Towards Automatic Identification of Activity Modes in Electricity Consumption Data for Small and Medium Sized Enterprises

Emil Holmegaard and Mikkel Baun Kjærgaard
Center for Smart Energy Solutions
Mærsk McKinney Møller Institute
University of Southern Denmark, Denmark
{em,mbkj}@mmmi.sdu.dk

***Abstract*—Small and Medium Sized Enterprises (SMEs) have a large share of the total electricity consumption in most countries and the efficiency of energy reduction efforts for SMEs by energy counselors depend on means for automatically identifying which SMEs are good candidates for energy reduction efforts. In this paper we propose a method based on statistics to identify activity modes in electricity consumption data of SMEs coarse-grained smart meter data. The method enables such automatic identification by computing a number of features on smart meter data including autocorrelation features. We also provide initial results for evaluating the method on a data set with smart meter data for 329 SMEs and identify options for further improving the method. Given the identification of the activity modes one can quantify the energy reduction potential of individual SMEs by comparing the load in the different activity modes and the variability of the load.**

Keywords*—*Disaggregation, Energy Efficiency, Data Analysis, Smart Meters

I. Introduction

In Denmark production enterprises consumes 32.6% of the total Danish electricity consumption [1]. Small and Medium Enterprises (SME) have a share of over 99% of Danish enterprises and their revenue is 55% of the total revenue of Danish companies [2]. Therefore SMEs consume a significant share of the electricity consumed in Denmark. To meet improvement goals in energy efficiency both at a national and an EU level it is important to foster SMEs to decrease their electricity consumption. One method for achieving reductions is government sponsored energy counseling and energy management initiatives [3]. However, to optimize such efforts it is important to know which SMEs would be good candidates for reduction efforts. Given that counselors have access to smart meter data from SMEs in their area it would improve their efforts if methods could automatically identify good candidates for reduction efforts.

One approach to identify good candidates for reduction efforts is to benchmark the consumption with regards to other enterprises in the same sector. However, for SMEs it is more difficult to compare and establish such benchmarks due to the large variety in consumption [4]. A second approach from a building perspective consider the comparison of baseline models normalized for temperature and lighting conditions [5] which, however, ignores the impact of enterprise activities.

In this work we follow another approach which relate consumption to enterprise activities to disaggregate the electricity consumption into loads that relate to activity modes of an enterprise. The argument for focusing on activity modes is that such modes are a good basis for identifying reduction options. For instance, that consumption while activities are idle should be low, cyclic repeating and stable. We here divide enterprise activities into three activity modes suggested to us by local energy counselors who apply this terminology in their daily work. The three activity modes are *idle mode* the minimum consumption with no activity in the enterprise, *base mode* the consumption used for enterprise activities in steady state and *process mode* the consumption used that correlate with production activities. Figure 1 illustrates the three modes on top of a weeks electricity consumption data from a smart meter installed at an SME. In regards to energy reduction efforts the idle and the base mode consumption are the first targets as changes here will not directly impact the activities of the enterprise, e.g., production processes.

The process of identifying activity modes is non-trivial due to the deviations in energy characteristics among SMEs [4]. Methods from non-intrusive load monitoring (NILM) for disaggregation of electricity consumption have been mainly focusing on households and commercial buildings, using pattern recognition to detect individual devices or overall loads [6], [7], [8]. Given our goal to apply our methods at scale the only type of data relevant to consider at present is course-grained smart meter data with a granularity in the range of 15-60 minutes. This rules out the ability to detect individual devices [6], however, individual device detection might also be more problematic in an SME context as they compared to households: i) utilize more diverse equipment making it harder to gather relevant training data and ii) experience a higher degree of activity variability meaning that more devices will be turning on and off nearly at the same time. Therefore we in this work focus on identification of activity modes and given the variability in consumption among SMEs focus on unsupervised methods. In comparison with existing unsupervised methods for NILM by Kolter et al. [9] we consider activity modes rather than individual device loads.

