

Unsupervised Non-Intrusive Load Monitoring of Residential Appliances

Karim Said Barsim, Roman Streubel, and Bin Yang
Institute for Signal Processing and System Theory
University of Stuttgart

Email: karim__said@hotmail.com, {roman.streubel; bin.yang}@iss.uni-stuttgart.de

Abstract—Non-Intrusive Load Monitoring (NILM) is a technique in which electrical appliances are monitored through the disaggregation of the aggregate electrical signal gathered from a single metering point. NILM has recently attracted wide interest in research for application in both the commercial and residential sectors. In this paper, a completely unsupervised NILM system is proposed for application in the residential sector. The proposed system consists of four stages. First, events are detected in real-time using a bucketing technique and EM clustering. Features are then extracted from the transient behavior of appliances. Afterwards, mean-shift clustering is utilized in the detection of recurrences. Finally, an initial transition matching approach is proposed. The system is then tested on the publicly available BLUED dataset and showed high event detection and clustering results and reported a possibility of complete disaggregation up to 92% of the total consumed energy.

Index Terms—Non-Intrusive Load Monitoring, Mode seeking, Mean-shift clustering, Bucketing technique, Event detection, BLUED dataset

I. INTRODUCTION

Even though electricity has become an integral form of energy to the everyday life, there are still misconceptions about the usage of the electrical energy specially among consumers in the residential sector [1–7]. This raises a clear need for a detailed understanding about the usage of the electrical energy. Energy conservation, load-dependent energy generation, activity sensing, and home automation are some of the several applications that also require such a detailed understanding. For instance, researchers in [8] have found experimentally that consumers are able to reduce their daily consumption level by 10-15% within one month if they are provided with a frequent written feedback about their energy consumption. Activity sensing, home automation, health care, and similar applications rely on the statement by [7] that location and activity of occupants can be inferred via monitoring specific appliances in the building.

Electrical loads can be monitored either intrusively or non-intrusively. In the intrusive approach, electrical signals are gathered from each appliance on its own. The Non-Intrusive Load Monitoring (NILM), on the other hand, gathers the electrical signal of simultaneously operating appliances. The signal is then disaggregated to infer information about the individual appliances. NILM is, therefore, less costly due to the less hardware used and more reliable as a result of the reduction of metering points. NILM systems are developed

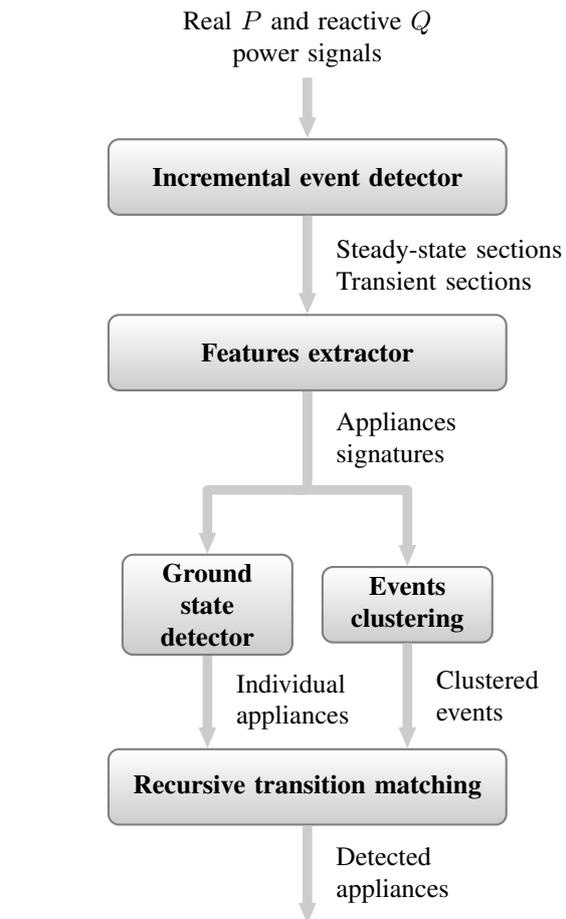


Fig. 1. A block diagram of the proposed unsupervised NILM system. All stages utilize only unsupervised techniques in order to disaggregate the individual appliances from the given real and reactive power signals.

specifically either for the residential sector or for the commercial one. The distinction between the two sectors stems from their difference in the consumers' behavior and the types and numbers of the underlying appliances. A good review of the available NILM systems can be found in [6, 9, 10].

