An appliance-driven, submeter-based energy disaggregation approach

Aimilia Papagiannaki[a]*, Stelios Koutroubinas[a], Spyros Sioutas[b], Konstantinos Tsichlas[b]

[a]Meazon S.A., 57 N.E.O Patron – Athinon str., 26442 Patras, Greece
[b]Department of Computer Engineering and Informatics, University of Patras, 26500 Rion, Patras, Greece

a.papagiannaki@meazon.com

Abstract

In the age of energy crisis, the scientific and industrial worlds come together to provide solutions and suggestions to reduce energy consumption and respective economic burden. In the energy efficiency direction, problems like non-intrusive load monitoring are still under intense investigation on how to be applied to real-world circumstances with undoubtable results. The current work proposes a framework to overcome existing obstacles in terms of IoT devices involved and algorithmic approaches. To be more precise, the utilization of separate appliance-oriented models is proposed, exploiting heuristic rules that emerge from the very electrical characteristics of the appliances. A hybrid solution is implemented and tested in-vivo for the case of fridges that operate within households, producing promising results and utilizing a powerful yet light smart meter.

1.   Introduction

During the last year, Europe has been standing in the middle of a forthcoming energy crisis, while the political and scientific community is trying to solve a difficult puzzle to disentangle the problem of seamless power supply. As the Earth is running out of fossil fuels, renewable energy sources are becoming more and more popular, although environmental issues are still to be solved. On top of that, industrial and domestic consumers are worried about the energy cost, thus they are motivated more than ever to reduce energy-consuming loads.

The above situation leads naturally to tackling the problem of non-intrusive load monitoring (NILM) that aims to identify underlying loads that are hidden in a measured aggregated power signal. This is useful since the goal of energy efficiency at the household level can be achieved through energy-saving policies that change the behavior of the consumers. This requires as a prerequisite to be able to measure the energy consumption at the appliance level.

The commercial impact of NILM at the level of the home customers [10], is the increased utility customer engagement and the reduced energy usage. The goal at this level is to itemize the consumer's energy bill [7], analyze the energy usage and cost per household appliance, make personalized and prioritized energy savings recommendations or even go as far as identify faulty appliances (e.g., frosting cycle of a fridge with a damaged seal is more frequent than a normal one [8]). All these should be viable through a single sensor per household that monitors the total

energy consumption and other related quantities [9]. The requirement in NILM is to identify the usage of an appliance merely by its signature on the aggregate energy waveform of the household.

Under that premise, significant effort has been made to develop smart plug-and-play solutions that combine little-to-no user engagement and low-cost IoT devices. To this end, the current article aims to present a work-in-progress that utilizes a powerful but low-cost smart meter to build appliance-driven models able to detect different loads within households. What makes the solution different than most of the existing approaches, is that no consumer engagement is required in the model building process, as well as the fact that partly high and partly low-frequency data is used to reduce network and memory-related burden.

The rest of the paper is structured as follows: Section 2 refers to previous work in the field, while the methodology and tools are described in Section 3. Early results are presented in Section 4 and we conclude with a discussion in Section 5.

2. Related Work

A significant number of research papers has been published in the field of energy disaggregation, with methodologies varying from neural network implementations [1] to traditional classification schemes [2], all tested against well-known open datasets available to the research community. In general, event-based approaches are employed or state-based approaches [3]. Usually, they are based on HMMs (Hidden Markov Models) and state transitions are obtained automatically from the hidden state estimation. The event-based methods can be further categorized into three classes: expert-heuristic based (state changes are detected by predefined empirical rules), probabilistic model based (e.g., state changes are detected by probabilistic models - HMMs) and matched-filter based (state changes are detected by signal processing methods).

Although a great deal of the approaches is supervised and utilize ground truth information from household appliances to develop the respective models, multiple papers [4, 5, 6] emphasize on the advantages of following unsupervised approaches that reduce the effort to collect separate data from electric devices. In particular, [4] presents a set of statistical and heuristic rules that determine the electrical profile of fridges and refrigerators, managing in this way to detect all respective instances of this type of appliances regardless of the specific model/manufacturer. Similarly, the authors in [5] break down the particular operating cycles of washing machines and define the respective active power limits to build an appliance signature able to capture instances of washing machine functioning within a household. Lastly, in [6] the target appliances are the air conditioning units, a puzzling category of electrical appliances with great deviations in operating active power amplitude and respective behavior. Again, the problem of appliance identification is approached through the discovery of generalized rules that can discriminate the air condition from similar devices. Needless to say, that these methodologies are also accompanied by puzzling challenges with regards to algorithm accuracy.

