Simple Event Detection and Disaggregation Approach for Residential Energy Estimation

Awet Abraha Girmay[1,2]
[1]Department of Electronics
and Telecommunication Engineering
Polytechnic University of Turin, Italy
Email: awetabraha.girmay@studenti.polito.it

Christian Camarda[1,2]
CEO and Responsible R&D of Midori Srl
[2]I3P, the Innovative Enterprise Incubator of the
Polytechnic University of Turin, Italy
Email: christian@midorisrl.eu

***Abstract*—Non-Intrusive Appliances Load Monitoring systems are crucial for augmenting virtuous energy saving behavior and providing residential energy monitoring solutions. NILM is aimed to accurately account for energy costs and load distribution by reporting where exactly the energy is going and what kind of devices are using it. To address this issue, much of existing methods require a lot of time consuming training, complex optimization algorithms and does not focus on the actual energy estimation problem rather concentrate on classification problem. In this paper, we propose a simple active window based NILM (AWB-NILM) approach that relies on an unsupervised localized events clustering, pairing and self-learning using automatic evolutionary clustering methods. We have tested our approach on a real residential power consumption data. The F-measure attained by the event detector is 94.6%. In this paper we showed accuracies of appliance events correctly classified and energy correctly assigned.**

***Index Terms*—Non-intrusive appliance load monitoring, Active window, Geometric features, Genetic algorithm, Automatic clustering.**

I. INTRODUCTION

In recent times, utility companies have shown keen interest in NILM and are making efforts for the implementation and deployment of smart meter and smart grid technologies so as to harness their many fold advantages as these technologies play a vital role in achieving advanced energy billing systems and increase the energy usage feedback to both utilities and customers. NIALM, a.k.a NILM was pioneered by G. Hart [1], in which an aggregate usage of active and reactive power is used to identify when appliances are turned on/off and estimate their energy usage. Identification as well as monitoring of each appliance is made possible by using unique signatures from each type of devices. Hence, employing such systems would enable utilities to have effective demand and supply management and customers to avoid wasteful practices.

A nice review of the different NIALM approaches can be found in [2] and [3]. Besides, several NIALM implementation approaches exploit various appliance features, [8] – [10], for instance the active and reactive power, the harmonic distortion, the transient behavior, and even the voltage distortion in order to obtain unique and typical characteristics of appliances but they require a high sampling rate in the order of kHz.

In our ongoing research, we propose an event-based NILM technique using simple events detection, feature extraction, and AWB-clustering suitable for energy disaggregation of the most energy-consuming devices based on active and reactive power measurements taken at 1 Hz at the mains entry using our own Ned-meter, a smart meter prototype built by us. We also provide evaluation results of the event detector as well as disaggregated energy estimation.

The organization of the paper is as follows. Section II gives a brief review of background. Section III provides the descriptions of the event detection, feature extraction, AWB-clustering, transition matching and energy estimation of appliances. In Section IV the experimental results along with the bench mark quality metrics are described. Lastly, in Section V conclusive remarks on the current work as well as future research path is presented.

II. BACKGROUND

Almost two decades back Hart [1] proposed a method for dis-aggregating electrical loads by examining only the appliance specic step-like changes present in the aggregate power consumption. Moreover, several event detection techniques have been proposed [4], [5]– [6]and steady-state and transient event-based feature mining approaches are developed so as to typify the detected events.

Similarly, researchers [6], [12] have used real power as a solo feature to dis-aggregate large-power consuming appliances and seem to ignore low power devices. In [9] disaggregation using simple power model and the sparsity property of appliance activities is proposed nevertheless the types and the number of appliances in a house are known a prior.

III. EVENT DETECTION, FEATURE EXTRACTION AND MATCHING

A typical event-based NIALM system incorporates event detection, feature extraction, classification (appliance labeling), and disaggregated energy estimation. Descriptions of the algorithms are discussed in the following subsections.

A. Event Detection

In the event detection stage, the electrical signal is split-up into contiguous and transient data points. The proposed event detector has low complexity and is able to suppress noisy data points by specifying the minimum time duration of the contiguous power levels and applying a smoothing filter.