We noticed that most existing NILM systems require a pre-training process in at least one of its monitoring stages. This

imposes a complex installation process and constraints on adding new appliances to the building. Unsupervised NILM systems are, therefore, expected to have a wider applicability and even less intrusion.

In this paper, we describe our work in developing a completely unsupervised NILM system targeted for the residential sector. In developing such a system only unsupervised detection and recognition tools are utilized.

Figure 1 shows a general view of the proposed NILM system. Each stage of the system is described in a separate section in this paper. Therefore, this paper is organized as follows. Section II describes the event detection stage in which the raw power signals are segmented into steady-states and transient sections. In Section III features extracted from the detected events are introduced. Features are then utilized in clustering events for the detection of recurrences which is described in Section IV. In Section V the detection of the ground-state is described together with its advantages. Section VI proposes an initial transition matching stage. The system was tested on one of the publicly available power datasets which is introduced in Section VII together with the test results. Finally, Section VIII concludes this paper.

II. EVENT DETECTION

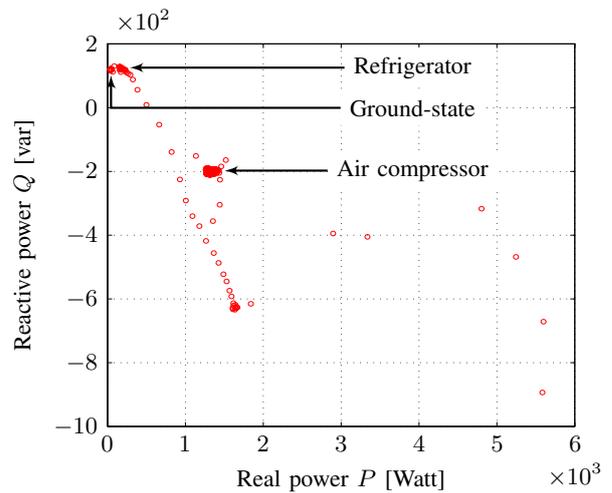
In the event detection stage, the signal is segmented into steady-state and transient sections. Since features are going to be extracted from the transient sections, two constraints on the event detection algorithm have to be satisfied. First, it has to be a change-interval detector. In contrast to common change-point detectors, the required event detector must accurately define the time limits of the whole transition interval. Second, it has to preserve all characteristics in the transient sections. Therefore, no filtering technique is used during the whole detection stage.

A good review about the event detection approaches can be found in [11] which also classified detection algorithms into three categories: (a) expert heuristics, (b) probabilistic models, and (c) matched filters. The proposed detection algorithm is a combination of the first two categories. This is because the event detector consists of two steps: the coarse event detection followed by a refining step on each detected event. To enhance the performance, the detection is performed on the logarithmically transformed signals rather than the raw ones.

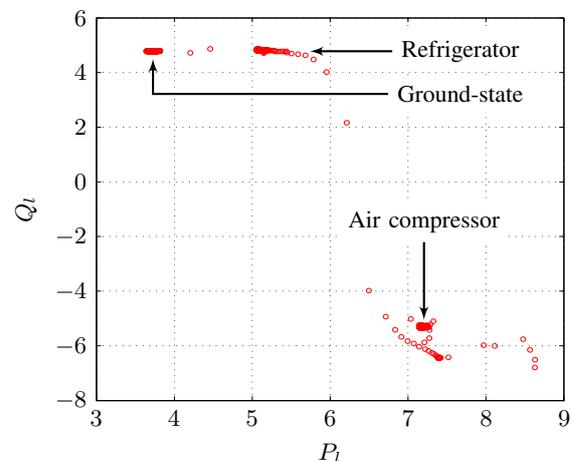
A. Logarithmic transform

We observed that very large spikes specially during the turn-on transient of some appliances significantly degrades the performance of the event detection algorithm. This is because such spikes produce wide empty areas in the \overline{PQ} -Plane.