This paper argue for the development of NILM methods for identifying electricity activity modes from coarse-grained electricity consumption data. The methods we propose for disaggregation build on computing a set of features designed to identify the three electricity activity modes. We report initial results for applying our methods to electricity consumption data for SMEs from the Irish CER data set [10] which contains smart meters readings for more than a year with at a granularity of thirty minutes. Our initial results highlight that the method can identify the activity modes in several cases but also highlight some challenges due to large temporal variability in SME electricity consumption.

II. Automatic Identification of Activity Modes

Given electricity consumption data with a resolution of 15-60 minutes, Carrie et al. [6] defines three categories that can be detected from such data: loads that correlate with outdoor temperature, loads that are continuous and loads that are time dependent. For the purpose of finding potential energy reduction candidates among SMEs, we will look into the loads that are continuous and time dependent. In particular we focus on disaggregating data into three activity modes as mentioned in the introduction: *idle mode*, *base mode* and *process mode*. The three activity modes are illustrated in Figure 1 and further described here:

Idle Mode The minimum consumption with no activity in the enterprise. Examples of individual loads could be light for emergency exits, standby lights, network equipment or electric heating.

Base Mode The consumption used for enterprise activities in steady state including the base consumption of any production process, e.g. the electricity consumed regardless of if one thousand or one piece is produced. Examples of individual loads include electricity used for conveyor belts or lights in production areas.

Process Mode the consumption used that correlate with production activities. Examples of individual loads include boilers, where the consumed electricity relates to how much that should be boiled. Another example could be a compressor, which would show a high frequency of repeated consumption when a process produces a high amount of items, and have a low frequency then the process produces a low amount of items.

Given the consumption of the three activity modes, the total aggregated electricity consumption E_{total} for the time t is given by equation (1):

$$E_{total}(t) = E_{idle}(t) + E_{base}(t) + E_{process}(t) \qquad (1)$$

Figure 1 shows E_{total} over a week as measured by a smart meter for the example *SME A* taken from the CER data set. In addition we have highlighted the consumption in regards to the different activity modes and we can observe that in this case the idle mode consumption (a) is almost constant, and the base mode load (b) is near identical each day during production activity, whereas the production consumption (c and d) shows a large variability. For SME A it is easy for the untrained eye to identify such modes. As stated in [4] diversity in SMEs makes it hard to use an one fits all approach. Furthermore the field of business can have process cycles that differ from a hourly or daily cycle. An example of this phenomenon is given for *SME B* in Figure 2 where it is more difficult with the above definitions to identify what is the idle, base and process

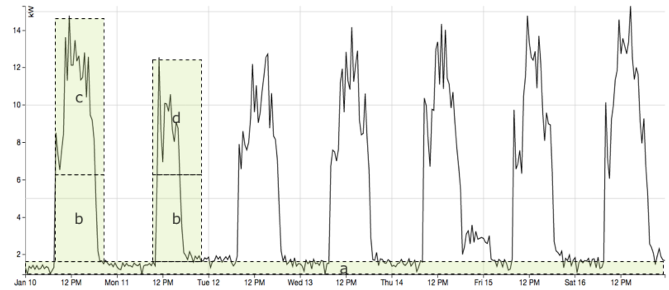

Fig. 1. SME A - aggregated electricity consumption over seven days, (a) is the idle consumption, (b) is the base consumption, (c) and (d) are process consumption. The boxes (a) to (d) was set by the authors, based on a interview with an energy counsellor.

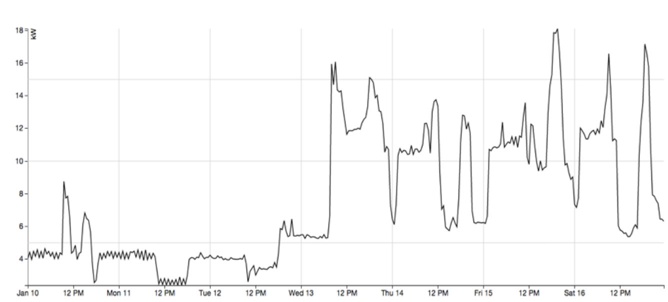

Fig. 2. SME B - aggregated electricity consumption over seven days, no daily pattern present

