2013 4th IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe), October 6-9, Copenhagen

Real-time and low cost energy disaggregation of coarse meter data

Emmanouil Vogiatzis, Georgios Kalogridis and Stojan Z. Denic
Toshiba Research Europe Limited, Telecommunications Research Laboratory, 32 Queen Square, Bristol, BS1 4ND, UK
Email: {emmanouil.vogiatzis, george, stojan.denic}@toshiba-trel.com

***Abstract*—A novel real-time non-intrusive appliance load monitoring algorithm is introduced. By applying finite state machine training, dynamic appliance disaggregation, rule-based filtering and discrete Fourier analysis of coarse data, several improvements are achieved: 1) reduced appliance model training complexity as compared to existing algorithms, and 2) novel and augmented detection stage. The application of this algorithm on real world data demonstrates that efficiency does not have to be compromised by the relatively lower complexity.**

I. Introduction

Energy disaggregation, also referred to as Non-intrusive Appliance Load Monitoring (NALM), is a method that helps identify appliance usage patterns that constitute a measured energy profile. Energy usage analysis, described by some authors as the "holy grail of energy efficiency", involves a combination of power control systems, and information and communication technologies (ICTs) to help develop advanced metering functionalities for the smart grid (SG), including demand response (DR) and demand side management (DSM), and applications such as load balancing and privacy [1].

Our proposed system is based on NALM principles developed in [2] and combines a) an offline and low complexity training method, and b) fast processing and low memory/cost real-time detection. We focus on the quite challenging problem of reliable software disaggregation of low sampling frequency energy data. Such data might be readily available using standard (nationwide deployment of) home smart metering systems. The developed algorithms operate autonomously, with no a-priori knowledge of appliance profiles, and the real-time detection component is designed to require low processing and memory. Our prototype implementation uses only active power measurements obtained every 20 seconds from a set of houses in Bristol City (UK), and it exhibits promising results.

The most important contributions of our proposed algorithms are summarised as follows.

- We modify the FSM optimization algorithm introduced in [3] to help speed up the training process and cope with lower frequency measurements.
- We modify the Knapsack algorithm introduced in [4] to help address a range of problems with real data, such as classification errors, variations in appliance consumption, errors due to low and unreliable measurement granularity, and the occurrence of unknown events including 'vampire power' and noise.
- We introduce a lookup table in order to improve the speed and reduce the cost of the recursive detection algorithm.
- We introduce a discrete Fourier analysis to help improve fridge detection.
- We use state transition rules obtained from the FSM table to further enhance the performance of the main algorithm and mitigate erroneous detection results.

The rest of the paper is organised as follows. Section II reviews the NALM literature, §III discusses a system architecture and §IV presents our technical contributions. Numerical results are given and discussed in §V, and, finally, §VI concludes this paper.

II. Related Work

The original NALM method is attributed to Hart [2] and it involves a custom recoding system and a cluster analysis of active and reactive power data. Since this original work, the field of NALM has been extensively studied in e.g. [3]–[13].

In [5], data is sampled with very high frequency (≥ 8 kHz); here, the observation of higher harmonics and transient behavior in the aggregated signal is utilized to generate appliance signatures. In [6], appliance disaggregation is made possible by observing the noise produced by the appliance switches. That is, the switching noise characteristics generated by different appliances are considerably different enabling the NALM through appliance switching noise modelling. The noise characteristics are analyzed by the Fourier transform and support vector machines. Gupta et al., [7] propose a voltage sampling frequency of 500 Hz or more, claiming 85-90% accuracy. Other granular methods build appliance fingerprints by sampling the waveform in scales of milliseconds [8]. Such proposals report 80% to 95% accuracy. A limitation of such systems involving an analysis of signal high harmonics is a high deployment cost and a manual calibration process.