Figure 2(a) shows a sample of two appliances, namely a refrigerator and an air compressor, which is sampled from the BLUED dataset [12] and plotted on the \overline{PQ} -Plane. As shown, transition spikes reach up to 6 kW in the real power P which is much larger than the maximum steady consumption level (~ 1.1 kW). This results in a very large \overline{PQ} -Plane. Moreover, steady-states, which are observed as clusters in the \overline{PQ} -Plane, are hardly separable. In order to overcome these



(a) \overline{PQ} -Plane



(b) $\overline{P_l Q_l}$ -Plane

Fig. 2. The logarithmic transform when applied on the P and Q signal. The figures show (a) the raw real and reactive power signals of two appliances plotted on the \overline{PQ} -Plane (b) the logarithmically transformed signals plotted on the $\overline{P_l Q_l}$ -Plane where empty areas are suppressed and clusters become more separable.

limitations, the raw real power P signal is transformed to the logarithmic real power P_l defined as

$$P_l = \begin{cases} \log(P) & : P > 0 \\ 0 & : P = 0 \\ -\log(-P) & : P < 0 \end{cases} \quad (1)$$

prior to applying the event detection algorithms. The same transform is applied on the reactive power signal Q in order to obtain the logarithmically transformed reactive power signal Q_l .

Figure 2(b) shows the same signal in (a) after applying the previously-mentioned transform when plotted on $\overline{P_l Q_l}$ -Plane. As can be observed, the ranges of the axes of both the real P and reactive Q powers are significantly reduced, empty areas are suppressed rather than trimmed, and clusters become more

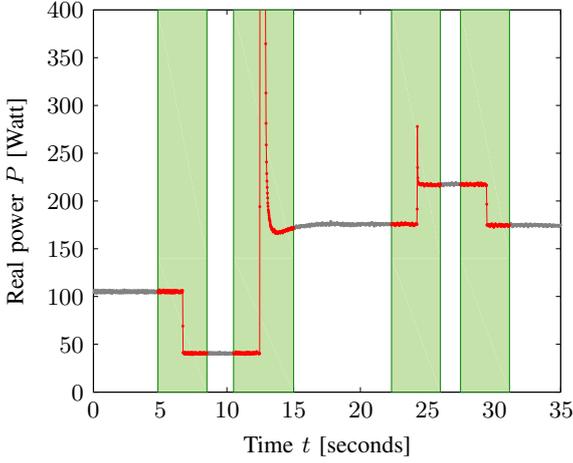


Fig. 3. A 35-seconds real power P signal sampled from the BLUED dataset. Shaded areas define the limits of the detected coarse transitions.

separable. The modified signals P_l and Q_l are then fed to the coarse transition detection.

B. Coarse event detection

During this step, events are detected one at a time together with few samples from the surrounding steady-states. Segments from the surrounding steady-states are equal to the minimum detectable steady-state length N_{th} more or less depending on the noise. The minimum detectable steady-state length in the case of the $f_s = 1$ Hz signal was selected to be $N_{th} = 3$ seconds whereas in the $f_s = 60$ Hz signal it was $N_{th} = 0.67$ seconds. Since events are detected one at a time, they are detected in real-time. This results in a real-time event detector even though the whole proposed NILM system does not preserve this characteristic.

The algorithm of the event detection algorithm is as follows

- Step 1: receive $n < N_{th}$ data samples.
- Step 2: estimate the model order \hat{M} .
- Step 3: if $\hat{M} = 1$, return to step 1.
- Step 4: mark last received sample t_2 .
- Step 5: read $n < N_{th}$ data samples backwards from t_2 .
- Step 6: estimate the model order \hat{M} .
- Step 7: if $\hat{M} = 1$, return to step 5.
- Step 8: mark last read point t_1 .

The output from this step is the interval $[t_1, t_2]$ and is referred to as a coarse event. The utilized model order estimator is the one that has been developed in a previous work in [13]. The model order estimator in [13] is based on a bucketing technique from image processing.