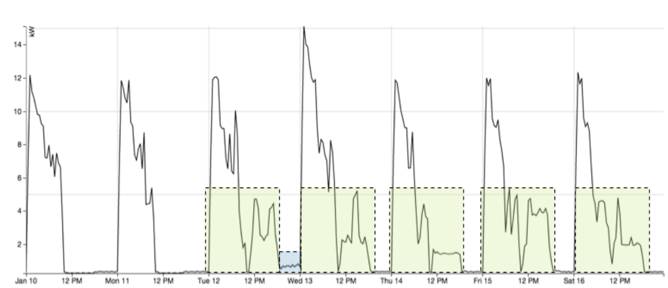

Fig. 3. SME C - aggregated electricity consumption over seven days, the green boxes highlight periods with different processes before and after midday. The blue box indicates an abnormality in the idle mode consumption.

mode consumption, e.g., the enterprise might have a production process spanning several weeks using different equipment that make it difficult to identify the modes on a daily basis. Figure 3 gives a third example for *SME C* where the duration of the load differ from week to weekends and the process load is different before and after noon (highlighted with green boxes).

The result of the activity mode classification enables several types of input to energy counseling activities. Given the identified activity modes enterprises might be ranked according to the following criteria to optimize the energy reduction efforts. Firstly, a high idle or base mode consumption might indicate that there exists options for reducing consumption by decreasing activation of equipment. Secondly, deviations in idle mode consumption might indicate that equipment is unnecessarily activated. An example of such deviations can be observed in Figure 3 highlighted with a blue box. In this case it, e.g., could be lights that has not been turned of when the employees left. Thirdly, increase in the base mode consumption can be due to equipment malfunctioning.

A. *Features*

In the following we will discuss the design of features to identify the three activity modes and the classification heuristic applied to the features.

1) Idle Mode: For finding idle mode consumption in SMEs, there must be a period without any production. This includes holiday periods and for some SMEs weekend periods and over night, e.g., as in Figure 1. The main property that we would like to utilize in our features is the stability of the consumption during such periods. A common approach for analysing the stability of signals is to use autocorrelation. Autocorrelation is a mathematical representation of the degree of similarity between a time serie and lagged copies of it. The autocorrelation can be calculated for different temporal lags. Autocorrelation takes a value between -1 representing a negative correlation and 1 representing a perfect correlation. Therefore a high autocorrelation in a certain time span, will indicate large stability and be a potential idle mode consumption. In addition we apply two heuristics to identify idle mode periods: i) we consider only entries e where the autocorrelation is above α in a time window of t around e; ii) we filter out periods where the average consumption in the period is higher that the daily average consumption to enforce that the electricity consumption is low during idle mode.

2) Base Mode: For identifying base mode consumption, there most be processes active meaning the consumption most increase and decrease with a similar level over a period of time. Furthermore, such processes are often periodic repeating. Therefore we are looking for a period with both an increase and a decrease in consumption higher than δ in kW within a sliding window of w readings. Finding periods with processes active, corresponds to finding the width of (b) in Figure 1. After we have found some candidates for potential periods we control that the period are periodic, meaning that at least two periods have starting points in the same hour of day. For finding the actual base consumption we take the arithmetic mean of for the electricity consumption in all the matching periods with active processes.

3) Process Mode: Given that we identify the idle and base mode consumption the process mode consumption will be the remaining consumption, as seen on Figure 1 and in equation (1). It is more challenging to directly identify the process mode consumption without having more fine-grained data that could directly identify the use of specific equipment.