All prior research in the field is highly restricted by the hardware & cloud system constraints that arise when the suggested solutions are applied in-vivo. The current work aims to address these restrictions in both directions, by exploiting a light-yet-powerful edge device, coupled with a hybrid data transmission plan.

3.   Methodology

The current approach is inspired by the work described in [4, 5, 6], where separate appliance models are developed based on their generalized characteristics, and respective heuristic rules are defined. Instead of collecting big bulks of ground truth data from different models of different appliances, which is a time-consuming and high-cost procedure, a breakdown of the separate appliances' signatures is performed to describe the respective electric profiles. A schematic illustration of the above framework is depicted in Figure 2.

The first attempt to implement such a solution is described in the current paper, where we adapted and implemented part of the procedure described in [4], concerning the characterization of fridges that operate within households. In the following subsections we describe the corresponding steps of the implementation.

Before diving into the details of the methodology, it's crucial to present the standard electrical pattern of any common fridge. A usual motor turn-ON of a fridge starts with an instant spike in active power that may reach some hundredWatts (even kWs), followed by a gradual smooth reduction of power until it reaches a steady state. The above phenomenon is illustrated in Figure 1.
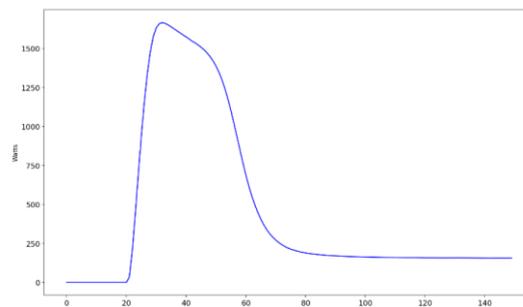

Figure 1: Common turn-ON of fridge's motor

3.1   Event-Detection

The first step towards the identification of fridge is to detect the respective events within a timeseries, and more specifically the 'On' event which is quite characteristic in the case of fridges. As described in [4], this is accomplished by computing the 1st derivative of the active power and respective 3 local maxima, on a sliding window of high frequency. The derivative of a single point k is defined as the difference between the exact next and previous points:

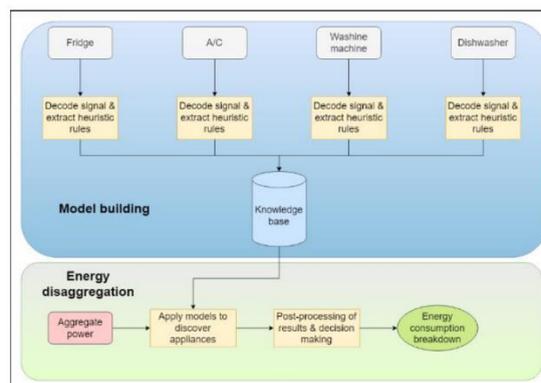$$\delta_k = P_{k+1} - P_{k-1} \qquad (1)$$

Figure 2: Generic framework for appliance detection & energy disaggregation

The respective 3 local maxima that emerge from the 1st derivative within each sliding window separate the time-series in 4 distinct sub-areas. As explained in [4], we calculate the average power of each sub-area and define some "rules" (see equations (3), (4), (5) and (6) of [4]) that need to be met in order for a sliding window to be considered as an "On" fridge event. When an "On" event is detected, the average delta active and reactive power are calculated, to serve as reference values for the upcoming events.

In parallel with the above procedure, we also check whether there is an "Off" event detected within a window. To eliminate the number of events detected, we narrow down the search to a delta active power that lies within the range of [30, 200] Watts. This range includes the normal values of standard fridges when the motor turns off. Here again the average delta active and reactive power are calculated for further use.

3.2 Coupling of On-Off events

To acquire a full operating cycle of a fridge, the detected "On" events must be coupled with respective "Off" events. To achieve that, we take advantage of the average active and reactive power of the "On events". Knowing that the steady state of the fridge right before it turns off is slightly lower that the steady state after it turns on, we define the following thresholds to discover the fridge "Off" events:

$$P_{off} < \widetilde{P_{on}} \qquad (2)$$
$$P_{off} > 0.7 * \widetilde{P_{on}} \qquad (3)$$
$$Q_{off} < 1.1 * Q_{on} \qquad (4)$$
$$Q_{off} > 0.8 * Q_{on} \qquad (5)$$