Using less granular measurements, e.g. obtained once a second, [9] propose pattern matching assuming a (static) appliance database. M. Baranski and J. Voss [3] further propose a combination of statistical and optimization techniques to build appliance databases. Requiring zero knowledge about user behavior or other information relevant to appliances, this algorithm offers a variety of parameters to be tuned. At the beginning, the edge events are clustered, and a genetic algorithm is employed to determine best possible patterns of such clustered events. Once these patterns are established, for each pattern a dynamic programming algorithm finds

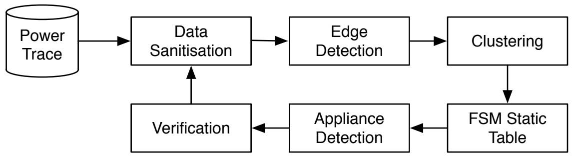

Fig. 1. Appliance disaggregation process.

the best sequence of clusters to model an appliance. At the end, possible conflicts among finite state models for different appliances are resolved. Bergman et al. [4] expanded this work by adding a distributed mechanism and automated re-learning to help improve the accuracy of appliance models and provisions for changes in the operating set of appliances. While reporting an accuracy of 77% to 90%, [4] indicate that their method is successful for loads of at least 1 kW.

In [10], NALM is used in commercial buildings; new challenges are identified, such as the existence of many similar loads, loads that vary smoothly over time, and loads which do not exhibit reactive power.

Some recent papers employ more complex probabilistic models such as hidden Makov models (HMMs) and their extensions. For example, in [11], factorial HMMs are proposed and are tested for publicly available energy consumption data. In [12], a more complex factorial HMM has been provided to include additional appliance features such as time of the day and dependency among appliances. An unsupervised two phase learning technique has been proposed which relies on an HMM structure. Parson et al. [13] addresses appliance model learning by tuning a HMM model from sections of the aggregate consumption data.

Our work extends [4] and it exhibits good detection results with real household data, even for appliance loads less than 100 Watts, such as of a fridge. Further, our method copes with measurements sampled every 20 seconds, which is less granular than the 1 second readings used in [3], [4]. This is due to a number of techniques introduced such as dynamic estimation of power noise and frequency analysis.

III. System Architecture

NALM is a key technology for the SG with applications in load control, load-shed verification, DR/DSM and privacy protection [1]. The type of NALM application determines a structure of overall system architecture.

With regards to privacy, NALM algorithms might be used to test the strength of protection algorithms; the role of protection here is inverse to NALM, i.e. its purpose is to scramble consumers' metering data such that this makes energy disaggregation more difficult [14]. Another role of NALM is to enable load monitoring and control, which helps improve energy efficiency and reliability. Such controls will typically be centrally managed and may involve generating instructions (or recommendations) for SG stakeholders within the generation, distribution, or consumption power network. It is noted that without NALM, monitoring appliance loads may alternatively require the use of intrusive methods (e.g. smart plugs).

As noticed in [4] different architectural solutions can be envisioned for incorporating NALM into the smart grid; e.g. NALM can be placed next to either the central controller, the smart meter, or both (in a distributed manner). In [4], the learning phase takes place in the central controller: edge events detected by the smart meter are communicated to the central controller that builds an appliance FSM table. For the detection phase, the smart meter uses the FSM table to identify large appliances in real time.

IV. Disaggregation Process

The overall process of our disaggregation algorithms is given in Fig. 1. It consists of a training phase (data sanitisation, edge detection, clustering, and FSM table construction) and an (appliance) detection phase. Further, we discuss the main steps of our algorithms in more detail.

A. Training

The training process involves the construction of appliance models that might characterize different households. We propose a modified and less complex FSM representation than the one in [3], which results in faster selection of appliance state permutations. Further, we emphasize several main steps of our FSM model design.