III. Initial Results

For evaluating the proposed methods we present initial results for idle, base and process mode identification from electricity consumption data of SMEs. For the evaluation we utilize a data set covering 329 SMEs collected by the Irish Commission for Energy Regulation (CER) [10] in a smart meter trial. The data set consists of electricity consumption in kilowatts for 1.5 years of 329 SMEs with a granularity of 30 minutes. For the initial results we focus on one week of data from week two of 2010. For the autocorrelation feature

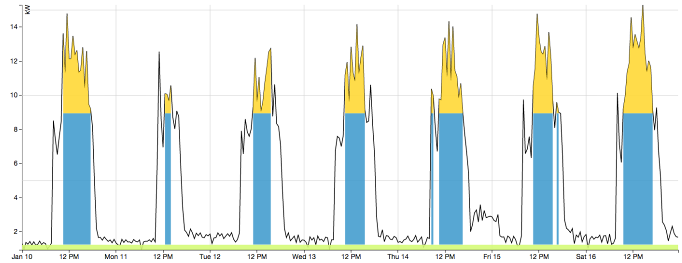

Fig. 4. Consumption for SME A with classified activity modes where green is idle mode at 1.2 kW and blue is base mode at 9.0 kW.

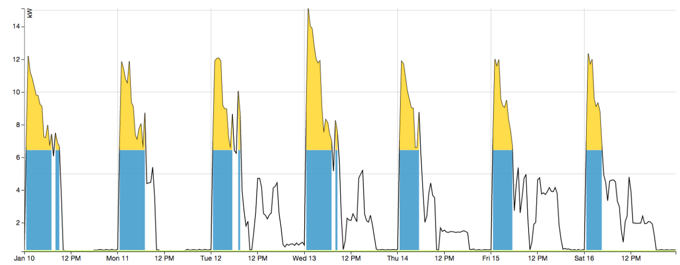

Fig. 5. Consumption for SME C with classified activity modes where green is idle mode at 0.3 kW and blue is base mode at 6.5 kW.

we use a lag of 1 which for this data set is equal to 30 minutes and for the heuristics we use $\alpha = 0.9$ and $t = 90$ minutes. For the base mode consumption feature we are using sliding window $w = 4$ readings and $\delta = 1.0$ kW.

For the initial results we will, firstly, consider if the features for each SME is able to identify any periods where we have idle and process mode and, secondly, we will discuss some examples where it is possible and some where it is not. Evaluating the methods in terms of accuracy with a proper ground truth is part of our ongoing work.

A. *Identification Results*

For the 329 SMEs the features identify periods with idle mode consumption for 81.8% of the SMEs. In the case of base mode consumption the feature found the base mode consumption for 79.6% of the SMEs.

B. *Examples with Successful or Failed Identification*

In Figure 4, 5 and 6 the found idle, base and process mode consumption has been illustrated on top of the electricity consumption of three different SMEs.

Figure 1 and 4 depict data for the same SME with activity modes identified by the authors and the proposed methods, respectively. When one compares the two figures one can observe that the methods generally identified the same activity modes and consumption as the authors with the exception that the consumption during the base mode is slightly higher than the consumption identified by the authors. Furthermore the base mode are centralised around 12 PM, and it looks like there is some processes before and after indicating that there could be more than one base mode.

For Figure 5 the automatic detection of activity modes identified activity close to how the authors would have read the

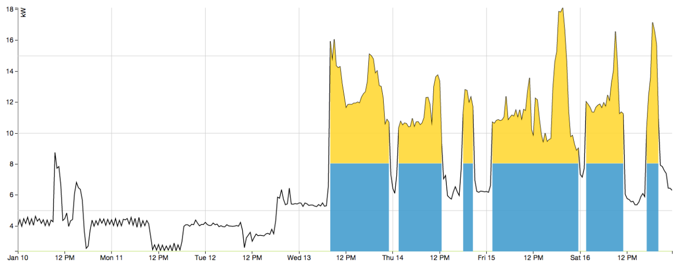

Fig. 6. Consumption for SME B with classified activity modes where green is idle mode which could not be identified and blue is basis mode at 8.1 kW.

consumption diagram. Only exception is that the base mode consumption for this SME does have a recurring pattern in the afternoon, which the feature does not catch, when only we are finding one base mode.

As mentioned earlier it was expected that the classification of SME B shown in Figure 6 would be difficult. It was not possible to identify any idle mode as there were no interval with equal or high similarity. For the base mode consumption the value seams to be almost as high as excepted, but again the pattern of SME C indicates a need for more than one base mode, when the pattern is not recurring at the same hour each day.