1) Contiguous Level Detector: In this contiguous level detector, aggregate power signal is preprocessed using median filtering of length 10 samples and is valid for the 1Hz sampling rate. This is a good compromise to keep relevant spin cycles of a washing machine and remove noisy spikes from TVs because measurement readings are affected by noise so as to get smoother signal and produce accurate abrupt edges. It then computes the difference δX, i.e. $\Delta p/\Delta t$, of the aggregate consumption in order to obtain the contiguous and transient levels of the power signal. So given a time series of both aggregate real and reactive power consumption $X(1:T)$ of length T,

$$\delta X(t) = \Big\{ X(t+1) - X(t), \quad \mathit{for}\ t = 1:T-1. \tag{1}$$

It then further labels the data points in $\delta X(t)$ as contiguous and transient based on a specified threshold which in this case is 30 Watts. The binary vector $\mathbf{a(t)}$ defines the labeling as:

$$\mathbf{a(t)} = \begin{cases} 1, & \text{if } -Thr \leq \delta X(t) \leq Thr. \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

In equation (2), 1 indicating the portion of the time series data is contiguous representing steady state portions and 0 corresponds to transient portions due to rising and falling edges.

Finally, the contiguous level detector obtains ω, the starting and ending time of each steady state power levels, using the algorithm in 1. In here, NumZeros represents the minimum span length of a power level which in our case is 5 seconds and Isabove is a flag indicating whether we want to get start and end of steady states or transient segments.

Algorithm 1 Get start and end times of contiguous power levels

Data: $\mathbf{a(t)}$, NumZeros, Isabove
Result: ω
belowThr = [0 $\mathbf{a(t)}$ 0]
edge = diff(belowThr)
rise = find(edge == 1)
fall = find(edge == -1)
spanWidth = fall - rise
if *Isabove* **then**
| wideEnough = spanWidth $\geq$ NumZeros
else
| wideEnough = spanWidth $\leq$ NumZeros
end
start = rise(wideEnough)
end = fall(wideEnough)-1
ω = [start;end]'

2) Background and Active Segment Labeling: In this stage, sections of the filtered P and Q components of the $\omega_i{}'s$ obtained in section III-A1 are labeled as either background or active segments. To achieve this, first the normal distribution of each contiguous power level is computed in order to get the μ_i and σ_i of each steady state segments. This results in a six dimensional feature vector of $[t_{starti}, t_{endi}, \mu_i^P, \mu_i^Q, \sigma_i^P, \sigma_i^Q]$ where t_{starti} and t_{starti} represent the star and end time index of the i^{th} power level.

A background-level is a a power level during which either standby or permanently on devices are operating and there is no detectable abrupt changes. Since the type of appliances being installed in households is not known a priori and not deterministic due to the fact that end users can change appliance or connect completely new device. So estimating the background power as the sum of standby powers of surveyed appliances may not be accurate. Hence, should be determined from the aggregate electrical signal. To better estimate a more realistic background power, we propose to cluster the contiguous power levels, $\mu_i^{P}{}'s$, using automatic and fast algorithms.

In our approach we utilized automatic clustering namely differential evolution(DE) algorithm originally proposed by [13] which is a very simple, yet very powerful and useful algorithm. The main advantages in DE are: it does not require number of clusters a priori, make few or no assumptions about the problem being optimized, does not require the optimization problem to be differentiable, and has self-adaptive mutation scheme. Applying DE to the $\mu_i^{P}{}'s$, the cluster members, c_k, of each $\mu_i^{P}{}'s$ is obtained. Finally, the labeling, ℓ, of the steady state power segments is achieved based on the cluster member, $c_{k_{\mu_{min}}}$, of the minimum μ_i as follows.

$$\ell_{\mathbf{i}} = \begin{cases} 1, & \text{if } c_{k_{\mu_i}} = c_{k_{\mu_{min}}}. \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Hence, the background power is obtained as,

$$P_{bgrnd} = mean(\mu_i \mid c_{k_{\mu_i}} = c_{k_{\mu_{min}}}) \tag{4}$$

3) Active Window Detection: The power consumption of a household varies over time depending on the activation of individual devices used by the occupants. In here, active windows are defined as the time periods in the aggregate signal over which the power stays well above the reference background power that is observed as a result of activation of single or multiple appliances. Once labeling of the steady state segments is done as in equation 3, the starting and ending time of active window periods are obtained by looking for portion of the aggregate signal surrounded by two background levels. Hence, the set of active windows, aw, are defined as:

$$aw = \{aw_1, aw_2, ..., aw_k \mid aw_k = [t^k_{start}, t^k_{end}]\} \tag{5}$$

where k corresponds to the number of active windows detected, t^k_{start} and t^k_{end} correspond to the starting and ending time of the k^{th} active window respectively.