Given a series of power measurements, P_t, an appliance profile may be characterized by a succession of power edges, $\Delta P_t = P_t - P_{t-1}$. Roughly speaking, if frequent measurements (e.g. 1 s) are available an event may be defined as a subsequence of ΔP_t; in this paper we consider that a ΔP_t constitutes an event S_i. Once the sequence of events $\mathbf{S} = (S_1, S_2, \cdots)$ is formed, the clustering is performed on $P(S_i)$, where $P(S_i)$ is the amplitude of S_i. The cluster centers C_r, $1 \leq r \leq N_C$, are used to represent the FSMs. If an FSM of the jth appliance has s states, it can be represented by a vector $\mathbf{Z}_j = (z_{j,1} \cdots z_{j,s})$, where $z_{j,i}$, $1 \leq i \leq s$, takes a value from the set $(C_1, C_2, \cdots, C_{N_C})$ of cluster centers.

Considering that $z_{j,i}$ can be positive or negative, the order of states in $\mathbf{Z}_j$ has to be chosen such that a corresponding power trace is positive, i.e. $\sum_{i=1}^{k} z_{j,i} \geq 0$ for $1 \leq k \leq s$ (condition 1). In addition, each cluster center may belong to only one appliance (condition 2). Using the frequency or appearance and zero-sum criteria introduced in [3], a 'quality' value for any $\mathbf{Z}_j$ may be readily evaluated. The above practical considerations form the basis of our low-complexity FSM optimization algorithm, which is summarized as follows:

1) Define the matrix with rows containing all permutations of cluster centers of length at most s, with the inclusion of the zero element.
2) Delete the rows which do not satisfy condition 1.
3) Sort the remaining rows according their quality, and exclude the worst 50% of rows.
4) By using the quality metric (defined in [3]), choose the best FSMs such that condition 2 is satisfied.

The Z_n remaining rows characterize the identified FSMs. An association of FSM rows with particular appliances is further discussed in §V, where we show, by example, that

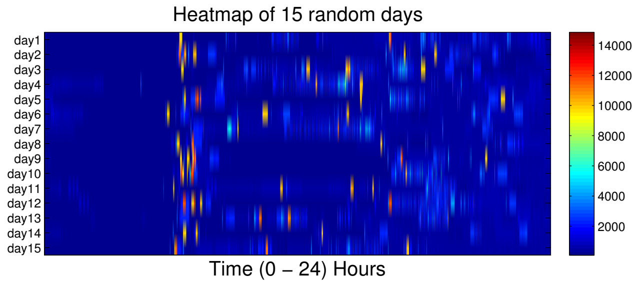

Fig. 2. 15 day heatmap of active power consumption (Watt).

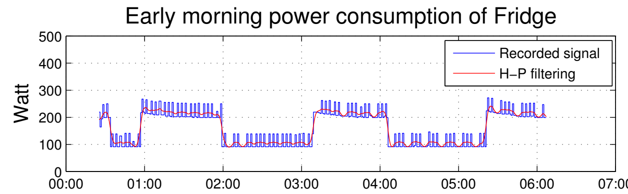

Fig. 3. Fridge measurements.

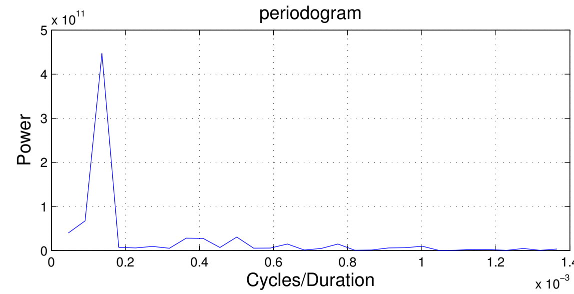

Fig. 4. Estimation of the fridge spectrum.

the modified training approach supports a very good detection performance.

We note that our proposed training process (FSM optimization) is simpler and faster than the generic one in [3].

B. Detection

We organise our technical contributions and optimisations for the detection process as follows.

