On the Optimization of Appliance Loads Inferred by Probabilistic Models

Marisa Figueiredo
Centre for Informatics and Systems
University of Coimbra
Polo II, 3030-290 Coimbra
Email: mbfig@dei.uc.pt

Bernardete Ribeiro
Department of Informatics Engineering
University of Coimbra
Polo II, 3030-290 Coimbra
Email: bribeiro@dei.uc.pt

Ana de Almeida
ISCTE-IUL
Universitary Institute of Lisbon
1649-026 Lisbon
Email: ana.almeida@iscte.pt

***Abstract*—Recent Non-Intrusive Load Monitoring (NILM) approaches consider probabilistic graphical models and statistical inference algorithms such as Hidden Markov Model (HMM). One interesting HMM based approach towards unsupervised energy disaggregation proposes prior models of general appliance types which are tuned to specific instances using only aggregated electrical consumption measurements. An essential step of this approach is the subtraction of the estimated usage from the aggregated load before the disaggregation of a new load. Then wrongly-detected states of a given device lead to errors that are disseminated by the subsequent disaggregation. In this paper we aim at investigating an unsupervised HMM based approach that overcomes this limitation. First, the general models are tuned for a selection of suitable periods of the signal in analysis. Second, the load disaggregation of each appliance is carried out. Third, given the inferred consumption of the whole set of devices and that the measured aggregated electrical usage is a linear combination of these loads, an optimization step regarding the adjustment of the calculated sources is performed. Although experiments yielded in the REDD data set show the adequacy of the approach, still further improvements are required.**

I. Introduction

Nowadays, electrical energy disaggregation is a research topic on its own. The emergence of smart grids and the energy efficiency concerns have driven the renewed interest on appliance-specific information and tools for perceiving how and when electrical loads are used without recurring to a swarm of individual meters, at least at the household level. Non-Intrusive Load Monitoring (NILM) systems [1] provide the appliance detailed information by acquiring the whole-home electrical demand signal at a single point (aggregated/mixed data) and breaking it up into the individual appliance/circuit signals in the electrical network requiring simple available hardware and complex software [1]. Currently, NILM literature has been focused in probabilistic graphic models able to represent the operation of appliances and in statistical inference algorithms used for the disaggregation of the whole-home electrical consumption. Hidden Markov Model (HMM) and their variants are the uppermost explored techniques [2], [3], [4], [5] mainly due to the assumption that a device can be characterized by a finite set of states and the search for unsupervised methods for disaggregation in order to avoid prior sub-metering. In fact, prior measurements of the individual consumptions may be required to model appliances [2] or general priors can be defined for each group of devices that are posteriorly adjusted [5].

This paper reports the exploration of an unsupervised HMM based approach. For each group of devices, a finite set of operation modes is defined based on the correspondent average power consumption. Following the approach presented in [5], the given general model is then adjusted for each appliance instance in a given house considering the measured aggregated signal. Additionally, at each instant of time, the sum of all power demands must correspond to the measured aggregated consumption. This requirement could be achieved by the separation of the power demand of a given equipment from the aggregated signal in analysis before the disaggregation of a new device, as proposed in [5]. Nevertheless, the inferred consumption of subsequent devices is negatively influenced if a state is wrongly-detected for the previous devices. For tackling this issue, the approach here presented considers the standard HMM and the same observed sequence for the load disaggregation for all the appliances. Additionally, the approach comprises also an optimization step where the power demand of each device is posteriorly adjusted.

The rest of this paper is organized as follows: next section introduces the necessary background, the HMM and appliance models, followed by describing the approach under consideration in Section III. The experimental setup designed is detailed in Section IV followed by the results, performance assessment and correspondent discussion. Lastly, conclusions and directions of future work are given.

II. Modeling Appliances

When load measurements are gathered only at a single point of an electrical network, the electrical energy disaggregation corresponds to the inference of appliance-detailed electrical consumption information over a period of time. Formally, given the measurements of the total electrical consumption, gathered during a period of time T, $\bar{x} = [\bar{x}(1), \bar{x}(2), \ldots, \bar{x}(T)]$, we seek for the loads associated with the consumption of each device or circuit $i, i = 1, \ldots, k$, *i.e.* for the sequence of power demands $x_i, i = 1, \ldots, k$, $x_i = [x_i(1), x_i(2), \ldots, x_i(T)]$. Assuming that the $\bar{x}$ is a linear combination of the k loads x_i then

$$\bar{x}(t) = \sum_{i=1}^{k} x_i(t). \tag{1}$$

Alternatively, given the whole-home power demand $\bar{x}$, the goal can be redefined as the uncovering of the states sequence