For instance Figure 1 depicts the aggregate power consumption of a pilot residential house taken by our smart meter over a course of one day. As can be seen from the plot, the blue contiguous segments represent active windows of single appliance operation surrounded by two background power levels and the red segments correspond to active windows in which the observed power curve is the result of either the operation of multi-state or multiple appliance in parallel.

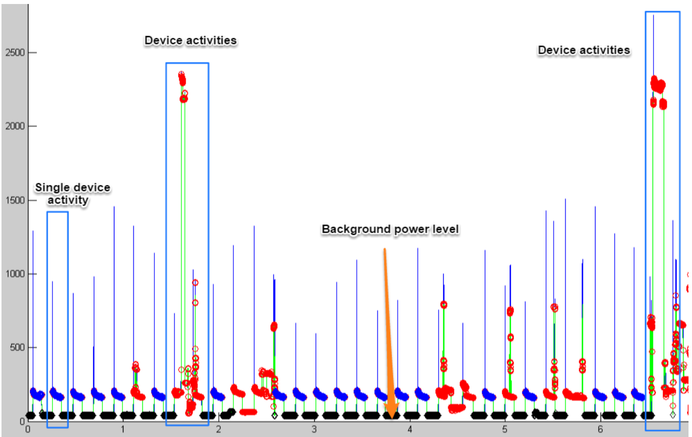

Fig. 1. This figure shows the detected activity windows(single or multiple appliances) and background level of portion of an aggregate power consumption of a household.

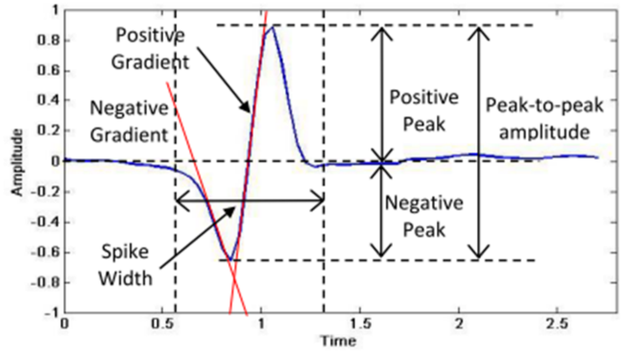

Fig. 2. Spike geometric features. spike features which characterize a spike shape include positive and negative gradients, spike width, positive and negative peaks, and peak-to-peak amplitude. [7]

The proposed approach is capable of labeling and obtaining activity windows due to single device operation automatically and lends itself to a simplified localized clustering and events pairing approach. In here, events are defined as the transition between two consecutive contiguous data segments.

B. Feature Extraction

In this section, features are extracted from each transient portions of the original unfiltered signal. Relevant features are the power change $\Delta(P, Q)$ and the geometric features, $\delta\varphi$, of the transient spikes on both real and reactive power.

1) Geometric Features: Geometric features are set of features that characterize the shape of a particular spike such as: positive and negative peak power amplitude, spike width, spike gradient, peak to peak power amplitude, and the peak amplitude to $\Delta(P, Q)$ ratios as shown in Fig. 2 and is extracted from the slope signal in Equation 1 around each event.

$$\delta\varphi = \{\delta X^{+}_{peak}, \delta X^{-}_{peak}, \frac{\delta X^{+}_{peak}}{\delta X^{-}_{peak}}, \frac{\delta X^{+}_{peak}}{\Delta P}, \frac{\delta X^{+}_{peak}}{\Delta Q}, \Delta\tau_{spike}\} \tag{6}$$

TABLE I
EXTRACTED VALUES OF THE DIFFERENT TURN-ON FEATURE PARAMETERS.

Event	t_{up}	$\delta\varphi(1)$	$\delta\varphi(2)$	$\delta\varphi^P(3)$	$\delta\varphi^Q(3)$	$\delta\varphi(4)$	$\delta\varphi(5)$	$\delta\varphi(6)$
1	3022	1097.0	-1002.9	1.09	1.11	13.64	10.70	6
2	8133	1092.2	-995.1	1.10	1.11	12.70	10.82	6
3	13267	1009.5	-920.5	1.10	1.12	11.72	9.48	6
4	18377	760.7	-676.2	1.12	1.17	9.53	6.92	6
5	23503	1035.9	-939.4	1.10	1.11	11.98	10.18	5

C. AWB-Clustering, Matching and Self Learning

In this stage, an active window based clustering is proposed in which localized events clustering within the detected active windows, $aw'_k s$, is performed. This approach is capable of avoiding mixing of events from different appliances into the same cluster which are temporally far apart, increases the accuracy of events pairing because of the narrower time window over which events matching is done, and doesn't assume what appliances exist within the households. Our proposed approach focuses on an active window based localized clustering of events residing in each active window comprised of either single or overlapped operation of on-off and multi-state appliances, even though the event clustering stage of the whole NILM system is based on batch-processing.