where $Poff$, $Qoff$ are the active and reactive power of "Off" events respectively, and $\widetilde{Pon}$, $\widetilde{Qon}$ are the average active and reactive power of "On" events. The rules emerged from a short exploratory data analysis and observation of the active and reactive power on steady state at the beginning of the operation cycle and at the end. More specifically, the active power at steady state right before the motor turns off, is always lower than the respective active power at the beginning of the motor operation. On the other hand, the reactive power before turn-off may reach the amplitude of the turn-on reactive but can't be much lower (e.g. half the amplitude). Therefore, we narrowed down to the specific parameters' values after experimenting with several thresholds, through a trial-and-error procedure. This step ensures that only the "Off" events that are similar to "On" events will be considered, supposedly indicating the end of a fridge's motor full operating cycle.

3.3 On/Off duration calculation

A significant step of the procedure is the estimation of time that the fridge remains at "On" mode. Through this step, we can eliminate false positives of "Off" events, while the duration will also be utilized to calculate the energy consumption of the fridge. To achieve that, we sort the events chronologically so that each "On" event is followed by an "Off" event and calculate the time difference between each couple of events. Since the entire approach is unsupervised and thus false positive events may occur, it's safe not to take into account all the time differences we calculated earlier. It is rather safer to take a percentile of the time differences, to serve as a representative value of the duration. Specifically, the 75th percentile of the "On" duration is considered, as well as the 75th percentile of the time difference between an "Off" and "On" (to

estimate how much time after a fridge has turned off, it will start operating again). These two values are denoted as Ton and $Toff$ to describe the duration (in seconds) of the "On" state and the time where the motor remains at "Off" state, respectively.

3.4 Energy consumption estimation

Up to this point, we have established that all the information needed to detect a full fridge operation cycle is available, along with the corresponding characteristic values with regards to active/reactive power range and duration. The last step of the approach is to estimate the energy consumption of the fridge within the detected operating cycles. Since the fridge is known to be a relatively stable load (after the initial spike when it turns on), we may estimate the energy consumption through the trapezoid area calculation. That is:

$$E = \frac{(P_{on}+P_{off})*\frac{T_{on}}{3600}}{2} Wh \qquad (6)$$

The above estimation is performed for all the time windows where the fridge was found to be in the "On" mode.

4. Experiments & Results
4.1 Hardware selection

This section is dedicated to the application of the aforementioned methodology to real conditions and the respective preliminary results that emerged from the experimentation. In detail, herein we describe the hardware equipment used for the experiments, as well as the setup where the measurements were acquired. Finally, the performance of the methodology is presented and discussed.

As previously mentioned, in order for the algorithm to work properly, high frequency data (tens of Hz) is necessary to be acquired, especially during the transition time slots. For that reason, the piece of hardware that was used to conduct the experiments is a smart meter designed and developed by Meazon (Meazon DinRail 3-phase Wifi), that is capable of measuring and transmitting data at two frequencies, providing a hybrid data acquisition solution. Specifically, the device is capable of measuring electrical variables such as active/reactive power, voltage, current, power factor and other, at frequency as high as ~33Hz (1sample/30msec). Although this frequency is sufficiently high to capture even the smallest changes in power, regularly transmitting at this sampling rate is not a realistic solution due to limitations in network and storage resources.

To overcome this challenge, a hybrid data transmission protocol has been implemented on the smart meter, to ensure that high-frequency data will only be transmitted when necessary. In detail, the smart meter is capable of detecting changes in power that exceed a predefined threshold, which in turn triggers the meter to start transmitting measurements at high rate for a few seconds, to ensure that the transition state of the power is accurately captured.

4.2 Household setup & data acquisition

The above device was installed at the central panel of an ordinary house (where also washing machine, oven, electric heater and microwave are present), to measure the aggregated power of the residence. The standard sampling rate was set to 1sample/minute, while the fast mode was set to 1sample/50msec for 3 seconds (after the 3 seconds the meter returns to its standard mode and transmits with lower frequency). The threshold of power to detect the events internally was

set to 100Watts. This installation made it easy to detect the respective events that take place during the day within the house. Although the approach is entirely unsupervised, we also installed a smart plug behind the fridge of the house to check the performance of the methodology. All measurements are transmitted in real-time and stored in Meazon's IoT platform for monitoring and visualization. The methodology was implemented as a scheduled process (in Python3) that runs within the virtual machine that hosts Meazon's platform. In particular, two discrete scenarios were defined, one as the first run of the methodology, where an unsupervised pseudo-training is performed, and one as the everyday run which detects the fridge's on/off and estimates the energy consumption. For the pseudo-training case, the process was set to retrieve data of the last 7 days, to ensure that adequate measurements will be available to extract fridge's electric signature. The necessary profile variables that emerge from this process (average active/reactive power, duration of "On" and "Off" states) are therefore stored for further use. Beyond this first run, the routine runs once a day to detect the fridge instances within the previous day's aggregated signal and estimate the corresponding energy consumption throughout the day.