$z_i = [z_i(1), z_i(2), \ldots, z_i(T)]$ for each device i and where $z_i(t)$ corresponds to the state observed in instant t, like 'on' or 'off'. In fact, for each appliance, a finite set of states can be defined corresponding to periods of operation for which the difference between power demands within this interval of time are approximately zero. For instance, for a refrigerator three states can be defined: on, off and peak [5]. Then, a mapping between $x_i = [x_i(1), x_i(2), \ldots, x_i(T)]$ and the sequence of states $z_i = [z_i(1), z_i(2), \ldots, z_i(T)]$ exists. Furthermore, as for the steady-state signatures, a change of state would correspond to a change in the aggregated consumption (observed sequence). Thereby, HMM is a suitable technique to describe the power demands of each device [6].

HMM is a statistical tool based on Markov chains that builds probabilistic models for sequences of observations of length T, $\bar{x} = [\bar{x}(1), \bar{x}(2), \ldots, \bar{x}(T)] \in \mathbb{R}^T$. The HMM consists of a discrete-time, discrete-state Markov chain with hidden states $z(t) \in \{1, \ldots, N\}$, where N is the number of possible discrete values for states, together with an observation model $p(\bar{x}(t) \mid z(t))$. In this extension of Markov chains the observation is a probabilistic function of a state. The model has a non-observable stochastic process that is 'accessible' via another set of stochastic processes that produce the sequence of observations [6].

A HMM of a given device i^* is fully described by: (i) the initial probabilities associated with each state, $\pi_j = p(z(1) = j)$; (ii) the transition matrix A comprising the transition probabilities between states j and w, $a_{jw} \equiv p(z(t) = w \mid z(t-1) = j)$ such that $a_{jw} \geq 0$ and $\sum_{j=0}^{N} a_{jw} = 1$; (iii) the emission probability considering continuous observations $\bar{x}$ such that $p(\bar{x}(t) \mid z(t) = j, \phi) = \mathcal{N}\left(\mu_j, \sum_j\right)$ where μ_j and $\sum_j$ are the mean and variance of the Gaussian distribution describing the power demand of appliance i at state j. In short, a HMM for an appliance i is expressed by $\theta_i = \{\pi_i, A_i, \phi_i\}$.

Solving the energy disaggregation problem via HMM involves then the training of θ_i. For this exploratory work, generic models of appliances as proposed in [5] are used: general models of the different device types are defined, which are then tuned for the specific appliances in each house, using only aggregate power consumption of that household. The generic models consist of prior information for each parameter in θ_i. These priors, a generalization of a vast number of instances of the same appliance, are then adjusted for each instance in study for a given home. This adjustment phase considers the aggregated consumption of the house and searches for periods where only a single appliance is most likely is being turn on and off. Towards this end, the parameters in θ_i are tuned by Expectation-Maximization (EM) algorithm, performed for several overlapped time-windows for the aggregated time series. For each time window, the EM is initialized with the prior information for the given appliance and the achieved likelihood for the current time-window, $\mathcal{L}$, is used to determine if it is a suitable period for training the appliance instance. A time-window is said suitable for that purpose if $\mathcal{L} > D$ holds, where D is the appliance specific threshold. The following step only considers the previously selected time-windows to tune the prior models θ_i via the EM algorithm, resulting in the $\hat{\theta}_i$ model for the device i. In this work we employed the standard HMM while the approach proposed in [5] studies the difference HMM, where a second observation sequence, derived from the first one, exists.

*For the sake of readability the index i is omitted when obviously implied.

III. Electrical Energy Disaggregation via Viterbi and Source Optimization

Given the tuned models $\hat{\theta}_i$ and the whole-home electrical consumption, the following stage is the disaggregation of electricity usage. In this work we investigate a three step approach: i) tune the general priors θ_i via EM for each $i = 1, \ldots, k$, as described in the previous section; ii) consumption inference for each device i and iii) optimization of the inferred loads. At the second phase the most probable sequence of states for the appliance i is calculated for a given sequence of aggregated usage samples and consequently, the inferred consumption of device i, s_i, is obtained, considering also the tuned model $\hat{\theta}_i$. Repeating the process for each device does not ensure that Equation 1 holds. Alternatively to the process proposed in [5] where the inferred usage of each device would be subtracted from the aggregate consumption previously to the state inference of a new appliance, we propose a posterior optimization that takes in consideration the requirement associated with Equation 1. With this approach we seek to overcome the limitations found in the 'subtract' approach for which the calculation of usage associated with subsequence appliances would disseminate the errors of wrongly-detected states for the current device.