Figure 3 shows the coarse events of a sample signal from the BLUED dataset with sampling frequency of $f_s = 60$ Hz. The output is highlighted in the shaded windows which define the time limits of each coarse event. Each window contains only one transient Ψ_i and two steady-states Π_i and Π_{i+1} . Therefore, if any of the coarse events is individually plotted on the \overline{PQ} -Plane, a model order estimation should always

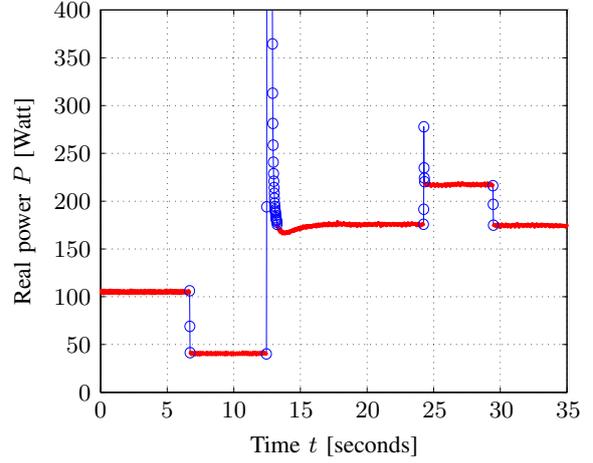


Fig. 4. The signal in Figure 3 after the application of the refining step. Highlighted with circles are the transients sections. As observed, transients are accurately defined while preserving their characteristics.

result in a model order of $\hat{M} = 2$. Each coarse event is then fed to the refining step once it is detected.

C. Fine event detection

In fine event detection, each received coarse event is fed to a more refining detection to accurately define the time limits of each transient section. In the previous work of [13] an unsupervised event detection technique was developed using a model order estimator followed by an Expectation Maximization (EM) clustering and Gaussian Mixture Models (GMM). It showed promising results but it had the limitation that it was developed under the assumption that only one appliance is operating at a time. Therefore, direct application on simultaneously-operating appliances was not possible. Rather than applying the detection of [13] directly on the whole signal, it is applied only on each detected coarse event for a more accurate segmentation. A coarse event is even simpler than a solely-operating appliance.

Figure 4 shows the sample signal previously coarsely-segmented in Figure 3 after applying the clustering algorithm of [13]. As a result, the signal is accurately segmented into steady and transient sections while preserving all characteristic spikes for accurate feature extraction.

III. FEATURE EXTRACTION

In this stage, features are extracted from the segmented signal. Selected features are the power change $\Delta\Psi$ and the transition spike $\delta\Psi$. Each feature is computed on both the real P and reactive Q power signals as follows

$$\Delta\Psi_i^X = \Psi_i^X(N_i^\Psi - 1) - \Psi_i^X(0) \quad (2)$$

and

$$\delta\Psi_i^X = \text{sign}(\Delta\Psi_i^X) \times \left(\max_n \Psi_i^X(n) - \min_n \Psi_i^X(n) \right) \quad (3)$$

where $X \in \{P, Q\}$ and N_i^Ψ is the number of data samples in the transient section Ψ_i . This results in a four dimensional

feature vector. The set of all vectors are then fed to the clustering stage.

IV. EVENTS CLUSTERING

In the clustering stage, events are grouped in separate clusters according to their extracted features. Since the number of the underlying appliances is not known in advance, the clustering algorithm has to be non-parametric. This limits the usage of all clustering schemes that are based on function optimization since they assume that the model order is known or require a model order estimation.

Two clustering schemes are recommended for this stage: the Basic Sequential Algorithmic Scheme (BSAS) and the mean-shift clustering algorithm. The BSAS clustering algorithm has the advantage that it is suitable for a real-time clustering. It is, on the other hand, dependent on the order in which data are presented to the algorithm. This is currently considered as a limitation. However, it will be taken into consideration in a future work where the joint modeling of the electrical loads and user intention is utilized as introduced in a previous work [14].

The utilized clustering algorithm is the mean-shift clustering. The mean-shift clustering has the advantage that it is completely non-parametric, independent of the underlying distribution and implicitly includes a mode-seeking algorithm. It is, on the other hand, more computationally expensive than the k-means algorithm. The mean-shift algorithm was first introduced by [15]. However, its usage have flourished only recently when it was widely used in computer vision and image processing [16–21]. It has also been implemented on several processing units such as Field Programmable Gate Arrays (FPGAs) [22] and Graphical Processing Unites (GPUs) [23]. Recently, it has been proposed for application in NILM systems and was proved to provide even better results than the k-means algorithm in special cases [24].