4.3   Preliminary Results

The results from online scheduled process described above are stored and visualized in Meazon's IoT platform. An example of the respective results is illustrated in Figure 3. It depicts the aggregated signal of the house (blue) with the detected points of ON/OFF (green bullets), as well as the ground truth signal of the fridge (red).
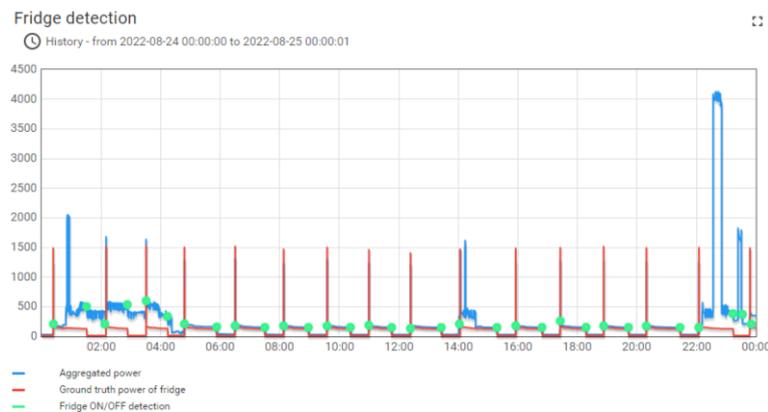

Figure 3: Online Detection of Fridge's ON/OFF

To evaluate the methodology, we monitored the process for approximately 3 weeks within August/September 2022. Although it's not a classification problem, we borrow the respective evaluation metrics to measure the performance of the event detection. Considering "Positives" all the instances of "On" and "Off", and "Negatives" all the other instances where the fridge remains at steady state, we define the following well known metrics:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{8}$$

$$Specificity = \frac{TN}{TN+FP} \tag{9}$$

TP refers to True Positives, TN to True Negatives, FP to False Positives and FN to False Negatives, respectively. The corresponding results of our experiments are depicted in Table 1 below.

Performance Metrics	Evaluation
Accuracy	94.86 %
Sensitivity	91.32 %
Specificity	96.32 %

Table 1. Performance of fridge "On" / "Off" detection

It is important to note here that the assumption we made about the "Negatives" being all the instances of steady state has a strong effect in the respective metrics containing "True Negatives" (Accuracy and Specificity), because the ratio of Negatives to Positives is very high in favor of Negatives (like having imbalanced classes). Thus, the metric that makes more sense to our analysis is sensitivity, where the number of correctly detected events to the number of ground truth events is calculated. Considering this metric, 91.32% is a very satisfying first result, suggesting that the generalized model we built is capable of correctly identifying a fridge without prior knowledge.

Regarding the energy consumption estimation, we evaluated the results through the percentage of relative error. The well-known formula is:

$$Relative\ error\ \% = \frac{real - estimated}{real} * 100 \tag{10}$$

Since the algorithm runs in daily basis, we calculated the daily relative error with respect to the ground truth, and then took the average of errors as an overall metric, along with the respective standard deviation. Table 2 illustrates the results. With an overall error up to 11.5% and standard deviation 8.58 it is safe to say that although energy estimation is satisfying, attention must be paid to the deviations between different days and the reasons behind this observation. There seems to be a direct relationship between the incidents of wrong event detection and high relative error. The reason behind false fridge detection is most likely to lie in the internal mechanism of the smart meter that identifies deltas in power and prompts a high-frequency data streaming. Potential differentiation in this internal process might give better results. Additionally, we might need to consider fine-tuning the parameters of the methodology that perhaps affect the performance.