1) Knapsack Adaptation: The knapsack optimization problem is formulated as: maximise $W(t) = \sum_{i=1}^{Z_n}\sum_{j=1}^{s} w_{i,j}(t)z_{i,j}$, subject to $W(t) \leq v(t) + T$, where $v(t)$ is the observed aggregate power, T is a tolerance value, and $w_{i,j}(t)$ takes values from $\{0,1\}$ with the constrain that $\sum_{i=1}^{s} w_{i,j}(t) = 1$, for all j, at any time t. We further set $T = x - B$, where B is a 'base' power and $0 \leq x \leq x_{\max}$. We set $x_{\max} = 500W$ to represent a maximum noise tolerance for the Knapsack algorithm. To calculate B we find the base (minimum) consumption in the training data, following the application of a Hodrick-Prescott (HP) filter.

2) Lookup Table: With regards to reducing the computational cost of finding Knapsack solutions (in real-time), we construct an array representing all plausible solutions. The lookup table may then be constructed by sorting the solutions that correspond to a choice of $w_{i,j}(t)$ and the associated $W(t)$.

3) Fridge Periodicity: With regards to the Fourier analysis discussed above, we observe that household appliances can exhibit two types of behavior, cyclic and acyclic. The most prominent cyclical appliance is a fridge (see, for example, Fig. 3). This appliance characteristic can be conveniently used to detect a cyclic appliance. One way to detect the cyclical behavior of the fridge is by applying the Fourier transform on the power signature during the night interval when the other appliances are not frequently active. The choice of a night interval may be justified, in this example, after observing the times of minimal power consumption in Fig. 2.

In Fig. 3 it can be observed that apart from the 'main' fridge signal, there is a 'parasitic' modulation which we want to suppress in order to isolate only the 'main' cycle and frequency. Therefore, the low-pass frequency HP filter is applied before the Fourier transform. After the application of the Fourier transform, the spectrum of the fridge signal is given in Fig. 4. This gives, in this particular example, a fridge cycle of approximately 2 hours and 2 minutes. Such an estimation of the fridge cycle duration is further used to improve our detection results, as discussed immediately below.

4) Overall Process: The operation of the real-time (and low-cost) detection algorithm can now be described as follows.

1) Given $v(t)$, find the Knapsack solution for $x = 0$.
2) Check the validity for all appliance state transitions, requiring such transitions to occur in a sequential order.
3) Check the validity of the fridge periodicity, requiring a transition to occur within an expected time window.
4) If the result is invalid, increase x by 25 and repeat from step 1, otherwise record the result and update the current state of each appliance.

V. Results and Discussion

Our results are based on energy data collected from real residencies in Bristol (UK) during the course of the '3eHouses' EU FP7 project [15]. An evaluation of the results is possible as particular appliance readings are also available. However, no information from particular appliance readings are used during the disaggregation (training and detection) phase.

For this discussion, results for two particular, albeit typical, houses are given. The dataset for each house contains (synchronized) aggregate and appliance instant power values sampled every 20 seconds covering a period of 10 days. The training algorithm uses the first five days of aggregate data, while the rest five days are used for detection. The overall process is repeated independently for each house.

For the training process we use a maximum of $s = 3$ states; the results are given in Table I. While the relation between real and hypothetical appliances is done by observation, we leave this issue for future work.

In Fig. 5, we highlight variations of power consumption of different appliances. This signifies the problem of building an FSM to represent a distribution of power values.

TABLE I
FSM TABLES FOR HYPOTHETICAL APPLIANCES, AS IDENTIFIED FROM DISAGGREGATION TRAINING, VS. REAL APPLIANCES, FOR TWO HOUSES.