The mean-shift clustering scheme is briefly described in the following. Given a kernel function such as

$$K(\theta) = \begin{cases} 1, & \text{if } \|\theta\| \leq \lambda \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where λ is the kernel bandwidth and given a random unclustered starting point \mathbf{x}_0 , the iterative calculation of the mean μ follows

$$\mu_{q+1} = \frac{\sum_{\mathbf{x}_i \in S} K(\mathbf{x}_i - \mu_q) \mathbf{x}_i}{\sum_{\mathbf{x}_i \in S} K(\mathbf{x}_i - \mu_q)} \quad (5)$$

where

$$\mu_0 = \mathbf{x}_0 \quad (6)$$

where S is the set of all sample points. The value $\mu_{q+1} - \mu_q$ is called the mean-shift. The algorithm converges when

$$\|\mu_{q+1} - \mu_q\| < \epsilon \quad (7)$$

The algorithm is repeated until all points are clustered. All points that converge to the same mean μ are merged into one cluster.

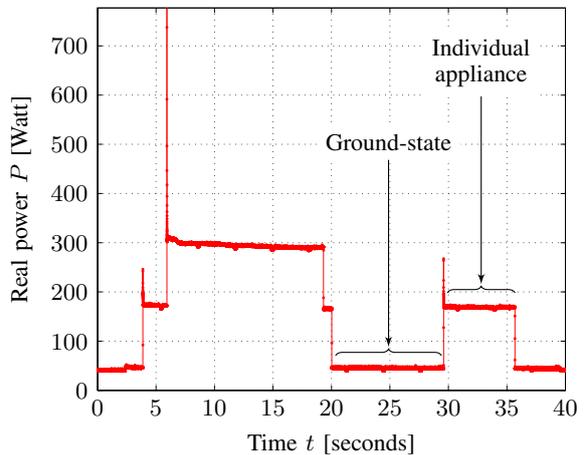


Fig. 5. A pair of on- and off-events surrounded by ground-states belongs to the same appliance.

V. GROUND-STATE DETECTION

The ground-state is the state during which no detectable appliance is operating. We observed that in the residential data there are several times when the occupants have limited activities. Detection of this stage can take the following schemes. In real-time processing, the ground-state could be defined as the first steady-state of the system. This imposes a constraint on the installation process that no appliance should operate until a predefined period after the meter has been connected to the circuit to be monitored. In the implemented NILM system, the ground-state is detected as the steady-state with the lowest power consumption level in a signal with the duration of at least one day.

Figure 5 shows a sample signal from the BLUED dataset where steady-states are highlighted. One of the advantages for detecting these ground-states is reducing the effect of faults occurring in the event detection stage. Another advantage is that it provides an opportunity for transition matching and self-learning as explained in the next section.

VI. TRANSITION MATCHING

In the transition matching stage, on- and off-events belonging to the same appliance are grouped together so that the whole operation interval of that appliance can be inferred. In this work, we only propose an initial transition matching stage that can be used to reduce the search space for further matching processes but does not guarantee high disaggregation ratios except in special cases. The matching process adopted in the proposed system is based on the ground-state detection. According the definition of the ground-state, if a pair of on- and off- events are surrounded by ground states (i.e. a ground-state before the on-event and another after the off-event) then these two events must belong to one appliance and the interval in between also belongs only to that appliance.

Figure 5 shows an example of a solely-operating appliance detected in between two ground-states. The on- and off-events are, therefore, matched and the interval in between can be utilized in self-training. Once these two events are matched,

the matching is triggered again based on the matching that occurred in the previous iteration. This is done recursively until no more matching is possible. In the Figure 5, the second iteration will match the remaining two events from the left-side operation interval.

VII. EXPERIMENTS AND RESULTS

The first stage of any machine learning application is always data acquisition, and NILM is not an exception. However, data acquisition in NILM systems is costly as it has to be gathered intrusively, time-consuming since it should be over long time periods, and faces several challenges [12]. Realizing this fact, researchers on NILM and similar energy applications started publishing their developed power datasets.