Date (2022)	Estimated energy (kWh)	Real energy (kWh)	Relative error
Aug 23	2016	2279	11.54 %
Aug 24	2043	2179	6.24 %
Aug 25	1912	2309	17.19 %
Aug 26	2137	2239	4.55 %
Aug 27	1738	2280	23.77 %
Aug 28	1775	2278	22.08 %
Aug 29	2385	2318	-2.89 %
Aug 30	1917	2281	15.95 %
Aug 31	2265	2231	-1.52 %
Sep 1	2235	2236	0.04 %
Sep 2	1705	2270	24.88 %
Sep 3	2003	2240	10.58 %
Sep 4	2229	2364	5.71 %
Sep 5	2027	2221	8.73 %

Sep 6	2306	2279	-1.18 %
Sep 7	1627	2217	26.61 %
Sep 8	1886	2143	11.99 %

Average relative error	**11.5** %
Standard deviation of error	8.58

Table 2: Evaluation of energy consumption estimation

5. Discussion

The current work presents an effort to build appliance-oriented unsupervised solutions to address the problem of non-intrusive load monitoring. The suggested framework consists of different signature models, one for each category of appliances, with emphasis on the characteristics of the operation cycle of each appliance. The entire methodology is centered around the idea of neither using ground truth data to train models, nor engaging the end user in the process.

The approach is based on different previous articles that follow this rationale, nevertheless adaptations are necessary in order for the solution to fit to the specifications of the smart meter involved.

The preliminary results that emerged from the implementation and respective experiments showed that there is a great potential of utilizing this methodology as an online solution in real-world scenarios. Nonetheless, corrective actions are yet to be applied to the existing solution to identify the reason(s) behind false event detection occurrences and corresponding energy estimation. Such actions include modification of the internal delta-power detection algorithm as well as tuning of the parameters involved in the generalized methodology.

As previously stated, this work reflects our intention to build a series of independent appliance models, the combination of which will result in an end-to-end solution that may be applied to households without previous training. Although the initial implementation is expected to be executed on cloud, the long-term objective is to move the entire solution on the edge (to be more precise on the smart meter).

The advantage of such a framework is its flexibility in terms of methodology, allowing the use of different approaches for each type of appliance. For example, though a heuristic approach might be utilized for one appliance, a neural network-based solution may fit better to another. This inter-modularity of methodologies can further be exploited in terms of on-edge implementation. Although up to date the energy disaggregation solutions are restricted to run entirely on cloud/edge/cloud & edge, our framework will provide the opportunity to implement each module at will. Thus, one appliance's model might be preferably implemented on-edge, while another may function better when it runs partly in cloud and partly on the edge device.

The limitations of our vision are yet to be determined, nonetheless we intend to keep working on our research and explore the capabilities of our ecosystem to incorporate a promising energy disaggregation solution.

6. Discussion

References

[1] Schirmer, P. A., Mporas, I., & Sheikh-Akbari, A. (2020). Robust energydisaggregation using appliance-specific temporal contextual information.EURASIP Journal on Advances in Signal Processing, 2020(1), 1-13.

[2] Azizi, E., Beheshti, M. T. H., & Bolouki, S. (2021). Event Matching Classification Method for Non-Intrusive Load Monitoring. Sustainability2021, 13, 693.

[3] Faustine, A., Mvungi, N. H., Kaijage, S., & Michael, K. (2017). A surveyon non-intrusive load monitoring methodies and techniques for energydisaggregation problem. arXiv preprint arXiv:1703.00785.

[4] Jacobs, G., & Maun, J. C. (2018, June). Identifying fridge consumption insupervised global electric consumption. In 2018 IEEE International EnergyConference (ENERGYCON) (pp. 1-6). IEEE.

[5] Jacobs, G., & Maun, J. C. (2019, June). Identifying washing machineconsumption in supervised global electric consumption. In 2019 IEEEMilan PowerTech (pp. 1-6). IEEE.

[6] Luan, W., Wei, Z., Liu, B., & Yu, Y. (2022). Non-intrusive powerwaveform modeling and identification of air conditioning load. AppliedEnergy, 324, 119755.

[7] Berges, M.E., Goldman, E., Matthews, H.S., Soibelman, L. (2010).Enhancing Electricity Audits in Residential Buildings with NonintrusiveLoad Monitoring. Journal of Industrial Ecology, 14:844–858, 2010.

[8] Najafi, B.,Moaveninejad, S.,Rinaldi, F. (2018). Chapter 17:Data Analytics for Energy Disaggregation: Methods and Applications. In Big DataApplication in Power Systems; pp. 377 – 408.

[9] Gupta, A., Chakravarty, P. (2016). Impact of Energy Disaggregation onConsumer Behavior. White Paper, pp.1–14.

[10] Fischer, C. (2008). Feedback on household electricity consumption: atool for saving energy? Energy Efficiency 1:79–104.