			Appliances	State 1	State 2	State 3
Real	(House 1)	(r1)	Fridge	122	22	0
		(r2)	Cooker	1994	17	0
		(r3)	Microwave	1218	28	0
		(r4)	TV/AV	282	222	120
		(r5)	Kettle	2896	12	0
		(r6)	XBox	28	128	107
Hypoth.	(House 1)	(h1)	Fridge?	130	0	0
		(h2)	Cooker + Microwave?	1960	1200	0
		(h3)	Unknown	550	790	82
		(h4)	TV/AV	280	105	0
		(h5)	Kettle?	2670	2570	0
		(h6)	Electric shower?	9620	9680	0
Real	(House 2)	(r1)	Fridge	64	9	0
		(r2)	Cooker	1756	63	0
		(r3)	Microwave	1391	30	0
		(r4)	Toaster	767	1	0
		(r5)	Kettle	2759	62	0
Hypoth.	(House 2)	(h1)	Fridge?	66	6	0
		(h2)	Cooker?	1767	1822	27
		(h3)	Toaster?	734	801	54
		(h4)	Kettle?	2996	3009	0

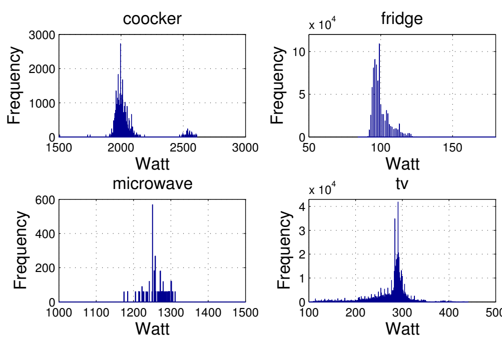

Fig. 5. Histogram of appliances for 15 days (house 1).

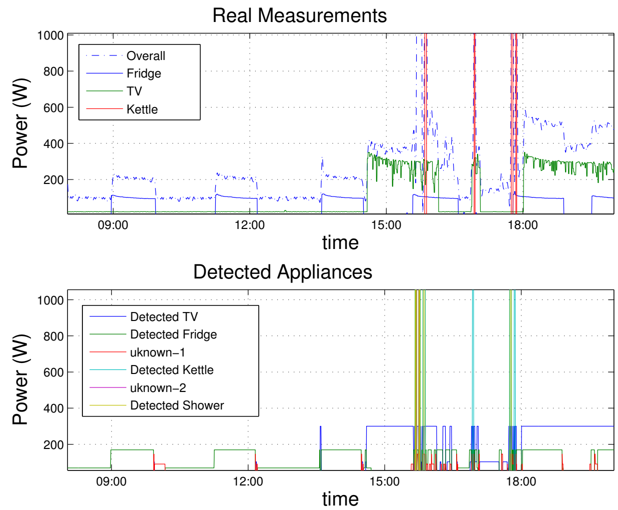

Fig. 6. Real measurements vs. detection results for house 1.

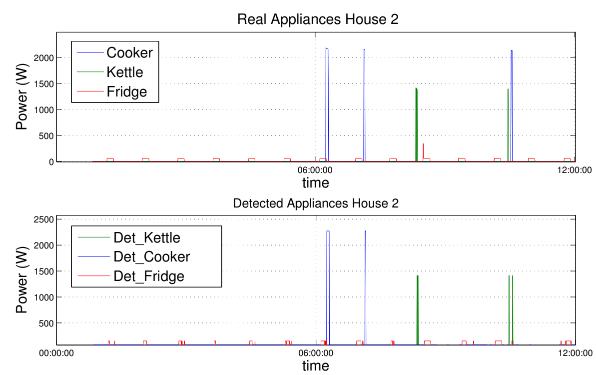

Fig. 7. Real measurements vs. detection results for house 2.

To estimate the base (noise) power level, B, we record the minimum power value between 1–5 am. For K-means clustering, we use $N_C = 18$. In general, N_C will affect a) the number of FSMs created, and b) the performance of the overall system. For a maximum of $s = 3$ states, typically the algorithm creates 4–7 hypothetical appliance FSMs.

