \geq 1 - \epsilon,$ for $i = 1, 2$

(c) clusters C_1 and C_2 do not interleave in the time domain, i.e.

$$\exists u, v > u: \boldsymbol{x}_n \notin C_1 \ \forall (n > u) \text{ and } \boldsymbol{x}_u \in C_1, \text{and}$$
$$\boldsymbol{x}_v \notin C_2 \ \forall (n < v) \text{ and } \boldsymbol{x}_v \in C_2$$

In this model, the limitation on the clustering cardinality is released and therefore a repeated noise pattern that eventually results in a wide (temporally wide) cluster would not mask events occurring in the same interval. Similar to $\mathcal{M}_2$, the sequence in this model contains two stationary segments $\mathbf{X}_{s1}$ and $\mathbf{X}_{s2}$ consisting of samples belonging to C_1 and C_2 respectively, and a change interval consisting of $\mathbf{X}_t = [\boldsymbol{x}_u, \boldsymbol{x}_{u+1}, \dots, \boldsymbol{x}_{v-1}, \boldsymbol{x}_v]$.

Figure 1 (b) and (c) show two event segments fit only by $\mathcal{M}_3$. Figure 1(a) shows the simplest event which is fit by all defined models. In Figure 1(b), the transient period as well as the noisy spikes can only be fit by $\mathcal{M}_2$ and $\mathcal{M}_3$. Finally, the repeated noise patten in Figure 1(c) or high fluctuations in Figure 1(d) only match the last event model $\mathcal{M}_3$.

2. DETECTION ALGORITHM

The main task of the event detection algorithm is to search for signal segments that match a given event model $\mathcal{M}_i$. This is achieved by applying a clustering algorithm on different segments and checking how much each segment matches the model. In all of the three models introduced in section 2, the clustering cardinality m is not known in advance. Therefore, a utilized clustering algorithm should either be nonparametric or a model order estimation step has to take place beforehand.

In our approach we utilized the commonly used Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [15]. The DBSCAN algorithm (or density-based clustering in general) has several advantages that make it the best candidate for a non-parametric sequential event detection. First, DBSCAN assumes no prior knowledge of the number of clusters. Second, DBSCAN supports arbitrarily shaped clusters with no constraints on their samples' distribution. In addition, DBSCAN is a noise-aware clustering algorithm and, therefore, can be utilized with any of the previously defined event models.

Ideally, the detection algorithm searches the input signal sequentially for segments that match a given event model. However, we control the matching process with a proximity measure that shows how much a segment matches the given model.

Definition 3: The model loss between an event model $\mathcal{M}_i$ and a signal segment **X** is defined as

$$\begin{aligned}\mathcal{L}(\mathcal{M}_\mathrm{i}, \mathbf{X}, u, v) = &|\{\boldsymbol{x}_n : n \leq u \text{ and } \boldsymbol{x}_n \in C_2\}| + \\ &|\{\boldsymbol{x}_n : n \geq v \text{ and } \boldsymbol{x}_n \in C_1\}| + \\ &|\{\boldsymbol{x}_n : u < n < v \text{ and } \boldsymbol{x}_n \in C_1 \cup C_2\}|\end{aligned} \tag{7}$$

where u and v are the indices of the first and last sample of the change-interval, respectively. In the case of $\mathcal{M}_1$ where $v = u + 1$, the last term in Equation 7 becomes zero regardless of u.

The model loss function counts the number of samples that need to be corrected (i.e. reassigned to a different set C_j of the clustering structure) in order for the segment **X** to match the event model $\mathcal{M}_i$. The lower the loss, the more the signal segment matches the event model.

The proposed detection algorithm can then be presented as to two sub-tasks, the forward detection step which is the main process for finding an event, and the backward reduction step that is responsible for a more accurate segmentation.

In the forward detection step, new samples are received one at a time and inserted into the clustering space. Upon insertion of a new sample, the clustering indices are updated and the model loss is re-estimated. Once a match is encountered (i.e. the model loss is zero or less than a predefined threshold λ), a detection is declared with the current change point u of the matched segment and the change-interval $\mathbf{X}_t = [\boldsymbol{x}_u, \boldsymbol{x}_{u+1}, \dots, \boldsymbol{x}_{v-1}, \boldsymbol{x}_v]$ where $\boldsymbol{x}_v$ is the first sample of the second stationary segment.