Four publicly available power datasets have been reviewed, namely the Tracebase dataset [25], the UMASS Smart* dataset [26], the Reference Energy Disaggregation Dataset (REDD) [27], and the Building-Level fully-labeled Electricity Disaggregation (BLUED) dataset [12]. Being the most suitable dataset for even-based NILM systems, the BLUED dataset has been chosen for the test of the proposed NILM systems.

The BLUED dataset is a one-week long measurements of a single family residence. Current and voltage signals are sampled at a frequency of $f_s = 12$ kHz. Post-computed real and reactive power signals at $f_s = 60$ Hz are accompanied with the dataset. Appliances in the residence are fed with two phases of the power source, phase A and phase B. Phase A consists mainly of on/off and FSM appliances whereas phase B contains, additionally, several continuously-variable appliances such as TVs, monitors, laptops, etc.

Table I provides detection results of the event detection stage on the both phases of the dataset. The only input to the event detector is the real power signal P with the sampling frequency $f_s = 60$ Hz. The True Positive Percentage (TPP) and the False Positive Percentage (FPP) are the second metric proposed for event detection in [11]. Complete detection of each phase was obtained in about 18 minutes of processing time and without a pre-training process. The minimum detectable steady-state length is $N_{th} = 0.67$ seconds.

TABLE I
EVENT DETECTION RESULTS

	TPP	FPP
Events	886	1579
Phase A	98.5%	0.55%
Phase B	70.5%	8.75%

Table II provides results of the event clustering stage. A simple test was used to evaluate the results of the clustering stage. For each possible combination of two events from the successfully detected events ($x_i, x_j \forall i \neq j$), the clustering is queried whether they belong to the same cluster or not. True Positive Rate (TPR) and False Positive Rate (FPR) are calculated accordingly. This is an unbalanced test as it gives

TABLE II
EVENT CLUSTERING RESULTS

	TPR	FPR
Phase A	99.3%	0.6%

priority to larger clusters. The clustering stage is only tested on phase A of the dataset.

Finally, the transition matching stage is able to match about 92% (i.e. 8.3 kW·H) of the total consumed energy in phase A (i.e. 8.9 kW·H). However, not all of the matching have been verified due to limitations of the BLUED dataset. The relatively high disaggregation ratio even with the initial transition matching approach is due to the refrigerator. The refrigerator consumes about 70% of the total energy and because it is always on, it was found solely-operating several times. Being the highest energy-consuming appliance, the refrigerator is one of the most important appliances for energy conservation applications while being a background appliance makes it the least important one for activity sensing applications.

VIII. CONCLUSION

We proposed an approach for a completely unsupervised NILM system for a simpler installation process, wider applicability in the residential sector, and less intrusion in the monitoring process. The proposed NILM system follows the stages of a typical NILM system. However, only unsupervised detection and recognition tools were utilized. The system has been tested on the publicly available BLUED dataset and showed high event detection and clustering results. Complete disaggregation is possible with the initial transition matching stage up to 92% of the total consumed energy.