The application of our developed algorithms (on different days of household data) is showcased in Table II for house 1. Further, a visualisation of such results for one day is given in Fig. 6, for house 1, and in Fig. 7 for house 2. In Table II, we note that the energy disaggregation detection error for a period of one day is reasonably low. In Fig. 6 it can be observed that appliances with low and variable power consumption (less than 500 Watts), such as fridge and TV, are detected with sufficient accuracy. This demonstrates the practical value of the contributions of this work (e.g. as compared with [4], which has success only with high consumption appliances such as the kettle). In Fig. 7 we further observe a) a successful detection of the fridge (despite the very low consumption value of 64 Watts), and b) a miss-detection of a kettle (around 11 am) instead of a (real) cooker. This may to some extend be due to distortion due to excessive noise or atypical consumption. A more particular visualization of a successful TV detection can be seen in Fig. 8. We observe that the detection of an oscillating on/off operation has little effect on the (daily) energy disaggregation results (see Table II). However, we note that while we observe success in this particular example, there are other cases where more significant errors occur.

A. Discussion

Most NALM algorithms employ a principle of observing changes in the power level of the appliance signals. To make a distinction between different appliances, the NALM algorithms count on differences among power signatures of different appliances. The larger the difference between two signatures, the larger the chance that a NALM algorithm will be able to disaggregate these two appliances. On the other hand, the chance for errors is larger if the algorithm is simple and only takes into account the start-up and shut-down edges of the power signatures. If the start-up edges of two appliances are similar this may lead to their misinterpretation.

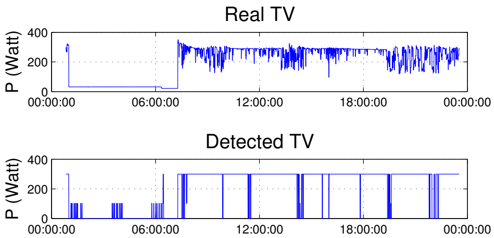

Fig. 8. Real measurements vs. detection results for TV.

TABLE II
DISAGGREGATION RESULTS (BASED ON A 5-DAY TRAINING DATASET, AND A DIFFERENT 5-DAY DATASET FOR REAL-TIME DETECTION).

	Day 1	Day 2	Day 3	Day 4	Day 5
Cooker (Wh)	1576.4	2647.1	2447.8	1354.0	2247.5
Detected (Wh)	1132.5	3034.8	2826.4	1645.1	2180.6
Time_real (sec)	227	487	474	124	303
Time_det (sec)	165	551	551	195	314
Error_Wh (%)	-28.2	14.6	15.5	21.5	-2.9
Error_time (%)	-37.6	11.6	16.2	57.3	3.6
Fridge (Wh)	2695.3	1767.3	2251.0	2375.9	3133.3
Detected (Wh)	2790.0	2133.5	2858.9	2720.5	3550.0
Time_real (sec)	3099	2930	2256	2433	3490
Time_det (sec)	2743	3231	2810	2674	3676
Error_Wh (%)	3.5	20.7	27.0	14.5	13.3
Error_time (%)	-11.5	10.3	24.6	9.9	5.3
TV (Wh)	2473.9	2766.6	2934.0	2669.5	4103.2
Detected (Wh)	2583.9	2323.6	2954.0	3550.0	3550.0
Time_real (sec)	1171	1989	1787	1719	3490
Time_det (sec)	1327	1618	1704	2490	2476
Error_Wh (%)	4.4	-16.0	0.7	32.3	-13.5
Error_time (%)	13.3	-18.7	-4.6	44.9	-29.1

This problem could be aggravated since the clustering methods are applied for building finite state machine representations of the appliances. The clustering method introduces additional noise, making the performance of simple algorithms worse. An ambiguity can also arise due to the close timing of start-up edges of two appliances; a low sampling rate of the NALM can cause that these two edges be understood as one appliance having an amplitude equal to the sum of the powers of two separate appliances. Then, the NALM algorithm may not be able to find a corresponding shut-down edge for this non-existing appliance. The situation becomes worse if there is a large number of possible appliances and the sampling frequency remains low, and if there is large invariability in the values of the states of appliances in operation.