Once an event is declared, the backward reduction step begins. In this step, samples are removed from the clustering space in a First-In-First-Out (FIFO) fashion while updating the clustering structure upon each deletion and re-estimating the model loss. The reduction ends by the last sample that satisfies the matching condition (i.e. if that sample is deleted, the segment will no longer matches the event model within the predefined threshold loss λ). The complete detection algorithm can be described as follows. Given an event model $\mathcal{M}_i$

1. Receive new sample $\boldsymbol{x}_{N+1}$ and append it to **X**
2. Update the clustering vector $\boldsymbol{y}$ and the clustering structure $\{C_j\}_{j=1}^{m}$
3. Check $\mathcal{L}(\mathcal{M}_\mathrm{i}, \mathbf{X}, u, v) \leq \lambda$ for all u, v, if not satisfied, go to step (1)
4. Declare event detection with change-interval $\mathbf{X}_t = [\boldsymbol{x}_u, \boldsymbol{x}_{u+1}, \dots, \boldsymbol{x}_{v-1}, \boldsymbol{x}_v]$ and change-point is u where u and v result in the minimum model loss between $\mathcal{M}_i$ and the current segment **X** (i.e. $\mathrm{argmin}_{u,v}\, \mathcal{L}(\mathcal{M}_\mathrm{i}, \mathbf{X}, u, v)$).
5. Delete oldest sample $\boldsymbol{x}_1$ from the segment
6. Update the clustering vector $\boldsymbol{y}$ and the clustering structure $\{C_j\}_{j=1}^{m}$
7. Check $\mathcal{L}(\mathcal{M}_\mathrm{i}, \mathbf{X}, u, v) \leq \lambda$ for all u, v, if satisfied, go to step (5)
8. Re-insert last sample and declare current segment **X** as a balanced event.

After each detection, the process restarts from the first sample of the second stationary segment $\boldsymbol{x}_v$. The main objective of the backward reduction step is to extract balanced stationary segments (i.e. $|C_1| \approx |C_2|$) around the transient interval. Balanced segments lead to more stable steady-state features as well as an enhanced robustness to missed detections (i.e. false negatives).

2. EXPERIMENTS AND RESULTS

The proposed event detection approach has been evaluated on different power datasets among them is the Building-Level fUlly labelled Electricity Disaggregation (BLUED) dataset [16]. In the following, we show the results of applying the event detection algorithm with event model $\mathcal{M}_3$ and the DBSCAN clustering scheme on the BLUED dataset. We only show evaluation of detection results. Evaluation of the accuracy of transient interval segmentation and the stability of extracted features is beyond the scope of this paper.

Table 1 shows the event detection results on the real and reactive power signals from the BLUED dataset. BLUED include aggregate measurements from a two-phase residential building (phase A and B) and each is evaluated separately. True Positives (TP) is the number of successful detections, False Positives (FP) is the number of detections that do not correspond to actual events, while False Negatives (FN) is the number of missed events. Finally, False Positive Percentage (FPP), precision, recall, and the F1-score measures are defined as

$$\text{FPP} = \frac{FP}{E} \tag{8}$$

$$\text{precision} = \frac{TP}{TP + FP} \tag{8}$$

$$\text{recall} = \frac{TP}{TP + FN} \tag{9}$$

$$F_1 - \text{score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{10}$$

where E is the number of events. Results show highly precise detection rates where the number of false positives is relatively low in both phases. It is also observed that, noise in the second phase (phase B) still masks a relatively large number of events.

Table 1. Event detection results on BLUED [16] dataset.

	Phase A	**Phase B**	**Total**
Number of events E	892	1609	2501
Number of detections	874	1176	2050
True Positives (TP)	867	1097	1964
False Positives (FP)	7	79	86
False Negatives (FN)	25	512	537
FPP	0.78%	4.91%	3.44%
precision	**99.20**%	**93.28**%	**95.81**%
recall (TPR)	**97.20**%	**68.18**%	**78.53**%
F_1-score	**98.19**%	**78.78**%	**86.31**%

3. CONCLUSIONS

We introduced a novel clustering-based approach for sequential event detection. The proposed algorithm features accurate segmentation of the stationary and non-stationary intervals for more stable feature extraction, support of arbitrarily shaped stationary segments with no prior assumptions on their sample distribution, and more robustness to noise as well as parameter variations.