REFERENCES

- [1] W. Kempton, C. K. Harris, J. G. Keith, and J. S. Weihl, "Do consumers know "what works" in energy conservation?" in *Marriage and Family Review*, vol. 9, 1985.
- [2] D. Parker, D. Hoak, A. Meier, and R. Brown Venue, "How Much Energy Are We Using? Potential of Residential Energy Demand Feedback Devices," in *Proceedings of the 2006 Summer Study on Energy Efficiency in Buildings*, American Council, 2006.
- [3] W. Kempton and L. Montgomery, "Folk Quantification of Energy," *Energy*, vol. 7, October 1982.
- [4] S. Darby, "The Effectiveness of Feedback on Energy Consumption," Environmental Change Institute, University of Oxford, Oxford, UK, Tech. Rep., April 2006.
- [5] C. Fischer, "Feedback on household electricity consumption: a tool for saving energy?" *Energy Efficiency*, Feb. 2008.
- [6] J. Froehlich, E. Larson, S. Gupta, G. Cohn, M. Reynolds, and S. Patel, "Disaggregated End-Use Energy Sensing for the Smart Grid," *Pervasive Computing, IEEE*, 2011.
- [7] 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 *Proceedings of the 9th international conference on Ubiquitous computing*, Berlin, Heidelberg, 2007.
- [8] R. A. Winett, M. S. Neale, and H. Cannon Grier, “Effects of self-monitoring and feedback on residential electricity consumption,” *J. Applied Behavior Analysis*, 1979.
- [9] H. Najmeddine, K. El Khamlichi Drissi, C. Pasquier, C. Faure, K. Kerroum, A. Diop, T. Jouannet, and M. Michou, “State of art on load monitoring methods,” in *2nd IEEE International Conference on Power and Energy (PECon 08)*, Johor Baharu, Malaysia, 2008.
- [10] M. Zeifman and K. Roth, “Nonintrusive appliance load monitoring: Review and outlook,” *Consumer Electronics, IEEE Transactions on*, 2011.
- [11] K. Anderson, M. Berges, A. Ocneanu, D. Benitez, and J. Moura, “Event detection for Non-Intrusive load monitoring,” in *IECON 2012 - 38th Annual Conference on IEEE Industrial Electronics Society*, October 2012.
- [12] 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 KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, Beijing, China, Aug. 2012.
- [13] R. Streubel and B. Yang, “Identification of electrical appliances via analysis of power consumption,” in *Universities Power Engineering Conference (UPEC), 2012 47th International*, sept. 2012.
- [14] B. Yang and Z. Zhou, “Joint Modeling of Device Load and User Intention,” in *Universities’ Power Engineering Conference (UPEC), Proceedings of 2011 46th International*, 2011.
- [15] K. Fukunaga and L. Hostetler, “The estimation of the gradient of a density function, with applications in pattern recognition,” *Information Theory, IEEE Transactions on*, 1975.
- [16] Z.-Q. Wen and Z. xing Cai, “Mean Shift Algorithm and its Application in Tracking of Objects,” in *Machine Learning and Cybernetics, 2006 International Conference on*, 2006.
- [17] J. G. Choe, J.-H. Seok, and J.-J. Lee, “Mean-shift tracker with face-adjusted model,” in *Control, Automation and Systems, 2008. ICCAS 2008. International Conference on*, 2008.
- [18] X. Chen, S. Yu, and Z. Ma, “An improved mean shift algorithm for moving object tracking,” in *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*, 2008.
- [19] W. Changjun and Z. Li, “Mean shift based orientation and location tracking of targets,” in *Natural Computation (ICNC), 2010 Sixth International Conference on*, 2010.
- [20] M. Deilamani and R. Asli, “Moving object tracking based on mean shift algorithm and features fusion,” in *Artificial Intelligence and Signal Processing (AISP), 2011 International Symposium on*, 2011.
- [21] T. Liu and X. ping Cheng, “Improved mean shift algorithm for moving object tracking,” in *Computer Engineering and Technology (ICCET), 2010 2nd International Conference on*, 2010.
- [22] D. B. K. Trieu and T. Maruyama, “An Implementation of the Mean Shift Filter on FPGA,” in *Field Programmable Logic and Applications (FPL), 2011 International Conference on*, 2011.
- [23] J. Zhang, S. Luo, and X. Liu, “Weighted Mean Shift Object Tracking Implemented on GPU for Embedded Systems,” in *Control Engineering and Communication Technology (ICCECT), 2012 International Conference on*, 2012.
- [24] W. Zhenyu and Z. Guilin, “The application of mean-shift cluster in residential appliance identification,” in *Control Conference (CCC), 2011 30th Chinese*, 2011.
- [25] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, and R. Steinmetz, “On the accuracy of appliance identification based on distributed load metering data,” in *Sustainable Internet and ICT for Sustainability (SustainIT), 2012, 2012*.
- [26] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht, “Smart*: An Open Data Set and Tools for Enabling Research in Sustainable Homes,” Beijing, China, Aug. 2012.
- [27] J. Kolter and M. Johnson, “REDD: A Public Data Set for Energy Disaggregation Research,” 2011.