The second step of this approach consists on the calculation of the most probable sequence of states $\tilde{z}_i$, for each instant of time $t = 1, \ldots, T$, for each device $i = 1, \ldots, k$ given the observed sequence of whole-home consumptions $\bar{x}$, *i.e.*, for each i we must compute: $\tilde{z}_i = \operatorname{argmax}_{z(1:T)} p(z(1:T) \mid \bar{x}(1:T))$. Within the HMM context this sequence is achieved via Viterbi algorithm which considers the previously tuned model $\hat{\theta}_i$ and $\bar{x}$ to determinate the probability of the possible states for each t. Then, the inferred electrical usage associated with appliance i, s_i, is computed considering the sequence $\tilde{z}_i$ and the mean power for each state provided by $\hat{\theta}_i$.

Given then the inferred consumption for every device in study as a result of the Viterbi algorithm, the following stage (third step) considers that at each instant of time, the aggregated consumption is the linear mixture of the k consumptions of the appliances and where the s_i represents an initial approximation to the usage of appliance i. Thereby we add weights $m_i \in \mathbb{R}_+^T$ to tune $s_i, i = 1, \ldots k$ which must be calculated such that

$$\bar{x} = \sum_{i=1}^{k} s_i \otimes m_i \quad (2)$$

holds (where $\otimes$ stands for the elementwise multiplication). Alternatively, considering the sparse matrix $A \in \mathbb{R}_+^{T \times kT}$ with lines $a_t = [\underbrace{0, \ldots, 0}_{(t-1)\,k}, s_1(t), s_2(t), \ldots, s_k(t), \underbrace{0, \ldots, 0}_{(T-t)\,k}]$,and the vector $w \in \mathbb{R}_+^{kT}$,

$$w = [m_1(1), \ldots, m_k(1), m_1(2), \ldots, m_k(2), \ldots \ldots, m_1(T), \ldots, m_k(T)]^T$$

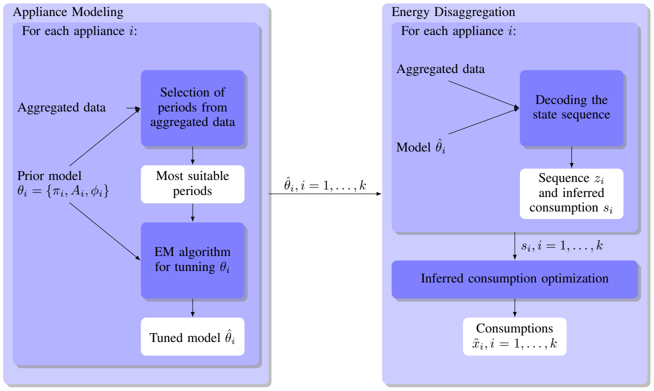

Fig. 1. Illustration of the proposed approach.

then the Equation 2 can be re-written as

$$\bar{x} = Sw. \tag{3}$$

Additionally, assuming that s_i is a good approximation to the actual consumption of the device i, the tuned load $\tilde{x}_i = s_i \otimes m_i$ should not differ significantly from s_i. Hence, it must be ensured that

$$\sum_{i=1}^{k} \|\tilde{x}_i - s_i\|_2^2 \equiv \sum_{i=1}^{k} \|s_i \otimes (m_i - 1)\|_2^2 \tag{4}$$

is minimum *i.e.* m_i should be as close as possible to the unit vector. Thus, the third step of the disaggregation approach consists of the following optimization problem:

$$\begin{aligned} \min \quad & \sum_{i=1}^{k} \|s_i \otimes (m_i - 1)\|_2^2 \\ \text{subject to} \quad & \bar{x} = Sw \\ & S_{tj} \geq 0, \quad \forall t = 1, \ldots, T, \quad \forall j = 1, \ldots, kT \\ & w_j \geq 0, \quad \forall j = 1, \ldots, kT. \end{aligned}$$

The overall approach schema is presented in Figure 1. Note that the optimization step allows for the calculation of the power consumption variations, which were “smoothed” by the definition of a finite set of states characterized by mean power consumption as required by the HMM. Thereby, a more accurate computation of consumption associated with the several appliances, in particular for appliances with continuous variable loads, is possible. This particularity distinguishes the proposed approach from the Factorial Hidden Markov Model, composed by several parallel and independent HMMs.