1) Events Clustering: Since the number of clusters is not know in advance, we utilize an automatic clustering as in [14], namely the genetic algorithm (GA) in which clustering itself is treated as an optimization problem. These algorithms were proven to have better performance over some of the classic clustering approaches and does not require number of clusters a priori.

2) Matching and Self-Learning: In the matching process, all rising and falling events are checked for matching pairs so as to infer the usage interval of each appliance. The matching process in the proposed system is based on the background-level detection and active window approach. Hence, this processing stage is performed on each identified active window times. It supports self-learning in operation to create a pool of possible appliance power states and build updateable appliance database. Since the characteristics of upward and downward transition events of appliances could slightly different due to voltage fluctuation or noise, threshold based Euclidean distances of the extracted features is performed to check if the new event clusters from the current active window also exists in the previously processed active windows. This method also utilizes simple transient related features and generic appliances usage patterns for labeling the category of appliances observed over a single or multiple active windows depending on the operation duration of appliances. In this stage, some clusters of matched pairs which does not belong to the target appliances are labeled as unknown high and low power appliances.

Once the operational activities of the target appliances are

TABLE II
QUALITY MEASURES FOR EVENT DETECTION

Appliance	N_e	N_d	N_{TP}	TPR	FPR	FNR	F-score
Fridge	1856	1818	1801	0.97	0.0092	0.0296	0.980
Washer	3800	3684	3500	0.92	0.0484	0.0789	0.935
Dishwasher	84	90	84	1	0.071	0.000	0.966
Microwave	69	72	66	0.956	0.087	0.0435	0.936
Stove	429	438	422	0.983	0.037	0.0163	0.973

detected and identied by using the proposed method, their energy estimation can be simply computed by taking the On-Off time and the steady-state power of each operation using Equation 7.

$$E^j = \sum_{z=1}^{Z} P_{ss}^j * (t_{off}^z - t_{on}^z) \quad (7)$$

where P_{ss}^j is the average steady state power, j refers to appliance number, z indicates the number of identied paired events and t_{off} and t_{on} are the on/off times of each operation cycle.

IV. EXPERIMENTAL RESULTS

Experimental analysis is performed on real households which consists of active and reactive power as well as voltage values. The measurement was carried out for one month at 1Hz from 10 pilot households with sub-metered ground truth data on some target appliance of our interest. Some of the measurements were taken using our own Ned-meter, a new and low cost smart meter prototype built by us, and others by a well known utility supplier in Italy.

To evaluate the performance, we utilize the test bench and quality measures for NILM as presented in [11]. Table II shows the performance metrics of the proposed event detection algorithm of two weeks data from one house. The true positive rate, false positive rate, the false negative rate, and the F-score are all computed as shown in Equation 8 and 9. The overall detection accuracy for all the measurements is more than 94%.

$$TPR = \frac{N_{TP}}{N_e}, \quad FPR = \frac{N_{FP}}{N_e}, \quad FNR = \frac{N_{FN}}{N_e} \quad (8)$$

$$F-score = 2 * \frac{Prec * Recall}{Prec + Recall}, \quad where$$
$$Prec = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad Recall = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (9)$$

Further more, Equations 10, 11 and 12 represent appliance event classification rate, appliance cycles classification rate and assigned energy rate respectively.

$$AECR = \frac{\sum \widehat{n}_{j,e}}{\sum n_{j,e}} = 0.948 \quad (10)$$

$$ACCR = \frac{\sum \widehat{n}_{j,a}}{\sum n_{j,a}} = 0.925 \quad (11)$$

$$AER = \frac{\sum \vec{\widehat{w}}_a}{\sum \vec{w}_a} = 0.915 \quad (12)$$

The results in Equations 10, 11, and 12 indicate an overall events, cycles classification accuracy and assigned energy rate of 94.8%, 92.5% and 91.5% respectively.