IV. Discussion

The initial results both show encouraging results and highlight options for improvements. In our ongoing work we are improving the methods along the following four dimensions.

1) Improving Idle Mode Detection: A problem observed for the idle mode detection is that 18.2% of the 329 SMEs were not identified to have any periods of idle mode time. This might be due to that we only include one week of data for the initial results, e.g., without any holiday periods. However, we expect that some SMEs might have enterprise activities nearly all 24 hours which means that the method has to be less restrictive, e.g., by lowering thresholds to identify idle mode time.

2) Improving Base Mode Detection: For base mode the method was not able to identify any base mode for 20.1% of the 329 SMEs. One issue explaining this trend is the SMEs without recurring processes at the same hour each day. A solution could be to add a filter for finding similar consumption based on different patterns e.g. monthly or weekly patterns. As seen in Figure 4 and 5 it could indicate that our approach of having one base mode, should be reconsidered.

3) Lower Granularity of Consumption Measurements: We are also considering if given measurements with lower granularity could improve the results and create opportunities for extending the methods. For instance, a lower granularity will enable more options for identifying individual processes and equipment.

4) Normalisations of SMEs for Comparison: For using this approach there is missing some mechanism for comparing different SMEs, too find the one with largest potential for reducing there energy consumption. To make this comparison, there must be a function for normalizing the energy consumption for a better comparison. For normalizing it could be good to know e.g. the number of employees, the size of buildings and type of business, etc.

V. Conclusion

In this paper we have proposed a method to disaggregate the electricity consumption of SMEs into the three activity modes: idle mode, base mode and process mode. The proposed method apply statistics for identifying these modes. Initial results on seven days of 30 minutes consumption data for 329 SMEs showed that in over eighty percent of the cases the method could identify such modes. By presenting individual SME examples a discussion was given of the accuracy of the methods for individual cases and ideas for further method improvements were discussed. Future work include improvements of the individual features, considering data with a lower granularity and an evaluation of the methods with proper ground truth provided by energy counselors. The proposed methods will when fully extended provide better data analytics to energy counselors thereby improving their energy reduction efforts.

Acknowledgment

The authors would like to thank The Commission for Energy Regulation (CER), Electricity Customer Behavior Trial for providing data. European Regional Development Fund (The Region of Southern Denmark) and Regional Commercial Development Fond for funding the Micro Grid Living Lab project.

References

[1] Danish Energy Agency, “Energy Statistic for Denmark 2012 - data, tables, statistics and maps,” 2012.

[2] P. V. Freytag, K. Klyver, and S. Løwe Nielsen, “CESFO Yearly Report 2012,” 2013.

[3] E. Abdelaziz, R. Saidur, and S. Mekhilef, “A review on energy saving strategies in industrial sector,” *Renewable and Sustainable Energy Reviews*, vol. 15, no. 1, pp. 150 – 168, 2011.

[4] J. Rawlings, P. Coker, J. Doak, and B. Burfoot, “Do smart grids offer a new incentive for sme carbon reduction?” *Sustainable Cities and Society*, vol. 10, pp. 245–250, 2014.

[5] P. P. Granderson, J., “Evaluation of the predictive accuracy of five whole-building baseline models,” August 2012.

[6] K. Carrie Armel, A. Gupta, G. Shrimali, and A. Albert, “Is disaggregation the holy grail of energy efficiency? the case of electricity,” *Energy Policy*, vol. 52, pp. 213–234, 2013.

[7] O. Parson, S. Ghosh, M. Weal, and A. Rogers, “Non-intrusive load monitoring using prior models of general appliance types.” in *AAAI*, 2012.

[8] 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 and Buildings*, vol. 30, no. 3, pp. 245–259, 1999.

[9] J. Z. Kolter and T. Jaakkola, “Approximate inference in additive factorial hmms with application to energy disaggregation,” in *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2012*, 2012, pp. 1472–1482.

[10] Commission for Energy Regulation, “Data from the commission for energy regulation.” [Online]. Available: http://www.ucd.ie/issda/data/commissionforenergyregulationcer/