This paper makes some progress in mitigating such problems (e.g. by dynamically adjusting the tolerance in the Knapsack algorithm, and by employing a set of rules with regards to the order of state transitions and, in particular, with the periodicity of the fridge). However, it is clear that the problem of miss-classification and miss-detection results remains, and further research is required.

VI. CONCLUSION

We provide a novel low complexity real time energy disaggregation algorithm for household appliance detection. We have reduced the complexity of the training stage as compared to existing algorithms and introduced several improvements in the detection stage of the algorithm. By employing the algorithm on the real world data, we have demonstrated that efficiency of the algorithm has not been compromised by its low complexity. Further research is required to address a number of challenges identified during the experimentation with the developed algorithms, and offer a level of reliability required in commercial applications.

ACKNOWLEDGMENT

The authors would like to thank the Directors of the Toshiba Telecommunications Research Laboratory for their support, as well as Dr. Zhong Fan for his invaluable comments.

REFERENCES

[1] G. Kalogridis and D. Saraansh, "PeHEMS: Privacy enabled HEMS and load balancing prototype," in *IEEE International Conference on Smart Grid Communications (SmartGridComm'12)*, Oct 2012, pp. 232–237.

[2] G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec 1992.

[3] M. Baranski and J. Voss, "Genetic algorithm for pattern detection in NIALM systems," in *IEEE International Conference on Systems, Man and Cybernetics*, vol. 4, Oct 2004, pp. 3462–3468.

[4] D. C. Bergman, D. Jin, J. P. Juen, N. Tanaka, C. A. Gunter, and A. K. Wright, "Distributed non-intrusive load monitoring," in *IEEE PES Innovative Smart Grid Technologies (ISGT)*, Jan 2011, pp. 1–8.

[5] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong, "Power signature analysis," *IEEE Power and Energy Magazine*, vol. 1, no. 2, pp. 56–63, 2003.

[6] 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*, ser. UbiComp '07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 271–288.

[7] S. Gupta, M. S. Reynolds, and S. N. Patel, "Electrisense: single-point sensing using emi for electrical event detection and classification in the home," in *Proceedings of the 12th ACM international conference on Ubiquitous computing*, ser. Ubicomp '10. New York, NY, USA: ACM, 2010, pp. 139–148.

[8] A. Reinhardt, D. Burkhardt, M. Zaheer, and R. Steinmetz, "Electric appliance classification based on distributed high resolution current sensing," in *37th IEEE Conference on Local Computer Networks Workshops*, 2012, pp. 999–1005.

[9] M. Berges, E. Goldman, H. S. Matthews, and L. Soibelman, "Training load monitoring algorithms on highly sub-metered home electricity consumption data," *Tsinghua Science and Technology*, vol. 13, no. S1, pp. 406–411, 2008.

[10] 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 and Buildings*, vol. 24, no. 1, pp. 51–64, 1996.

[11] J. Z. Kolter and M. J. Johnson, "Redd: A public data set for energy disaggregation research," in *Workshop on Data Mining Applications in Sustainability (SIGKDD)*. Curran Associates, Inc., 2011.

[12] H. Kim, M. Marwah, M. F. Arlitt, G. Lyon, and J. Han, "Unsupervised disaggregation of low frequency power measurements," in *The 11th SIAM International Conference on Data Mining*, 2011, pp. 747–758.

[13] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Using hidden markov models for iterative non-intrusive appliance monitoring," in *Neural Information Processing Systems workshop on Machine Learning for Sustainability*, 2011.

[14] G. Kalogridis and S. Z. Denic, "Data mining and privacy of personal behaviour types in smart grid," in *IEEE 11th Int. Conference on Data Mining Workshops*, ser. ICDMW'11, December 2011, pp. 636–642.

[15] "3eHouses European Union FP7 project," http://www.3ehouses.eu/, 2013, [Online].