IV. Preliminary Computational Experiments

The freely-available Reference Energy Disaggregation Dataset (REDD) comprises whole-home and circuit/device specific electricity consumption for six real houses over several months. For each monitored house, the whole-home electricity signal and individual circuits in the home, each labeled with its category of appliance, were recorded [7]. The prior models were defined considering the whole dataset, yet for assessing the proposed approach, only the data associated to House 1 was used. A preprocessing phase selected only the common periods of sampling for both aggregated and individual signals. They were then downsampled using a median filter such that each sample became spaced 1 minute apart. Based on the percentage that each group represents in the total electricity consumed, the five ones presenting higher impact were selected. The sixth one contains the remaining power. The signals were subjected to a normalization step assuring that the relative importance of each group in the aggregated signal was maintained. The elements of each signal (aggregated and device ones) were normalised regarding the norm of the aggregated signal. In short, we focused on Refrigerator, Dishwasher, Kitchen Outlets, Lighting, Washer Dryer and Others. The post-processed data comprises 23 daily signals, which 16 were used for tuning the prior appliance models and the remaining was only used at the disaggregation step. The setup is similar to previous works allowing us to perform further comparisons.

The review of the NILM related literature reveals that a variety of performance indicators are employed and usually selected in accordance with the designed solution. In this exploratory work, the performance was measured in terms of the root-mean-square errors (RMSE), a measure between the inferred values and the truly observed amount. The RMSE was computed concerning both (i) an overview of the error and (ii) a detailed error assessment by appliance. For the former analysis, the RMSE associated with the aggregated signal $\bar{x}$ and its predicted version $\tilde{\bar{x}} = \sum_{i=1}^{k} \tilde{x}_i$, $RMSE(\bar{x}, \tilde{\bar{x}}) = \sqrt{\frac{\sum_{t=1}^{T} \left(\bar{x} - \tilde{\bar{x}}\right)^2}{T}}$, was computed. For the latter, the RMSE corresponding to each device i between the measured consumption x_i and its predicted version $\tilde{x}_i, i = 1, \ldots, k$, was also calculated using the above Equation with the appropriated adjustments.

In this computational experiment we explore the suitability of the proposed approach, the third step in particular. Figure 2 reports the average of the RMSE values for 5 runs, calculated in accordance with the above description. Namely the RMSE values computed concerning the inferred loads without the optimization step (denoted as ‘RMSE after Viterbi’) and the RMSE values calculated regarding the adjusted loads after the optimization step (denoted as ‘RMSE after optimization’). The leftmost columns report the RMSE associated with the aggregated signal while the remaining report the RMSE values by appliance.

A notorious decrease on the RMSE value correspondent to the aggregated signal occurred. Indeed, considering the inferred aggregated consumption as the sum of the inferred loads without the optimization step, the RMSE is above 0.018 while this error was virtually null for the aggregated signal calculated after the adjustment resultant from the optimization ($1.14 \times 10^{-13} \pm 5.09 \times 10^{-14}$). Note that the built formulation ensures the equality between the loads of appliances and the total measured electricity (see Equation 3) thereby a low error value was expected. A more detailed analysis is provided by the RMSE associated with each group in study. Once again, a prominent decreased between the inferred consumption after the disaggregation via the Viterbi algorithm and the ‘tuned’ usage demands by the optimization step is observed for all the devices in analysis. Moreover, the decrement in the error was, with exception of the group ‘Others’, always superior to 50%. The error reduction is specially prominent for the ‘Lighting’ and ‘Washer Dryer’ whose RMSE corresponds to only 21% and 25%, respectively, of the error calculated for the loads

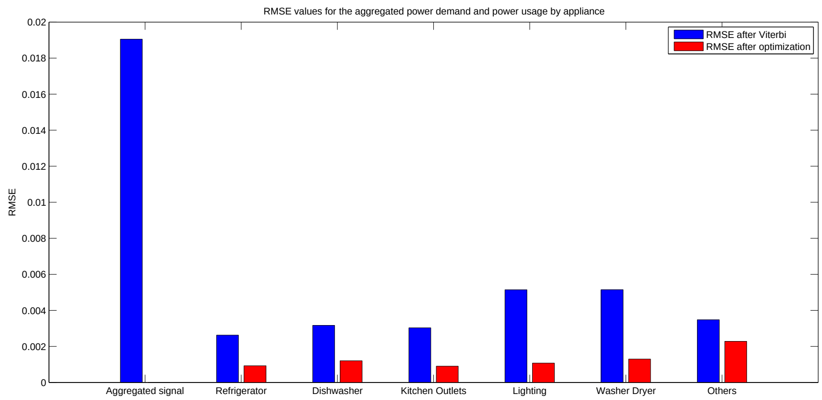

Fig. 2. RMSE values considering the assigned aggregated consumption and the power usage of each appliance in study.