REFERENCES

[1] G. W. Hart, "Nonintrusive appliance load monitoring", in proceedings of the IEEE: vol.80, no.12, pp. 1870-1891, Dec. 1992. doi:10.1109/5.192069

[2] A. I. Cole and A. Albicki, "Data extraction for effective non-intrusive identification of residential power loads", in proceedings of the Instrumentation and Measurement Technology Conference (IMTC) 1998 IEEE: vol.2, pp.812-815, May 1998. doi:10.1109/IMTC.1998.676838

[3] S. Drenker and A. Kader, "Nonintrusive monitoring of electric loads", in Computer Applications in Power, IEEE: vol.12, no.4, pp.47-51, Oct 1999. doi:10.1109/67.795138

[4] M. El Hachemi Benbouzid, "A review of induction motors signature analysis as a medium for faults detection", IEEE Transactions on Industrial Electronics, vo.47, no.5, pp.984-993, Oct 2000.

[5] S. N. Patel, T. Robertson, J. A. Kientz, M. S. Reynolds, and G. D. Abowd, "At the Flick of a Switch: Detecting and Classifying Unique Electrical Events on the Residential Power Line", in UbiComp 2007: Ubiquitous Computing, vol.4717, pp.271-288, 2007.

[6] M. Basseville and I. V. Nikiforov. Detection of Abrupt Changes: Theory and Application. Prentice Hall, 1993.

[7] M. Berges, E. Goldman, L. Soibelman, H. S. Matthews, and K. Anderson, "User-centred non-intrusive electricity load monitoring for residential buildings", Journal of Computing in Civil Engineering, vol.25, no.1, 2011.

[8] K. D. Anderson, M. E. Berges, A. Ocneanu, D. Benitez, and J. M. F. Moura, "Event Detection for Non-Intrusive Load Monitoring", in IECON 2012, 38th Annual Conference on IEEE Industrial Electronics Society, October 2012.

[9] K. N. Trung, E. Dekneuvel, B. Nicolle, and O. Zammit, "Event Detection and Disaggregation Algorithms for NIALM System", in the 2nd International Non-Intrusive Load Monitoring (NILM) Workshop, Jun 2014.

[10] Y. Jin, E. Tebekaemi, M. Berges, and L. Soibelman, "A time-frequency approach for event detection in nonintrusive load monitoring" in proceedings of the Signal Processing, Sensor Fusion, and Target Recognition, Orlando, Florida, USA, 2011.

[11] M. Volpi, D. Tuia, G. Camps-Valls, and M. Kanevski, "Unsupervised Change Detection With Kernels", in Geoscience and Remote Sensing Letters, IEEE: vol.9, no.6, pp.1026-1030, Nov. 2012. doi: 10.1109/LGRS.2012.2189092

[12] M. Luong, V. Perduca, and G. Nuel, "Hidden Markov Model Applications in Change-Point Analysis", arXiv Journal, preprint arXiv:1212.1778v1, 2012.

[13] G. L. Grinblat, L. C. Uzal and P. M. Granitto, “Abrupt change detection with one-class time-adaptive support vector machines”, Expert Systems with Applications Journal, vol. 40, pp. 7242–7249, 2013.

[14] S. B. Leeb, S. R. Shaw, and Jr. Kirtley, J. L. “Transient event detection in spectral envelope estimates for nonintrusive load monitoring”, IEEE Transactions on Power Delivery, 10(3):1200–1210, July 1995.

[15] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, “A density based algorithm for discovering clusters in large spatial databases with noise” in proceedings of Knowledge Discovery and Data mining (KDD), 1996.

[16] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, and M. Berges, “BLUED: a fully labeled public dataset for Event-Based Non-Intrusive load monitoring research”, in proceedings of the 2nd Knowledge Discovery and Data mining (KDD) Workshop on Data Mining Applications in Sustainability (SustKDD), Beijing, China, August 2012. URL: http://nilm.cmubi.org/