TABLE III
ENERGY ESTIMATION OF TARGET APPLIANCE OVER THE COURSE OF ONE DAY

App-name	active-time	mean-power	Ener-total	Ener-spin	Ener-heat	num-Heat	num-Spin
Fridge	520.38	130.98	1154.77	-	1154.77	32	-
washer	90.48	1762.03	388.18	106.92	281.27	2	250
Dishwasher	77.55	1757.80	740.14	-	740.14	2	-
Microwave	15.12	1019.34	127.305	-	127.305	7	-
Stove	63.82	2254.31	963.14	-	963.14	48	-
Dryer	56.38	1951.84	1554.63	-	1554.63	26	-

The results in table III show energy estimation, 4^{th} column, of target appliances with the 7^{th} column showing number of pair of heating events or fridge cycles in a day for each appliance. The last column shows the number spin cycles detected in one washing operation. When compared to the ground truth energy vector $E_{GT} = \{1211.35, 453.468, 892.728, 168.724, 1016.457, 1643.246\}$ Watt-hours of the target appliances, the energy estimation accuracy attained is around 91.5%.

V. CONCLUSION

This paper introduces a spike-like geometric features of the slope for appliance identification and presents an active window based unsupervised localized clustering and matching technique for monitoring home appliances utilizing appliances geometric features and usage patterns. The proposed approach for event detection, feature extraction and active window based appliance classification technique along with energy estimation are described in the paper. We developed a simple NILM system and then evaluate its accuracy using a test bench quality measures. Our experimental results clearly show that the disaggregation algorithms perform well.

VI. FUTURE WORK

As previously mentioned, this is an ongoing research and our work on NILM is still in progress. In our future work we envision to improve the various stages of the algorithms and perform continuous evaluation and results verification of the proposed approach. At the moment testing is done on our own dataset, but in the future we will continue testing in various publicly available dataset. In the future, further research is required to increase the overall accuracy of the disaggregation results.

ACKNOWLEDGMENT

We gratefully acknowledge the support received from Midori S.r.l.

REFERENCES

[1] G. W. Hart, ”Non intrusive appliance load monitoring,” Proc. IEEE, vol. 80, no. 12, 1992.

[2] Zoha, A.; Gluhak, A.; Imran, M.A.; Rajasegarar, S. ”Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey.” Sensors 2012, 12, 16838–16866.

[3] M. Zeifman and K. Roth, ”Nonintrusive appliance load monitoring: Review and outlook,” 2011 IEEE Int. Conf. Consum. Electron., pp. 239–240, 2011.

[4] M. Berges, E. Goldman, H. S. Matthews, L. Soibelman, and K. Anderson, ”User-Centered Nonintrusive Electricity Load Monitoring for Residential Buildings,” Journal of Computing in Civil Engineering, vol. 25, no. 6. pp. 471–480, 2011.

[5] S. B. Leeb, S. R. Shaw, and J. L. Kirtley, ”Transient event detection in spectral envelope estimates for nonintrusive load monitoring,” IEEE Trans. Power Deliv., vol. 10, no. 3, pp. 1200–1210, 1995.

[6] L. K. Norford and S. B. Leeb, ”Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms,” Energy Build., vol. 24, no. 1, pp. 51–64, 1996.

[7] Clinton Zheng. Spike sorting report v2.1.

[8] W. L. Chan, A. T. P. So, and L. L. Lai, ”Harmonics load signature recognition by wavelets transforms,” DRPT2000. Int. Conf. Electr. Util. Deregul. Restruct. Power Technol. Proc. (Cat. No.00EX382), 2000.

[9] Tang, Guoming, Wu, Kui, Lei, Jingsheng, and Tang, Jiuyang, Plug and Play! A Simple, Universal Model for Energy Disaggregation, Mar. 2014.

[10] H.-H. Chang, C.-L. Lin, and J.-K. Lee, ”Load identification in nonintrusive load monitoring using steady-state and turn-on transient energy algorithms,” Comput. Support. Coop. Work Des. (CSCWD), 2010 14th Int. Conf., 2010.

[11] P. Klein, J. Merckle, D. Benyoucef, and T. Bier, ”Test bench and quality measures for non-intrusive load monitoring algorithms,” in IECON 2013 – 39th Annual Conference of the IEEE Industrial Electronics Society, 2013, pp. 5006–5011.

[12] L. Farinaccio and R. Zmeureanu, ”Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses,” Energy Build., vol. 30, no. 3, pp. 245–259, 1999.

[13] Rainer Storn, Kenneth Price, ”Differential Evolution-A Simple and Efficient Heuristic for global Optimization over Continuous Spaces”, Journal of Global Optimization, December 1997, Volume 11,Issue 4, pp 341-359.

[14] Liu, Y., Wu, X., Shen, Y. ”Automatic clustering using genetic algorithms. Applied Mathematics and Computation”, vol. 218, no. 4, pp. 1267-1279 (2011) Show original message