based only on the Viterbi outcome. Regarding the magnitude of the errors, we observed that the post-optimization RMSE associated with both the aggregated signal and the several devices in analysis has a similar magnitude to the obtained results for other techniques as presented in [8] showing the effectiveness of the optimization step. On the other hand, the pre-optimization error (RMSE after Viterbi) is substantially higher than RMSE values calculated for the optimized loads. Additionally, the performed empirical analysis revealed the need of further improvement of the computed loads based on the Viterbi algorithm outcome, namely concerning the appliance modeling and the observation sequence used. Note that the outcome of the Viterbi algorithm is crucial for the optimization step: the function being minimized is based on the distance between the 'tuned' loads and the inferred usage based on the decoding process (see Equation 4). That is, the approach assumes that the outcome of the second step is a relative good approximation to the actual consumption of each device. Nevertheless, the achieved results lead to the conclusion that this assumption is not verified which implies further improvements to the approach are of upmost importance.

V. Conclusion and Future Work

The relevance and pertinence of research on Non-intrusive Load Monitoring (NILM) is undeniable. Currently several researchers pursue unsupervised techniques in order to avoid the requirement of prior individual measurements of appliances' consumption. For this purpose, NILM studies are, at the moment, focused on probabilistic graphic models, namely HMM.

Inspired by the generic models of different appliance types proposed in the related literature, this exploratory work studies an unsupervised HMM based approach. For a given house, these general models are adjusted for the particular instance of the appliances there used, requiring only measurements of the aggregated electrical consumption of the house. Next, the load disaggregation occurs taking in consideration the previously tuned appliance models. Still the Viterbi algorithm *per si*, used to calculated the most probable sequence of states for a given appliance, does not take into consideration that the observation sequence of the aggregated signal is an linear combination of the appliance loads. For that purpose, the approach could 'subtract' the computed power usage of a given appliance from the aggregated time-series previous to the adjustment and disaggregation of a new device. Yet errors would be disseminated by the subsequent power disaggregation. Alternatively, this exploratory work proposes a posterior optimization step to tune the computed electrical usage associated with each appliance, ensuring that the linear combination of these loads corresponds to the measured aggregated signal. Experiments used the REDD dataset and demonstrated the feasibility of the approach. Still further improvements are required, namely in terms of appliance modeling via HMM. In fact, the loads computed according to the most probable sequence of states are the base of the optimization step, thereby the appliance modeling via HMM must be as accurate as possible. Further improvements and on-going work consider, for instance, the so-called delta coefficients, widely used in speech recognition, and the use of an observation sequence derived from the subtraction of a baseline power draw from the aggregated signal.

Acknowledgment

FCT is acknowledged for the grant SFRH/BD/68353/2010.

References

[1] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. of the IEEE*, vol. 80, pp. 1870–1891, 1992.

[2] T. Zia, D. Bruckner, and A. Zaidi, "A hidden markov model based procedure for identifying household electric loads," in *Proc. of Annual Conf. on IEEE Industrial Electronics Society*, 2011, pp. 3218–3223.

[3] H. Kim, M. Marwah, M. F. Arlitt, G. Lyon, and J. Han, "Unsupervised disaggregation of low frequency power measurements," in *Proc. of the SIAM Intl. Conf. on Data Mining*, 2011, pp. 747–758.

[4] J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation," *Journal of Machine Learning Research*, vol. 22, pp. 1472–1482, 2012.

[5] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Non-intrusive load monitoring using prior models of general appliance types," in *Proc. of Conf. on Artificial Intelligence*, 2012.

[6] L. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proc. of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[7] J. Z. Kolter and M. J. Johnson, "REDD: A public data set for energy disaggregation research," in *Proc. of SustKDD Workshop on Data Mining Applications in Sustainability*, 2011.

[8] M. Figueiredo, B. Ribeiro, and A. de Almeida, "Electrical signal source separation via nonnegative tensor factorization using on site measurements in a smart home," *IEEE Trans. on Instrumentation and Measurement*, vol. 63, no. 2, pp. 364–373, 2014.